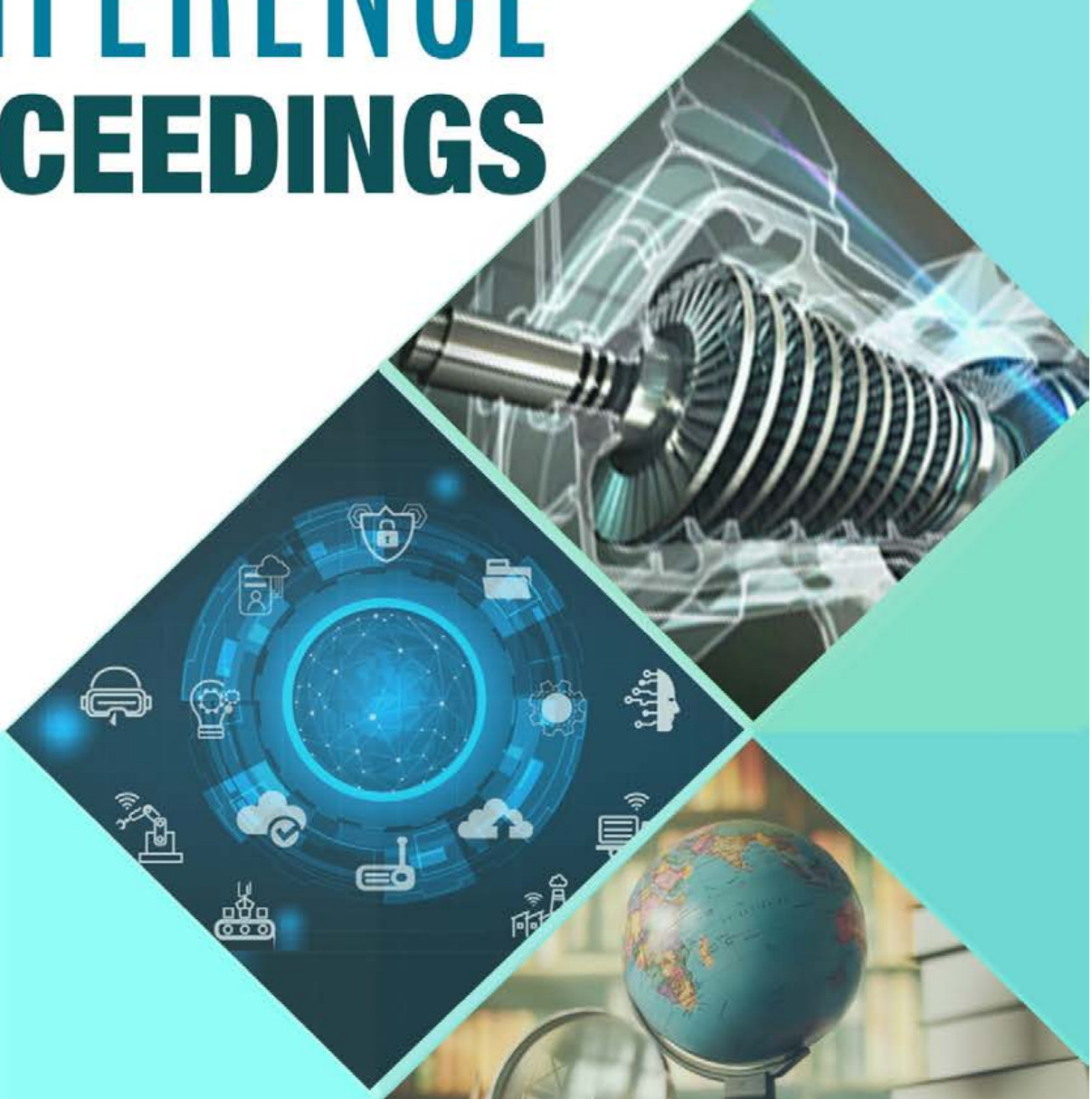
ISSN 2756-9160

# ICATC-2021

THE 6TH INTERNATIONAL CONFERENCE IN ADVANCES OF TECHNOLOGY AND COMPUTING

# CONFERENCE PROCEEDINGS



Faculty of Computing and Technology,
University of Kelaniya, Sri Lanka


***"DEFINING THE ROLE OF TECHNOLOGY FOR ECONOMIC DEVELOPMENT OF A COUNTRY"***

# International Conference on Advances in Technology and Computing

# (ICATC–2021) Proceedings

18th December 2021

Faculty of Computing and Technology

University of Kelaniya

Sri Lanka

2021

ISSN 2756 - 9160

International Conference on Advances in Technology and Computing (ICATC–2021) Proceedings



International Conference on Advances in Technology and Computing Secretariat
Faculty of Computing and Technology
University of Kelaniya.
Sri Lanka

Tel / Fax : +94112912709
email : icatc@kln.ac.lk
web : http://conf.kln.ac.lk/icatc

**Table of Contents**



**Research Papers: Titles and Authors**

**A: Advancements in Applications of Computing Technology**

**B: Advancements in Applications of Engineering and Science in Technology**

**C: Advancements in Applications of Technology in Teaching and learning**

# Organizing Committee of the International Conference on Advances in Computing and Technology (ICACT–2021)

## Conference Chair

Dr. Kasun Fernando

## Chief Editors

Dr. (Ms) Thilini Rupasinghe
Dr. Laalitha S. I. Liyanage

## Editorial Advisory Chairman

Professor Nilanthi de Silva, Vice Chancellor, University of Kelaniya
Professor Kapila Seneviratne, Chairman of the Research Council, University of Kelaniya

## Editorial Board

Prof. N. G. J. Dias
Dr. Chamli Pushpakumara
Dr. (Ms) Carmel Wijegunasekara
Dr. Gamini Wijayarathna
Dr. Hesiri Dhammika Weerasinghe
Dr. Sidath Liyanage
Dr. (Ms) Devindri Perera
Dr. (Ms) Thilini Rupasinghe
Dr. Laalitha S. I. Liyanage
Dr. Pradeep Samarasekere
Dr. Kasun Fernando
Ms. Shanika Arachchi
Dr. (Ms) Anuradhi Welhenge
Dr Chanaka Udayanga
Dr. (Ms) Shakila Pathirana
Dr. (Ms) Sandeli Kasthuriarachchi
Dr. Rasika Rajapaksha

## Reviewers

Meta-Reviewer
Dr. Rasika Rajapaksha

Internal Reviewers – FCT, University of Kelaniya, Sri Lanka

Dr. Sidath Liyanage
Dr. (Ms) Thilini Rupasinghe
Dr. (Ms) Anuradhi Welhenge
Dr. (Ms) Sandeli Kasthuriarachchi
Dr. (Ms) Shakila Pathirana
Dr. Chanaka Udayanga
Dr. Laalitha Liyanage
Dr. Kasun Fernando
Dr. (Ms) Amila Jeewandara
Dr. Chathuranga Hettiarachchi
Dr. (Ms) Carmel Wijegunasekara
Dr. (Ms) Devindri Perera
Dr. Dhammika Weerasinghe
Ms. Shanika Arachchi
Ms. Udari Perera
Mr. Tharuka Sarathchandra
Mr. Pathum Udukumburage
Ms. Malsha Mahawatta
Ms. Hansi Udapola
Mr. Nisal Udawatta
Mr. Rajitha Peiris
Ms. Udara Rathnayake
Mr. Krishnadeva Kesavan
Ms. Niromi Chathurika
Mr. Nisal Geelaka
Ms. Darsha Jayamini
Mr. Thisara Pathirana
Mr. Chathura Priyankara

External Reviewers – Sri Lanka

| | |
|---|---|
| Dr. Sameera Viswakula | : University of Colombo, Sri Lanka |
| Dr. (Ms) Aashani Tillekaratne | : University of Colombo, Sri Lanka |
| Dr. Indrajith Nissanka | : University of Moratuwa, Sri Lanka |
| Dr. (Ms) Damayanthi Herath | : University of Peradeniya, Sri Lanka |
| Dr. (Ms) M.R.N Yatigammana | : University of Kelaniya, Sri Lanka |
| Dr. Asitha Kulasekera | : University of Moratuwa, Sri Lanka |
| Dr. Uditha Jarayathne | : DMS Software Technologies (Pvt) Ltd, Sri Lanka |

External Reviewers – Foreign

| | |
|---|---|
| Mr. Susara Thenuwara | : University of Tokyo, Japan |
| Eng. Rochelle Silva | : Nanyang Technological University, Singapore |
| Mr. Chamin Nalinda | : University College Dublin, Ireland |
| Mr. Renxiang Lu | : University of Wyoming, United States |
| Dr. Saliya Kirigeeganage | : Hydromax, United States |
| Dr. Gayan Abeynayake | : Cardiff University, United Kingdom |
| Dr. Buddhika Abeyweera | : Intel Corporation, United States |

**Publication committee**

Publication Chair
Dr. Chathuranga Hettiarachchi
Dr. Anuradhi Welhenge
Mr. Rajitha Peiris

CMT System Management
Ms. Navodi Mekhala
Ms. Sulani Kavirathne

Proceeding Editors
Ms. Hansi Yashodara
Ms. Sujani Perera
Ms. Thishakya Hettiarachchi

Publication supporting Staff
Ms. Lakmali Gamage
Ms. Nishavi Samaraweera
Ms. Niromi Chathurika
Ms. Prabha Mihirani
Mr. Nisal Geelaka
Ms. Nadeesha Tharundi
Ms. Thilini Tharanga
Ms. Chamishka Jayathilake

**Additional Contributions**

Cover Page & other Creative Designs
Ms. Manuja Wickramasinge

Web Site Management
Mr. Pathum Udukumburage

Event Management
Mr. Thisara Pathirana
Mr. Akash Perera
Ms. Iyanthika Basnayake

# Message from the Vice Chancellor of University of Kelaniya

It is with great pleasure that I write this message for ICACT-2021, the 6th International Conference on Advances in Computing and Technology, organized by our University's Faculty of Computing & Technology.

I am informed that this year's conference, which is based on the theme "*Defining the role of technology for economic development of a country*", will involve several firsts. A Faculty Research Symposium that marks the graduation of the very first batch of students from the Faculty will help them to move into an international level. Another first is the Pitch 90 program that involves an official collaboration with the Ministry of Education, to reach school children studying in the Technology Stream for the GCE Advanced Level examination, which is meant to encourage them to work on technical solutions to problems in their own communities.

The international conference remains the main event, however, providing a forum for the presentation of the 40 best manuscripts submitted to the conference. Over the last year, we have all discovered that online conferences offer a great platform for the dissemination of knowledge, not just in times of travel restrictions and lockdowns, but even at other times, because it enables us to do away with a lot of tedious travel and communicate much more directly with each other.

I congratulate the organizing committee for putting together a conference programme that brings so many different stakeholders together to share their ideas and present their findings, at a time when technology has never been more important in getting our economy out of the dire straits that we face, both nationally and internationally. I am sure that the Organizing Committee has put in a lot of effort into organizing the different events that form part of this year's conference, and thank all of them most sincerely for their hard work.

Finally, I would like to take this opportunity to wish all the participants of ICACT 2021 fruitful deliberations!

**Senior Professor Nilanthi de Silva**
MBBS (Colombo), MSc (London), MD (Colombo)
Senior Professor of Parasitology and Vice Chancellor

# Message from the Dean of the Faculty of Computing and Technology

It is an honor to write this message for the International Conference on Advances in Computing and Technology 2021 (ICACT 2021) as the Dean of the Faculty of Computing and Technology, University of Kelaniya, Sri Lanka. Technology invades our day to day lives. The Covid-19 pandemic has catalyzed the adoption of technology into day-to-day activities of the citizens around the Globe. A special attention should be given to the technology-based food supply chain, health, and education of every citizen. This is an opportunity for innovations with lot of challenges which need to be discussed and argued enough to ensure the health and safety of individuals, and to identify the social and cultural impact. The opportunities come infrequently with challenge. It is the duty of the research and development community to identify the opportunities immediately and grab them for innovations for the betterment of each member of the society. I am sure that ICACT 2021 has opened a wide discussion forum for the researchers and other stakeholders from different geographical locations to meet virtually and have a fruitful discussion in line with the conference theme "Defining the role of Technology for Economic Development of a Country". I am also sure that the keynote speakers will provide the insight into the theme of the conference ICACT 2021 and set the stage for the three conference tracks of the ICACT2021.

The three additions to the ICACT 2021: The Faculty Research Symposium, Pitch 90 and Workshop target the postgraduate, undergraduate and school student communities. These open room for student community to listen to and work with experts in the respective fields and learn. We believe that these new additions will have a significant impact on student research development work. Hope that ICACT 2021 will help the faculty to produce right human resources for the socio-economic development of the county by producing at right time.

I take this opportunity to wish all the best for all the presenters and the researchers.

**Dr. Gamini Wijayarathna**
DrEng (Electro-Communications, Japan)
MEng (Electro-Communications, Japan), BSc (Kelaniya, Sri Lanka)

# Keynote Speech – “Data Science and Analytics during Gaze-based Interactions”

“With the rise of affordable eye tracking technology, it is now possible to follow a person’s gaze patterns on a computer screen, i.e., revealing exactly where on the screen a person is/is not looking at a given time (and for how long), as well as the person’s movements from one area to another. This in turn enables the analysis of a plethora of aspects about a person and their interaction using machine learning and data science techniques. For example, we study how to develop eye tracking technology to identify variety of aspects that can relate to the student learning process (e.g., moments of confusion, boredom, delight), as well as aspects about the multi-user interactions (e.g., joint attention). Data Science & analytics during gaze-based interactions will be useful in investigating the search as a learning process and to employ these sensory data as assessment of reading, mind-wandering and inferring metadata features for machine learning models.”

**Dr. Sampath Jayarathna**
PhD (Texas A&M)
Assistant Professor
Department of Computer Science
College of Science
Old Dominion University
Norfolk, VA 23529
sampath@cs.odu.edu

# Keynote Speech – “Adoption of Technologies in Agricultural Development”

“Any techniques, skills, methods, and processes used in the production and services are considered as technological applications. Progress in production and services could be achieved through technological innovations that would make efficient use of limited resources. In the context of economic growth and development, a country should use suitable technologies to achieve development goals by having a proper understanding of the demand, the supply, and the market mechanism. Appropriate technologies mean the technologies which are suitable to the social and economic conditions of the geographic area in which those can be applied, are environmentally sound, as well as promote self-sufficiency. Thus, considering agricultural development in a country like Sri Lanka, small-scale, locally affordable, decentralized, labor-intensive, energy-efficient technology applications, environment friendly and locally autonomous are much more beneficial. Thus, the challenge is to work for appropriate technologies while checking such effective applications in the agriculture sector around the world. The rural farming sector in the United States provides good examples that could be studied.”

**Dr. Saman Bandara,**
PhD (NRE) , MBA (Fin)
Associate Professor of Agribusiness and Economics
University of Mount Olive
Office: 116, Glover Hall
janaranjanaherath@gmail.com

# Keynote Speech – “Robotics and AI for the Benefit of Humanity”

“Human ingenuity has shown a knack for automation of tasks for many centuries. Robots are a natural progression of this trend. Modern day robots go beyond the automation of repetitive tasks that were commonly assigned to robots such as work on factory assembly lines. There are many instances in various industries and scenarios where humans have to put themselves in harm’s way, either to perform a routine task or to respond to an emergency. Industrial examples include humans having to enter extremely confined spaces such as the wing cavity of an aircraft and inspecting sewerage pipes or working at extreme heights such as power line inspection and building construction. Other examples come from emergency response and underground mining where humans have to enter environments that potentially have hazardous gases. These types of jobs are typically classed as ‘Dull’, ‘Dirty’ and ‘Dangerous’ (Three ‘D’s) jobs and would benefit from the introduction of robots who can remove humans from harm’s way.

There are some anxieties in society about ‘robots taking away our jobs’. Similar to what happened with the industrial revolution, with the advancement of robotics and AI, the way we work is changing. This will continue with the improvements made in hardware and software going into the future. As with any technological development, negative effects are inevitable. However, the potential benefits far outweigh the drawbacks. In reality, completely new industrial sectors would be created as a result of adopting robotics and automation going into the future.”

**Dr. Navinda Kottege**
PhD (Australian National University)
CSIRO Queensland Centre for Advanced Technologies (QCAT)
1 Technology Court, Pullenvale
QLD 4069
Australia
Navinda.Kottege@data61.csiro.au

## Conference Tracks

**Track A:**

Advancements in Applications of Computing Technology

**Track B:**

Advancements in Applications of Engineering and Science in Technology

**Track C:**

Advancements in Applications of Technology in Teaching and Learning

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index N0: CT-01-29*

# Convolutional Neural Network for Classification and Value Estimation of Selected Gemstones in Sri Lanka

Samali Amarasekara (1st Author)
*Apple Research and Development Centre, Department of Computer Science, Faculty of Applied Sciences, University of Sri Jayewardenepura,*
Nugegoda, Sri Lanka.
samali.r.amarasekara@gmail.com

Ravinda Meegama (2nd Author)
*Apple Research and Development Centre, Department of Computer Science, Faculty of Applied Sciences, University of Sri Jayewardenepura,*
Nugegoda, Sri Lanka.
rgn@sjp.ac.lk

***Abstract*— Gemstone classification and value estimation are considered to be tedious tasks encountered in the gem industry all over the world. This happens due to colour variations found in the same gem type which is often difficult to detect with the naked eye. This paper presents a machine learning approach to automatically classify the gem type by using an image and also to estimate the value of the stone using a few measurements. The proposed technique uses a microscopic image of a gemstone which is taken using a gemological microscope. A Convolutional Neural Network (CNN) is trained to classify gem type while features such as type, colour palette, shape and weight are used to estimate the value of the. This work creates a system that is capable of classifying and estimating the value of four types of gemstones, namely, Blue Sapphire, Yellow Sapphire, Amethyst and Cat's eye. The results indicate that the proposed technique managed to classify the gemstones with the highest accuracy of 87% for yellow sapphires and 77% for blue sapphires. The yellow sapphires produced the highest accuracy in colour categorization which can be attributed to the high contrast of the images available. As such, it can be concluded that the quality of the original image is important in correctly identifying the exact colour of a gemstone.**

***Keywords - Gem classification, gem value estimation, artificial intelligence, machine learning***

## I. INTRODUCTION

The gem industry in Sri Lanka is reported to be existent for over 2500 years consisting of about thirty popular and rare gem stones [1]. Over the past decades, only people who have gathered sufficient experience from their generations in this industry can categorize and make a valuation by examining the gem using the naked eye and a few tools and therefore,the value of a gem is highly dependent on the person who examines the stone and often varies by a substantial margin when valuations are made by different persons. The National Gem and Jewelry Authority [2] is a state institute in Sri Lanka that provides a credible valuation for a gem and clients from other countries depend on this entity to obtain the correct value of a gem.

While some gemstones such as blue sapphires, pink sapphires, yellow sapphires, ruby, spinel and chrysoberyl show extreme colour variations from each other, gemstones such as blue sapphires, blue spinels, tanzanites and aquamarines show minor variations of the same colour. As shown in Fig. 1, the blue sapphires have about seven colour categories [3]. This problem of varying valuations can be eliminated by modern technology and automated tools that utilizes machine learning techniques to classify and make a valuation for a particular gemstone.

## II. RECENT WORK

A labeling technique for amber gemstones have been suggested where image processing and machine learning are used for identification [4]. Amber pieces are identified and labeled to one of 30 colour classes or one of 20 geometric shape classes. Experimental results have shown that the technique is effective for inorganic shape classification within a selected geometric shape seven if there is high ambiguity between organic shapes.

A method is presented for detecting synthetic gemstones using image processing where the gemstone is illuminated using a laser beam to produce a reflection/refraction pattern on a screen [5]. Then, an image of the reflection pattern is captured to determine the plurality of white spots. Further, the coordinates of these white spots are used to determine whether the gemstone is synthetic or not.

A study is carried out to identify three types of gemstone, namely, ruby, sapphire and emerald using the HSV colour space, image processing techniques and an Artificial Neural

Network (ANN) based backpropagation algorithm that learns by examples [6].

A machine vision system is developed to automate grading of opals, referred to as gemological digital analyzer (GDA) which analyzes the opal images to extract a summary of the flash, body tone, and other characteristics to automatically classify the opal into one of several opal classes [7]. The experimental results have shown that the opals in the training set could be classified with over a 90% correct classification rate.

An invention to automate gem evaluation and reporting is presented that uses predefined criteria for the estimation [8]. The relationship between the selling price of diamonds and their weight in carat is investigated by collecting more than 100,000 certified data about diamonds from the site: www.infodiamond.com [9].

To estimate the correct value of a gem, we mostly consider the category, colour, clarity, cut, carat, rarity and weight [10]. However, as of today, only a few experiments have been carried out in the gem industry in Sri Lanka for gemstone identification and value estimation. The research presented in this paper proposes a method to identify, categorize the colour and subsequently, compute a value for a given gemstone.

## III. METHODOLOGY

Although Sri Lanka has approximately twenty gem types, it is difficult to obtain high-resolution images of good quality stones.The scope of this research is limited to identifying only four types of gem types, which are, Amethyst, Blue Sapphire, Cat's eye and Yellow Sapphire as depicted in Table I. The most tedious task of this research is to collect a data set that can effectively be used to train the neural network models. The microscopic images, often given with a verification issued to clients who wish to purchase the stones, were collected from the National Gem and Jewelry Authority of Sri Lanka. This image data is used to obtain accurate information about the colour, shape, weight and price of each stone.

Fig. 2 shows different colour palettes attributed to each gem type. In this study, the images are separated into three colour palettes as dark, middle and light.

The flow chart in Fig. 3 provides the stages involved in training CNN and the regression model using Python programming and Keras libraries. Some researchers have performed color categorization of the amber gemstone using pruned decision tree classifier in which the mean, standard deviation, kurtosis, and skewness are calculated on amber pixels from grayscale and HSV color spaces were selected for the classification [20]. This technique requires a pipeline of routines starting from pre-processing and subsequent morphological operations to extract features of interest and as such, the final output is heavily biased towards image processing algorithms used in the said sub-tasks resulting in an overall accuracy of about 72%.


Fig. 1: Color variations of blue colour gems (top) and blue sapphire (bottom)

The following steps are used to train the CNN model11 given in Fig. 4 to obtain four classes of gemstones.

i. Images having a 2D resolution of 64x64 are used for training.
ii. Applied MaxPoolingwith a pool size of (2,2).
iii. Applied fully connected layer of 128 neurons and activation Relu (Rectified Linear Unit)
iv. Applied another fully connected layer of 4 neurons and activation Softmax.
v. Applied image augmentation methods with (shear, zoom and scale).
vi. Trained the model for 250 epochs.

Table 1. Four types of gemstones and the collected sample size of each.

| Class | Image | Gem name | # of images |
|---|---|---|---|
| 1 |  | Amethyst | 82 |
| 2 |  | Blue Sapphire | 85 |
| 3 |  | Cat's eye | 65 |
| 4 |  | Yellow Sapphire | 80 |


Fig. 2: Classified colour palette of gem types

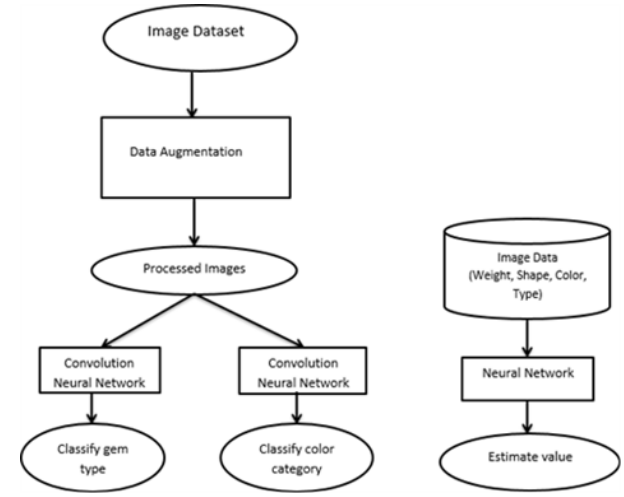



Fig. 3. Flow chart of training the neural network used to train the model.

After the images are classified into 4 classes, it is required to perform colour categorization for each class to determine whether each image belongs to dark, mid or light colour. The following steps are used to train the CNN model [12] as in Fig. 5 for colour categorization:

i. Images having a 2D resolution of 64x64 are used for training.
ii. Applied MaxPooling with a pool size of (2,2).
iii. Applied fully connected layer of 128 neurons and activation Relu.
iv. Applied another fully connected layer of 3 neurons and activation Softmax.
v. Applied image augmentation method with (shear, zoom and scale).
vi. Trained the model for 250 epochs.

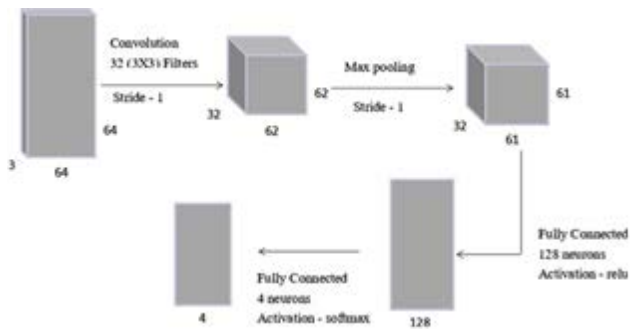



Fig. 4. CNN model structure for gem classification into four classes

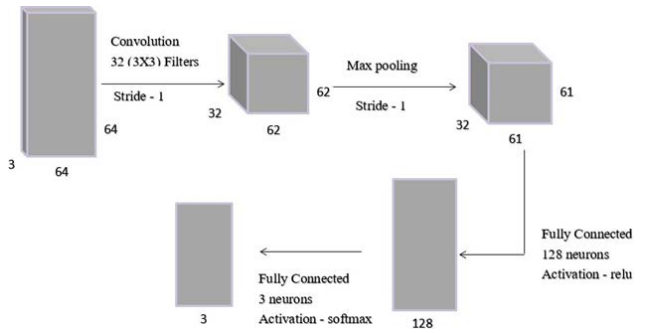



Fig. 5. CNN model structure for gem categorization into three colours.

To estimate the value, the type, shape, colour category and the weight of the gem are considered as inputs to the data pre-processor that uses Min-MaxScaler and Label encoder. Using a quarter of the data randomly for testing, the regression model with one hidden layer of 128 neurons and ReLU activation function is used for the inputs while the output layer is made up of 1 neuron and ReLU activation function [13]. As our task is a multiclass classification problem, we used a confusion matrix to determine incorrect classifications [14].

## IV. RESULTS AND DISCUSSION

A confusion matrix, as shown in Table II, is used to measure the performance of the classification model where the actual and predicated colour variation is indicated. In this study, a quarter of the dataset is randomly used for testing the model for 993 epochs. A confusion matrix is specifically used as the output may fall into several classes. As seen in Table II, gemstones with dark colour texture managed to come up somewhat closer to the predicted values.

The 30% testing dataset was selected randomly before starting the training. The accuracy of colour categorization in each of the four gemstone types is given in Table III. It is observed variations in accuracy over different gem types. The highest accuracy is exhibited for yellow sapphires while the blue sapphires exhibit the lowest. This can be attributed to the high contrast of the images available for yellow sapphires than others. The cat's eye gemstone too is uniquely identifiable from its sparkling spot and thus, produced a similar accuracy in colour categorization to the yellow sapphire.

The main objective of this study is to identify gem type and estimate the value. Related works on this study were not found for gemstones, especially for colour classification in yellow and blue sapphires. Our machine learning approach yielded good results in colour categorization of amethyst, blue and yellow sapphires while an accuracy of 77% was obtained for the cat's eye.

Table 2. Confusion matrix of gem categorization based on color (n indicates sample size).

| Gem type | Colour variation (actual) | Colour variation (Predicated) | | |
|---|---|---|---|---|
| | | Dark | Middle | Light |
| Blue sapphire (n = 30) | Dark | 6 | 2 | 0 |
| | Middle | 0 | 10 | 1 |
| | Light | 0 | 2 | 9 |
| Cat's eye (n = 22) | Dark | 3 | 1 | 1 |
| | Middle | 2 | 3 | 1 |
| | Light | 0 | 0 | 11 |
| Yellow sapphire (n = 44) | Dark | 13 | 0 | 0 |
| | Middle | 0 | 24 | 0 |
| | Light | 0 | 0 | 17 |
| Amethyst (n = 30) | Dark | 9 | 0 | 0 |
| | Middle | 0 | 6 | 4 |
| | Light | 0 | 0 | 11 |

Table 3. Accuracy of color categorization in four gem types.

| Gem type | Accuracy of colour categorization |
|---|---|
| Amethyst | 80% |
| Blue sapphire | 77% |
| Cat's eye | 86% |
| Yellow sapphire | 87% |

As gemstones are often transparent, human experts use color coverage, brilliance and dispersion to make an accurate evaluation and color categorization [15]. As there is no clear standard to obtain a valuation, certain gemstones are heat-treated to obtain color variations. Such treatment requires expertise knowledge in advance so that proper parameters can be set before heating because even a small error can make a yellow sapphire turn orange making it a raw stone that does not have any value. Availability of such precise treatment methods can be the reason for obtaining higher accuracy in colour categorization for yellow sapphires [16].

Colors of blue and yellow gemstones can also be evaluated by using fluorescence spectroscopy [17]. Using short and long pass filters, the gemstones can be screened for color categorization in both treated and un-treated stones. However, as spectroscopy results in color-bias, it is difficult to make a general identification rate for each gemstone type other than simply stating identification rates for individual colors for each gemstone. Also, Cat's eye and Amethyst gemstone types do not respond to fluorescence spectroscopy to produce a color separation resulting in less than 30% accuracy for classification.

Reflectance spectroscopy, coupled with an artificial neural network (ANN), can also be used for the identification of certain types of gemstones [18,19]. The reflectance spectral information of a gemstone can be fed into an ANN during the training stage. An inherent limitation of using such spectroscopy-based methods is that identification of gems of the same type, whether natural or synthetic, is not effective thus producing poor results. For instance, the yellow and blue sapphire exhibit several shades of the primary color and in such cases, using the reflectance spectroscopy will not guarantee an accurate result in the classification process.

The blue sapphires are produced across the world in a variety of colour saturations. The actual color of a natural blue sapphire is so complex that even with a human expert with considerable experience, classification tends to be a tedious task [21].

While a majority of blue sapphires produced, show a very dark colour that hides the actual radiant nature of the stone, the Sri Lankan blue sapphires exhibit bright colour due to their low iron content. The high iron content of the blue sapphires can result in low accuracy in colour categorization in the experiments. A previous study had used K-means algorithm for color categorization of blue sapphires producing accuracies of 98.4%, 97.9%, 97.8 %, 98.7 % and 99.0%, respectively of blue, cyan, light blue, dark blue, and gray blue. However, it does not produce an overall color categorization for the blue gemstone [22].

As our machine learning technique utilizes the colour pixel values of the images, images with low contrast produced lower accuracy in colour categorization. As such, the requirement to have high contrast images to obtain high accuracy results is a limitation of the proposed methodology. Further, the biggest obstacle that was encountered during the research is collecting approximately 300 high-resolution images of gems. As images of all the colour pallets of each gem are not available, the gem colours are divided into three categories as dark, medium and light using expert knowledge.

## V. CONCLUSION

To arrive at an approximate estimate of a gemstone, an appraiser would manually review and analyze each gemstone separately and generate a report indicating the value. As this process is time-consuming, the proposed technique allows us to automate this task of estimating the value using a machine learning approach. Although there are more than seventy varieties of gems in Sri Lanka, approximately twenty types of stones are found commonly in the industry. For experiments conducted in this paper, only four gem types are considered which are widely available in the local industry. The quality of the original image is important in correctly identifying the exact colour of a gemstone. Extending the present research for the identification and value estimation of other gem types is possible with a large number of data sets having good quality images.

## REFERENCES

[1] J.W. Herath, "Geology and occurrence of gems in Sri Lanka," The Journal of the National Science Council of Sri Lanka, vol. 12, no. 2, pp. 257-271, 1984.
[2] National Gem and Jewelry Authority of Sri Lanka, Available: http://www.ngja.gov.lk/en/.
[3] S. De Meo, A. Plutino and A. Rizzi, "Assessing Color of Gemstones," Color Research & Application, vol. 45, no. 2, pp. 224-234, 2020.
[4] F. Emmanuel and R.M. George, "An update on color in gems. Part 3: Colors caused by band gaps and physical phenomena," Gems & Gemology, vol. 24, no. 2, pp. 81-102, 1998.
[5] A.T. Collins, "The detection of colour-enhanced and synthetic gem diamonds by optical spectroscopy," Diamond and related materials, vol. 12, pp. 1976-1983, 2003.
[6] X. Yang, "Application and Development of Digital Image Processing Technology in Jewelry Design," International Conference on Machine Learning and Big Data Analytics for IoT Security and Privacy. Springer, Cham, 2020.
[7] D. Wang et. al., "Automated Opal Grading by Imaging and Statistical Learning," IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 46, no. 2, pp. 185–201, 2016.
[8] P.J. Malnekoff, "Automated gemstone evaluation system," 2011. US6304853B1161 [Online]. Available from: https://patents.google.com/patent/US6304853B1/en [Accessed:11-Feb-2020]
[9] N.G. Vaillant and W. François-Charles, "Understanding diamond pricing using unconditional quantile regressions," Hal archives, 2013, halshs-00853384f. https://halshs.archives-ouvertes.fr/halshs-00853384/
[10] M. Januz, "Application of Neural Networks in Regression Problems with Qualitative Data on an Example of Gemstone Valuation," Scientific Conference on Social Sciences and Arts SGEM, 2018. DOI: 10.5593/sgemsocial2018/1.5/S05.010
[11] T. Guo, J. Dong, H. Li and T. Gao, "Simple convolutional neural network on image classification," In: IEEE 2nd International Conference on Big Data Analysis (ICBDA), pp. 721-724, 2017, DOI: 10.1109/ICBDA.2017.8078730.
[12] X. Lei, H. Pan X. and Huang, "A Dilated CNN Model for Image Classification," IEEE Access, vol. 7, pp: 124087-124095, 2019.

[13] J. Xu, Z. Yuanjian and M. Duoqian, "Three-way confusion matrix for classification: A measure driven view," Information sciences. vol. 507, pp. 772-794, 2020.
[14] I. Maula, et al., "Development of a Gemstone Type Identification System Based on HSV Space Colour Using an Artificial Neural Network Back Propagation Algorithm," In: International Conference on Science and Technology, and Technology for Development, 2017.
[15] M. Silvia, A. Plutino and A. Rizzi, "Assessing Color of Gemstones," Color Research & Application, vol. 45, pp. 224-234, 2020.
[16] K. Kanyarat, et al, "Heat Treating Blue Sapphire of Pailin, Cambodia," Chian Mai Journal of Science, vol. 45, no. 1, pp. 588-600, 2018.
[17] T. Tsung-Han and U. D'Haenens-Johansson, "Rapid Gemstone Screening and Identification Using Fluorescence Spectroscopy," Applied Optics, vol. 60, no. 12, pp. 3412-3421, 2021.
[18] Q. Jun-Ting, L. Qiu, and H. Mu, "Feasibility of Gem Identification Using Reflectance Spectra Coupled with Artificial Intelligence," Spectroscopy Letters, vol. 52, no. 9, pp. 520-532, 2019.
[19] D. Kun, et al., "Price Estimation Model of Jade Bracelet Based on RBF Neural Network," IOP Conference Series: Materials Science and Engineering, vol. 81, no. 81(1), 2018.
[20] S. Saulius, A. Lipnickas and K. Rimkus, "Automatic Amber Gemstones Identification by Color and Shape Visual Properties," Engineering Applications of Artificial Intelligence, vol. 37, pp. 258-267, 2015.
[21] J. Han, Y. Guo and S. Liu, "Environmental Issues on Color Quality Evaluation of Blue Sapphire based on GemdialogueTM Color Comparison Charts," Ekoloji, vol. 106, pp. 1365-1376, 2018.
[22] P. Aaron, "Coexisting rubies and blue sapphires from major world deposits: A brief review of their mineralogical properties," Minerals, vol. 10, no. 5, pp. 472, 2020.

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index N0: CT-02-31*

# Database Management System Deployment on Docker Containerization for Distributed Systems

W.M.C.J.T. Kithulwatta ( 1st Author)
*Faculty of Graduate Studies*
*Sabaragamuwa University of Sri Lanka*
Belihuloya, Sri Lanka
*Faculty of Technological Studeis*
*Uva Wellassa University of Sri Lanka* Badulla, Sri Lanka
chirantha@uwu.ac.lk

K.P.N. Jayasena (2nd Author)
*Department of Computing and Information Systems*
*Faculty of Applied Sciences*
*Sabaragamuwa University of Sri Lanka*
Belihuloya, Sri Lanka
pubudu@appsc.sab.ac.lk

B.T.G.S. Kumara (3rd Author)
*Department of Computing and Information Systems*
*Faculty of Applied Sciences*
*Sabaragamuwa University of Sri Lanka*
Belihuloya, Sri Lanka
kumara@appsc.sab.ac.lk

R.M.K.T. Rathnayaka (4th Author)
*Department of Physical Sciences and Technology*
*Faculty of Applied Sciences*
*Sabaragamuwa University of Sri Lanka*
Belihuloya, Sri Lanka
kapilar@appsc.sab.ac.lk

***Abstract*—Containerization is a novel technology that brings an alternative for virtualization. Due to the most infrastructure-based features, most computer system administration engineers use Docker as the infrastructure level platform. On the Docker containers, any such kind of software service can be deployed. This study aims to evaluate Docker container based relational database management system container behavior. Currently, most scholarly research articles are existing for the database engine performance evaluation under different metrics and measurements of the database management systems. Therefore, without repeating them: this study evaluated the data storage mechanisms, security approaches, container resource usages and container features on the launching mechanism. According to the observed features and factors on the containerized database management systems, containerized database management systems are presenting more value-added features. Hence containerized database management system Docker containers can be recommended for the distributed computer systems for getting the benefit of effectiveness and efficiency.**

***Keywords - containers, database management systems, distributed Systems, Docker, MySQL, PostgreSQL***

## I. INTRODUCTION

Virtualization is an old concept that provides on premise or cloud-based virtual machines to deploy any such software applications or system services. Virtual machines provide the facility to optimize the host server capacity by launching multiple different operating systems. Virtualization carries additional overhead since one virtual machine consists of a fully installed operating system. To optimize the whole virtual machine process, reduce the virtual machine weight and enhance the infrastructure performances: an alternative technology called container virtualization arrived.

Within the containerization, containers allow to deploy and run software applications and services without using or creating separate virtual machines. By sharing the host computer operating system kernel, separated multiple containers are executing on the infrastructure. To execute any software application or service, all necessary software dependencies, libraries and binaries are packaged into each container [1].

For the secure execution of the containers, basic Linux features are used for the containers. Those are *cgroups*, *chroot* and *namespaces*. Since containers are not using full operating system instances, containers require less CPU, memory and storage capacity according to the fundamental theory of container virtualization [1]. Fig. 1 presents the container architecture as a graphical notation.

Fully packaged independent containers are running on own container engine. Each container consists of its own independent subsystem for the file system, memory and network. The container engine is the component that has the authority to manage containers. Containers of the same container engine share the same host operating system. Therefore, the infrastructure supports launching a massive number of containers on a single operating system [2].

Within the practitioner of containerization, various container management technologies are available. Docker, Rkt and Linux containers are a few container management technologies [1]. Among them, Docker is the most trending and most popular container management technology [3]. According to the official Docker documentation, currently, eleven million developers are engaged with Docker and thirteen billion of Docker images have been downloaded [4]. Database Management Systems are the specific software packages that provide the dedicated technology and facility to store and retrieve data in efficiently and appropriately [5]. A Database Management System stores data in a most prominent way to retrieve, manipulate, manage and produce information.

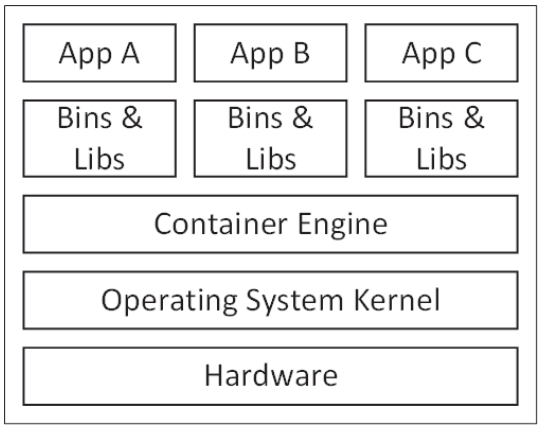



Fig. 1: Container architecture [2]

Relational Database Management Systems (RDBMS) are a specific database management system specification, which are based on the relational model and Structured Query Language (SQL). Most modern database systems are RDBMSs. MySQL, PostgreSQL, IBM DB2, MS SQL Server and Oracle are the best examples for the RDBMSs [6].

Within the existing research studies, the authors have evaluated the database management systems by considering the taken time to particular SQL queries and response time in commonly.

Currently, there is a trend to shift software services and applications to the container-native methodology. Hence database management systems are also shifting to the containers. There are a lot of scholarly research articles to evaluate and measure the performances of databases and database management systems. Therefore, this research activity was aimed to do an experimental study on the database management system deploying infrastructures for the distributed systems.

The overall research study provides answers to the below research questions.

**RQ1:** How to mechanize the Docker based Database Management Systems to have a persistence data storage approach?

**RQ2:** What kind of container-based infrastructure level security approaches can be applied to the Docker based Database Management Systems?

**RQ3:** How are the container resource usage and utilization from the host computer infrastructure for Database engine activities?

**RQ4:** What are the differences among manually deployed Database Management System containers over ready-made Docker image approaches?

## II. METHODOLOGY

For the study, the experimental platform was launched on a cloud-based Linux environment. To launch the Docker service, an Ubuntu computer host was used. Canonical Ubuntu 18.04 operating system was used for the host computer infrastructure. Host computer-based architecture was GNU/Linux 4.15.0-112-generic x86_64. As well, the host computer was with two virtual CPUs, 15 GB internal memory and 1 Gbps network bandwidth. An external block volume was attached to the host computer. The directory path of the host computer: ***/home/$user*** was mounted to the block volume.

On the above-mentioned host computer, the Docker platform was launched. The launched Docker configurations are mentioned in table 1. (Within table 1, these abbreviations are used for the easiness of the representation: API = Application Programming Interface, OS = Operating System)

It presents the Docker version, Docker API version supporting based operating system architecture.

Table 1: Configured Docker details

| **Option** | **Detail for configuration** |
|---|---|
| Client: Docker Engine - Community | |
| Version | 19.03.9 |
| API version | 1.40 |
| OS/Architecture | Linux/amd64 |
| Server: Docker Engine - Community | |
| Version | 19.03.9 |
| API version | 1.40 |
| OS/Architecture | Linux/amd64 |

Within the above Docker platform, two internal Docker networks were established using the bridge drivers. One network ($N_1$) was with 172.17.0.0/16 as the subnet and 172.17.0.1 as the gateway. The second network ($N_2$) was with 172.22.0.0/16 as the subnet and 172.22.0.1 as the gateway. For the experiment study, two RDBMSs were used: MySQL and PostgreSQL. Those two RDBMSs were launched on two separate Ubuntu bionic Docker containers within the above mentioned two Docker networks. Table 2 presents the infrastructure details for the distributed RDBMS Docker farm. The term IP was abbreviated for the internet protocol.

Table 2: Docker container details

| Container | Network | Container IP | Container port | Host port |
|---|---|---|---|---|
| MySQL conatiner | $N_1$ | 172.17.0.2 | 3306 | 3300 |
| PostgreSQL container | $N_2$ | 172.22.0.2 | 5432 | 5400 |

For the experimental study, MySQL version 5.7.30 and PostgreSQL version 12.6 was used. For the MySQL RDBMS, a database with 24 tables was used. For the PostgreSQL RDBMS, a database with 30 tables was used.

For each container, the specific data and file paths were mounted to the path, ***/var/lib/docker/volumes*** on the host computer. By default, MySQL and PostgreSQL RDBMSs are serving through the port 3306 and 5432. Within the experimented approach, each port was mapped to 3300 and 5400 respectively for the host ports. By performing data retrieval operations on the RDBMSs, the host container resource usage and utilization was evaluated. To launch each RDBMS Docker container, two approaches were followed. Each approach is defined below, and each approach was evaluated with their behaviors.

Approach 1: Launch Ubuntu bionic Docker container and install the respective RDBMS service (Used as the main approach for the study)

Approach 2: Launch RDMS Docker container using Docker images/templates.

## III. RESULTS AND DISCUSSION

After launching the experimental platform with the RDBMS Docker containers, specific operations and analyzing the proposed disturbed mechanism, was performed.

### *A. Persistence data storage/archiving*

The Docker container-based platform was launched on a host computer infrastructure. To keep a more data persistence, the main directory path of the Docker ***(/var/lib/docker/)*** was linked to the path, ***/home/$user*** (to the block volume path) of the host computer infrastructure. Hence, specific objects and configurations of the Docker could be attached to the block volume. Those very specific Docker objects and configurations were container, images, volumes, network, Docker swarm, plugins, temporary files, etc. Therefore, as the primary mechanism, the whole Docker based infrastructure was with persistence data storage mechanism.

Furthermore, to keep a persistence data storage and archiving approach for each container separately, data volumes were mounted to each container. For the MySQL RDBMS Docker container, the paths ***/var/lib/*** and ***/var/log/*** are carried more specific data and configurations for the database engine. Hence those two directory paths were mounted to Docker data volumes. To mount those two paths, two different approaches were used, and those two approaches are defined below.

- Both ***/var/lib/*** and ***/var/log/*** directory paths were mounted to one Docker data volume.
- For ***/var/lib/*** and ***/var/log/*** data directory paths were mounted two separate Docker data volumes.

Between above two approaches, mounting two Docker data volumes was more strategic since if a Docker volume crashed, it does not affect the rest of Docker data volume.

Same as above, for the PostgreSQL RDBMS Docker container, the data directory path, ***/var/lib/postgresql/*** was identified as carrying most key data and configurations of the service. Hence, the identified directory path was mounted to a Docker data volume.

After making a stable database management system on two Docker containers, two Docker containers were crashed by stopping the containers and jamming with installing different unwanted packages and dependencies. Thereafter, respective data volumes were re-attached for the new Docker containers which carry the RDBMS service. Then without losing any data or configurations, the new Docker container was restored to previous data and configurations.

Without detaching the previously attached Docker data volumes from containers, attachments for new Docker containers were possible. But assignment of the previous host port to the new Docker container's host port was not possible even if the previous container was stopped on the Docker engine. Therefore, essentially, previous container needed to be removed to assign the host port for the new Docker container same as the previous container.

Those all-mounted Docker data volumes were directly linked with host computer infrastructure. Hence any file or directory of those data paths, could be faced for any such operation on the file system (copy/rename/delete/move).

### *B. Secured approchess for the infrastructure*

For the experimental setup, each container port was mapped for a host port. For the MySQL RDBMS Docker container perspectival: 0.0.0.0:3306 → 3300/tcp and for the PostgreSQL RDBMS Docker container perspectival: 0.0.0.0:5432 → 5400/tcp was applied as the port mapping. As a usual practice, attacks or vulnerabilities are looking for default ports of any services. Hence by mapping each service port (same as container port in this case) was mapped to arbitrary port value. Hence guessing port values are not possible for this case and the proposed approach is bringing a valuable security concern for the infrastructure.

As well, the default service port of each container was changed for arbitrary values by changing the service installation configurations at the next stage. Then newly assigned ports were mapped for a new set of host ports. Therefore, default port values are not available for any of the layers. Hence, guessing port values from an external person or device was mitigated.

Both RDBMS Docker containers are in two isolated Docker networks. By using static network address translation (NAT) for each multiple private IP addresses of the containers, all inward and outward traffic was handled in a more securely. Without exposing the container IP

address to the outside world, the public IP address of the host computer was exposed to the outside world. Therefore, without translating the container IP addresses to the external IP addresses, internal IP addresses could not be routed to the external world. Furthermore, the NAT mechanism was assured that all outbound traffic is from valid and known external IP addresses. Therefore, the approach helps to enhance the infrastructure security, all incoming and outgoing requests go through a translation process. The process ensures to qualify and/or authenticate all incoming traffic.

As described in section III.A, the main data and file directory of Docker was linked to the path: ***/home/$user.*** That directory path was privileged only for the super user. Therefore, without any command or operations could not be done for that directory path without super user credentials. Hence that ensures the security of the approach.

### *C. Container resource usage & utilization*

To measure and evaluate the internal resource consumption of each RDBMS Docker container from the host computer infrastructure, main resource metrics were measured for the idle state and data operating states.

Table 3 presents primary details of each Docker container and resource usage from the host computer infrastructure. For the ease of documentation purposes, below abbreviations were used for table 3. [ Container ID = the unique identifier or the Docker container within the launched Docker platform, Name = assigned container name, CPU% = the percentage of the container consuming CPU from the host computer, Memory% = the percentage of the container consuming memory from the host computer, MEM_Usage/Limit = total memory which is used by the container and the allowed total memory to use, NET I/O = the overall amount of the data which the container has sent and received over the network interface and PIDs = created the total amount of the processors or threads by the container ][7].

Table 3: Container resource usage and utilization

| **Container measurement** | **Measurement value of each container** | |
|---|---|---|
| | **MySQL container** | **PostgreSQL container** |
| Container ID | c35b95254633 | db7f7cee8390 |
| Name | MySQL Container | PGSQL Container |
| CPU% (for idle states) | 0.25% | 0.38% |
| Memory% | 0.07% | 0.15% |
| MEM_Usage/Limit | 10.51MiB / 14.68GiB | 21.85MiB / 14.68GiB |
| NET I/O | 12.9GB / 2.23GB | 98.8GB / 5.85GB |
| PIDs | 50 | 8 |

The table 4 presents how the container host computer infrastructure is behaving for host resources while running Docker engine and other embedded services.

Table 4: Host computer resource usage for container host

| **Host computer measurement** | **Measurement value of single host computer** |
|---|---|
| CPU% (for idle states) | 0.9% |
| Memory% | 0.4% |
| PIDs | 197 |

For further evaluation purposes, the same MySQL and PostgreSQL RDBMS services were launched on a separate computer instance. For that computer instance was with the same configurations of the Docker hosted computer infrastructure. As well, the same databases were used for the computer instance-oriented study. This case was named as the β case. The table 5 presents the measurement values for the computer host for the β case for the idle state of the RDBMSs.

Table 5: Host computer resource usage for case β

| **Host computer measurement** | **Measurement value of single host computer** |
|---|---|
| CPU% (for idle states) | 2.1% |
| Memory% | 2.7% |
| PIDs | 124 |

According to table 4 and table 5, Docker host computer infrastructure is consuming lower resource usage from the host computer. However, when compared with Docker host and case β host, slightly lower usage is for Docker host.

### *D. Docker conatiner expanding & shrinking*

Docker containers are running by using minimal resources from the host computer infrastructure. For any such kind of heavy processes, Docker containers consume higher resources from the host computer infrastructure. Within the experimental study, for data retrieve operations, container expanding and shrinking was visualized.

For the MySQL RDBMS container, four hundred forty-eight thousand data records were retrieved. The fig. 2 presents CPU usage: before data retrieval, while data retrieving and after retrieving the data.

In the fig. 2, the x-axis presents the time in the GMT +5.30-time zone and the y-axis presents the CPU usage as the percentage.

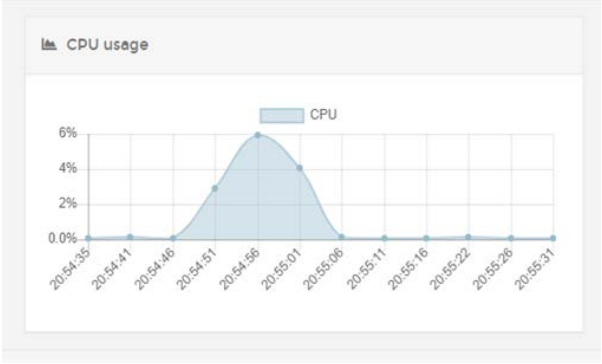



Fig. 2: CPU usage for data retrieving for MySQL container

Up to the peak point of the graph (20:54:56), the container was expanded for the operation of data retrieval. After generating the results, the container was shrunken.

For the PostgreSQL RDBMS container, three hundred thousand data records were retrieved. The fig. 3 presents the CPU usage: before data retrieval, while data retrieving and after retrieving the data for the PostgreSQL Docker container.

In fig. 3, the x-axis presents the time in the GMT +5.30-time zone and the y-axis presents the CPU usage as the percentage.

Up to the peak point of the graph (21:57:43), the container was expanded for the operation of data retrieval. After generating the results, the container was shrunk the same as for the MySQL Docker container.

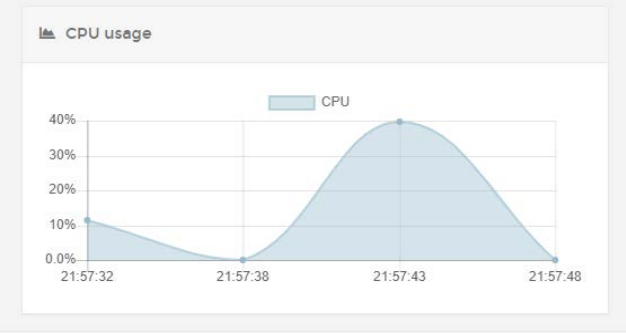



Fig. 3: CPU usage for data retrieving for PostgreSQL container

According to the fig. 2 and fig. 3, the two graphs are with few differences due to the internal architectural differences of database management engines.

Most scholarly articles were presented that the Docker containers are shrinking and expanding while shipping and operating on the container engine. Hence above graphical representation of the container expanding and shrinking are the most theoretical proof of the container stretching feature.

### *E. Feature differentiate between Approach 1 & 2*

Within the experimental evaluation, to launch the RDBMS Docker container, two approaches were applied. Those are presented in Approach 1 and Approach 2 under section II. The table 6 presents the feature to differentiate between the launched container approaches for the Approach 1 & 2.

Table 6: Feature difference between approach 1 & 2

| **Approach 1** | **Approach 2** |
|---|---|
| Difficult to lauch | Easy to launch |
| Easy to cutomize the installation | Customized installation is difficult |
| Need to install external dependencis | No need to install external dependencies |

## IV. CONCLUSION

Docker containers are high trending computer infrastructure technology. To launch any software service or application is possible on the Docker based infrastructure. After launching the Docker engine on the Ubuntu host computer, MySQL and PostgreSQL RDBMS Docker containers were launched within two networks separately to obtain answers to the pre-defined research questions.

The experimental platform was with host computer perspective, Docker engine perspective and container perspectival different data storage mechanisms. Those are with mounting block volumes for the host computer and Docker data volumes for Docker containers. Therefore, those aspects are answered for the RQ1. The experimented setup was with different secured approaches for ports, network and data directory perspectival. Therefore, it denotes that, the established platform is with more secured approach. Hence those are answered for the RQ2.

For the Docker container perspectival, those containers used only limited and minimal resources from the host computer infrastructure. As well, only for the higher operations, the containers expanded and other idle states, containers were shrunken. Therefore, those are answered for the RQ3. To deploy service on Docker, image usage or deploying the service from scratch on an operating system container are possible. Most of the pros and cons are available for both mechanisms and those are answered for the RQ4.

Docker containers are currently used mostly for testing and deployment of software applications. But today the world is moving with data science, image processing, artificial intelligence, Internet of Things and etc. for containerization. Therefore, containers will play a major role in the Information Technology era.

## REFERENCES

[1] John Paul Martin, A. Kandasamy, and K. Chandrasekaran. 2018. Exploring the support for high performance applications in the container runtime environment. Hum.-centric Comput. Inf. Sci. 8, 1, Article 124 (December 2018), 15 pages. DOI:https://doi.org/10.1186/s13673-017-0124-3

[2] B. I. Ismail et al., "Evaluation of Docker as Edge computing platform," 2015 IEEE Conference on Open Systems (ICOS), 2015, pp. 130-135, doi: 10.1109/ICOS.2015.7377291.

[3] 8 surprising facts about real Docker adoption, 2021. [Online]. Available: https://www.datadoghq.com/docker-adoption/. [Accessed: 25- Jun- 2021].

[4] Empowering App Development for Developers | Docker, Docker, 2021. [Online]. Available: https://www.docker.com/. [Accessed: 25- Jun- 2021].

[5] Database Management System Tutorial - Tutorialspoin, Tutorialspoint.com, 2021. [Online]. Available: https://www.tutorialspoint.com/dbms/index.htm. [Accessed: 25- Jun- 2021].

[6] SQL - RDBMS Concepts - Tutorialspoint, Tutorialspoint.com, 2021. [Online]. Available: https://www.tutorialspoint.com/sql/sql-rdbms-concepts.htm. [Accessed: 25- Jun- 2021].

[7] docker stats, Docker Documentation, 2021. [Online]. Available: https://docs.docker.com/engine/reference/commandline/stats/. [Accessed: 25- Jun- 2021].

[8] F. Paraiso, S. Challita, Y. Al-Dhuraibi and P. Merle, "Model-Driven Management of Docker Containers," 2016 IEEE 9th International Conference on Cloud Computing (CLOUD), 2016, pp. 718-725, doi: 10.1109/CLOUD.2016.0100.

[9] J. Stubbs, W. Moreira and R. Dooley, "Distributed Systems of Microservices Using Docker and Serfnode," 2015 7th International Workshop on Science Gateways, 2015, pp. 34-39, doi: 10.1109/IWSG.2015.16.

[10] Peinl, R., Holzschuher, F. & Pfitzer, F. Docker Cluster Management for the Cloud - Survey Results and Own Solution. J Grid Computing 14, 265–282 (2016). https://doi.org/10.1007/s10723-016-9366-y

[11] D. Liu and L. Zhao, "The research and implementation of cloud computing platform based on docker," 2014 11th International Computer Conference on Wavelet Actiev Media Technology and Information Processing(ICCWAMTIP), 2014, pp. 475-478, doi: 10.1109/ICCWAMTIP.2014.7073453.

[12] M. T. Chung, N. Quang-Hung, M. Nguyen and N. Thoai, "Using Docker in high performance computing applications," 2016 IEEE Sixth International Conference on Communications and Electronics (ICCE), 2016, pp. 52-57, doi: 10.1109/CCE.2016.7562612.

[13] M. G. Xavier, I. C. De Oliveira, F. D. Rossi, R. D. Dos Passos, K. J. Matteussi and C. A. F. D. Rose, "A Performance Isolation Analysis of Disk-Intensive Workloads on Container-Based Clouds," 2015 23rd Euromicro International Conference on Parallel, Distributed, and Network-Based Processing, 2015, pp. 253-260, doi: 10.1109/PDP.2015.67.

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index N0: CT-03-37*

# Hybrid Movie Recommendation System

Sinthujan Punitharasa (1st Author)
*Faculty of Information Technology*
*University of Moratuwa*
Moratuwa, Sri Lanka
sinthujanspp@gmail.com

Kirisanthi Selvanajagam (2nd Author)
*Faculty of Information Technology*
*University of Moratuwa*
Moratuwa, Sri Lanka
kirisanthiselvanajagam@gmail.com

Thamilini Ramakrishnan (3rd Author)
*Faculty of Information Technology*
*University of Moratuwa*
Moratuwa, Sri Lanka
thamiliniram96@gmail.com

Amalraj Chelvarajah (4th Author)
*Faculty of Information Technology*
*University of Moratuwa*
Moratuwa, Sri Lanka
amalraj@uom.lk

W.M.R.M. Weerasinghe (5th Author)
*Faculty of Information Technology*
*University of Moratuwa*
Moratuwa, Sri Lanka
ruviniw@uom.lk

***Abstract*- Movie recommendations play a great part in the aspects of our social life. Such a system allows users to recommend a group of films based on their interests or the popularity of the movies. This research was conducted to study different approaches to movie recommendation and discusses a hybrid approach that combines a content-based filter, a collaborative-memory-based filter, and a collaborative-model-based filter. The proposed system aims to reduce the issues with existing movie recommendation systems by enhancing performance. The content-based filter is based on a TF-IDF classifier with cosine similarity. The collaborative-memory-based filter is based on truncated SVD with Pearson correlation. A collaborative model-based filter is based on improved SVD matrix factorization.**

***Keywords- Movie Recommendation System; Content-based filtering; Collaborative- memory-based filtering; Collaborative- model-based filtering; Hybrid approach.***

## I. INTRODUCTION

Due to the growth of the Internet and Information Technology, the data transactions volumes that are happening every second have increased dramatically. Currently, movie releases per year have increased with the number of movie watchers. However, every movie watcher has different interests. Therefore, not all data related to movies available on the internet gives results that give satisfaction to the movie watchers. Movie data in such a large amount will often become irrelevant, and without proper processing of this information, it will be wasted. In such cases, movie watchers have to run their search several times before they get the first look. Movie watchers have to spend a lot of time finding the desired/ interesting movies. But due to workloads, they are not willing to spend much time searching. Researchers have come up with movie recommendations to overcome this problem. A movie recommendation system provides needed information by considering the movie watcher's past choices. Movie data is filtered and customized to suit the needs of the movie watcher.

The movie recommendation systems have proven to be a powerful tool in providing satisfactory movie suggestions for movie watchers. The recommendations are made to help movie watchers efforts to handle information loads and to help them find the right movies quickly and easily [1]. It is a kind of filtration system, which tries to predict a movie watcher's preferences and recommends movies. As with movie recommendation systems, there are many other types of applications for recommendation systems. For example, Netflix can track users' interactions with various kinds of stories in users' feeds to find out what 2 kinds of stories/movies attract users. Sometimes, movie recommendation systems can improve their system by monitoring the activities of a large number of people. The movie recommendation system is an implementation of the machine learning algorithms [2, 3].

Movie Recommendation Systems are classified into four broad groups, namely Collaborative filtering systems, Content filtering systems, Knowledge-based systems, and Hybrid systems [4]. The Content-based filtering movie recommendation system works by assessing the behavioral needs of new movie watchers. This is a kind of keyword-specific recommendation system where specific keywords are used to identify movies. The collaborative filtering-based movie recommendation systems are based on similarity measures between the movie watcher's data/information and the movies. Movies are suggested to the new movie watcher by considering the other movie watcher's past browsing history. The collaborative filtering system uses the mean rating of the movies, recognizes similarities based on movie watchers' reviews, and generates new recommendations based on comparisons between movie watchers [5]. The knowledge-based movie recommender system attempts to recommend movies based on assumptions about a movie watcher's needs and preferences [6]. The hybrid movie

recommendation system performs its functions considering the integrated behavior of content-based and collaborative filtering techniques to suit a particular film. The hybrid movie recommendation system is considered the best movie recommendation system because of its ability to eliminate weaknesses. These hybrid movie recommendation systems usually build upon the programming languages like python / C++ [3].

## II. RELATED WORK

Over the past decade, several recommendation systems have been developed for various fields like books, cosmetics, novels, movies, and are in use. These recommendation systems are implemented based on a variety of approaches, including the content-based filtering approach, the collaborative-based filtering approach (memory-based & model-based), the geographic approach, the knowledge-based approach, the utility-based approach, and the hybrid approach.

### *A. Content-based filtering*

In movie recommendation, content-based filtering is a moderated learning algorithm [11]. This filtering technique gives recommendations for a movie by comparing the movie profile and the customer profile. Movies are recommended to users based on what the users like [11]. SRS Reddy et. al. proposed a content-based filtering system using genre correlation [12]. This system aims to recommend movies to users based on the similarity of genres. If the user rates highly for a particular movie, then this system recommends other movies with similar genres to that user. They used the MovieLens dataset in their research as and R is used as the data analysis tool. In this research, the main dataset is divided into two subsets. Its subgroups also have a list of movies with the genres in which they are classified. The other subset contains a list of user-rated movies on a scale of 1 to 5, with 5 being the highest [12].

Ramni Harbir Singh et. al. suggested a method for recommending movies to users using cosine similarity and KNN [13]. This system provides general recommendations for each user, based on popularity and/or the type of movie. They used cosine similarity and content-based filtering to predict the outcome and recommend a movie to the user by running the code in python using the NumPy and Panda libraries. This system can store a large amount of data and give effective results. This recommendation system searches for the best movies similar to the movie users watched based on genre and gives results. The researchers used cosine similarity to measure the similarity between movies based on different properties. Along with cosine similarity, they used KNN to find the nearest neighbor. It gives more accuracy than other distance metrics and the complexity is relatively low [13].

### *B. Collaborative Memory-based filtering*

Ramil G. Lumauag et. al. proposed an improved collaborative memory-based algorithm by formulating a similarity measure to identify the number of co-rated movies, calculate the user's similarity by selecting the nearest neighbor. They used the Movielens 100K dataset for their research. This dataset includes 943 users and 1682 movies and also it contains 100,000 rating records on a scale of 1 to 5 points, with 1 being the lowest and 5 being the highest. They found that the MAE of their improved algorithm has a small prediction error compared to the traditional collaborative filtering algorithm [14].

Kyung Soo Kim et. al., proposed a content metadata-based approach that uses content metadata efficiently. In their research, they collected a large amount of user movie ratings and user trust network information from Epinions.com and crawled movie metadata from the web. The dataset collected consists of 91,735 users, 611,741 user trust links, 26,527 movie content items along with their metadata, and 170,797 user content reviews. Furthermore, they provided two content representations that were generated from content metadata such as content-metadata TF-IDF vector and content link network. Content metadata was used to improve the process of finding live content for a target content element in the context of an object-based CF. Content metadata exploited with current user content rating scores and confidence in networks to predict classification results more accurately for target contents [15].

### *C. Collaborative Model-based filtering*

Rohan Mehter et. al. suggested a movie recommendation engine using collaborative filtering with two algorithms namely Alternative Least Square (ALS) and Single Value Decomposition (SVD) [16]. This movie recommendation system was once created with an Alternative Least Square algorithm and again with the Singular Value Decomposition algorithm. The researchers used Apache Spark, a machine learning library, to train the dataset more efficiently and with flexible search to retrieve results quickly. Researchers used the MovieLens dataset to recommend movies. This movie data set contains an anonymous movie rating. They have used Jupyter notebook for machine learning, data visualization, and code performance. They integrated the recommendation system using the Python programming language. Apache Spark is used to process big data while HDFS used to store a large amount of data which is processed with spark according to their study [16].

Mirza Ilhami et. al. proposed a system for recommending films using the Matrix factorization and collaborative filtering. They have used the MovieLens dataset in their research which contains 855,598 ratings for 10,197 movies and 2,133 users. They labeled the rating data as a three-way list representing user ID, item ID, and ratings. The researchers have split the data set into two parts, 80% for training and 20% for algorithm testing. The dataset contains actors, countries, directors, and genres. However, for

experiments, the researchers used the rating data from users to make recommendations. In this approach, first, normalize the data to remove bias. Second, perform a matrix analysis to obtain forecasts and recommendations, and it has also been used to remove element space dimensions and restore inherent relationships between dataset elements [17].

To overcome the limitations of the content-based filtration approach and the collaborative-based filtration approach, a hybrid approach was introduced. Rahul Katarya et al., proposed a hybrid and optimization-based technical suite to enhance the predictive accuracy of films. And this proposed solution helps to reduce the limitations of the collaborative filtering-based recommendation system. They use K-mean as the Clustering algorithm and Cuckoo search as an optimization algorithm. In their research, they apply these two algorithms to the MovieLens dataset to optimize the effective recommendation systems [18].

## III. PROPOSED SYSTEM

There are a lot of approaches to recommend movies to users. In our hybrid recommendation approach, we combine content-based filtering and collaborative filtering to provide better results to users. Fig.1 elaborates the top structure of the proposed system.

There are 3 main modules in this system named collaborative memory-based filtering, content-based filtering, and collaborative model-based filtering.

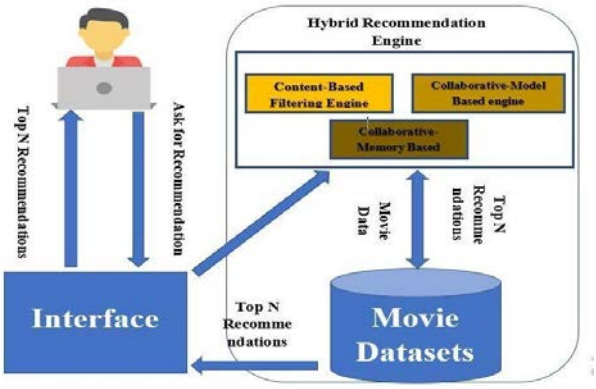



Fig. 1. Top Level Architecture for Proposed System

### A. *Content-based filtering*

The content-based filter uses data about movies, users and looks for similarities before recommending movies. The similarity of different movies computed based on user past and current data. There are different techniques or similarity measures used to calculate similarity. In this paper, we use the cosine similarity technique to find similar users. This scale is used to calculate textual similarity data. We convert this text movie data into vectors and check the cosine angle between these two vectors. If the angle between them is 0 or 1, then they are the same or not.

There are various concepts to build content-based filters. Term Frequency-Inverse Document Frequency (TF-IDF) is one of those concepts used in our study. Fig.2 shows the content-based filter approach.

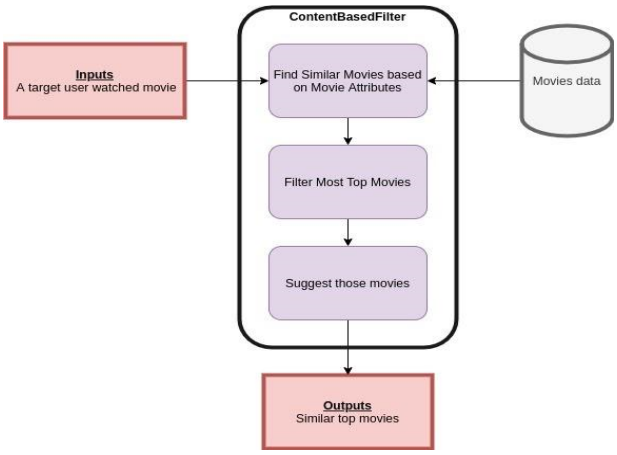



Fig. 2.Content-based filter

### B. *Collaborative- Memory-based filtering*

The memory-based collaborative filtering approaches can be further subdivided into two main parts namely user-based filter and item-based filter. User-based filtering takes a specific user, searches for users similar to that user, based on rating similarity, and recommends movies depends on similar users' interest. But, item-based filtering takes a pair of movies and finds the similarity of those movies by considering the rating of those pairs of movies by all users.

There are various concepts to build this module. Truncated SVD is one of those concepts which is used in our study. Unlike regular SVDs, a truncated SVD generates a factorization where the number of columns for multiple disconnections can be specified. It works with sparse matrices efficiently. Along with Truncated SVD, the Pearson correlation similarity technique is used to find similar items. Fig.3 shows the collaborative- memory-based filter approach.

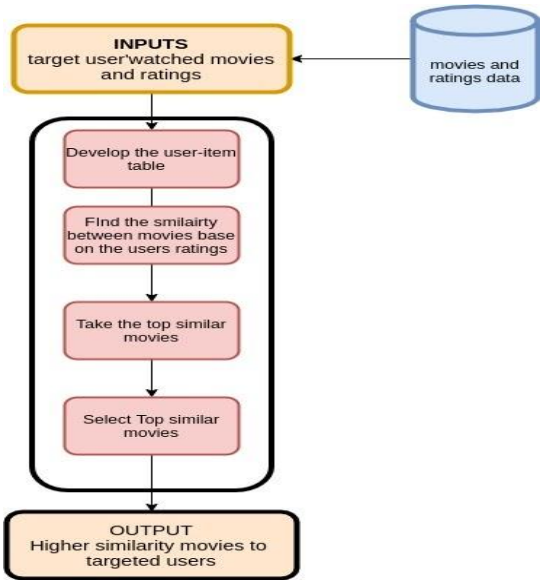



Fig. 3. Collaborative- Memory-based filter

### C. *Collaborative- Model-based filtering*

The model-based filter learns a user's preferences by directing itself on past user interactions and creating a user profile. This model-based approach considers a basic "generative" model, which explains user-movie interactions and attempts to detect them to make new predictions. This model is trained to reconstruct the values of user movie interactions from the own representation of users and movies. New movie recommendations can be made based on this model.

There are various concepts to build this module such as matrix factorization, clustering, and deep learning. In this paper, SVD, which is a classical method of linear algebra is used to implement a model-based filter. Matrix factorization works by decomposing a user-item interaction matrix into the product of two low-dimensional rectangular arrays. Not every user gives ratings to all movies, there are many missing values in the matrix and this results in a sparse matrix. SVD provides a solution for the sparsity problem. It uses a matrix framework where each row represents a user, and each column represents a movie. Elements of this matrix are the ratings that users give to movies. Fig.4 shows the collaborative-model-based filter approach.

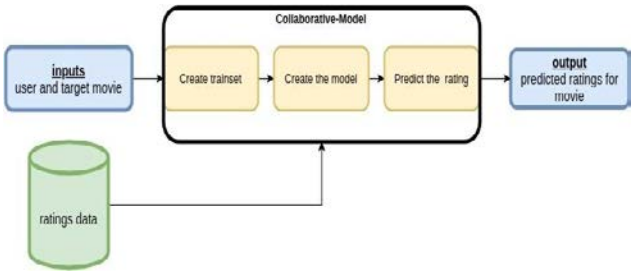



Fig. 4. Collaborative- Model-based filter

## IV. IMPLEMENTATION

To implement the proposed system, we used the TMDB dataset which combines movie plot description, Metastore ratings, comments and user ratings and reviews, release dates and many more aspects. In this data, we have four sub-datasets which are the movies dataset includes 45,000 movies with attributes such as movie description, average rating, number of votes, genre, etc. The rating dataset includes 26 million rating details from 270,000 users. 10000 movie data and 100000 rating data are used in our system.

### A. *Content-based filter*

Under the content-based filtering approach, the title, cast, crew, genres, keywords, and the original language of movies are extracted and TF-IDF vector matrix is created to get rid of weight down (lose the accuracy) the features of movies and based on that matrix find the similarity of the target movies whose have watched by target user by using similarity algorithms and suggest the top similarity movies to that user.

To find the similarity between the movies a similar matrix is created by using TF-IDF classifiers and cosine similarity metrics (based on TF-IDF vector).

### B. *Collaborative-Memory-based filter*

An item-based memory filter is also considered in this study. Movie rating data that was given by users is also used. To provide the recommendation to the target user, the user-movies matrix which has a similarity ratio of each movie for that target user based on the ratings of other users is used. Then, the top most similar movies are considered to find their top similar movies which are suggested to the target user.

To find the similarity between the movies, Pearson's correlation coefficient matrix is used. It is calculated as the covariance of the two variables divided by the product of the standard deviation of each data sample. The Truncated SVD classifier is used to create a user-item matrix for finding the similarity between movies.

### C. *Collaborative-Model-based filter*

Panda and surprise python libraries are used under the collaborative-model-based filter. The panda package is used to read the ratings which are in CSV format and convert them into a python data frame. The surprise dataset function is used to convert the rating data frame as a train set for the model-based predictions algorithms. Once the train set is created based on the users' ratings over the movie, the train set is used to predict the target user's ratings over a movie as a test set by using the surprise SVD functions. Surprise SVD classifier is used to create the model based on the train data of user rating, and predict the rating for the test data. The hyper-parameters of the SVD were tuned to minimize the error values.

### D. *Hybrid Approach*

By adding either a content-based filter with the collaborative-model-based filter or a collaborative-memory-based filter with a collaborative-model-based filter, the hybrid approach is implemented. If it is the content-model hybrid approach, the user is selected first and find the high rated movies of that user then find the most similar movies and for those movies then take those most similar movies to predict ratings on behalf of that user and out of those prediction ratings, then top rated movies are selected and suggest the top similar and high rated movies to users.

For the item-model hybrid approach, the user is selected first and find the top-rated movies and for those top rated movies find the similar movies then took those most similar movies and predict rating and suggest the top similar and high rated movies to users.

## V. EVALUATION

Cross-validation is performed to assess the effectiveness of the model. This is a statistical method used to evaluate the performance of our models. It is used to protect against the overfitting of the predictive model, especially when the amount of data is small. In cross-validation, the system generates a certain number of folds, in this study it is five, which carry out the analysis in each part and then calculate the average total error rate. Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) were selected to perform cross-validation on Collaborative- memory & model filters. With the use of metrics in Sci-kit learn, MAE and RMSE are calculated. And precision, recall, and F1 are used to evaluate content-based filters.

The precision is calculated considering predicted values and true values. Equation 1 shows the formula of precision. The recall means how many values are true in prediction out of real true values. Equation 2 shows the formula of recall. F1 score is the weighted average of Precision and Recall. Therefore, this score takes both false positives and false negatives into account. Intuitively it is not easy to understand as accurate, but F1 is usually more useful than accuracy. Equation 3 shows the formula for the F1 score.

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \quad (1)$$

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \quad (2)$$

$$F_1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (3)$$

MAE is the difference between the actual value (rating) and the predicted value. Equation 4 shows the formula of the MAE matrix. In 4, $\mathbf{y_i}$ denotes the observed value for the $i^{th}$ observation, $\mathbf{x_i}$ denotes the predicted value for the $i^{th}$ observation and **n** denotes the total number of observations. RMSE is similar to MAE but the only difference is that the absolute value of the residual is squared and the square root of the whole term is taken for comparison. Equation 5 shows the formula of the RMSE matrix. The advantage of using RMSE over MAE is that it penalizes the term more when the error is high.

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n} \quad (4)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (Predicted_i - Actual_i)^2}{N}} \quad (5)$$

### A. *Content-based filter*

Table 1 shows the testing results of content-based filters.

Table 1. Content-based filter results

| Filter System | **Precision** | **Recall** | **F1 Score** |
|---|---|---|---|
| Content-Based | 0.83 | 0.83 | 0.84 |

From Table 1, we get to know about the prediction accuracy of our content-based recommendation system based on the Precision, Recall, and F1 scores. Where it shows the Precision and Recall as 0.83 and the F1 score as 0.84. The F1 score provides a more Standard way of accuracy measurement by doing the weights of Precision and Recall.

### B. *Collaborative- Memory based filter*

The following Table 2 shows the testing results of a collaborative- memory-based filter where each fold contains a different combination of test and train data. Based on the train and test data the average RMSE and MAE of the

Collaborative model are identified as 0.8999, 0.6939 respectively.

Table 2. Testing Results of Collaborative Memory

| Metrics | RMSE | MAE |
|---|---|---|
| Fold 1 | 0.9045 | 0.7003 |
| Fold 2 | 0.8886 | 0.6864 |
| Fold 3 | 0.9070 | 0.6995 |
| Fold 4 | 0.8886 | 0.6780 |
| Fold 5 | 0.9108 | 0.7053 |
| Mean | 0.8999 | 0.6939 |

*a.*

### C. *Collaborative- Model-based filter*

Table 3 shows the testing results of the collaborative-model-based filter where each folds folding a different combination of test and train data. Based on the train and test data the average RMSE and MAE of the Collaborative model are identified as 0.8359, 0.6356 respectively.

Table 3. Testing Results of Collaborative Model-Based Filters

| Metrics | RMSE | MAE |
|---|---|---|
| Fold 1 | 0.8349 | 0.6348 |
| Fold 2 | 0.8370 | 0.6362 |
| Fold 3 | 0.8384 | 0.6379 |
| Fold 4 | 0.8358 | 0.6349 |
| Fold 5 | 0.8336 | 0.6343 |
| Mean | 0.8359 | 0.6356 |

### D. *Hybrid System*

In this study, each filter is improved, when these filters are combined as a hybrid system which can be either a content-based filter with a collaborative-model-based filter or a collaborative- memory-filter with a collaborative-model filter, the performance of the hybrid system will be increased..

## VI. CONCLUSION & FUTURE WORK

The recommendation system proposed in this paper is a Hybrid type of recommendation system where each of the filters has been analyzed with the possible classifiers and created based on the best and improved method according to the evaluations. The content-based filter has been made by adding more movie features. This will improve the accuracy of the filter to align and recommend a similar movie to users. The item-based collaborative- memory filter has been implemented and this will improve its accuracy and scalability compared with the user-based approach. The collaborative- model has been created with an improved model. It will lead reducing the difference between the actual and predicted movies based on user ratings. Further, for future work, the execution time of the system has to be reduced to provide immediate response with high accuracy.

## REFERENCES

[1] G. Rodríguez, “Introduction to Recommender Systems,” Tryolabs.com, 09-May-2018. [Online]. Available: https://tryolabs.com/blog/introduction-to-recommender-systems/. [Accessed: 13-Dec-2021].
[2] B. C. A. G. P. S. Vishwa Gosalia, "Movie Recommendation System," International Research Journal of Engineering and Technology (IRJET), vol. 05, no. 02, pp. 300-302, Feb-2018. [Accessed: 13-Dec-2021].
[3] Medium.com. [Online]. Available: https://medium.com/@madasamy/introduction-to-recommendation-systems-and-how-to-design-recommendation-system-that-resembling-the-9ac167e30e95. [Accessed: 13-Dec-2021].
[4] “Classifying different types of recommender systems,” Bluepiit.com, 14-Nov-2015.
[5] dataaspirant, Editorial Team, K. Taylor, and K. Wiggers, “An introduction to recommendation engines - dataconomy,” Dataconomy.com, 13-Mar-2015. [Online]. Available: https://dataconomy.com/2015/03/an-introduction-to-recommendation-engines/. [Accessed: 13-Dec-2021].
[6] S. M. F. C. S. D. Geetha G, "A Hybrid Approach using Collaborative filtering," National Conference on Mathematical Techniques and its Applications (NCMTA 18), 2018.
[7] Y. S. P. A. Prateek Sappadla, "Movie Recommender System," Search Engine Architecture, Spring 2017, NYU Courant, 2017.
[8] D. Y. A. S. V. K. G. Manoj Kumar, "A Movie Recommender System: MOVREC," International Journal of Computer Applications (0975 – 8887), vol. 124, no. 3, August 2015.
[9] S. K. A. A. N. P. Gharbi Alshammari, "Improved Movie Recommendations Based on a Hybrid," Vietnam Journal of Computer Science, vol. 6, no. 3, p. 363–376, 2019.
[10] Geeksforgeeks.org. [Online]. Available: https://www.geeksforgeeks.org/python-implementation-of-movie-recommender-system/. [Accessed: 13-Dec-2021].
[11] S. M. S. S. K. S. Roy, 2019 Global Conference for Advancement in Technology (GCAT), pp. 1-5, 2019.
[12] S. N. S. K. S. A. SRS Reddy, "Content-Based Movie Recommendation System Using Genre Correlation," 2018.
[13] S. M. T. T. T. N. G. S. Ramni Harbir Singh, "Movie Recommendation System using Cosine Similarity and KNN," International Journal of Engineering and Advanced Technology (IJEAT), vol. 9, no. 5, June 2020.
[14] A. M. S. R. P. M. Ramil G. Lumauag, "An Enhanced Memory-Based Collaborative Filtering Algorithm Based on User Similarity for Recommender Systems," International Journal of Recent Technology and Engineering (IJRTE), vol. 7, no. 6s, March 2019.
[15] D. S. C. Y. S. C. Kyung Soo Kim, "Boosting Memory-Based Collaborative Filtering Using Content-Metadata," April 2019.
[16] D. P. G. Rohan Mhetre, "Movie Recommendation Engine using Collaborative Filtering with Alternative Least Square and Singular Value Decomposition Algorithms," International

Journal of Advanced Research in Computer and Communication Engineering, vol. 8, no. 2, February 2019.

[17] S. Mirza Ilhami, "Film Recommendation Systems using Matrix," International Conference on Information Technology Systems and Innovation (ICITSI) 2014, pp. 1-6, November 2014.

[18] O. V. Rahul Katarya, "An effective collaborative movie recommender system with cuckoo search," Egyptian Informatics Journal, vol. 18, pp. 105-122, July 2017.

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index N0: CT-04-39*

# A Systematic Approach to Identify the Breast Cancer Grades in Histopathological Images Using Deep Neural Networks

S.H.S. Silva (1st Author)
*Department of Computer Science*
*University of Sri Jayewardenepura*
Gangodawila, 10250, Sri Lanka
hasinisandunika.usjp@gmail.com

T. M. K. K. Jinesena (2nd Author)
*Department of Computer Science*
*University of Sri Jayewardenepura*
Gangodawila, 10250, Sri Lanka
kasun@sjp.ac.lk

***Abstract* — Breast cancer can be recognized as one of the most well-known and life-threatening cancers impacting women and this has been identified as the second most common cancer across the world. According to registered data, there were over 2 million newly reported cases in 2020. The Deep Convolutional Neural Network has been identified as one of the most dominant and powerful deep learning approaches involved in the analysis of visual imagination. There are many shreds of evidence that indicate the appropriateness of this in medical imaging including breast cancer detection, classification, and segmentation with higher accuracy rates. The main intent of the research is to develop an automated application that can determine the Nottingham Histologic Score of a given input histopathological image obtained from breast cancer or healthy tissues with DenseNet based architecture. Healthy or benign tissues are categorized as zero and cancerous tissues are categorized based on the grade obtained as one, two, or three. In this study, we were able to obtain more than 94% accuracy rates for each trained model including 2-predict, 3-predict, and 4-predict networks. Further, a desktop-based inference tool that allows us to perform breast cancer grading was also developed as a result of this study.**

***Keywords* — *Breast Cancer Grading, Computer-Aided Diagnosis, DenseNet, Histopathological Images, Nottingham Histologic Score***

## I. INTRODUCTION

Breast cancer is identified as the second most common cancer and this holds the highest mortality rate due to cancers among women worldwide. As per the Union for International Cancer Control organization, 2.3 million new cases were reported in 2020 [34]. Once the breast cancer cells are formed, that will lead to the spread of malignant cells to other parts of the body by making it life-threatening. Breast cancers are often found earlier when they are small and before they spread. Due to technical advancement, early detection, and procedures, the survival rate of breast cancer patients has significantly improved across the world within the last few years. Breast cancer grade classification is a labor-intensive task and needs human experts and time to do the diagnoses. The grade of breast cancer tissue is calculated according to the way that cancer cells look under the microscope. For the laboratory breast cancer grade analysis purpose, it is best known that specimens of the tissues with the affected areas are used. Histopathological images are extracted as a result of this. It is known that histopathological images can be extracted with different High-Power Field (HPF) values. But images with 40X HPF values are well defined and illustrate the key features. The classified grade is used to figure out what treatments work best.

Densely Connected Convolutional Network (DenseNet) is one of the discoveries in deep neural networks for visual object recognition. Deep Convolutional Neural Networks (DCNN) with DenseNet architecture can be identified as one of the most dominant and powerful deep learning approaches involved in the analysis of visual imagination that provide excellent performance in medical imaging including breast cancer detection, classification, and segmentation with higher accuracy rates of more than 90% [32] [33]. This research aims to introduce a new computer-aided diagnosis approach that can continue the existing laboratory grading procedure with a higher accuracy rate.

## II. OBJECTIVES

Current study intends to develop an application that can determine the grade of a given histopathological image of breast cancer or healthy tissue as 0-benign, 1, 2, or 3 using a DenseNet based DCNN.

## III. METHODOLOGY

Below mentioned are the major tasks that had to be carried out during our research work.

- Data preprocessing

To complete the data preprocessing task, several mechanisms had to be applied and they are discussed under data preprocessing.

- Model building and evaluation

After completing the dataset preparation, built three DCNN models with DenseNet architecture as,

❖ 2-predict: classifies benign-0 and malignant-1

❖ 3-predict: classifies grades into grade 1, grade 2, and grade 3
❖ 4-predict: classifies as 0-benign, grade 1, grade 2, and grade 3.
● Implementation of the inference tool

To visualize the results obtained as well as to make the models available for use, an inference tool was built.

### *A. Data Preprocessing*

Initially, we received a graded (1, 2, and 3) dataset [11] with different HPF values (4X, 10X, 20X, 40X), resolutions, and image formats. To complete the data preprocessing task, we used a subset of the initial image dataset with 40X HPF value and increased that by applying several data augmentation techniques including image cropping, rotating, etc. Later we were able to create NumPy arrays for each grade with dimensions 700*128*128*3.

Again, we have used another dataset [14] to train the 2-predict model and after applying preprocessing techniques on the received benign 40X HPF dataset, we have created a NumPy array for that with dimensions 700*128*128*3. Fig. 1. illustrates an application of image augmentation by recreating images with cropping and rotating.

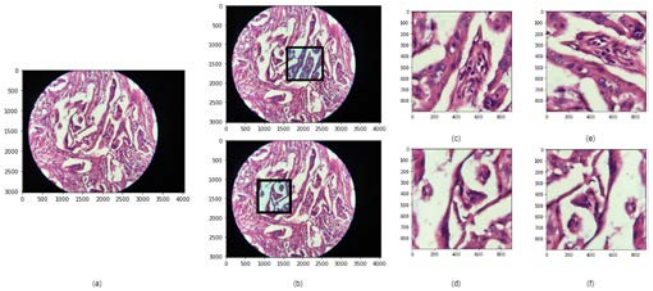


Fig. 1. Illustrates an application of Image Augmentation-(a) Original image, (b) Sub images to be extracted, (c,d) Extracted images, (e) Application of rotation up to 90 degrees for c, (f) Application of rotation up to 180 degrees for d.

For the 2-predict model implementation, the malignant dataset was created by applying the simple random sampling technique on the dataset of the 3-predict model. Dataset of the 4-predict model was created by combining both the benign image dataset of the 2-predict model and the dataset of the 3-predict model. Table 1. defines the dimensions of the training and testing datasets used in each model.

Table 1. Tabular Representation of the Datasets.

| Model | Training Dataset | Testing Dataset |
|---|---|---|
| 2-predict | 600*128*128*3 for 0<br>600*128*128*3 for 1 | 100*128*128*3 for 0<br>100*128*128*3 for 1 |
| 3-predict | 600*128*128*3 for 1<br>600*128*128*3 for 2<br>600*128*128*3 for 3 | 100*128*128*3 for 1<br>100*128*128*3 for 2<br>100*128*128*3 for 3 |
| 4-predict | 600*128*128*3 for 0<br>600*128*128*3 for 1<br>600*128*128*3 for 2<br>600*128*128*3 for 3 | 100*128*128*3 for 0<br>100*128*128*3 for 1<br>100*128*128*3 for 2<br>100*128*128*3 for 3 |

### *B. Model Building and Evaluation*

To build each model, the following key aspects were considered; Grade classification was carried out by considering two scenarios. In the first scenario, a single neural network model was implemented to classify the histopathological images as benign-0, grade 1, grade 2, and grade 3. In the second scenario, two deep neural network models were trained to achieve the same objective as in the first scenario whereas implementing the initial model to classify the histopathological image as benign or malignant and at the second step, to classify the grade based on the prediction obtained at the first step. As in Fig. 2., the initial model classifies the input histopathological image as benign-0 or malignant-1 and if it is malignant, obtains the prediction from the second model as 1, 2, or 3 otherwise the grade will be remaining as zero.

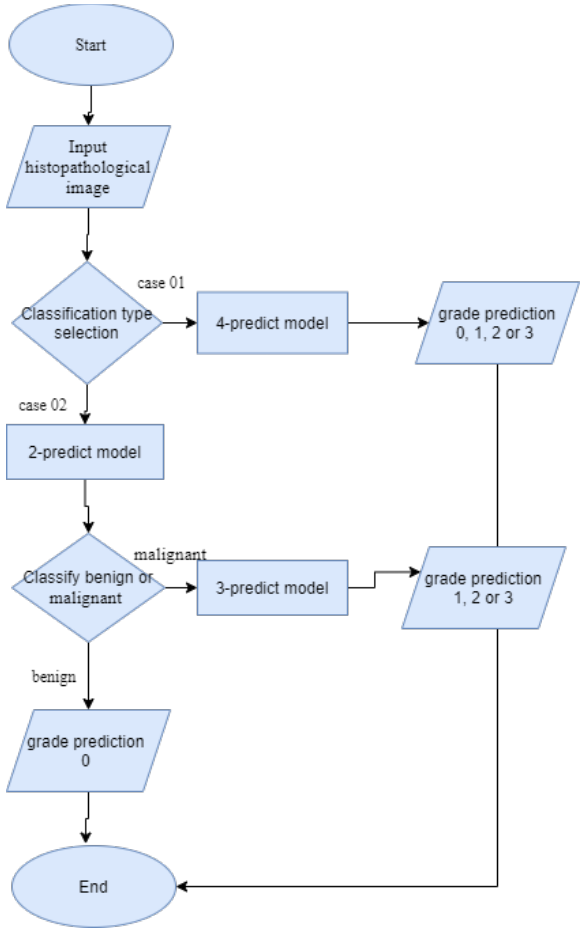



Fig. 2. Logical Overview of the Proposed Solution.

As per the research objectives that were mentioned in previous sections, this research is targeted at building a DCNN with transfer learning to predict the grades (0-benign, 1, 2, and 3) of histopathological images. Below define the architectures of transfer learning models, and the process of the evaluation carried out to achieve the above-defined goal.

*1. Model Building*

The transfer learning technique was applied to implement 2-predict, 3-predict, and 4-predict models for 128*128 RGB image datasets, and we were able to receive better test accuracy rates. This model is a composition of,

- Pre-trained DenseNet-201 weights
- 2D global average pooling layer
- Dropout layer with 0.5 probability
- Batch normalization layer
- Densely connected layer with softmax operation

The following Fig. 3. illustrates the architecture of the transfer learning model. This was implemented with Spyder and Anaconda framework and the libraries, dependencies used were TensorFlow, Keras, sklearn, and OpenCV.

```
Layer (type)                 Output Shape              Param #
=================================================================
densenet201 (Functional)     (None, 4, 4, 1920)        18321984
_________________________________________________________________
global_average_pooling2d (Gl (None, 1920)              0
_________________________________________________________________
dropout (Dropout)            (None, 1920)              0
_________________________________________________________________
batch_normalization (BatchNo (None, 1920)              7680
_________________________________________________________________
dense (Dense)                (None, 2)                 3842
=================================================================
Total params: 18,333,506
Trainable params: 18,100,610
Non-trainable params: 232,896
```

Fig. 3. The Architecture of the Transfer Learning Model.

Apart from that, a Flask Application Program Interface with Anaconda framework was developed to obtain the prediction by considering each case.

*2. Evaluation*

This was performed on the reserved test datasets of each model. For each model, there were around 14% of allocated histopathological images for testing purposes out of all images.

The evaluation was carried out by changing parameters including batch size, dropout probability, and the number of epochs. But, all three models provided significant responses only for the varied number of epochs rather than batch size and dropout probability. Mainly we have performed evaluation operations on test datasets including test accuracy rate calculation, confusion matrix illustration, and classification reports.

*C. Implementation of the Inference Tool*

This was carried out to visualize the results obtained and to make the trained models available for use. The inference tool was implemented as a desktop graphical user interface application and has added some basic graphical user interface functions such as uploading an image, viewing results, saving an image with the classification, searching a record by ID, deleting a record by ID, converting a record into PDF format with a suitable database connection.

The summarized architecture of the proposed solution is illustrated in Fig. 4.

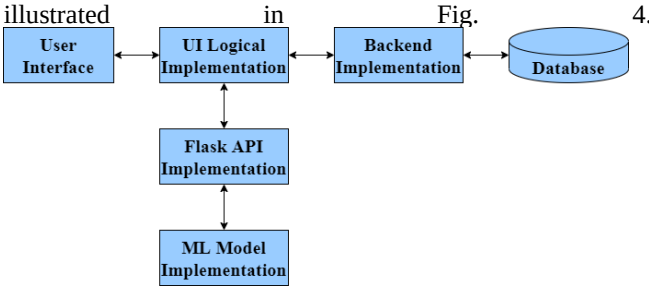



Fig.4. Summarized Architecture of the Proposed Solution.

This was implemented with Apache NetBeans with Maven dependencies and MongoDB was used as the database.

Apart from that, an Application Program Interface integration was carried out to obtain the predictions by considering each case as 1and 2.

## IV. RESULTS AND DISCUSSION

To analyze the results generated by each model, the main method that followed was the test data-based results generation. The following includes the results obtained and the comparison between the obtained results.

*A. Two-Predict Model*

Fig. 5. shows the loss vs epochs received by applying transfer learning during the process of training for the two-predict model with 20 epochs. As per this, curves of both validation and training are getting converged with the number of epochs. We could see that the model attains to the minimum loss rate at around the 4 th epoch and continues the same until the 20th epoch. This indicates that the model shows a good fit with the training dataset.

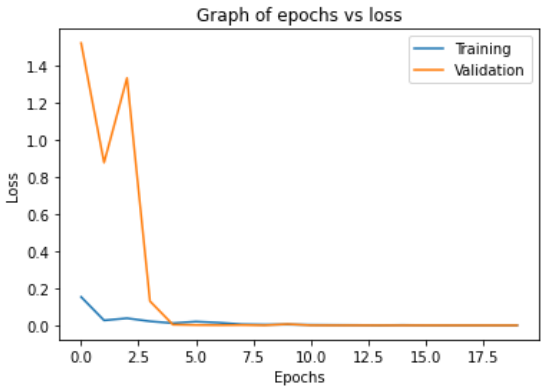



Fig. 5. Loss vs Epochs for Two-Predict Model.

Fig. 6. shows the normalized confusion matrix received by applying transfer learning from the testing dataset for the two-predict model with 20 epochs.

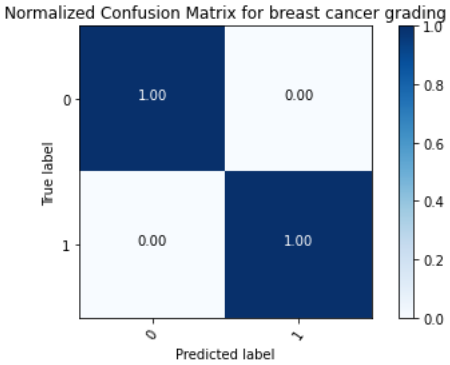



Fig. 6. Normalized Confusion Matrix for Two-Predict Model.

Fig. 7. shows the classification report received by applying transfer learning from the testing dataset for the two-predict model with 20 epochs. According to this result, precision, recall, and F1-score have obtained the highest values as can be obtained. Also, the accuracy received for the test dataset is 100%.

```
In [13]: print(classification_report( true_y.argmax(axis=-1), y_predict))
              precision    recall  f1-score   support

           0       1.00      1.00      1.00       100
           1       1.00      1.00      1.00       100

    accuracy                           1.00       200
   macro avg       1.00      1.00      1.00       200
weighted avg       1.00      1.00      1.00       200
```

Fig. 7. Classification Report for Two-Predict Model.

### *B. Three-Predict Model*

Fig. 8. shows the loss vs epochs received by applying transfer learning during the process of training for the three-predict model with 30 epochs. This model was previously trained for 20 epochs. As per that, the trained model with 30 epochs tends to converge with the number of epochs rather than the model with 20 epochs. This indicates that the model with 30 epochs shows a better fit with the training dataset than the model with 20 epochs.

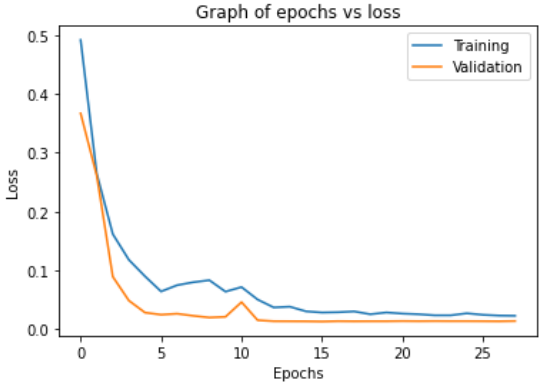



Fig. 8. Loss vs Epochs for Three-Predict Model.

Fig. 9. shows the normalized confusion matrix received by applying transfer learning from the testing dataset for the three-predict model with 30 epochs.

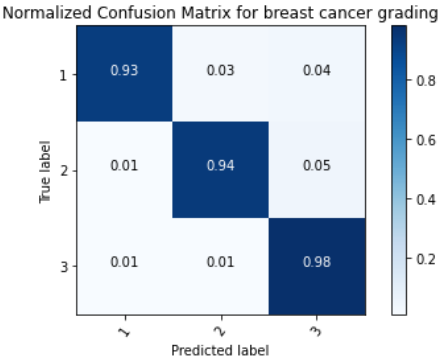



Fig. 9. Normalized Confusion Matrix for Three-Predict Model.

Fig. 10. shows the classification report received by applying transfer learning from the testing dataset for the three-predict model with 30 epochs. As this was trained for 20 epochs previously, the trained model with 30 epochs holds maximum values for precision, recall, and F1-score. The test accuracy rates of trained models with 20 and 30 epochs are 94.33% and 94.99% respectively.

```
In [47]: print(classification_report( true_y.argmax(axis=-1), y_predict))
              precision    recall  f1-score   support

           0       0.98      0.93      0.95       100
           1       0.96      0.94      0.95       100
           2       0.92      0.98      0.95       100

    accuracy                           0.95       300
   macro avg       0.95      0.95      0.95       300
weighted avg       0.95      0.95      0.95       300
```

Fig. 10. Classification Report for Three-Predict Model.

### *C. Four-Predict Model*

Fig. 11. shows the loss vs epochs received by applying transfer learning during the process of training for the four-predict model with 25 epochs. As this was previously trained for 20 and 30 epochs, the trained model with 25 epochs shows much convergence between validation and training curves. This indicates that the model with 25 epochs shows a good fit with the training dataset.

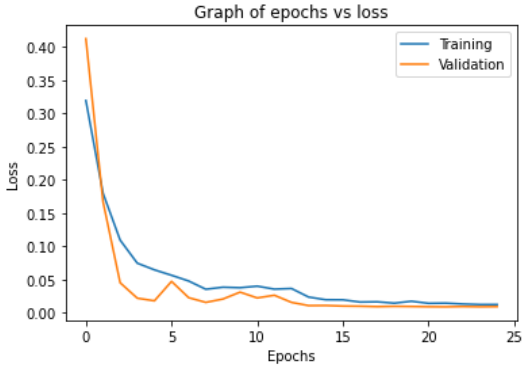



Fig. 11. Loss vs Epochs for Four-Predict Model.

Fig. 12. shows the normalized confusion matrix received by applying transfer learning from the testing dataset for the four-predict model with 25 epochs.

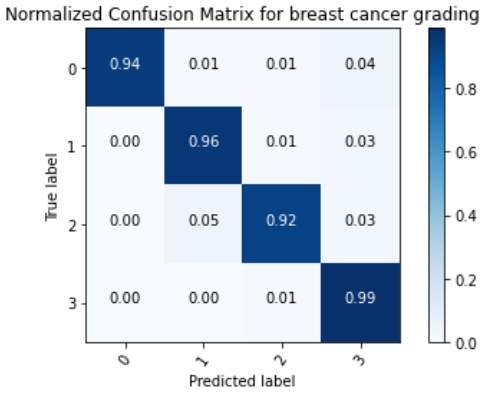



Fig. 12. Normalized Confusion Matrix for Four-Predict Model.

Fig. 13. show the classification report received by applying transfer learning from the testing dataset for the four-predict model with 25 epochs. As per previous results for 20 and 30 epochs, the trained model with 25 epochs holds maximum values for precision, recall, and F1-score. The test accuracy rates of trained models with 20, 25, and 30 epochs are 92.75%, 95.2499%, and 94.7499% respectively.

```
In [61]: print(classification_report( true_y.argmax(axis=-1), y_predict))
              precision    recall  f1-score   support

           0       1.00      0.94      0.97       100
           1       0.94      0.96      0.95       100
           2       0.97      0.92      0.94       100
           3       0.91      0.99      0.95       100

    accuracy                           0.95       400
   macro avg       0.95      0.95      0.95       400
weighted avg       0.95      0.95      0.95       400
```

Fig. 13. Classification Report for Four-Predict Model.

*D. Model Type (Case 1 and 2)*

Fig. 14. shows the test accuracy rates received by performing case 1 and 2 operations for the test dataset of the 4-predict model. This was implemented with the 2-predict model with 20 epochs, the 3-predict model with 30 epochs, and the 4-predict model with 25 epochs.

```
Size of the output testng data set:  (400, 4)
Size of the input testing data set:  (400, 128, 128, 3)
13/13 - 63s - loss: 0.0602 - accuracy: 0.9525
True count:  372
False count:  28


Testing accuracy score for case 1:  95.24999856948853 %
Testing accuracy score for case 2:  93.0 %
```

Fig. 14. Test Accuracy Rates of Case 1 and 2.

According to the results obtained,

- Two-predict model: When Fig. 5 is considered, we could see that the graph is getting converged whereas increasing the number of epochs. Fig. 6 and 7 also evidently express the appropriateness of the model by providing a 100% accuracy rate for the test dataset.
- Three-predict model: When Fig. 8 is considered, we could see that we can obtain a good fit by increasing the number of epochs up to 30. Fig. 9 and 10 also provide some evidence, providing a 94.99% accuracy rate for the test dataset for 30 epochs.
- Four-predict model: By considering Fig. 11, we can conclude that we can increase the fitness of the model by increasing the number of epochs up to 25. Fig. 12 and 13 also show a higher test data accuracy rate around 95.249% for 25 epochs.
- Model type (case 1 and 2) comparison: According to Fig. 14, can conclude that both approaches are good enough, and also both are applicable for the desktop graphical user interface prediction purpose.

By referring to the above results, we can conclude that it is appropriate to train the two-predict model up to 20 epochs, the three-predict model up to 30 epochs, and the four-predict model up to 25 epochs. Furthermore, both model types (case 1 and 2) are applicable for the desktop graphical user interface prediction purpose.

## V. CONCLUSION

In this research, three DCNN based models with transfer learning have been implemented to classify the grades of breast cancer. These DCNN models have been trained and tested with histopathological image datasets with the 40X HPF value. 2-predict, 3-predict, and 4-predicts models were able to obtain test accuracy rates of 100%, 94.999%, and 95.2499% respectively. Further, this research confirms the successfulness of using transfer learning with DenseNet architecture whereas providing test accuracy rates of more than 94% for all three trained models. Moreover, a desktop application was developed to infer the solution. Ultimately, accuracy improvements, implementation of the developed system as a web-based application could be identified as work that could be carried out as further improvements of this research.

## REFERENCES

[1] T. Mahmood, M. Arsalan, M. Owais, M. B. Lee, and K. R. Park, “Artificial intelligence-based mitosis detection in breast cancer histopathology images using Faster R-CNN and deep CNNs,” J. Clin. Med., vol. 9, no. 3, p. 749, 2020.

[2] D. C. Cireşan, A. Giusti, L. M. Gambardella, and J. Schmidhuber, “Mitosis detection in breast cancer histology images with deep neural networks,” Med. Image Comput. Comput. Assist. Interv., vol. 16, no. Pt 2, pp. 411–418, 2013.

[3] T. Mahmood, S. Ziauddin, A. R. Shahid, and A. Safi, “Mitosis detection in breast cancer histopathology images using statistical, color and shape-based features,” J. Med. Imaging Health Inform., vol. 8, no. 5, pp. 932–938, 2018.

[4] M. A. Aswathy and M. Jagannath, “Detection of breast cancer on digital histopathology images: Present status and future possibilities,” Inform. Med. Unlocked, vol. 8, pp. 74–79, 2017.

[5] A. Tashk, M. S. Helfroush, H. Danyali, and M. Akbarzadeh-jahromi, “Automatic detection of breast cancer mitotic cells based on the combination of textural, statistical and innovative mathematical features,” Appl. Math. Model., vol. 39, no. 20, pp. 6165–6182, 2015.

[6] H. Wang et al., “Mitosis detection in breast cancer pathology images by combining handcrafted and convolutional neural network features,” J. Med. Imaging (Bellingham), vol. 1, no. 3, p. 034003, 2014.

[7] T. X. Jian, M. Nazahah, M. M. Yusoff, and A. R. K. Shakir, “Mitotic cell detection in breast histopathology image: A review,” J. Phys. Conf. Ser., vol. 1372, p. 012026, 2019.

[8] S. Rao, “Mitos-RCNN: Mitotic figure detection in breast cancer histopathology images using region-based convolutional neural networks,” 2018.

[9] T. K. ten Kate, J. A. Beliën, A. W. Smeulders, and J. P. Baak, "Method for counting mitoses by image processing in Feulgen stained breast cancer sections," Cytometry, vol. 14, no. 3, pp. 241–250, 1993.
[10] M. Veta et al., "Assessment of algorithms for mitosis detection in breast cancer histopathology images," Med. Image Anal., vol. 20, no. 1, pp. 237–248, 2015.
[11] H. Bolhasani, E. Amjadi, M. Tabatabaeian, and S. J. Jassbi, "A histopathological image dataset for grading breast invasive ductal carcinomas," Inform. Med. Unlocked, vol. 19, no. 100341, p. 100341, 2020.
[12] "Staging & Grade," Jhu.edu. [Online]. Available: https://pathology.jhu.edu/breast/staging-grade/. [Accessed: 12-Jun-2021].
[13] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "Breast cancer histopathological image classification using Convolutional Neural Networks," in 2016 International Joint Conference on Neural Networks (IJCNN), 2016.
[14] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "A Dataset for Breast Cancer Histopathological Image Classification," IEEE Trans. Biomed. Eng., vol. 63, no. 7, pp. 1455–1462, 2016.
[15] A. Fernando, U. Jayarajah, S. Prabashani, E. A. Fernando, and S. A. Seneviratne, "Incidence trends and patterns of breast cancer in Sri Lanka: an analysis of the national cancer database," BMC Cancer, vol. 18, no. 1, 2018.
[16] Deep Learning by Ian Goodfellow, Yoshua Bengio, and Aaron Courville published by MIT Press, 2016.
[17] Stanford University's Course — CS231n: Convolutional Neural Network for Visual Recognition by Prof. Fei-Fei Li, Justin Johnson, Serena Yeung.
[18] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," Proc. IEEE Inst. Electr. Electron. Eng., vol. 86, no. 11, pp. 2278– 2324, 1998.
[19] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Commun. ACM, vol. 60, no. 6, pp. 84–90, 2017.
[20] H. Peiris, L. Mudduwa, N. Thalagala, and K. Jayatialake, "The value of Nottingham grade in breast cancer re-visited in the Sri Lankan setting," Malays. J. Pathol., vol. 39, no. 2, pp. 141–148, 2017. 9
[21] C. W. Elston and I. O. Ellis, "Pathological prognostic factors in breast cancer. I The value of histological grade in breast cancer: experience from a large study with long-term follow-up," Histopathology, vol. 19, no. 5, pp. 403–410, 1991.
[22] A. Paul and D. P. Mukherjee, "Mitosis detection for invasive breast cancer grading in histopathological images," IEEE Trans. Image Process., vol. 24, no. 11, pp. 4041–4054, 2015.
[23] M. Talo, "Automated classification of histopathology images using transfer learning," Artif. Intell. Med., vol. 101, no. 101743, p. 101743, 2019.
[24] X. Wu, T. Yang, and Z. Xia, "Gait recognition based on DenseNet transfer learning," Ijset.net. [Online]. Available: https://www.ijset.net/journal/2458.pdf. [Accessed: 12-Jun-2021].
[25] Q. Cai, X. Liu, and Z. Guo, "Identifying architectural distortion in mammogram images via a SE-DenseNet model and twice transfer learning," in 2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), 2018.
[26] S. Minaee, R. Kafieh, M. Sonka, S. Yazdani, and G. Jamalipour Soufi, "Deep-COVID: Predicting COVID-19 from chest X-ray images using deep transfer learning," Med. Image Anal., vol. 65, no. 101794, p. 101794, 2020.
[27] X. Xu, J. Lin, Y. Tao, and X. Wang, "An improved DenseNet method based on transfer learning for fundus medical images," in 2018 7th International Conference on Digital Home (ICDH), 2018.
[28] Y. Celik, M. Talo, O. Yildirim, M. Karabatak, and U. R. Acharya, "Automated invasive ductal carcinoma detection based using deep transfer learning with whole-slide images," Pattern Recognit. Lett., vol. 133, pp. 232–239, 2020.
[29] F. Imrie, A. R. Bradley, M. van der Schaar, and C. M. Deane, "Protein family-specific models using deep neural networks and transfer learning improve virtual screening and highlight the need for more data," J. Chem. Inf. Model., vol. 58, no. 11, pp. 2319–2330, 2018.
[30] S. Pal and M. L. Gupta, "Correlation between cytological and histological grading of breast cancer and its role in prognosis," J. Cytol., vol. 33, no. 4, pp. 182–186, 2016.
[31] S.-H. Wang and Y.-D. Zhang, "DenseNet-201-based deep neural network with composite learning factor and precomputation for multiple sclerosis classification," ACM trans. multimed. comput. commun. appl., vol. 16, no. 2s, pp. 1–19, 2020.
[32] Y. Yari, T. V. Nguyen, and H. T. Nguyen, "Deep learning applied for histological diagnosis of breast cancer," IEEE Access, vol. 8, pp. 162432–162448, 2020.
[33] A. Rakhlin, A. Shvets, V. Iglovikov, and A. A. Kalinin, "Deep convolutional neural networks for breast cancer histology image analysis," in Lecture Notes in Computer Science, Cham: Springer International Publishing, 2018, pp. 737–744.
[34] "Breast cancer," Uicc.org. [Online]. Available: https://www.uicc.org/what-we-do/thematic-areas-work/breast-cancer. [Accessed: 12-Jun-2021

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index N0: CT-05-40*

# Aspect Based Multi-Class Sentiment Dataset for Bilingual eWOM of Commercial Food Products

Binguni Edirisuriya (1st Author)
*Department of Industrial Management*
*University of Kelaniya*
Dalugama, Sri Lanka.
edirisur_im16018@stu.kln.ac.lk

Dinesh Asanka (2nd Author)
*Department of Industrial Management*
*University of Kelaniya*
Dalugama, Sri Lanka.
dasanka@kln.ac.lk

Janaka Wijayanayake (3rd Author)
*Department of Industrial Management*
*University of Kelaniya*
Dalugama, Sri Lanka.
janaka@kln.ac.lk

***Abstract* — Aspect-Based Sentiment Analysis for product review opinion analysis is commonly utilized by the commercial food products manufacturing businesses to drive decisions regarding products. However, the general consumers are not facilitated with decision-making ready end-user applications which generates insights to arrive at the purchase decision at the time of purchase due to the unavailability of products' attribute-wise analysis-ready data. Although Electronic Word of Mouth (eWOM) platforms are comprised of opinions with a diversity of languages and expression formats, themselves do not generate any value to make comparable decision making. Hence, there is an existing gap of impactful information retrieval by the consumer to aid the purchase. Therefore, creating a publicly available analysis-ready dataset for the commercial food product domain contributes significantly to the Sri Lankan consumers and Government organizations. Through our research work, a manually annotated bilingual eWOM opinion text dataset for selected commercial food products categories has been delivered in which the opinions expressed in the Sinhala language have been translated into English language and each opinion has been manually rated into five levels by two domain experts. Two product attributes, "Price of the product", "Safeness of product" have been considered as aspects to conduct the Aspect-Based Sentiment Analysis. This study describes the sub-tasks performed to conduct the Aspect-Based Sentiment Analysis on the dataset along with the basic statistical evaluation of the dataset. We have presented results on the performance of the dataset by utilizing an existing Long Short-Term Memory Model.**

***Keywords* — *aspect-based sentiment analysis, commercial food domain***

## I. INTRODUCTION

Commercial food products play a vital role in fulfilling the consumption needs of individuals to manage food requirements with their busy lifestyles. In Sri Lanka, the commercial food market has been occupied by many local and multinational companies who are manufacturing the same varieties of food products with varied product quality attributes and products' prices. The production quality, packaging, and ingredients of the product reveal the "consumption-safeness of the product" which the consumer must focus on when making the purchase decision, as unsafe food product consumption may lead to health threats including Non-Communicable diseases. Consumer surveys reflect typical Sri Lankan consumer faces many challenges when making a purchase decision of commercial food products due to the lack of perceived knowledge on food hygiene and nutritional information analysis, being caught under influential, deceive marketing campaigns, promotions executed by business [1], inability to conduct a quick critical review on economical decision making when other alternative products are remained unobserved, unrevealing accurate information by the responsible government authorities and unavailability of a suitable unbiased supportive information resource which can be incorporated in analytical end-user applications which can aid the purchase decision at the time of purchase.

Nowadays, eWOM has become a prominent information resource that utilizes for business analytics purposes in many domains. Since eWOM comprises of diverse knowledge base, ideologies, expressions on experience with a high volume of user-base, with more than 6.40 million users in Sri Lanka, it can be recognized as a rich information resource.

### *A. Knowledge level on Commercial Food Products*

The stakeholders associated with the commercial food product domain are the general consumers, commercial food manufacturers, Government Policy authorities like Consumer Affairs Authority, and Sri Lankan Medical Authorities. Consumer survey results and results of the interviews conducted in Sri Lanka in 2020, to assess the knowledge of food stakeholders [2],[3] depicted the fact, the lack of literacy exists in food stakeholders to recognize, distinguish, and select the most suitable food products to consume.

### *B. Text Mining in eWOM on Purchase Decision*

The definition for Electronic Word of Mouth (eWOM) can be identified as "consumer-generated, consumption related communication that employs digital tools and is directed primarily to other consumers", which are in different

forms as reviews, opinions, feedback where they become valuable, persuasive mode of information since most of the reviews are experienced-based, open-connected nature, active participation and affects the purchase decision making process of the consumer [4].

Text mining has been utilized for decision-making by the food products-related businesses, food science, and nutrition domain. Different data gathering mechanisms, data analysis methods have been discussed in the review paper analysis [5],[6] The evaluation of the review analysis was able to reveal that the key data source which can be used to identify consumers' opinions, food issues and the food consumption behavior is "social media".

### C. *Aspect-Based Sentiment Analysis*

Aspect-Based Sentiment Analysis related research domain comprises many dedicated research studies targeting different domains. Publicly accessible improved datasets have been incorporated to evaluate the process and sub-tasks of Aspect-Based Sentiment Analysis. Using the publicly available restaurant review dataset, [7] study delivered annotated dataset considering six aspect categories and provide overall sentiment polarity and not focusing on aspect-wise sentiment polarity consideration. SemEval 2014 [8], [9] introduced laptops, restaurants datasets annotated with four fields as aspect term, aspect term polarity, aspect category, and aspect category polarity to be consumed in Aspect-Based Sentiment analysis. Driving a new approach SemEval 2016 [10] presented Multilingual datasets from 7 domains and 8 languages in which all were involved in Aspect-Based Sentiment Analysis tasks. The domains included in the datasets were hotels, laptops, mobile phones.

The research study conducted by [11] has focused on aspect-based sentiment analysis in the legal domain by fetching data from an available public dataset. They have taken 2000 legal statements and involved domain experts to annotate data manually by aspect identification through the legal statement and assign the sentiment level for each party by deciding the level of the beneficial party. They have focused on sub-tasks of sentiment analysis when annotating the dataset.

The research study conducted by the scholars, [12] has focused on analyzing the reviews given by university students regarding the learning experience at the University. The study has been driven considering the 22 different aspect levels of the different parties associated with the teaching and learning process. The reviews have been classified based on a weighted score for aspects to deliver unbiased results corresponding to each aspect.

This study would drive research attention on how the Aspect-Based Sentiment Analysis process can be incorporated to construct an analysis-ready user opinion dataset for the commercial food products domain. Parallel to that, methods for eWOM data extraction, data analysis, aspect-wise weight values assignment would be considered referring to the literature.

## II. OBJECTIVES

This section discusses the main aim and objectives of constructing the dataset and its significance. As discussed in Section I, it is obvious that there is an existing gap for the retrieval of "decision-making" ready information by the Sri Lankan consumer, to make a beneficial commercial food product purchase for the wellbeing of the purchaser. To achieve that latency consumers should be supplied with an accurate public information retrieval source. Our main aim is to construct a publicly available dataset that can be utilized for opinion analysis purposes related to the food products domain. To achieve the main aim, several sub-objectives are executed to methodically gather dynamic, unstructured commercial food product related posts' data in Facebook eWOM platform, to perform sub-tasks of Aspect-Based Sentiment Analysis for aspect category identification and categorization of posts and aspect term identification with appropriate multi-class sentiment polarity assignment for each eWOM comment in the dataset, which would be utilized to conduct experiments.

Hence, this research would be significant, since it benefits the Sri Lankan community, with an analysis-ready data resource that can be incorporated as the analytical resource to develop analytical end-user applications to aid the purchase decision-making process. Thus, the researchers, business analytical officers, economists, health officials, and other government authorities such as Consumer Affairs Authority, Sri Lanka Standard Institute would gain the opportunity to conduct social-opinion-oriented surveys.

## III. METHODOLOGY

### A. *Data Source Selection*

This section describes the criterion considered for selecting the data sources for data collection. As stated in *StatCounter Global Stats*, 57.28% of social media users in Sri Lanka have been actively sharing opinions on the Facebook platform between September 2020 to September 2021. Two public Facebook discussion groups with members count over 100,000 and which focus on consumer experience on purchasing commercial food products were selected as the data collection sources.

### B. *Product Aspects Selection*

This section describes how the products' aspects were elected to be considered in the research study. The *Consumer awareness Survey on Sustainable Consumption – 2018: Sri Lanka* has presented a consumer survey conducted targeting purchase decision making of food products where 423 individuals from five main densely populated districts in Sri Lanka were involved. The results of the survey revealed that the "price of the food product" was the main determinant for purchase decision with a weighted mean average of 3.78 while the "consumption safe quality of the product" was 3.56. In our study, the above values were taken into consideration as the aspect weight values to be assigned into opinions.

Table 1. Product Aspects of Commercial Food Products

| Product Aspect | Opinion Consideration by eWOM users |
|---|---|
| Safety of Food Product | Ingredient analysis, Packaging, Nutrition levels, Harmful effects, Experience on consumption, Health issues |
| Price of Food Product | Price among alternatives, Economical benefits, Inflation, Beneficial Purchase Decision |

### C. *Facebook Post data Crawling*

This section discusses the process of collecting data from Facebook posts with the mechanism used to crawl data. Observing the eWOM discussions in the public Facebook discussion groups selected food product categories were "Diary Products" and "Beverages" concerning timely discussions that occurred during the covid19 pandemic period. To conduct the Facebook data crawling, a python script was developed using Python 3 programming language and *Facebook scraper* library available in python language. About 1574 comments were collected along with the post details.

**Algorithm 1: Facebook post details extraction**

```
Input : page_id
Input : no_of_posts_crawl
Function scrape(page_id, no_of_posts_crawl):
    page = get_page (page_id, cookies = "cookie_file_name.txt")
    all_posts = []
    posts = get_posts (page_id, no_of_posts_to_scrape)
    i = 0
    Loop For post in posts:
        all_posts [i] ['post'] . add(post.id AND post.content)
        all_posts [i] ['comments'] . add(post.comments)
        Do increase i by 1
    End Loop
End Function
```

Fig. 1. Algorithm: Extraction of Facebook Post Details & Comments

### D. *Dataset Construction*

This section describes how the dataset construction was done with the extracted Facebook data. The post details were retrieved in nested JavaScript Object Notation (JSON) format and must be transformed into a data frame with post ID, post text, product category (Diary product or Beverage), product ID where the product's brand name was masked to preserve the unbiasedness of the comments' annotation process by the annotators, post aspect (price or safety of the product) and the comment.

### E. *Comments Preprocessing Steps*

This section discusses the preprocessing steps conducted for bilingual comments pre-processing. Facebook eWOM comments are ambiguous with different linguistics, different variety of opinions expression modes with media types (images, videos, pictorial emojis), hashtags (#), punctuations, URL s, mentioning of an entity with @ sign and pictorial emojis. Therefore, preprocessing steps were conducted with Pandas python library to eliminate the above figures from comments.

.

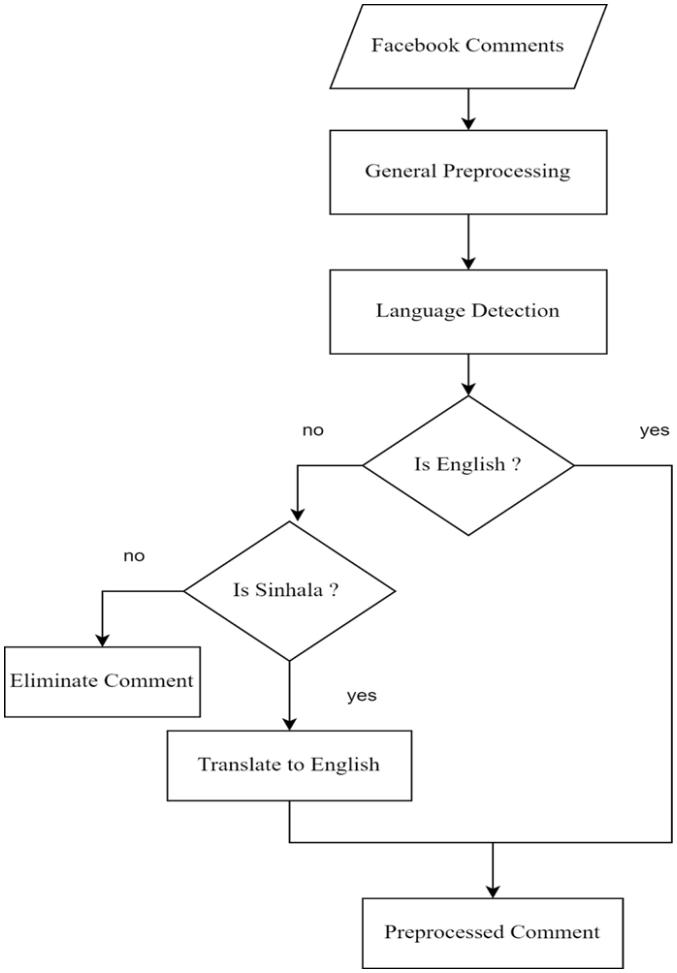



Fig. 2. Flowchart : Language Detection & Translation

The results elaborated that the comments which are expressed in "Singlish" format could not be recognized as one categorization. Currently, there's no publicly available tool to translate the "Singlish" comments to the English language. Therefore, it reveals that the semantic orientation of Singlish opinion expression could not be captured considering the whole sentence level. Hence, for this level of the study comments which have been expressed in neither English nor Sinhala were eliminated from the consideration. The comments which were detected as "Sinhala" language were translated into the English language, using Google Translator API since sentiment analysis studies have incorporated it for language translation purposes [15].

### G. *Manual Data Annotation*

This section discusses the process driven by the two annotators to manually annotate the Facebook eWOM

comment data. Reviewing the past literature, we were able to find that there is no publicly available tool to conduct bilingual eWOM text data annotation. As annotators, one graduate and one undergraduate from the University of Kelaniya were involved in the data annotation process. The annotator's task was to figure out the aspect categories expressed in each comment and assign a polarity rating for each aspect. To address the ambiguities in language and opinion expression, the need for a wider range of rating levels was obvious [16], [12]. Therefore, the annotators conducted the rating assignment adhering to the rating levels mentioned in Table 1.

Table 2. Rating levels of Annotation

| Rating Level | Sentiment level assigned for rating |
|---|---|
| 1 | Very Negative |
| 2 | Negative |
| 3 | Neutral |
| 4 | Positive |
| 5 | Very Positive |

Example 1

- *Comment 1.1 : D001milk powder contains toxic matter so , useless to buy at a high price.*

Table 3. Rating assignment for the comment by Annotator

| Aspect: Sentence | Opinion Word | Rank Assignment | | | | |
|---|---|---|---|---|---|---|
| | | 5 | 4 | 3 | 2 | 1 |
| Price | Toxic | | | | | ✓ |
| Safety | High Price | | | | ✓ | |

The two human annotators assigned ratings for the Facebook eWOM comments separately, by identifying the aspect category expressed through the comment and assigning it with a rating value [11], [12], [17]. A single comment is assigned with two sentiment rating levels by a single annotator, since one comment may consist of opinions expressed concerning both aspects "product safety" and "product price". The comments which have been expressed an opinion only on a single aspect were assigned the rating level 3 (*Neutral*) for the aspect which is not included in the comment. The above example demonstrates how the comment annotation process was conducted by one single annotator. Consider the comment stated in Example 1, the annotator reads the comment and identifies the aspect categories expressed by the comment, next concerning each aspect annotator identifies the sentiment-bearing words and assigns two ratings per each comment under the two product aspects. Above Table 2 shows how the annotator assigns ratings for the above comment in Example 1.

After the two annotators have assigned ratings, the agreement level between the two annotators was assessed with "Inter- Annotator agreement level" to verify the interpretation assigned by the annotators. For that, "Cohen's Kappa" was calculated [18], which is a quantitative measure to measure the chance of how often that raters are agreeing on the same thing/opinion.

$$K = Pa - Pc / 1-Pc \quad (1)$$

The Cohen's kappa value was calculated separately for the two aspects "product price" and "product safety" and further the results are presented in Section IV.

### *H. Aspect-wise Class label Assignment for comments*

This section discusses how the class labels were assigned for each comment concerning two product aspects considered in the study. Considering the results of Cohen's Kappa scores presented in Section IV, the final aspect rating for a particular comment was obtained by calculating the average rating score assigned by the two annotators since the Cohen's Kappa score revealed that the agreement level between two annotators is relatively same.

In the next step, the average rating score obtained for a certain comment was weighted with the weighted mean value considered for each aspect. The average rating score received for the "price aspect" was weighted with 3.78, while the average rating score received for the "safety aspect" was weighted with 3.56. Further, the weighted rating scores were subjected to Min-Max Scaling to derive a score between 0 and 1, to generalize the process of splitting the multiple-sentiment classes. Table 4 shows the counts of the comments received for each scaled value under the two aspects.

Table 4. Statistics: Comments for respective scaled values

| Scaled Value | Count: Safety | Count: Price |
|---|---|---|
| 0.0 | 4 | 4 |
| 0.1 | 0 | 0 |
| 0.2 | 61 | 122 |
| 0.3 | 1 | 9 |
| 0.4 | 92 | 139 |
| 0.5 | 2 | 4 |
| 0.6 | 478 | 670 |
| 0.7 | 11 | 3 |
| 0.8 | 255 | 120 |
| 0.9 | 46 | 12 |
| 1.0 | 249 | 116 |

The next step was to define the multiple sentiment classes considering the scaled value distribution. The scaled value range was sliced into five ranges with similar sizes and defined class labels corresponding to each range considering the lower score values to high score values as shown in Table 5. The weighted rating scores for each aspect of each comment were classified into sentiment classes using the "Binning Technique" in Pandas library.

Table 5. Statistics: Comments for respective scaled values

| Range | Sentiment level of rating |
|---|---|
| 0.0 <= score < 0.2 | Very Negative |
| 0.2 <= score < 0.4 | Negative |
| 0.4 <= score < 0.6 | Neutral |
| 0.6 <= score < 0.8 | Positive |
| 0.8 <= score < 1.0 | Very Positive |

Section IV presents the aspect-wise sentiment classification of the comments considered in the dataset.

## IV. RESULTS AND DISCUSSION

This section has presented the statistics of the dataset as a baseline evaluation, aspect-wise consideration of the statistics with some experiments driven to evaluate the performance of the manually annotated dataset.
Aspect-Based Sentiment Analysis is associated with several sub-tasks and in this study, we consider several sub-tasks as Aspect category identification, Aspect term identification, Aspect category polarity of the sentence. The dataset comprised of eWOM text with translated opinions that were expressed in the Sinhala language before. Hence, it adds complexity to the sub-tasks due to the ambiguity of eWOM text.

### *A. Aspect Category Identification*

Initially, the extracted posts were categorized into the aspects "safety" and "price" considering the content of the post. One single comment may contain opinions for both aspects and some may not belong to any of the considered aspects. Annotators identify the aspects manually and decide on assigning a rate for the comment.

### *B. Aspect Term Identification*

For the given eWOM comment, there can be several terms, a combination of terms that imply the polarity of the comment concerning the considered aspects. Dataset presents the terms which were concerned in the polarity assignment for the comment.

### *C. Aspect Polarity of the Sentence*

Each comment in the dataset comprises of two rating levels between 1-5 scale, for each aspect "safety" and "price" corresponding to the polarity level of the comment. The two rating levels assigned have been utilized to conduct the experiments in the dataset.

### *D. Dataset Statistics*

This section presents preliminary results of statistical analysis done on the collected data. A total of 1574 comments were collected initially and only 1197 comments were concerned for the data annotation process, since 377 comments which were expressed in "Singlish" format and other media formats, were eliminated from the consideration. Table 6 represents the post-aspect-wise statistics of the dataset, in which the respective word counts have been considered after eliminating "Stop-Words" using the NLTK library.

Table 6. Categorization of items as an example

| Post Aspect | Comments Count | Word Count |
|---|---|---|
| Safety of Food Product | 651 | 3832 |
| Price of Food Product | 546 | 2557 |

To minimize the effect of bias in the interpretation by a single annotator, two annotators were involved in the rating process, and an agreement analysis was conducted using "Cohen's Kappa". The objective of the analysis was to evaluate the reliability of the annotation process. Table 7 represents the Cohen's Kappa scores obtained respectively to each aspect of eWOM opinion. As the Kappa measure for both aspects was above 0.80 [19], the agreement level between the annotators revealed a high level of confidence in the annotation process[18].

Table 7. Cohen's Kappa Score for Annotator rating

| Inter-Annotator Agreement | Sentiment level of rating |
|---|---|
| Cohen's Kappa: Safety Aspect | 0.925 |
| Cohen's Kappa: Price Aspect | 0.956 |

Following the process described in Section H in Methodology, aspect-wise sentiment classification was conducted based on the average weighted score received for each eWOM opinion. Table 8 and Table 9 represent the statistics of the classification, respective to each product aspect, "Product Price" and "Product Safety".

Table 8. Class Label Assignment : Product Price Aspect

| Sentiment Class Label | Overall Polarity Score | Comments Count |
|---|---|---|
| Very Positive | +2 | 128 |
| Positive | +1 | 793 |
| Neutral | 0 | 4 |
| Negative | -1 | 148 |
| Very Negative | -2 | 122 |

Table 9. Class Label Assignment : Product Safety Aspect

| Sentiment Class Label | Overall Polarity Score | Comments Count |
|---|---|---|
| Very Positive | +2 | 295 |
| Positive | +1 | 266 |
| Neutral | 0 | 480 |
| Negative | -1 | 93 |
| Very Negative | -2 | 61 |

Generally, comments/opinions expressed by the users in eWOM platforms are shorter in length since the eWOM user tends to express ideas relevant to the specific Facebook post content. In our study, the average word frequency received for comment was 5.3 approximately. Figure 2 and Figure 3 represent the frequency of the words present in comments of the considered dataset relative to each aspect, after "Stop Word Removal" using the NLTK library.

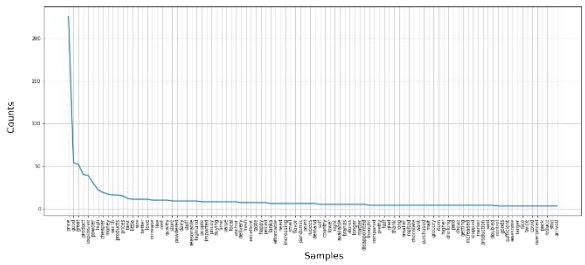



Fig. 3. Word Frequency Plot (100 samples): Price Aspect

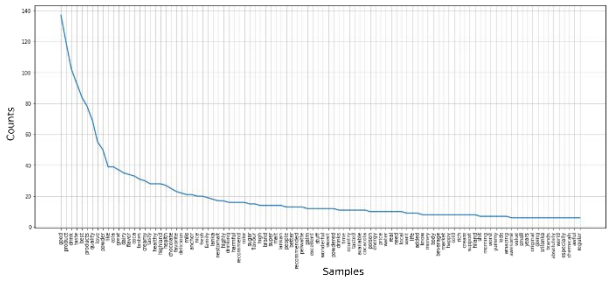



Fig. 4. Word Frequency Plot (100 samples): Safety Aspect

### *E. Experiments*

In this section, the results of the basic experiment conducted with an existing Deep Learning-based Bi-Directional Long Short-Term Memory (Bi-LSTM) to evaluate the performance of the dataset have been presented.

As the initial step, the dataset was subjected to preprocessing. Sequences were generated to be utilized as features for the model. The 80% of the dataset randomly was split in to train split and 20% of the dataset was split as the test set to conduct the experiments with the model. One Bi-Directional Layer was considered with a total of 240 neurons and finally classified the polarity labels by feeding to a softmax activation function, while total trainable parameters for the embedding layer, Bi-Directional layer, and dense layer was 164280, 231360, and 329929 respectively. The model training was conducted considering up to 200 training epochs on the training dataset for both aspects separately. Figure 5 and Figure 6 respectively represent the accuracy score plot obtained for both aspects.

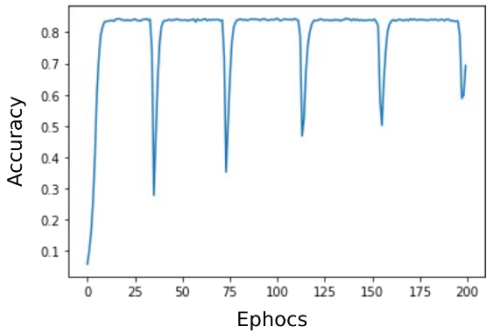



Fig. 5. Accuracy Plot: Safety Aspect

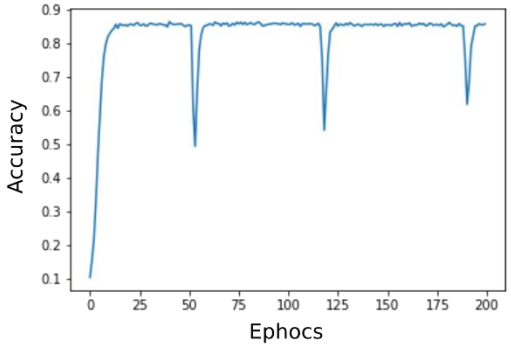



Fig. 6. Accuracy Plot: Price Aspect

At this level of study, a manually annotated eWOM dataset consisting of 1197 opinions was considered for the above experiment. Using the above-explained Bi-LSTM model we were able to obtain an average accuracy score of 0.83 (83%) for the price aspect and 0.85 (85%) for the safety aspect. This dataset can be further improved to utilize for analysis purposes by the parties stated in Section II.

## V. CONCLUSION

This study has been driven with the objective to construct a human-annotated dataset for Aspect-Based Sentiment Analysis in the commercial food products domain with the utilization of the Facebook eWOM entity which is comprised of diverse opinions. The current version of the study presented three sub-tasks of Aspect-Based Sentiment Analysis as aspect category extraction, aspect term recognition, aspect category-wise polarity detection for each comment. Since this version of the dataset performed with a satisfiable accuracy for both "product price" and "product safety" aspects, it can be justified that, the study achieved the main aim of constructing an analysis-ready resource for the commercial food products domain.

## REFERENCES

[1] Shivany, S. and Jayasinghe, S., Critical Discourse Analysis of Sinhala Cultural Identity in Celebrity Endorsement in Advertisements. *Journal of Business Studies*, 3(2), p.56, 2018.

[2] Hettiarachchi, C., Food Act of Sri Lanka: A Systematic Review on Awareness and Practice. *World Nutrition*, *11*(4), 56-65. doi: 10.26596/wn.202011456-65, 2020.

[3] Talagala, I., & Arambepola, C., Use of food labels by adolescents to make healthier choices on snacks: a cross-sectional study from Sri Lanka. *BMC Public Health*, *16*(1). doi: 10.1186/s12889-016-3422-1, 2016.

[4] Hewage, W., & Madhuhansi, T., Impact of Social Media Characteristics on Consumer Buying Decision 47 Process in Western Province, Sri Lanka. *Global Journal Of Management And Business Research: E Marketing*, *19*, 2019.

[5] Tao, D., Yang, P., & Feng, H., Utilization of text mining as a big data analysis tool for food science and nutrition. *Comprehensive Reviews In Food Science And Food Safety*, *19*(2), 875-894. doi: 10.1111/1541-4337.12540, 2020.

[6] Chua, A., & Banerjee, S., Helpfulness of user-generated reviews as a function of review sentiment, product type and information quality. *Computers In Human Behavior*, *54*, 547-554. doi: 10.1016/j.chb.2015.08.057, 2016.

[7] Ganu, G., Elhadad, N., & Martin, A., Beyond the stars: Improving rating predictions using review text content. *Webdb*, *9*, 1-6, 2009.

[8] Oepen, S., Kuhlmann, M., Miyao, Y., Zeman, D., Flickinger, D., & Hajic, J. et al., SemEval 2014 Task 8: Broad-Coverage Semantic Dependency Parsing. *Proceedings Of The 8Th International Workshop On Semantic Evaluation (Semeval 2014)*. doi: 10.3115/v1/s14-2008, 2014.

[9] Pontiki, M., Galanis, D., Papageorgiou, H., Manandhar, S., & Androutsopoulos, I., SemEval-2015 task 12: Aspect based Sentiment anaysis. *Proceedings Of The 9Th International Workshop On Semantic Evaluation (Semeval 2015)*, 2015.

[10] Pontiki, M., Galanis, D., Papageorgiou, H., Androutsopoulos, I., Manandhar, S., & AL-Smadi, M. et al., SemEval-2016 Task 5: Aspect Based Sentiment Analysis. *Proceedings Of The 10Th International Workshop On Semantic Evaluation (Semeval-2016)*. doi: 10.18653/v1/s16-1002, 2016.

[11] Mudalige, C., Karunarathna, D., Rajapaksha, I., Silva, N., Ratnayaka, G., Perera, A., & Pathirana, R., SigmaLaw-ABSA: Dataset for Aspect-Based Sentiment Analysis in Legal Opinion Texts. *2020 IEEE 15Th International Conference On Industrial And Information Systems (ICIIS)*. doi: 10.1109/iciis51140.2020.9342650, 2020.

[12] Valakunde, N., & Patwardhan, M., Multi-aspect and Multi-class Based Document Sentiment Analysis of Educational Data Catering Accreditation Process. *2013 International Conference On Cloud & Ubiquitous Computing & Emerging Technologies*. doi: 10.1109/cube.2013.42, 2013.

[13] Dashtipour, K., Poria, S., Hussain, A., Cambria, E., Hawalah, A., Gelbukh, A., & Zhou, Q., Multilingual Sentiment Analysis: State of the Art and Independent Comparison of Techniques. *Cognitive Computation*, *8*(4), 757-771. doi: 10.1007/s12559-016-9415-7, 2016.

[14] Asanka, D., & →, V., Language Detection in Azure Machine Learning with basic Text Analytics Techniques. Retrieved 2 November 2021, from https://www.sqlshack.com/language-detection-in-azure-machine-learning-with-basic-text-analytics-techniques/, 2021

[15] Fernando, A., & Wijayasiriwardhane, T., Identifying religious extremism-based threats in SriLanka using bilingual social media intelligence. *2020 International Research Conference On Smart Computing And Systems Engineering (SCSE)*. doi: 10.1109/scse49731.2020.9313010, 2020.

[16] Kaur, W., Balakrishnan, V., Rana, O., & Sinniah, A., Liking, sharing, commenting and reacting on Facebook: User behaviors' impact on sentiment intensity. *Telematics And Informatics*, *39*, 25-36. doi: 10.1016/j.tele.2018.12.005, 2019

[17] Rahman, M., & Kumar Dey, E., Datasets for Aspect-Based Sentiment Analysis in Bangla and Its Baseline Evaluation. *Data*, *3*(2), 15. doi: 10.3390/data3020015, 2018.

[18] Alqaryouti, O., Siyam, N., & Shaalan, K., A Sentiment Analysis Lexical Resource and Dataset for Government Smart Apps Domain. *Advances In Intelligent Systems And Computing*, 230-240. doi: 10.1007/978-3-319-99010-1_21, 2018.

[19] Takala, P., Malo, P., Sinha, A., & Ahlgren, O., Gold-standard for topic-specific sentiment analysis of economic text. *LREC*, 2152-2157, 2014.

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index N0: CT-06-44*

# Challenges of Adopting Blockchain Technology to Pharmaceutical Supply Chain – A Case Study From Sri Lankan Health Sector

Pavani Wasana Paththinige (1st Author)
*Department of Industrial Management*
*University of Kelaniya*
Dalugama, Sri Lanka
pavaniwasana22@gmail.com

Dr. Chathura Rajapakse (2nd Author)
*Department of Industrial Management*
*University of Kelaniya*
Dalugama, Sri Lanka
chathura@kln.ac.lk

***Abstract* — Ensuring the transparency of pharmaceutical supply chains is an important task to control the adverse health effects of counterfeit drugs. Blockchain technology has been widely recognized among supply chain researchers as a useful emerging technology to enhance transparency and security of various supply chains. However, the adoption of the blockchain technology in pharmaceutical supply chains is still in its infancy with only a handful of research reported to date. This paper presents the details of a conceptual model developed to explore the challenges of adopting blockchain technology to manage pharmaceutical supply chain while combatting against the flow of counterfeit drugs. The proposed conceptual model, which is based on a comprehensive review of literature, encapsulates the complex linkages between seven influencing factors namely 1) Relative advantage 2) Upper management support 3) Human Resource 4) Compatibility 5) Cost 6) Complexity, and 7) Technological Infrastructure and Architecture. The factors evaluated in the framework interact and impact one another. The proposed framework can be utilized as a starting point for implementing blockchain applications in the pharmaceutical supply chain as well as by academics to develop, refine, and assess blockchain based research. As factors have been identified, practitioners will be able to develop a strategy for implementing blockchain in the pharmaceutical supply chain**

***Keywords* — *Blockchain Technology, Blockchain Adoption, Drug Counterfeiting, Pharmaceutical Supply Chain***

## I. INTRODUCTION

Counterfeit and pharmaceuticals are a widespread issue that has a negative impact on public health. It has become a worldwide problem hence, counterfeit pharmaceuticals are on the rise, making it difficult to keep track of them. Counterfeit pharmaceuticals are common in developing countries including Sri Lanka. According to the World Health Organization WHO, counterfeit medications account for 25% of all medicines in developing countries and 10% of all medicines worldwide [1]. Supply management is a critical issue in all industries, but it is especially important in healthcare because of its increasing complexity. This is because any disruption in the healthcare supply chain has an impact on a patient's health. As the complexity is high, pharmaceutical supply chains are insecure and contain openings for fraudulent attacks.

. The deletion of sensitive information from the Sri Lankan Government Cloud is a severe data loss recently occurred at National Medicine Regulatory Authority, NMRA. The unexpected deletion of sensitive data held by the Drug Regulatory Authority is a contradiction of rules and regulations. It was one of the most serious data corruptions in Sri Lanka, as there was no data backup. These circumstances develop as a result of a lack of transparency and a centralized structure. These situations arise as a result of a lack of transparency and the lack of a decentralized system. Currently By bringing better data transparency and improved product traceability, blockchains enable a safe and secure platform to solve this problem and, in some situations, prevent fraud from occurring. Thus, the aim of this research was to explore all of these challenges and giving insights about the feasibility of adopting blockchain technology to Sri Lankan Pharmaceutical Supply Chain. Built on the findings of this research, a conceptual framework was proposed to be used in the future as well in order to develop strategy for implementing blockchain in pharmaceutical supply chain.

### *A. Why Blockchain*

Blockchain technology came into the world with the introduction of Bitcoin. With the invention of this technology (blockchain platform), it stated that it could be used to perform peer to peer transactions without going through a trusted party [2]. Blockchain is a decentralized and distributed ledger technology that maintain transactional data and is governed by a consensus mechanism. In the blockchain, users can make and verify transactions instantly without the need for a central authority.[3] In general, transaction validation is done by

network members coming to an agreement. It has blocks, which are a collection of transactions linked together by a cryptographic hash. Data that is significant to all parties can be updated in real time, eliminating the need for time-consuming and eliminating fraud. As a result, each member of the network has a far better and more timely view of what is going on in the network.

Blockchain technology is increasingly being used in innovative ways that are relevant to the challenges posed by the COVID-19 pandemic. The inadequacy of traditional systems to provide dependable and effective solutions to challenges caused by the global crisis has increased the value of blockchain applications. It becomes much easier to instantly see the origin of items and to search for complete supplies of an item without developing a centralized database by keeping suppliers and buyers on a shared ledger and recording all transactions on the chain as immutable records. Not only are centralized databases difficult to keep up to date, but they can also be easy targets for hackers [4]. The blockchain's trustless system can help tremendously in decreasing supply chain failures, especially when the epidemics like Covid-19 progresses.

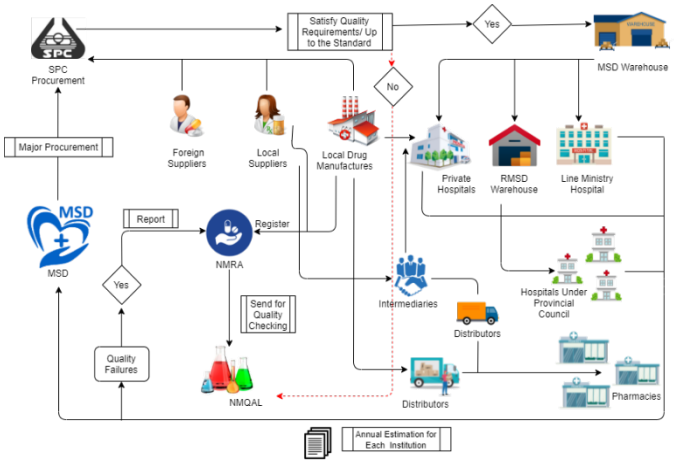



Fig.1. Pharmaceutical supply chain in Sri Lanka

### *B. Pharmaceutical Supply Chain and Reasons for increase in counterfeit drugs*

Fig. 1 depicts the pharmaceutical supply chain in Sri Lanka. The National Medicines Regulatory Authority (NMRA), which is part of the ministry of health, nutrition, and indigenous medicine, as well as the State Pharmaceutical Corporation (SPC), Medical Supply Division (MSD), National Medicines Quality Assurance Laboratory (NMQAL) are in charge of overseeing the procurement, quality assurance, and distribution of pharmaceutical products, including medicines, in Sri Lanka. Under the Ministry of Health, the NMRA and NMQAL are separate authorities. Only drugs that have been registered with the NMRA can be manufactured, imported, and sold in Sri Lanka, according to the NMRA Act. To increase the quality, efficacy, and safety of the medication, there are acts and several health institutes in place. Local manufacturers are inspected by a team of officials linked to the Office of Medical Technology & Supplies and the NMQAL, while overseas manufacturers are evaluated by assessing their company profiles. Counterfeit medicine can enter the pharmaceutical supply chain through foreign suppliers, local suppliers, local drug manufacturers, and supply chain intermediaries, as shown in Fig. 1

### *C. Improve Pharmaceutical supply chain with blockchain*

As there are information losses and barriers at every step of the supply chain, it is extremely difficult to track the supply chain to investigate issues. Buyers and sellers require a trusted way to verify the real value of a product or service. [8]. Through the use of immutable records of data, distributed storage, and regulated user accesses, blockchain technology has the potential to improve traceability and transparency difficulties within the manufacturing supply chain. It would be possible to not only track the production and location of drugs using blockchain as an approach, but also to improve the traceability of falsified drugs [9]. Medical data sharing, privacy, and security can be considerably improved amongst healthcare professionals and the pharmaceutical industry by adopting blockchain technology [10]. Other than that blockchain can face cyber attacks/threats, which improve the reliability of the pharmaceutical supply chain. According to a statistical survey conducted by Statista, the financing investment in blockchain technology has increased dramatically worldwide since 2014 [11].

## II. OBJETIVES

The pharmaceutical sector needs an effective supply chain management system to prevent counterfeit pharmaceuticals, and blockchain technology is the best available strategy for developing a faultless supply chain management system. When data security and privacy are the most crucial factors, it is the best alternative. This study aims to provide a thorough understanding of the numerous barriers to blockchain adoption in the Sri Lankan pharmaceutical sector. The study's findings will allow practitioners to better understand problems and their impact on the adoption process. Researchers and practitioners will be able to better grasp problems and their solutions as a result of the study's findings. This study aims to contribute to two different areas. Firstly, the existing pharmaceutical supply chain in Sri Lanka and potential usage of blockchain technology on supply chain will be identified. In second phase of the study barriers for the adoptability of blockchain will be identified. To the best of the knowledge, no research papers evaluating the adaptability of blockchain technology to the Sri Lankan context have been published. This study attempts to fill the current research gap about the adoptability of blockchain technology to the Sri Lankan Pharmaceutical Supply Chain

## III. METHODOLOGY

### A. Data Collection

The initial step of the analysis was to search the articles relevant to the study. The search criteria included keywords like “Blockchain”, “Supply Chain”, “Pharmaceutical Industry”, and the like. 155 related articles were found through the investigation. This search strategy was needed because the studies on pharmaceutica industry and supply chain management have appeared in a narrow range of publications. A comprehensive literature review was done to discover the knowledge available in the related fields. In databases like as Web of Science, Business Source Complete, Scopus, and Google Scholar, a simple search for the keywords "blockchain" and "adoption" generated many papers. Each paper was thoroughly examined to filter the irrelevant content to the work and finally, 78 articles were selected for the analysis. After analyzing the abstracts of each of the articles, it was discovered that only few of the of them have focused on the adoptability of blockchain. The important parameters were discovered after reading all those papers. At the end 27 papers were selected for the final review

In terms of data collection, the data is gathered through a survey approach and interviews. When it comes to examining a phenomenon of interest, the survey approach and interviews are appropriate (in our case, blockchain adoption in pharmaceutical supply chain). This study incorporated extra items from the existing literature and changed them to match our research environment, similar to other recent studies that employed a survey technique approach and interviews to collect data. Interviews and questionires had been conducted among six groups who are pharmaceutical product manufacturers, pharmaceutical companies (importers), Medical Supply Division, NMRA, Sri Lanka Chamber of the Pharmaceutical Industry (SCLPI), Ministry of Health, and Technical experts. Employees at the director, manager, and executive levels with more than one year of experience were chosen for the data gathering.

### B. Data Analysis

A 7-point Likert scale (range from "strongly disagree" to "strongly agree") is used to assess all constructs. Both qualitative data analysis techniques and quantitative data analysis techniques will be carried out to analyze the data precisely in a “Mixed Method Approach”. The developed hypothesis is tested using PLS SEM. The sample is gathered to meet the ten times rule's minimum sample size criterion for performing the PLS-SEM analysis. The respondent groups for the data gathering are chosen according to their areas of competence.

### C. Model Development

To build the conceptual framework, the authors went through existing technology and literature to determine the most important aspects of blockchain technology. Expert opinions were taken by conducting structured interviews with professionals in the fields of pharmaceutical manufacturing industry and government organizations Formed on the analysis of the data collection and the systematic literature review of the literature, the conceptual model was developed by obeying to all those discoveries. Based on the data collection and the systematic review of the literature, the factors which may have an influence towards the adoption of blockchain technology to Pharmaceuical Supply Chain have been identified as the following. This research shows a number of different elements/factors that need to be considered when evaluating the adoptability of blockchain technology.

- Cost
- Complexity
- Compatibility
- Architecture
- Human Resources
- Relative Advantage
- Technological Infrastructure
- Upper Management Support

### D. Factors affecting adoption of Blockchain Technology

The factors that affect the adoption of blockchain in the pharmaceutical supply chain have been compiled based on a survey of published literature and answers from professionals with 1-20 years of experience in the pharmaceutical industry. A number of researchers have discovered substantial links between supply chain performance and other crucial characteristics, according to a recent study.

*Upper Management Support*

Managerial roadblocks have a significant impact on adoption decisions and are frequently linked to a company's strategic goals, particularly during technological implementation. Upper management commitment, on the other hand, can help spread technology, but it must remain active in order to accomplish the necessary results [12. The provision of training facilities is an important component in this context, as it ensures that employees have access to appropriate training in order to adapt to blockchain technology within the sector. [13].

*Human Resources*

Completely digitizing all supply chain activities utilizing multiple data transmission systems is a significant decision. This will alter the entire way of operation, exacerbating the fear of change is a major reason why businesses are hesitant to accept 'new' technologies. This implies that once the owners believe in change, they will be able to guide their workforce toward effective blockchain adoption [14]. Additionally, management and employee reluctance to change might hinder blockchain technology adoption. Furthermore, organizational culture, which is

defined as the pattern of people's behaviors and practices inside the supply chain, is a crucial component that can impact the adoption of blockchains for performance enhancement [15].

*Compatibility*

Compatibility, which may be described as the ease with which blockchain technology can be integrated on essential platforms, is a key feature in this context [16]. With blockchain's growing popularity and the benefits it can provide to supply chains, pharmaceutical companies might be able to integrate it into their existing infrastructures [17].

*Cost*

One of the most common causes for adoption resistance is financial concerns [18]. Hardware, software, recruitment, and in-house training are just a few of the essential aspects that might affect implementation costs, which include both opportunity costs and accounting costs [19]. Although it has cost-cutting benefits, blockchain is thought to be a technology with high up-front investment costs. Blockchain technology is regarded as an investment because it necessitates the purchase of new gear and software [20].

*Complexity*

From the standpoint of adopters, a blockchain is tough to comprehend in practice. Algorithms may make mistakes that people are unaware of until it is too late to correct. Potential adopters' hesitancy due to concerns about blockchain's complexity may exacerbate opposition, outweighing implementation intentions [21]. In pharmaceutical supply chain operations, supply chain performance is crucial, and achieving it has become increasingly difficult due to the growing complexity of operations in the digital age [22].

*Relative Advantage*

Blockchain has recently attracted a lot of researchers and practitioners' interest and is seen to have the potential to disrupt a variety of industries [23]. "A decentralized, encrypted electronic distributed ledger that operates as an immutable, incorruptible linear event database of information/ transactions shared across networked participants," according to the definition of blockchain. [24]

*Technological Infrastructure*

The availability of infrastructure is another crucial aspect that might influence blockchain adoption, as it ensures that existing technologies are maintained and capable of meeting current infrastructure requirements [26].

*Architecture*

The technology's decentralized nature, as well as its potential for transparency and accountability, can create new environments in which people are less reliant on centralized, sometimes inefficient services provided by related and intermediary service providers [27]. The application of blockchain in supply chain is still in its early stages of acceptance, requiring a protracted deployment lifetime and a lot of unknowns about its significance for business processes.

Information transparency is a major benefit of blockchain in the supply chain. Once data is uploaded to the blockchain, it cannot be withdrawn from the chain and remains accessible for the duration of the blockchain. Despite the fact that blockchains create a decentralized network, they are nevertheless managed by service providers, raising worries about possible illegal spying and data misuse [28].

### *B. Conceptual Framework*

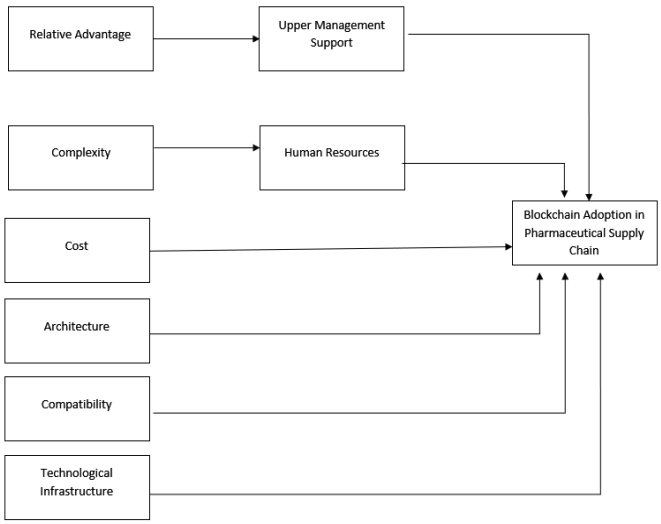



Fig .2. Conceptual Framework

Fig. 2 depicts the conceptual framework. The authors of this study created a conceptual framework for blockchain technology adoption in sri lankan pharmaceutical supply chain based on the literature (Table 1) and the insights from the experts of the Pharmaceutical Industry. The factors affecting the adoption of blockchain technology was examined through systematic literature review and interviews with experts and consultants.

Table 1. Summary of main factors affecting adoption of blockchain

| **Factors affecting adoption of blockchain** | | **References** |
|---|---|---|
| **Relative Advantage** | Increase traceability of supply chain | [5] [6] [9] [10] [11] [13] [14] [16] [17] |
| | Increase transparency of supply chain | |
| | Enhance the visibility and privacy of supply chain | |
| | Enhance trust with descentralized nature | |
| **Upper Management Support** | Upper managers actively respond and pay attention when blockchain based supply chain management project is initiated | [1] [5] [9] [14] [13] [15] [16] |
| | Upper managers support by providing labour resources, materials for blockchain based supply chain project | |
| | Upper managers are willing to accept risks when adopting Blockchain to supply chain | |
| **Human Resource** | Existing knowledge gap | [5] [9] [13] [14] [16] |
| | Mindset of people | |
| | Fear of change | |
| | Organization culture | |
| | Lack of understanding of blockchain technology | |
| **Compatibility** | Compatible with the way we work | [1] [13] [9][15] [16] |
| | Compatible with our operations in supply chain | |
| | Compatible with the supply chain management process | |
| | Compatible with the other information systems used (ERP) | |
| **Cost** | High investment cost | [1] [13] [14] [15] [16] |
| | High cost for training and recruiting | |
| | Maintenance Cost | |
| | Hardware and Software facility cost | |
| **Complexity** | Blockchain is difficult to understand from a technical perspective | [1] [3] [5] [15] [16] |
| | Blockchain is conceptually difficult to understand from supply chain management perspective | |
| | Interacting with the system does not require a lot of mental effort | |
| **Technological Infrastructure** | The current technological structure adequate for blockchain | [1] [9] [11] [15] [16] |
| | Current internet service is efficient enough for blockchain | |
| | There is sufficient infrastructure to adopt blockchain | |
| | Sufficient large computing power | |
| **Architecture** | Lack of Technological maturity | [3] [5] [6] [9] [10] [11] [14] |
| | Data Security | |
| | Performance and Scalability | |
| | Immutability Challenge | |

When addressing blockchain implementation in the pharmaceutical supply chain in Sri Lanka, the conceptual framework suggested in Fig. 2 is a complete conceptual framework that provides an overview of components and their relationships. The framework can be utilized by government bodies and pharmaceutical companies to gain a better understanding of blockchain technology. It emphasizes the need of understanding the factors that influence technology adoption. This highlights the need of comprehending the interaction of factors and materiality during the transformation process, which ultimately influences the adoption of blockchain technology. The framework can be used to comprehend the ramifications of adoption on a larger scale.

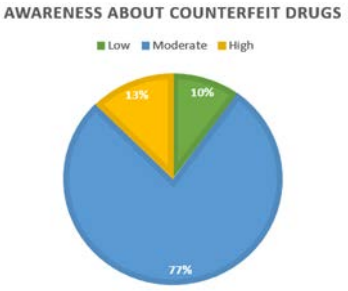



Fig. 3. Awareness about the prevalence of counterfeit drugs in Sri Lanka

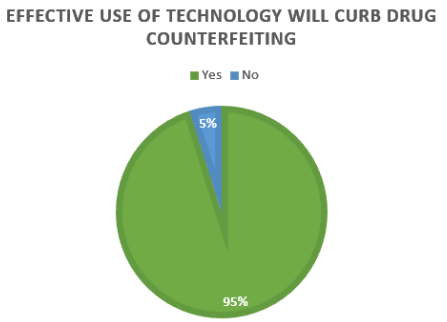



Fig. 4. Effective use of technology will curb counterfeit drugs in Sri Lanka

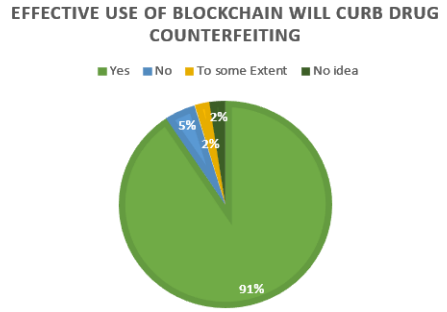



Fig. 5. Effective use of blockchain will curb counterfeit drugs in Sri Lanka

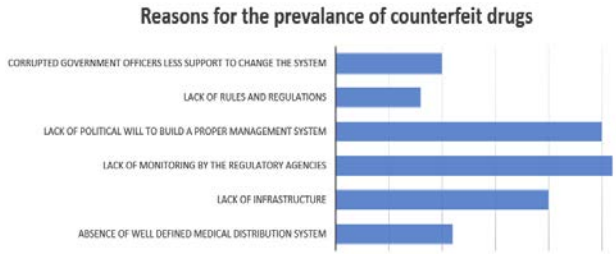



Fig. 6. Reasons for the prevalence of counterfeit drugs in Sri Lanka

In addition to the factors identified for the adoption of blockchain technology to the Sri Lankan pharmaceutical supply chain, 90 % of respondents believe that the availability of counterfeit drugs is moderate/ high in Sri Lanka, according to the investigation (Fig. 3). Moreover, according to the fig. 6, most of the respondets have identified “Lack of monioring by the regulatory agencies” is the main reason for the prevalance of counterfeit drugs in Sri Lanka. Other than that, “Lack of political will to build a proper management system”, “lack of infrastructure”, “Corrupted government officers less support to change the system”, and “Lack of rules and regulations” are the major reasons for the prevalance of counterfeit drugs. As fig. 4 depicts 95% of the respondents believe that effective technology will curb the counterfeit drugs in Sri Lanka and according to the figure 4, 91% of the people believe tha blockchain technology will curb drug counterfeiting. Only 5% of people believe that blockchain technology will not be able to curb drug counterfeiting in Sri Lanka. All of these findings, and the factors identified, can be taken into account when developing a high-quality strategy to combat counterfeit pharmaceuticals from Sri lanka.

## IV. CONCLUSIONS

In contrast to earlier research that focused on blockchain adoption from larger supply chains and sustainable supply networks, the findings of this study provide a better knowledge of the characteristics that promote/enable blockchain adoption in the pharmaceutical supply chain in the context of Sri Lanka. The availability of specific blockchain tools, supporting infrastructure, and government regulation and support for the freight logistics business all influence adoption and implementation. Adoption factors in the pharmaceutical supply chain are discussed in the research as an important part of blockchain technology adoption. This research proposed a new framework of essential criteria that drive blockchain technology adoption in the pharmaceutical supply chain. These constructs could be used as a starting point for more in-depth theoretical research into the elements that influence blockchain adoption in the pharmaceutical supply chain.

From a managerial standpoint, the findings of this study are expected to aid the Sri Lankan pharmaceutical sector, blockchain service providers, and government agencies in focusing on the elements identified in this study in order to successfully implement blockchain. Furthermore, the findings of his research could be applied to other companies in emerging and developing economies and environments that are interested in digitizing their processes to improve transparency and competitiveness, reinforcing their utility. Future and further research into the area will benefit from these additional research avenues.

## REFERENCES

[1] Orji, I.J., Kusi-Sarpong, S., Huang, S. and Vazquez-Brust, D., 2020. Evaluating the factors that influence blockchain adoption in the freight logistics industry. Transportation Research Part E: Logistics and Transportation Review, 141, p.102025.J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[2] Nakamoto, S., 2008. Bitcoin whitepaper. URL: https://bitcoin. org/bitcoin. pdf-(Дата обращения: 17.07. 2019).

[3] Wamba, S.F., Queiroz, M.M. and Trinchera, L., 2020. Dynamics between blockchain adoption determinants and supply chain performance: An empirical investigation. International Journal of Production Economics, 229, p.107791.K. Elissa, “Title of paper if known,” unpublished.

[4] Duan, J., Zhang, C., Gong, Y., Brown, S. and Li, Z., 2020. A content-analysis based literature review in blockchain adoption within food supply chain. International Journal of Environmental Research and Public Health, 17(5), p.1784.

[5] Janssen, M., Weerakkody, V., Ismagilova, E., Sivarajah, U. and Irani, Z., 2020. A framework for analysing blockchain technology adoption: Integrating institutional, market and technical factors. International Journal of Information Management, 50, pp.302-309.

[6] Kamble, S., Gunasekaran, A. and Arha, H., 2019. Understanding the Blockchain technology adoption in supply chains-Indian context. International Journal of Production Research, 57(7), pp.2009-2033.

[7] Karamchandani, A., Srivastava, S.K. and Srivastava, R.K., 2020. Perception-based model for analyzing the impact of enterprise blockchain adoption on SCM in the Indian service industry. International Journal of Information Management, 52, p.102019.

[8] Queiroz, M.M., Fosso Wamba, S., De Bourmont, M. and Telles, R., 2020. Blockchain adoption in operations and supply chain management: empirical evidence from an emerging economy. International Journal of Production Research, pp.1-17.

[9] Francisco, K. and Swanson, D., 2018. The supply chain has no clothes: Technology adoption of blockchain for supply chain transparency. Logistics, 2(1), p.2.

[10] Upadhyay, N., 2020. Demystifying blockchain: A critical analysis of challenges, applications and opportunities. International Journal of Information Management, 54, p.102120.

[11] Werner, F., Basalla, M., Schneider, J., Hays, D. and Vom Brocke, J., 2021. Blockchain adoption from an interorganizational systems perspective–a mixed-methods approach. Information Systems Management, 38(2), pp.135-150.

[12] Zhao, J.L., Fan, S. and Yan, J., 2016. Overview of business innovations and research opportunities in blockchain and introduction to the special issue.

[13] Upadhyay, A., Ayodele, J.O., Kumar, A. and Garza-Reyes, J.A., 2020. A review of challenges and opportunities of blockchain adoption for operational excellence in the UK automotive industry. Journal of Global Operations and Strategic Sourcing.

[14] Caldarelli, G., Zardini, A. and Rossignoli, C., 2021. Blockchain adoption in the fashion sustainable supply chain: Pragmatically addressing barriers. Journal of Organizational Change Management.

[15] Alazab, M., Alhyari, S., Awajan, A. and Abdallah, A.B., 2021. Blockchain technology in supply chain management: an empirical study of the factors affecting user adoption/acceptance. Cluster Computing, 24(1), pp.83-101.

[16] Mathivathanan, D., Mathiyazhagan, K., Rana, N.P., Khorana, S. and Dwivedi, Y.K., 2021. Barriers to the adoption of blockchain technology in business supply chains: a total interpretive structural modelling (TISM) approach. International Journal of Production Research, pp.1-22.

[17] Mathivathanan, D., Mathiyazhagan, K., Rana, N.P., Khorana, S. and Dwivedi, Y.K., 2021. Barriers to the adoption of blockchain technology in business supply chains: a total interpretive structural modelling (TISM) approach. International Journal of Production Research, pp.1-22.

[18] Siyal, A. A., Junejo, A. Z., Zawish, M., & Ahmed, K., Applications of Blockchain Technology in Medicine and Healthcare : Challenges and Future Perspectives. 1–16https://doi.org/10.3390/cryptography3010003, 2019.

[19] Aderibole, A., Aljarwan, A., Habib, M., & Rehman, U. R., Blockchain Technology for Smart Grids : Decentralized NIST Conceptual Model, 8, 43177–43190, 2020.

[20] Khurshid, Anjum., Applying Blockchain Technology to Address the Crisis of Trust During the COVID-19 Pandemic, 8, 1–9. https://doi.org/10.2196/20477, 2020.

[21] Khurshid, Anjum., Applying Blockchain Technology to Address the Crisis of Trust During the COVID-19 Pandemic, 8, https://doi.org/10.2196/20477, 2020.

[22] Khurshid, Anjum., Applying Blockchain Technology to Address the Crisis of Trust During the COVID-19 Pandemic, 8, 1–9. https://doi.org/10.2196/20477, 2020.

[23] Kulkarni, N., Shaikh, A., Kurkure, N., & Bagul, U., A Secure Healthcare System using Blockchain Technology, 964–966, 2019.

[24] Haq, I., & Esuka, O. M., Blockchain technology in pharmaceutical industry to prevent counterfeit drugs. International Journal of Computer Applications, 180(25), 8– 12, 2018.

[25] Jamil, F., Hang, L., Kim, K., & Kim, D., A Novel Medical Blockchain Model for Drug Supply Chain Integrity Management in a Smart Hospital, 1–32. https://doi. Choi, D., Chung, C. Y., Seyha, T., & Young, J., sustainability

[26] Factors A ff ecting Organizations ' Resistance to the Adoption of Blockchain Technology in Supply Networks, 1–37, 2020.org/10.3390/electronics8050505, 2019.

[27] Adsul, K. B., Kosbatwar, S. P., Kajal, M., & Adsul, B., A Novel Approach for Traceability & Detection of Counterfeit Medicines Through Blockchain, 20220.

[28] Labaran, M. J., & Hamma-adama, M., The Nigerian Pharmaceutical Supply Chain : Blockchain Adoption , Counterfeit Drugs and Successful Deployment of COVID-19 Vaccine in. 27(2), 20–36. https://doi.org/10.9734/JSRR/2021/v27i230356,2021.

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index N0: CT-07-45*

# Software Test Effort Estimation Using Machine Learning Techniques

Miyushi Perera (1st Author)
*Department of Computing & Information Systems*
*Wayamba University Of Sri Lanka*
Kuliyapitiya, Sri Lanka
miyushiperera@gmail.com

VGTN Vidanagama (2nd Author)
*Department of Computing & Information Systems*
*Wayamba University Of Sri Lanka*
Kuliyapitiya, Sri Lanka
tharinda@wyb.ac.lk

***Abstract* — Software testing is the method of verifying a software product to recognize any errors, gaps, or missing requirements versus the exact requirements. Manual testing and automation testing are the two strategies of software testing. Testing requires a good amount of time and effort in the software development life cycle. The Software Development Life Cycle includes Planning, Designing, Developing, Testing, and Deploying. Software testing is acknowledged as an essential part of the software development life cycle since it concludes whether the software is ready to be delivered. This paper presents several machine learning techniques for test effort estimation. Support Vector Machine (SVM), KNearest Neighbour (KNN), and Linear regression are the techniques considered for the public dataset namely Desharnais.**

***Keywords* — *software testing, machine learning, effort estimation***

## I. INTRODUCTION

The project manager usually faces the problem of estimating the effort needed to develop a project. This task depends on the software engineers in the team but it can be measured with various procedures mainly Expert estimation, Top-down estimation, Bottom-up estimation, and parametric estimation methods. Some of these procedures are straightforward to apply but require many additional data while the others are time-consuming and difficult to use [1].

Indeed we make fundamental judgments, there are various drawbacks in effort estimation like the software engineers involved in the project and their expertise. But, businesses require to perform effort estimation. And usually, all these estimators are done to estimate the effort for a brand-new project by project managers who use their experience-based judgments and knowledge from previous projects. [2]

It is a great challenge for a software company to develop a new software project of high quality within a predetermined budget and time. Many datasets have been used to estimate software effort that is publicly available in the PROMISE repository which is one of the most famous used repositories in the Software Engineering Community to estimate effort. Desharnais dataset[3] consists of 81 projects collected by J.M. Desharnais in the late 1980s from a Canadian software house. The original dataset consists of 12 attributes but in this study, the ProjectID attribute was omitted from the original dataset because it has no meaning to the study, the left are ten independent attributes and one dependent attribute (effort), all the values in this dataset are numeric but only one nominal attribute that is Language.

Table 1. List of variables in Desharnais dataset

| Symbol | Name |
|---|---|
| TeamExp | Team experience – measured in years |
| ManagerExp | Manager experience – measured in years |
| YearEnd | Year project ended |
| Entities | The number of entities in the systems data model (function points) |
| Transactions | A count of basic logical transactions in the system (function points) |
| Length | Actual project schedule in months |
| PointsNonAjust | Transactions + Entities (function points) |
| PointsAdjust | Function points adjusted by the Adjustment factor = 0.65 + (0.01 * PointsNonAdjust) |
| Adjustment | Function point complexity adjustment factor (Total Processing Complexity) |
| Effort | Actual Effort is measured in person-hours (Dependent) |
| Language | Programming Language |

In this study, we will discuss three machine learning techniques that could be used to predict the accuracy of effort estimation. KNN is one of the techniques used for classification problems, and it is one of the most simple classification techniques that should be the first option for a classification study. KNN works first by computing the distance between an instance with other instances and finding the k-nearest neighbor for that instance, then it estimates the effort.

SVM is a set of machine learning methods used in many areas, such as classification and regression. SVM classifier

separates the instances from two different classes by using a hyper-plane which tries to maximize the margin. This increases the generalization capability of the classifier [4].

Classification and Regression both are types of supervised learning algorithms. Both are working on labeled data set and used for predicting the output. Regression analysis is a predictive modeling technique that estimates the relationship between two or more variables. Regression analysis focuses on the relationship between a dependent (target) variable and an independent variable(s) (predictors). Linear regression is one of the most commonly used predictive modeling techniques. It is represented by,

$$Y = a + b_1 X_1 + b_2 X_2 + \ldots + b_n X_n \quad (1)$$

where a is the intercept, b's are the slopes of the line. This equation can be used to predict the value of a target variable based on a given predictor variable(s).

## II. OBJECTIVES

Machine learning algorithms find natural patterns in data that generate insight and help make better decisions and predictions. They are used every day to perform critical decisions in medical diagnosis, stock trading, energy load forecasting, and more. This research focuses on selecting appropriate machine learning technique for software testing effort estimation and fitting a model.

## III. METHODOLOGY

In this research, SVM, KNN, and Linear regression techniques were used, and the accuracy for each model was compared applying to the Desharnais data.

The final model was obtained for the highest accuracy. The independent variables were identified according to the high correlation coefficient values.

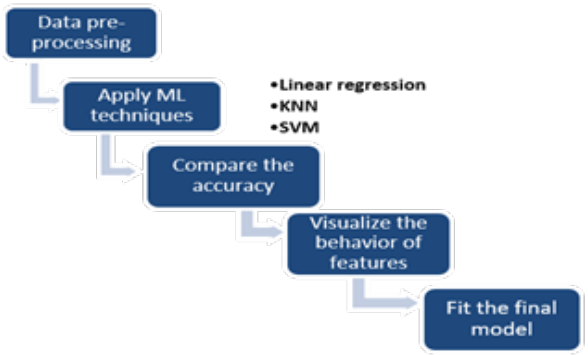



Fig. 1. How the estimation works

## IV. RESULTS AND DISCUSSION

In this section, we will discuss the results of three machine learning techniques.

We get positive correlation coefficient values for the variables Length, Transactions, Effort, PointsNonAdjust, and PointsAdjust among 11 variables described in Table 1.

Table 2. Coefficient of determination ($R^2$) values

| Model | Coefficient of determination ($R^2$) |
|---|---|
| Linear Regression | 0.84 |
| KNN | 0.71 |
| SVM | 0.79 |

We get the highest accuracy for the Linear regression model. The below table represents the coefficient values of each independent variable for the linear regression model.

Table 3. Coefficient values of each variable

| Variable | Coefficient |
|---|---|
| Length | 253.35 |
| Transactions | -16.09 |
| Effort | -4.39 |
| PointsNonAdjust | -20.49 |
| PointsAdjust | 39.29 |
| Intercept | 394.02 |

## V. CONCLUSION

This study was done to evaluate the machine learning techniques by applying them to the Desharnais data set to predict software test effort for a new project.

The results describe the possibility of using the Linear regression method to predict the software effort with a coefficient of determination of 84%, the KNN method of 70%, and SVM with 79%.

We have seen from Table 1, that when we apply the Linear Regression model it has the highest accuracy for effort estimation. According to Equation 1, finally, we can fit the model for effort estimation where Y is Effort, X's are Length, Transactions, PointsNonAdjust & PointsAdjust, b's are the coefficient values and a is the Intercept listed in Table 3.

As mentioned earlier Desharnais data has only 81 project data. Therefore, a large data set can be used to train and test the model with better accuracy.

## REFERENCES

[1] Omar Hidmi, a. B. (2017). 'Software Development Effort Estimation Using Ensemble Machine Learning'. *Int'l Journal of Computing, 4*, 143-147.
[2] Radlinski, L. &. (2010). 'predicting software development effort using machine learning techniques and local data'. *International Journal of Software, 2(2).*
[3] http://promise.site.uottawa.ca/SERepository/
[4] Sayyad Shirabad, J. T. (2005). *'The PROMISE Repository of Software Engineering Databases'*. Retrieved from http://promise.site.uottawa.ca/SERepository/

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index N0: CT-08-51*

# A Deep Neural Network Approach for Analysis of Firewall Log Data

Chandesh Lillmond (1st Author)
*School of Innovative Technologies and Engineering (SITE)*
University of Technology Mauritius (UTM)
Pointe-aux-Sables, Republic of Mauritius
lillmond@umail.utm.ac.mu

Geerish Suddul (2nd Author)
School of Innovative Technologies and Engineering (SITE)
University of Technology Mauritius (UTM)
Pointe-aux-Sables, Republic of Mauritius
g.suddul@umail.utm.ac.mu

***Abstract* — In this paper, we propose an intelligent approach for the classification of incoming and outgoing firewall traffic packets. A firewall is a quintessential tool that ensures the control of traffic over machines' communication over a network. It uses a set of specific rules to define the traffic and thus assists in avoiding cyber-attacks which can be very costly to an organization. Our intelligent approach is mainly through the application of the Deep Neural Network (DNN) Machine Learning algorithm so that packets going through the firewall can be automatically classified as either allow, deny or drop. Our experiments demonstrate a classification accuracy of around 94%, which is higher when compared with other approaches.**

***Keywords — Firewall, DNN, Classification***

## I. INTRODUCTION

A firewall is software that is used to detect malicious data from incoming and outgoing packets on the internet. A firewall normally sets up a wall between a trusted network and an untrusted network, like the Internet It is important to have a firewall installed on a computer as it helps to protect the user from cyber-attacks and corrupted data. The main job of a firewall is to inspect packets transferred between computers. The firewall has a set of rules which indicate the appropriate action to each packet. There are 3 main actions regarding a packet is "allow", "deny" and "drop". With the help of machine learning, the firewall can become more accurate to classify the packets to their corresponding set of rules, which can result in a more secure network everywhere.

The main objective of this research is (i) to propose the modelling and implementation of a classifier model based on Deep Learning techniques and, (ii) to perform a comparative study that evaluates the performance of the proposed model.

## II. LITERATURE REVIEW

The following segments portray different work completed by researchers utilizing different machine learning algorithms and methodologies carried out. This study [2] analysed data of a network using supervised machine learning techniques. To classify the data set obtained from the UCI machine learning repository, a self-organizing feature map (SOFM) and K-means algorithms were used. The authors achieved an accuracy of 97%.

This paper [3] proposed a classification framework that can be utilized in the firewall frameworks to create a legitimate classification for every transmitted packet by breaking down packet parameters by using shallow neural network (SNN) and optimizable decision tree (ODT) as machine-learning methods. In particular, the proposed models were utilized to prepare and group the Firewall dataset into three classes: "allow, "deny," and "drop/reset." The experiment scored an accuracy of 98%, and 99% for SNN, and ODT respectively.

[4] Analysed 500,000 instances using 6 features, which have been generated from Snort and TWIDS. The Action parameter was selected as the class attribute. The "Drop" and "Allow" parameters have been specified for the Action class. The firewall logs dataset was analysed and the features were inserted into machine-learning classifiers including Naive Bayes, kNN (k-Nearest Neighbours), One R and J48 using Spark in the Weka tool. The authors also compared the classification accuracy of these algorithms in terms of measurement metrics including Accuracy, F-measure and ROC values.

This experiment [5] proposed predictive models for predicting the work status at the finishing stage in the HPC framework. The model can be utilized as a device for checking the jobs in the HPC framework. The authors developed and built three models including HPC-CNN, HPC-AlexNet, and HPC-VGG16 based on the two machine learning techniques, which involved Initial and Transfer Learning of Convolutional Neural Network based on the HPC-work load dataset. Moreover, the three state-of-the-art Machine Learning methods: Artificial Neural Network (ANN), Classification and Regression Tree (CART) and Support Vector Machine (SVM) are used as the baseline models for performance comparison. The results show that the model that performs the best predictive performance is the proposed HPC-CNN model. The authors achieved 76.48% accuracy with the HPC-CNN

model followed with the CART model (75.60%), while the SVM model performs lowest the accuracy at 66.80%.

This study [6] classified some data from the Firewall Device used at a university, using a machine learning algorithm called multi-class support vector machine (SVM) classifier. As activation function for SVM classification they used linear, polynomial, sigmoid and Radial Basis Function (RBF). The authors measured the performance of the classifier by observing the estimation esteems of F1 score, recall and precision. The Action column is selected as the class attribute. The “deny”, “drop”, “allow” and “reset-both” attributes were implemented for the Action class. To obtain the maximum precision value in the classifier, SVM responses have been evaluated. The authors attempted to acquire the best activation function for the F1 score value. For each class, Receiver Operating Characteristic curves were also done.

## III. METHODOLOGY

In this study, the Internet Firewall Dataset from UCI Machine Learning Repository was used [8]. The classification process is commonly done by coordinating the network packets against a set of guidelines and rules to block digital dangers from accessing the network. Subsequently, the firewall framework continues with either to "allow," "deny," or "drop/reset-both" the approaching packet. [1]

To classify the firewall log information, 11 of the characteristics in the informational index were chosen. While choosing information, it is critical to choose credits that have more numerical qualities. The action parameter has been acknowledged as a class. Table 1 shows the parameters and their description. [6]

Table 1. Dataset Columns and Description

| Columns | Description |
|---|---|
| Source Port | The client source port number |
| Destination Port | The client destination port number |
| NAT Source Port | Network Address Translation Source Port Number |
| NAT Destination Port | Network Address Translation Destination Port Number |
| Action | allow, deny, drop, reset-both |
| Bytes | Total Bytes |
| Bytes Sent | Bytes Sent |
| Bytes Received | Bytes Received |
| Packets | Total Packets |
| Elapsed Time (sec) | Elapsed Time |
| Pkts_sent | Packets Sent |
| Pkts_received | Packets Received |

There are 4 parameters in the action attribute used as a class. Descriptions of these parameters are shown in Table 2. [6]

Table 2. Dataset Action and Description

| Action | Description |
|---|---|
| Allow | Explicitly allows traffic that matches the rule to pass |
| Deny | The firewall sends an ICMP type 3 (destination unreachable) message response back |
| Drop | The packet will be drop |
| Reset-both | A TCP reset is sent to both the client-side and server-side devices |

Imbalanced classification represents a test for predictive modelling as the vast majority of the machine learning utilized for classification was planned around the assumption of an equivalent number of attributes for each class. This outcome in models that have poor predictive execution, explicitly for the minority class. As shown in Fig. 1, the dataset is imbalanced.

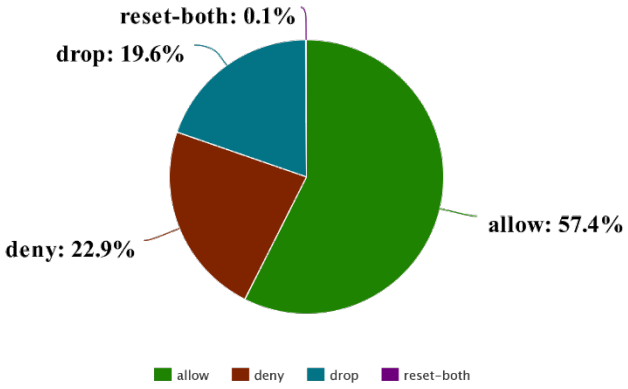



Fig. 1. Unbalanced Dataset

To counter this problem, the dataset needs to be balanced with resampling methods. The “reset-both” class has too few data to be re-sampled, so we merged “reset-both” and “deny” classes into one class. Undersampling was used to balance the uneven dataset by keeping each of the data in the minority class and diminishing the size of the larger part class.

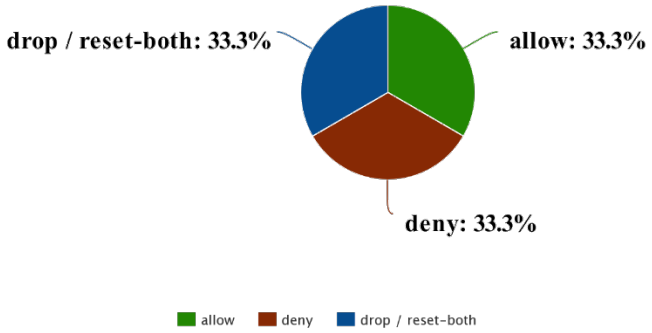



Fig. 2. Balanced Dataset

To have this prediction take place, we had to remap the variables of the Action column to 0, 1, 2 because the model does not work well with string values.

Table 3. Action and Mapped Value

| Action | Mapped Value |
|---|---|
| Allow | 0 |
| Deny | 1 |
| Drop / Reset-both | 2 |

For this model, an artificial neural network classifier with a standard scaler was used. Two hidden layers were configured with 10 and 5 neurons respectively. The L2 penalty was also configured to reduce the risk of overfitting.

The study was carried out on a laptop with specifications shown in Table 4. Anaconda [9] was used as software and scikit-learn [10] and imbalanced-learn [11] as main libraries.

Table 4. Laptop Specifications

| | |
|---|---|
| CPU | Intel(R) Core (TM) i5-1035G1 CPU @ 1.00GHz, 4 Cores, 8 Logical Processors |
| GPU | Intel (R) UHD Graphics |
| RAM | 8 GB DDR4 |
| Hard Disk | 1TB 5.4K RPM SATA Hard Drive |

To build a more generalized model which can perform well on unseen data, k-fold cross-validation was used to partition the data. In this model, the data is split into 5 folds as this value have been shown exactly to yield test error rate estimates that experience neither from excessively high bias nor from very high variance

To avoid overfitting, a simple hidden layer was used with the default iteration and the L2 penalty parameter. The most common term for L2 penalty is L2 regularization. L2 regularization attempts to decrease the chance of overfitting by keeping the values of the weights and biases small.

To assess the performance of the models, four evaluators including accuracy, precision, recall, and F1-score were selected. All evaluators are figured from the confusion matrix table. [1]

The true positive (TP) is the number of the predicted data is “True”, and the actual data is “True”. The false-negative (FN) is the number of the predicted data is “False”, while the actual data is “True”. The false positive (FP) is the number of the predicted data is “True”, while the actual data is “False”. The true negative (TN) is the number of the predicted data is “False”, and the actual data is “False”. [1] [12]

The accuracy (1) is an evaluator that assesses the overall performance of the model. The recall (2) regards the model performance based on the actual value point view. Meanwhile, precision (3) observes the model performance based on the predicted value point of view. The F-measure or F1 score (4) is a harmonic mean of precision and recall. [1]

$$\text{Accuracy} = \frac{\text{TN} + \text{TP}}{\text{TN} + \text{FN} + \text{TP} + \text{FP}} \quad (1)$$

$$\text{Recall} = \frac{\text{TP}}{\text{FN} + \text{TP}} \quad (2)$$

$$\text{Precision} = \frac{\text{TP}}{\text{FP} + \text{TP}} \quad (3)$$

$$\text{F}_1 = 2\,\text{X}\,\frac{\text{Recall X Precision}}{\text{Recall} + \text{Precision}} \quad (4)$$

## IV. RESULTS AND DISCUSSION

Fig. 3 shows the confusion matrix which predicts an evaluation metric for the classification model of train values.

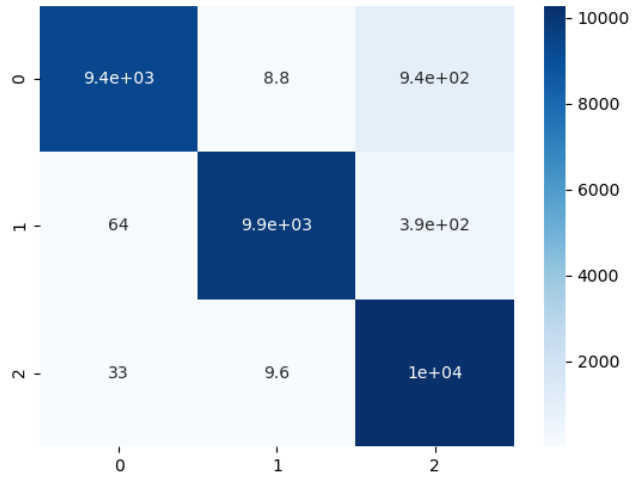



Fig. 3. Confusion Matrix for Train data

Table 5. Train Accuracy

| Precision | Recall | F1-score | Accuracy |
|---|---|---|---|
| 96.17% | 95.81% | 95.84% | 95.81% |

Fig. 4 shows the confusion matrix which predicts an evaluation metric for the classification model of test values.

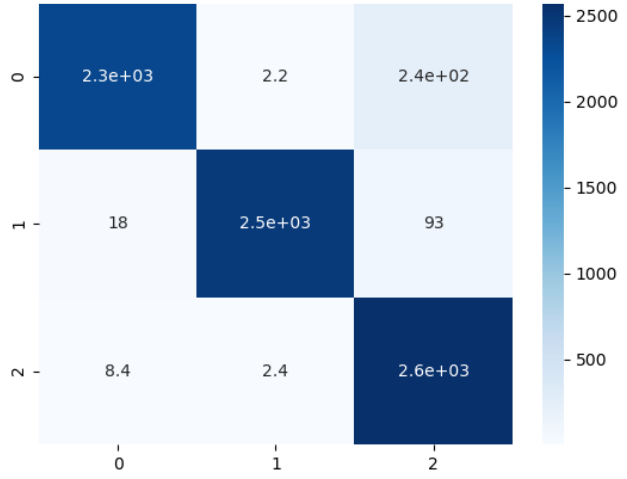



Fig. 4. Confusion Matrix for Test data

Table 6. Test Accuracy

| Precision | Recall | F1-score | Accuracy |
|---|---|---|---|
| 95.45% | 94.93% | 94.97% | 94.49% |

Fig. 5 shows the train v/s test accuracy during the 5-flod cross-validation.

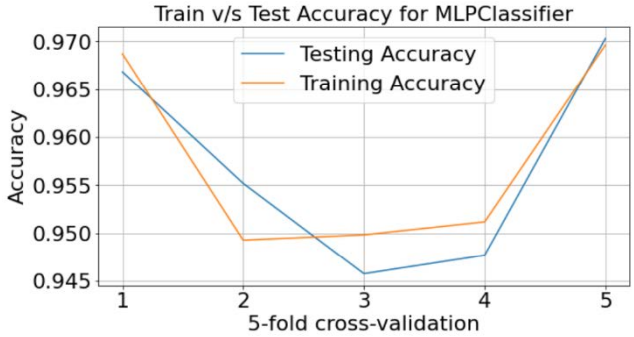



Fig. 5. Train v/s Test accuracy

Table 6 illustrates the test result for classifying the 3 classes log event. The model is giving accuracy up to 94.49%. Hence using our model, the network log data auditing and analysis can be done most optimally.

To acquire an understanding of the proposed model's benefits, our model was analysed by contrasting its accuracy with other best AI-based firewall-activity classification models in terms of classification performance metrics. The examinations are given in Table 7 beneath.

Table 7. Existing models and their accuracy

| Research | Year | ML Technique | Accuracy |
|---|---|---|---|
| Fatih Ertam et.al [6] | 2019 | Support Vector Machine (SVM) | 79.40% |
| Anupong Banjongkan et.al [5] | 2020 | Convolutional Neural Network (CNN) | 76.50% |
| Shridhar Allagi et.al [2] | 2020 | Self-Organizing Feature Map (SOFM) | 97.20% |
| Adrian Pıru et.al [7] | 2019 | Deep Neural Network (DNN) | 92.82% |
| Our Model | 2021 | Deep Neural Network (DNN) | **94.49%** |

Our model has a higher classification accuracy than most of the other existing related machine learning-based models.

## V. CONCLUSION

The main objective of this paper was to develop a machine learning model, using Deep Neural Network representations. 11 features from the Internet Firewall Dataset from UCI Machine Learning Repository were used. As the dataset was imbalanced, undersampling method was used to create a generalized model. Our experiments made use of the k-folds technique and regularization to avoid overfitting. The proposed model has achieved 94.49% accuracy with testing data and 95.81% with training data. When comparing the accuracy of our approach with other research works, we noticed that ours performs better.

As future work, the plan is to integrate this trained model into a live system to better assess its generalization performance.

## REFERENCES

[1] Qasem Abu Al-Haija and Abdelraouf Ishtaiwi,"Machine Learning Based Model to Identify Firewall Decisions to Improve Cyber-Defense," International Journal on Advanced Science, Engineering and Information Technology, vol. 11, no. 4, pp. 1688-1695, 2021. [Online]. Available: http://dx.doi.org/10.18517/ijaseit.11.4.14608.

[2] Shridhar Allagi, Rashmi Rachh. "Analysis of Network log data using Machine Learning" , 2019 IEEE 5th International Conference for Convergence in Technology (I2CT), 2019

[3] Abu Al-Haija, Qasem & Ishtaiwi, Abdelraouf. (2021). Machine Learning Based Model to Identify Firewall Decisions to Improve Cyber-Defense. International Journal on Advanced Science, Engineering and Information Technology. 11. 1688. 10.18517/ijaseit.11.4.14608.

[4] As-Suhbani, Hajar & Khamitkar, S. (2019). Classification of Firewall Logs Using Supervised Machine Learning Algorithms. International Journal of Computer Sciences and Engineering. 7. 301-304. 10.26438/ijcse/v7i8.301304.

[5] A. Banjongkan, et. al., "A Comparative Study of Learning Techniques with Convolutional Neural Network Based on HPC-Workload Dataset" Inter. Journal of Machine Learning and Computing, vol. 10, no.1, 2020.

[6] Fatih Ertam, Mustafa Kaya. "Classification of firewall log files with multiclass support vector machine" , 2018 6th International Symposium on Digital Forensic and Security (ISDFS), 2018

[7] A. I. Pîrîu, M. Leonte, N. Postolachi and D. T. Gavrilut, "Optimizing Cleanset Growth by Using Multi-Class Neural Networks," in Proc. Of 20th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC), Romania, pp. 425-429, 2018.

[8] Archive.ics.uci.edu. 2021. UCI Machine Learning Repository: Internet Firewall Data Data Set. [online] Available at: https://archive.ics.uci.edu/ml/datasets/Internet+Firewall+Data [Accessed 13 November 2021].

[9] Anaconda. 2021. Anaconda | The World's Most Popular Data Science Platform. [online] Available at: https://www.anaconda.com/ [Accessed 13 November 2021].

[10] Scikit-learn.org. 2021. scikit-learn: machine learning in Python — scikit-learn 1.0.1 documentation. [online] Available at: https://scikit-learn.org/stable/ [Accessed 13 November 2021].

[11] Imbalanced-learn.org. 2021. imbalanced-learn documentation — Version 0.8.1. [online] Available at: https://imbalanced-learn.org/stable/ [Accessed 13 November 2021].

[12] Kenta Sakamoto, Masaaki Okabe, Hiroshi Yadoshisa. "Generalized canonical correlation analysis for labeled data" , Procedia Computer Science, 2021

[13] Ranganathan, G., Fernando, X., Shi, F. and El Allioui, Y., 2021. Soft computing for security applications.

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index N0: CT-09-52*

# An Approach to Detect Fileless Malware that Maintains Persistence in Windows Environment

Malmi Atapattu (1st Author)
*Department of Industrial Management*
*University of Kelaniya*
Dalugama, Sri Lanka
malminatasha@gmail.com

Buddhika Jayawardena (2nd Author)
*Department of Industrial Management*
*University of Kelaniya*
Dalugama, Sri Lanka
buddhikaj@kln.ac.lk

***Abstract* — The rapid enhancement of the Internet in the past few years has increasingly impacted the general public's work and life. As a drawback, this enhancement has also led to a major increase in malicious software on the internet causing great security threats to the consumers of the internet. Currently, a new type of malware class called Fileless malware has come into action causing more destructive damages. As the name Fileless suggests, these types of malware programs are not files or executables, but a malicious activity that runs entirely in the memory, leaving the slightest evidence on the targeted host machine. Microsoft Windows is one of the most widely used operating systems both in personal desktop computers and enterprise computer systems and is highly targeted by Fileless malware. This paper provides an approach to detect fileless malware that maintains persistence in the Windows environment using Fileless malware behavioural data and deep learning-based classification models.**

***Keywords — fileless malware, windows, deep learning***

## I. INTRODUCTION

It is predicted that the cost of global cybercrime will increase by 15 percent annually over the next five years which is estimated to reach $10.5 trillion USD annually by 2025. This projected cost includes theft of intellectual property, loss of personal and financial data, stolen money, restoration of destructed infrastructure and reputational harm [1]. The defensive systems, techniques and tools are improved continuously to avoid cybercrimes and to reduce the harm of consequences. But cybercrimes are consistently growing mainly due to the facts like an increase in the number of internet users, increase in the number of connected devices, increase in the variety and complexity of services, and increase of digitalization [2]. Almost every day technology consumers are introduced to new types of malware that cause them severe destructions including privacy-related issues. According to AV-TEST security reports, nearly 140 million new malware pieces have been introduced only during 2021.

Malware is basically a piece of software or a code that is inserted into a computer system with the objective of compromising the computer functions, stealing data, or evading access control without user knowledge. In order to evade detection, a malware uses different types of more advanced and sophisticated techniques. Thus, some variant of malware uses a technique that is running within the memory rather than from a file and those are called fileless malware [3]. For ease of reference, fileless malware can be categorized in various ways such as initially by their entry point, secondly by the form of the entry point, and finally by the type of the infected host [4]. A fileless malware is developed to be concealed from antivirus software making the detection difficult. The attack lifecycle of a fileless malware consists of various stages where stage 1 focuses on the initial delivery of attack, stage 2 targets the persistence, stage 3 is the execution and completes the attack once the objectives are met [5]. These stages are illustrated in Fig. 1.

| **Stage 1** Delivery | **Stage 2** Persistence | **Stage 3** Execution | **Achieve Objective** |
|---|---|---|---|
| • Exploits<br>• Malicio-us Macro<br>• Scripts<br>etc | • WMI subscriptions<br>• Registry<br>• Task scheduler<br>etc | • PowerShell<br>• JavaScript<br>• VBScript<br>• Command line tools<br>etc | |

Fig. 1. Lifecycle of filelesss malware attack.

Stage 1. Delivery: Social engineering methods are often used by these attackers to commence the initial delivery. In this stage, two main strategies are used to evade detection by antivirus signature scanning which are, downloading directly into memory and use of trusted applications. Attackers tend to get a whitelisted and approved application; thus security software will not inspect legitimate software [5].

Stage 2. Persistence: To achieve persistence, attackers store the malicious code in unusual locations in the operating systems or common utilities such as the Windows registry, Windows Management Instrumentation (WMI) store, SQL tables or scheduled tasks. Code is injected into a system process and will seem like coming from a legitimate process [5].

Stage 3. Execution: After the persistence, malware will depend on Windows internals like PowerShell, JavaScript and Macro execution of the official documents and other legitimate executables to commence the execution [5].

Fileless malware attacks are initiated directly on most Windows applications and system administration tools like PowerShell and Windows Management Instrumentation (WMI) to exploit and spread the infection [6]. One of the main concerns behind Windows being a victim of fileless malware is the development of the Microsoft .NET framework. Though .NET was able to provide a significant contribution to enhance the software development process, unintentionally revolutionized the malware industry by providing a much easier surface for the malware coders to develop and spread malware and attain their unethical goals [5]. Windows administrative tools which are already installed on the victim's software can be easily used to launch the initial infection and also to commence the attack which is the main reason for the fileless malware to be a serious concern for Windows users [7].

According to WatchGuard's Internet Security Report for Q4 2020, fileless malware attack rates had grown by nearly 900% in 2020 compared to 2019 [8]. Due to the unforeseen fileless malware rising trends, many organizations are at a higher risk. Because of the fileless nature of these types of malware pieces, they are difficult to be detected by traditional anti-virus software. Windows OS is more vulnerable to fileless malware because of its system administration tools like PowerShell, WMI and also applications like Microsoft Office Macros. Though Microsoft has implemented next-gen anti-virus capabilities like Antimalware Scan Interface (AMSI), according to Sophos, AMSI is not a remedy for fileless malware since attackers are continuously finding methods to obfuscate malicious codes and bypass these kind of anti-virus solutions [9]. Therefore, same as other complex malware, fileless malware detection has been a serious concern for today's technological world.

Various traditional mechanisms [5], signature-based [10],[11] and heuristic-based techniques have been used to detect fileless malware. Due to the considerable drawbacks of signature-based techniques, some techniques have been implemented using heuristic-based machine learning and deep learning approaches. As machine learning models, classification algorithms like Perceptron [7], SVM, RF, XGB [12] and as deep learning models, MLP and CNN [13] have been mainly used. Most of the research works are limited to detecting PowerShell based malware and also there is a least deliberation towards malware persistence. Because of the freshness of this topic and the complexity of these types of malware, there is a lack of research regarding fileless malware detection and prevention techniques [7]. Thus, there is a noticeable gap in proper mechanisms to detect fileless malware that maintain persistence in the Windows environment.

The purpose of this research is to get a systematic approach to develop a solution that detects fileless malware which maintain persistence in the Windows environment. The research objectives are as follows: To identify the current techniques and mechanisms used by intruders to initiate fileless malware attacks on the Windows environment. Then, to identify the static and dynamic behaviours of fileless malware on Windows environment. Next, to develop a model to detect fileless malware that maintain persistence in the Windows environment using the identified behavioural data. Finally, to identify a suitable mechanism to test the validity of the proposed model.

In this study, a systematic approach is taken to address the concern of how to detect fileless malware that maintain persistence in the Windows environment. This approach includes the analysis of the current fileless malware techniques and the development of a mechanism to detect Windows fileless malware. The remainder of the paper is organized as follows: In section 2, a discussion is made on the review on the existing literature. Then in section 3, the methodology is described. Results and discussions are explained in section 4. Finally, in section 5, the conclusion is included with limitations and future work.

## II. LITERATURE REVIEW

This study is an approach to investigate and understand the techniques that are being used to initiate fileless malware in the Windows environment by reviewing the existing literature. In addition, the identified techniques are mapped into an industry-standard framework called MITRE ATT&CK framework for further analysis.

In [5], fileless malware techniques are categorized based on the evasion techniques that are targeted by the malware authors which are malicious documents, malicious scripts, living off the land, and malicious code in memory. According to [14], fileless malware attack techniques are categorized into memory-resident malware, Windows registry malware, rootkits, process hollowing/injection attacks, reflective DLL injection, dynamic data exchange attacks, and dual-use tool attacks. Since this study mainly focuses on malware persistence, further discussion is made on persistence attack techniques.

Persistence attacks are mainly targeting to existing inside the victim's system for a longer period of time until the attack's goal is accomplished. To gain persistence, attackers use many evasion mechanisms such as malicious documents and scripts and living off the land techniques. Adversaries often compromise the legitimate Windows administration tools like PowerShell and WMI to evade detection and maintain persistence inside the system [15]. Further, these malware pieces are typically stored in unusual locations in the system like operating system utilities, WMI store, SQL tables, Windows registry, or OS task scheduler. While maintaining persistence, attackers look for the vulnerabilities in the system to exploit and steal data. Most of the time living in the system and running the fileless malware in the background as a service leads to a successful attack since the detection is difficult.

In Windows registry malware attacks, attackers mainly target to embed the malware deeply inside the Windows registry [14]. Windows registry is a system-level database that stores settings which are required for the operations of Windows OS and some applications. The malicious payload is injected into the registry and after accomplishing the malicious objective it can destruct itself without leaving traces. There are several attack methods in Windows registry malware. One is adding JavaScript code into the registry and

the code is executed when a legitimate application is running. Another method is process hollowing which is replacing the legitimate process in the memory with a malicious payload by injecting. More sophisticated attacks are capable of performing multiple process injections thus, the detection is entirely difficult.

In addition to the review on literature, these identified techniques were mapped with the MITRE ATT&CK framework which is a behavioural model and a curated knowledgebase created by the MITRE cooperation that includes information regarding cyber adversary behaviour signifying the different stages of an adversary's lifecycle and their targeted platforms. With this analysis, new possibilities of fileless malware techniques, tools, attack procedures are identified along with examples of the malware types that use the techniques. Table 1 defines the findings of the analysis.

Table 1. Techniques, tools, attack procedure and malware examples from ATT&CK framework

| Technique | Tool | Attack Procedure | Malware Examples |
|---|---|---|---|
| Use of command and scripting interpreter | PowerShell | Abusing PowerShell commands and scripts for execution of malware, information discovery, download and run executables from the Internet. | APT19, APT28, APT29, PowerSploit, njRAT |
| | Windows Command Shell (cmd) | Batch scripting to automate the execution of the malware. Opening a reverse shell or a remote shell on the system to execute commands. | 4H RAT, ABK, abdupd, admin@338, APT1, APT18, APT41 |
| | JavaScript | Abusing various implementations of JavaScript to execute malicious behavior like hosting malicious scripts on websites or downloading and executing scripts as secondary payloads. | Poweliks, Astaroth, FIN6 |
| Process Hollowing | API calls | Creating a process in a suspended state and hollowing its memory which then be replaced with malicious payload. Use of native Windows API calls like CreateProcess, ZxUnmapViewOfSection, NtUnmapViewOfSection, VirtualAllocEx | Agent Tesla, Astaroth, Bazar, Smoke Loader |
| Process Doppelganging | Windows Transactional NTFS (TxF) | Exploiting TxF to replace the memory of a legitimate process. | Bazar, Leafminer, SynAck |
| DLL/ PE Injection | PE files | Before loading the DLL, writing the path in the virtual address space of the target process. | Aria-body, ComRAT, PowerSploit |
| | | Use of Windows native API calls such as VirtualAllocEx, WriteProcessMemory and CreateRemoteThread. | |
| Modify Registry | Command line utilities | Hiding configuration information within Registry keys. Gaining persistence in the system. Use of Reg command line utility for registry modification. Adding an entry to the "run keys" in the Registry hives or in startup folder. | ADVSTORES HELL, BADCALL, Cardinal RAT, Netwalker |
| Rootkits | Master Boot Record, System Firmware | Intercepting and modifying system API calls while living in the kernel level | ZeroAccess |
| Dynamic Data Exchange | Visual Basic for Applications (VBA) macros | Infecting MS Office applications with DDE commands | APT28, APT37 |

## III. METHODOLOGY

This section describes the methodology utilized to perform the research, including the behaviour analysis of fileless malware using a sandbox environment and data collection. Finally, to determine a suitable deep learning based fileless malware classification model.

### *A. Fileless Malware Behavior Analysis*

In this section further discussion is made on the behaviour of fileless malware using the behaviour reports of Poweliks malware from the Cuckoo sandbox.

Cuckoo sandbox is an open-source automated malware analysis sandbox [16], which consists of features like tracing API calls and general behavior of the files, and advanced memory analysis [17]. When creating the sandbox environment for the analysis, Cuckoo sandbox was chosen since its advanced features and functionalities. For the host environment a PC with 8GB memory, 1TB hard disk and 4 core processer was used. Ubuntu LTS 18.04 was installed as the operating system for the host environment. The guest environment was created using Virtual Box 5.2.4 with hardware specifications of 2GB memory and 32GB hard disk. Windows 7 operating system was used for the virtual machine and Google Chrome, MS Office applications, and Adobe Reader was installed as an additional software.

Poweliks malware is one of the first fileless malware that has the ability to do a Ransomware infection on a computer system [18]. According to Symantec report [19] on Poweliks malware was an evolution of the file-based malware called Wowliks. This malware uses registry manipulation and persistence techniques during the attack. When the JavaScript script which contains the malicious payload runs, a new registry entry will be added to the registry. This is then used

for the fileless execution. Using an alternate data stream, the original file is deleted leaving no traces in the system but the registries will be maintaining the persistence continuously contacting the malicious command and control server for information stealing or other purposes.

A sample Poweliks trojan from 'VirusShare' was executed on the Cuckoo sandbox environment for behaviour analysis. Using the JavaScript command, it had created a registry key to a long series of bytes to store the malicious configurations. As a persistence technique, it had installed itself for autorun at Windows startup which is shown in Fig. 2.

| | | | |
|---|---|---|---|
| reg_key | HKEY_CURRENT_USER\Software\Microsoft\Windows\CurrentVersion\Run\ | reg_value | rundll32.exe javascript:"\..\mshtml, RunHTMLApplication ";document.write("\74script language=jscript. encode>"+(new ActiveXObject("WScript. Shell")). RegRead("HKCU\\ software\\microsoft\\windows\\currentversion\\run\\\" )+"\74/ script>") |

Fig. 2. Installation at Windows Startup for autorun.

Moreover, it allocated execute permission to another process and resumed a suspended thread in a remote process which is potentially indicative of possible code injection. Windows command-line utility had used for the command execution and other applications were not used during the execution. Further, this malware had tried to reach some IP addresses rather than a domain which did not respond back that can be potentially indicative of command-and-control traffic. To conclude, this behaviour of Poweliks malware which is a real-world example of fileless malware indicates previously discussed evasion techniques.

### B. Data Collection

'VirusShare' malware repository and 'theZoo' malware repository was used to get the malware samples. The selected malware sample includes malicious Windows PE files, .NET binaries, known Windows fileless malware, PowerShell and WMI based malware. To create the dataset, initially, each malware was executed in Cuckoo sandbox environment and collected the reports. For this purpose, the same environment that was used for the behavior analysis was used with same conditions. Each collected report is in JSON format and contains all the information regarding the execution process including API calls, memory buffer and string values. All together 3085 malware sandbox reports were collected. Benign malware samples were also collected from Portable Freeware collection [20]. The collected sample also executed in the same sandbox environment and 1100 benign malware reports were collected.

### C. Data Preprocessing

For data preprocessing, API calls were selected from registry and system categories because according to the identified fileless malware techniques system and registry related API sequence will be effective when determining fileless persistence attacks. API call sequences for each malware was converted to the given unique index value of the sandbox report. The same technique was used for the benign samples as well. After the sequential data, in the last column of the CSV file '1' was given if it was malware and '0' was given if it was benign malware. The prepossessed data was then used to train the sequence data classification model.

### D. Sequence Data Classification

API call sequence represents the behaviour of the malware across certain period. This behavioural data can be trained using a classification algorithm to develop a model that detects fileless malware from its behaviour. For this purpose, deep learning classification was chosen because, as specified by the authors, there is a higher detection ability for deep learning when compared to traditional shallow learners such as SVM [21]. For this paper, two deep learning-based sequential classification models that are LSTM and BI-LSTM models were used to train the dataset.

*1) Long short-term memory (LSTM):* LSTM is a variant of the recurrent neural network (RNN) which was proposed by Hochreiter and Schmidhuber who had applied three gates to solve the vanishing gradient problem of RNN [22].

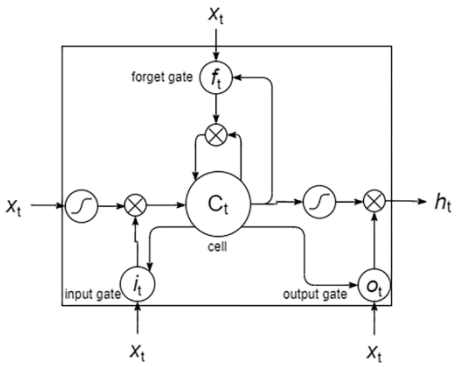



Fig. 3. A Long Short-Term Memory Cell.

LSTM networks are similar to RNNs except for the updates of the hidden layer are replaced by purpose-built memory cells. Therefore, LSTM provides better results for a long range of data with dependencies. The structure of the LSTM cell [23] is represented in Fig. 3 and it can be implemented as follows:

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \quad (1)$$

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f) \quad (2)$$

$$c_t = f_t c_{t-1} + i_t tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad (3)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_{t-1} + b_o) \quad (4)$$

$$h_t = o_t tanh(c_t) \quad (5)$$

In Equation (1), (2), (3), (4), (5), where □ is the logistic sigmoid function, and $i$, $f$, $o$, $c$, are input gate, forget gate, output gate and cell vectors respectively. The cell vectors and hidden vector $h$, all are in the same size. $W$ is the weight matrix, and $b$ is the bias vector. The activation value $h_t$ of the hidden unit at time step $t$ can be calculated using the information at various times [23]. It is possible because the gating mechanism works through storing the historical memory.

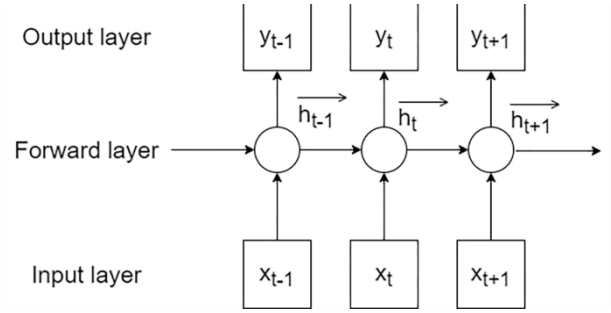



Fig. 4. A Long Short-Term Memory Model.

*2) Bi-directional LSTM (BI-LSTM):* BI-LSTM inputs run in two ways, that are in forward and backward directions. Therefore, forward states and backward states for a specific time frame can be utilized efficiently.

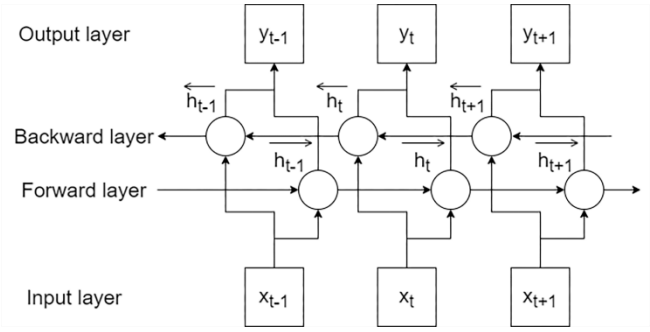



Fig. 5. A Bi-directional Long Short-Term Memory Model.

BI-LSTM connects the two hidden layers to the same output layer which is shown in Fig. 5 [24].

### *E. Model Trainig*

For the model training, preprocessed dataset was used. The dataset includes 3050 malware API sequences and 1050 benign malware API sequences. The dataset was split to approximately 6:1 ratio for training and testing respectively.

Table 2. Dataset information

| | **Training** | **Testing** | **Total** |
|---|---|---|---|
| Data (malware & benign) | 3500 | 600 | 4100 |

Both LSTM and BI-LSTM models were trained using the dataset. In Table 3 it shows the parameters that were used in model training.

Table 3. Parameter table

| **Training Parameters** | **LSTM** | **BI-LSTM** |
|---|---|---|
| Embedding | Word2vec | Word2vec |
| Batch size | 64 | 256 |
| Steps per Epoch | 250 | 300 |
| Optimizer | Adam | Adam |
| Activation Function | Sigmoid | Sigmoid |
| Loss Function | Binary cross entropy | Binary cross entropy |

## IV. RESULTS AND DISCUSSION

This section includes the findings of the research and performance evaluation. The sequence labeled dataset was trained using LSTM and BI-LSTM algorithms to select the suitable model for the fileless malware detection.

Table 4. Confusion Matrix

| **Reality / Predicted** | **Malware** | **Benign** |
|---|---|---|
| **Malware** | TP | FP |
| **Benign** | FN | TN |

Based on the confusion matrix for binary classification that is shown in Table 4, *Accuracy* of each model can be obtained (6).

$$Accuracy = TP + TN / (TP + FP + FN + TN) \quad (6)$$

Tensorflow and Keras libraries have been used to construct, train, and evaluate the fileless malware and benign API sequence classification.

### *A. LSTM Performance*

For LSTM the highest accuracy received is 0.88 when using 3 LSTM layers. Fig. 6 shows the variation of training accuracy and average validation accuracy of fileless malware API sequence classification.

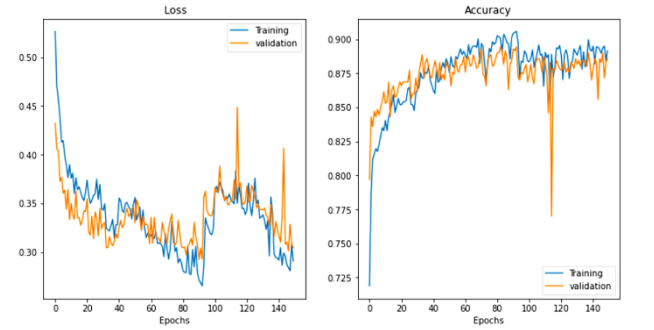



Fig. 6. Training and validation accuracy, average training and validation loss comparison of LSTM classification.

### *B. BI-LSTM*

For BI-LSTM highest accuracy received is 0.92 when using 3 layered BI-LSTM model. Fig. 7 shows the variation of training accuracy and average validation accuracy of fileless malware API sequence classification using BI-LSTM.

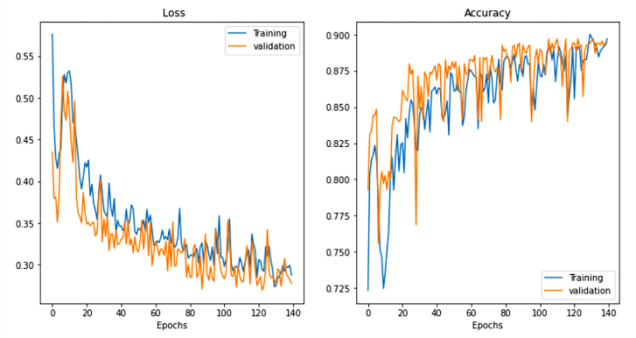



Fig. 7. Training and validation accuracy, average training and validation loss comparison of BI-LSTM classification.

From the above comparison it is clearly visible that BI-LSTM shows a higher accuracy in detecting fileless malware API sequences rather than LSTM model.

Proposed fileless malware detection mechanism is compared with a similar study that uses fileless malware API sequences in a dynamic signature model [10]. As a major drawback, it is stated that there is a higher false positive rate. Moreover, the database should be updated regularly with signatures of API call sequences which is difficult because of the current rapid evolvement of fileless malware. Therefore, our model is more applicable for detecting fileless malware using the API sequences.

## V. CONCLUSION

In this paper, we presented a model for detecting fileless malware that maintains persistence in the Windows environment. For this purpose, initially, we identified the techniques and mechanisms used by intruders to initiate fileless malware using the existing literature and mapping the findings with the MITRE ATT&CK framework. According to the findings we identified that malware API sequence would be suitable for detecting fileless malware that maintains persistence. After extracting fileless malware API sequence data from sandbox reports, we trained two models which are LSTM and BI-LSTM to determine the most suitable model. Based on the highest accuracy we concluded that the BI-LSTM model performs efficiently in detecting fileless malware API sequences.

As future work, improvements will be made on the model by combining with other variants of deep learning classification methods.

## REFERENCES

[1] S. Morgan, “Cybercrime To Cost The World $10.5 Trillion Annually By 2025,” 2020. https://cybersecurityventures.com/cybercrime-damages-6-trillion-by-2021/ (accessed Aug. 20, 2021).

[2] L. Caviglione *et al.*, “Tight Arms Race: Overview of Current Malware Threats and Trends in Their Detection,” *IEEE Access*, vol. 9, pp. 5371–5396, 2021.

[3] A. Alshamrani, S. Myneni, A. Chowdhary, and D. Huang, “A Survey on Advanced Persistent Threats: Techniques, Solutions, Challenges, and Research Opportunities,” *IEEE Commun. Surv. Tutorials*, vol. 21, no. 2, pp. 1851–1877, 2019.

[4] “Fileless threats.” https://docs.microsoft.com/en-us/windows/security/threat-protection/intelligence/fileless-threats (accessed May 06, 2021).

[5] B. N. Sanjay, D. C. Rakshith, and V. H. Vinay, “An Approach to Detect Fileless Malware and Defend its Evasive mechanisms,” *2018 3rd Int. Conf. Comput. Syst. Inf. Technol. Sustain. Solut.*, pp. 234–239, 2018.

[6] V. Khushali, “A Review on Fileless Malware Analysis Techniques,” vol. 9, no. 05, pp. 46–49, 2020.

[7] A. G. Bucevschi, G. Balan, and D. B. Prelipcean, “Preventing File-Less Attacks with Machine Learning Techniques,” *Proc. - 21st Int. Symp. Symb. Numer. Algorithms Sci. Comput. SYNASC 2019*, pp. 248–252, 2019.

[8] “New Research: Fileless Malware Attacks Surge by 900% and Cryptominers Make a Comeback, While Ransomware Attacks Decline,” 2021. https://www.watchguard.com/wgrd-about/press-releases/new-research-fileless-malware-attacks-surge-900-and-cryptominers-make (accessed Jun. 25, 2021).

[9] C. Osborne, “This is how attackers bypass Microsoft’s AMSI anti-malware scanning protection,” 2021. https://www.zdnet.com/article/this-is-how-attackers-bypass-icrosoft-antimalware-scan-software-amsi/ (accessed May 20, 2021).

[10] R. Tarek, S. Chaimae, and C. Habiba, “Runtime API Signature for Fileless Malware,” *Adv. Inf. Commun. Proc. 2020 Futur. Inf. Commun. Conf.*, vol. 1, no. AISC 1129, pp. 645–654, 2020.

[11] N. R. Mistry and M. S. Dahiya, “Signature based volatile memory forensics: a detection based approach for analyzing sophisticated cyber attacks,” *Int. J. Inf. Technol.*, vol. 11, no. 3, pp. 583–589, 2019.

[12] M. Mimura and Y. Tajiri, “Static detection of malicious PowerShell based on word embeddings,” *Internet of Things*, vol. 15, p. 100404, 2021.

[13] X. Huang, L. Ma, W. Yang, and Y. Zhong, “A Method for Windows Malware Detection Based on Deep Learning,” *J. Signal Process. Syst.*, vol. 93, no. 2–3, pp. 265–273, 2021.

[14] A. Afreen, M. Aslam, and S. Ahmed, “Analysis of Fileless Malware and its Evasive Behavior,” *1st Annu. Int. Conf. Cyber Warf. Secur. ICCWS 2020 - Proc.*, 2020.

[15] S. Quintero-Bonilla and A. M. del Rey, “A new proposal on the advanced persistent threat: A survey,” *Appl. Sci.*, vol. 10, no. 11, 2020.

[16] Heena, “Advances In Malware Detection- An Overview,” 2021, [Online]. Available: http://arxiv.org/abs/2104.01835.

[17] “Cuckoo.” https://cuckoosandbox.org/ (accessed Aug. 20, 2021).

[18] B. Alsulami and S. Mancoridis, “Behavioral Malware Classification using Convolutional Recurrent Neural Networks,” *MALWARE 2018 - Proc. 2018 13th Int. Conf. Malicious Unwanted Softw.*, pp. 103–111, 2019.

[19] L. O’murchu and F. P. Gutierrez, “SECURITY RESPONSE The evolution of the fileless click-fraud malware Poweliks,” 2015. [Online]. Available: https://autoblog.postblue.info/autoblogs/lamaredugoffrblog_a1de86d064e376dc283723997fd86bde6ba2d492/media/456c7abf.evolution-of-poweliks.pdf.

[20] “Portable Freeware.” https://www.portablefreeware.com/ (accessed Oct. 10, 2021).

[21] D. Javaheri, P. Lalbakhsh, and M. Hosseinzadeh, “A Novel Method for Detecting Future Generations of Targeted and Metamorphic Malware Based on Genetic Algorithm,” *IEEE Access*, vol. 9, pp. 69951–69970, 2021.

[22] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” *Neural Comput.*, vol. 9, pp. 1735–1780, 1997.

[23] Y.-W. Lu, C.-Y. Hsu, and K.-C. Huang, “An Autoencoder Gated Recurrent Unit for Remaining Useful Life Prediction,” *Processes*, vol. 8, no. 1155, 2020.

[24] Z. Huang, W. Xu, and K. Yu, “Bidirectional LSTM-CRF Models for Sequence Tagging,” 2015, [Online]. Available: http://arxiv.org/abs/1508.01991.

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index N0: CT-10-57*

# A Consortium Blockchain Model to Overcome Issues in the Global Patent Authentication and Management Process

Prageeth Thilina Gunasekara (1st Author)
*Department of Industrial Management*
*University of Kelaniya*
Dalugama, Sri Lanka
prageeththilina8@gmail.com

Chathura Rajapakse (2nd Author)
*Department of Industrial Management*
*University of Kelaniya*
Dalugama, Sri Lanka
chathura@kln.ac.lk

***Abstract*—Patent authentication and management is an important activity to protect the intellectual property rights of individuals and organizations globally. As the patents are territorial and the data management related to patent applications mostly happens with local intellectual property (IP) offices, the patent authentication and management (PAM) process is significantly inefficient and time-consuming on the global scale. Moreover, the data management issues in the patent process often lead to conflicts and legal actions among competing parties. Blockchain technology is widely recognized for its potential use in creating decentralized, secure, and transparent systems with immutable records. It hence seems to be a useful technology to overcome the issues in the patent domain. This paper explores the adaptability of blockchain technology in the patent domain. It presents the design of a consortium blockchain system, which is proposed as a solution to numerous issues stemming from inefficient data management. The proposed consortium blockchain design is based on the Ethereum architecture and is equipped with smart contracts to ensure the reliability of patent data as well as the real-time update of records. This paper further discusses the potential testing and validation strategies for the proposed model.**

***Keywords — blockchain, patent authentication and management***

## I. INTRODUCTION

A patent is an exclusive right granted to an inventor that precludes others from commercializing, using, distributing, importing, or selling the invention without permission from the patent holder. According to the World Intellectual Property Statistics, 3 million patent applications were filed in 2019. As a result, an increasing number of people are eager to file patent applications for their inventions. A patent, on the other hand, is territorial, which means that its rights are restricted to a single country. This restriction is incompatible with the fast-paced corporate culture of today's economy. Even with the support of international accords like the Patent Cooperation Treaty, registering a patent in another country requires multiple patent filings in each target patent office, which is inefficient and costly (PCT). Figure 1 depicts the current patent authentication and administration system. It takes 3–6 months to complete a patent application, and 3–5 years to examine patentability. This is a long period during which new ideas can be introduced to the market, lowering the likelihood of this invention being commercialized in the future.

At the national level, intellectual property rights protection is seen as a major issue. This has an impact on the intensity of innovations and is critical in enhancing the drive of entrepreneurs to innovate. To safeguard these rights, the existing intellectual property rights system needs a tremendous amount of money and work. The existing PAM system has issues such as the length of time required for patenting, the absence of evidence of ownership before requesting legal action, low-quality patent approval, high search expenses, a high risk of expropriation, limited access for small inventors, and the complexity of undertaking a patent landscape of a given technology.[1], [2]

The objective of this study is to see if blockchain technology can be used to solve problems in the global patent authentication and management process. Through smart contracts in safe document management, blockchain can significantly influence immutability, trustworthiness, security, and federated features, and so will have a good impact on patent authentication and management. We plan to incorporate all of these capabilities into the PAM by utilizing the benefits of blockchain technology.

### *A. Blockchain Technology*

Blockchain technology is a decentralized, peer-to-peer network that creates a transparent, encrypted, and immutable digital federated ledger system. The users themselves maintain this shared ledger, which does not require the involvement of a third party. Because data entered into the blockchain cannot be changed without being identified, the system is impenetrable. It has a wide range of applications because it can store any type of data, including cryptocurrency, transactional data, contractual data, design data, etc. Blockchain technologies are interesting in today's era, which is prone to cyber-attacks on centralized systems, because of their immutability, traceability, and lack of third-party invention.

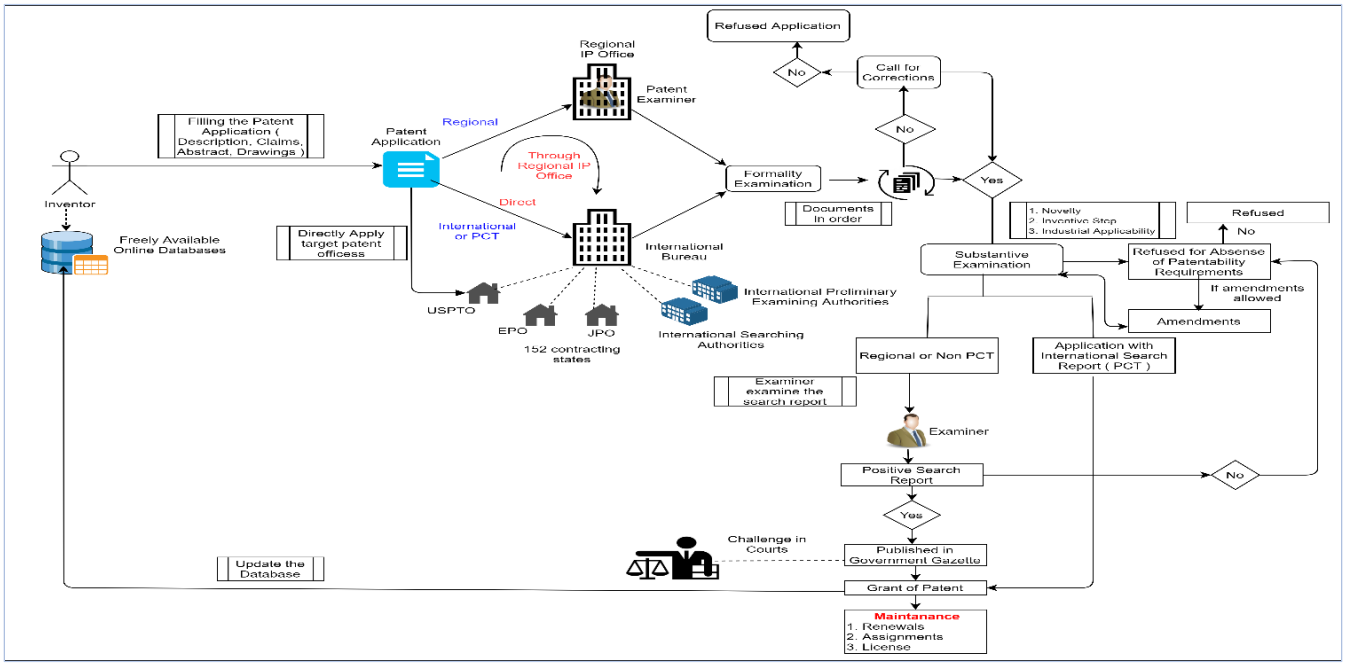



Fig. 1 The existing patent authentication and management system

Information is stored in discrete encrypted blocks that are linked by hash references in blockchain technology. These hash points build a chain or blockchain by connecting one block to the next. Individual transactional data is included in these blocks, which are time-stamped separately. A consensus mechanism ensures that all nodes agree on a unique sequence for adding entries to the blockchain and that all nodes verify the data that will be appended to the blockchain. Self-autonomy, trust, transparency and provenance, immutability, disintermediation, and collaboration are some of the benefits of blockchain technology. [3] Blockchain technology can also be used to create a distributed database, which will be extremely helpful in the preservation of public data. It can also be utilized in government processes to optimize and cross-validate systems. This technology appears to be best suited for key government processes like digital identity management, secure record-keeping, and document management.

Public permissionless blockchain, public permission blockchain, consortium blockchain, and private blockchain are the four types of blockchain. [4],[5]

**Public permissionless blockchain**: Anyone can connect to the network and write and access data in this sort of blockchain (e.g., Bitcoin). In this type of blockchain, there is no single owner of the networks, ensuring anonymity. Although this sort of blockchain is thought to be extremely secure, validating transactions take a long time because the amount of power required to verify each subsequent transaction grows. In this situation, proof of work has a greater influence than in other types of blockchain, where more nodes participate in transaction verification.

**Public permission blockchain**: In this sort of blockchain, everyone can connect to the network, but only a select group of people can write to it.

**Consortium blockchain**: This is a hybrid blockchain that combines the features of both public and private blockchains. Only a tiny set of authorized members in this network have access to the information. Therefore, only authorized users are allowed to make changes to the network.

**Private blockchain**: The data in this blockchain network can only be viewed by a select set of authorized nodes. Only the network operator can write to the network.

*B. Potential to use Blockchain in Patent authentication and Management*

Along with the patent authentication and management process, the immutability, trustworthiness, reliability, and federated features of Blockchain can be gained at any level. At the moment, each intellectual property office's patent records are housed in databases. Such information may go out of sync. To combat this, data must be validated and updated frequently. The advancement of blockchain technology has the potential to greatly assist IP offices all over the world. The integration of these technologies into patent databases helped them. The Blockchain has a lot of potentials when it comes to establishing proof of original inventorship. It will aid in the reduction of lawsuits while also identifying the proper inventor/creator/proprietor. The filling data kept on blockchain will have sufficient explicit significance in deciding the rights of the first filler under a

"Priority Date" notion. Blockchain technology can be used to synchronize internal and external search databases. Patent investigators will be able to find anticipation in inventions by using a single unified platform for patent search.

Patent data is stored either on paper or electronically by intellectual property offices. In a court of law, those registries have legal validity in certifying the holders' rights. Blockchain technology can be used to keep records that ensure data accuracy. Furthermore, these registers enable real-time record updating in the event of a right transfer. Because blockchain is a linked list that employs hash pointers instead of normal pointers, data verification would be much easier and tamper-proof. Each blockchain node is given the task of locating the next node and determining if it has changed. Blockchain technology has a lot of potential for interacting with IP offices all over the world. Faster research, more reliable patent data management, advanced licensing, and increased credibility will all benefit from incorporating this technology into IP offices' day-to-day operations as soon as possible.

### *C. Blockchain Usability*

Blockchain technology got its start in 2008 with the launch of Bitcoin. In a white paper, Satoshi Nakamoto titled Bitcoin: A Peer-to-Peer Electronic Cash System. [6] [3] gives an overview of blockchain technology, including its present applications and ramifications. According to the report, no third-party mediators control data on a blockchain network, and all transactions must go through a cryptographic validation mechanism. In a decentralized and distributed network, blockchain is a chain of blocks that holds information with digital signatures.

Decentralization, immutability, and transparency are all characteristics of blockchain technology that make systems safer and more resistant to tampering. Blockchain technology has applications in financial and social services, risk management, healthcare facilities, and other industries, in addition to cryptocurrency. [7] At the moment, it looks that blockchain technology is best suited for digital identity management, secure records management, and document processing, all of which are vital governmental functions. A blockchain is a secure, verifiable record of all monetary and governmental transactions. Furthermore, blockchain technology ensures that document management system administration is secure. [8]

### *D. Consortium Blockchain*

A consortium blockchain is a type of blockchain that enlists the help of authorized nodes to keep the distributed ledger up to date. In this type of blockchain deployment, only approved nodes are included. Consortium blockchains have emerged as a significant concept and architecture for integrating the freedom and anonymity of private blockchain transactions with public blockchain governance. [9] In research, the topic of consortium blockchain has gotten very little attention. This method can be used in PAM systems to improve infrastructure while also enhancing service quality.

Consortium blockchain will ensure that:

- Transaction limitations will be decreased as a result of transaction confirmation delays
- No transactions fee
- Blockchain data privacy will be improved.
- The special pre-approved set of nodes does the verification of the operations.

### *E. Review of Smart Contract in the context of blockchain*

According to [10], a “smart contract is a type of computer program which runs on the blockchain and is executed by all consensus nodes. It is made up of program code and a storage file.” Any user can create a contract by uploading a transaction to the blockchain. A contract's program code is fixed when the contract is created and cannot be modified.

The combination of blockchain technology and smart contracts allows for more flexibility in inventing, designing and addressing real-world problems at a cheaper cost and faster than traditional third-party-based systems. The importance of blockchain technology's smart contract integration has been a focal area for development since it allows for peer-to-peer transactions and databases to be maintained open in a secure and trustworthy environment. Smart contracts can be traced and are not subject to alteration. A smart contract is a computer program that stores all transaction data and executes it automatically. [11]

### *F. Blockchain consensus mechanism and Practical Byzantine Fault Tolerance (PBFT)*

Consensus algorithms are critical in assuring the security and efficiency of the blockchain network. A blockchain application's performance can be greatly improved with the right algorithm implementation. [12]. In a distributed, untrusted environment, this strategy allowed all nodes on the public ledger to reach a consensus. The blockchain system's security, availability, and performance are all decided by the consensus process.

Practical Byzantine Fault Tolerance (PBFT) is presently being used as a consensus approach in many blockchain projects since it has various advantages, including high transaction throughput and reduced energy loss. [13] The PBFT consensus algorithm is the first high-performance consensus protocol with optimum byzantine fault tolerance. The PBFT protocol is a Byzantine Fault Tolerance mechanism with a simple algorithm and a large number of applications in distributed systems. PBFT ensures that during each round of consensus, nodes maintain a shared state and take consistent action. Because

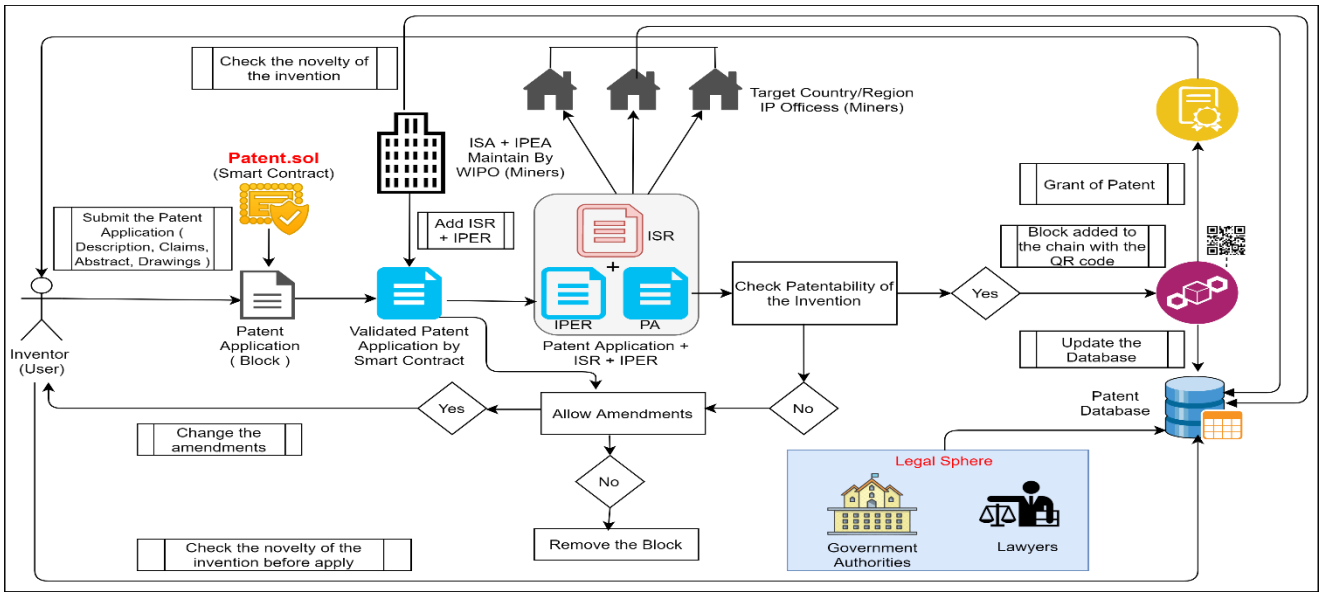



Fig. 2 Blockchain based patent authentication and management ecosystem

it meets the goal of great consistency, PBFT is referred to as an absolute-finality consensus protocol. [14]

PBFT is a state machine replication-based consensus mechanism. Services are reproduced in numerous nodes of a distributed network as a state machine. The state of the service, as well as the activities it performs, are saved in each replica of the state machine. This method can ensure that the system runs normally when the number of defective nodes does not exceed one-third of the total number of nodes. A message inquiring about the content of messages received by other nodes should be sent to each node. There is one primary node out of n nodes in the PBFT method, and the other backup nodes are referred to as replicas. The view-change component of the PBFT algorithm is also crucial. When the primary node goes down and is unable to finish the data request within a certain amount of time, the other replicas initiate a view-change, and the replacement primary node takes over once the conversion is performed. By broadcasting to the entire network in each round and enabling each node to vote for the primary node, the PBFT approach can handle both non-Byzantine and Byzantine problems at the same time. This complex technique assures that PBFT maintains its consistency, availability, and anti-fraud attack features.

*G. Related Works*

Wei-Tek Tsai et al. (China, 2017) presented blockchain integrated microfilm IP protection framework that focused primarily on scripts and names, which were used to distinguish one microfilm from another and could be kept in a blockchain database. [15] Savelyev and Alexander Ivanovitch noted significant legal concerns that arise when misusing blockchain technologies in the copyright elements in 2017. They analyzed existing issues with digital availability of copyrighted works, solutions available in blockchain technology, and related issues that must be taken into account in this topic. [16]

In 2018, Kensuke and Marcus presented a copyright protection mechanism based on blockchain technology. Martin Holland et al. [17] discussed the use of blockchain technology. In addition, digital rights management was highlighted as a key technology bottleneck in the transition to additive manufacturing processes. It was also a step toward defining and preventing IP theft in the 3D printing supply chain. They proposed the Secure Additive Manufacturing Platform (SAMPL), which created a secure and trusted foundation for additive manufacturing technologies. [18] Fran Casino et al. stressed the inclusion of information and transactions in blockchain integrity verification applications that are associated with the development and longevity of products and services when reviewing the systematic literature of blockchain-based applications. Provenance, counterfeiting, and IP management are just a few of the major uses. [19]

Gonenc Gurkaynak et al. (2018) used a blockchain-based architecture to improve the functioning of IP offices, improve customs processes for detecting counterfeit products, and increase the efficiency of the right holder's IP rights management. They took many steps to help the growth of blockchain technology, including promoting it among the general public and successfully integrating it into a variety of services and registration/transaction channels in IP administration. [20] Alexander Schönhals and colleagues have emphasized blockchain-based strategy to safeguard innovative ideas and the first product design and concept development. [21] Martin Zeilinger explored the use of blockchain technologies to improve the exclusive digital art market in 2018 [22].

The majority of current research, on the other hand, focuses on strategies to protect copyrights and trademarks. Because the applicant is the only user type in their suggested solutions, applying these tactics to patent protection is difficult. Examiners are required in patent applications, unlike copyright and trademarks, because patents are often technical. For patent authentication and management, we need to design a more appropriate blockchain-based system paradigm.

## II. METHODOLOGY

This research is being done to propose and examine a new consortium blockchain model to address issues with worldwide patent authentication and management. Expert comments were obtained by conducting structured interviews with specialists in the intellectual property office to identify the process of the existing system and its limitations. A thorough literature review was also conducted to determine the state of knowledge in the associated domains. The identified challenges, constraints, and flaws are then solved through the creation of unique architecture. The potential testing and validation procedures for the proposed model are also discussed in this study.

### *A. Blockchain-based patent authentication and management system*

As previously noted, there are significant technological, legal, and economic difficulties with the current PAM system. All of these problems led to major defects that impacted the business environment and the innovation cycle. The existing PAM system, it is said, lacks the efficiency and flexibility needed to respond to market changes. Incorporating blockchain technology into the PAM system would address all of these concerns while also enhancing the efficiency of innovation processes in the country's economy, for all of the reasons stated above. Everyone who needs to patent their inventions will be encouraged by the recommended process.

Figure 2 depicts the proposed blockchain-based patent authentication and management ecosystem. All stakeholders are included, including inventors, the World Intellectual Property Organization (WIPO), the International Search Authority (ISA), the International Preliminary Examination Authority (IPEA), global IP offices (USPTO, EPO, JPO), lawyers, government officials, and any potential relationships between participants in this platform and the system's workflow.

Inventors may conduct a preliminary search for already claimed patents and the patentability of their concept before filing for a patent. A single unified platform for patent search is given since international IP offices maintain (examine and update data regularly) the same database. An inventor who wants to file for a patent submits an e-request with the invention's information. Smart contracts look for

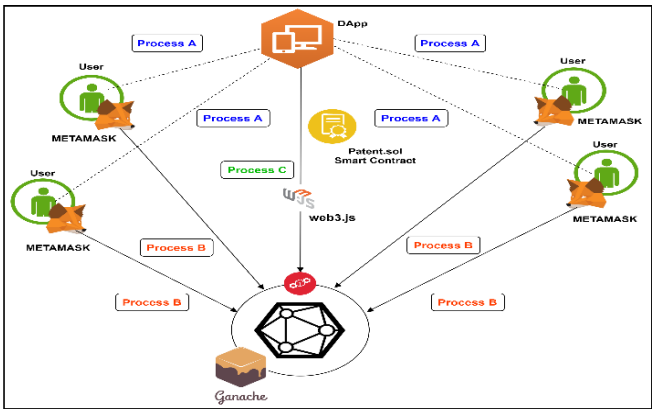



Fig. 3 Overview of the Decentralized App

already claimed patents, which will be incredibly useful in compiling the ISA report, according to the invention's keywords. The patent smart contract stores the validation e- request on the blockchain when it has been validated. The International Search Report (ISR) and International Preliminary Examination Report (IPER) are then added by the ISA and IPEA, respectively. Then, at the same time, all of the target patent offices take such e-requests and validate the details by approving or rejecting them. The patent is given once they are authorized, and it is transmitted to the smart contract holding the patent data, as well as the owners' information. Before storing the information in the database, the smart contract verifies that the user who initiated the request is the patent owner. It also verifies the legitimacy of the IP offices that verify patent data.

The owner creates an e-request in the event of a change of ownership, which is subsequently stored on the blockchain in the patent smart contract. The patent examiner then authorizes the e-request, and patent information is sent to the smart contract, which holds patent ownership information, and the owners' information is updated. The smart contract confirms that the individual who generated the request is the actual owner of the patent and that the patent examiner who approved the transaction is legitimate before updating the owner's data.

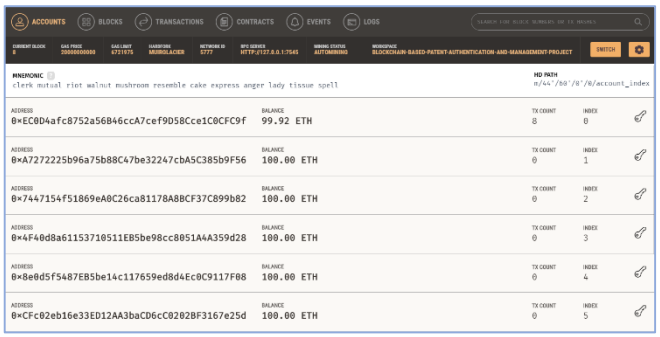



Fig. 4 Ganache Test Client with generated accounts

*B. Testing and Validation Strategies*

A prototype of the Ethereum blockchain-based decentralized application (DApp) shown in Figure 3 is being developed to demonstrate the architecture and overcome the highlighted hurdles and constraints. The difference between this model and the Public Ethereum blockchain is that under the Bitcoin design, Proof of Work is carried out by network miners who are compensated in 'ether' for their efforts. In this case, the PBFT consensus is given to a collection of authorities that includes the World Intellectual Property Organization (WIPO), IP offices, and others. Consortium Blockchains are the name for these blockchains. As a result, when a new transaction occurs and a block is created, the authorities must validate the transaction before the block can be added to the blockchain. The patent details may be inserted as data into the block by the patent.sol smart contract. After the block has been verified, it is uploaded to the blockchain.

DApps in Ethereum are programs having a graphical user interface that make it easier for users to interact with the blockchain. Smart contracts handle the application's business logic, while HTML, JavaScript, and other front-end technologies are used. Because contract code is redundant and expensive to run, DApps are meant to do only the essential operations on the blockchain, while anything else that could be done on the client-side is handled by the application's front end.

**Process A -** For the authentication process, the system users' (Inventors/WIPO/IP offices) metamask browsers will connect to the Blockchain.
**Process B -** To complete the PAM transactions, the users will communicate with the DApp.
**Process C -** To write transactions to the blockchain, the DApp will link to the Ganache blockchain.

To get results, it is also tested in a ganache simulated environment with produced actors and data. Ganache is a blockchain simulator that may be used to create a private Ethereum blockchain. It allows you to inspect the state of the blockchain while also controlling the chain's operations. It will make it possible to do all actions on the main Blockchain for free. Figure 4 depicts the wallets that were created, together with their wallet addresses. Case studies such as accepting or rejecting patents, adding remarks to the patent application by target patent offices, verifying patent ownership, transferring patent ownership, and conducting a patent landscape are used to assess the platform's capabilities. Mock data is used to simulate the input and output data that is required during testing.

Ganache is a private blockchain for testing purposes. On a public Ethereum network like Rinkeby, you can try this. The Proof-of-Authority consensus mechanism is provided by the Rinkeby network, whereas the Proof-of-Work consensus mechanism is used by the Ethereum Ropsten network.

## III. RESULTS AND DISCUSSIONS

A blockchain-based PAM system provides live data on patents with easy searchability because the blockchain network contains all data related to patents registered on the system. This also assists in determining who owns a patent. By offering a comprehensive overview of certain technology in the framework, it prevents other patents from being replicated. As a result, our technology will search for similar patents in a speedy, methodical, and straightforward manner. Because every transaction is permanently recorded and cannot be edited or altered in this system, the blockchain-based PAM architecture would make the transition of patent ownership easier. Examiners, inventors, and lawyers would benefit from the blockchain-based PAM system because it would save costs. In legal procedures, lawyers might use the system's unchanging data to prove ownership and the presence of eligible patents. As a result, this system unifies several PAM-related services into a single interface.

The novel technology for manipulating the PAM system must be implemented at the national level, with a system peer node included in each IP office in the country region reaching agreements in all patent applications. The management of this framework, including the WIPO, should be opposed by the government. As a result, we're pushing for the creation of a consortium blockchain platform that oversees the consensus process and includes regulated bodies like the WIPO and regional IP offices. If this does not happen, the private sector will gain control of many platforms, making dispute resolution and legal empowerment much more difficult. The system's viability, dependability, and efficiency are all enhanced by this platform. As a result, there will be no need for a third party to oversee patent-related transactions because the platform's scalability has been assured. To develop a regionally coordinated national database of patents that may be used as evidence in national courts, the system should be linked to all national IP offices throughout the world.

## IV. CONCLUSIONS

This study looked into the problems with the current PAM system, which are mostly connected to time, cost, security, and dependability. Combining blockchain technology and the PAM system, in response to the aforementioned problems, would solve all of the system's flaws and boost performance capabilities, minimizing bottlenecks in the present system. To address these issues, a blockchain-based PAM platform was proposed, which would bring together all PAM system stakeholders, including inventors, WIPO, regional IP offices, ISA, and IPEA, to ensure that all patent information is shared among all IP offices while maintaining the security provided by the blockchain.

It is capable of meeting the requirements of IP offices all around the world. It will be much easier for IP offices to

accomplish results in terms of faster inspection, reliable record management, and smart licensing if they can quickly integrate this technology into their daily operations. Patent prosecution, protection of patent rights, and the resolution of patent infringement issues will all benefit from this technology. The importance of blockchain as a widely used database was also investigated in this study (which is critical in cross-validated reliable systems like digital identity management, secure record-keeping, and document processing in government operations).

## REFERENCES

[1] Alnafrah Ibrahim, and Bogdanova Elena L., "A new holistic approach for studying blockchain-based intellectual property rights ecosystem," vol. 19, no. 1 pp. 133-140. 2019.

[2] Lin, J., Long, W., Zhang, A. and Chai, Y., "Blockchain and IoT-based architecture design for intellectual property protection", International Journal of Crowd Science, Vol. 4 No. 3, pp. 283-293. 2020.

[3] Holotescu, C.understanding blockchain technology and how to get involved. In: The 14th International Scientific Conference eLearning and Software for Education, April 2018.

[4] Hileman, G., and Rauchs, M. *Global Blockchain Benchmarking Study*, Cambridge Centre for Alternative Finance, pp.1–122, 2017.

[5] Mougayar, W. The Business Blockchain: Promise Practice, and Application of the NextInternet Technology, John Wiley & Sons, New Jersey. 2016.

[6] Nakamoto, S. Bitcoin: A Peer-to-Peer Electronic Cash System, 2008.

[7] A. A. Monrat, O. Schelén, and K. Andersson, "A Survey of Blockchain From the Perspectives of Applications, Challenges, and Opportunities," in IEEE Access, vol. 7, pp. 117134-117151, 2019.

[8] Jansen A., Olnes S. Blockchain Technology as s Support Infrastructure in e-Government. In: Janssen M. et al. (eds) Electronic Government. EGOV 2017. Lecture Notes in Computer Science, vol 10428. Springer, Cham. 2017.

[9] Dib, O., Brousmiche, K. L., Durand, A., Thea, E., & Hamida, E. B. Consortium blockchains: Overview, applications, and challenges. International Journal On Advances in Telecommunications, 11(1&2), 51-64, 2018.

[10] Delmolino, K., Arnett, M., Kosba, A., Miller, A., & Shi, E., Step by step towards creating a safe smart contract: lesson and insights from a cryptocurrency lab. International Conference on Financial Cryptography and Data Security, pp. 79-94, 2016.

[11] B. K. Mohanta, S. S. Panda and D. Jena, "An Overview of Smart Contract and Use Cases in Blockchain Technology,", 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), pp. 1-4, 2018.

[12] D. Mingxiao, M. Xiaofeng, Z. Zhe, W. Xiangwei and C. Qijun, "A review on consensus algorithm of blockchain," 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 2567-2572, 2017

[13] M. Kim, Y. Kwon, and Y. Kim, "Is Stellar As Secure As You Think?" IEEE European Symposium on Security and Privacy Workshops (EuroS&PW), pp. 377-385, 2019.

[14] Shijie Zhang, Jong-Hyouk Lee, Analysis of the main consensus protocols of blockchain, ICT Express, Volume 6, Issue 2, ISSN 2405-9595, Pages 93-97, 2020

[15] Tsai, W.-T., Feng, L., Zhang, H., You, Y., Wang, L., and Y. and Zhong, "Intellectual-property blockchain-based protection model for microfilms", Proceedings of 2017 IEEE Symposium on Service-Oriented System Engineering (SOSE), IEEE, San Francisco, CA, April 6-9 2017.

[16] Savelyev, A., and Ivanovitch, "Copyright in the blockchain era: promises and challenges (November 21, 2017)", Higher School of Economics Research Paper. No. WP BRP 77/LAW/2017, available at SSRN. 2017.

[17] Ito, K., & O'Dair, M. A Critical Examination of the Application of Blockchain Technology to Intellectual Property Management. Business Transformation through Blockchain, 317–335. 2018.

[18] Holland, M., Stjepandi_c, J., and C. and Nigischer, "Intellectual property protection of 3D print supply chain with blockchain technology", Proceedings of 2018 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC), IEEE, Stuttgart. 2018.

[19] Casino, F., Dasaklis, T.K. and Patsakis, C., "A systematic literature review of blockchain-based applications: Current status, classification, and open issues", Telematics and Informatics, Vol. 36, pp. 55-81. 2019.

[20] Gurkaynak, G., Yılmaz, _I., Yes_ilaltay, B. and Bengi, B. (), "Intellectual property law and practice in the blockchain realm", Computer Law and Security Review, Vol. 34 No. 4, pp. 847-862, 2018.

[21] Schönhals, A., Hepp, T., and B. and Gipp, "Design thinking using the blockchain: enable traceability of intellectual property in problem-solving processes for open innovation", Proceedings of the 1st Workshop on Cryptocurrencies and Blockchains for Distributed Systems (CryBlock'18). ACM, New York, NY, pp. 105-110, 2018

[22] Zeilinger, M., "Digital art as `monetized graphics': enforcing intellectual property on the blockchain", Philosophy and Technology, Vol. 31 No. 1, pp. 15-41, 2018.

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index N0: CT-11-74*

# Audio Steganography using LSB Technique to Embedding Data

Dayanantha Shanmugaradnam ( 1st Author)
*Department of Computer Systems Engineering*
*University of Kelaniya.*
Kelaniya, Sri Lanka
dayanant_ct15009@stu.kln.ac.lk

Dr. Hesiri Dhammika Weerasinghe (2nd Author)
*Department of Computer Systems Engineering*
*University of Kelaniya*
Kelaniya, Sri Lanka
hesiri@kln.ac.lk

***Abstract* - Data protection is a major concern on the internet medium. Data needs to be protected from intrusion, penetration, and data theft. Audio steganography is used to hide secret messages in an audio file. This method was intended to secure the secret message. The secret message was protected using the hashing and encryption technique and the Least Significant Bit substitution was performed to hide the data. The Stego audio files were analyzed by Signal-to-Noise Ratio and Mean Opinion Score. The Stego audio files had an 80dB average Signal-to-Noise Ratio and the overall Mean Opinion Score was 4.4 out of 5. It proves that this method helped to improve the robustness and imperceptibility. Using the proposed method higher security can be achieved.**

***Keywords — data hiding, audio steganography, lsb***

## I. INTRODUCTION

Data can be hidden using various methods. Steganography is one of them. In other data hiding methods, the existence of the data can be easily discovered and revealed [1]. Steganography is the art of concealing data. In steganography, only the sender and the receiver know the existence of the hidden data. As a result, steganography has emerged as the most reliable protection method [2].

The basic scheme behind the audio steganography process is to conceal a message secretly on an audio signal called cover audio [3]. The resultant audio after this process is called “stego” audio, which is sent via a protected channel to the recipient. At the receiver’s end, the hidden data will be extracted from the audio by applying a series of pre-defined steps [4].

Audio data concealing is very difficult than other stenographic methods. Because it relies on the human ear's dissimilar sensitiveness to a higher and lower intensity of sounds [5]. In audio steganography, several methods and several techniques were used. The Least Significant Bit coding method is coming under the Spatial Domain Technique [6].

The least significant Bit based audio steganography methods have made a significant contribution. However, because of their lack of robustness, the security of the Least Significant Bit based methods could be quickly violated to obtain the hidden data [7].

Therefore, the hidden or the embedded data needs to be secured. Using Cryptography along with Steganography will increase security.

## II. OBJECTIVES

The following objectives were focused on throughout this study.

- A detailed study on Steganography
- Identify the Techniques and Methods
- Study on the Least Significant Bit
- Securing the secret message
- Testing the stego audio

## III. METHODS

### *A. Adapted Technologies*

#### *Cryptography*

Cryptography is the process of altering the structure of a file before it is sent. Through cryptography, the golden triangle of Confidentiality, Integrity, and Authentication can be achieved [8]. The secure data hashing and the symmetric key encryptions are used in this method.

#### *Audio Signal Processing*

Through signal processing, the character of the audio file can be changed. This audio processing can be used for different purposes [9]. Here it is used to enhance the quality of audio, create a new effect, compression and store and transmit data and information.

### *B. Embedding Process*

A key and salt were generated from the password using SHA 256 hashing algorithm. At the same time, the secret message and its length were used to create a formatted message. Then the generated key was encrypted using the fernet symmetric key encryption. Later, a token

was created by encrypting the formatted message using the encrypted key. Finally, the embedding process started in the least significant bit of the cover audio from its offset. The length of the token and salt values was embedded in the first 8 bytes of the cover audio. From the next bytes, the token value string is embedded in the cover audio.

### C. *Extraction Process*

In the stego audio from its offset length of the token and the salt were extracted. As the next step token value was extracted using its length. Then the key was generated from the given password and salt using SHA 256 algorithm. Later, the generated key was encrypted using fernet symmetric key encryption. The token was decrypted using the encrypted key to get the formatted message. From the formatted message length of the message was extracted and the secret message was extracted.

Hashing the password and the salt are increasing the security of the embedded message. From this method, Confidentiality, Integrity, and Authentication are archived.

## IV. RESULT AND DISCUSSION

The proposed method is tested on four Wav files. And three different secret messages are used to embed. These are evaluated using the Signal to Noise Ratio (SNR) and Mean Opinion Score (MOS).

### A. *Signal to Noise Ratio*

The ratio between the cover audio signal and the stego audio signal are used to evaluate the SNR value. The bigger the SNR ratio is better the sound quality.

Table 1. Signal-to-Noise Ratio Values

| **Cover Audio (.wav)** | **Cover audio size (KB)** | **Sampling Frequency (kHz)** | **Time (Sec)** | **SNR (dB)** | | |
|---|---|---|---|---|---|---|
| | | | | **456 bytes** | **796 bytes** | **1480 bytes** |
| 1.Violin | 2585 | 44 | 15 | 84.18 | 79.74 | 68.89 |
| 2.MakeItUp | 2585 | 44 | 15 | 85.84 | 80.56 | 70.12 |
| 3.Takeaway | 3446 | 44 | 20 | 89.65 | 85.49 | 79.38 |
| 4.Believer | 3446 | 44 | 20 | 87.37 | 84.61 | 77.42 |

### B. *Mean Opinion Score (MOS)*

MOS is commonly used to evaluate the quality of sound. Group of individuals is rated the quality of the stego audio file by listening and compare with the cover audio file.

Table 2. Mean Opinion Score Values

| **Stego Audio (.wav)** | **Mean Opinion Score (scale 1 to 5)** | | |
|---|---|---|---|
| | **456 bytes** | **796 bytes** | **1480 bytes** |
| 1.Violin | 4.1 | 3.7 | 3.2 |
| 2.MakeItUp | 4.6 | 4.6 | 4.6 |
| 3.Takeaway | 4.8 | 4.8 | 4.7 |
| 4.Believer | 4.8 | 4.7 | 4.8 |

Experimental study results clearly show the differences between the cover audio file and how the secret message affects the quality of the stego audio file.

The audios 1. Violin and 2. MakeItUp are equal in file size and length. But the SNR values are different with the size of the embedded data. It is because these files have different frequency ranges. The audios 3. Takeaway and 4. Believer are bigger than the previous audio samples and also both are equal in file size and length. These files also have different SNR values. Therefore, it is concluded that the SNR value depends on not only the sampling frequency, length, size, the frequency range of cover audio but also the size and the content of the hidden data. To avoid detection, a user should listen to a stego audio before sending it to a receiver.

The hashing algorithm enhances security. If only the encryption algorithm is used, the encrypted message can be retrieved when the password key is found. But in hashing the hash values has a fixed length and it cannot be reversible. In addition to that, a hash algorithm can be created using the salt and the password. Salt is a randomized number added with the password before the hashing process starts. Even though two people use the same password its hash value will differ because of the salt added. It is the best way to store passwords.

The larger the file size is carrying the larger amount of secret data without any major change. It is proven from the sample audios and the different secret messages.

## V. CONCLUSIONS

Information security is very important since information is exchanged on the publicly accessible Internet. Both cryptography and steganography can be used to provide information security. Steganographic techniques allow hiding valuable information in a normal file. While the cryptography technique is used to create confidentiality, integrity and authentication to the hided information.

This proposed method increases the security of the embedded data using encryptions and hashing techniques. The SNR values show the quality of the stego audio file. And the MOS values ensured the quality of the stego audio files. The quality of the audio does not drastically change when a secret message is embedded into the audio by using the proposed solution.

Even though the steganography breaks and the hidden chunk is discovered, it is still useless because it is encrypted using a password. When these two systems are combined, a high level of protection is achieved.

## REFERENCES

[1] S. P. Rajput, K. P. Adhiya, and G. K. Patnaik, "Hide Text in Audio," *2017 Int. Conf. Comput. Commun. Control Autom.*, pp. 1–6, 2017.

[2] R. Doshi, P. Jain, and L. Gupta, "Steganography and its Applications in Security," *Int. J. Mod. Eng. Res.*, vol. 2, no. 6, pp. 4634–4638, 2012.

[3] A. Binny, "Hiding Secret Information Using LSB Based Audio Steganography," pp. 1–4, 2014, doi: 10.1109/ISCMI.2014.24.

[4] V. Sharmal, "LSB Modification based Audio Steganography using Trusted Third Party Key Indexing Method," pp. 403–406, 2015.

[5] H. Dudhwal, A. Prof, and J. Boaddh, "A Review of an Extensive Survey on Audio Steganography Based on LSB Method," vol. 7, no. 1, pp. 177–180, 2021.

[6] S. Mishra, V. K. Yadav, M. C. Trivedi, and T. Shrimali, "Audio steganography techniques: A survey," *Adv. Intell. Syst. Comput.*, vol. 554, no. February, pp. 581–589, 2018, doi: 10.1007/978-981-10-3773-3_56.

[7] W. A. Shah, D. Shehzad, A. I. Umar, J. Hussain, and A. Qadir, "Audio Steganography Based on Lsb Msb," vol. 15, no. 6, 2017.

[8] S. K. Moudgil, A. K. Goel, and M. Sharma, "Steganography on Audio Wave Tenth Layer by Using Signal to Noise Ratio Test and Spectrogram Analyses," *Int. J. Appl. Eng. Res.*, vol. 13, no. 4, p. 1931, 2018, doi: 10.37622/ijaer/13.4.2018.1931-1935.

[9] H. Dutta, R. K. Das, S. Nandi, and S. R. M. Prasanna, "An Overview of Digital Audio Steganography," *IETE Tech. Rev. (Institution Electron. Telecommun. Eng. India)*, no. December, 2019, doi: 10.1080/02564602.2019.1699454.

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index No: ES-12-20*

# Field Testing of Highway Bridges Enhanced by Assumptions of Composite Action

Renxiang Lu (1st Author)
*Department of Civil and Architectural Engineering*
*University of Wyoming*
Laramie, Wyoming, USA
rlu@uwyo.edu

Johnn Judd (2nd Author)
*Department of Civil and Construction Engineering*
*Brigham Young University*
Provo, Utah, USA
johnn@byu.edu

***Abstract* — A load rating procedure that involves field testing and composite action considerations for girders affected by external effects is presented in this paper. In the proposed procedure, the critical vehicle sequence for the bridge is determined and the actual response of the bridge is measured using a series of runs by driving a vehicle of calibrated weight. To replace the data readings affected by external effects, the position of the neutral axis corresponding to a fully composite action is assumed. After applying this correction, the actual load rating is discretized and compared with the analytical load rating so that different contributions to the loading capacity are quantified. A three-span non-composite steel girder highway bridge was used to illustrate the procedure. Results indicate that although the unintended composite action was the dominant contribution, the contribution was unreliable for loads beyond the linear elastic regime. It was observed that corrections made based on composite action assumptions improve the understanding of the case study because the contributions due to additional stiffness lateral and longitudinal distribution would have been unrealistic otherwise.**

***Keywords* — *Field testing, composite action, load rating***

## I. INTRODUCTION

Highway bridges are essential to allow people to commute over physical or man-made obstacles, but the loading capacity of highway bridges can diminish over time for a variety of reasons. Fatigue, caused by quotidian traffic, damages caused by accidents, as well as impacts due to adverse environments, are common causes for a reduction in bridge load capacity [1]. Fortuitously, there can be additional inherent "reserve" load capacity in bridges which is not accounted for in the routine design. For example, the additional stiffness contributed by curbs, barriers and railings can change the load path and can increase load capacity [2]. Similarly, the actual load distribution within the bridge span, between the piers and abutments, can be significantly different from the theoretical distribution [3]. Partial restraint of the supports due to accumulation of debris can alter the distribution of load [4]. Composite action between the bridge girders and the bridge deck can also occur, even if the bridge was designed to act non-compositely [5]. Although partial restraint of the supports and unintended composite action can be unreliable contributions to capacity, an understanding of these effects is necessary to determine the overall load-carrying capacity of the highway bridge. In contrast to conventional design practice, i.e., application of theoretical equations given in the codes and specifications, non-destructive field testing is an effective way to evaluate unintended composite action and to determine the reserve load capacity.

## II. OBJECTIVES

This paper presents a procedure to obtain the actual loading capacity of a typical highway bridge based on experimental non-destructive field testing coupled with a determination of composite action. In brief, the procedure consists of determining the critical vehicle sequence for the bridge, obtaining the actual internal response of the bridge under vehicular loads, and comparing the measured response with the analytical response to systematically identify reliable and unreliable contributors to the loading capacity. The procedure is demonstrated using a three-span non-composite steel girder highway bridge with a concrete deck.

## III. METHODOLOGY

In this section, the procedure of a non-destructive bridge field testing coupled with assumptions of composite action is described and applied to a case-study bridge located in Laramie, Wyoming (United States). After identifying the relevant properties of the bridge for the analysis, girders were instrumented with strain gauges at the maximum positive and negative moment positions. A vehicle with calibrated weight is driven across the entire bridge at distinct transverse locations in to order to capture the response for every girder. The responses are collected in form of strain histories in which the corresponding time to the peak strain is identified. The strains associated with this time are used to compute strain profiles which are therefore employed for the calculation of internal moments and axial forces, separating the non-composite and composite action components. For the collected data that was affected by external effects unanticipated in the field testing, corrections on the load rating calculation are made based on fully composite action assumptions. The fully composite neutral axis is used to replace the strain readings that were impaired. With that, the actual load rating is calculated, and its corresponding live load effects are discretized. By comparing it with the existing load rating, the contributions due to longitudinal and lateral

distributions, additional stiffness in the system, deck flexure, and unintended composite action are disaggregated.

### *A. Laramie Bridge*

The methodology described above was illustrated using the westbound highway bridge on Interstate 80 over the Laramie River in Laramie, Wyoming. The bridge is non skewed and has three spans (Fig. 1). The outer spans are 18.3 m (60 ft) each, and the inner span is 22.9 m (75 ft). The original bridge was built in the 1960s with four nominally identical I-shape steel girders that were designed to be non-composite (Girders 1 through 4 in Fig. 2). The bridge was later widened by adding a fifth I-shape steel girder (Girder 5 in Fig. 2). This girder was still designed to act non-compositely, but a metal sheet was poured below the deck which connects it with Girder 4 to confer a higher bonding. All girders are bisymmetric, and the dimensions are identified in Table 1. The clear roadway width is 12.2 m (40 ft), and the deck is 191 mm (7.5 in.) thick in addition to an overlay of 25.4 mm (1 in.) throughout the whole bridge. The girders are supported by expansion bearings, except at the outer pier on the west side which is supported by a fixed bearing. Typical X-bracing cross-frames are present at the piers and steel diaphragms are present on the abutments. Also, X-bracing cross-frames are spaced 6.10 m (20 ft) on the outer spans and 7.62 m (25 ft) on the inner span.

The critical moment positions for the Laramie Bridge occur at the western outer span at the 0.4 times of the outer span measured from the abutment for the positive moment, and at the pier for the negative moment. For comparison purposes, the actual responses of the same longitudinal positions for positive and negative moments were determined. Thus, all girders at the corresponding positions were instrumented with ST350 strain gauges [6]. At the positive moment, each girder received three strain gauges: at the center of the bottom flange, on the web at a position 254 mm (10 in.) and 508 mm (20 in.) above the bottom of the bottom flange. At the negative moment, each girder received two strain gauges: at the center of the bottom flange, and on the web at a position 508 mm (20 in.) above the bottom of the bottom flange. It is recognized that this setup only allows the calculation of the live load effects.

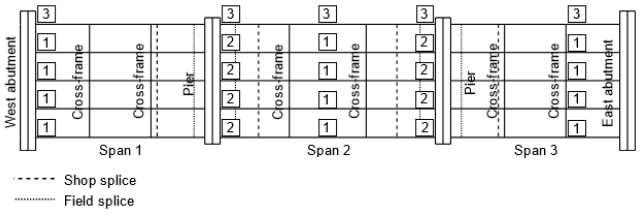



Fig. 1. Plan view of the Laramie bridge.

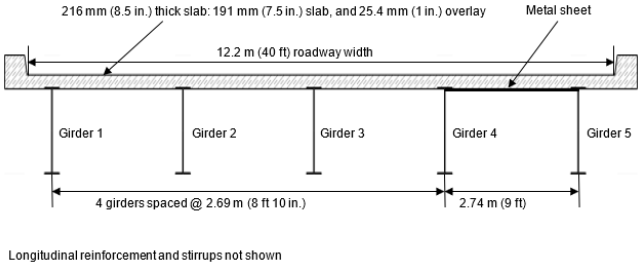



Fig. 2. Cross-sectional view of Laramie bridge (direction of traffic).

Table 1. Nominal section girder dimensions, mm (in.) for locations identified in Fig. 1.

| Location | Flange | | Web | |
|---|---|---|---|---|
| | Width | Thickness | Width | Thickness |
| 1 | 305 (12) | 19.1 (0.75) | 1300 (51) | 9.53 (0.38) |
| 2 | 305 (12) | 25.4 (1.00) | 1300 (51) | 9.53 (0.38) |
| 3 | 305 (12) | 19.1 (0.75) | 1300 (51) | 11.1 (0.44) |

### *B. Instrumentation and Field Testing*

In the field test, a vehicle with calibrated weight was driven at crawl speed over the bridge in successive runs that traverse the width of the roadway according to the requirements given in *The Manual for Bridge Evaluation* [7]. A total of fifteen runs were conducted to obtain the maximum responses for every girder. The runs were conducted from right to left, relative to the direction of travel. In the first run, the center of the right front wheel was positioned 0.91 m (3 ft) from the right curb. The position of each successive run was 0.61 m (2 ft) offset to the left of the previous run (Fig. 3). As a non-destructive field test, the vehicular load was selected so that responses generated on the bridge are within the linear elastic regime. To simulate side-by-side vehicle loadings, responses obtained in individual runs are superimposed.

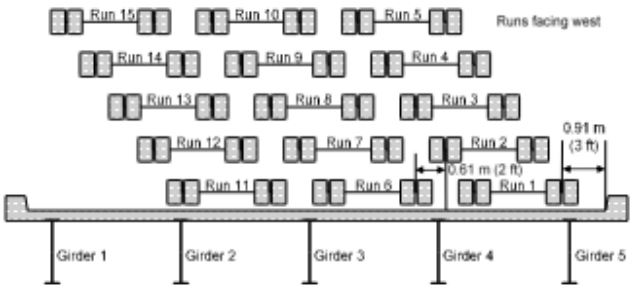



Fig. 3. Transverse position of successive runs of the field testing.

During the field test, a strain history (i.e., a graph that illustrates the strain variation over time due to the vehicle) was collected for every girder and run. For each run, the girder with the peak strain and the corresponding time of occurrence were selected. This girder was identified as the critical girder for the run. The time that the peak strain occurred was also applied to the other strain gauge of the critical girder so that a linear strain profile was computed

(Fig. 4). The same selected time was also applied to all the non-critical girders to obtain their respective strain profiles. For the negative moment, the peak strain when the vehicle was at the outer span was selected instead. Although a larger strain occurs when the vehicle is on the inner span, this time is not used because it does not allow the statical moment to be calculated, which is used later to determine the actual live load. At the positive moment, the strain profiles were obtained based on the best fit least square regression line since three strain gauges were mounted on each girder, while at the negative moment, the strain profiles were obtained by linearly connecting the strain readings on the two strain gauges mounted on each girder.

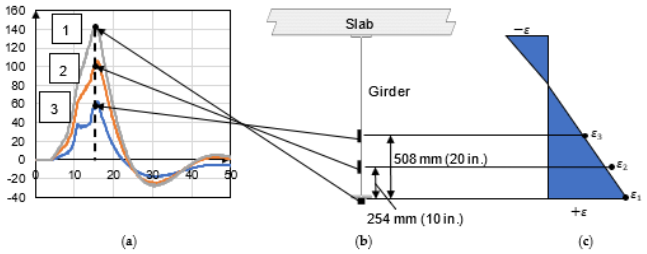



Fig. 4. Computation of a strain profile at the positive moment location: (**a**) strain history and maximum strains; (**b**) bridge cross-section; (**c**) strain profile.

### *C. Data Reduction*

Live load internal stresses were calculated based on the measured strain profile and the material and geometric properties of the girder and the deck. The calculated total live load internal stress profile (Fig. 5(a)) was decomposed into a live load axial component (Fig. 5(b)) and a live load flexural component (Fig. 5(c)). The girder live load axial stress, $\sigma_{cg}$ is equal to the total stress at the center of gravity of the girder. The maximum girder live load flexural stress is equal to the live load total stress at the extreme fiber of the bottom flange, $\sigma_0$ minus $\sigma_{cg}$. The live load axial force in the girder, $N$ is equal to the product of $\sigma_{cg}$ and cross-sectional area of the girder. The live load axial force in the deck was assumed to be equal to $N$ to maintain the equilibrium of the system. The live moment of the girder, $M_{girder}$ was obtained as the product of the girder live load flexural stress and the girder section modulus. The live load moment in the deck, $M_{deck}$ is the product of $M_{girder}$ and the ratio of the deck and girder flexural stiffnesses. The live load moment due to interaction between the girder and the deck was calculated as the product of $N$ and the distance from the girder to deck centers of gravity, $a$. The total live load moment, $M_{total}$ for the girder is the sum of the live load internal moments.

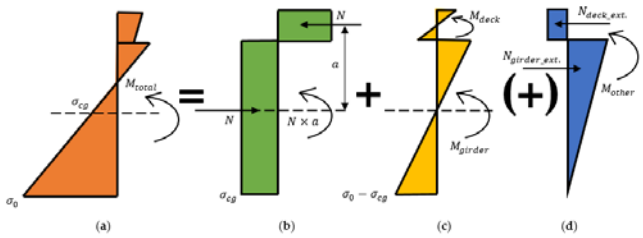



Fig. 5. Decomposition of a stress profile due to live load: (**a**) total stress; (**b**) axial stress; (**c**) flexural stress; (**d**) stress due to external effects.

The procedure of obtaining $M_{total}$ was modified as some of the collected data seemed to be affected by external effects. It was observed that the neutral axis of Girder 1 through 4 was below the theoretical neutral axis even for runs that are transversely close to the respective girders. As an example shown in Fig. 6(a) and Fig. 6(b), which shows the stress profiles of Girder 2 and Girder 3 (most stressed girders according to Fig. 3 for Run 9), the neutral axes fall below the strain gauge mounted on the web at a position 508 mm (20 in.) (as a bisymmetric girder, the theoretically estimated non-composite neutral axis for Girders 1 to 4 at the positive moment is at the half-center line, a position 667 mm [26.3 in.] above the bottom of the bottom flange). This is contrary to the expectations since, even within the linear elastic regime, some composition action is expected [5]. The accidental placement of strain gauges at the area of influence of additional stiffeners present on the bridge is the most likely cause for these discrepancies (Fig. 7). It is likely that the out-of-plane restraints, as well as residual stresses accumulated between the stiffener and girder web may have impacted the readings on the strain gauges mounted on the girders' web [8], although strain gauges mounted on the bottom flanges were possibly affected in a lesser degree. This speculation is confirmed by the fact that the neutral axis on Girder 5 at the positive moment and on any girder at the negative moment were not affected, as no stiffeners are present on those positions.

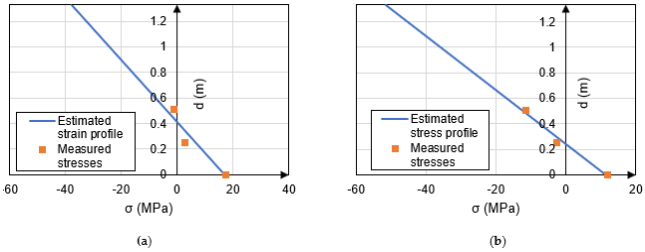



Fig. 6. Stress profiles affected by the transverse stiffeners for Run 9. (**a**) Girder 2. (**b**) Girder 3.

The inconsistencies showed above can be corrected by modifying the live load internal stress decomposition shown in Fig. 5(a), Fig. 5(b), and Fig. 5(c) by adding a new component as shown in Fig. 5(d). In this case, although the measured live load stress profile is unchanged, the live load axial component and the live load flexural component are obtained based on the measured stress at the bottom of the bottom flange and the assumption of a theoretically estimated

composite action to define the neutral axis position (considering, *a priori*, that the live load stress profile was affected by external effects). Under this condition, and assuming that the bottom strain gauge was not affected by the stiffener, $\sigma_{cg}$, the live load maximum girder flexural stress ($\sigma_0$ minus $\sigma_{cg}$), $N$, $M_{girder}$, and $M_{deck}$ are calculated the same way as previously. Despite not being necessary for the computation of the internal moments for this study, a stress profile caused by the external effect is developed resulting from the balance with the assumed condition and measured stress profile. The geometry of the stress profile caused by the external effect is consistent with the hypothesis as the strain gauges on the web are expected to read additional stresses not accounted by the live load flexural nor axial components. Thus, a resultant force, $N_{girder\_ext.}$ and a resultant moment, $M_{other}$ on the girder are shown. Since composite action is expected, a resultant force on the deck, $N_{deck\_ext.}$ may also exist to equilibrate the system.

In this approach, it was assumed that the readings for the strain gauge installed at the bottom of the bottom flange at Girder 1 through 4 were not affected by the transverse stiffeners. The readings of the strain gauges on the web were substituted by a theoretical neutral axis assuming a fully composite action behavior. Although the assumption is not conservative, it best describes the behavior of non-composite girders when subjected to loads within the linear elastic regime, as the bonding and friction between girder and deck are still not overcome [4]. It is recognized that the assumption does not cover the possibility of the girders having a partial composite behavior nor the actual behavior of the girders falling beyond the theorical limits. The theoretical fully composite neutral axis is located at the center of gravity of the combined system (girder plus deck). This is calculated based on principles of mechanics of materials, and the effective tributary width of the deck is estimated and converted into an equivalent width of steel, based on the modular ratio of steel to concrete. The fully composite neutral axis is equal to 1.19 m (46.8 in.) for Girder 1 and equal to 1.21 m (47.8 in.) for Girders 2, 3 and 4, measuring form the bottom of the bottom flange.


Fig. 7. Placement of strain gauges near transverse stiffeners.

### *D. Limit States Considerations*

According to AASHTO, the loading capacity of highway bridges are obtained through calculation of the rating factor, *RF*, for each individual girder, with the lowest individual girder rating factor controlling the load rating of the bridge [9]. Among the different limits used to load rate bridges, the strength inventory *RF* is generally the most critical, and thus adopted in this research. As part of a research sponsored by the WYDOT (Wyoming Department of Transportation), since the bridge was originally designed and evaluated using the Load Factor Design (LFD) method and Load Factor Rating (LFR) limit states, respectively, the *RF* in this paper is also determined using the LFD method. Thus, all the analytical values are based on *Standard specifications for highway bridges* [10].

*RF* is defined as the ratio between the reserve capacity for live load and the limit state design live load and, in the LFD method, it is expressed in terms of the number of HS20 standardized design truck loads,

$$RF = (R_n\text{-}1.3D)/(2.17LL[1+I]) \quad (1)$$

where $R_n$ is the capacity of the member, $D$ is the dead load effect on the member, $LL$ is the live load effect of the member, and $I$ is the dynamic impact factor. If *RF* is greater than 1.0, the bridge is deemed adequate for the design load, and if *RF* is less than 1.0, the bridge is deemed inadequate. It was obtained by WYDOT that the analytical load rating at the critical positive and negative moment positions are 0.92 and 1.05, respectively.

## IV. RESULTS AND DISCUSSION

In this section, the data reduction for the actual load rating is presented. Additionally, the disaggregated ratio of load ratings corrected based on fully composite action assumptions is also discussed.

### *A. Actual Load Rating*

In this study, no inspections were conducted *in-situ* to determine the actual dead load effects of the bridge. As a result, it was assumed that the actual $R_n$ and $D$ are equal to the values obtained analytically. In addition, since no dynamic tests were conducted, the actual $I$ was also assumed to be equal to the analytical.

The actual live load accounting for dynamic impact, $LL(1+I_E)$, is calculated as follows:

$$LL(1+I_E) = (M_{HS20}/M_{TRK\_OS})M_{girder} \\ (M_{TRK}/M_{TRK\_OS})m_E(1+I_E) \quad (2)$$

where $M_{HS20}$ is the maximum analytical moment due to the HS20 design truck at the location of interest, $M_{TRK}$ is the analytical maximum moment (positive or negative) due to the vehicle used in the field test, $M_{TRK\_OS}$ is the analytical maximum moment due to the vehicle used in the field test when the vehicle is at the outer span, $m_E$ is the actual live load multi-presence factor, and $I_E$ is the actual impact factor. The values of $M_{TRK}$ and $M_{TRK\_OS}$ are theoretical values that were obtained using a line-girder analysis in which the vehicle used in the field was modeled as three concentrated loads and

placed at the critical moment positions (the same model that WYDOT used to obtain the analytical load rating). It was determined that $M_{TRK}$ and $M_{TRK\_OS}$ are the same (since both occur at the outer span) and equal to 693 kN-m (511 k-ft) for the positive moment. For the negative moment, $M_{TRK}$ (which occurs at the inner span) is equal to -455 kN-m (-335 k-ft), and $M_{TRK\_OS}$ is equal to -381 kN-m (-281 k-ft). As defined in the *Standard Specifications for Highway Bridges*, $m_E$ accounts for the improbability of coincident maximum loading and is equal to 1.0 for single and 2-vehicle loadings, and equal to 0.9 for 3-vehicle loadings.

Based on the transverse positions of successive runs of the field testing shown in Fig. 3, a total of ten 2 side-by-side vehicle loadings and five 3 side-by-side vehicle loadings were considered. According to the experimental data, the two side-by-side vehicle loadings involving Runs 1 and 6 superimposed controlled (i.e., yielded the greatest $LL(1+I_E)$) at the positive moment, and the three side-by-side vehicle loadings involving Runs 3, 8 and 13 superimposed controlled at the negative moment. The controlling girder at the positive moment is Girder 5, and the controlling girder at the negative moment is Girder 3. The values of each variable to determine the load rating at the critical positive and negative moment positions are shown in Table 2. Thus, it was determined that $LL(1+I_E)$ is equal to 469 kN-m (346 k-ft) and -546 kN-m (-403 k-ft) at the positive and negative moment positions, respectively. As a result, the actual load rating at the critical positive and negative moment positions are equal to 1.75 and 1.53, respectively. The actual load ratings show that the bridge is deemed adequate at both critical moment positions.

Table 2. Values of variables for the calculation of the actual load ratings using Eq. (1) and Eq. (2).

| Variables | Positive Moment | Negative Moment |
|---|---|---|
| $R_n$, kN-m (k-ft) | 2300 (1700) | -2870 (-2110) |
| $D$, kN-m (k-ft) | 406 (299) | -807 (-595) |
| $M_{HS20}$, kN-m (k-ft) | 871 (642) | -784 (-578) |
| $M_{TRK}$, kN-m (k-ft) | 693 (511) | -455 (-335) |
| $M_{TRK\_OS}$, kN-m (k-ft) | 693 (511) | -381 (-281) |
| $M_{girder}$, kN-m (k-ft) | 294 (217) | -197 (-145) |
| $m_E$ | 1.0 | 0.90 |
| $I_E$ | 0.27 | 0.26 |

### *B. Disaggregated Load Rating Comparison*

The experimental load rating ($RF_E$) relative to the analytical load rating ($RF_A$) is defined as follows:

$$RF_E/RF_A = (M_{TRK\_OS}/M_{TRK})((\sum STAT_A/\sum STAT_E)((DF_A \times m_A)/(DF_E \times m_E))(M_{LE}/M_{total})(M_{total}/(M_{girder}+N\times a))(M_{girder}+N\times a)/M_{girder}) \quad (3)$$

where $\sum STAT_A$ and $\sum STAT_E$ are the analytical and actual statical moment, respectively, $DF_A$ and $DF_E$ are the analytical and actual live load distribution factor, respectively, $m_A$ is the analytical live load multi-presence factor, and $M_{LE}$ is the elastic longitudinal adjustment moment. The values of these variables at the critical positive and negative moment positions are shown in Table 3. Details for the calculation of each variable were shown in previous studies [3, 4].

Table 3. Values of the variables for the actual and analytical load rating comparison using Eq. (3).

| Variables | Positive Moment | Negative Moment |
|---|---|---|
| $\sum STAT_A$, kN-m (k-ft) | 1670 (1230) | 2470 (1820) |
| $DF_A$ | 0.81 | 0.81 |
| $m_A$ | 1.0 | 1.0 |
| $\sum STAT_E$, kN-m (k-ft) | 1320 (974) | 2350 (1730) |
| $DF_E$ | 0.81 | 0.91 |
| $M_{LE}$, kN-m (k-ft) | 443 (327) | -331 (-244) |
| $M_{total}$, kN-m (k-ft) | 454 (335) | -315 (-232) |
| $N\times a$, kN-m (k-ft) | 150 (111) | -112 (-82.4) |

Eq. (3) also shows the comparison of different contributions between the actual and analytical load ratings. The term $(M_{TRK\_OS}/M_{TRK})$ is the contribution of the critical span adjustment, $(\sum STAT_A/\sum STAT_E)$ is the contribution of the additional stiffness in the system, $((DF_A\times m_A)/(DF_E\times m_E))$ is the contribution of the differences in lateral distribution, $(M_{LE}/M_{total})$ is the contribution of the differences in longitudinal distribution, $(M_{total}/(M_{girder}+N\times a))$ is the contribution of deck flexure, and $((M_{girder}+N\times a)/M_{girder})$ is the contribution of the unintended composite action. The disaggregation of the actual load rating is carried out based on principles of mechanics of materials, linear responses of the bridge, and equilibrium [3]. Note that in this case study, the same value is obtained by taking the inverse ratio of live load effects because $R_n$ and $D$ were kept constant. The values for each contribution are shown in Table 4.

Table 4. Values of the contributions for the actual and analytical load rating comparison using Eq. (3).

| Contributions | Positive Moment | Negative Moment |
|---|---|---|
| Critical span adjustment | -- | 0.839 |
| Additional stiffness | 1.266 | 1.053 |
| Lateral distribution | 0.995 | 0.979 |
| Longitudinal distribution | 0.974 | 1.050 |
| Deck flexure | 1.023 | 1.019 |
| Unintended comp. action | 1.513 | 1.568 |

The contribution due to critical span adjustment applies only if the span instrumented in the field test differs from the

critical span determined for the analytical rating. It does not apply for the positive moment because the strain gauges were mounted at the same span (outer) as to where the maximum moment occurred. This contribution is less than 1.0 at the negative moment because the negative moment from the inner span loading is larger than the negative moment from the outer span loading. The contribution due to additional system stiffness accounts for other components in the bridge, such as curbs and railings. These components are not considered in the analysis and drive away the live load resisted by the girders. Therefore, the results are expected to be higher than 1.0. Although the values on the negative moment seemed to be accurate, the 27% increase observed on the positive moment may seem to be an overestimation. This inaccuracy may be caused by changes in the actual support condition due to accumulation of debris, which has a direct impact on the statical moment. If this value were true, the bridge owner could solely rely on this contribution to strengthen the bridge. The contribution due to lateral distribution accounts for the difference of the fraction of the moment carried by the critical girder relative to the total moment across the bridge in the actual and analytical load ratings. Minor differences between $DF_A$ and $DF_E$ are expected because the former is conservatively expressed as a function of the transverse girder spacing, whereas the actual lateral distribution of load also depends on the girder and edge stiffnesses [11]. The contribution due to longitudinal distribution accounts for the difference of moment distribution to the positive and negative moment positions in the actual and analytical load ratings. At the positive and negative moment positions, it was shown that the actual moment was almost as stiff as (3% and 5% difference, respectively) compared to the line-girder analysis. These differences show that more statical moment is going to the positive moment and less to the negative moment than predicted in the analysis. The contribution due to deck flexure represents the contribution besides the live load non-composite girder moment ($M_{girder}$) and interaction component of the live load moment ($N \times a$). This contribution is expected to be small (1% to 3%) because the curvature of the girder and deck deflection is the same, but the flexural stiffness of the former is much larger than the latter. The contribution due to unintended composite action accounts for the composite action developed between girder and deck for girders designed to be non-composite. In this study, it was observed that this contribution is dominant for both positive and negative moment positions. Although this effect is acceptable for a linear-elastic regime, it would become unreliable for loads beyond this limit. As a result, to determine a reliable ratio of load ratings, this contribution is divided out from the final value. Based on the load ratings calculated above, the ratio of the actual to the analytical load rating is equal to 1.90 at the positive moment and equal to 1.45 at the negative moment. The value is interpreted as the number of times that the actual load rating is relative to the analytical. If the contribution due to unintended composite action is removed, the ratio of the load ratings is then equal to 1.26 at the positive moment and equal to 0.93 at the negative moment. The reduction in the product for the negative moment reflects conservation of the statical moment (i.e., the load rating cannot simultaneously increase for both positive and negative moments). As a result, the actual reliable load rating is equal to 1.16 at the positive moment and 0.98 at the negative moment.

### *C. Comparison with Uncorrected Data*

If the data were not corrected based on the fully composite action assumptions for Girders 1 through 4 at the positive moment, the ratio of load ratings would be equal to 1.90 at the positive moment and equal to 1.41 at the negative moment with the same respective controlling side-by-side vehicle loadings and critical girders. The values for each contribution are shown in Table 5 (the values of the variables for each contribution were not shown).

Table 5. Values of the contributions for the actual and analytical load rating comparison using Eq. (3) if no assumptions of composition action were made.

| Contributions | Positive Moment | Negative Moment |
|---|---|---|
| Critical span adjustment | -- | 0.839 |
| Additional stiffness | 1.761 | 1.920 |
| Lateral distribution | 0.677 | 0.968 |
| Longitudinal distribution | 1.027 | 0.567 |
| Deck flexure | 1.023 | 1.019 |
| Unintended comp. action | 1.513 | 1.568 |

Although the values of the ratio of load ratings do not differ much from the values of the proposed procedure, it is possible to see that the contribution of additional stiffness, lateral distribution (positive), and longitudinal distribution (negative) are not realistic if no composite action corrections are made. For Girders 1 through 4, since the neutral axes of the strain profiles are below the girder centers of gravity in addition to a positive strain at the bottom of the bottom flange (Fig. 5), $\sigma_{cg}$ is negative and so is $N \times a$. Even though a negative $\sigma_{cg}$ contributes to the magnitude of $M_{girder}$, it is overweighted by the negative $N \times a$ which leads to a negative $M_{total}$. This reduces the values of $\sum STAT_E$ and $DF_E$ and consequently affects the contributions due to additional stiffness, lateral and longitudinal distributions. As shown in Table 5, the contributions are unrealistic as they are significantly different from the analysis (it shows that the actual additional stiffness is almost twice the analytical, while the actual lateral and longitudinal distributions are almost cut in half). The contributions due to the deck flexure and unintended composite action were not affected because they only depend on the properties of the critical girder, which were not changed relative to the results of the proposed procedure.

## V. CONCLUSION

In this paper, a procedure for a bridge field testing coupled with assumptions of composite action was described and

applied to a three-span five girder non-composite steel girder bridge located in Laramie, Wyoming. In the proposed procedure, the critical vehicle sequence for the bridge is determined, and the elastic response is measured for a series of vehicle tests by installing strain gauges at the critical positive and negative moment positions. To replace the data readings affected by the external effects, the position of the neutral axis corresponding to a fully composite action is assumed. After applying this correction, the actual load rating is discretized and compared with the analytical load rating so that different contributions to the loading capacity are quantified. For the bridge analyzed in this study, the results showed that the reliable load rating was equal to 1.26 and 0.93 at the positive and negative moment positions, respectively. Although the unintended composite action was the dominant contribution, it was removed because the bonding between girders and deck will likely be overcome for loads beyond the linear elastic regime. In addition, the results indicated that if assumptions of fully composite action for girders affected by external effects were not considered, the contributions due to additional stiffness, lateral and longitudinal distributions would have been unrealistic.

The study demonstrated that the proposed load rating procedure is a promising method to understand the behavior of highway bridges. In the proposed method, different contributions to the loading capacity are quantified and can be used in the decision-making process for maintenance, rehabilitation, and posting of highway bridges. The correction applied based on composite action is not only regarded as an alternative to replace data affected by external effects, but also can be adapted for field testing of highway bridges where mounting strain gauges on the web is not feasible (e.g., due to damages caused on the surface, lack of access, or desire to avoid impacting traffic flow). Additional analysis using a finite element model of the bridge is needed to accurately determine the degree of composite action.

## REFERENCES

[1] D. B. Colleti, B. Chavel, and W. J. Gatti, “Challenges of skew in bridges with steel girders,” Transp. Res. Rec., vol. 2251, pp. 47-56, 2011.

[2] M. G. Barker, “Steel girder bridge field test procedures,” Constr. Build. Mater., vol. 13, pp. 229-239, 1999.

[3] R. Lu, Experimental load rating of skewed steel girder highway bridges, PhD dissertation, Dept. of Civil and Architectural Engineering, Univ. of Wyoming, 2020.

[4] M. G. Barker, “Quantifying field-test behavior for raitng steel girder bridges,” J. of Bridge Eng., vol. 6, pp. 254-261, 2001.

[5] M. Yarnold, T. Golecki, and J. Weidner, “Identification of composite action through truck load testing,” Front. Built Environ., vol. 4, pp. 74, 2018.

[6] BDI (Bridge Diagnostics Inc.), ST350 Operation Manual, Boulder, CO, 2012.

[7] AASHTO (American Association of State Highway and Transportation Officials), The Manual for Bridge Evaluation, 2nd ed., Washington, DC, 2011.

[8] C. R. Hendy, and F. Presta, “Transverse web stiffeners and shear moment interaction for steel plate girder bridges,” The Struct. Eng., vol. 86, pp. 017, 2008 [Proceedings of 7th International Conference on Steel Bridges, 4-6 June 2008, Guimarães, Portugal. ECCS, pp. 8, 2008].

[9] M. Sanayei, M., A. J. Reiff, B. R. Brenner, and G. R. Imbaro, “Load rating of a fully instrumented bridge: Comparison of LRFR approaches,” J. Perform. Constr. Facil., vol. 30, pp. 04015019, 2016.

[10] AASHTO (American Association of State Highway and Transportation Officials), Standard Specifications for Highway Bridges, 16th ed., Washington, DC, 1996.

[11] F. N. Catbas, H. B. Gokce, and M. Gul, “Practical approach for estimating distribution factor for load rating: Demonstration on reinforced concrete T-beam bridges,” J. Bridge Eng., vol. 14, pp. 652-661, 2012.

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index No: ES-13-33*

# Modeling the Relationship between Profitability and Market Share of Licensed Commercial Banks in Sri Lanka

Imesha Hettiarachchi (1st Author)
*Department of Mathematical Sciences*
*Wayamba University of Sri Lanka*
Kuliyapitiya, Sri Lanka
shana.hettiarachchi@gmail.com

Piyal Ekanayake (2nd Author)
*Department of Mathematical Sciences*
*Wayamba University of Sri Lanka*
Kuliyapitiya, Sri Lanka
piyalekanayake@wyb.ac.lk

Lakshan Aponsu (3rd Author)
*Department of Mathematical Sciences*
*Wayamba University of Sri Lanka*
Kuliyapitiya, Sri Lanka
lakshana@wyb.ac.lk

***Abstract* — Sri Lankan banking industry comprises of two major types of establishments viz. Licensed Commercial Banks (LCB) and Licensed Specialized Banks of which the profitability of the former being determined by the size of the firm, indicated by total assets held by the bank. In this study, the relationship between market share and profitability of LCBs was investigated using Multiple Linear Regression (MLR) analysis and considering the variables unique to the Sri Lankan banking industry. Quarterly data were obtained from the Central Bank of Sri Lanka from 2008 to 2020. In addition to MLR, analysis of variance and time-series analysis was also carried out to ensure the reliance of the results and conclusions. It was found that the market share represented by Deposit Customers and Loan Customers have a positive relationship with the profitability measured by Profit After Tax (PAT). Further, market share explained a substantial 96.5% of the variability in profitability. The model predicted that a unit increase in deposits and loans lead to an increase in profitability by 6.1% and 10.7% respectively. Finally, it could be concluded that the loans granted to bank customers accounted for increasing the profitability of LCBs in Sri Lanka than the deposits made by the customers. Accordingly, the LCBs can review their strategies to optimize profitability based on the loans granted to and deposits received from the customers.**

***Keywords — Deposit, Loan, Profitability***

## I. INTRODUCTION

The banking sector of any country determines the financial stability of economic activities while providing funds for long term investments. The high profitability of banks accounts for financial stability which needs consistency over long periods for ensuring sustainable growth of the financial system. The success of the banking industry is presented to the public mostly through the profitability statistics and their consistency over the years is achieved through prudent decisions by the top management. Precise strategies implemented to achieve effectiveness and efficiency in dealing with deposit portfolios and loans will earn customer confidence which is crucial to boost profitability. Therefore, the managers should possess updated knowledge on present and future trends not only in the banking sphere but also in the competitive business world. It is reported in the literature that some researchers: [1], [10], [8], and [7] have focused on modelling the correlation between market share and profitability in order to expose potential avenues for achieving reliable and sustainable profitability. More specifically the following works have presented the most relevant findings to the present study.

A study, [6], on the impact of bank capital on profitability and risk was carried out by applying the Generalized Method of Moment (GMM) technique on bank level data from 42 countries in Asia covering the period from 1994 to 2008. It was pointed out that different profitability variables influence the persistence of profit while showing a positive (negative) relationship between capital and profitability (risk) in the Asian region. The relationship between the market share and profitability of the banking sector in Nigeria was examined using nine years of secondary data obtained from the Stock Exchange in Nigeria from 2003 to 2011, [2]. Applying MLR with PAT as the dependent variable and both deposit and loan customers as independent variables, it was shown that the market share represented by those customers is positively correlated with the PAT of the banks in Nigeria. While concluding that loans granted to customers can increase the profitability in Nigerian banks, the authors recommended the use of other modelling techniques like multivariate adaptive regression splines (MARS) and conic multivariate adaptive regression splines (CMARS) in future studies.

The factors influencing the profitability of banks can be broadly classified into macroeconomic-, industry-, and bank-specific levels. Bank specific factors are confined to the functions carried out within the institution some of which are indicated by deposit ratio, capital ratio, liquidity ratio, overhead cost of management, and the bank size. Considering some of the above factors as independent variables, an MLR was used on the STATA statistical package to investigate their impact on the bank's profitability based on data from 12 commercial banks in Sri Lanka covering the period from 2011 to 2015, [5]. It was found that a positive relationship exists between the bank profitability and the bank specific factors: firm size, capital ratio, and deposit ratio while the liquidity was shown to be an insignificant factor which is negatively related to the profitability of Sri Lankan

commercial banks. The impact of market share and concentration ratio on the profitability of banks in Bulgaria in Eastern Europe was studied by [4] using balanced panel data of 22 banks during the 5-years from 2006 to 2010 and choosing return on equity as the profitability indicator. This study found a statistically significant positive correlation between market share and profitability but no such relationship between the concentration and profitability. Further, having applied the pooled Ordinary least squares (OLS) estimation and random-effects methods, it was empirically shown that the performance of Bulgarian banks depends solely on the prudence of managerial decisions within the institutions and not on either macroeconomic or industry specific external issues.

The factors determining the performance of Spanish banks during the period 1999-2009 were examined in [9] using the GMM estimator which revealed that more loans in total assets and customer deposits, efficiency in transactions, and low uncertain assets ratio are associated with high bank profitability. Further, it was concluded that higher returns of the bank is proportional to the higher capital ratios subject to the profitability measured by the return on assets. Addressing a wider range of determinants within all bank-specific, industry-specific, and macro level spheres, the dynamic estimation technique was applied [3] to identify probability models for selected banks in Croatia for the period from 2007 to 2014. Projecting the profitability in terms of return on assets (ROA) and the ratio net interest margin (RNIM), they proved that asset size, loan portfolio, and gross domestic product create a positive impact on the banks' profitability while risks and administrative costs create a negative impact. They also showed a correlation between market concentration and profitability in addition to the influence of the capital adequacy ratio on ROA and RNIM.

In Sri Lanka, the banking industry is one of the most significant sectors of the economy and also it dominates the financial sector. The rapid booming of this sector began with the liberalization of the financial sector under the open economic policies introduced in the late 1970s. As of now, there are three major categories of institutions that are subject to licensing, regulation, and supervision by the Central Bank of Sri Lanka in the public interest because they accept deposits from the general public. They are licensed commercial banks, licensed specialized banks, and registered finance companies. With the establishment of private banks and branches of leading foreign banks over the years, there are 23 licensed commercial banks of which 11 are locally incorporated while the other 12 are local branches of foreign banks. They all have a total of 1711 branches and other service outlets spread throughout the country.

The Sri Lankan banking industry comprises two major types of banks namely licensed commercial banks and licensed specialized banks. A licensed commercial bank is a banking institution issued with a license by the Central Bank to carry on among other things, maintaining current accounts for customers from which money could be transferred by cheques and withdrawn on demand. The licensed specialized banks are different from the licensed commercial banks since they are not authorized to accept demand deposits from the public and therefore they do not maintain current accounts for customers. Also, they are not authorized to deal in foreign currency. The profitability of Sri Lankan commercial banks are mainly determined by the size of the firm indicated by the total assets held by the institution. Though the notion that higher market share is positively associated with higher profits are generally accepted in virtually all studies, it is still studied in every country due to the diverse and unique nature of the business environment in different countries.

Although the above scholars have examined the relationship between market share and profitability, it has not been possible to generalize their findings to the banking sector due to the dynamic business environments that change over the years with the change of policies. Some researchers have considered only the domestic commercial banks in Sri Lanka which may restrict the representativeness of the findings. This research considered all the commercial banks in Sri Lanka as it is important to identify high-net-worth customers with integrity and good credit rating whose patronage immensely contributes to the growth of banks and a developing economy. Having considered such uncertainties, this study was designed to address the relationship between market share and profitability in LCBs using variables peculiar to the Sri Lankan context. More specifically, we set our research objectives to identify how various deposits and loans affect the bank profitability and to develop a model that can be used to define the relationship between deposit customers, loan customers, and the PAT. Further, a predictive model for forecasting the Profitability of the ensuing year is also attempted. The specificity of the prediction model is that it takes into account deposit and loan indicators of bank profitability. Quarterly data were collected from the Central Bank covering a period of twelve years from 2008 to 2020 and multiple regression analysis was used to test the hypotheses.

## II. OBJECTIVES

Our objectives are to identify how various deposits and loans affect the profitability of LCBs; develop a model for presenting the relationship between deposit customers, loan customers, and profit after tax; and to explore the possibility for a time-series model to forecast realization.

## III. METHODOLOGY

Multiple linear regression analysis was used to identify the relationship between market share and profitability of licensed commercial banks using quarterly records from 2008Q4 to 2020Q4 which were obtained from the official website of the Central Bank of Sri Lanka. Deposit customers (DC) and loan customers (LC) were used to represent the market share of banks with a quantitative approach to analyze the data.

Bank Profitability (P), which is measured by the Operating Profits After Corporate Tax (PAT), was used as the dependent variable of the multiple linear regression. Deposits made by customers (DC), which was quantified by Deposits to Total Assets, and the loans granted to customers (LC), measured by Total Loans and Advances to Total Assets were

used as the independent variables in the regression model as presented in Table 1 and Equation (1).

Table 1. Operationalization of Variables

| Concept | Variable | Indicator | Measurement |
|---|---|---|---|
| Profitability | P- Profitability | Number | Operating Profits After Corporate Tax (PAT) |
| Market share | DC-Deposit Customer | Ratio | Total deposits/Total assets |
| | LC-Loan Customer | Ratio | Total Loans and Advances /Total assets |

$$P = \beta_0 + \beta_1 *(DC) + \beta_2 *(LC) + \varepsilon \quad (1)$$

where; $\beta_0$ is a constant, $\beta_1$and $\beta_2$ are the coefficients of the independent variables, and ε is the associated error term.

Apart from identifying the relationship between bank profitability and market share, it is vital to understand the expected profitability which enables bank managers to successfully implement their future plans. Therefore, Box-Jenkins approach of time series analysis was used to forecast the Bank Profitability by using the quarterly data from 2008Q4 to 2020Q4 obtained from the official website of the Central Bank of Sri Lanka.

## IV. RESULTS AND DISCUSSION

### *A. Regression Analysis*

The following hypothesis was assumed to assess the significance of correlation between profitability and explanatory variables.

$H_0$: Correlation is not significant

$H_1$: Correlation is significant.

The correlation between the variables was examined using the Pearson's Correlation coefficient at 5% level of significance, summarized in Table 2.

Table 2. Correlations

| | | P | DC | LC |
|---|---|---|---|---|
| **P** | Pearson Correlation | 1 | .172 | .424 |
| | Sig. (2-tailed) | | .023 | .002 |
| | N | 52 | 52 | 52 |
| **DC** | Pearson Correlation | .172 | 1 | .148 |
| | Sig. (2-tailed) | .023 | | .026 |
| | N | 52 | 52 | 52 |
| **LC** | Pearson Correlation | .424 | .148 | 1 |
| | Sig. (2-tailed) | .002 | .026 | |
| | N | 52 | 52 | 52 |

The results in Table 2 indicate that there are significant positive correlations between the dependent variable and the independent variables at 5% significance level. Therefore, $H_0$ is rejected and it can be concluded that the correlation is statistically significant.

The normality of the variables was examined with the following hypothesis using the Kolmogorov-Smirnov Test and the Shapiro-Wilk Test, which yielded the values summarized in Table 3.

$H_0$: Data are normally distributed.

$H_1$: Data are not normally distributed

Table 3. Tests of Normality

| | **Kolmogorov-Smirnov[a]** | | | **Shapiro-Wilk** | | |
|---|---|---|---|---|---|---|
| | Statistic | df | Sig. | Statistic | df | Sig. |
| P | .129 | 52 | .031 | .922 | 52 | .002 |
| DC | .119 | 52 | .064 | .962 | 52 | .097 |
| LC | .122 | 52 | .051 | .952 | 52 | .065 |

According to the results in Table 3, profitability (P) is not normally distributed (p = 0.002) but DC (p = 0.97) and LC (p = 0.065) are normally distributed at 5% significance level. Therefore, the dependent variable was log transformed. The performance of the model was evaluated in terms of the coefficient of determination, which resulted in the figures in Table 4.

Table 4. Interpretation of Model Summary

| **Model** | **R** | **R Square[b]** | **Adjusted R Square** | **Std. Error of the Estimate** | **Durbin-Watson** |
|---|---|---|---|---|---|
| 1 | .978[a] | .965 | .965 | .72846 | 1.220 |

According to Table 4, the coefficient of determination is .965 which means that 96.5% of the variation of bank profitability can be explained by the amount of deposits made by customers (DC) and the amount of loans granted to customers (LC).

Analysis of Variance test was also employed to test the statistical significance of the model using the following hypothesis and the results are shown in Table 5.

$H_0$: Model is not significant

$H_1$: Model is significant

Table 5. ANOVA Table

| Model | | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| 1 | Regression | 5763.980 | 2 | 2881.990 | 431.085 | .000[b] |
| | Residual | 26.199 | 50 | .531 | | |
| | Total | 5790.513[d] | 52 | | | |

It is clear from Table 5 that p-value of the F-statistic is less than 0.05 and accordingly $H_0$ is rejected at 5% significance level. Therefore, it is evident that the independent variables significantly predict the dependent variable.

The non-zero existence of model coefficients was also looked into with the following hypothesis and the values obtained are given in Table 6.

$H_0$: $\beta_i = 0$

H1: $\beta_i \neq 0$ ;i =1,2.

Table 6. Coefficients[a]

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
|---|---|---|---|---|---|---|
| | | B | Std. Error | Beta | | |
| 1 | DC | .059 | .026 | .395 | 2.305 | .025 |
| | LC | .102 | .029 | .603 | 3.522 | .001 |

Results in Table 6 show that deposits made by customers (DC) (p = .025< 0.05) and amount of loans granted to customers (LC) (p = .001< 0.05) have a significant impact on the log of bank profitability. The model predicts that 1-unit increase in deposits leads to an increase in profitability by 6.1% and 1-unit increase in Loans increases the profitability by 10.7%. Therefore, it can be concluded that with the increment of the deposits and loans, the profitability of the bank also increases. The assumptions of regression analysis were checked to confirm the accuracy of the identified model and the results are given below.

a. Homoscedasticity of Residuals

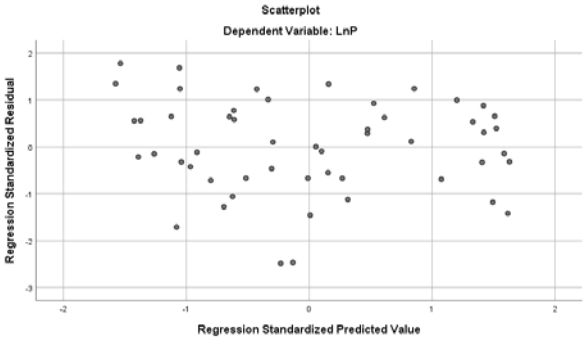



Fig. 1. Scatter plot of Residuals against the predicted value

According to Figure 1, the scatter plot of residuals does not show any visible pattern and the points that are randomly scattered, which confirm the constant variance of the residuals.

b. Normality of Residuals

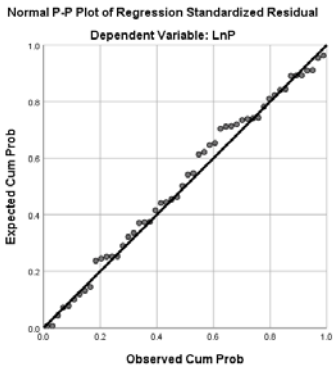



Fig. 2. Normality of residuals

According to Figure 2, the standardized residuals follow an approximately straight line indicating the normally distributed nature of residuals.

The multi-collinearity between the independent variables was checked according to the values of Variance Inflation Factor (VIF) in Table 7.

Table 7. Coefficients determining multicollinearity

| Model | Collinearity Statistics | |
|---|---|---|
| | Tolerance | VIF |
| DC | .978 | 1.022 |
| LC | .868 | 1.262 |

Table 7 clearly indicates that the values of VIF of both independent variables are less than 5 confirming the absence of multi-collinearity between the independent variables.The best fitted regression model can thus be presented as follows.

log(Profitability) = .059*(Deposit Customers) +.102* (Loan Customers) (2)

## *B. Time Series Analysis*

A Time Series analysis was also carried out to explore the existence of temporal behavior of profitability

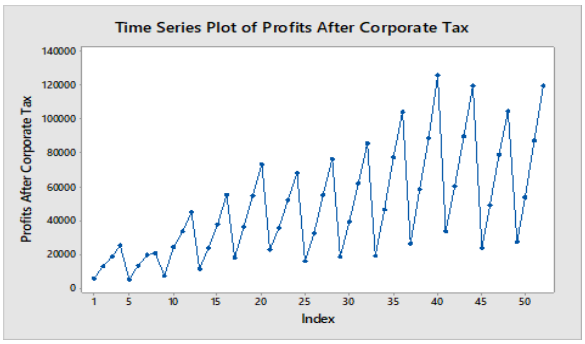



Fig. 3. Time Series of Profit After Corporate Tax

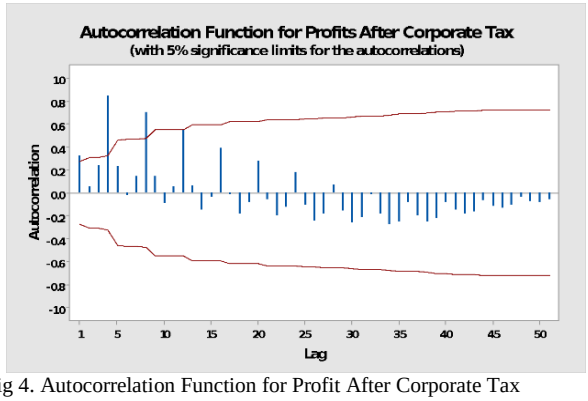



Fig 4. Autocorrelation Function for Profit After Corporate Tax

According to Figures 3 and 4, it can be observed that the series is not stationary as the plotted graph and the corresponding ACF show seasonal variation and an upward trend. Therefore, seasonal differencing was required to make the series stationary, as depicted in Figures 5 and 6.

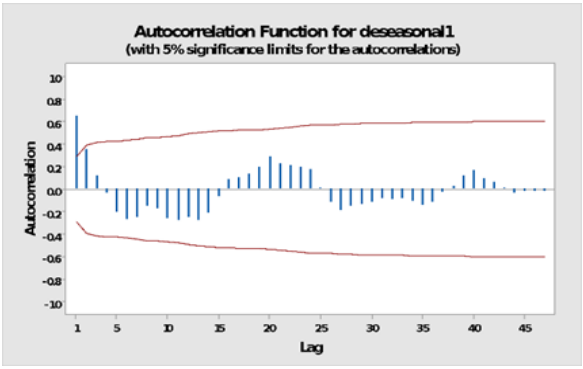



Fig. 5. ACF of first seasonal differenced series

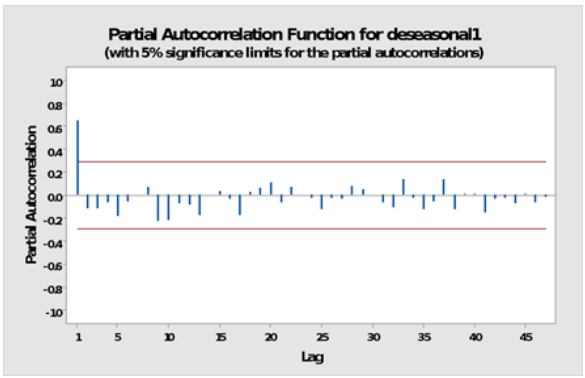



Fig. 6. PACF of first seasonal differenced series

According to Figures 5 and 6, it can be understood that the differenced series is stationary and in the non-seasonal area, both ACF and PACF cut off at lag 1. Therefore, the tentative model is identified as SARIMA (1,0,1) $(0,1,0)_4$. Based on this tentative model, three more sub tentative models were also identified as follows.

SARIMA (2,0,1) $(0,1,0)_4$

SARIMA (2,0,2) $(0,1,0)_4$

SARIMA (1,0,2) $(0,1,0)_4$

Having calculated parameter estimates, Box-Pierce (Ljung-Box) Chi-Square statistic, and correlation matrix and checked the normality of the residuals, it could be concluded that one model is adequate to forecast the profitability to a certain extent.

The model along with its MAPE and AIC values are shown in Table 8.

Table 8. Accuracy Measures for Forecast

| Model | MAPE value (%) | Accuracy | AIC |
|---|---|---|---|
| SARIMA $(1,0,0)$ $(0,1,0)_4$ | 35.6906 | 64.3094 | 23.1682 |

Based on the moderate accuracy it would be possible to use SARIMA (1,0,0) $(0,1,0)_4$ as a useful model to forecast the profitability.

## V. CONCLUSION

This study investigated the relationship between market share and profitability of licensed commercial banks in Sri Lanka. From the statistical analysis, it was found that market share significantly contributes to improve the performance of banks. Multiple linear regression proved that there is a positive relationship between profitability and market share. Analysis of Variance was conducted for testing the strength of the model, which concluded that the independent variables (DC and LC) significantly predict the dependent variable. Results further showed that, if the deposits made by customers are increased by 1 unit, the profitability will rise by 6.1% and if the loans granted to customers are increased by 1 unit, the profitability will increase by 10.7%.

Finally, it could be concluded that amount of loans granted to customers (LC) is most effective to influence the profitability of licensed commercial banks in Sri Lanka. The results of this study supported both theoretical and empirical evidence of prior studies as the loans granted to bank customers accounted for increasing the profitability of licensed commercial banks in Sri Lanka. According to the Time series Analysis, SARIMA $(1,0,0)\ (0,1,0)_4$ was found to be a reasonably accurate alternative model to forecast the profitability (Profits After Corporate Tax) of licensed commercial banks with moderate accuracy. These findings would be useful for the Licensed commercial banks to review their strategies for optimizing profitability based on strengthening the loan and deposit customer profiles.

## REFERENCES

[1] Buzzell, R. D., Gale, B. T., & Sultan, R. G. M. (1975). Market ShareA Key to Profitibility. Harvard Business Review, 53(1), 1–9

[2] Etale, L. M., Bingilar, P. F., & Ifurueze, M. S. (2016). Market Share and Profitability Relationship: A Study of the Banking Sector in Nigeria. International Journal of Business, Economics and Management, 3(8), 103–112. https://doi.org/10.18488/journal.62/2016.3.8/62.8.103.112

[3] Fidanoski, F., Choudhry, M., Davidovic, M., & Sergi, B. (2016). What does affect profitability of banks in Croatia Competitiveness Review: An International Business Journal. Competitiveness Review: An International Business Journal.

[4] Genchev, E. (2012). Effects of Market Share on the Bank's Profitability. Review of Applied Socio-Economic Research, 3(1), 87–94.

[5] Hirindukawshala, K. (2017). The Factors Effecting on Bank Profitability. International Journal of Scientific and Research Publications, 7(2), 212. www.ijsrp.org

[6] Lee, C. C., & Hsieh, M. F. (2013). The impact of bank capital on profitability and risk in Asian banking. Journal of International Money and Finance, 32(1), 251–281. https://doi.org/10.1016/j.jimonfin.2012.04.013

[7] McGahan, A. M., & Porter, M. E. (1997). How much does industry matter, really?. Strategic management journal, 18(S1), 15-30.

[8] Rumelt, R. P. (1991). How much does industry matter? Strategic Management Journal, 12(3), 167–185. https://doi.org/10.1002/smj.4250120302

[9] Trujillo-Ponce, A. (2013). What determines the profitability of banks? Evidence from Spain. Accounting and Finance, 53(2), 561–586. https://doi.org/10.1111/j.1467-629X.2011.00466.x

[10] Woo, C. Y. (1981). Market Share Leadership--Does It Always Pay Off? Academy of Management Proceedings, 1981(1), 7–11. https://doi.org/10.5465/ambpp.1981.4976490.

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index No: ES-14-35*

# Drone Technology for Rice Agriculture at the Fertilizer Spraying Process

Akarshani Amarasinghe (1st Author)
*University of Colombo School of Computing*
*University of Colombo*
*Colombo, Sri Lanka*
m.c.a.amarasinghe@gmail.com

Lakshman Jayaratne (2nd Author)
*University of Colombo School of Computing*
*University of Colombo*
Colombo, Sri Lanka
klj@ucsc.cmb.ac.lk

Viraj B. Wijesuriya (3rd Author)
*Department of Computer Science*
*University of Oxford*
Oxford, United Kingdom
viraj.wijesuriya@cs.ox.ac.uk

***Abstract*— This paper presents a drone simulator to spray fertilizer for rice farmland with minimum human intervention. The proposed drone simulator is capable of suggesting the optimal path to spray fertilizer, indicating the current altitude and the current battery level of the drone at flight. The obtained results at the evaluation stage show that the proposed path planning algorithm outputs the optimal path for given farmland with minimum execution time. This solution will cherish the use of drone technology for rice agriculture while supporting the economic growth of a country.**

***Keywords - drone simulator, q-learning, path planning algorithm***

## I. INTRODUCTION

Rice cultivation is one of the main income generating methods in most if the developing countries. Due to increasing demand for rice, farmers need to produce a higher quantity of rice every year. Fertilizing crops is the most difficult part of crop production.

A drone has considerably larger popularity rather than other unmanned equivalents due to its smaller size and flexibility. Also, drone technology has been successfully applied to agriculture through soil and field analysis, planting, crop spraying, crop monitoring, irrigation, and health assessment [1].

Using individual drones for spraying fertilizer is not a novel approach. Operating a drone in a suitable manner can minimize excessive fertilizer usage. Using an autonomous drone to spray fertilizer has several advantages over manual fertilizer spraying process as it allows to cover a larger area within a small time, which ultimately leads to low cost by reducing labor and time.

Through this research, we address the question of "How to develop a drone simulator for spraying fertilizer in a rice farmland?". This solution proposes a path planning algorithm for an autonomous drone to spray fertilizer appropriately to arable areas with minimum human intervention [11, 12]. The proposed solution has 4 modules: the 2D Grid Maker module, the Path Planning Algorithm module, the Altitude Generation Service module and the Battery Monitoring Service module. Here, when the farmer inputs the relevant details of the farmland that he/she wants to spray fertilizer, the proposed solution outputs the 2D grid of the farmland using the 2D Grid Maker module. Then the Path Planning Algorithm module outputs the optimal path for the generated 2D grid and the Altitude Generation Service module represents the current altitude of the drone when it flies along the proposed optimal path. Afterwards, the Battery Monitoring Service module monitors the current battery level of the drone in flight.

When using the proposed solution, the farmer only needs to identify the safe and unsafe areas of the farmland using its aerial image and input these details to the system. The farmer does not have to worry about the spraying process. He/she can monitor the overall spraying process through a mobile application while the drone is executing its task. Moreover, this solution can spray fertilizer to multiple arable areas with minimum human intervention.

## II. OBJECTIVES

Manually spraying fertilizer takes a considerable amount of time and manpower. Hence, using an autonomous drone for the fertilizer spraying process can cover large farmland within a minimum time while reducing human intervention. This approach can diminish the extravagant cost for the labourers while saving time and money of the farmers.

Since the proposed drone simulator is for a fully customized autonomous drone, it sprays only the approved amount of fertilizer. Hence, it helps to reduce fertilizer wastage, health and environmental issues and save money spent on excessive amounts of fertilizer.

Most of the available path planning drone solutions are available only to achieve a single target from a starting location [2]. However, our research solution is capable of visiting and covering multiple arable areas (multiple targets) in a single drone flight. As mentioned in [3], some

successful path planning methods are only for unknown environments. However, the proposed solution is for a fully observable environment. A key feature here in is the ability of the drone to avoid unsafe areas which can harm the drone flight [4]. Interestingly, the proposed drone simulator will be a better platform to the fertilizer industry since it can regulate the fertilizer usage, reduce fertilizer over usage, save farmer's time, speed up the fertilizer spraying process and minimize the human intervention at the spraying process.

## III. METHODOLOGY

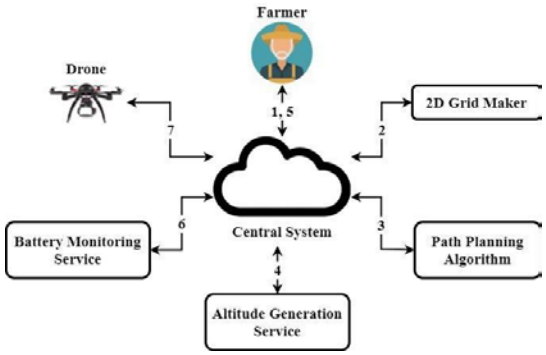



Fig. 1. The overview of the proposed solution

Fig. 1 illustrates the overview of the proposed solution, and the stepwise details are as follows:

*A.* There is a central system to manage drone operations and a mobile application for the farmer to communicate with the central system. The farmer has an aerial image of the farmland (Fig. 2(a)), and he/she identifies the safe and unsafe areas of the farmland and sends these details to the central system. Here, the safe areas are the areas with crops and the unsafe areas are traps, forbidden areas (dams, ponds, wastelands, streams, and private properties), and dead zones.

*B.* The central system sends the farmer's input to the 2D Grid Maker module. It generates the 2D grid based on the input as illustrated in Fig. 2(b). Here, the purple cell is the initial distribution (the starting location of the drone flight), the green cells are the safe areas, the red cells are the unsafe areas, and blue cells are the obstacle-free areas. The 2D grid details are passed to the central system.

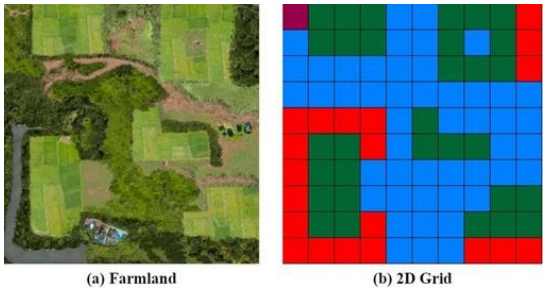



Fig. 2. The aerial image and the 2D grid of the farmland

*C.* The central system passes the obtained 2D grid details to the Path Planning Algorithm module. It outputs the optimal path as illustrated in Fig. 3. Here, the yellow line shows the proposed path for the drone flight, which starts from the initial distribution and covers all the safe areas in a single visit. A white dot in the middle of each safe area represents the fertilizer spraying locations. The drone needs to hover in-place at the location, which is represented by the white dot, spray the approved amount of fertilizer, and move forward along the proposed path.

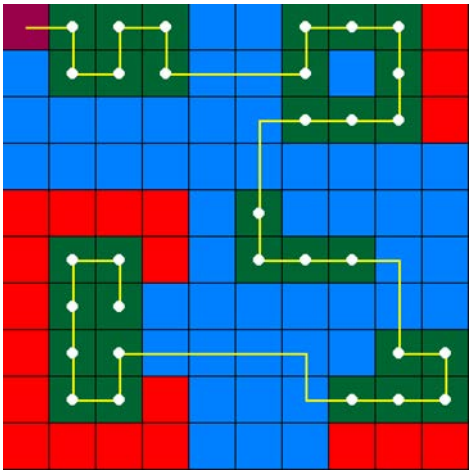


Fig. 3. The 2D grid with the optimal path

The steps of the Path Planning Algorithm module are as follows:

i. Create the transition matrix. Here, it contains the transitions from safe areas to obstacle-free areas and from obstacle-free areas to safe areas. We avoid the transitions to and from unsafe areas to speed up the learning and training process.

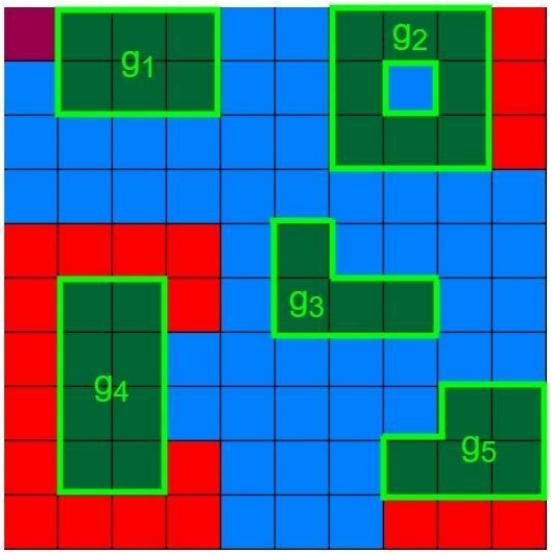



Fig. 4. The goal clusters

ii. Merge successive safe areas into a single goal cluster (Fig. 4). Otherwise, there are a lot of safe areas to cover by a single path from the initial distribution and it will take more time to give the final output.

iii. Acquire the relevant Linear Temporal Logic (LTL) [5] formula to cover all the goal clusters as represented by Equation (1).

$$\Diamond g1 \wedge \Diamond g2 \wedge \Diamond g3 \wedge \Diamond g4 \wedge \Diamond g5 \quad (1)$$

iv. Obtain the relevant Deterministic Büchi Automata (DBA) [5] to satisfy the obtained LTL formula.

v. Update the transition matrix according to the newly formed goal clusters.

vi. Apply the Q learning algorithm [6] to the updated transition matrix. The reason behind that is to identify a single optimal path from the initial distribution to multiple goal clusters.

vii. However, the Q learning algorithm can only approach the goal clusters and it cannot visit the safe areas inside a goal cluster. To solve that, we apply the Hamiltonian Cycle [7] for each goal cluster while the Q learning process is ongoing.

viii. Obtain the optimal path for the relevant 2D grid as shown in Fig. 3 and then send the path details to the central system.

*D.* According to the obtained path, the central system directs the Altitude Generation Service module to create the relevant altitude graph. Using this module, the farmer can monitor the flying altitude of the drone at the spraying process, in real time. This module helps to fly the drone above the ground level without colliding on the ground.

We identify the altitude of the proposed path according to the geo-locations with the help of Google Earth [8]. The farmer can monitor the current altitude of the drone using the altitude curve illustrated in Fig. 5 while the drone is in flight via the mobile application. Here, the x-axis and the y- axis represent geo-location (latitude and longitude) and altitude in meters of the flying path respectively. Also, the green curve shows the true elevation function (the altitude curve of the farmland) and the red curve shows the drone altitude function (the flying altitude of the drone). To obtain the drone altitude function, we apply a Polynomial Regression [9]. It helps the drone to fly a little bit higher than the ground level without colliding on the ground. Moreover, the purple dot in Fig. 5 shows the current altitude value of the drone. The farmer can observe this dot moving when the drone moves along the proposed path. However, the movement of the drone is shown only until the end of the proposed path. If the battery dies before finishing the spraying process, the system shows the movement of the drone only until that time. At the moment, the proposed solution does not have an option to identify the path from the stopping location of the drone to the initial distribution.

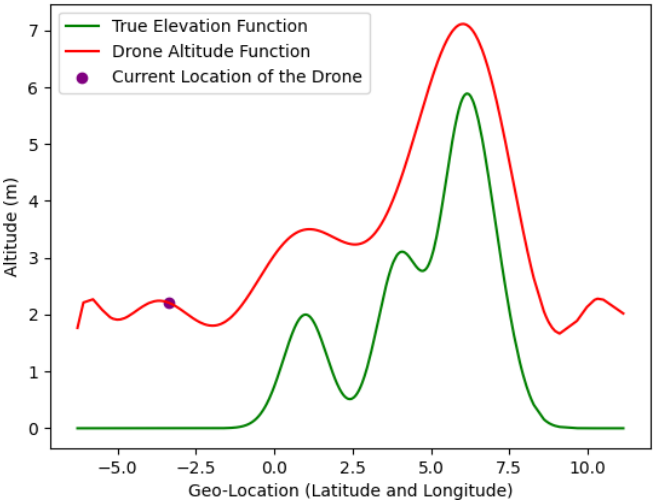



Fig. 5. A sample altitude curve

*E.* Afterwards, the central system informs the farmer, that the drone is ready to fly. The farmer can command the central system to fly the drone to spray fertilizer.

*F.* If the farmer wants to monitor the real-time battery level of the drone at its flight, he/she can enable the Battery Monitoring Service module through the mobile application. This module functions in real-time.

Fig. 6 illustrates a battery draining curve that we have used. Here, the x-axis represents the time in minutes and the y-axis represents the battery life (the battery level) as a percentage. Also, the green curve and the red curve show the true battery draining function and the 10% level of the battery respectively. We interpolate [10] the first 10 data points of the battery reading of the drone and extrapolate [10] the future battery readings based on the interpolated values.

Fig. 7 shows the battery reading at 34 minutes and the remaining cells (number of hops) are 9. The number of hops shows the remaining cell count that the drone can fly with the current battery level. The blue dotted line illustrates the current battery value reading and the purple dotted line shows the predicted battery values based on the extrapolation. If the battery level becomes 10%, the drone stops the flight, and the system stops its execution. Currently, we do not have an option to plan the path or follow any safety steps at the battery level at 10%.

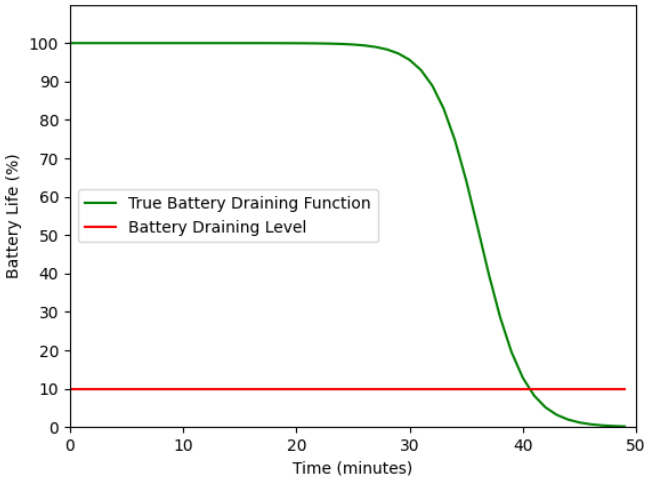



Fig. 6. The battery draining curve

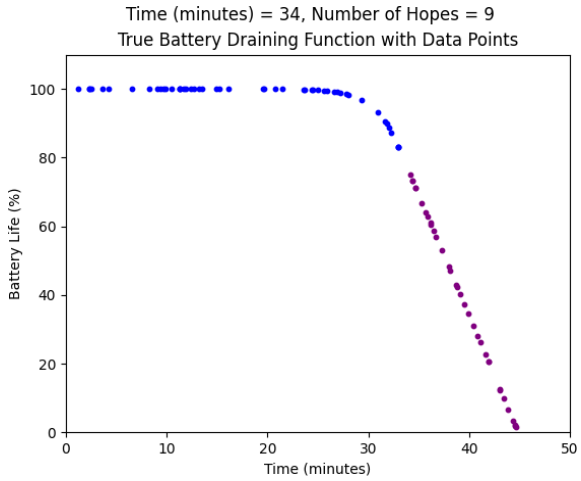



Fig. 7. Real-time battery reading

*G.* Finally, the central system commands the drone to spray fertilizer according to the proposed path. The farmer can view the spraying process in real-time via the mobile application. Here, the drone has a GPS sensor for geo-location identification, a General Packet Radio Service (GPRS) sensor to communicate via internet with the central system, a spray tank to spray fertilizer and a charging port to recharge its battery.

## IV. RESULTS AND DISCUSSION

The overall evaluation was carried out in 2 phases with the purpose of evaluating accuracy and speed. The evaluation phases are described as follows:

### *A. Suitability of the Proposed Algorithm*

The execution speed of our Path Planning Algorithm (PPA) was evaluated with the Without Clustering approach (WC). WC has similar steps as PPA except for steps 3(ii) and 3(vii) in Section III. In WC, successive safe areas were not merged into a single goal cluster, due to each safe area being considered as a separate goal and there was no need of using step 3(ii) in Section III. In PPA, the Hamiltonian Cycle was applied only to visit all the safe areas inside a goal cluster. However, in WC there are no goal clusters and it is useless to use step 3(vii) in Section III for a single goal (a safe area).

As illustrated in Fig. 8, PPA and WC for 4 test cases were evaluated. According to Fig. 8, (1)a and (1)b illustrates the obtained path for test case 1 using PPA and WC respectively. Other figures in Fig. 8 are similar to that. The relevant simulation results for each test case are represented by Table 1. It shows the goal count, the DBA states count, the path length and the execution time (in seconds) for each test case when using PPA and WC.

According to Table 1, we can see the goal count is relatively smaller in PPA. The reason behind this is the proposed algorithm merging the successive safe areas into a single goal cluster at step 3(ii) in Section III. As a result that, the total number of DBA states count is also decreased. Now the address space of the updated transition matrix is also relatively smaller than WC's updated transition matrix.

As all the safe areas has to be visited in a single goal cluster, the Hamiltonian Cycle was used for PPA. However, in WC, there aren't any goal clusters and it is useless to apply the Hamiltonian Cycle while the Q learning process is ongoing. According to that, WC should speed up the training process, since there is no step 3(vii) in Section III with step 3(vi) in Section III. Nevertheless, the execution time of PPA is relatively trivial. The reason behind that is step 3(ii) in Section III of PPA enables the capability of covering the successive safe areas as a single goal cluster and it enables the Hamiltonian Cycle to visit inside it. Unfortunately, WC does not have such a mechanism. Initially, it considers all safe areas as separate goals. As a result that, it has a higher number of DBA states. The increasing DBA states count also increases the address space in the updated transition matrix. So the state space becomes relatively larger and step 3(ii) in Section III has to train and learn based on this larger transition matrix. Therefore, WC has a higher execution time than PPA. Since WC has to visit every goal separately, it cannot traverse to a certain safe area through the successive safe areas. This procedure is the reason to increase the path lengths in WC relatively.

When carefully checking Fig. 8 and Table 1, Figure (3)b and its relevant results were not available. At the testing time, it took more than 5 hours for the execution using WC. Since WC increases the goal count and the DBA states count relatively, the transition matrix becomes complex. Furthermore, comparatively test case 3 has an enclosed environment to achieve safe areas. These aforementioned reasons cause to take more time for training using WC. Therefore, there is no output for Fig. 8 3(b).

According to the obtained simulation results, it can be concluded that PPA outputs a shorter path with a minimum

execution time since merging successive safe areas into a single goal cluster and utilizing the Hamiltonian Cycle. The results prove that the proposed path planning algorithm is more suitable for the given context.

Table 1. The simulation results with path lengths and execution times

| Test Case Number | Goal Count | | DBA States Count | | Path Length | | Execution Time (s) | |
|---|---|---|---|---|---|---|---|---|
| | PPA | WC | PPA | WC | PPA | WC | PPA | WC |
| (1) | 1 | 3 | 1 | 3 | 4 | 7 | 4.35 | 5.89 |
| (2) | 1 | 4 | 1 | 4 | 6 | 12 | 3.84 | 28.59 |
| (3) | 2 | 5 | 2 | 5 | 8 | - | 6.72 | - |
| (4) | 3 | 5 | 3 | 5 | 9 | 12 | 5.51 | 57.34 |

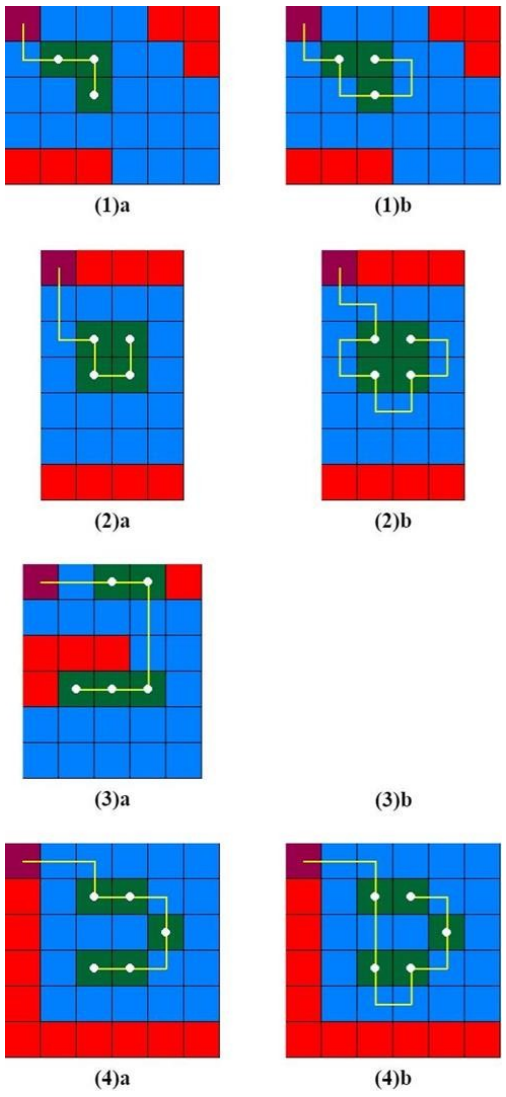



Fig. 8. The outcooms of the two evaluation approaches.

### *B. Accuracy*

In this phase we have checked whether the proposed path planning algorithm can output the optimal path for different 2D grids as illustrated in Fig. 9. We executed the proposed solution 10 times for every 2D grid in Fig. 9 and obtained results as shown in Table 2.

According to Table 2, Grid (a) has the lowest accuracy. The main reason behind that is that the goal count and DBA states count are relatively higher than Grid (b), and there are comparatively a lot of obstacle-free areas between the initial distribution and the goals clusters. When there are a lot of obstacle-free areas between the initial distribution and the goals clusters, it is very hard to find an optimal path. Since Grid (c) has many obstacle-free areas between the initial distribution and the goal clusters, it also has lower accuracy. Grid (b) has the highest accuracy, as there are fewer obstacle-free areas between the initial distribution and the goal clusters and it is easy the find a path to cover all the goal clusters from the initial distribution. Furthermore, the other reason for the highest accuracy of Grid (b) is its goal count and DBA states count is relatively lower than the other 2D grids. However, the proposed algorithm outputs the shortest path for all the grids at least 7 times out of 10 executions. As an average, we got an 86.67% accuracy for all the 2D grids in Fig. 9. According to these results, we can state that the proposed path planning algorithm outputs the optimal path for a given 2D grid.

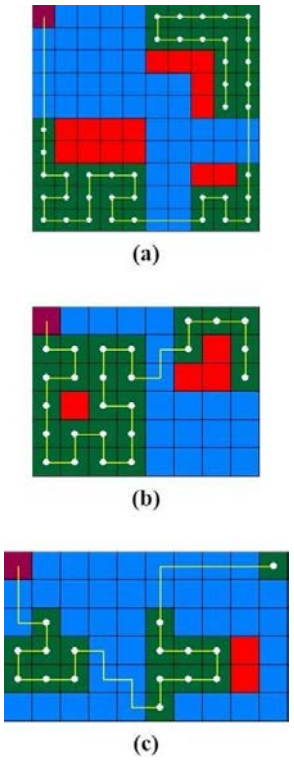



Fig. 9. The test results for different 2D grids

The obtained results from the 2 evaluation phases show the proposed path planning solution is capable to output the optimal path for a given farmland with a minimum execution time. Nevertheless, there can be some accuracy issues of the proposed path, if there are more obstacle-free areas among the initial distribution and the goal clusters. However, the proposed path planning algorithm has a few limitations: the goal cluster must have a Hamiltonian path and its shape must be a square or a rectangle.

Table 2. The path accuracy for different 2D grids

| Grid Name | Goal Count | DBA States Count | Accuracy (%) |
|---|---|---|---|
| (a) | 3 | 3 | 70.00 |
| (b) | 2 | 2 | 100.00 |
| (c) | 3 | 3 | 90.00 |
| **Overall Accuracy (%)** | | | 86.67 |

## V. CONCLUSION

In this paper, we have proposed a drone simulator with an optimal path planning algorithm for an autonomous drone. The main purpose for developing this, is reducing the human intervention at the fertilizer spraying process in rice agriculture. The proposed solution has the 2D Grid Maker module to convert an aerial image of the farmland that we want to spray fertilizer to the relevant 2D grid image, the Path Planning Algorithm module to generate the optimal path to cover all safe areas while avoiding unsafe areas from the initial distribution, the Altitude Generation Service module to identify the current flying altitude of the drone without colliding on the ground and the Battery Monitoring Service module to monitor the real-time battery draining of the drone. Since the Altitude Generation Service module and the Battery Monitoring Service module are at the development stage, we do not have the path planning solution and altitude changes after the drone battery dies.

We have evaluated the proposed path planning algorithm under 2 phases in Section IV: suitability-wise and accuracy- wise. According to the obtained results, it proves the algorithm outputs the optimal path within a minimum execution time while avoiding unsafe areas. The proposed path planning algorithm produces the optimal path from the initial distribution to cover all the safe areas in a single visit. However, there are several limitations in the proposed solution: there could be accuracy issues when there are more obstacle-free areas among the initial distribution and the goal clusters, there must be at least a single Hamiltonian path to visit inside the goal cluster and the goal cluster's shape should be a square or a rectangle.

Through the proposed drone simulator, we can cover large farmland within a minimum time and minimum human intervention. This helps to reduce the fertilizer wastage during the spraying process and speed up the spraying process. Also, through reducing the overuse of fertilizer the farmer can save money without spending extravagantly on fertilizer. Using drone technology for rice agriculture can save the farmer's money from expensive labourers as well.

In the future, we expect to complete the development of the ongoing modules and tackle the problems while the drone battery dies. Furthermore, we hope to test the drone simulator with an autonomous drone. We anticipate improving the proposed drone simulator not only for rice but also for other agricultural crops. Moreover, we look forward to selling or renting out the developed drone solution to the farmers as a package with a concessional price.

This research helps to promote the drone technology in developing countries and regulate the fertilizer usage. Through the regulation of fertilizer, the farmer can increase not only the quantity of the harvest, but also the quality. It will provide a great support to the economical development of the country.

## REFERENCES

[1] "Six Ways Drones Are Revolutionizing Agriculture - MIT Technology Review." https://www.technologyreview.com/s/601935/six-ways-drones-are- revolutionizing-agriculture/ (accessed Apr. 01, 2018).

[2] M. Jun and R. D'Andrea, "Path Planning for Unmanned Aerial Vehicles in Uncertain and Adversarial Environments," Springer, Boston, MA, 2003, pp. 95–110.

[3] M. Hasanbeig, A. Abate, and D. Kroening, "Logically-constrained neural fitted Q-iteration," in Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS, 2019, vol. 4, pp. 2012–2014.

[4] L. V. Nguyen, B. Hoxha, T. T. Johnson, and G. Fainekos, "Mission Planning for Multiple Vehicles with Temporal Specifications using UxAS," in IFAC-PapersOnLine, Jan. 2018, vol. 51, no. 16, pp. 67– 72, doi: 10.1016/j.ifacol.2018.08.012.

[5] C. Baier and J. Katoen, Principles of Model Checking (Representation and Mind Series). The MIT Press, 2008.

[6] P. R. Montague, "Reinforcement Learning: An Introduction, by Sutton, R.S. and Barto, A.G.," Trends Cogn. Sci., vol. 3, no. 9, p. 360, 1999, doi: 10.1016/s1364-6613(99)01331-5.

[7] R. J. Gould, "Advances on the Hamiltonian Problem-A Survey," 2002.

[8] "Google Earth." https://earth.google.com/web (accessed Jun. 23, 2021).

[9] Y.-W. Chang, C.-J. Hsieh, K.-W. Chang, M. Ringgaard, and C.-J. Lin, "Training and Testing Low-degree Polynomial Data Mappings via Linear SVM," 2010. Accessed: Jun. 23, 2021. [Online]. Available: http://jmlr.org/papers/v11/chang10a.html.

[10] R. Kress, Numerical Analysis, vol. 181. New York, NY: Springer New York, 1998.

[11] A. Amarasinghe, V. B. Wijesuriya, D. Ganepola, and L. Jayaratne, "A swarm of crop spraying drones solution for optimising safe pesticide usage in arable lands: Poster abstract," in SenSys 2019 - Proceedings of the 17th Conference on Embedded Networked Sensor Systems, Nov. 2019, pp. 410–411, doi: 10.1145/3356250.3361948.

[12] A. Amarasinghe, V. B. Wijesuriya, and L. Jayaratne, "A Path Planning Algorithm for an Autonomous Drone Against the Overuse of Pesticides," in 2021 10th International Conference on Information and Automation for Sustainability (ICIAfS), Aug. 2021, pp. 446–451, doi: 10.1109/ICIAFS52090.2021.9606118.

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index No: ES-15-38*

# Solar Power as a Sustainable Energy Source and Readiness Level in Sri Lanka: A Review

Lasny Manjitha (1st Author)
*University of Kelaniya*
Kelaniya, Sri Lanka
lasni91@yahoo.com

Amila Munasinghe (2nd Author)
*University of Kelaniya*
Kelaniya, Sri Lanka
amila@kln.ac.lk

***Abstract* — Rapid demand for energy has influenced engineers and scientists' investigation on renewable sustainable energy solutions. Although a wide variety of sustainable natural energy resources are available, usableness depends on technical feasibility and government intervention. Solar energy is a widely accepted solution for electricity generation due to its unique availability. With promotion of the solar power as a means for Sustainable Development Goal (SDG7) of the United Nations, this study is motivated to review information on solar power as a renewable energy source and to examine how Sri Lanka is ready for such move to relieve the economic burden from imported energy. The paper reveals government interventions in solar power initiatives and challenges towards energy sustainability and provides a future outlook.**

***Keywords — sustainability, solar power, Sri Lanka, energy***

## I. INTRODUCTION

With the complex lifestyles of humans and the development of economies, the energy demand is rapidly increasing globally. As the World Energy Council predicts [1], global electricity demand will peak in 2030. Increased use of conventional energy sources like fossil fuel over the years has caused several hazards such as water and air pollution, global warming, acid rain, land degradation, and all of these hazards transforming towards a burning global issue- Climate change. Fossil fuel is only one of the three main categories of global energy resources (oil, gas, coal) and global attention is focused on two other categories of energy resources; nuclear energy and renewable energy (wind, solar, geothermal, airpower, biomass, hydrogen, ocean) [2]. With the United Nations focus on Climate change through Framework Convention on Climate Change (UNFCCC), leading international bodies such as International Renewable Energy Agency (IRENA) promote the widespread adoption and sustainable use of renewable energy sources that includes bioenergy, geothermal, hydropower, ocean, solar and wind energy as sustainable & clean energy sources. United Nations Sustainable Development Goals (SDGs) set in 2015 “Ensure access to affordable, reliable, sustainable and modern energy for all” as the seventh goal (SDG7), has emphasized the sustainability in energy usage to achieve by 2030 [4].

Advancement in technology, proper regulation of policies [5] and enhancement of efficiency through research and development (R&D) [6] are some of the options to conserve the existing energy sources while the innovation of environment-friendly renewable resources. The sustainability concept concerns preserving natural resources for future generations to meet their own needs while fulfilling the present needs [7]. Among different renewable energy sources, solar power has become significant as it uses the energy of the sun which is giving a large amount of energy due to the fusion of hydrogen nuclei in the form of solar radiations.

This article focuses on answering the research question of "What do we know about solar power and how the government of Sri Lanka supports its adoption”. This study was motivated due to the scant literature that reveals the readiness and levels of adoption of sustainable energy sources by Sri Lanka. Accordingly, this paper aims to produce an overview of solar power as a means for reaching the Sustainable Development Goal (SDG7) of the United Nations and provide information on its readiness and adoption in Sri Lanka.

## II. SOLAR POWER AS A SUSTAINABLE ENERGY SOURCE

The Joint Sustainable Development Goals (SDG) Fund is one of the initiatives that transform policies and stimulate strategic investments that require meeting SDGs' needs [8]. Among 17 goals, the world is progressing towards Goal 7 with encouraging signs that energy is becoming more sustainable and widely available. As improvement proofs, access to electricity in poorer countries shows making impressive gains in energy efficiency in the electricity sector. As the name implies, it should have more focused attention to improve access to clean and safe energy sources. Its Energy Progress Report provides a global dashboard to register progress on energy access, energy efficiency and renewable energy. This assesses the progress of each country using these three main aspects how far they stand ahead for SDG. As at the website of United Nations, Department of Economic and Social Affairs, Sustainable

Development [9] still 13% of the global population lacks their access to modern electricity while 3 billion people are relying on wood, coal, charcoal or animal waste for cooking and heating. Usage of these conventional energy sources has become the dominant contributor to climate change with 60% of total global greenhouse gas emissions. According to the data found in 2016, the share of renewable energy sources has increased at its fastest rate since 2012 by the application of hydropower, wind, and solar [10].

Sustainability is not a new concept to the scientific community and it has become the main component of recent national policies, strategies and development plans of many countries [11].

Applications of renewable energy sources have shown not only environmental benefits but also economic development as well. Especially renewable energy development has become a trend globally to meet the need for electricity due to the rapid decrease of stored sources and negative environmental impacts [12]. According to David Pimentel, 2001 hydropower contributes significantly to world energy by providing 6.5% of the supply [13]. Biogas is alternative energy source for power generation and heat generation for household use [14]. Wind power has provided energy to pump water and to run mills and other machines, although it is limited by the availability of sites with sufficient wind [15].

Solar power is one of the easily accessible sources of energy that the sun emits at high-speed radiation which is available on large areas of Earth [16]. As reported usage of solar power has no harmful effects on the ecosystem and it is widely applicable efficiently in urban and rural areas, in industrial and domestic conditions due to its easy accessibility and usability. The applications of solar energy can be broadly classified into two categories such as thermal energy systems which convert solar energy into thermal energy and photovoltaic energy system which converts the solar energy into electrical energy.

In order to meet the SDG target by 2030 while increasing the share of renewable energy in the world energy mix with double the global rate of improvement in energy efficiency, it is suggested to focus on research for clean energy and technology, energy efficiency means and models and use of cleaner fossil-fuel technology.

## III. READINESS FOR SOLAR ENERGY – A GLOBAL PERSPECTIVE

Restricted access to energy sources is one of the major crises in the world. Energy supply through noncarbon energy renewable sources is considered as the most promising solution for this problem [17]. By considering the economic and environmental benefits, the world has initiated various programs to replace fossil fuels with renewable energy resources.

Research and development of solar energy have existed for many decades in finding new technologies such as solar heating, solar drying, solar thermo-electricity generation and photovoltaic solar energy conversion. According to Siripala, 2014 [18] there are main two categories of solar research. One is focusing on inventing efficient solar cells that can be made at a very low cost, and others have focused on developing very high-efficiency cells even though they are expensive in manufacturing. Among all these researches, one of the main challenges is invention of cost effective efficient solar cell comparing with fossil fuels. Although it is challenging majority of the universities and research institutes worldwide have prominent research programs on solar energy converting devices as they have understood damage done to the environment by using fossil fuels will be irrecoverable [19]. All these efforts require immense amounts of money and resources in order to find a suitable solution to the energy problem [20].

Another aspect is although there are a large number of research programs conducted throughout the world, outcomes with practical applications are expected. The main reason is that lack of knowledge and accumulation of knowledge generated by a large number of projects will be essential to develop a final product. Fundamental research on the related subject areas plays a very important role in developing final applicable devices. For instance, the discovery of the transistor by Bardeen, Shockley and Brattain in 1947 [21] has influenced understanding the fundamental principles of semiconductors.

Victoria *et al*., 2021 [22] has mentioned that solar photovoltaics (PV) which are installed has cumulative capacity at the end of 2019 accounted for more than 600 GW and PVs could become our majority global energy source and an improved representation of this technology with new innovations of solar cell and system levels could highly contribute world energy consumption. Countries like Africa has tend to apply these technologies as PV appear to offer the possibility of 'green' electricity for rural areas specially in Sub Saharan Africa [23].

Solar hybrid energy system is a combination of energy sources with different characteristics and an energy storage system. It is important when solar energy itself is not enough in generating constant power for consumption [24]. As an energy source with no carbon or greenhouse gas emission photovoltaic power generation systems has used as an alternative not only for cooking and day-to-day household activities but also in highway service areas especially in China as they are small in size, lightweight and its small occupational space. Usage as an energy source in industrial activities is important as this equipment has a long useful life period ranging from 20 to 50 years with zero pollution emissions as PV consumes no fuel without any noise. It power is generated without water supply is also important as PV systems can be applied for areas which are uninhabited [25].

Developing an environmental-friendly photovoltaic solar panel that can shut high temperature radiation within a panel box was done by Kohei *et al*., 2020 which they include a panel having a decompression-boiling heat collector that can absorb heat from the PV module. There

a single PV system could keep the heat radiation temperature from the PV solar panel surface around 45°C even in summer and able to provide hot water at least 60°C by controlling the flow rate. As mentioned, this system has high efficiency in the utilization of solar energy with a small environmental heat load [26]. Chemical storage of solar energy such as hydrogen production by water splitting using sunlight has been done by Boretti *et al.*, 2021 [27], is an active research area. Among all, photovoltaic solar cell research and development have a prominent place. As a result, semiconductor silicon-based solar cell devices are at the forefront today, while the emerging technology is on the development of thin-film solar cells using other semiconductor materials. China leads in solar power generation globally and India joins as a prominent user in Asia. Israel is one of the pioneer countries that use solar energy for domestic usage and for solar plants. There they have equipped all most all the new buildings to be with solar collectors for water heating while high rates of domestic solar water heating worldwide (about 75% of households).

Companies in the country have solar technologies such as large-scale solar-powered electricity generating plant installed in Southern California's Mojave Desert [28].

Increasing the technology of solar energy systems has been experimented by a number of experiments all over the world by in ceasing their capacities with minimal environmental impacts.

## IV. READINESS FOR SOLAR ENERGY IN SRI LANKA

With reference to SDG7, by 2030 it has expected to upgrade infrastructure and technology to enable a move towards sustainable energy sources. Sri Lanka has faced a major economic burden due to the dependence on imported energy resources and their high costs. This section reviews solar power initiatives and institutional and legal framework governing such initiatives in Sri Lanka. Solar energy, one of the clean and free renewable energy resources abundantly available in Sri Lanka that could be a major cornerstone of the future renewable energy systems in the country as the country is located within the equatorial belt. As mapped by the National Renewable Energy Laborite, USA solar radiation varies from 4.0 – 4.5 kWh/m2 /day [29]. According to the data on the CEB website (on 12th November 2021), small power producers contribute only 0.22 GWh countrys' daily electricity generation (Fig. 1.) [30].

When considered the institutional framework for Solar power in Sri Lanka, at present, the Energy and Business Development (MOPEBD), Ceylon Electricity Board (CEB) which contributes to the country's energy generation, transmission and distribution, Sri Lanka Sustainable Energy Authority (SLSEA) which plays a dual role as regulator and facilitator, Ministry of Power & Renewable Energy which involves in the planning of energy policies. Public Utilities Commission of Sri Lanka which is involved in sector governance are leading institutions for the development of this concept of solar energy of Sri Lanka [31]. Other stakeholder institutions and the Ministry of Finance & Planning, Ministry of Environment is helping in developing environmental policies while Central Environmental Authority involves in preparing environmental regulations [32].

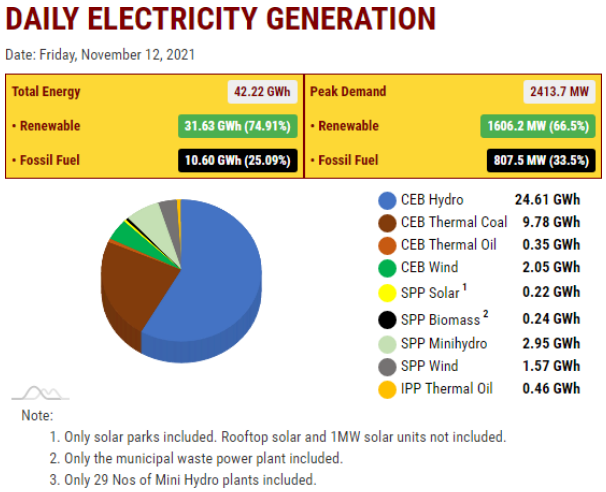



Fig. 1. Daily electricity generation of Sri Lanka (as at 12.11.2021)

Except for all these institutions, Universities and Research and Development (R&D) Institutions of the country are important in R&D and providing training services.Other organizations such as Non-Governmental Organizations and technology suppliers of Sri Lanka are also important in this aspect. CEB has taken the priority to boost access to renewable, clean or green energy power generation in Sri Lanka. As the initial step, the CEB has introduced the cost attract investments for the new technology [33]. As the solar power industry has become matured, CEB gradually introduced the competitive bidding process in line with the Sri Lanka Electricity Act. As found in December 2020, 414 MW of Solar power capacity has been grid-connected. A Gradual transformation from Feed-in-Tariff (FIT) scheme to a competitive bidding process has also been introduced by the CEB aimed at traffic reduction trends in the global market. CEB has already planned 7,000 ground-mounted solar power projects with an installed capacity of 75 kW to be positioned at rural and semi-urban areas of the country. The project was promoted under the theme of "Gamata Balagarayak" and aimed to attract local entrepreneurs for the project [34].

According to SLSEA, the solar resource atlas of Sri Lanka is an important addition to the existing knowledge on solar resources of Sri Lanka. The first solar atlas of Sri Lanka was prepared by the National Renewable Energy Laboratory (NREL) of the USA, in 2005 which has provided a great opportunity in exploring solar resources of the country, leading to gross estimates of solar potential.

The availability of accurate solar resource is crucial for the sustainable development of solar resources in Sri Lanka.

CEA as the key institution of preparing regulations and Environmental Impact Assessment (EIA), has contributed various national projects in this regard by evaluating its value with reference to the environment of the country. Ministry of Power, CEB and Sri Lanka Sustainable Energy Authority (SLSEA) initiated grid-scale solar PV power project, small scale distributed solar PV projects and rooftop solar PV installing commercial scale. As an example, CEA has joined with SLSEA for the EIA report of Proposed Solar Park in Siyabalanduwa, Sri March 2021 [35]. Review of the legal framework for Solar power in Sri Lanka shows that in recent past years the technology of solar energy has rapidly grown with the encouragement of government policies. Its low maintenance cost with low environmental issues makes it easy to cut down the electricity bills.

Sri Lanka Sustainable Energy Authority Act, No.35 of 2007 is one of the leading Act which was there for the establishment of the Sri Lanka Sustainable Energy Authority, to develop renewable energy resources, to declare energy development areas, to implement energy efficiency measures and conservation programs, to promote energy security and to evaluate the reliability and cost- effectiveness in energy delivery and information management [36]. Guideline for Renewable Energy Project Development and Circular: "Speeding up of the Renewable Energy Projects Development Process" is also two major regulatory framework for Sri Lankas' ground-mounted solar power projects. Supplying sustainable and affordable energy services to support socially equitable development in Sri Lanka [37]. In order to move toward energy independence and sustainable development, Sri Lanka should develop a technology mix using available indigenous energy sources such as hydro, solar, wind, biomass while reducing the usage of imported fossil fuel [38].

A review of publically available documents shows initiatives of solar power projects in the country. Rooftop solar systems have become popular with the availability of rooftop spaces in urban and rural areas [39]. The "Rooftop Solar PV Power Generation Project" has provided credit facilities by the Government of Sri Lanka (GoSL) through a loan from the Asian Development Bank (ADB). The credit line for the project is managed by the Ministry of Finance (MoF) and funds are channeled to the beneficiaries through selected Banks (Participating Financial Institutions -PFI). The technical support is provided by the Project Implementation Unit (PIU) in close collaboration with MoF, Ministry of Power, Energy and Business Development (MoPE) and Sri Lanka Sustainable Energy Authority (SLSEA) [40]. Power from the Rooftop Solar PV Installations can be integrated with the National grid through the Utility Providers in Sri Lanka, namely Ceylon Electricity Board and Lanka Electricity Company Private Limited.

Large scale development projects have been planned especially in areas such as Trincomalee, Ampara, Monaragala, Hambantota, Kurunegala and Anuradhapura [41] and in other urban cities such as Colombo. Distributed solar PV resource development has its own challenges such as institutional issues, financial issues, regulatory issues, awareness issues in a country like Sri Lanka. Research has been done by National Renewable Energy Laboratory, Colorado in 2003 on solar resource assessment for Sri Lanka and Maldives. They have applied a gridded cloud cover database at a 40-km resolution to produce updated monthly average daily total estimates of all solar resources (global horizontal, DNI, and diffuse) to input hourly or three-hourly cloud cover observations made at nine weather stations in Sri Lanka and two in the Maldives into a solar model that produces estimates of hourly solar radiation values [42]. There they have investigated the readiness of Sri Lanka for application of this technology by referring data obtained. The net-metering scheme was introduced in 2010 to serve the solar PV rooftop industry with large scale implementation across the country.

On September 6, 2016, the Government of Sri Lanka had launched an enhanced version of the rooftop solar PV program under the theme “Sooryabala Sangramaya” which means “Battle for Solar Power” [43]. This program attempts to encourage institutional users by exporting through a separate export meter without making any change to the electricity users' metering method. With the significant reduction of the cost of solar PV components, the service providers have quickly moved for large industrial customers who own large buildings with good roofs for solar PV systems.

Research on Solar energy has been going on in Sri Lanka for a considerable period, especially concentrating on solar thermal and photovoltaic systems, solar energy materials and solar water. Many universities and research institutes in Sri Lanka are involved in these research programs and among them most of the research are focused on development of low-cost thin-film solar cell devices and the search for low cost semiconductor materials for solar energy applications. Most of these research programs are funded by the National Science Foundation (NSF) and the National Research Council (NRC) [18]. Numbers of research publications are available on new concepts, methodologies, and materials which have been discovered by locals. Commercialization of these innovative products has not yet been realized from this research.

While achieving and realizing the capabilities of solar energy research in Sri Lanka, investment on solar energy research in Sri Lanka is an investment for the future. In order to get rid of the extremely high cost of fossil fuels, Sri Lanka will have to heavily depend on solar energy and other renewable energy resources for its future energy needs. Unless Sri Lanka develops the technology on its own for energy generation, in future it might have to heavily depend on foreign technology to harness the available solar energy. Therefore, continuation and further strengthening of

research and development on solar energy in Sri Lanka will be a meaningful investment for the future.

## V. CHALLENGES IN SOLAR ENERGY ADOPTION

Solar power systems face problems with weather conditions and the length of the day. The prices of the equipment in these systems are relatively high. PV system has a high capital cost compared to other conventional energy sources [37]. The surfaces of the photo panels and mirror photovoltaic should always be cleaned of different dust. Certain large areas of the solar photovoltaic station can be a challenge for cleaning.

The effectiveness of photovoltaic plants heating can be reduced due to various reasons. PV-panels may be inefficient in collecting all available solar energy due to a number of environmental factors such as intensity of the sun, cloud cover and wind speed [44].

Lack of awareness of the potential benefits of the solar industry, especially among the rural population around the world is another challenge. Competition with energy sources in the markets and toxicity of the production material of solar cells are also some significant challenges. Instability of the energy supply make it to combine the system with other source of energy. The inability to provide enough supply to meet the need is also a weakness [45]. Large PV systems can cause some undesirable environmental impacts as well.

## VI. FUTURE OF THE SOLAR ENERGY

To improve the capacity of current solar systems, new designs would need to be able to capture more light, transform light energy to electricity more efficiently and be less expensive. The cost for designing a product that can track the position of the sun accurately and consistently is an ongoing challenge but innovation on this front continues. As Emily Kerr from Harvard University, 2019 indicates, improving the performance of solar systems should target efficiency [46]. Solar systems with many layers of light- capturing material can capture more photons. As tested solar cells with four layers can capture 46% of the incoming light energy although it is too expensive. But ongoing research may be able to implement these super-efficient cells which will be cost-effective for commercialization. Major market players' involvement in investing more in research and development to increase the efficiency of this innovation is important while improving the sustainability of the resource.

## VII. CONCLUSION

Energy is inevitable for an economy of any country. Countries like Sri Lanka incurs huge costs on energy imports causing balance of payment deficits every year. This paper aimed to contribute to the literature through reviewing information on solar power as a renewable energy source and to reveal on adoption in Sri Lanka. Solar power generation demands the most renewable sources of electricity. It has several advantages compared to other forms of energy but is still challenging as a consistent and cost- effective energy source to meet the high energy demand. Several pieces of research have been done by researchers all over the world in order to make it real while overcoming current challenges. Applications of solar energy are popular even in developing countries as the country's location severely affects these solar systems' efficacy. However, improvements in solar systems and their applications will take time and are possible only if they undergo bulk manufacturing and installation. Sri Lanka, as a middle income country needs the integration of policies and strong legal framework in order to achieve the SGD Goal7 and for the effective use solar power as a sustainable energy source. This study emphasizes the need for further research studies on solar power adoption by different countries, issues, and innovations to support reaching sustainable energy goals by 2030.

## REFERENCES

[1] World Energy Scenarios Composing energy futures to 2050, World energy Council, 2013.

[2] Choifin, M., Rodli, A. F., Sari, A. K., Wahjoedi, T. and Aziz, A., A study of renewable energy and solar panel literature through bibliometric positioning during three decades, University of Nebraska – Lincoln, 2021.

[3] Sambhar, N. Internet-only Bank Industry Research Report, Growth Trends and Competitive Analysis 2021-2027: Atom Bank, WeBank (Tencent HoldingsLimited), *The Premium Sample Internet-only Bank Market Research Market Report*, 2021

[4] Saxena, A., Ramaswamy, M., Beale, J., Darcy Marciniuk, D. and Smith, P., Striving for the United Nations (UN) Sustainable Development Goals (SDGs): what will it take? *Discov Sustain 2, 20*, 2021. https://doi.org/10.1007/s43621-021-00029-8

[5] Schmid, G., The development of renewable energy power in India: which policies have been effective? *Energy Policy*, 2012; 45:317–326.

[6] Khare, V., Nema, S., Barreda, P., Status of solar-wind renewable energy in India, Renewable and Sustainable Energy Reviews. 2013; 27: 1-10.

[7] Kuhlman T., Farrington J. What is Sustainability? *Sustainability*. 2010; 2(11):3436-3448. https://doi.org/10.3390/su2113436.

[8] SDF Fund. Our Programmes. SDG Fund UN Initiative. 2019. Available from: https://www.sdgfund.org/programmes. Accessed 2021 November 10. PubMed/NCBI Google Scholar

[9] MAKE THE SDGS A REALITY, United Nations, Department of Economic and Social Affair, Sustainable Development, https://sdgs.un.org/ Accessed on 10.11.2021

[10] Owusu, P. A. and Asumadu-Sarkodie, S, A review of renewable energy sources, sustainability issues and climate change mitigation, 2016 https://doi.org/10.1080/23311916.2016.1167990

[11] Owusu, P. A., and Asumadu-Sarkodie, S., A review of renewable energy sources, sustainability issues and climate change mitigation. Cogent Engineering, 2016; 3 (1), 1-14. https://doi.org/10.1080/23311916.2016.1167990

[12] Widjaja, M., Risiko, Investment-State Dispute Settlement Dalam Pemberian Insentif Bagi Investasi Asing Di Sektor Energy Terbarukan Indonesia. UNIVERSITAS AIRLANGGA, 2020

[13] Pimentel, D., McNair, S., Janecka, J., Wightman, J. and Simmonds, C., O'Connell, C., Wong, E., Russel, L., Zern, J. and Aquino, T., Economic and environmental threats of alien plant, animal, and microbe invasions. *Agriculture, Ecosystems & Environment*, 2021; 4. 1-20.

[14] gde Haryanto, M. U., Ibadillah, A. F., Alfita, R., Aji, K. and Rizkyandi, R., Smart aquaponic system-based Internet of Things (IoT), *Journal of Physics Conference Series*, 2019.

[15] Smith D. R. and Ilyin , M. A., Wind and solar energy: Costs and value, *The American Society of Mechanical Engineers 10th Annual Wind Energy Symposium*, 1991

[16] Stanojevic, A. B., Application of photovoltaic technology in the use of solar energy, *Ann Environ Sci Toxicol*. 2021; 5(1): 094-09.

[17] Boyle, G., Renewable Energy Power: Power for a Sustainable Future. *Oxford University Press, London*, UK, 2012.

[18] Siripala, W., Solar cell development research: A Sri Lankan perspective, Article in Journal of the National Science Foundation of Sri Lanka; 2014.

[19] Hamakawa Y., Thin Film Solar cells: Next Generation Solar Cells and its Applications. *SpringerVerlag, Berlin, Germany*, 2014.

[20] Kurdgelashvili ,L., A Review of Solar Energy, Market, Economic and Policies, Policy Research Working Paper 5845. Environment and Energy Team, Development Research Group, The World Bank, Washington DC, USA 2011, Available at elibrary.worldbank.org/doi/pdf/10.1596/1813- 9450-5845

[21] Norman, J. M., Bardeen, Brattain & Shockley Invent the Point Contact Transistor, Jeremy Norman's HistoryofInformation.com, https://www.historyofinformation.com/detail.php?id=662, Accessed on 10.11.2021

[22] Victoria, M., Haegel, N., Peters, I. M. Sinton, Ron, Jäger-Waldau, A., del Cañizo, C., Breyer, C., Stocks, M., Blakers, A., Kaizuka, I., Komoto, K., and Smets, A. Solar photovoltaics is ready to power a sustainable future, Joule, 2021. Web. https://doi.org/10.1016/j.joule.2021.03.005.

[23] Rhys, G., Charles, Matthew, L. D., Peter, D., Ingrid, L. and Hallin, I. M., Sustainable energy storage for solar home systems in rural Sub-Saharan Africa – A comparative examination of lifecycle aspects of battery technologies for circular economy, with emphasis on the South African context, Energy, 2019; 166: 1207-1215, https://doi.org/10.1016/j.energy.2018.10.053.

[24] Kannan, N., Vakeesan, D., Solar energy for future world: A review. Renew Sust Energ Rev., 2016; 62: 1092-1105. Link: https://bit.ly/3xkw0CJ

[25] Qin, X., Shen, Y. and Shao, S. The Application Study in Solar Energy Technology for Highway Service Area: A Case Study of West Lushan Highway Low-Carbon Service Area in China, *International Journal of Photoenergy*. 2015; 1-8. 10.1155/2015/703603.

[26] Terashima, K., Sato, H. and Ikaga, T., Development of an environmentally friendly PV/T solar panel. Solar Energy, 2020; 199, 510-520. https://doi.org/10.1016/j.solener.2020.02.051

[27] Boretti, A., Nayfeh, J. and Al-Maaitah, A., Hydrogen Production by Solar Thermochemical water-splitting Cycle via a Beam Down Concentrator, Front. Energy Res., 2021; 9 (116) https://doi.org/10.3389/fenrg.2021.666191

[28] Statement by Ambassador Uzi Manor, Coordinator for Sustainable Development, Ministry of Foreign Affairs, Jerusalem, the Intergovernmental Preparatory Meeting 15th Commission on Sustainable Development at United Nations, New York, 28 February 2007.

[29] Wijesena, G. H. D. and Amarasinghe, Solar Energy and its Role in Sri Lanka, 2018.

[30] De Silva, K. P. S. D, Opportunities and challenges of solar energy application in energy sector of Sri Lanka, Bulletin of the Korea Photovoltaic Society 2020 (1).

[31] MoPE, "Sri Lanka Energy Sector Development Plan for a Knowledge- based Economy 2015-2025", Ministry of Power & Energy, (MoPE), Government of Sri Lanka. http://powermin.gov.lk/sinhala/wp content/uploads/2015/03/ENERGY_EMPOWERED_NATION_2015_2025

[32] National Energy Policy & Strategies for Sri Lanka (amended 2019), Ministry of Power & Energy.

[33] Central Bank of Sri Lanka Statistics Department. 2018. Economic and Social Statistics of Sri Lanka 2018

[34] Sri Lanka Energy Sector Assessment, Strategy and Road Map (2019), Asian Development Bank

[35] George, R., Renné, D.,Marion, B., Heimiller, D., and Gueymard, C. Solar Resource Assessment for Sri Lanka and Maldives, 2003.

[36] Wijesena, G. H.D. and Amarasinghe, A. R. *International Journal of Engineering Trends and Technology (IJETT,)*, 2018; 65 (3)

[37] Ugli, T. J. T. The Importance of Alternative Solar Energy Sources and the Advantages and Disadvantages of Using Solar Panels in this Process *International Journal of Engineering and Information Systems (IJEAIS)*, 2019; 3 (4): 70-79.

[38] Hernandez, R., Easter, S.B., Murphy-Mariscal, M. L., Maestre, F. T., Tavassoli M., Environmental impacts of utility-scale solar energy. *Renew Sust Energ Rev*, 2014; 29: 766-779. Link: https://bit.ly/3fso8sD

[39] Project on Electricity Sector Master Plan Study in Democratic Socialist Republic of Sri Lanka, Ministry of Power and Renewable Energy

[40] Kerr, E, The Future of Solar is bright, Blog, sitn20: celebrating 20 years, March 21, 2019.

[41] Larson, A., Future of Solar Energy Brighter Than Sun: Rapid Innovations Pave the Way, Solar, Power, Jan 24, 2020.

[42] Ugli, T. J. T. The Importance of Alternative Solar Energy Sources and the Advantages and Disadvantages of Using Solar Panels in this Process *International Journal of Engineering and Information Systems (IJEAIS)*, 2019; 3 (4): 70-79.

[43] Hernandez, R., Easter, S.B., Murphy-Mariscal, M. L., Maestre, F. T., Tavassoli M., Environmental impacts of utility-scale solar energy. *Renew Sust Energ Rev*, 2014; 29: 766-779. Link: https://bit.ly/3fso8sD

[44] Project on Electricity Sector Master Plan Study in Democratic Socialist Republic of Sri Lanka, Ministry of Power and Renewable Energy

[45] Kerr, E, The Future of Solar is bright, Blog, sitn20: celebrating 20 years, march 21, 2019

[46] Larson, A., Future of Solar Energy Brighter Than Sun: Rapid Innovations Pave the Way, Solar, Power, Jan 24, 2020.

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index No: ES-16-43*

# Waste to Value through Rice Hull to Nano-silica: A Novel Paradigm of Sustainable Agronomy

Imasha K. Anuththara (1st Author)
*Department of Chemistry*
*The Open University of Sri Lanka*
Nawala, Sri Lanka
imashakavindi29@gmail.com

Hasanka Kumara (2nd Author)
*Department of Chemistry*
*The Open University of Sri Lanka*
Nawala, Sri Lanka
kalinguarachchi195@gmail.com

Lahiru Wijenayaka (3rd Author)
*Department of Chemistry*
*The Open University of Sri Lanka*
Nawala, Sri Lanka
lawij@ou.ac.lk

Darshani Weerahewa (4tht Author)
*Department of Botany*
*The Open University of Sri Lanka*
Nawala, Sri Lanka
weerahewa@gmail.com

***Abstract* — Although Silicon (Si) in its bulk phases has long been used as a fertilizer in agriculture, specifically for those in the Poaceae family, the unique role their nanoscale counterparts may adopt in modern day agriculture remains mostly unknown. *Oryza sativa* (rice) is among the main agricultural crops in Sri Lanka. The rice plants accumulate and greatly benefit from Si, where it is deposited beneath the cuticle as cuticle-Si double layer in the form of silicic acid. Further, Si may also interact favorably with other applied nutrients, hence improving the agronomic performance, crop yield, as well as the tolerance of rice plants to abiotic and biotic stresses, thus making Si essential in sustainable rice production. The hull, or the hard protective layer of the rice grain, is mostly regarded as a bulk-scale waste produced during the post-harvest processing of rice. Nevertheless, it is notable that Si is a prominent constituent in rice hull. Hence, there have been recent attempts to utilize rice hull as a precursor for producing Si nanoparticles. However, such approaches, still being in their infancy, requires further systematic investigations to optimize the nanoparticle preparation, in terms of the ensuing nanoparticle properties (i.e., size, morphology, and porosity etc.) as well as the process economies themselves. Hence, in this study, a novel, facile, efficient, and scalable strategy was systematically developed for the preparation of Si nanoparticles from waste rice hull. Notably, pre-processing conditions, synthetic parameters, and the chemical usages were optimized to allow scalability and sustainability to ensue. Initial characterization of the nanoparticles synthesized under aptly optimized conditions indicated the ability of this novel approach to synthesize nanoparticles of smaller dimension that could be stably suspended in aqueous media for prolonged durations without any signs of instability. Further investigations are ongoing in terms of creating and controlling the porosity of these nanomaterials, which will further enhance their carrier properties and thus the applicability as an effective nano-fertilizer. Overall, the knowledge imparted by this study will be significant in intensifying agricultural practices, specifically in many developing parts of the world, where rice remains to be a prominent crop, thus indicating a novel paradigm of sustainable agronomy.**

***Keywords — Rice hull, Nano-silica, Synthesis, Stability, Nano-fertilizer***

## I. INTRODUCTION

Mineral nutrition is well known to contribute towards improving the quality and yield of cereal crops. Silicon; a quasi-essential element abundant in many soils has proven to be effective in producing (1) increased resistance to pathogens and insects, (2) increased resistance to strong wind and rain, (3) reduced lodging, (4) enhanced uptake of K, P and Ca, as well as (5) the alleviation of P deficiency, drought stress, salt stress, and metal toxicity (Mn, Cd, As, Al, Zn, Fe etc).[1]–[3] However, the ability of bulk Si to produce the above effects would depend on their transport through the plant. Hence, it is intuitive to believe that if Si in nanoscale is incorporated into fertilizer applications, the transport process will be momentously enhanced, thereby greatly amplifying the above benefits.[4]

Additionally, with the recent advent of nanotechnology, it is well-known that matter may adopt unique properties, specifically, some of which might even be significantly different from those in the bulk-scale, due to the distinctive quantum confinement effects that prevails in the nanometer regime.[5], [6] As a result, some materials have been found to produce novel properties, specifically resulting from their nanoscale dimensions. Scientific research in this regard is mostly ongoing, and of note, some recent findings have proven that nanoparticles of Si or nano-silica are a powerful

and versatile tool that is greatly underutilized in contemporary agriculture, where the anticipated unique benefits may include (1) enhanced regulation of plant growth, (2) enhanced resistance to environmental stress, (3) significant improvement in soil quality (mainly the water holding capacity), and (4) the ability to serve as potential carries for fertilizers, pesticides, herbicides, or even DNA and proteins.[1]–[3]

*Oryza sativa* (rice) is among the main agricultural crops in Sri Lanka. The rice plants accumulate and greatly benefit from Si, where it is deposited beneath the cuticle as cuticle-Si double layer in the form of silicic acid.[7] Further, Si may interact favorably with other applied nutrients, hence improving the agronomic performance, crop yield, as well as the tolerance of rice plants to abiotic and biotic stresses, thus making Si essential in sustainable rice production.[8] However, the effect of this imperative nutrient in rice cultivation at the nanoscale has not been explored thus far.

The hull, or the hard protective layer of the rice grain, is mostly regarded as a bulk-scale waste produced during the post-harvest processing of rice. Nevertheless, it is notable that Si is a prominent constituent in rice hull.[7] Hence, there have been recent attempts to utilize rice hull as a precursor for the production of Si nanoparticles.[9]–[11] However, such approaches, still being in their infancy, requires further systematic investigations to optimize the nanoparticle preparation, in terms of the ensuing nanoparticle properties (i.e. size, morphology, and porosity etc) as well as the process economies themselves. Notably, if novel, facile, efficient, and scalable strategies could be developed for the preparation of Si nanoparticles from the waste rice hull, and the ensuing shape size and concentration effect of the same in rice cultivation is systematically evaluated, the knowledge acquired would be greatly significant in intensifying agricultural practices, specifically in many developing parts of the world, where rice remains to be a prominent crop.

## II. OBJECTIVES

The focus of this study was to develop a novel method to effectively prepare Si nanoparticles using rice hull as the Si precursor, and to assess the effect of the same in rice cultivation. Notably, the initial investigations have revealed that nano silica synthesized under aptly optimized conditions indicated the ability of this novel approach to synthesize nanoparticles of smaller dimension that could be stably suspended in aqueous media for prolonged durations without any signs of instability.

## III. METHODOLOGY

*Synthesis of nano-silica*: Synthesis of nano-silica was conducted using a modified procedure based on methods previously reported in literature.[12]–[14] First, the rice hull (RH) was cleaned with tap water and deionized water, dried in an oven. Then, it was soaked in 1 M hydrochloric acid (HCl) for 2 hours, washed with deionized water and dried in an oven again. The as-prepared RH was sintered in a muffle furnace at a standard temperature and time to obtain rice hull ash (RHA). Then, RHA was ground into a powder, which was washed again with HCl by mixing for 2 hours followed by overnight soaking. Then, the solution was centrifuged, and the supernatant was replaced with deionized water until any excess acid is removed, as confirmed with the help of pH papers. Then, a standard solution of sodium hydroxide (NaOH) was added into the RHA and the solution was heated while stirring until the ash was completely dissolved, followed by hot filtration. Once the filtrate was cooled to room temperature, a known volume of deionized water was added to the solution and finally a standard solution of sulfuric acid ($H_2SO_4$) was added dropwise under continuous stirring until the solution was completely neutralized and a turbidity was observed. The as-synthesized particles were purified via repeated centrifugation followed by resuspension in deionized water. The cloudy suspensions obtained were transferred into reagent bottles and stored until further use. Above procedure was repeated multiple times to ensure repeatability of the procedure.

*Optimization of synthetic parameters*: In optimizing the scalability and the economic viability of the procedure developed for synthesizing nano-silica, (1) the need for acid washing, (2) reusability of the acid wash solution, and (3) the stability of the nanoparticles based on diluent volume used during the synthesis were optimized here. The above parameters were decided based on the cost-factors that would likely be associated in commercializing the scientific findings and the stability of the nanoparticles which is deemed to alter the behavior of the nanomaterials upon eventual application. The effect of acid washing on the synthesis of nano-silica was evaluated by conducting the above outlined synthesis procedure with and without the acid washing step, while all other conditions and parameters were held constant. The reusability of the acid solution for washing RH was evaluated by repeating the above outlined synthesis procedure with RH washed with fresh and used acid solutions, while all other conditions and parameters were held constant. The effect of diluent volume on the ensuing stability of the nano-silica was evaluated by repeating the above outlined synthesis procedure with different volumes of deionized water (50, 100, and 150 mL) being used as the diluent, while all other conditions and parameters were held constant.

## IV. RESULTS AND DISCUSSION

All synthesis and analysis conducted here were done using RH obtained during the processing of a single batch of samba rice, while the pre-cleaned RH was mixed to ensure homogeneity within the sample. It was observed that the synthetic approach developed here could produce a fine aqueous suspension of nano-silica which could be stably suspended for prolonged durations. Hence, in optimizing the synthetic parameters further, the effect of acid washing, reusability of the acid solution for washing, and effect of diluent volume were systematically evaluated to understand their role in the synthesis process and/or in dictating the properties of the ensuing nanoparticles.

Figure 1 (A) shows an image of RH after acid-washing and drying. To evaluate the effect of acid washing, multiple samples prepared in this manner was sintered after being treated with concentrated acid, while control samples were used as is. Images of RHA obtained for the acid washed and

unwashed RH are indicated in Figure 1 (B) and (C) respectively. As can be clearly seen, the acid washing was essential in (1) allowing complete combustion taking place during the sintering process, while (2) assisting the disintegration of the RH structure during sintering; both of which are important for the efficient pre-synthetic processing of RH. Clearly, for the controls where acid-washing was not conducted, incomplete combustion was observed as clearly visible by the appearance of the sintered samples, while the ensuing disintegration of the RH structure during the sintering was also minimal. Hence, it was inferred that acid-washing is essential to remove any impurities that are likely present in RH thus leading to incomplete combustion, while the same assists the efficient disintegration of the carbonaceous backbone of RH during sintering.


Fig. 1. Images of (A) acid-washed and dried RH, and RHA (i.e. sintered RH) (B) with and (C) without acid washing

As per the overall objective of developing a novel nano-fertilizer with silica, process scalability, economic viability, as well as environmental sustainability were considered as important considerations. Hence, to assess the reusability of the acid solution for washing RH experiments were conducted to observe RHA washed with fresh and reused acid solutions of a standard solution. Leaching of phytochemicals and/or impurities into the acid solution was clearly apparent after even after the first round of washing, hence indicating that the efficiency of the acid solution would decrease with successive use.

However, as observed by the images of RHA washed using fresh, twice, and thrice used acid solutions in Figure 2 (A), (B), and (C) respectively, efficient sintering was apparent with acid solutions reused up to three rounds of washing. All samples washed with acid solutions reused beyond this level indicated inefficient processing as was clearly apparent by the hue and the presence of blackish RH resides which were apparent post-sintering. Hence, in optimizing the reusability of the acid solution for washing RH during the pre-processing, it was inferred that the acid washing could be done with fresh, once-used, or even twice-used acid solutions without hampering the anticipated pre-processing of RH thereby assisting the economic viability as well as environmental sustainability of the synthetic approach developed herein.

Eventually, the RHA obtained in this manner was used in the nano-silica synthesis where the formation of nanoparticles was apparent by the appearance of a cloudy turbidity in the otherwise clear solution during the final neutralization step (Figure 2 (D)). However, it was apparent that the formation of nanoparticles was accompanied by the observation of a thick gel in the synthetic medium which is likely to hinder any successive nucleation and growth of nanoparticles that may ensue. Hence, to minimize this, dilution of the synthetic medium was conducted using deionized water right before the final acid addition step, thus allowing the solution viscosity to be remain intact during the nanoparticle formation. As anticipated, dilution of the synthetic medium led to the formation of nanoparticles as evidenced by the formation of a cloudy suspension upon neutralization, without any gel formation.


Fig. 2. Images of RHA washed using (A) fresh, (B) twice, and (C) thrice used acid, and (D) the milky suspension of ensuing nano-silica

Hence, to optimize the diluent (i.e. deionized water) volume used during the synthesis, the synthesis procedure was repeated with different volumes of deionized water being used as the diluent, while all other conditions and parameters were held constant. This indicated that the synthetic medium could be efficiently diluted with deionized water thereby diminishing any gel formation, while still leading to the formation of nanoparticles. However, further dilution of the medium indicated that nucleation may be hindered due to decreased precursor concentration hence dampening the nanoparticle formation. Thus, the experiment conducted here allowed to identify the optimum diluent volume to be used for the nano-silica synthesis.

The nanoparticles prepared in this manner indicated significantly stable dispersion stability where no signs of precipitation were apparent for prolonged durations of up to few months. Further, in the absence of microscopic characterization, the appearance of the suspension as well as the prolonged suspension stability serves as evidence for the smaller dimensions of the nanoparticles formed via this novel synthetic approach. Nevertheless, such dispersion stability is deemed important in utilizing the nanoparticles in fertilizer applications where the suspensions could be directly applied onto a desired crop, while the stability and the smaller dimensions of the nanoparticles will increase the uptake as well as the nutrient delivery that the nanoparticle fertilizer is anticipated to perform.

Further detailed characterization of the nanoparticles developed here, and the investigation of their fertilizing efficacy are ongoing. However, with the initial observations reported here, the study will likely result in a process by which rice hull could be efficiently transformed into an effective and sustainable fertilizer starting with waste rice hull, indicating competent use of technology in facilitating sustainable agricultural practices.

## V. CONCLUSION

A novel, facile, and controlled synthetic approach was developed for the RH mediated preparation of nano-silica. The salient process parameters such as acid washing, reusability of acid was solution, as well as diluent volume were aptly investigated and systematically optimized via carefully designed investigations. Overall, the nano-silica

obtained via the optimized synthetic pathway indicated exceptional dispersion stability, serving as initial evidence for the smaller dimensions as well as the appreciable stability of the nanoparticles produced here. Thus, the initial investigations conducted suggests that the process developed here is facile, scalable, and controllable in producing nano-silica that may efficiently be incorporated as a novel nano-fertilizer for the many crops that essentially requires silicon as a notable micronutrient. Investigations are ongoing in terms of creating and controlling the porosity of these nanomaterials, which will further enhance their carrier properties and thus the applicability as an effective nano-fertilizer. Overall, the knowledge imparted here is significant in intensifying agricultural practices, specifically in many developing parts of the world, where rice remains to be a prominent crop, while unfortunately RH is considered only as mere agricultural waste, thus indicating a novel paradigm of sustainable agronomy.

## REFERENCES

[1] H. Jayawardana, H. L. D. Weerahewa, and M. Saparamadu, "Effect of root or foliar application of soluble silicon on plant growth, fruit quality and anthracnose development of capsicum," 2014.

[2] D. Weerahewa and K. Somapala, "Role of Silicon on Enhancing Disease Resistance in Tropical Fruits and Vegetables: A Review," 2016.

[3] D. Weerahewa and D. David, "Effect of silicon and potassium on tomato anthracnose and on the postharvest quality of tomato fruit (Lycopersicon esculentum Mill.)," *J. Natl. Sci. Found. Sri Lanka*, vol. 43, p. 3, 2015.

[4] S. Mor, C. K. Manchanda, S. K. Kansal, and K. Ravindra, "Nanosilica extraction from processed agricultural residue using green technology," *J. Clean. Prod.*, vol. 143, pp. 1284–1290, 2017.

[5] H. M. El-Bendary and A. A. El-Helaly, "First record nanotechnology in agricultural: Silica nano-particles a potential
new insecticide for pest control," *App. Sci. Rep.*, vol. 4, no. 3, pp. 241–246, 2013.

[6] A. Mullen, "Expectations from nano in agriculture," *Nat. Nanotechnol.*, vol. 14, no. 6, pp. 515–516, 2019.

[7] S. Liu and C. Ho, "Study in the nature of silicon in rice hull. I. solubility of the silicon part," *J. Chinese Chem. Soc.*, vol. 6, no. 2, pp. 141–153, 1960.

[8] L. E. Datnoff, G. H. Snyder, and G. H. Korndörfer, *Silicon in agriculture*. Elsevier, 2001.

[9] M. M. Lu, D. M. De Silva, E. Peralta, A. Fajardo, and M. Peralta, "Effects of nanosilica powder from rice hull ash on seed germination of tomato (Lycopersicon esculentum)," *Philipp. e-Journal Appl. Res. Dev.*, vol. 5, pp. 11–22, 2015.

[10] W. Jansomboon, K. Boonmaloet, S. Sukaros, and P. Prapainainar, "Rice hull micro and nanosilica: synthesis and characterization," in *Key Engineering Materials*, 2017, vol. 718, pp. 77–80.

[11] N. R. Dhineshbabu, R. Vettumperumal, and N. V Selvam, "Study the Optical Properties of Amorphous Silica Nanoparticles from Rice Hull," *Adv. Sci. Eng. Med.*, vol. 8, no. 9, pp. 730–735, 2016.

[12] I. A. Rahman and V. Padavettan, "Synthesis of Silica Nanoparticles by Sol-Gel: Size-Dependent Properties, Surface Modification, and Applications in Silica-Polymer Nanocomposites—A Review," *J. Nanomater.*, vol. 2012, p. 132424, 2012, doi: 10.1155/2012/132424.

[13] U. Zulfiqar, T. Subhani, and S. Wilayat Husain, "Synthesis of silica nanoparticles from sodium silicate under alkaline conditions," *J. Sol-Gel Sci. Technol.*, vol. 77, no. 3, pp. 753–758, 2016, doi: 10.1007/s10971-015-3950-7.

[14] M. I. Vlasenkova, E. S. Dolinina, and E. V Parfenyuk, "Preparation of mesoporous silica microparticles by sol-gel/emulsion route for protein release.," *Pharm. Dev. Technol.*, vol. 24, no. 2, pp. 243–252, Feb. 2019, doi: 10.1080/10837450.2018.1457051.

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index No: ES-17-47*

# Modelling and Forecasting the Volatility of Daily Exchange Rate Using GARCH Model

Madhubhashitha WKR (1st Author)
*Department of Mathematical Sciences*
*Wayamba University of Sri Lanka*
Kuliyapitiya, Sri Lanka
rameshmadhubhashitha01@gmail.com

Aponsu GMLM (2nd Author)
*Department of Mathematical Sciences*
*Wayamba University of Sri Lanka*
Kuliyapitiya, Sri Lanka
lakshana@wyb.ac.lk

***Abstract*— The volatility of daily exchange rate is a significant economic indicator for open economic countries like Sri Lanka, when considering the international level trade. Therefore, it is very important to be aware of the behavior of future fluctuations of the exchange rate volatility, even though accurate volatility forecasting is challenging. The purpose of this study is to model and forecast the volatility of the US Dollar against the Sri Lankan Rupee (USD/LKR) daily exchange rate. The daily USD/LKR exchange rate data from 1st January 2015 to 30th April 2021 were used in this study and it was found that the exchange rate was continuously increasing throughout the period and stationarity of the daily exchange rate return series was confirmed by the Augmented Dickey-Fuller (ADF) test. Volatility of the daily exchange rate returns were modeled using Generalized Auto-Regressive Conditional Heteroscedastic (GARCH) models. ARMA(2,1) was found to be the most preferable model for the mean equation, and GARCH(1,1) was identified to be best to capture the conditional volatility of the residuals of the ARMA(2,1) model. In addition, Lagrange Multiplier (LM) test clearly showed that ARCH effect no longer exists in the residuals of ARMA(2,1) - GARCH(1,1) model and the Sign Bias test indicated that there was no asymmetry effect in residuals. Therefore ARMA(2,1) - GARCH(1,1) was identified as the best model to forecast the USD/LKR daily exchange rate with Mean Squared Error (MSE), Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) equal to 0.047695, 0.218391, and 0.047695 respectively. The findings of this study can be used in decision and policy making stages to minimize the risk associated with exchange rate volatility.**

***Keywords — Exchange rate, GARCH, Volatility***

## I. INTRODUCTION

From the earliest stages of globalization, many countries tended to international trade where the transactions mostly depend on the USD and the foreign exchange market has been expanded with time. The foreign exchange rate is the relation of the value of two currencies which is regarded as the relative value. Volatility of the exchange rate is indicated by the relation of the demand and supply of the currency in the foreign exchange market. Therefore exchange rate is a crucial economic factor to continue the economic stability of all countries and its fluctuation significantly impacts the macroeconomic variables like gap between net exports and imports, interest rate, debt payments installment of the countries, inflation, unemployment, etc. However the global economic crisis leads to the volatility of the exchange rate fluctuations which mainly affect the economic activities of the developing open economy countries like Sri Lanka than other developed countries.

Modeling and forecasting of exchange rate volatility are great interest to decision and policy making at different levels to minimize the risk associated with exchange rate volatility. Most researchers have studied the non-constant variance in the financial time series like exchange rate, and uncertainties in prices and returns. Therefore, the behavior of exchange rate volatility is explained using time series econometric models. A study on analyzing daily closing prices for four currency pairs viz. Euro, Pound, Swiss Franc and yen against the US Dollar [1], using data from January 2002 to December 2011 forecasted the volatility of each foreign exchange rate by using univariate GARCH models. It was compared the in-sample forecasts from symmetric and asymmetric GARCH models with the implied volatility derived from currency options for each dollar parity. A low volatility was noted during the period of 2002 to 2007, and there was high volatility during the rest of the period. Finally, it was concluded that the volatility forecasts significantly outperform the GARCH models in both low and high volatility periods.

Univariate nonlinear time series models such as Auto-Regressive Conditional Heteroscedastic (ARCH) and GARCH models have been used to examine the behavior of the daily (TZS/USD) exchange rate [2] by using data from January 2009 to July 2015. In the same study, Exponential GARCH (EGARCH) model also has been used to capture the asymmetry in volatility clustering and the leverage effect in exchange rate. In a study of identifying the volatility of the RM/Sterling exchange rate by using GARCH models [3], the maximum likelihood method and several goodness-of-fit statistics were used to estimate the parameters of these models and diagnosed the performance of the within-sample

estimation. Also the accuracy of the out-of-sample and one-step-ahead forecasts were evaluated using mean squared error. In this study the stationary GARCH-M outperforms other GARCH models in out-of-sample and one-step-ahead forecasting. When using random walk model as the naive benchmark, all GARCH models outperform this model in forecasting the volatility of the RM/Sterling exchange rate.

To understand the theoretical and empirical performance of the GARCH class of models and to exploit the potential gains in modeling conditional variance, a study was carried out to forecast monthly exchange rates of Pakistan for the period ranging from July 1981 to May 2010[4]. ARMA, ARCH, GARCH, and EGARCH models were used in this study and found that GARCH (1, 2) was best to remove the persistence in volatility while EGARCH (1, 2) successfully overcome the leverage effect in the exchange rate returns. When it comes to the context of Sri Lanka, the behavior of the daily USD/LKR exchange rate was examined by using several ARCH and GARCH models[5], and found that ARMA(1,1) – ARCH(6) is the best fitted model to forecast the daily USD/LKR Exchange rate volatility.

Even though above researchers have studied about forecasting exchange rates in various contexts, with the dynamic nature of the volatility of exchange rate, it is vital to study further and identify better forecasting models to forecast exchange rate and it will assist to reduce the risk associated with investment and policy making decisions. Therefore, this research is focused on identifying a forecasting model to forecast USD/LKR daily exchange rate volatility using GARCH model.

## II. OBJECTIVES

The main objective of this research is to analyze the behavior of the daily USD/LKR exchange rate over the study period and to construct forecasting (ARCH family) model to forecast the exchange rate while evaluating the forecasting performance of the best fitted model.

## III. METHODOLOGY

The data used in this study consists of daily exchange rates of USD/LKR from 1st January 2015 to 30th April 2021 obtained from the Central Bank of Sri Lanka which comprises 1520 observations. Exchange rate return was calculated from the following Equation (1):

$$R_t = ln\left(\frac{X_t}{X_{t-1}}\right) = ln\ ln\ (X_t)\ - ln(X_{t-1}) \quad (1)$$

, where $R_t$ is the daily return of USD/LKR exchange rate at time $t$ and $X_t$ is the daily USD/LKR exchange rate at time $t$.

The data set was divided into two sets namely training and testing where the training set ranging from 1st January 2015 to 31th May 2019 (1062 observations) and the testing set ranging from 3th June 2019 to 30th April 2021 (458 observations).

First, the stationarity of the return series of daily exchange rate was checked by using Augmented Dickey-Fuller (ADF) test [6]. Box Jenkin's methodology was applied next to model the conditional mean equation of the exchange rate return series [7]. The combination of AR(p), MA(q) and ARMA (p, q) models were used and latter is expressed in Equation (2):

$$R_t = c + \sum_{i=1}^{p} \varphi_i R_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \quad (2)$$

, where $\varphi_i$ and $\theta_i$ are parameters, $c$ is a constant, and $\varepsilon_t$ is the residual at time t. The best model for the mean equation was selected by comparing the accuracy measures such as Akaike Information Criterion (AIC), Bayesian Information Criteria (BIC), Durbin-Watson (DW) statistic and R-square value of identified adequate models. The residual diagnostics of the mean equation were tested by correlogram Q-statistic and Jarque-Bera test. The presence of ARCH effect of the residuals of the conditional mean model was tested by using the Lagrange Multiplier (ARCH-LM) test [8]. The conditional variance of the residuals of the conditional mean model was modeled by using two time-varying volatility models namely ARCH and GARCH. Engle (1982) [9] developed ARCH model to capture the volatility of the financial time series. The general ARCH(q) model assumes normally distributed residuals where the current conditional variance depends on the first q previous squared innovations as follows:

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 \quad (3)$$

, where $\omega>0, \alpha_i \geq 0$ *and* $i>0$.

Bollerslev (1986) [10] developed the GARCH model and the general GARCH (p, q) model assumes normally distributed residuals where the current conditional variance depends on the first p previous conditional variances and the first q previous squared innovations as follows:

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \quad (4)$$

, where $\omega>0, \alpha_i \geq 0, \beta_j>0$ *and* $i, j>0$.

The best model for the variance equation was selected by comparing the accuracy measures such as AIC, BIC, DW statistic and R-square value of identified adequate models. Then, the ARCH-LM test was conducted to test the ARCH effect of the residuals of selected conditional variance model and also the serial correlations, normality and asymmetric effect of residuals were tested. Finally, the forecasting performance of the best fitted model was evaluated by using MSE, RMSE and MAE.

## IV. RESULTS AND DISCUSSION

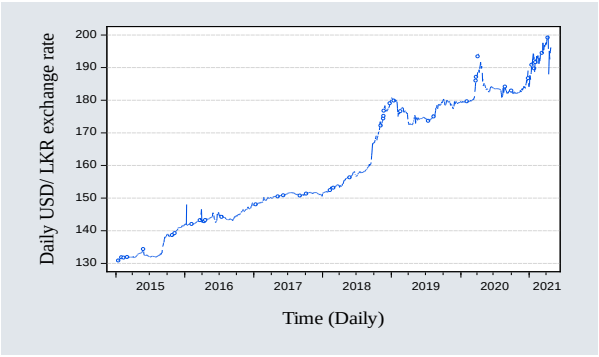



Fig. 1. The time series plot of the daily USD/ LKR exchange rate

Figure 1 shows the time series plot of the daily USD/LKR exchange rate which gives a comprehensive knowledge about the daily USD/LKR exchange rate variation during the past six years. From 2015 to 2021, the exchange rate has been increased dramatically from Rs. 130 to Rs.200. The global economic crisis due to Covid-19 pandemic is the main reason for the unexpected increment of the exchange rate during last year.

A forecasting model can be fitted only for a stationary time series. The stationarity of the exchange rate return series was tested by using ADF test based on the following hypothesis and the results are shown in Table 1.

$H_0$: return series is not stationary
$H_1$: return series is stationary

Table 1. ADF test

| | | T - Statistic | Prob.* |
|---|---|---|---|
| Augmented Dickey – Fuller test statistic | | -40.93181 | 0.0000 |
| Test critical values: | 1% level | -3.434454 | |
| | 5% level | -2.863240 | |
| | 10% level | -2.567723 | |

According to Table 1, the null hypothesis is rejected as the p-value of the ADF test statistic (0.000 < 0.05) is significant at 5% significance level and it can be concluded that the daily exchange rate return series is stationary. Then, several tentative conditional mean models were identified by using Box Jenkin's methodology and the best mean equation was identified by comparing AIC, BIC, DW and R-square values. The accuracy measures of the adequate mean models are shown in Table 2.

Table 2. Accuracy measures of adequate mean models

| Model | AIC | BIC | $R^2$ | DW |
|---|---|---|---|---|
| AR(1) | -9.026841 | -9.017471 | 0.015821 | 1.984824 |
| MA(1) | -9.023793 | -9.014423 | 0.012811 | 2.023697 |
| AR(2) | -9.027237 | -9.013182 | 0.01807 | 2.003955 |
| MA(2) | -9.028238 | -9.014183 | 0.019054 | 1.995333 |
| ARMA(1,2) | -9.037846 | -9.019106 | 0.030457 | 2.017056 |
| ARMA(2,1) | -9.039877 | -9.021138 | 0.032424 | 1.996509 |

Table 2 indicates the results of the accuracy measures of the identified tentative conditional mean models. By comparing the AIC, BIC, R-square and DW statistics of each model, ARMA (2,1) model was selected as the best conditional mean model and the corresponding equation is as follows.

$$R_t = 0.791211\, R_{t-1} + 0.190915\, R_{t-1} - 0.962305\, \varepsilon_{t-1} \quad (5)$$

The serial correlation of the residuals of mean equation was tested using the correlogram Q-statistic and the results are shown in Figure 2.

$H_0$: There is no autocorrelation in the residuals
$H_1$: There is autocorrelation in the residuals

| Autocorrelation | Partial Correlation | | AC | PAC | Q-Stat | Prob |
|---|---|---|---|---|---|---|
| | | 1 | -0.001 | -0.001 | 0.0004 | |
| | | 2 | 0.008 | 0.008 | 0.0634 | |
| | | 3 | 0.003 | 0.003 | 0.0705 | |
| | | 4 | 0.014 | 0.014 | 0.2805 | 0.596 |
| | | 5 | -0.007 | -0.007 | 0.3380 | 0.845 |
| | | 6 | -0.028 | -0.029 | 1.1969 | 0.754 |
| | | 7 | -0.026 | -0.026 | 1.9354 | 0.748 |
| | | 8 | 0.003 | 0.004 | 1.9481 | 0.856 |
| | | 9 | -0.001 | -0.000 | 1.9488 | 0.924 |
| | | 10 | -0.050 | -0.050 | 4.6733 | 0.700 |
| | | 11 | -0.057 | -0.057 | 8.1116 | 0.423 |
| | | 12 | -0.032 | -0.033 | 9.1786 | 0.421 |
| | | 13 | 0.026 | 0.026 | 9.9091 | 0.449 |
| | | 14 | 0.019 | 0.021 | 10.312 | 0.503 |
| | | 15 | 0.024 | 0.025 | 10.938 | 0.534 |
| | | 16 | 0.031 | 0.028 | 11.944 | 0.532 |
| | | 17 | 0.027 | 0.020 | 12.706 | 0.550 |
| | | 18 | -0.005 | -0.009 | 12.729 | 0.623 |
| | | 19 | -0.001 | -0.002 | 12.731 | 0.692 |
| | | 20 | -0.004 | -0.005 | 12.753 | 0.753 |
| | | 21 | -0.005 | -0.009 | 12.783 | 0.804 |
| | | 22 | 0.021 | 0.018 | 13.267 | 0.825 |
| | | 23 | 0.011 | 0.013 | 13.407 | 0.859 |
| | | 24 | -0.003 | 0.002 | 13.415 | 0.893 |
| | | 25 | -0.010 | -0.004 | 13.521 | 0.918 |

Fig. 2. The Correlagram Q-Statistic of ARMA (2, 1) model

All p values are greater than 0.05 according to Figure 2 and null hypothesis is not rejected at 5% level of significance confirming the non-existence of serial correlation in residuals of the best fitted conditional mean model. The normality of the residuals was tested based on the following hypothesis [11].

$H_0$: Residuals are normally distributed

$H_1$: Residuals are not normally distributed

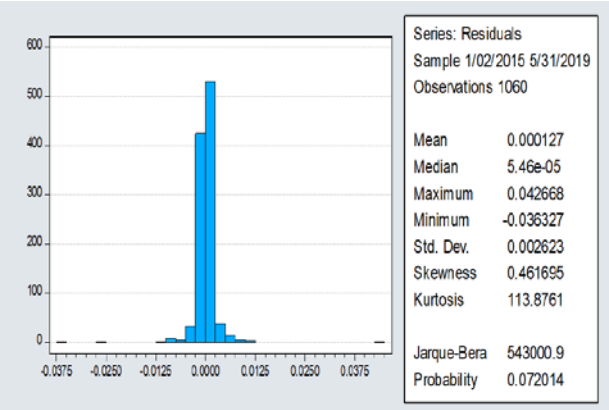



Fig. 3. The histogram of the residuals of ARMA (2, 1) model

According to Figure 3, it is clear that the Jarque-Bera test statistic is not significant (0.07 > 0.05) at 5% significance level and null hypothesis is not rejected, confirming the normality of residuals. ARCH-LM test was used to check the homoscedasticity of residuals using the following hypothesis and based on the results of Table 3, it can be concluded that there is an ARCH effect in the residuals as the p-value of the ARCH - LM test (0.0000 < 0.05) is significant at 5% significance level and the null hypothesis is rejected.

$H_0$: There is no ARCH effect in the residuals
$H_1$: There is an ARCH effect in the residuals

Table 3. LM test on ARMA (2, 1) residuals

| | | | |
|---|---|---|---|
| F - statistic | 268.7309 | Prob.F(1,1057) | 0.0000 |
| Obs* R-squared | 214.6635 | Prob. Chi-Square(1) | 0.0000 |

As the residuals of the mean equation are heteroscedastic, ARCH and GARCH models were used to estimate the variance equation. The best model for the variance equation was selected by comparing the accuracy measures such as AIC, BIC, DW statistic and R-square value of adequate models and the results are summarized in Table 4.

Table 4. Accuracy measures of adequate models

| Model | AIC | BIC | $R^2$ | DW |
|---|---|---|---|---|
| ARMA(2,1) - ARCH(1) | -9.750007 | -9.726613 | -1.652382 | 3.36569 |
| ARMA(2,1) - GARCH(1,1) | -10.05404 | -10.02589 | 0.012868 | 1.98407 |

Based on the results of Table 4, the accuracy measures of tentative models were compared and ARMA(2,1) – GARCH (1,1) model was selected as the most preferable model for the variance equation.

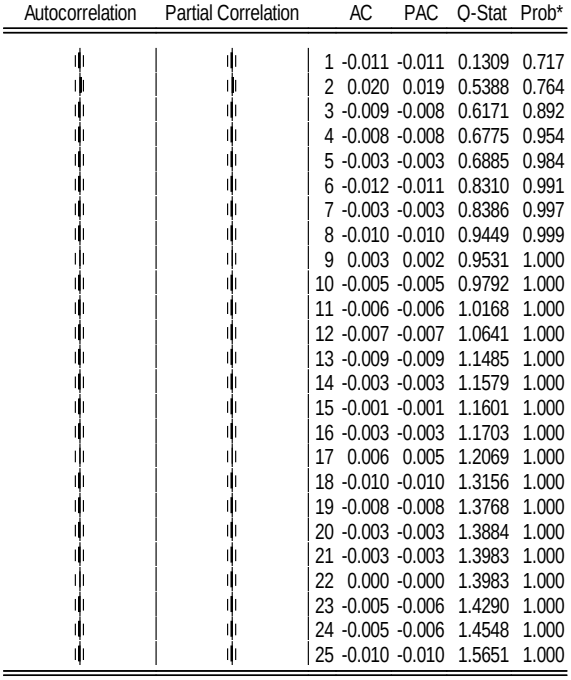



Fig. 4. The Correlogram of square residuals of ARMA (2, 1) – GARCH (1,1) model

$H_0$: There is no autocorrelation in the residuals
$H_1$: There is autocorrelation in the residuals

The serial correlation of the residuals of variance equation was tested using the correlogram Q-statistic of the squared residuals and the results are shown in Figure 4.

All p values are greater than 0.05 according to Figure 4 and null hypothesis is not rejected at 5% level of significance confirming the non-existence of serial correlation in residuals of the variance equation. ARCH-LM test was used to check the homoscedasticity of residuals and based on the results of Table 5, it can be concluded that there is no ARCH effect in the residuals as ARCH - LM test statistic (0.7183 > 0.05) is not significant at 5% significance level and the null hypothesis is not rejected.

$H_0$: There is no ARCH effect in the residuals
$H_1$: There is an ARCH effect in the residuals

Table 5. LM test on ARMA (2, 1) - GARCH(1,1) residuals

| | | | |
|---|---|---|---|
| F - statistic | 0.130202 | Prob.F(1,1057) | 0.7183 |
| Obs* R-squared | 0.130433 | Prob. Chi-Square(1) | 0.7180 |

Asymmetric effect of the conditional variance model was tested using following hypothesis [12] and the results are shown in Table 6.

$H_0$: There is no asymmetric effect
$H_1$: There is an asymmetric effect

Table 6. Sign Bias test results

| Sign Bias Test | t - value | prob. sig. |
|---|---|---|
| Sign Bias | 1.147660 | 0.2512 |
| Negative Sign Bias | 0.002899 | 0.9977 |
| Positive Sign Bias | 0.020700 | 0.9835 |
| Joint Effect | 1.426738 | 0.6993 |

According to Table 6, it is clear that the null hypothesis is not rejected at 5% significance level as the p value of Sign Bias test is greater than 0.05. Therefore, it can be concluded that there is no asymmetric effect in residuals of the ARMA(2, 1) - GARCH (1, 1) model.

Table 7. Estimated coefficients of the ARMA(2, 1) - GARCH (1, 1) model

| Variable | Coefficient | Prob. |
|---|---|---|
| AR(1) | 0.719951 | 0.0000 |
| AR(2) | 0.286423 | 0.0000 |
| MA(1) | -0.969023 | 0.0000 |
| $\omega$ | 1.81E-07 | 0.0000 |
| RESID(-1)^2-($\alpha$) | 2.153248 | 0.0000 |
| GARCH(-1)-($\beta$) | 0.272999 | 0.0000 |

The Estimated coefficients of the ARMA(2, 1) - GARCH (1, 1) model is given in Table 7 and it indicates that the coefficients of both mean equation and variance equation are significant at 5% level of significance. The positive coefficients confirm the non-negative constraints of the model. The significance of $\alpha$ and $\beta$ indicated that the conditional variance depends on the lagged squared residuals and the lagged conditional variance. That is the news about volatility from previous periods have a significant impact on the current volatility. Finally, forecast performance of the best fitted ARMA(2, 1) - GARCH (1, 1) model was evaluated using MAE, RMSE and MAE and the results are shown in Table 8.

Table 8. Forecast performance of ARMA (2, 1) - GARCH(1,1) model

| Accuracy matrix | Value |
|---|---|
| mean square error (MSE) | 2.0457529e-05 |
| root mean square error (RMSE) | 0.004525 |
| mean absolute error (MAE) | 0.002192 |

According to the results of Table 8, it is clear that the forecasting accuracy of the ARMA(2,1) - GARCH(1, 1) model is excellent as MAE, RMSE and MAE are very close to 0.

## V. CONCLUSION

This study was focused on building a forecasting model to forecast the daily USD/LKR exchange rate returns using GARCH model. The stationarity of the daily exchange rate return series was examined using ADF test. ARMA(2,1) model was identified as the best mean equation of the exchange rate returns based on the results of AIC, BIC, DW statistic and R-square value. The ARCH-LM test confirmed the presence of ARCH effect in the residuals of the conditional mean equation. After comparing accuracy measures of adequate ARCH and GARCH models, GARCH(1, 1) model was identified as the adequate model to capture the remaining conditional heteroscedastic effect of the mean model. Moreover, the ARCH-LM test was performed to test the additional ARCH effect in the residuals of ARMA(2,1) - GARCH(1, 1) model and resulted with no ARCH effect. Further, Sign-Bias test confirmed that there is no asymmetric effect in the residuals of ARMA(2,1) - GARCH(1, 1) model. Therefore ARMA(2,1) - GARCH(1, 1) model was identified as the best model to forecast the volatility of USD/LKR exchange rate return series. Finally, the forecast performance of the identified model was evaluated and MSE, RMSE and MAE were found to be to 0.047695, 0.218391, and 0.047695 respectively. Therefore it can be concluded that the prediction accuracy of the ARMA(2,1) - GARCH(1, 1) model is ample to make better decisions to minimize risk associated with USD/LKR exchange rate volatility.

## REFERENCES

[1] Keith, P., & Langeland, K. N. (2014). Forecasting exchange rate volatility: GARCH models versus implied volatility forecasts. *International Economics and Economic Policy*, 12(1). http://dx.doi.org/10.1007/s10368-014-0289-4

[2] Epaphra, M. (2017). Modeling Exchange Rate Volatility: Application of the GARCH and EGARCH Models. *Journal of Mathematical Finance*, 07(1), 121-143. http://dx.doi.org/10.4236/jmf.2017.71007

[3] Choo, W., Loo, S., & Muhammad, I. A. (2002). Modelling the Volatility of Currency Exchange Rate Using GARCH Model. *Pertanika J. Soc. Sci. & Hum.* 10(2), 85-95

[4] Ramzan, S., Ramzan, S., & Zahid, F. (2012). Modeling and Forecasting Exchange Rate Dynamics in Pakistan using ARCH Family of Models. *Electronic Journal of Applied Statistical Analysis.* 05(1), 15-29. http://dx.doi.org/10.1285/i20705948v5n1p15

[5] Gunarathna, K. V. N. C., & Aponsu, G. M. L. M. (2018). Forecasting USD/LKR Exchange Rate Using Time Series Analysis. *The 10th Symposium on Applied Science, Business & Industrial Research* . pp.50-51.

[6] D. A. Dickey & W. A. Fuller. (1979) Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*. 74, 427–431.

[7] Rocha, L., Souza, A., Santos, E., & Ferreira, N. (2014). Box-Jenkins and Volatility Models for Brazilian 'Selic' Interest and Currency Rates. *The International Journal of Latest Trends in Finance and Economic Sciences.* 4, 766-773.

[8] Catani, P.S. & Ahlgren, N.J.C. (2017). Combined Lagrange multiplier test for ARCH in vector autoregressive models. *Econometrics and Statistics.* 1, 62-84. https://doi.org/10.1016/j.ecosta.2016.10.006

[9] R. F. Engle. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*. 50(4), 987–1007.

[10] T. Bollerslev. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*. 31(3), 307–327.

[11] Brys, G. Hubert, M. & Struyf, A. (2004). A robustification of the Jarque-Bera test of normality. *COMPSTAT'2004 Symposium*. 753-760.

[12] Akpan, E. A., & Moffat, I. U. (2017). Detection and Modeling of Asymmetric GARCH Effects in a Discrete-Time Series. *International Journal of Statistics and Probability*. 6(6). 111 – 119. https://doi.org/10.5539/ijsp.v6n6p111

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index No: ES-18-50*

# Techno-Economic Assessment of Struvite Phosphorus Recovery from Sewage Sludge Treatment in Sri Lanka

Thejanie Jayasekara (1st Author)
*Department of Chemical and Process Engineering*
*University of Moratuwa*
Katubedda, Sri Lanka
thejaniej@uom.lk

Mahinsasa Rathnayake (2nd Author)
*Department of Chemical and Process Engineering*
*University of Moratuwa*
Katubedda, Sri Lanka
mratnayake@uom.lk

***Abstract* — Phosphorous is a nutrient that is considerably available in treated sludge coming out from municipal sewage treatment. In conventional sewage treatment plants, a larger portion of phosphorous remains unrecovered in the treated sludge. In this study, the techno-economic feasibility of phosphorous recovery in the form of struvite from treated sewage sludge is assessed for an already functioning sewage treatment plant in Sri Lanka. For comparative assessment between possible struvite phosphorous recovery technologies, two scenarios are considered, i.e., Scenario (A): wet chemical method with incineration and Scenario (B): wet chemical method without incineration. Composition data of treated sewage sludge is obtained from an existing sewage treatment plant, and design data for conceptualized struvite phosphorous recovery processes are retrieved from the published literature. According to the assessment results, scenario (A) provides a greater phosphorous recovery of 16.2 kg/day. However, the total energy consumption of 50,907.04 MJ/day in Scenario (A) is approximately six times higher than that of Scenario (B). The economic parameters, as well as the freshwater eutrophication impact reduction potentials for both scenarios are further analyzed. The most appropriate struvite phosphorous recovery technology for an existing or future sewage treatment plant in terms of techno-economic aspects are discussed.**

***Keywords* — *Struvite phosphorous recovery, Sewage sludge treatment, Techno-economic assessment***

## I. INTRODUCTION

Phosphorous is an essential element for food production and agricultural industry which cannot be replaced by other elements. Phosphorous is generally produced from natural phosphate rocks that exist in the form of $P_2O_5$, corresponding to almost 7,000 million tonnes of world reserves [1]. The major application of globally mined phosphate rocks (accounting for 80% in quantity) is fertilizer production [2]. Even though global phosphorous demand is rapidly increasing due to the expansion of agri-food production, it is arguab about the existence of phosphorous reserves as it is estimated that currently minable phosphorous resources only last for next 50-100 years [3]. With this resources limitation, agricultural industry and food production would be in a huge risk that enforces the inevitable requirement of alternative phosphorous-based sources for fertilizer production. Thus, it is essential to develop sustainable methodologies for phosphorous recovery from possible waste materials that do not induce excessive resources depletion.

Sewage sludge is defined as a semi-solid fluid with around 20% of solids and 80% of water content, which is produced in large quantities as a by-product from municipal wastewater treatment processes [4]. Further, sewage sludge is a rich organic source of phosphorous and other nutrients, such as calcium, aluminum, potassium, and sodium. Therefore, proper nutrient recovery is required due to the possibility of eutrophication if treated sewage sludge is disposed of without nutrient recovery. Moreover, the recovered phosphorous and other nutrients from sewage sludge can be converted to value added products, including fertilizers and organic acids which could create business opportunities with additional revenue generation from waste. Around 90% of phosphorous composition available in sewage sludge is in the solid part of the sludge. Hence, the sludge solid is required to process before converting into valuable products and releasing the treated sludge into the environment [5].

Recovery of phosphorous as struvite from sewage sludge provides a sustainable phosphorous source as other available waste sources of phosphorous, including bone meal, animal manure, etc. are insufficient in available quantities to cater the current phosphorous demand [6]. Struvite is magnesium ammonium phosphate ($MgNH_4PO_4$), and it forms a hard crystalline deposit when the molar ratio of Mg: $NH_4$: $PO_4$ becomes greater than 1: 1: 1 [7]. Several methods are available to treat sewage sludge and to recover phosphorous from sludge solids, including biological treatment method, wet-chemical treatment method, and incineration method [8]. Even though these technologies have been practiced all around the world, no such study is reported or a real phosphorous recovery plant is unavailable in the Sri Lankan context. Thus, it is necessary to evaluate the applicability of

such technologies in Sri Lanka concerning the techno-economic approach before implementing struvite phosphorous recovery for sewage sludge treatment.

Hence, this paper focuses on the technical and economic feasibility assessment for application of struvite phosphorous recovery technologies by selecting an already functioning sewage treatment plant in Sri Lanka as the case study. This study would support future policy decision making for the implementation of possible struvite phosphorous recovery plant facilities for sewage treatment plants that improve sustainable waste management in the country.

## II. OBJECTIVES

This study aims at the goals and objectives as follows.

- Determination of the level of phosphorous availability in thickened sewage sludge from a selected sewage treatment plant in Sri Lanka.
- Theoretical evaluation of the phosphorous recovery potential from a selected sewage treatment plant in Sri Lanka and related techno-economic feasibility.
- Evaluation of eutrophication impact reduction potential by struvite phosphate recovery from sewage sludge.

## III. METHODOLOGY

### *A. Raw data collection from sewage sludge treatment*

An already functioning sewage treatment plant in Sri Lanka was selected as the sampling location to obtain quality parameters of influent (raw sewage) and effluent (wastewater). This plant has a capacity of 17,000 $m^3$/day with 6000 $m^3$ of generated wastewater per day. Parameters, including Biological Oxygen Demand ($BOD_5$), Chemical Oxygen Demand (COD), Total Suspended Solids (TSS), total phosphorous, and total nitrogen were basically considered as the raw data to be collected considering a reasonable time period to include parameter variations.

Table 1 lists the obtained quality parameters of influent and effluent flows of the selected sewage treatment plant. The corresponding influent and effluent quality parameter values were utilized for calculations as necessary.

Table 1. Influent and effluent quality parameters of sewage treatment

| Parameter | Unit | Influent (Raw sewage) | Effluent (Treated water) |
|---|---|---|---|
| $BOD_5$ | mg/l | 99 | 3.6 |
| COD | mg/l | 260 | 42 |
| Total phosphorus | mg/l | 2.7 | 0.71 |
| Total nitrogen | mg/l | 54 | 12 |
| Total suspended solid | mg/l | 340 | 6.3 |

### *B. Conceptual design of struvite recovery*

A process flow diagram with the conceptual design of a struvite recovery plant was developed studying the existing process of sewage treatment. According to the current practice, raw sewage is first subjected to a primary treatment process applying gravity separation, then biologically treated by the activated sludge process as the secondary treatment. Therefore, to evaluate the feasibility of the use of phosphorous recovery, the conceptually designed struvite recovery plant was placed in between the secondary treatment process and discharge of sludge into the environment. For identification of the most suitable conceptual design, a qualitative comparison was carried out among the currently practicing methodologies for the struvite phosphorous recovery in the published literature concerning its suitability to establish for the Sri Lankan context. Accordingly, the wet chemical treatment method was identified as the widely applied struvite recovery method from sewage sludge [8]. Moreover, the effect of incineration before the wet chemical treatment method was also identified as important to study. Hence, a scenario-based conceptual design and techno-economic assessment was conducted in this study.

### *C. Scenario description*

Two scenarios were developed for struvite recovery from the wet chemical treatment method with and without incineration, i.e., Scenario (A): wet-chemical treatment of secondary sewage sludge after thickening and incineration, Scenario (B): wet-chemical treatment of secondary sewage sludge without incineration.

Material and energy flow calculations were performed for the selected two scenarios, including detailed calculations for individual unit processes, i.e., sewage sludge thickening decanter, precipitation reactor, and struvite dryer units. Each unit process is considered to be operated under the ambient conditions.

In the considered two scenarios, dewatered sludge undergoes wet chemical treatment and incineration according to the stoichiometrically balanced chemical reactions given by Equation (1) and Equation (2), respectively.

$$\text{Dewatered sludge} \rightarrow Mg^{2+} + NH_4^+ + PO_4^{3+} + 6H_2O \rightarrow MgNH_4PO_4 \cdot 6H_2O \text{ (struvite)} \quad (1)$$

$$\text{Dewatered sludge} \rightarrow \text{Ash} \rightarrow Mg^{2+} + NH_4^+ + PO_4^{3+} + 6H_2O \rightarrow MgNH_4PO_4 \cdot 6H_2O \text{ (struvite)} \quad (2)$$

The initial design parameter values were considered for mass and energy flow calculations of unit processes in the conceptual struvite recovery process based on the published literature. Table 2 lists the considered design parameters for the struvite recovery process.

Table 2. Design parameters for struvite recovery process

| Parameter | Unit | Value | Reference |
|---|---|---|---|
| Phosphate concentration | mg/l | 190 | [9] |
| pH value | - | 8.7 | [10] |
| Magnesium concentration | mg/l | 67 | [9] |
| Conversion (%) | - | 90 | [10] |
| Reaction time | hours | 7.9 | [10] |
| Moisture content of filtered struvite | g water /g dry solids | 1.5 | [10] |

All the detailed calculations were carried out with the aid of a MS excel model. Equipment sizing was also computed and obtained resultant values are presented under the Results and Discussion.

### *D. Freshwater eutrophication impact assessment*

The freshwater eutrophication impact due to unrecovered phosphorous that could be released to the environment from conventional sewage sludge treatment without struvite recovery was evaluated. The eutrophication impact reduction through struvite phosphorous recovery in sewage sludge treatment was calculated by comparison of the eutrophication impact due to unrecovered phosphorous in the scenario A and B with respect to that of conventional sewage sludge treatment. Freshwater eutrophication impact was calculated with the aid of Equation (3), and results are given by the unit of kg P equivalents.

$$\text{Freshwater eutrophication impact} = EP_i \times m_i \quad (3)$$

Where, $m_i$ = Eutrophication potential of substance i.

The ReCiPe world (H) V1.12 impact assessment method in the SimaPro Life Cycle Assessment (LCA) software was used to retrieve the mid-point characterization factors required for the freshwater eutrophication impact calculations.

### *E. Economic assessment*

In the economic assessment of the conceptual struvite recovery process, capital expenditure for the integration (CAPEX), operating expenditure (OPEX), revenue from struvite production (revenue) were considered to evaluate the economic feasibility of the considered two scenarios, in a comparative manner. Equations (4), (5), and (6) were applied to calculate CAPEX, OPEX, and revenue, respectively.

CAPEX=(Property, plant, and equipment at the end of the year – Property, plant, and equipment at the Beginning of year) + Depreciation expense (4)

OPEX=Operating Expense = Salaries + Sales Commissions +Promotional & Advertising Cost + Rental Expense + Utilities (5)

Revenue from struvite production=Sales ×Average price of sales (6)

Table 3 lists the selected economic data and parameter values for the economic assessment of the two considered scenarios. One US Dollar (USD) was considered to be equivalent to 200 Sri Lankan Rupees (LKR) in the economic calculations.

Table 3. Economic data and parameter values required for economic assessment

| Parameter | Unit | Value | Reference |
|---|---|---|---|
| Cost of land | USD per square rod | 3,303.96 | Estimated |
| Monthly labor wages | USD per person | 165.2 | Estimated |
| Tariff rate | USD/kWh | 0.074 | [11] |
| Magnesium Chloride price | USD/kg | 0.22 | [12] |
| Ammonium Chloride price | USD/kg | 0.14 | [13] |
| Market price of Kerosine | USD/l | 0.86 | [14] |
| Market price of struvite | USD/tonne | 1061.11 | [15] |

## IV. RESULTS AND DISCUSSION

### *A. Mass and energy flow results*

Table 4 indicates the mass and energy flow results obtained for the individual unit processes in Scenario (A) and Scenario (B) of the conceptualized struvite recovery process.

Table 4. Mass and energy flow results of unit processes in Scenario (A) and Scenario (B)

| Unit process | Mass/energy flow parameter | Unit | Scenario (A) | Scenario (B) |
|---|---|---|---|---|
| Sewage sludge thickening decanter | Phosphate concentration in the output | mg/l | 1,080 | 190 |
| | Wastewater flow rate | $m^3$/day | 15 | 63.08 |
| | Power consumption | kW | 12.45 | 1.9 |
| Precipitation reactor unit | Struvite production | kg/day | 55.4 | 38.81 |
| | Ammonia requirement | kg/day | 4.28 | 3.17 |
| | Magnesium requirement | kg/day | 5.71 | 4.22 |
| | Design flow rate | $m^3$/day | 90 | 70 |
| | Phosphorous availability | kg/day | 16.2 | 11.98 |
| | Power consumption of agitators | kW | 1.1 | 0.81 |
| Struvite dryer unit | Output flow rate of sewage sludge | $m^3$/day | 4.26 | 3.15 |
| | Power consumption | kW | 126.88 | 93.83 |

According to the material flow results listed in Table 4, struvite production in Scenario (A) (i.e., 55.4 kg/day) is greater than that of Scenario (B) (i.e., 38.81 kg/day). In Scenario (A), incineration of treated sewage sludge before struvite recovery helps to get rid of harmful substances, such as pharmaceutical residues, multi-resistant bacteria, microplastics, and harmful organic substances that would inhibit phosphorous recovery [16]. This could be the main reason for the considerable increase of struvite production capability of incinerated sewage sludge in Scenario (A) than non-incinerated sewage sludge in Scenario (B). In contrast, total energy consumption of the unit processes in Scenario (A) is far greater than that of Scenario (B). Table 5 reports the energy flow results of each unit process in the conceptualized struvite recovery process.

Table 5. Energy consumption of unit processes in Scenario (A) and Scenario (B)

| Unit Process | Unit | Scenario (A) | Scenario (B) |
|---|---|---|---|
| Sludge thickening decanter | MJ/day | 1,075.68 | 164.16 |
| Precipitation reactor unit | MJ/day | 95.04 | 69.98 |
| Struvite dryer | MJ/day | 11,165.32 | 8,270.84 |
| Incinerator | MJ/day | 38,571.00 | 0.00 |
| Total | MJ/day | 50,907.04 | 8,504.98 |

According to the energy consumption results indicated in Table 5, sludge thickening decanter in Scenario (A) consumes 1,075.68 MJ/day which is around seven times greater than that of Scenario (B) (i.e., 164.16 MJ/day). Further, energy consumption of the precipitation reactor unit in Scenario (A) (i.e., 95.04 MJ/day) is slightly greater for that of the same unit process in Scenario (B) (i.e., 69.98 MJ/day). Similarly, struvite dryer in Scenario (A) also consumes more energy (i.e., 11,165.32 MJ/day) than the struvite dryer in Scenario (B) (i.e., 8.270.84 MJ). Thus, Scenario (A) exhibits a clear increase in energy consumption in all unit processes compared with energy consumption of Scenario (B). The requirement of sludge thickening at a higher level in the decanter, high extent of reaction in the precipitation reactor, and the lowered dry matter content for the struvite dryer in Scenario (A) would be the potential reasons for the increment of energy consumption than of Scenario (B) [17]. Moreover, energy requirement for the incineration process in Scenario (A) is significantly higher which is not applicable for Scenario (B). Therefore, the total energy consumption of the conceptualized struvite recovery process in Scenario (A) is extensively higher than the total energy consumption of Scenario (B).

Fig. 1 illustrates the contribution of unit processes for the total energy consumption in Scenario (A) and Scenario (B).

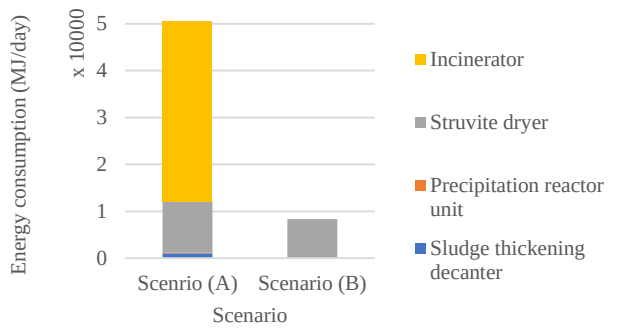



Fig. 1. Contribution of unit processes for total energy consumption in Scenario (A) and Scenario (B)

Fig.1 clearly depicts the huge difference in total energy consumption in Scenario (A) compared with Scenario (B). In consideration of energy consumption by each unit process, the sludge thickening decanter and precipitation reactor units have very low share of energy consumption. However, energy consumptions of drying units in both scenarios are significant and almost contributes to the total energy consumption of Scenario (B). In addition, energy consumption of the incineration unit in Scenario (A) is far greater compared to the other unit processes in the conceptualized struvite recovery plant. Elevated process energy consumption implies a higher quantity of associated environmental emissions as well as increased operating cost that would affect the sustainability of a process. Therefore, Scenario (B) with significantly lower energy consumption is preferrable for struvite recovery from sewage sludge even though the amount of struvite recovery is comparatively higher in Scenario (A).

### *B. Economic assessment results*

Table 5 shows the results obtained for the economic parameters, including CAPEX, OPEX, and revenue of Scenario (A) and Scenario (B).

Table 5. Economic assessment results of Scenario (A) and Scenario (B)

| Economic parameter | Unit | Scenario A | Scenario B |
|---|---|---|---|
| CAPEX | USD/year | 256,085.18 | 76,332.28 |
| OPEX | USD/year | 188,730.55 | 13,859.05 |
| Revenue | USD/year | 21,458.35 | 15,030.70 |

According to the economic parameter results, it is obvious that the annual revenue from the struvite recovery in Scenario (A) (i.e., USD 21,458.35) is around one-third greater than the annual revenue from struvite phosphorous recovery in Scenario (B) (i.e., USD 15,030.70). The significant increase in revenue is due to the enhancement of phosphorous recovery in Scenario (A) compared to that of Scenario (B). Nevertheless, other economic parameters including, CAPEX and OPEX are tremendously higher in Scenario (A). CAPEX of Scenario (A) is USD 256,085.18 per year which is almost three times that of Scenario (B) (i.e., USD 76,332.28 per year).

Similarly, OPEX in Scenario (A) shows a drastic rise indicating USD 188,730.55 per year which is nearly thirteen times greater than that in Scenario (B) (i.e., USD 13,859.05 per year). This is resulted mainly because of the cost of capital equipment required and the operating expenses associated with the greater energy consumption for the incineration process and other unit processes in Scenario (A).

Fig. 2 illustrates comparison of economic parameters in Scenario (A) and Scenario (B) in graphical mode.

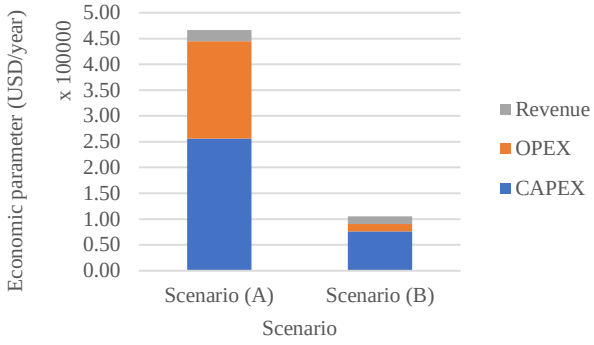



Fig. 2. Economic parameters of Scenario (A) and Scenario (B)

According to the graph, the greatness of CAPEX and OPEX with respect to revenue in Scenario A is obvious in comparison to Scenario B. Therefore, Scenario B is preferable for struvite phosphorous recovery from sewage sludge in terms of techno-economic perspective.

### C. *Freshwater eutrophication impact assessment results*

Table 6 shows the calculated freshwater eutrophication impact results for the case of conventional sewage treatment, and struvite phosphorous recovery scenarios, i.e., Scenario (A) and Scenario (B). In conventional sewage treatment, there is no phosphorous recovery from treated sewage sludge that would cause freshwater eutrophication impact when discharged to the environment. In Scenarios (A) and (B) with struvite phosphorous recovery, the unrecovered phosphorous due to the recovery efficiency of the respective technology can cause the freshwater eutrophication impact.

Table 6. Freshwater eutrophication impact results

| Description | Freshwater eutrophication impact (kg of P eq.) |
|---|---|
| Conventional sewage treatment | 17.10 |
| Scenario (A) | 0.90 |
| Scenario (B) | 5.12 |

According to the results in Table 6, both scenarios (A) and (B) with struvite phosphorous recovery have significantly reduced freshwater eutrophication impact values compared with the impact of the conventional sewage treatment without phosphorous recovery. Fig.3 illustrates the freshwater eutrophication impact reduction potentials of Scenarios (A) and (B) in graphical form.

As depicted in Fig.3, Scenario (A) has the highest freshwater eutrophication impact reduction potential that corresponds to nearly 95%. However, Scenario (B) is also capable of reducing freshwater eutrophication impact by around 70% via struvite phosphorous recovery. In overall consideration of trade-off among material and energy consumption, revenue/expenses difference, and environmental impact reduction, Scenario (B) surpasses Scenario (A) in this techno-economic assessment. Therefore, struvite phosphorous recovery using the wet chemical method without incineration (i.e., Scenario (B)) could be suggested as the most appropriate technology scenario for future phosphorous recovery integrated sewage treatment plants in Sri Lanka.

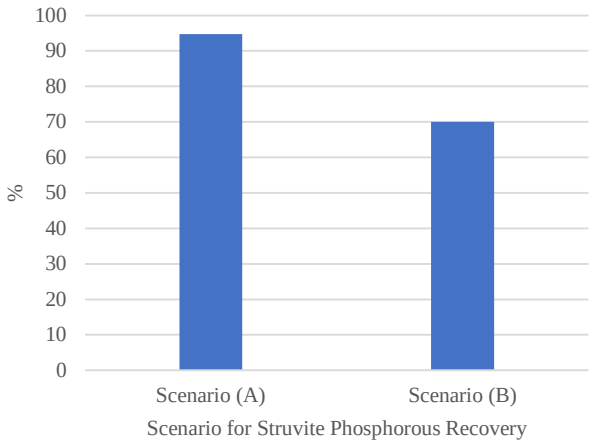



Fig.3. Freshwater eutrophication impact reduction potentials of Scenarios (A) and (B)

## V. CONCLUSION

In this study, a techno-economic assessment was conducted considering material and energy utilization, economic analysis, and environmental benefit of two technology scenarios for struvite phosphorous recovery from treated sewage sludge. The findings from this study reveal that the wet chemical method with incineration provides enhanced struvite phosphorous recovery. In contrast, overall comparison of techno-economic aspects demonstrates the appropriateness of the wet chemical method without incineration for struvite phosphorous recovery. Accordingly, the results from this assessment conclude that further treatment of conventionally treated sewage sludge for struvite phosphorous recovery is feasible and beneficial for existing sewage treatment plants. Thus, this assessment would support future decision making for implementation of new sewage treatment plants or modification of existing sewage treatment plants with struvite phosphorous recovery.

## REFERENCES

[1] D. B. Colleti, B. Chavel, and W. J. Gatti, “Challenges of skew in D. Seyhan, "Country-scale phosphorus balancing as a base for resources conservation", *Resources, Conservation and Recycling*, vol. 53, no. 12, pp. 698-709, 2009. Available: 10.1016/j.resconrec.2009.05.001.

[2] L. Shu, P. Schneider, V. Jegatheesan and J. Johnson, "An economic evaluation of phosphorus recovery as struvite from digester supernatant", *Bioresource Technology*, vol. 97, no. 17, pp. 2211-2216, 2006. Available: 10.1016/j.biortech.2005.11.005.

[3] Y. Liu, J. Chen, A. Mol and R. Ayres, "Comparative analysis of phosphorus use within national and local economies in China", *Resources, Conservation and Recycling*, vol. 51, no. 2, pp. 454-474, 2007. Available: 10.1016/j.resconrec.2006.10.012.

[4] S. Baroutian, N. Eshtiaghi and D. Gapes, "Rheology of a primary and secondary sewage sludge mixture: Dependency on temperature and solid concentration", *Bioresource Technology*, vol. 140, pp. 227-233, 2013. Available: 10.1016/j.biortech.2013.04.114.

[5] C. Blöcher, C. Niewersch and T. Melin, "Phosphorus recovery from sewage sludge with a hybrid process of low pressure wet oxidation and nanofiltration", *Water Research*, vol. 46, no. 6, pp. 2009-2019, 2012. Available: 10.1016/j.watres.2012.01.022.

[6] "Scientists Warn of Lack of Vital Phosphorus as Biofuels Raise Demand - Schwartzreport", *Schwartzreport*, 2021. [Online]. Available: https://www.schwartzreport.net/2008/06/24/scientists-warn-of-lack-of-vital-phosphorus-as-biofuels-raise-demand/. [Accessed: 10- Nov- 2021].

[7] Y. Jaffer, T. Clark, P. Pearce and S. Parsons, "Potential phosphorus recovery by struvite formation", *Water Research*, vol. 36, no. 7, pp. 1834-1842, 2002. Available: 10.1016/s0043-1354(01)00391-8.

[8] M. Sikosana, D. Randall and H. Von Blottnitz, "A technological and economic exploration of phosphate recovery from centralised sewage treatment in a transitioning economy context", *Water SA*, vol. 43, no. 2, p. 343, 2017. Available: 10.4314/wsa.v43i2.17.

[9] P. van Rensburg, E. Musvoto, M. Wentzel and G. Ekama, "Modelling multiple mineral precipitation in anaerobic digester liquor", *Water Research*, vol. 37, no. 13, pp. 3087-3097, 2003. Available: 10.1016/s0043-1354(03)00173-8.

[10] M. Rahman, M. Salleh, U. Rashid, A. Ahsan, M. Hossain and C. Ra, "Production of slow release crystal fertilizer from wastewaters through struvite crystallization – A review", *Arabian Journal of Chemistry*, vol. 7, no. 1, pp. 139-155, 2014. Available: 10.1016/j.arabjc.2013.10.007.

[11] A. Dzyuba and I. Solovyeva, "PRICE-BASED DEMAND-SIDE MANAGEMENT MODEL FOR INDUSTRIAL AND LARGE ELECTRICITY CONSUMERS", *International Journal of Energy Economics and Policy*, vol. 10, no. 4, pp. 135-149, 2020. Available: 10.32479/ijeep.8982.

[12] "Buy 99 Anhydrous Magnesium chloride mgcl2 white in China on Alibaba.com", *Alibaba.com*, 2021. [Online]. Available: https://www.alibaba.com/product-detail/99-Anhydrous-Magnesium-chloride-mgcl2-white_62459136551.html. [Accessed: 10- Nov- 2021].

[13] "Buy Ammonium Chloride 99 5 min in China on Alibaba.com", *Alibaba.com*, 2021. [Online]. Available: https://www.alibaba.com/product-detail/Ammonium-Chloride-99-5-min_62152385558.html. [Accessed: 10- Nov- 2021].

[14] "Kerosene prices around the world, 08-Nov-2021 | GlobalPetrolPrices.com", *GlobalPetrolPrices.com*, 2021. [Online]. Available: https://www.globalpetrolprices.com/kerosene_prices/. [Accessed: 10- Nov- 2021].

[15] D. Kok, S. Pande, J. van Lier, A. Ortigara, H. Savenije and S. Uhlenbrook, "Global phosphorus recovery from wastewater for agricultural reuse", *Hydrology and Earth System Sciences*, vol. 22, no. 11, pp. 5781-5799, 2018. Available: 10.5194/hess-22-5781-2018.

[16] A. Bhasin, "Recovery of Phosphorus from Incineration of Sewage Sludge", Degree project in Chemical Science and Engineering, second cycle, KTH Royal Institute of Technology School of Chemical Science and Engineering, 2017.

[17] D. Đurđević, P. Blecich and Ž. Jurić, "Energy Recovery from Sewage Sludge: The Case Study of Croatia", *Energies*, vol. 12, no. 10, p. 1927, 2019. Available: 10.3390/en12101927.

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index No: ES-19-53*

# Advancements in Environmental Technologies for Sustainable Urban Regeneration: A Comparative Assessment

Ruwan Danuka Edirisinghe (1st Author)
*Faculty of Graduate Studies*
*University of Colombo*
Colombo, Sri Lanka
ruwanedirisinghe1@gmail.com

***Abstract* — The present study aims to appraise advancements in environmental technologies applicable to urban regeneration, with a special focus on urban brownfield redevelopment. The rapid literature review technique was employed as the research strategy, in the mixed method research design. Technological solutions proposed in the selected articles were comparatively assessed their practicality in an urban setting, in terms of cost, efficacy, physical space required and potential harm to the neighboring environment, by using a five-point scale scoring system. In this study, nanoremediation, thermal remediation methods (i.e. electrical resistance heating, thermal conduction heating and steam enhanced extraction), non-thermal physical remediation methods (electrokinetic remediation, non-thermal plasma technologies, air sparging, soil washing and replacement and passive treatment technologies such as permeable reactive barriers), chemical oxidation (advanced chemical oxidation and Fenton process), and nature-based solutions or bioremediation or gentle remediation technologies (biodegradation processes methods such as bioaugmentation, bioventing, bioprecipitation, biostimulation, landfarming, and phytoremediation methods such as phytostabilization, phytovolatilization and phytoextraction or phytomining and monitored natural attenuation) are presented. Each environmental restoration strategies provided has its own set of limitations, application possibilities and future development potential, as evidenced by this study. Nanoremediation, bioremediation and radio frequency heating in the current state of the art are found to be feasible for an urban area. Property developers and urban authorities could consider the application potential of these technologies in urban brownfield redevelopment in urban regeneration. An integrated approach for addressing the limitations of these technologies may be worth considering in research and developments in the urban sector.**

***Keywords* — *Decontamination, Environmental-Technology, Green, Redevelopment, Remediation, Urban***

## I. INTRODUCTION

Despite the inequalities in socioeconomic development, urbanization is happening at a fast pace over the world. It is debatable whether these human-centred developments will coexist with the natural environment in the long run. Land resources, on the other hand, are limitedly available, particularly in urban areas. The author of this study supports the idea that urban development should not lead to gentrification. Furthermore, the urban environments are deteriorating, turning them into inhabitable places (i.e. urban brownfields) because of loss of vegetation, air pollution, water scarcity, contamination of lands, to name a few [1]. The process of urban redevelopment can be significantly more complicated than addressing physical aspects like renovating structures (e.g. neighborhood and quality of life) [2]. These concerns have raised the importance of sustainable renovation of urban environments to meet the present and future challenges in making the urban sector more livable to humans and other living beings while remaining in optimum harmony with natural environment.

Infrastructure expansion and industrialization in metropolitan areas at the expense of ecosystem health, as well as derelict properties that have become brownfields, are posing risks to landscape resilience, which necessitate redevelopment. However, regeneration of urban areas having poor environmental quality can be a practically challenging task in terms of investment and technology [3]. A soil is deemed polluted when the contaminant levels surpass the natural assimilation capacity of the soil system. Urban soils become contaminated by polychlorinated biphenyls (PCBs), heavy metals, hydrocarbons (e.g. Polycyclic Aromatic Hydrocarbons), phthalate, alkylbenzene, microplastics, and persistent organic pollutants (POPs) [4-7]. The sources of pollution can be of point and or non-point (e.g. industrial parks versus sediments transported through urban waterways). Toxins in the urban soils and aquatic systems lead to biomagnification, affecting human health [8]. Indicators such as the Bio Concentration Factor are used to estimate the transfer of contaminants from soil to plants. The World Health Organization has defined permissible limits for common contaminants in soil ($mg.kg^{-1}$) of 20 for arsenic, 5000 for iron and 100 for lead [9].

Previous studies propose interdisciplinary approaches and innovative technological solutions to remediate contaminated lands in an urban setting. The remediation process is started by characterizing the contaminated site, which may involve analysis of operating history of the site, invasive drilling for hydrogeological assessments, and use of unmanned aerial vehicles, Digital Elevation Models, LiDAR and GIS based approaches for soil mapping to comprehend the fate and transport of contaminants [10-11]. Contaminants are analyzed by using laboratory methods such as Titrimetry, Spectrophotometry, Graphite Furnace Atomic Absorption Spectrometry, Gas Chromatography, Mass Spectrometry and Inductively Coupled Plasma Mass Spectrometry [12].

Soil remediation technologies include nanoremediation, phytoremediation, physical and chemical methods such as subsurface heating technologies, soil replacement and chemical oxidation, by on-site or off-site basis, reducing the bioavailability of toxins, lessening the risks to the environment. Importantly, choice of remediation technologies would determine the technical and economic feasibilities [13]. Therefore, the understanding of advancements in environmental technologies is unarguably valuable for policymakers, authorities, urban designers and developers to make cities more sustainable. The present study explores advancements in technologies for environmental remediation in urban redevelopment. This critical appraisal from an interdisciplinary perspective shows the application and further research potentials of these technologies under different scenarios.

## II. OBJECTIVES

This study aims to appraise technological advancements in contaminated land redevelopment, delimiting the study to identify technological solutions applicable to revitalization of urban brownfields, and to identify the future developmental potentials of these technologies.

## III. METHODOLOGY

The present exploratory study adopted a rapid literature review technique [14] as the research strategy, with a deductive approach in a mixed method design. The search strategy comprised of defining inclusion and exclusion criteria, key words and synonyms and formulating search strings. Inclusion criteria comprised of contaminated land remediation, research papers and grey literature published in English and journal pre-proofs, whereas the exclusion criteria consisting the non-urban studies, non-English publication and abstract only publications. Predatory journals, publications published before year 2010, research papers with low scientific quality were eliminated from the primary screening process. The Google Scholar, Elsevier, ScienceDirect, Emerald, ResearchGate, Web of Science and MDPI were among the prominent databases employed in the search strategy. A total of 467 publications were screened and 68 were used in the analysis. The environmental technologies considered in this study were comparatively assessed in terms of cost, efficacy, physical space required and potential harm to the neighboring environment. The assessment criteria were ranked using a five-point-scale scoring system (i.e. 0,1,2,3 and 4 for no, low, medium, high and extremely high, respectively) [15]. The 'Impacts on Neighborhood' score was subtracted from the algebraic sum of the scores of the remaining three criteria in each remediation technology to arrive at the overall possibility ranking (Table 1).

## IV. RESULTS AND DISCUSSION

Decontamination of urban lands to restore the ecological health has been addressed in literature mainly by integrating the applications of life sciences, physical sciences and engineering. In this study, nanoremediation, thermal remediation, air sparging, non-thermal physical remediation methods inclusive of electrokinetic remediation, chemical oxidation and bioremediation technologies are presented, with especial focus on their applicability to land redevelopment in the urban context.

The use of nanotechnology for remediation (i.e. nanoremediation) of polluted soils is becoming a promising solution. The underline principle of nanoremediation is the prevention of migration of contaminants by solidification and stabilization using engineered nanoparticles of having high surface area and reactivity (e.g. acting as adsorbents and reductants). Nanoparticles such as Nanoscale Zerovalent Iron (nZVI) can detoxify soils by immobilizing toxic ions such as heavy metals by adsorption (Fig. 1) [16-17]. Estimation of adsorption capacity is given in (1) [18].

$$q_e = \frac{(C_0 - C_e)}{m} V \quad (1)$$

Where, $q_e$ adoption capacity at the equilibrium state (mg.g$^{-1}$), $C_0$ initial concentration of pollutant (mg. L$^{-1}$), $C_e$ equilibrium concentration of pollutant (mg. L$^{-1}$), $V$ volume of solution, $m$ mass of nanomaterial (g)

Nanoscale elemental iron particles that are pyrophoric are coated with a passivating oxide (e.g. FeO(OH)) or natural or synthetic polymers (e.g. Carboxymethyl Cellulose or Poly Acrylic Acid) [19]. Commercially, nZVI are available in emulsified, aqueous or dry form (e.g. *NANO IRON s.r.o.* in Czech Republic). Nanomaterials of desired properties are mechanically injected into contaminated soils.

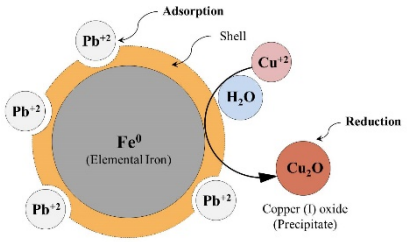



Fig. 1. The mechanism of nanoparticles immobilizing heavy metals

However, biological processes (e.g. microbial activity) in the soil can be negatively affected, despite the effective remediation provided by nZVI [20-21]. A few possible reasons for this may be the concentration of nZVI applied and the physical contact with nZVI (i.e. with the disruption of microbes' cell membranes), resulting in alterations in the composition of microbial communities in terms of species

diversity and functionalities [22-24]. Encapsulation of nZVI with shells made out of soluble compounds such as Magnesium Hydroxide ($Mg(OH)_2$) may be worth considering in manipulating the reactivity and release of nZVI into the soil environment [25]. Alternatively, bioremediation (e.g. Phytoremediation) to decontaminate soils can be assisted with reactive nanomaterials such as nanoscale Titanium Dioxide ($TiO_2$ NPs) which is a photocatalyst [26]. Similar to nZVI, $TiO_2$ NPs could adversely affect the soil health [27], which could be moderated by encapsulation with Carboxymethyl Cellulose [28]. Other widely used nanomaterials include Multiwalled Carbon Nanotube (MWCNT) and magnetic nanoparticles applied for decontamination of soils polluted by heavy metals [29]. However, the efficiency of remediation could depend on the MWCNT concentration [30]. Other factors affecting the efficacy of the nanoremediation include the pH, temperature and physical properties of the contaminated soil (e.g. porosity and soil texture), and initial concentration, contact time and efficiency of diffusion of nanomaterials [31]. Nanoparticles are employed in environmental monitoring (i.e. tracer) apart from remediating degraded environments [32]. Overall, nanomaterials appear to be a viable option for soil remediation in an urban setting, as long as the potential risks to soil ecology are considered.

Soil thermal remediation is another approach to decontaminate polluted urban soils. In this method, subsurface soils are heated by employing Electrical Resistance Heating (ERH), Thermal Conduction Heating (TCH) and Steam Enhanced Extraction (SEE) methods on site to fluidize or vaporize pollutants, and Soil Vapor Extraction (SVE) assemblies are used to recover the vaporized pollutants (i.e. vapor collection wells or vacuum wells secured with vapor cap) for subsequent treatment. In the ERH, electrical current (i.e. alternative current) is applied through electrodes inserted into subsoil from which heat is produced by the electrical resistivity of soils, whereas TCH comprises electrical heaters to raise the temperature of subsoil by thermal conduction. Similar to TCH, the SEE heats subsurface soils by thermal conduction, using steam piping systems installed in soil layers. Alternatively, Low Temperature Thermal Desorption is used for ex-situ treatment of contaminated soils. However, previous authors present supporting and opposing arguments on soil remediation by thermal methods. Hydrocarbons such as Polycyclic Aromatic Hydrocarbons (PAHs) bound to clayey matter and organic carbons have been found effectively desorbed when the soil is thermally treated [33]. Pesticides such as Hexachlorocyclohexanes (HCHs) are degraded with thermal remediation [34]. Efficacy of thermal treatment for soils contaminated with PCBs can be improved by adding calcium hydroxide ($Ca(OH)_2$) [35]. On the other hand Microwave Coupled Infrared Radiation (MCIR) and Radio Frequency Heating (RFH) as non-ionizing radiation methods would provide comparatively lessened adverse impacts on microbial activities in the soil environment owing to thermal remediation [36]. RFH could be a low-cost and option [37]. Efficient decontamination of heavy metals such as Mercury (Hg) by thermal treatment at low temperatures ranging from 100 – $400^0$C has been established in literature [38]. Flame retardants like Decabrominated Diphenyl Ether (BDE 209) in soils can be effectively treated thermally [39]. Importantly, removal of polyester microfibers in soils by thermal remediation has shown improved soil microbial activities [40]. Soils polluted with oily matter (e.g. lubricants) can be remediated by employing thermal desorption methods at low temperatures and Fluidized Bed Reactors [41]. Nonaqueous Phase Liquid (NAPLs) in soils can be removed thermally by vaporization [42-43]. However, this empirical evidence mostly resulted in temperatures above $300^0$C may lead to waste of energy and detrimental effects on soil ecology.

Air Sparging (soil venting) is a physical treatment method to remove contaminants in the saturated zone of the soils. In this method, pressurized air is injected into the contaminated area, allowing volatile contaminants to volatilize, which are collected from vapor extraction vacuum wells situated in the vadose zone of the soil profile [44]. Contaminated air is then treated by means of biofiltration, adsorption and combustion. Air sparging method is applied for groundwater remediation as well. A fundamental limitation of the air sparging remediation method could be that the efficacy of the process can be affected by the permeability of the soil layers, which permits contaminants to migrate even more in the soil profile. However, compared to potential heat stress resulting from thermal methods, air sparging process may harmless to the soil's physical, chemical and biological properties.

Among the non-thermal physical remediation methods, Electrokinetic Remediation (ER) has been widely discussed in previous literature. In ER, an electric field is generated to remove contaminants using electromigration and electroosmosis phenomena, which can be coupled with biological and chemical methods as well [45-46]. Advanced oxidation process such as Non-thermal Plasma Technologies (NPT), for example, Dielectric Barrier Discharge (DBD), Pulsed Corona Plasma and Non-thermal Plasma Fluidized Bed (PFB) are emerging technologies in soil remediation [47-50]. The process of in situ vitrification offers a wider range of applications for treating soils contaminated with organic, radioactive and inorganic hazardous wastes [51]. Ultrasonic Desorption and coal agglomeration are employed in soils contaminated with oils [52]. Despite the sophistication of these technologies, pragmatic aspects in field implementation can be debatable. Similarly, conventional methods such as soil replacement and surfactant-aid soil washing [53] in an urban setting might not be practical due to potential adverse effects on neighboring load bearing structures supported on soils and other physical, legal and administrative constraints. However, passive treatment technologies such as Permeable Reactive Barriers (PRB) which are used for treating groundwater in a contaminated site and soil remediation by encapsulation (e.g. silica encapsulation) is another approach that can be adopted in metropolitan areas [54].

Urban soils can be decontaminated by chemical methods. More prevalent method is the In Situ Chemical Reduction (ISCR) or chemical oxidation, which encompasses hydrolysis, advanced oxidation, redox and mineralization. In some instances, the same chemical method is applied to remediate both contaminated soils and waters. Chemicals such as hydrogen peroxide, persulfate, potassium dichromate and alkali (e.g. NaOH) are typically used [55]. Although the Fenton oxidation as a pre-oxidative method to treat organic

pollutants is not a novel approach in today's context, Photo-Fenton, Chelate Modified Fenton and Sono-Fenton processes are being reemerged in the field of soil remediation. In the Fenton method, in an aqueous medium, ferrous (Iron (II)) ions act as a catalyst to generate hydroxyl radicals and hydroxide ions by reacting with hydrogen peroxide (i.e. Haber-Weiss Reaction), from which resultant hydroxyl radicals oxidize the pollutant. The Fenton process can be applied to remediate soils contaminated with hydrocarbons [56]. Contrastingly, contaminates in dry soil may require wetting the soil to apply the Fenton process [57]. Studies have shown that, hydrocarbons such as PHAs in soils can be oxidized by ammonium persulfate, assisted with subsequent microbial degradation [58]. Despite the reported microbial activity, the Fenton method may be deemed more ecologically sound than ammonium persulfate-based treatment due to residual sulphate concerns. Chemical methods, on the other hand, can compromise the functionality of soil microorganisms [59-60]. Moreover, certain chemical remediation approaches seem to necessitate precise conditions, therefore the expected treatment efficacy in an urban context may be debatable.

Nature-based solutions (i.e. bioremediation) to remediate contaminated soils are becoming increasingly popular as a green initiative. The underlying principles of nature-based solutions are the stabilization and accumulation of contaminants with the aid of plants and microorganisms, and biomonitoring. Nature-based solutions widely discussed in the literature include biodegradation process such as bioaugmentation, bioventing, bioprecipitation, biostimulation and landfarming, and phytoremediation methods such as phytostabilization, phytovolatilization and phytoextraction or phytomining, and monitored natural attenuation. In bioaugmentation, cultured microorganisms are introduced into contaminated soils to accelerate the biodegradation rate. For example, *Sphingomondas* and *Mycobacterium* species can be used to remediate soils containing PAHs [61]. By supplying air, the bioventing technique, on the other hand, allows microorganisms already present in the soil to breakdown pollutants. Similarly, with the biostimulation approach, nutrients are provided to indigenous microbes to manage the limiting factors on biodegradation. With the help of microorganisms, the bioprecipitation process induces pollutants, particularly heavy metals, to precipitate. Phytostabilization is the process of immobilizing contaminants utilizing plants while reducing bioavailability and eliminating the means for contaminant migration (i.e. soil erosion). Because no organism is purposefully introduced to the soil environment, bioventing and biostimulation can be considered environmentally safe. In phytovolatilization, on the other hand, pollutants are absorbed by plants and then released into the environment as gases through evapotranspiration. Bioaccumulation of heavy metals such as Cd by hyperaccumulators like *Malva rotundifolia* has shown a Bioaccumulation Coefficient greater than 1.0 [62]. Empirical evidence shows that *Thlaspi elegans* grown in serpentine soils has over 15000 $mg.kg^{-1}$ (dry weight basis) accumulation rate of nickel (Ni) [63]. Urban soils contaminated with radionuclides can be effectively remediated by phytoextraction [64]. Genetic engineering approaches such as recombinant RNA and DNA technologies are applied to improve the efficiency of biological agents used in bioremediation measures [65-66]. However, despite the environmental friendliness of phytoremediation technologies; physical space, plant and microbial responses to environmental vulnerabilities, and the time required to remediate contaminated soils can all be limiting factors in an urban setting, compared to other biodegradation methods.

Table 1. Comparison of practicabilities of technologies in an urban setting

| Env. Technology | Limitations | | | | Overall Possibility Rank | Reference |
|---|---|---|---|---|---|---|
| | Cost Effectiveness | Physical Space Saving | Efficacy | Impacts on Neighborhood | | |
| Nanoremediation | 3 | 4 | 3 | 2 | **8** | [67-68] |
| Thermal ERH, TCH | 2 | 1 | 1 | 2 | 2 | [69] |
| Thermal - RFH | 3 | 3 | 3 | 1 | **8** | [70] |
| Air Sparging | 2 | 2 | 2 | 2 | 4 | [71] |
| Chemical - ISCR | 1 | 1 | 3 | 3 | 2 | [72] |
| Chemical - Fenton | 3 | 1 | 2 | 3 | 4 | [73] |
| Electrokinetic | 3 | 2 | 1 | 2 | 4 | [74-75] |
| Non-thermal Plasma | 3 | 1 | 1 | 2 | 5 | [76] |
| Nature Based | 3 | 2 | 3 | 0 | **9** | **[77]** |
| Soil Replacement | 1 | 0 | 4 | 4 | 1 | [78] |
| Soil Washing | 2 | 0 | 3 | 4 | 1 | [79-80] |
| Vitrification | 2 | 1 | 2 | 3 | 2 | [81] |
| Encapsulation | 2 | 3 | 3 | 1 | 7 | [82] |
| Passive PRB | 3 | 2 | 2 | 1 | 6 | [83] |

It is obvious from this research that the environmental restoration methods presented have their own set of limitations, application possibilities and further developmental potential. However, nanoremediation, radio frequency heating and bioremediation can be regarded as feasible for urban regeneration process, with especial focus on redevelopment of contaminated lands (Table 1).

## V. CONCLUSION

By devising a rapid literature review process and adopting a mixed method research design, the present study assessed advancements in environmental technologies applicable to urban regeneration, with a special focus on land redevelopment. For an urban setting, nanoremediation, radio frequency heating, and bioremediation have all been recognized as viable solutions. However, the proposed environmental remediation methods have their own set of limitations, application possibilities and potential for future research. The application potential of these technologies in urban brownfield redevelopment toward sustainable urban regeneration can be considered by the property developers and urban authorities. Researchers in the field of urban development can explore the developmental potential of new technological avenues by integrating several methods discussed in this paper, addressing the limitations of the technologies, interdisciplinary manner.

## REFERENCES

[1] R.F.M. Ameen and M. Mourshed, Urban environmental challenges in developing countriesd A stakeholder perspective. Habitat International, vol. 64, pp. 1-10, 2017, doi: 10.1016/j.habitatint.2017.04.002.

[2] E. Korol, Y. Gaydysheva and D. Passmore, Integration of organizational-technological and social aspects in the realization of the program of renovation of residential development. MATEC Web of Conferences, vol. 251, 2018, doi: 10.1051/matecconf/201825106031

[3] C.C. Egolum and F.I. Emoh, The Issues and Challenges of Urban Renewal in a Developing Economy. International Journal of Development and Economic Sustainability, vol.5, no.1, pp.32-44, 2017.

[4] H. Yu, Y. Liu, X. …and Y. Pan, Assessment of the spatial distribution of organochlorine pesticides (OCPs) and polychlorinated biphenyls (PCBs) in urban soil of China, Chemosphere, 2019, doi: 10.1016/j.chemosphere.2019.125392.

[5] D. Castillo-Nava, M. Elias-Santos, …and H. A. Luna-Olvera, Heavy metals (lead, cadmium and zinc) from street dust in Monterrey, Mexico: ecological risk index. International Journal of Environmental Science and Technology, pp. 1-10, 2020, doi: 10.1007/s13762-020-02649-5.

[6] T. Minkina, S. Sushkova, … and E. Antonenko, Polycyclic Aromatic Hydrocarbons in Urban Soils Within the Different Land Use: A Case Study of Tyumen, Russia. Polycyclic Aromatic Compounds, pp.1-16, doi: 10.1080/10406638.2018.154099.

[7] L. Yang, Y. Zhang, S. Kang, Z. Wang and C. Wud, Microplastics in soil: A review on methods, occurrence, sources, and potential risk. Science of The Total Environment, vol. 780, 2021, doi: 10.1016/j.scitotenv.2021.146546.

[8] V. Pecina, M. Brtnický, T. Baltazár, D. Juřička, J. Kynický and M.V. Galiová, Human health and ecological risk assessment of trace elements in urban soils of 101 cities in China: A meta-analysis. Chemosphere, 2020, doi: 10.1016/j.chemosphere.2020.129215.

[9] J. M. Nyika1, E. K. Onyari, M.O. Dinka and S.B. Mishra, Heavy Metal Pollution and Mobility in Soils within a Landfill Vicinity: A South African Case Study. Oriental Journal of Chemistry, vol. 35, pp. 1286-1296, 2019, doi: 10.13005/ojc/350406.

[10] C. Swyngedouw, Sampling Methods for Site Characterization. Environmental Geochemistry (Second Edition), pp. 13-24, 2018, doi: 10.1016/B978-0-444-63763-5.00003-3.

[11] A. Capolupo, S. Pindozzi, …and L. Boccia, Photogrammetry for environmental monitoring: The use of drones and hydrological models for detection of soil contaminated by copper. Science of the Total Environment, vol.514, pp.298–306, 2015, dio: 10.1016/j.scitotenv.2015.01.109.

[12] M. Zeiner, R. Pirkl and I. JuranovićCindrić, Field-Tests versus Laboratory Methods for Determining Metal Pollutants in Soil Extracts. Soil and Sediment Contamination, 2019, doi: 10.1080/15320383.2019.1670136.

[13] New Jersey Department of Environmental Protection, Findings & Recommendations for the Remediation of Historic Pesticide Contamination - Final Report March 1999. Retrieved from: https://www.nj.gov/dep/special/hpctf/final/costs.htm

[14] G. McCartney, W. Hearty, M. Taulbut, R. Mitchell, R. Dryden and C. Collins, Regeneration and health: a structured, rapid literature review. Public Health, vol. 148, pp. 69-87, 2017, doi: 10.1016/j.puhe.2017.02.022.

[15] M. Naseri-Rad, R. Berndtsson, K.M. Persson, and K. Nakagawa, An efficient guide for sustainable remediation practice in addressing contaminated soil and groundwater. Science of The Total Environment, 2020, doi:10.1016/j.scitotenv.2020.13

[16] M. Gil-Díaz, S. Diez-Pascual, A. Gonzalez, J. Alonso, E. Rodríguez-Valdes, J.R. Gallego and M.C. Lobo, A nanoremediation strategy for the recovery of an As-polluted soil. Chemosphere, vol. 149, 137-145, 2016, doi: 10.1016/j.chemosphere.2016.01.106.

[17] M. Radziemska, Z.M.Gusiatin, …and M.Brtnicky, Nano Zero Valent Iron (nZVI) as an Amendment for Phytostabilization of Highly Multi-PTE Contaminated Soil. Materials, vol.14, pp.1-10, 2021, doi:10.3390/ma14102559.

[18] M. Gil-Díaz, M.A.Álvarez, J.Alonso and M. C. Lobo, Efectiveness of nanoscale zero-valent iron for the immobilization of Cu and/or Ni in water and soil samples. Scientific Reports, vo.10, 2020, doi: 10.1038/s41598-020-73144-7.

[19] Z. Dongsheng, G. Wenqiang, … and L. Youzhi, Removal of heavy metal lead(II) using nanoscale zero-valent iron with different preservation methods. Advanced Powder Technology,vol.30,pp. 581-589, 2019, doi:10.1016/j.apt.2018.12.013.

[20] M. Anza, O. Salazar, …and C. Garbisu, The Application of Nanoscale Zero-Valent Iron Promotes Soil Remediation While Negatively Affecting Soil Microbial Biomass and Activity. Frontiers in Environmental Science, vol. 7, 2019, doi: 10.3389/fenvs.2019.00019.

[21] C. Fajardo, J. Garcıa-Cantalejo, … and M.Martin, New insights into the impact of nZVI on soil microbial biodiversity and functionality. Journal of Environmental Science and Health Part A, 2018, doi: 10.1080/10934529.2018.1535159.

[22] C. Fajardo, S. Sánchez-Fortún, …and M. Martín, Evaluation of nanoremediation strategy in a Pb, Zn and Cd contaminated soil. Science of the Total Environment, vol.706, pp.1-11, 2020, doi: 10.1016/j.scitotenv.2019.136041.

[23] Z. Shen, Z. Chen, Z. Hou, T. Li and X. Lu, Ecotoxicological effect of zinc oxide nanoparticles on soil microorganisms. Front. Environ. Sci. Eng., 2015, doi: 10.1007/s11783-015-0789-7.

[24] Y. Hu, M. Zhang, R. Qiu and . Li, Encapsulating nanoscale zero-valent iron with a soluble Mg(OH)2shell for improved mobility and controlled reactivity release. Journal of Materials Chemistry A, 2017, doi: 10.1039/c7ta08605a.

[25] A.D. Zand, A.M. Tabrizi and A.V. Heir, Co-application of biochar and titanium dioxide nanoparticles to promote remediation of antimony from soil bySorghum bicolor: metal uptake and plant response. Heliyon, vol. 6, 2020, doi: 10.1016/j.heliyon.2020.e04669.

[26] M. Simonin, A.Richaume, J.P.Guyonnet, A. Dubost, J.M.F. Martins and T.Pommier, Titanium dioxide nanoparticles strongly impact soil microbial function by affecting archaeal nitrifiers. Scintific Reporters, vol.6, pp. 1-10, doi: 10.1038/srep33643.

[27] J. Virkutyte, S. Al-Abed, H.Choi and C.Bennett-Stamper, Distinct structural behavior and transport of TiO2nano- and nanostructured particles in sand. Colloids and Surfaces A: Physicochem. Eng. Aspects, vol. 443, pp. 188–194, 2014, doi: 10.1016/j.colsurfa.2013.11.004.

[28] D.Baragañoa, J.Alonsob, J.R.Gallegoa, M.C.Lobob and M.Gil-Díazb, Magnetite nanoparticles for the remediation of soils co-contaminated with As and PAHs. Chemical Engineering Journal, vol. 399, 2020, doi: 10.1016/j.cej.2020.125809.

[29] X. Chen, J. Wang, K. Hayat, D. Zhang and P. Zhou, Small structures with big impact: Multi-walled carbon nanotubes enhanced remediation efficiency in hyperaccumulatorSolanum nigrumL. under cadmium and arsenic stress. Chemosphere, vol. 276, pp.1-12, 2021, doi: 10.1016/j.chemosphere.2021.130130.

[30] Y. Gong, Y. Liu, Z. Xiong and D. Zhao, Immobilization of Mercury by Carboxymethyl Cellulose Stabilized Iron Sulfide Nanoparticles: Reaction Mechanisms and Effects of Stabilizer and Water Chemistry. Environ. Sci. Technol., vol. 48, 2014, doi: 10.1021/es404418a.

[31] X. Gong, H. Danlian, Y. Liu, Z. Guangming, R. Wang, P. Xu, C. Zhang, M. Cheng, W. Xueab and S. Chenab, Roles of multiwall carbon nanotubes in phytoremediation: cadmium uptake and oxidative burst in Boehmeria nivea (L.) Gaudich. Environ. Sci.: Nano, vol. 6, pp. 851-862, 2019, doi: 10.1039/C8EN00723C.

[32] S. Su, W. Wu, J. Gao, J. Lub and C. Fan, Nanomaterials-based sensors for applications in environmental monitoring. Journal of Materials Chemistry, vol.22, 2012, doi:10.1039/C2JM33284A.

[33] Y. Cheng, H. Sun, E. Yang, J. Lv, B. Wena, F. Sun, L. Luo and Z. Liu, Distribution and bioaccessibility of polycyclic aromatic hydrocarbons in industrially contaminated site soils as affected by thermal treatment

. Journal of Hazardous Materials, vol. 411, 2021, doi: 10.1016/j.jhazmat.2021.125129.

[34] C. M. Dominguez, A. Checa-Fernandez, A. Romero and A. Santos, Degradation of HCHs by thermally activated persulfate in soil system: Effect of temperature and oxidant concentration. Journal of Environmental Chemical Engineering, pp.1-10, 2021, doi:10.1016/j.jece.2021.105668.

[35] U. Roland, F. Holzer and F. Kopinke, Influenceof Low-and High-frequency Heating on Biodegrading Microorganisms in Soil: Microbial Degradation, Journal of Microwave Power and Electromagnetic Energy, vol.47, pp.280-289, 2013, doi:10.1080/08327823.2013.11689865.

[36] K. Yang, M. Fan, …and Y. Wang, An effective thermal treatment of toluene-contaminated soil by uniform heating from microwave coupled infrared radiation (MCIR). Journal of Soils and Sediments,vol.21, pp.1329–1336, 2021, doi: 10.1007/s11368-020-02855-1.

[37] S. Srirattana, K. Piaowan, G.V. Lowry and T. Phenrat, Electromagnetic induction of foam-based nanoscale zerovalent iron (NZVI) particles to thermally enhance non-aqueous phase liquid (NAPL) volatilization in unsaturated porous media: Proof of concept. Chemosphere, vol. 183,pp.323-331,2017,doi: 10.1016/j.chemosphere.2017.05.114.

[38] M. Park, M. Kim and Y. Hong, The kinetics of mercury vaporization in soil during low-temperature thermal treatment. Geoderma, vol.363, 2020, doi: 10.1016/j.geoderma.2019.114150.

[39] C. Ko, Y. Yang, J.I.W.W. Liu and Y. Shih, Thermal treatment of decabrominated diphenyl ether in its highly contaminated soil in Taiwan. Chemosphere, vol.287, 2021, doi: 10.1016/j.chemosphere.2021.131924.

[40] Q.Q. Guo, M.R. Xiao and G.S. Zhang, The persistent impacts of polyester microfibers on soil bio-physical properties following thermal treatment. Journal of Hazardous Materials, vol. 420, 2021, doi: 10.1016/j.jhazmat.2021.126671.

[41] U. Samaksaman, T. Peng, J. Kuo, C. Lu and M. Wey, Thermal treatment of soil co-contaminated with lube oil and heavy metals in a low-temperature two-stage fluidized bed incinerator, Applied Thermal Engineering, 2015, doi: 10.1016/j.applthermaleng.2015.09.024.

[42] Q. Xie, K.G. Mumford, B.H. Kueper and C. Zhao, A numerical model for estimating the removal of volatile organic compounds in laboratory-scale treatability tests for thermal treatment of NAPL-impacted soils. Journal of Contaminant Hydrology, vol.226, 2019, doi:10.1016/j.jconhyd.2019.103526.

[43] G.C. Scholes, J.I. Gerhard, G.P. Grant, D.W. Major, J.E. Vidumsky, C. Switzer and J.L. Torero, Smoldering Remediation of Coal-Tar-Contaminated Soil: Pilot Field Tests of STAR. Environ. Sci. Technol., vol. 49, pp. 14334−1434, 2015, doi: 10.1021/acs.est.5b03177.

[44] M.L. Brusseau, Soil and Groundwater Remediation. Environmental and Pollution Science, pp. 329-354, 2019, doi:10.1016/B978-0-12-814719-1.00019-7.

[45] C. Cameselle, S. Gouveia, … and B. Belhadj, Advances in Electrokinetic Remediation for the Removal of Organic Contaminants in Soils. Organic Pollutants - Monitoring, Risk and Treatment, 2013, doi: 10.5772/54334.

[46] D. Liyanage and J. Walpita, Organic pollutants from E-waste and their electrokinetic remediation. Handbook of Electronic Waste Management, pp. 171-189, 2020, doi: 10.1016/B978-0-12-817030-4.00006-1.

[47] C. Du and J. Yan, Remediation of Phenanthrene-Contaminated Soil Using Non-thermal Plasma Fluidized Bed. Plasma Remediation Technology for Environmental Protection, pp.7-20, 2017, doi: 10.1007/978-981-10-3656-9_2.

[48] H. Xu, L. Cang, Y. Song and J. Yang, Influence of electrode configuration on electrokinetic-enhanced persulfate oxidation remediation of PAH-contaminated soil. Environmental Science and Pollution Research, 2020, doi: 10.1007/s11356-020-10338-6.

[49] J. Zhan, A. Zhang, …and Y. Liu, Remediation of Perfluorooctanoic Acid (PFOA) polluted soil using Pulsed Corona Discharge Plasma. Journal of Hazardous Materials, 2019, doi: 10.1016/j.jhazmat.2019.121688.

[50] C.A. Aggelopoulos, P. Svarnas, M.I. Klapa and C.D. Tsakiroglou, Dielectric barrier discharge plasma used as a means for the remediation of soils contaminated by non-aqueous phase liquids. Chemical Engineering Journal, vol.270, pp. 428–436, 2015, doi: 10.1016/j.cej.2015.02.056.

[51] M.D.F. Rodríguez, M.C.G. Gómez, …and J.V.Tarazona, Soil Pollution Remediation. Encyclopedia of Toxicology (Third Edition), pp. 344-355, 2014, doi: 10.1016/B978-0-12-386454-3.00579-0.

[52] J. Li , X. Song , G. Hu and R. W. Thring, Ultrasonic desorption of petroleum hydrocarbons from crude oil contaminated soils. Journal of Environmental Science and Health, Part A: Toxic/Hazardous Substances and Environmental Engineering, vol.48, pp.1378-1389, doi:10.1080/10934529.2013.781885.

[53] Y. Li, X. Liao, … and Q. Lin, The combined effects of surfactant solubilization and chemical oxidation on the removal of polycyclic aromatic hydrocarbon from soil. Science of the Total Environment, vol. 647, pp. 1106–1112, 2019, doi: 10.1016/j.scitotenv.2018.07.420.

[54] V.O. Akpoveta, Process optimization of silica encapsulation technique as a unique remediation technology for the treatment of crude oil contaminated soil. Egyptian Journal of Petroleum, vol. 29, pp. 113–119, 2020, doi: 10.1016/j.ejpe.2019.12.004.

[55] E. J. Ordoñez-Frías, V. I. Domínguez-Rodríguez and R. H. Adams, Remediation of soil contaminated by di(2-ethylhexyl) phthalate by chemical oxidation or bioremediation. nternational Journal of Environmental Science and Technology, pp. 1-10, 2020, doi: 10.1007/s13762-020-03007-1.

[56] A. Gharaee, M. R. Khosravi-Nikou and B. Anvaripour, Hydrocarbon contaminated soil remediation: A comparison between Fenton, sono-Fenton, photo-Fenton and sono-photo-Fenton processes. Journal of Industrial and Engineering Chemistry, vol. 79, pp.181–193, 2019, doi: 10.1016/j.jiec.2019.06.033.

[57] M. Farzadkia, M. Dehghani and M. Moafian, The effects of Fenton process on the removal of petroleum hydrocarbons from oily sludge in Shiraz oil refinery, Iran. Journal of Environmental Health Science & Engineering, vol.12, pp.1-7, 2014,doi: 10.1186/2052-336X-12-31.

[58] R. Medina, P.M.D. Gara, A. J. Fernández-González, J. A. Rosso and M.T.D. Pann, Remediation of a soil chronically contaminated with hydrocarbons through persulfate oxidation and bioremediation. Science of the Total Environment, vol. 618, pp. 518–530, 2018, doi: 10.1016/j.scitotenv.2017.10.326.

[59] E. Kakosová, P. Hrabák, … and L. Vrtoch, Effect of various chemical oxidation agents on soil microbial communities. Chemical Engineering Journal, 2016, doi:10.1016/j.cej.2016.12.065.

[60] E. Martinez-Pascual, T. Grotenhuis, …and M. Vinas, Coupling chemical oxidation and biostimulation: effects on the natural attenuation capacity and resilience of the native microbial community in alkylbenzene-polluted soil. Journal of Hazardous Materials, 2015, doi: 10.1016/j.jhazmat.2015.06.061.

[61] L.K. Redfern, C.M. Gardner, E. Hodzic, P.L. Ferguson, H. Hsu-Kim and C.K. Gunsch, A new framework for approaching precision bioremediation of PAH contaminated soils. Journal of Hazardous Materials, vol. 378, 2019, doi: 10.1016/j.jhazmat.2019.120859.

[62] M. Wu, Q. Luo, S. Liu, Y. Zhao, Y. Long and Y. Pan, Screening ornamental plants to identify potential Cd hyperaccumulators for bioremediation. Ecotoxicology and Environmental Safety, vol. 162, 35–41, 2018, doi: 10.1016/j.ecoenv.2018.06.049.

[63] N. Koleli, A. Demir, … and R. Binzet, Heavy Metal Accumulation in Serpentine Flora of Mersin-Findikpinari (Turkey) – Role of Ethylenediamine Tetraacetic Acid in Facilitating Extraction of Nickel. Soil Remediation and Plants, pp. 629–659, 2015, doi:10.1016/b978-0-12-799937-1.

[64] L. Yan, Q.V. Le, … W. Peng, Phytoremediation of radionuclides in soil, sediments and water. Journal of Hazardous Materials, 2020, doi:10.1016/j.jhazmat.2020.1247

[65] P. Kumar and G. Baul, Recombinant DNA Technology for Bioremediation of Pollutants. In: Fulekar M.H. (eds) Bioremediation Technology, 2010, doi: 10.1007/978-90-481-3678-0_8.

[66] A. Kumar, S. Devi and D. Singh, Significance and Approaches of Microbial Bioremediation in Sustainable Development. Microbial

Bioprospecting for Sustainable Development, 2018, doi:10.1007/978-981-13-0053-0_5.

[67] Araújo, R., Castro, A. C. M., & Fiúza, A. (2015). The Use of Nanoparticles in Soil and Water Remediation Processes. Materials Today: Proceedings, 2(1), 315–320. doi:10.1016/j.matpr.2015.04.055

[68] G. Medina-Pérez, F. Fernández-Luqueño, ... and M. Miranda-Arámbula, Remediating Polluted Soils Using Nanotechnologies: Environmental Benefits and Risks. Pol. J. Environ. Stud. Vol. 28, (3), 2019, pp. 1013-1030, doi: 10.15244/pjoes/87099

[69] P.L. O'Brien, T.M. DeSutter, ...A.F. Wick, Thermal remediation alters soil properties – a review. Journal of Environmental Management, vol. 206, pp. 826–835, 2018, doi:10.1016/j.jenvman.2017.11.0

[70] G. Huon, T. Simpson, ...U. Roland, In Situ Radio-Frequency Heating for Soil Remediation at a Former Service Station: Case Study and General Aspects. Chemical Engineering & Technology, vol. 35(8), pp. 1534–1544, 2012, doi:10.1002/ceat.201200027

[71] S. Khalid, M. Shahid...C. Dumat, A comparison of technologies for remediation of heavy metal contaminated soils. Journal of Geochemical Exploration, vol. 182, pp. 247–268, 2017, doi:10.1016/j.gexplo.2016.11.021

[72] L. Huysegoms, S. Rousseau and V. Cappuyns, Chemical or Natural? Including LCA in Social CBA to Compare Remediation Alternatives for a Dry-Cleaning Facility. Sustainability, vol. 11(7), 2019, doi:10.3390/su11071975

[73] Z. Zhou, X. Liu, ...W. Ouyang, Persulfate-based advanced oxidation processes (AOPs) for organic-contaminated soil remediation: A review. Chemical Engineering Journal, vol. 372, pp. 836-851, 2019, doi:10.1016/j.cej.2019.04.213

[74] B. Gidudu and E.M.N. Chirwa, The combined application of a high voltage, low electrode spacing, and biosurfactants enhances the bio-electrokinetic remediation of petroleum contaminated soil. Journal of Cleaner Production, 2020, doi:10.1016/j.jclepro.2020.1227

[75] D. Wen, R. Fu and Q. Li, Removal of Inorganic Contaminants in Soil by Electrokinetic Remediation Technologies:A Review. Journal of Hazardous Materials, 2020, doi:10.1016/j.jhazmat.2020.123345

[76] H. Zhang, D. Ma, R. Qiu, Y. Tang and C. Du, Non-thermal plasma technology for organic contaminated soil remediation: A review. Chemical Engineering Journal, vol. 313 pp. 157-170, 2017, doi: 10.1016/j.cej.2016.12.067.

[77] X. Wan, M. Lei and T.Chen, Cost–benefit calculation of phytoremediation technology for heavy-metal-contaminated soil. Science of The Total Environment, 2016, doi:10.1016/j.scitotenv.2015.12

[78] D. Lavee, T. Ash and G. Baniad, Cost-benefit analysis of soil remediation in Israeli industrial zones. Natural Resources Forum, vol. 36, pp. 285–299, 2012.

[79] W. Chen and H. Li, Cost-Effectiveness Analysis for Soil Heavy Metal Contamination Treatments. Water, Air, & Soil Pollution, vol. 229(4), 2018, doi:10.1007/s11270-018-3784-3

[80] Voglar, D., & Lestan, D. (2014). Chelant soil-washing technology for metal-contaminated soil. Environmental Technology, vol. 35(11), 1389–1400. doi:10.1080/09593330.2013.869265

[81] J.E. Vidonish, K. Zygourakis, ...P.J.J. Alvarez, Thermal Treatment of Hydrocarbon-Impacted Soils: A Review of Technology Innovation for Sustainable Remediation. Engineering, vol. 2(4), pp. 426–437, 2016, doi:10.1016/j.eng.2016.04.005

[82] L.Liu, W. Li, W. Song and M. Guo, Remediation techniques for heavy metal-contaminated soils: Principles and applicability. Retrieved_from:https://www.sciencedirect.com/science/article/pii/S0048969718309215

[83] S.W. Jeen, R.W. Gillham and A. Przepiora, Predictions of long-term performance of granular iron permeable reactive barriers: Field-scale evaluation. Journal of Contaminant Hydrology, vol. 123(1-2), pp. 50–64, 2011, doi:10.1016/j.jconhyd.2010.12.00

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index No: ES-20-61*

# Agricultural Call Centres: An Overview Of 1920 Agricultural Advisory Service Usage, Sri Lanka

Sanduni Rathnayake (1st Author)
*Department of Agricultural Extension*
*University of Peradeniya*
Peradeniya, Sri Lanka
sandunianu@agri.pdn.ac.lk

T.A. Kamiss (2nd Author)
*Department of Agriculture*
*National Agricultural Information Communication Centre*
Peradeniya, Sri Lanka
kamiss.doa@gmail.com

U.I. Dissanayeke (3rd Author)
*Department of Agricultural Extension*
*University of Peradeniya*
Peradeniya, Sri Lanka
uvasarad@gmail.com

P.C.B. Alahakoon (4thAuthor)
*Department of Agricultural Extension*
*University of Peradeniya*
Peradeniya, Sri Lanka
prasad_uop@yahoo.com

M.F.M. Rizwan (5th Author)
*Department of Agriculture*
*National Agricultural Information Communication Centre*
Peradeniya, Sri Lanka
rizwan.doa@gmail.com

***Abstract* — The 1920 Agricultural Advisory Service (AAS) was introduced in 2006, as an ICT initiative of agricultural extension in Sri Lanka. There are no studies assessing the long-term usage of AAS by the community. Therefore, secondary research was carried out to identify the trends and patterns in 1920 AAS usage. Data were collected from internal databases, records, seasonal reports of National Agriculture Information and Communication Centre (NAICC), journal articles and web sources. Results revealed an increasing trend of receiving queries from 2006 to 2012 and, from 2016 to 2018. The trend was decreasing from 2012 to 2015. By the end of 2018, AAS has received 546,473 queries since its inception. Queries on fruits, vegetables and subject-specific matters have been highest for several years. The highest number of queries was regarding 'cultivation' matters and queries from Colombo district was the highest in every year. There was no seasonal variation of the total number of queries received. Peak hours of AAS were from 8.00 am to 10.00 am where it served 31.6% of the total queries received per month. This study revealed several avenues for primary research. A detailed field level studies on the parameters claimed for the highest number of queries will be helpful to plan tailor-made extension programs. Mapping of queries received by AAS will provide a valuable information source for the policy makers in designing field extension programmes in Sri Lanka.**

***Keywords* — *1920 Agricultural Advisory Service, Agricultural Community, Call center solutions, ICT for agriculture, Sri Lanka***

## I. INTRODUCTION

Rapid spread of mobile phone usage has accelerated the use of ICT in agricultural extension in developing countries [1,2]. In line with that, Call Centre models have been successfully used in offering demand driven agricultural advisory services to farmer communities [3,4]. These call centres were initiated mainly targeting small scale farmers those who had limited access to information [5,4]. Further, such centres were also helpful in supplementing the traditional extension systems which had inadequate resources to meet urgent information needs of the farmers [5]. The food crop sector in Sri Lanka employs more than 2500 field extension workers, and an individual extension worker has to serve about 3000-4000 farm families [6] which is not adequate at all. Therefore, by identifying the need of an alternative mechanism to cater farmer information needs, government interest was shifted to incorporate ICTs in agricultural extension [7]. Accordingly, Call Centre model was initiated by the Department of Agriculture, Sri Lanka (DOASL) to address day-to-day information needs of farmers [6,8].

In 2006 "Govi Sahana Sarana" was initiated by the DOASL where people can call and clarify farming related problems through a hotline 1920 [8]. The service was completely free of charge [8] and named as 1920 Toll Free Service (TFS). It operated daily from 8.00 am to 6.00 pm. Later, from 2011 onwards, callers had to bear the cost of telephoning, while the advisory service remained free of charge. From 2007 February onwards, TFS operated from 8.30 am to 4.15 pm during only weekdays. In October 2016, the location of the service was shifted to the National Agriculture Information & Communication Centre (NAICC), Gannoruwa which is the communication arm of DOASL. Simultaneously, the service was re-named as "1920 Agricultural Advisory Service" (AAS). During the study period, there were 18 call centre agents who are trained

Agricultural Instructors (AI s), attending for the telephone queries and providing required information to the callers.

According to the system plan of AAS in Fig.1, 1920 is receiving telephone calls through various service providers and those calls are connected to Public Switched Telephone Network (PSTN). Through Session Initiation Protocol (SIP trunk), calls are diverted to Call Centre server and answered by Call Centre agents. Callers also can communicate with 1920 through Facebook, WhatsApp, Imo, Skype, Viber and emailing which facilitates detailed information sharing (eg. Images). According to the DOASL, for 2006 to 2008, 40,000-50,000 queries have been annually made by clients [9]. This service became very popular as it was free of charge and delivered in local language [8]. High level of mobile phone usage by the farmers in the country to access the information [10] may also have triggered this success of AAS. There are several micro-level studies conducted on AAS at the field level [9,11]. However, there are no studies available, investigating the long-term patterns and trends in AAS use by the island wide farmer community. Therefore, this study will try to fill that research gap by identifying the trends and patterns of AAS use, for the period of 2006 to 2018. The findings of this study will be helpful to identify the real information requirement of the farmers and to plan tailor-made advisory programmes.

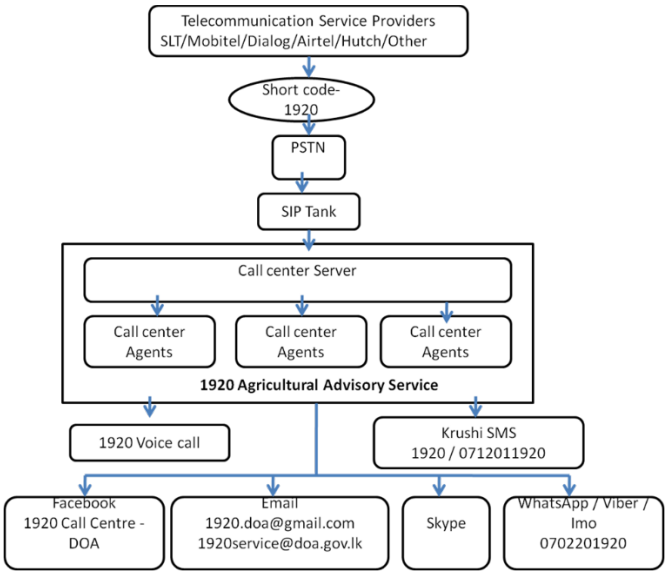



Fig. 1. System plan of 1920 AAS

## II. OBJETIVES

The general objective of this secondary research is to identify trends and patterns of usage of AAS by the community in the period of 2006 to 2018.

## III. METHODOLOGY

The study was conducted as a secondary research (desk research) [12,13] in 2019. Data were collected from the internal databases, records, seasonal (Yala and Maha ) reports maintained by NAICC for a period of 12 years from 2006 to 2018. Annual reports, books, and web sources were also used to collect the data. Data were analyzed using MS Excel and SPSS version 16.

## IV. RESULTS AND DISCUSSION

### *A. Trends in total number of queries received (2006-2018)*

The average number of calls received during the period of 2006 to 2018 was 41,565 (Fig. 2). Based on the service demand analysis, three distinctive phases can be identified in relation to the number of calls received by the AAS by the end of 2018. The demand for AAS has changed time to time. As depicted in Fig. 2, phase I shows an increasing trend of receiving queries during the initial 6 years till 2012. In 2006, when the service was first established, the number of calls received were quite small compared to subsequent years. This was the first attempt of using mobile phones for agricultural information dissemination in Sri Lanka. Since the community was at the early stage of the adoption process, number of calls may not be very high. But, when the level of awareness is increased, people automatically started to recognize 1920 as an important information source. According to Kumari et.al. (2009), 8.8% (n=80) of the vegetable farmers in Kandy district have identified 1920 as their main information source for farming activities while 86% of contacted growers (n=50) have informed fellow farmers about 1920 [11]. Hence, awareness of 1920 AAS within the community might have been increased through fellow farmers and other sources of information. Subsequently, the number of queries received may be increasing gradually till 2012.

During phase II which lasted from 2012 to 2015, there was a decline in the number of calls received by the AAS. One possible reason for the reduced numbers would be the implementation of a cost-sharing approach in 2011, where the callers must bear the cost for the phone call. When the awareness increased about the expenses of calling to AAS, it might have decreased the service usage by the community.

In the phase III, there was an increase in number of calls during the period of 2016 to 2018 where AAS underwent various changes. The number of telephone lines and number of call agents were increased from 4 to 20 and from 6 to 18 consequently. As a result, 1920 AAS was able to accommodate more calls at a given time. A huge promotional campaign of 1920 AAS was conducted by using posters and through DOASL mass media programs. These measures resulted increased public awareness of the service. Until 2016, 1920 AAS used to be in an office area with a limited space and later, it was shifted to a new, spacious location. It reduced the external disturbances and enabled clear communication between the caller and call agent. All these strategies may have combined effect for the increase in number of calls received by 1920 AAS following the year 2015.

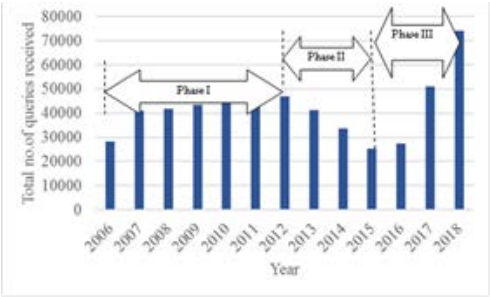



Fig. 2. Total number of queries received by 1920 AAS from 2006 - 2018

The average number of queries received during the period from 2006 to 2017 was 39,358. In 2018 alone, the AAS received 74,168 of queries which is 88.4% increase of queries compared to the average. It was the highest figure recorded in its 12 years of history. In 2018, Fall Army Worm (FAW) pest attack appeared in Maize cultivations [14] and people were educated to contact 1920, if they encounter this pest [15]. Farmers were mindful about the unidentified pests in their fields. It increased the number of calls received by AAS asking verifications about the pests they observed. This was coupled with the photo sharing through Viber and WhatsApp, which were newly added social media platforms into 1920 AAS.

The forecasting of annual number of calls using exponential smoothing method revealed that the number of calls has been steadily increasing from the inception and then there was a minor drawback (Fig. 3). Now, the service is recovering and it is heading towards a rapid increase of queries in the coming years.

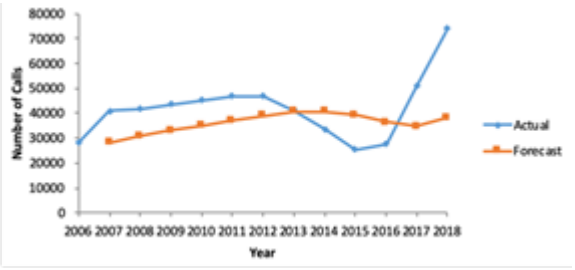



Fig. 3. Forecasting trend of total annual number of calls

*B. Trends in crop categories claimed for queries:*

The queries received by 1920 AAS were classified into 11 categories. DOASL is following general agricultural extension approach where it covers advisory work for a wide spectrum of crops. Therefore, the number of queries can be answered directly by the DOASL call agents, is high. The queries on other types of crops (eg. tea, coconut) or group of crops (eg. export agricultural crops) are directed to the relevant institution where necessary. It was evident that queries related to floriculture, medicinal plants, roots and tuber crops and livestock is relatively lesser than queries in other crop categories. Queries related to fruits, vegetables and subject-specific problems was the highest in several years (Table 1) covering more than one fifth of the total queries raised for corresponding year. Until 2016, the lowest number of queries had been raised regarding medicinal plants (average- 153 queries) and the queries related to livestock was the lowest from 2016 to 2019 April (average-261 queries). Exponential smoothing analysis of annual number of calls on fruits, vegetables and paddy (Fig. 4, Fig. 5, and Fig. 6 consequently) showed a similar pattern to that of the total number of annual calls (shown in the Figure 3).


Fig. 4. Forecasting the trend of annual number of calls on fruits

Table 1. Crop categories claimed for highest number of queries in each year (2006-2018).

| Year | Crop category with highest number of queries | Number of queries received | Percentage out of total queries received for the corresponding year |
|---|---|---|---|
| 2006 | Fruits | 6743 | 23.8 |
| 2007 | Vegetables | 11438 | 27.9 |
| 2008 | Vegetables | 11439 | 27.3 |
| 2009 | Fruits | 10598 | 24.4 |
| 2010 | Fruits | 12550 | 27.8 |
| 2011 | Vegetables | 11217 | 23.9 |
| 2012 | Fruits | 12809 | 27.3 |
| 2013 | Fruits | 10705 | 25.9 |
| 2014 | Subject-specific problems | 7866 | 23.4 |
| 2015 | Subject-specific problems | 6105 | 23.9 |
| 2016 | Subject-specific problems | 6691 | 24.2 |
| 2017 | Fruits | 11938 | 23.4 |
| 2018 | Subject-specific problems | 16910 | 22.7 |

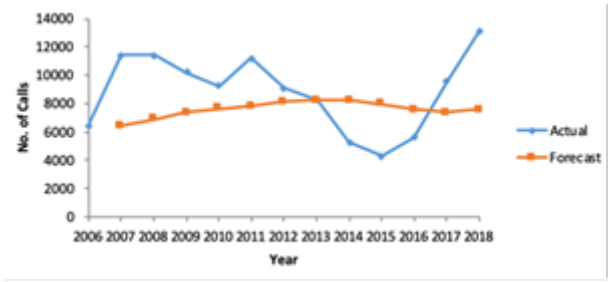



Fig. 5. Forecasting the trend of annual number of calls on vegetables

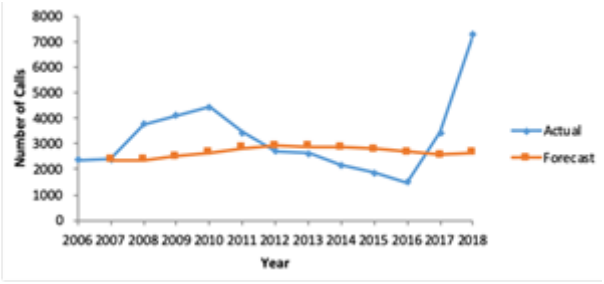



Fig. 6. Forecasting the trend of annual number of calls on Paddy

Further detailed analysis of repeatedly queried crop categories will be beneficial in developing crop-specific extension programmes

The 1920 AAS database has seasonal reports for both cultivation seasons in Sri Lanka; Yala and Maha [16]. Analysis of seasonal data applicable for 2006 to 2016 showed that, in both seasons, vegetables, fruits and subject-specific problems have claimed for the highest number of queries, interchangeably, which is similar to the findings of analysis of total number of queries received per year. Paddy is the main food crop in the country, cultivated in both seasons. Hence, the average amount of queries received on Paddy was also considered with vegetables and fruits (Table 2).

Table 2. Average amount of queries received by crop category in two seasons (2006 - 2018).

| Crop category | *Maha* | *Yala* |
|---|---|---|
| Paddy | 1790 | 1329 |
| Fruits | 4868 | 4993 |
| Vegetables | 4443 | 4326 |
| Total | 20,699 | 20,005 |

The total number of queries and the queries about paddy, fruits and vegetables were higher in Maha than Yala in between years 2006 to 2018 (Table 2). The Mann-Whitney U test results showed that there is no significant difference between total numbers of queries received for two seasons ($U = 63$, $p = 0.854$). There is a significant difference ($\alpha$=0.10) between total numbers of queries received for two seasons about paddy ($U = 35$, $p = 0.056$). In Maha season, Paddy is cultivated in many parts in the island and in Yala it is limited mostly to the certain areas. As a result, the number of queries received in Maha is higher than in Yala season.

There is no significant difference between total numbers of queries received for two seasons about fruits ($U = 62$, $p = 0.806$) and about vegetables ($U = 61$, $p = 0.758$). It appears that the problems related to key crop categories are not much affected by the seasonal variations.

### *C. Trends in agronomic practices and other related information claimed for queries:*

For 2006 to 2016 period, highest no. of queries has been received regarding 'cultivation' issues. According to Fig. 7, except in 2009, not much seasonal variation can be observed with respect to no. of queries related to 'cultivation'. It will be important to further study on the crop categories claimed for high amount of 'cultivation-related' information as it will indicate the real information requirement, which would help to develop crop-specific extension programmes.

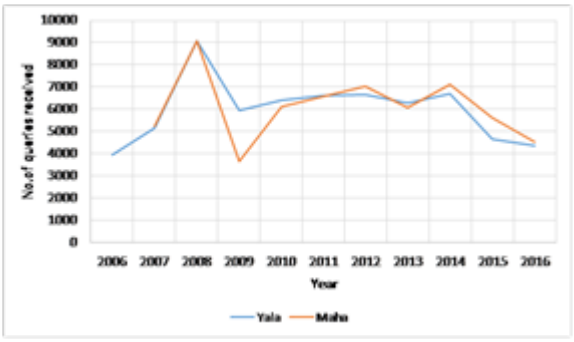



Fig.7. Seasonal variation of queries received regarding 'cultivation' issues (2006-2016).

### *D. Trends in geographical area claimed for queries:*

According to Fig. 8, for the period of 2006 to 2016, the highest amount (14.7%) of queries have received from Colombo district (Average 5509 queries), followed by Kurunegala (11.5%) and Kandy (9.8%) districts. In 2008, there were 8778 queries from Colombo district which covered 16.7% of the total queries received and 26% of queries were related to 'subject-specific problems'. Home gardening' is also included in 'subject-specific problems' category of AAS. Part-time farming is common in Colombo district which is basically an urban area. Highest number of queries about fruits and vegetables, and higher no.of queries from Colombo district reflects that people do home gardening (where fruits and vegetables are common) as part-time or secondary activity, would contact AAS very often. However, a detailed field level study is required to validate this assumption. Reasons behind the highest number of calls from such an urban area may be due to easy access to mobile phones, lack of contacts with the field level extension officers, etc. It will be worthwhile to study whether AAS is performing far better in extension and advisory work in urban agriculture in Sri Lanka.

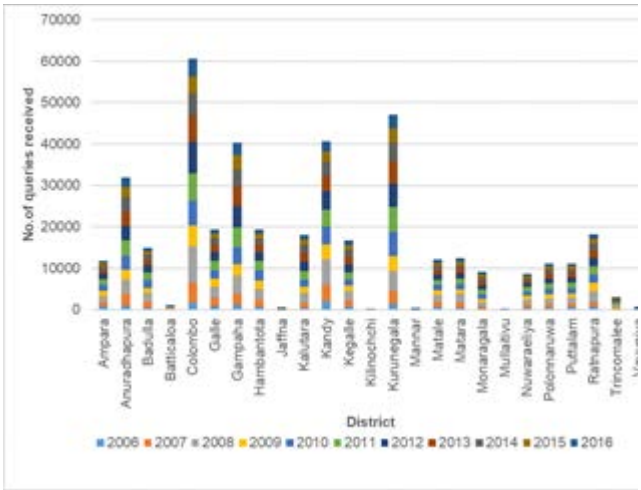



Fig. 8. Geographical variation of total no. of queries received (2006-2016)

According to the Figure 8, a lesser number of queries has been received from Mulativu (14), Mannar (28), Killinochchi (18), and Jaffna (43) districts where Tamils makes the majority of population. Lack of Tamil speaking Call Center agents in 1920 AAS to respond the queries, lack of awareness about the 1920 AAS or availability of good agricultural extension set up in the area, may also have some effect on this. Further field level study will provide more insights on how to make these 'less-active' areas more engaged with AAS.

Interestingly, Anuradhapura and Polonnaruwa districts, which are mainly Paddy cultivating areas in the country, [17] claimed for 7.7% and 2.7% of queries consequently, which is far behind the values for Colombo district. It can be assumed that AAS would not play a significant role in providing extension and advisory services to Paddy farmers. It may be due to availability of strong field level extension network at the field where no other alternative information source is needed, or lack of acceptance of new communication tools by the traditional Paddy farmers. However, a detailed field level study will be required to explain this scenario. A detail analysis on the "problematic" districts will provide a good foundation to plan "site-specific" agricultural information services.

### *E. Peak hours of 1920 AAS*

It was observed 31.6% of the total queries received per month (average) in the period of 2017 to 2019 had been received in between 8.00am to 10.00am. Identifying these peak hours is beneficial for the decision makers in AAS to make sure that, all the call agents are at their seats during that time, to receive the call. It avoids the caller keep waiting and enable smooth functioning of the service. It will be interesting to study the relationship between the peak hours of 1920 and the types of queries raised, so that, call agent would know 'what type of queries comes in what time', in advance.

## V. CONCLUSION

AAS is providing different types of agricultural information to the farmers all over country. However, in its 12 years history, there are not many studies available assessing the AAS at the field level as an ICT initiative for agricultural extension in Sri Lanka. Being secondary research, this study provides an overview of the AAS usage over the years and it bring forward some multiple avenues for primary researches [12] in different aspects which would help to plan tailor-made extension programs. Moreover, this study has identified geographical distribution of clients and queries, therefore mapping of the queries received by AAS, will be possible. Based on that, more ICT tools can be tested in high potential areas for similar purpose. It will also make a good source of information for all the personnel involved in field level extension and advisory work in Sri Lanka. Overall, this study contributes to the academic literature by producing a valuable snapshot of the AAS usage by the Sri Lankan community, as a novel ICT tool for agricultural extension.

## REFERENCES

[1] Aker, J.C. 2011. Dial "A" for agriculture: a review of information and communication technologies for agricultural extension in developing countries. *Agricultural economics*, 42(6) : 631-647.

[2] Steinke, J., van Etten, J., Müller, A., Ortiz-Crespo, B., van de Gevel, J., Silvestri, S.2020. Tapping the full potential of the digital revolution for agricultural extension: an emerging innovation agenda. *International Journal of Agricultural Sustainability*, p 1-17, doi: 10.1080/14735903.2020.1738754

[3] Wen, G., Zetian, F.,, Daoliag, L., Longyong, Y., Jian, Z., Xiaoshuan, Z. 2007. Agrilnfo: An agricultural information system based on a call centre in China. *New Zealand Journal of Agricultural Research*, 50(5): 797-806

[4] Koshy, S., Husain,S., Kumar, K. 2015. Agricultural information delivery mechanism using ICT: A case study from Kerala, India. doi:10.1109/ISTAS.2015.7439413.

[5] Cole, Shawn, A., A. Nilesh Fernando. 2012. The Value of Advice: Evidence from Mobile Phone-Based Agricultural Extension. *Harvard Business School Working Paper*, no. 13–047

[6] Mudannayake, I. 2006. Forging Partnerships between Libraries and Extension Services for Improved Access to Agriculture Information: A case study in Sri Lanka. p.1-9. In Proc. *World Library and Information Congress*, 20-24 August.,2006. Seoul, Korea.

[7] Sivayoganathan, C. and Wirasinghe, S. 1992. Agriculture extension in the food crop sector of Sri Lanka: Past present and future. Jubilee year publication of the faculty of agriculture and the postgraduate institute of agriculture. p.174-195

[8] Wijekoon, R. and Rizwan, M. 2011. Cyber Extension: An ICT Initiative for Strengthening Agricultural Extension in Sri Lanka. *Success Stories on ICT/ICM in AR4D in Asia and the Pacific Region* Retreived from https://d1wqtxts1xzle7.cloudfront.net/51816419/Success_Stories_on_ICTICM_in_AR4D_in_AP_Region_2011-with-cover-page-v2.pdf?Expires=1628625987&Signature=cyzaeTuCHYSnCWyCovsYgDBKhlCtIqTFRCbHbyGt8o18RoCdHdLGz~81QartFOKXPWzgy6Od~RHw5ueL5KNFFQ9tVCn~aLZwPo-SPG-JefTCsaYw6diq2uJ~nmaTr0suQ8WixGTRFNi2fN4ux~3XYfZWy5aaN7T0S4pauG2M4tKknWO-HNloUpe0Ce~6dbpGDNjT3lJFRuJlclLcFpVzF4GE4LoPJ94rPS1KEKj-JofzrslDvaWAdhErrACOX83p0OPor0hBdXpzRMrRMIGtSUNJGd3fgnQOGzfvMRsK92ApJboWVr7fv2lGIuU5TVuwJnR1JQ9O5rM5~xY~BVh~Hw__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA#page=36

[9] Madana, M.H.B.P.H., Dissanayake, D.M.L.B. 2009. Field assessment of toll free information delivery service: Case study in Anuradhapura district of Sri Lanka. p. 341-354. In Proc. *Experiences & Challenges in Agricultural Extension: Meeting Farmer Needs*, 27-28 August., 2009. PGRC Gannoruwa, Peradeniya, Sri Lanka.

[10] Dissanayeke, U. and Wanigasundera, W. 2014. Mobile Based Information Communication Interactions among Major Agricultural

Stakeholders: Sri Lankan Experience. *The Electronic Journal on Information Systems in Developing Countries*, 60(1): 1-12

[11] Kumari, B.S.A, Sivayoganathan,C., Kumara, W.A.G.S. 2009. Effectiveness of toll free agricultural advisory service to vegetable growers in Kandy district. p.355-371. In Proc. *Experiences & Challenges in Agricultural Extension: Meeting Farmer Needs*, 27-28 August., 2009. PGRC, Gannoruwa, Peradeniya, Sri Lanka

[12] Stewart, D.W. and Kamins, M.A. 1993. Secondary research: Information sources and methods. 2nd ed. Sage,USA

[13] QuestionPro. 2021. Secondary Research- Definition, Methods and Examples.Retreived May 14,2021 from https://www.questionpro.com/blog/secondary-research/.

[14] Perera, N., Magamage, M., Kumara, A., Galahitigama, H., Dissanayake, K., Wekumbura, C. 2019. Fall Armyworm (FAW) Epidemic in Sri Lanka: Ratnapura District Perspectives. I*nternational Journal of Entomological Research*, 7(1): 9-18.

[15] Mudugamuwa, I. 2018. Battle against Sena intensifies. Daily News. Retreived May 13, 2021 from https://dailynews.lk/2018/11/10/local/168055/battle-against-sena-intensifies.

[16] Department of Census and Statistics Sri Lanka. 2021. Paddy Statistics. Retreived May 14, 2021 from http://www.statistics.gov.lk/Agriculture/StaticalInformation/rubpaddy .

[17] Thiruchelvam, S. 2005. Efficiency of rice production and issues relating to cost of production in the districts of Anuradhapura and Polonnaruwa. *J. Nat.Sci. Foundation of Sri Lanka*. 33(4): 247-256.

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index No: ES-21-65*

# Evaluation of Cooling Techniques for Personal Cooling Garments in Hot and Humid Climates

D.O.D.P. Fernando (1st Author)
*Department of Textile and Apparel Engineering*
*University of Moratuwa*
Katubedda, Sri Lanka
dodpfernando@gmail.com

A. N. Ahmad (2nd Author)
*Department of Textile and Apparel Engineering*
*University of Moratuwa*
Katubedda, Sri Lanka
ahmd.nabrees@gmail.com

Gayani K. Nandasiri (3rd Author)
*Department of Textile and Apparel Engineering*
*University of Moratuwa*
Katubedda, Sri Lanka
gayanin@uom.lk

Indrajith D. Nissanka (4th Author)
*Department of Mechanical Engineering*
*University of Moratuwa*
Katubedda, Sri Lanka
nissankai@uom.lk

***Abstract*— Personal cooling garments have become a major requirement to cater thermal discomfort, caused mainly due to the rise in global temperatures. Various cooling techniques have been used for specific applications, mainly focused on thermal protection than thermal comfort. Thermoelectric cooling, phase change materials, and evaporative cooling techniques were selected and further analyzed through mathematical modeling to assess their suitability for personal cooling garments. Phase change materials can provide 2 hours of cooling using 46 capsules, while evaporative cooling can provide cooling for 36.1 minutes and thermoelectric cooling can provide continuous cooling. The study concluded that a combination of thermoelectric cooling with phase change materials could remove required heat from the top part of the body which is emitted during medium level activities. Two circuits, each with three Peltier modules of the TEC1-127 series supplied with a direct current of 1.5A and supplied voltage of 9V could provide efficient heat pumping ability, with the coefficient of performance of 1.17 each. The macro capsules with eicosane as the core material act as the heat sink when used in combination with thermoelectric modules. Hence, it concludes that a combination of thermoelectric modules with phase change materials integrated into a personal cooling garment could effectively provide the required thermal comfort.**

***Keywords – personal cooling garments, thermal comfort, phase change materials, thermoelectric cooling***

## I. INTRODUCTION

The current temperature increment due to global warming is a raising environmental issue that highly attracted the focus of almost all the nations. It is also stated by NASA that the highest temperature was marked in 2020 [1]. Due to this temperature increment, thermal comfort of humans was compromised and the heat related illnesses also increased, along with the raise in the death percentages [2]. The effect of global warming and the raise of temperature can have more effect on tropical countries compared to others. The recent COVID-19 pandemic is another severe problem for the frontline medical staff who are required to wear a sealed type of clothing for longer durations, which are mostly made with polyolefin fibres that increase thermal discomfort. Human thermal comfort depends on four environmental parameters: ambient and radiant temperature, relative humidity and air velocity and two human parameters: activity level and clothing insulation [3].

In general, a centralized cooling system is used in commercial buildings and split type air conditioners into residential units. However, the cooling effect will only be limited to the indoor environment, and it is also reported that the cooling efficiencies are comparatively lower with regards to the high energy consumption [4]. Therefore, considering the above limitations, personal cooling garments (PCG) will provide the best solution to human thermal comfort. It is also stated that PCGs can reduce up to 25% of the energy consumed for Heating, Ventilation, and Air Conditioning (HVAC) [5].

During the last nine decades research on protective garments used in high temperature environments, paving the way towards the introduction of cooling techniques namely air cooling (AC), liquid cooling (LC), phase change materials for cooling (PCM), and evaporative cooling techniques (EC). These are also known as the traditional cooling techniques which are used for garments in the fields such as military, sports, healthcare, firefighting and industrial applications. Studies on PCGs are still in progress and providing new cooling techniques such as the enhanced conductive cooling with boron nitride (BN), vacuum desiccant cooling (VDC), which is also an advanced evaporative cooling and thermoelectric cooling using Peltier effect (TEC). The

effectiveness of these cooling techniques varies depending on the climate conditions they are being used. Thus, precise selection has to be made when it comes to the selection of a suitable cooling technique for PCGs.

Thermal comfort is a subjective sensation, and therefore the individual requirement towards it may also differ. Hence, providing the individual requirement on thermal comfort is possible by the use of PCGs. Considering the impacts of global warming, the limitations of centralized cooling and the capabilities of PCGs, it could be envisaged that a PCG with precise cooling technique/ techniques can create a new market segment in technical textiles which has a high potential for future growth of wearables and smart textiles.

In normal circumstances energy from the human body is transferred to the environment in different modes. Accordingly, out of the total energy transferred, more than 70% is through convection and radiation, 2% through respiration and conduction, and the remaining percentage by the evaporation of sweat [6]. Even though the heat is getting transferred to the environment, there is a huge percentage of heat getting stored as thermal energy in the microclimate between the body and the clothing contributing to increase in temperature. The functionality of the PCG is primarily to remove the stored thermal energy while providing the required thermal comfort for personal wear in daily activities. The main heat transfer modes can be described mathematically as stated by Equation (1) the Fourier's law of heat conduction, Equation (2) the convection by Newton's law of cooling, and Equation (3) the evaporative heat transfer.

$$q_{conduction} = -K\frac{dT}{dx} \quad (1)$$

$$q_{convection} = h(T_\tau - T_\infty) \quad (2)$$

$$q_{evaporation} = \varepsilon(P_\tau - P_\infty) \quad (3)$$

where, T is the temperature and K is the thermal conductivity of the material, $h$ is the convective heat transfer coefficient, $T_\tau$ is the outer surface temperature of the clothing and $T_\infty$ is the temperature of the ambident environment, $\varepsilon$ is the evaporative heat transfer coefficient, $P_\tau$ – partial pressure of the ambient water vapor and $P_\infty$ - partial pressure of the water vapor at the skin.

Even though there are different cooling techniques used for specific heat protection applications as mentioned above, out of those, only a few could effectively be used for PCGs. The air cooling garments (ACG) functions with the circulation of the air through the porous tubing arrangements. The cold air gets blown to the body through the pores and increases the convective heat transfer. Similarly, the liquid cooling garments (LCG) functions by circulating the coolant or liquids in place of air in ACGs. However, the tubing arrangement is of sealed type restricting any leakages. This liquid circulating around absorbs the energy stored in the microclimate through conduction. Meanwhile, garments incorporating PCM (PCMG) functions according to the theory of latent heat where the temporary cooling is provided during the phase change: energy is transferred from human body to the PCM via conduction. Similarly, in evaporative cooling garments (ECG) the cooling is due to the result of evaporation of the liquid which is mostly water, where the liquid absorbs the energy in the microclimate and thereby change its phase to vapour.

Considering the novel cooling techniques, conductive clothing can enhance efficient heat removal from the microclimate through BN/PVA composite [7], also VDC can function more efficiently in any environmental condition [8]. The TEC is a technique, where the temperature difference can be created when the current goes through the module (Peltier effect) and provide cooling [4]. A critical comparison of these cooling techniques for the suitability of use in a PCG, by considering the requirements in a daily wear are described in Table 1. The AC and LC techniques have shown the highest cooling efficiencies during the long time of wearing opposed to other traditional cooling techniques. Nevertheless, the flexibility and freedom of movement, low weight and less bulkiness which are the major requirements in PCGs , get compromised in AC and LC techniques due to the auxiliary equipment used to obtain its functionality [6]. a similar difficulty is also observed in VDCs as the thickness and bulkiness of cooling pads used in VDC are high.

This shows that the AC and LC techniques are not suitable for the PCG even though they provide high cooling efficiencies in long duration of wearing. The PCGs also need to have the properties similar to a garment that is used as daily wear and these are worn in normal environments for a short period of time, unlike the garments worn in hostile environment. Similarly, the bulkiness and the resistance to movements of VDCs makes it less desirable for PCGs. The remaining four techniques listed in Table 1 can be selected for the PCGs, out of which enhanced conductive cooling with BN/PVA is still under research. Also, with BN/PVA it is possible that there are more chances in tropical countries where the radiant temperature gets increased more than skin temperature resulting the reverse heat flow from the environment to the human body.

Focusing on PCM, the main limitation is related to the duration of cooling. To provide a longer duration of cooling, it required to incorporate larger mass of PCM into the PCG. Researchers have used microcapsules, macro capsules and macro packets to incorporate PCMs to the garments. Out of these, the mostly used type is the microcapsules. Nevertheless, due to the micro scaled size of the capsule, it was not an effective method of cooling for longer durations due to the less amount of PCM mass in the garment. As a solution, to incorporate more PCM in the garment, some researchers have used macro packets of PCM, which resulted in increase in weight and limited the freedom of movement[9]. Considering these limitations, the new area of research have focused on the macro capsules which are in the scale of 3mm – 5mm [10]. Usage of these macro capsules can have higher mass of PCM compared to the micro capsules

thus support the longer duration of cooling. The other major issue with respect to PCM is the durability of the capsules. Due to the external forces, it was found that there are more chances for the capsules to get damaged.

Table 1. Comparison of cooling techniques

| **Description** | **ACG** | **LCG** | **PCMG** | **ECG** | **BN/PVA Composite** | **TEC** | **VDC** |
|---|---|---|---|---|---|---|---|
| Requirement of external equipment | For the circulation of air | For the circulation of liquid | No | No | No | For cooling purposes and to power the modules | For cooling |
| Uniqueness | No | Highest efficiency and best for longer durations | Quick response & high efficiency for short term cooling | Best suitability for less humid environment | Only technique to use conductive heat transfer efficiently | Shortest response time and quickest cooling | Functions in any environment |
| Major limitation to be used for cooling in PCG | High weight, bulkiness due to external elements | High weight, bulkiness due to external elements | Duration of cooling time, and durability of the PCG | Not suitable in high humid environment | Under research, heat from environment can transfer towards the body.. | Current flow in the garment | Bulkiness and the increased thickness of the pads. |

Considering this problem, the researchers have investigated the strength and durability with the different core to shell ratios. With the increase of the core to shell ratio, the encapsulation efficiency is not as expected due to possibility of incomplete coating [11]. Nevertheless, in decreasing the ratio, the heat capacity is also compromised. Thus the best core to shell ratio suggested in the literature is 3:1 [12]

TEC is a new cooling technique when it comes to textiles, the problem related to this method is possibility of current leakages. It is also stated the temperature difference between the two sides of the TEC depends on the current that flows through it [4], [13]. Therefore, a precise calculation on the current that can provide cooling without harm, need to be carried out and accordingly the suitable module should be selected [4].

Focusing on the evaporative cooling, the functionality highly depends on the relative humidity of the environment [14]. Therefore, it can efficiently function in lower humid environments. For the tropical countries where the humidity is higher [15], this technique is not highly efficient.

The critical review of the available techniques suggested the possibility to use three main techniques for cooling garments, as listed below.

- PCM technique
- Liquid Mediate EC technique
- TEC using Peltier modules

## II. OBJECTIVES

Hence, this study focuses on qualitative and quantitative analysis of different cooling techniques, and their combinations to evaluate the suitability to be used in personal cooling garments for hot and humid climatic conditions. This study aims to achieve following objectives,

- to review different cooling techniques and evaluate their effectiveness for the use in personal cooling garments,
- to evaluate the effectiveness of combined cooling techniques for personal cooling garments, using analytical/ numerical methods,
- to propose suitable combination of cooling techniques for personal cooling garments for hot and humid climates.

## III. METHODOLOGY

With the identification of the cooling techniques, that can be incorporated in the cooling garments through the review, separate mathematical models were developed, for the combined cooling techniques by adopting the available heat transfer equations for individual cooling techniques by considering the enhanced performance of each technique separately.

### *A. Human thermoregulation*

According to [3], thermodynamics between the environment and the human body is described with the steady-state energy balance model, where heat transfer between the body and the environment can be expressed using Equation (4).,

$$0 = M_r - W_r - C_{res} - E_{res} - C_b - R_b - E_b - S_t \quad (4)$$

where, $M_r$ - metabolic rate, $W_r$ - work rate performed on the environment, $C_{res}$ - convection respiratory heat transfer, $E_{res}$ - latent respiratory heat transfer, $C_b$ - body convection heat transfer, $R_b$ - body radiation heat transfer, $E_b$-body evaporative heat transfer, $S_t$ - heat storage rate by the body. The heat storage $S_t$ in Equation (4) defines the temperature of the microclimate between the body skin and the fabric. The thermally comfortable skin temperature in tropical countries varies in the temperature range of 304.5K to 305.5K [16].

### B. Mathematical representation of PCM technique

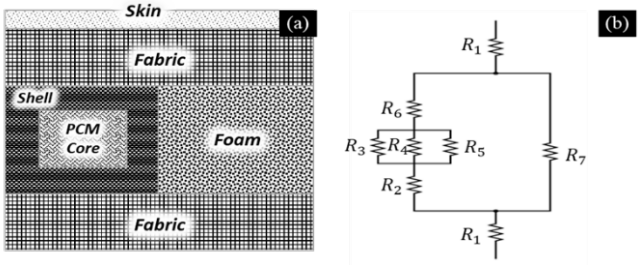



Fig. 1 – Model of a garment incorporated with PCM (a) Schematic representation (b) Equivalent thermal resistance diagram. (R1 -6, represent the thermal resistance of each component).

Mathematical equations were derived from the basic heat transfer equation available in the literature, for a sandwiched structure that contains PCM macro capsules embedded in flexible foam material in between two fabric layers. The heat stored in the microclimate gets reduced during the phase changing period of PCM, and during this period the wearer feels comfortable. The proposed mathematical model, for the heat transfer calculations of this system, was developed based on the schematic presented in Fig. 1.

During the heat transfer calculations, the following assumptions were made,

- The phase change of PCM occurs uniformly (No temperature gradient and moving boundary problem)
- The spherical capsules are assumed to be cube shaped to reduce the complexity of modelling.
- The heat transfer is assumed to be in one dimension (perpendicular to the fabric surface).

Therefore, the total heat flow per unit area $Q_{Tot}$ through the design could be calculated as,

$$Q_{Tot} = \frac{(T_b - T_e)}{R} \quad (5)$$

where, $T_b$ - Temperature of the body, $T_e$-Temperature of the environment, $R$ - equivalent thermal resistance of resistors $R_1$ to $R_7$, where $R_2 = R_6$ and $R_3 = R_5$ as shown in Fig. 1.

Substituting the expression for $R$ in the Equation (5), the $Q_{Tot}$ could be expressed as;

$$Q_{Tot} = \frac{2(T_b - T_e)\{R_1[(R_3+2R_4)(2R_2+R_7)+R_3R_4]+2R_2R_7(R_3+2R_4)+R_3R_4R_7\}}{(R_3+2R_4)(2R_2+R_7)+R_3R_4} \quad (6)$$

The time period of cooling t can be obtained by Equation (7),

$$t = n * \frac{V\rho L_f}{Q_s} \quad (7)$$

where, $n$ - no. of capsules required, $V$ - volume of a capsule, $\rho$ - density of the core material, $Q_s$ - heat flux to the core and $L_f$ - latent heat of fusion.

### C. Mathematical representation for water-mediate EC technique

Fig. 3 shows the schematic representation and the illustration of the thermal resistance for the water-mediated EC model. The sweat evaporated from the skin and the evaporation of moisture present on the fabric contributes to the cooling effect produced. The fabric layer consists of a three-layered structure as shown in Fig. 2.

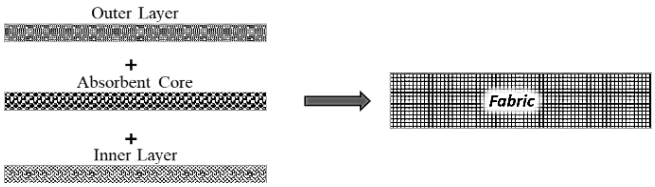



Fig. 2. – The model representing the sandwiched fabric layer.

This cooling technique functions after the liquid water is introduced to the absorbent core. The cooling effect is created by the evaporation of the liquid water which occurs at any temperature, if the air is unsaturated. Accordingly, when the fabric is worn near to the skin, the energy stored within the microclimate is absorbed, which results in evaporation. Evaporation of one liter of water takes away approximately 2400 kJ, which can effectively reduce the heat stress as described in Equation (4).

During the model formulation, the following assumptions were made

- The three-layered fabric has been modelled as one fabric layer.
- The system behaves in isothermal conditions.
- The fabric surface is covered by a continuous water film.
- The partial pressure of water vapor at the skin surface reaches the saturated level.

The first assumption to represent a 'three-layered fabric as one layer of fabric' was due to the fact that the inner layer and the outer layer do not contribute to holding water, thus, the effect created towards the water film is negligible.

Based on the literature, the heat flow that causes the skin cooling $q_{fabw,sk}$ can be expressed as,

$$q_{fabw,sk} = \frac{\beta(P_{sat,fab} - P_{air})}{(1 + \alpha R_{ct}(1 - kU) + \alpha R_{gap})} \quad (8)$$

where, $\beta$ - convection mass transfer coefficient, $P_{sat.fab}$ - saturated water vapour pressure on the fabric surface, $P_{air}$ - water vapour pressure of the environment, $\alpha$ – convective heat transfer coefficient which increases with air velocity, $R_{ct}$ - the thermal resistance of fabric in the ultra-dry state, $k$ - experimentally determined constant characterizing the decrease of thermal resistance caused by the increased moisture of the fabric and $U$ - relative mass increase of the fabric with moisture content (%), $R_{gap}$ - evaporative resistance of the air layer.

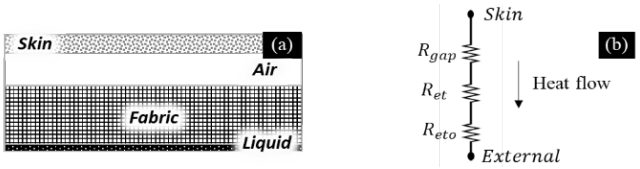



Fig. 3 - Model of a water-mediate EC garment (a) Schematic representation (b) Equivalent resistance diagram ($R_{gap}$, $R_{et,}$ $R_{eto}$, represent the thermal resistance of each component).

The heat flow coming from the skin $q_{skin}$, can be described by Equation (9),

$$q_{skin} = \frac{P_{sat} - P_{air}}{(R_{gap} + R_{et} + R_{eto})} \quad (9)$$

where, $P_{sat}$ - saturated vapour pressure in the skin surface, $R_{et}$ - experimentally determined constant for evaporative resistance of the fabric, $R_{eto}$ - evaporative resistance of the boundary layer.

The total heat flow $q_{tot}$ through the fabric surface can be calculated by the addition of Equations (8) and (9).

$$q_{tot} = \frac{P_{sat} - P_{air}}{R_{gap} + R_{et} + R_{eto}} + \frac{\beta(P_{sat.fab} - P_{air})}{1 + \alpha R_{ct}(1 - kU) + \alpha R_{cgap}} \quad (10)$$

*D. Mathematical model for TEC technique*

Fig. 4 shows the schematic representation and the representation of the thermal resistance circuit created for the model on Peltier arrangements. According to the model, one Peltier element is considered, which is sandwiched in between two fabric layers. A flexible Peltier is considered since it is worn closer to the skin in a personal cooling garment, and therefore it should support the body movements.

The Peltier effect occurs if and only if a current passes through the thermoelectric module. Due to this current, heat energy gets absorbed from the connective points of thermoelectric pillars to the external circuits. Since the heat gets absorbed in these positions, the temperature gets reduced. The absorbed heat flows through the doped *p* and *n* type semiconductors and gets released at the next end, therefore the temperature in those places gets increased. The resultant cooling will depend on the balance of the cooling power, heating power, and heat produced due to the current flow (due to charge carrier movement), etc. Therefore, according to the literature [20], resultant cooling of a Peltier module can be described using Equation (11),

$$Q_{RC} = IT_c(\alpha_P - \alpha_n) - (T_H - T_C)(K_p + K_n)\left(\frac{A}{l}\right) - \frac{I^2}{2}(R_p + R_n) \quad (11)$$

where, $I$ - current through the Peltier, $\alpha_P$ & $\alpha_n$ - Seeback coefficients of branches and $T_c$ - temperature of cooler side of Peltier module, $T_H$ - temperature of the hotter side of the Peltier module and $K_p$ & $K_n$ - thermal conductivity of the branches, $A$ - cross-section of a leg, $l$ - height of leg in the module, $R_p$ & $R_n$ - Thermal resistance of the branches

For the cooling to occur, the power supplied to the Peltier should be large enough to cater the Joule heating and to create the temperature difference. Accordingly, Equation (12) describes the required power $W$ for the Thermoelectric module to create the Peltier effect [20].

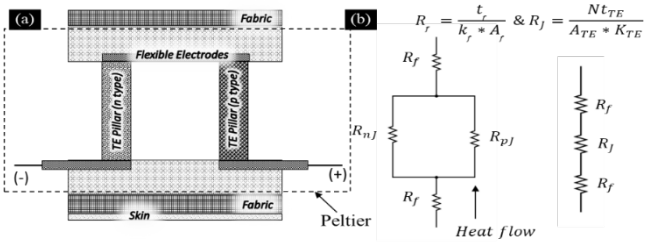



Fig. 4 – Model of TEC using Peltier - Modules (a) Schematic representation (b) Equivalent resistance diagram ($R_f$, $R_{pj,}$ $R_{nj}$, $R_j$, represent the thermal resistance of each component).

$$W = (\alpha_P - \alpha_n) * I * (T_H - T_c) + \frac{I^2}{2} * (R_p + R_n) \quad (12)$$

Considering the function of the Peltier modules, it functions as the heat pump and therefore the term efficiency is not appropriate. Based on the second law of thermodynamics the suitable term to assess the performance of the Peltier is the coefficient of performance (COP) and can be calculated using Equation (13).

$$COP = \frac{Q_{Res\ Cooling}}{W} \quad (13)$$

## IV. RESULTS AND DISCUSSION

The results of this study fall under two categories: the analytical model results, and the suitability of the cooling techniques, which have been carried out with the physical boundary conditions suited for a high humid high temperature environment.

*A. Results of the Mathematical analysis*

*1) Results of the study of the PCM cooling*

According to the model described in Fig. 1, in an environment with a temperature in the range of 303K to 306K, nonadecane can be used as the PCM core encapsulated by melamine formaldehyde as the shell material. The core to shell ratio of the PCM capsule used was 3:1 with the length and width of 3 mm on each side. These capsules were included in the flexible polyurethane foam which gets sandwiched by high density polyethylene (HDPE) fabric. The analysis was made with relation to a repeat unit where the fabric thickness considered was 0.5 mm.

Thermal conductivity values of HDPE, polyurethane foam, and melamine formaldehyde were 0.481, 0.028, 0.5$Wm^{-1}K^{-1}$, respectively. The core of the macro capsule has the latent heat of fusion of 222 kJ/kg. According to the analysis with the one-dimensional heat flow, it was found that the heat that passes through the shell material of a PCM capsule was 5.197 mW, when temperatures $T_b$ = 306K and $T_e$ = 303K. Furthermore, it was calculated that 46 macro capsules of diameter 3mm can provide 2 hours of sufficient cooling.

However, it was found that the temperature of the skin was maintained at 305K (the phase change temperature of the nonadecane) during the long period of wearing. Therefore, with respect to the tropical countries where the radiant temperature is higher, selected PCM alone can only provide the cooling with limited temperature reduction in the microclimate

*2) Theoretical results based on the study of the water mediate evaporative cooling technique*

Considering the environmental parameters in high temperature and high humid environments the following were used for the analysis,

- the environmental temperature is $304 \pm 3\ K$ [17],
- relative humidity varies between $60\ \% - 100\ \%$ [17],
- wind speed varies between $2\ ms^{-1} - 5\ ms^{-1}$ [18].

At the above stated environmental parameters, considering a three-layered fabric with 20ml of water absorbed in the core layer, the calculation results showed that the total heat flux going through the fabric was around 330 W/m$^{-2}$. Furthermore, it was found that the composite structure was able to cool for 35 minutes through the evaporation of the water available in the middle layer.

*3) Results based on the study of thermoelectric cooling using Peltier modules.*

When a person is performing a light activity, the energy emitted by the body to the environment is 87W/m$^2$ [3]. The average surface area of the human body is 1.8m$^2$ and it is also stated that the upper back part of the body between the neck and waist, emits a larger amount of heat to the environment and has a surface area of 0.3m$^2$ [19]. Considering the function of a TEC as a heat pump/refrigerator, by pumping the heat emitted by the body which is stored on the microclimate (colder side) to the environment (hotter side), thermoregulation can be achieved. To pump the heat emitted by the body to the environment, a Thermoelectric module (model no: TEC1-12706), with its parameters such as conductivity of branches $K_{TE} = 1.50 Wm^{-1}K^{-1}$, Seeback coefficients $\alpha_m$ = 0.200VK-1, the temperature of the hotter side $T_H$ = 300K, resistance $R_{module}$ = 1.98Ω was considered for the analysis.

According to the calculations carried out, to cater the stated requirement with achieving a temperature difference of 10K between the two sides of the module, using 6 Peltier modules with 3 in one circuit is more realistic. It is calculated that 1.5A current with a supply voltage of 9V for one circuit will be required. The calculated COP of one module was 1.17 for this mentioned situation as calculated from (13).

### B. *Suitability of these cooling techniques in hot and humid environments.*

All these cooling techniques are suitable to be used in a PCG. A comparison of the theoretical calculations made based on models derived from heat transfer principles discussed in section B, are summarised and tabulated in Table 2, which depicted that PCM could provide 2 hours of cooling by simply having 46 macro capsules. This has the advantage as the capsule can be restored back to its initial state with the reduction of temperature. Thus, there is no requirement of any additional inbuilt energy source to be incorporated into the PCG. Nevertheless, the temperature reduction during the cooling period by PCM is lesser compared to the other cooling techniques. The water mediate evaporative cooling technique also does not require any additional elements to power or any additional energy sources. As shown in Table 2, this can provide cooling for 36 minutes by simply evaporating 20 ml of water. However, this evaporation is dependent on the relative humidity and in humid environments, the effect created by this cooling technique is comparatively low. Thermoelectric cooling using Peltier modules is an advanced method that uses current, and can reduce the temperature to any required temperature range within a very short time period. The heat can be pumped effectively out to the environment even by using small currents and voltages, yet with a higher COP value. Nevertheless, the disadvantage of this method is that at lower temperature differences created, the back heating gets more prominent. Also, looking at the combinations possible, it is clear that there are three combinations, where, the TEC and EC methods are impossible as it uses water and electricity together. Comparing the other two combinations, the PCM and EC methods could provide lesser cooling compared to the combination of the Pelter & PCM technique. The back heating effect created by the TEC could be absorbed by the PCM, which can act as a heat sink when eicosane is used. As the phase change temperature of the material is 309K which is closer to the temperature of the hot side of the Peltier, it can provide an efficient heat sinking process. Using more than one Peltier and placing them on the backside of the cooling panel could help in removing larger amounts of heat as well as to avoid effects due to the back heating, thus providing the optimum cooling.

Table 2. Analysis of cooling techniques

| Cooling technique | Dependent | Cooling duration |
|---|---|---|
| PCM | 46 capsules | 2 hours |
| EC | 20 ml | 36.1 minutes |
| TEC | 1.5A/3V | continuous |

## V. CONCLUSION

The temperature recording by NASA and the consensus related to the heat related illnesses shows that there is no proper thermal comfort achieved in clothing globally and when it comes to the tropical countries this effect is severe. This proves the requirement for a personal cooling system in clothing since it is considered as the second skin. The requirement of PCG was further proven by the survey conducted in Sri Lanka. Hence this study investigated the suitability of different cooling techniques and their combinations to use in personal cooling garments in tropical countries.

To provide thermal comfort via PCG, the qualitative analysis made on the available cooling techniques suitable for textiles concluded that cooling using Phase Change Material

(PCM), Evaporative Cooling (EC) and Thermoelectric Cooling (TEC) was found to be more effective in both cooling and to cater the requirements from the day-to-day wear. The mathematical analysis of these three techniques showed that TEC provided continuous cooling with a 10 °C temperature difference between the two sides of it, while having problems with the back heating as it functions as the heat pump. PCM was able to provide 2 hours of cooling with 46 capsules, although the temperature reduction it can provide in the microclimate was not adequate. EC was having a limitation in cooling due to the barrier caused for the water to evaporate in high humid environments. Therefore, considering the pros and cons of each of these three techniques, it is recommended that the combination of PCM & TEC can provide optimum cooling for a PCG in tropical climates. These results obtained from the mathematical analysis are further required to validate with experimental evidence, which would be the next phase of this research. Also, the final garment design should consider other factors such as heaviness of the cloth, energy consumption etc.

## REFERENCES

[1] "Global Temperature | Vital Signs – Climate Change: Vital Signs of the Planet." https://climate.nasa.gov/vital-signs/global-temperature/ (accessed Feb. 27, 2021).

[2] "WHO | Information and public health advice: heat and health." https://www.who.int/globalchange/publications/heat-and-health/en/ (accessed Feb. 27, 2021).

[3] ASHRAE-55, "Thermal Environmental Conditions for Human Occupancy," Encycl. Financ., pp. 227–227, 2010, doi: 10.1007/0-387-26336-5_1680.

[4] S. Hong et al., "Wearable thermoelectrics for personalized thermoregulation," Sci. Adv., vol. 5, no. 5, 2019, doi: 10.1126/sciadv.aaw0536.

[5] J. Wu et al., "Flexible and Robust Biomaterial Microstructured Colored Textiles for Personal Thermoregulation," ACS Appl. Mater. Interfaces, vol. 12, no. 16, pp. 19015–19022, 2020, doi: 10.1021/acsami.0c02300.

[6] M. Yazdi and M. Sheikhzadeh, "Personal cooling garments: a review," J. Text. Inst., vol. 105, no. 12, pp. 1231–1250, 2014, doi: 10.1080/00405000.2014.895088.

[7] B. Yang et al., "Three-Dimensional Printed Thermal Regulation Textiles," ACS Nano, vol. 11, no. 11, pp. 11513–11520, 2017, doi: 10.1021/acsnano.7b06295.

[8] Y. Yang, "Vacuum Desiccant Cooling for Personal Heat Stress Management," 2016.

[9] M. Mokhtari Yazdi and M. Sheikhzadeh, "Personal cooling garments: a review," J. Text. Inst., vol. 105, no. 12, pp. 1231–1250, 2014, doi: 10.1080/00405000.2014.895088.

[10] D. Colvin, "Body heat stress measurements with MacroPCM cooling apparel," ASME Int. Mech. Eng. Congr. Expo. Proc., pp. 37–44, 2002, doi: 10.1115/IMECE2002-33332.

[11] A. Hassan, M. S. Laghari, and Y. Rashid, "Micro-encapsulated phase change materials: A review of encapsulation, safety and thermal characteristics," Sustain., vol. 8, no. 10, 2016, doi: 10.3390/su8101046.

[12] J. Giro-Paloma, C. Alkan, J. M. Chimenos, and A. I. Fernández, "Comparison of Microencapsulated Phase Change Materials prepared at laboratory containing the same core and different shell material," Appl. Sci., vol. 7, no. 7, 2017, doi: 10.3390/app7070723.

[13] R. A. Kishore, A. Nozariasbmarz, B. Poudel, M. Sanghadasa, and S. Priya, "Ultra-high performance wearable thermoelectric coolers with less materials," Nat. Commun., vol. 10, no. 1, pp. 1–13, 2019, doi: 10.1038/s41467-019-09707-8.

[14] Y. Yang, G. Cui, and C. Lan, "Developments in evaporative cooling and enhanced evaporative cooling - A review," Renew. Sustain. Energy Rev., vol. 113, no. June 2016, p. 109230, 2019, doi: 10.1016/j.rser.2019.06.037.

[15] A. Nabrees, D. O. D. P. Fernando, G. K. Nandasiri, and I. D. Nissanka, "A study of thermal comfort in clothing and the need for personal cooling garments in Sri Lanka," vol. 1, no. NERS, pp. 149–154, 2020.

[16] A. Gagge, J. Stolwijk, and J. Hardy, "Comfort and thermal sensations and associated physiological responses at various ambient temperatures," Environ. Res., vol. 1, no. 1, pp. 1–20, 1967, doi: 10.1016/0013-9351(67)90002-3.

[17] "Department of Meteorological Sri-Lanka, Climate of Sri-Lanka," 2021. www.meteo.gov.lk.

[18] W. L. S. Maduranga and C. S. Lewangamage, "Development of Wind Loading Maps for Sri Lanka for use with Different Wind Loading Codes," Eng. J. Inst. Eng. Sri Lanka, vol. 51, no. 3, p. 47, 2018, doi: 10.4038/engineer.v51i3.7305.

[19] P. Tikuisis and D. Canada, "Human body surface area : Measurement and prediction using three dimensional body scans," no. September 2001, 2014, doi: 10.1007/s004210100484.

[20] Z. Ouyang and D. Li, "Modelling of segmented high-performance thermoelectric generators with effects of thermal radiation, electrical and thermal contact resistances," Scientific Reports 2016 6:1, vol. 6, no. 1, pp. 1–12, Apr. 2016, doi: 10.1038/srep24123.

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index No: ES-22-72*

# Development of an Autonomous Agricultural Drone and Real - Time Monitoring Interface for Aerial Seeding

M.R.Mohamed (1st Author)
*Department of Applied Computing*
*Faculty of Computing and Technology*
University of Kelaniya
Kelaniya, Sri Lanka
rushdimohamed09@gmail.com

D.M.T Priyankara (2nd Author)
*Department of Applied Computing Faculty of Computing and Technology*
University of Kelaniya
Kelaniya, Sri Lanka
tpriyankara10@gmail.com

C.N.Kaluwandara (3rd Author)
*Department of Applied Computing*
*Faculty of Computing and Technology*
University of Kelaniya
Kelaniya, Sri Lanka
chanukan60@gmail.com

S.Bandara (4th Author)
*Department of Applied Computing*
*Faculty of Computing and Technology*
University of Kelaniya
Kelaniya, Sri Lanka
shakya@kln.ac.lk

***Abstract* - Applications of Unmanned Aerial Vehicles (UAVs) or commonly known as drones, have been considerably growing during the past few decades in various fields such as agriculture, aerial mapping, surveillance, disaster management, and military applications. Drones are used in agriculture to monitor, harvest, and apply pesticides and fertilizers [1]. An autonomous agricultural drone (quadrotor) was developed for aerial seeding and a web interface using IoT technology with a raspberry pi microcontroller. A low-cost and open-source flight controller was used [2]. The real-time unmanned aerial vehicle monitoring system was developed using PHP/MySQL as the backend technologies on the website and Bootstrap, AJAX, JQuery for the frontend and to communicate with the Pixhawk and Raspberry Pi, Python was used. To enhance the continuous flying time with load, proper weight calculations were done before it flies with the load. CAD software was used to design the drone from the top, bottom, right and left views with exact calculations [3]. The design calculations, along with thrust values, also calculated for payload.**

***Keywords* - *Agriculture, Drone, IoT, Pixhawk, Raspberry Pi, RPA, UAV, Quadrotors***

**Abbreviations: BEC (Battery Elimination Circuit); *BLDC, Brushless Direct Current Motor; CAD, Computer-Aided Design and drafting; ESC, Electronic Speed Controller; IoT, Internet of Things; RPA, Remotely Piloted Aircraft;* UAV, *Unmanned Aerial Vehicle.***

## I. INTRODUCTION

UAVs or UAS and RPAs popularly known as drones. Normally these referred, as UAVs, are principally connected with the military, industry, and other specific activities, yet with ongoing improvements in the space of sensors and data innovation over the most recent twenty years the extent of drones has been augmented to different areas like Agriculture [4]. UAV will soon be an important applicable tool for conservation and revegetation practices. A drone can provide efficient and effective methods and low-cost and low-impact solutions to environmental managers working in a variety of ecosystems. Their agility, image quality, and logistic abilities make them become valuable tools. Cutting-edge drones are getting more astute by incorporating open-source innovation, intelligent sensors with better mixes of the most recent invention and more flight time [5]. These drones can have used for a vast range of applications. Here, some of them reported as forest-fire fighting, police surveillance and protection, environmental factors, such as radiation and infectious diseases, monitoring and control, search and rescue missions, ecological research studies, oil and gas industry security. In addition, mapping and surveying, weather monitoring, seismic and geothermal monitoring, to deliver medical supplies and products in times of critical demands across both accessible and inaccessible or dangerous locations. and other calamities the board applications [6].

Concerning the agricultural economic sector, replacing large and unwieldy aerial or land vehicles for UAVs in agricultural operations is now a reality. The use of UAVs in advanced agriculture meets Eco-innovation requirements helping to reduce the negative impact that agricultural

activity causes in ecosystems, in particular by reducing the use of fossil fuels and their replacement by electric energy that can be provided from renewable sources [7][8]. UAVs present greater operating flexibility, lower initial and operating investment costs, reduced size and increased profitability by reducing expenses related to the reduction of human resources since a single operator can control several simultaneous UAVs in a safe and comfortable way [9][10].

The drone has been applied in many agricultural operations such as the prediction of fungi appearing and movements weed mapping and control, application of fertilizers, insecticides, pesticides, fungicides spraying, or other spraying chemicals. This research intended to extend the application of drone technology in agriculture by proposing an autonomous drone for aerial seeding and customized web interface. The proposed drone consists with an on-board microcomputer to communicate with the flight controller and the web interface [11][12]. As compared to existing unmanned vehicles, this drone has the capability to real-time monitoring and analyze visually the data in a later time to understand the performance and the behavior on a web portal which could be accessed via a smartphone [13].

Fig. 1. Completed drone attached with the Raspberry Pi, seed dispenser and the transmitter.

## II. OBJECTIVES

Development of an autonomous drone with a seed dispenser to hold the seeds and drop the seeds accordingly. Also connecting the drone with a microcontroller such as raspberry pi to communicate with the internet and display analytics of the drone via a web portal with real-time drone path tracking.

## III. MATERIALS AND METHODS

The drone plays a significant role in the project, and the drone designed according to the X-frame type since the X-frame type provides more balance. To align with the project's objective, the drone must have the ability to lift heavy weights and be portable and light (which are some attributes of X-frame type drones) [14]. This section describes the peripherals, application software, and seed dispenser unit used to conduct this project. Following Fig. 2. is the detailed CAD drawing of the drone with the seed dispenser.

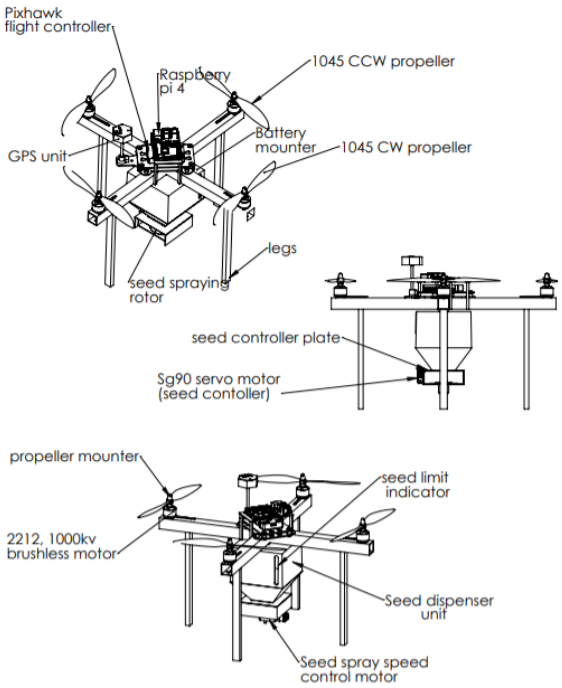



Fig. 2. Detailed drone with the seed dispenser

### A. *Components selected for the drone*

Four brushless motors (A2212/13T 1000Kv), four electronic speed controllers (40A), four 10inch carbon fiber propellers with 45-degree pitch angle, Pixhawk flight controller with SE 100 GPS unit and 11.1V, 5500mAh Lithium polymer battery. Also, X-type drone frame, Raspberry pi 4, 4G dongle to provide the internet connection for the Raspberry Pi Microcontroller, one servo motor(5v) and one DC gear motor(5v), seeds dispenser unit, seed dispenser unit mounter [15].

### B. *Hardware Implementation*

An aluminum bar used in order to minimize costs. In addition, the landing gear been made with plastic that will be used to land the machine softly and spread the landing force over the body. In addition, there are two aluminum bars that were been linked together in the X-shape and on top of the bars in each corner, and a brushless motor was mounted. The middle part of the body contains all the payloads (ESC, Controller, RF receiver, battery and mobile device) [16][17].

### C. *The Drone Thrust Calculations*

The total weight and the thrust calculations done to make sure the user could have the drone under proper stability.

Table 1. Weights of the individual components.

| Drone part | Weight |
|---|---|
| Drone frame | 200g |
| 5500mah battery | 288g |
| Flight controller | 40g |
| 40A ESC (4 pcs) | (38 * 40) 152g |
| A2212 1000kv motor (4 pcs) | (64 * 4) 256g |
| Seeds caring unit | 200g |
| Raspberry Pi 4 | 50g |
| Other payload | 100g |
| Total | 1286g |

When calculating the thrust of the drone, two times of the drone's weight used to hover the drone in the air by half thrust, which enables the remaining thrust for stability and control movements of the drone. [18]. Therefore, the calculation is for the thrust is as follows,

$$T = 2 * (Wd + Wpl) \quad (1)$$

Where, (T) thrust, (Wd) weight of the drone, (Wpl) payload.

$$Total\ trust = T + [T * 0.20] \quad (2)$$

According to the calculation, the required thrust is 2572g to get it off the ground. An additional 20 % (514 grams) added to that total to ensure that the drone could hover. After dividing the total thrust by the number of motors, the drone requires.

771.6g (3086.4/4 = 771.6) of thrust per motor.

Drone weight = 1286g

Double weight with seed dispenser = 2572g + 514g

When dividing it into four motor = 3086.4/4

So, one motor trust =771.6g

So A2212 1000kv motors with 1045 propellers and 40A ESC have been used for each motor [18].

(Each motor can carry 800g approximate weight, finally, able to get 800*4 = 3200g trust).

### *D. Assembling*

As the first step, the project developed by using a designing tool (i.e., engineering drawing software). Then the box bar cut and assembled in the shape of "X" to create the drone frame. The speed controllers and the four motors nailed to the box bar to tighten the grip. After that, the flight controller and other related GPS units installed to the frame. After finishing the wiring works flight controller was programed and run pilot tests.[14]. In addition, the following Fig. 3. shows the assembled drone.

Fig.3. Completed drone attached with the Raspberry Pi and the transmitter

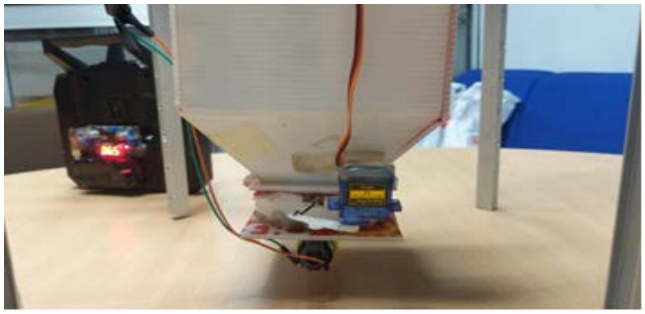


Fig. 4. Speed Dispenser

According to the 3D design, the PVC foam board (5mm) cut to the relevant measures and chambers created to enter the seeds and display the level. The cutting sketch glued as a funnel shape according to the above Fig. 4. Then, the seed control mechanism was created using a servo motor and aluminum plate (2in*2in)[19]. The bottom part of the seed dispenser unit created to spread the seeds widely using a gear motor. Servomotor and gear motor connected to the flight controller and tested. To switch on the gear motor, a relay (transistor embedded) used. The seed dispenser unit attached to the drone (which also can be detached easily). Moreover, tested the flight with seeds. Then a stand placed under the drone to protect the seeds dispenser unit. A test run conducted to measure the seeds' flow rate prior to connecting the raspberry pi to the flight controller.

### *E. Internet of things*

To read the flight controller's data, a raspberry pi installed. Raspberry pi was linked to a website using the python programming language[10] [20]. The flight data history of the drone and the accuracy followed according to the given instructions checked through the website. Here for the data communication JSON API requests will be used and those data will be sent to the web site and once the web page captured the data, those data will be saved in the database and those data will be processed and later on, displayed on a web site [21].

## IV. Results

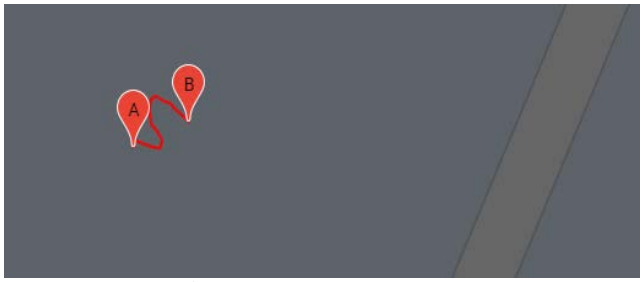



Fig. 5. Path tracking of the drone

The above Fig. 5. displays a sample movements of the drone which is plotted over Google Maps using google APIs, which was live tracked when the drone flew to drop seeds. In addition, the following Fig. 6. is a dashboard that shows the summary of the paths that the drone been traveled.

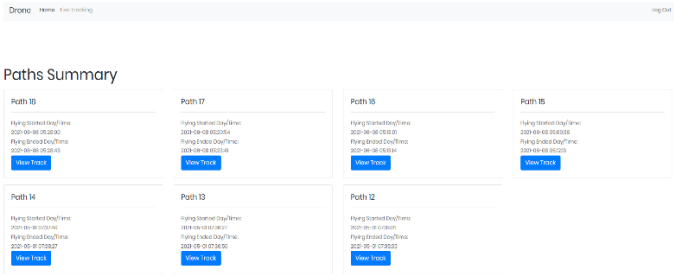



Fig. 6. Website Homepage with paths traveled history

Fig. 7. Drone way point. (Drone planner) Location – University of Kelaniya, Ground.

- Total way points – 14
- Maximum Altitude – 3ft (Feet)
- Latitude measure north-south position, with the equator at 0 degrees and the North Pole at 90 degrees north.
- Longitude measure east-west position, with 0 degrees at Greenwich, England.
- Home location (H) –
  - Latitude - 6.974836
  - Longitude - 79.914406

Table 2. Waypoint table with longitude and latitude.

| Waypoint | Flying altitude(ft.) | Latitude | Longitude |
|---|---|---|---|
| H | 2.992 | 0 | 0 |
| 2-3 | 3.775 | 6.9747367 | 79.9144127 |
| 3-4 | 2.551 | 6.974645 | 79.9145016 |
| 4-5 | 3.225 | 6.9746356 | 79.9144886 |
| 5-6 | 3.735 | 6.9747278 | 79.9143992 |
| 6-7 | 3.756 | 6.9747385 | 79.9143959 |
| 7-8 | 2.890 | 6.9746262 | 79.9144755 |
| 8-9 | 3.560 | 6.9746168 | 79.9144624 |
| 9-10 | 3.460 | 6.97471 | 79.914372 |
| 10-11 | 2.450 | 6.9747011 | 79.9143584 |
| 11-12 | 2.557 | 6.9746074 | 79.9144493 |
| 12-13 | 3.330 | 6.974598 | 79.9144362 |
| 13-14 | 3.775 | 6.9746922 | 79.9143448 |
| 14-H | 2.545 | 6.9746545 | 79.9145147 |

Here the max altitude has been defined as 3ft, but when the drone actually flying there was a fluctuation been the real value such as around ± 25.83% altitude due to external factors and imbalance due the weight. Also focusing on reducing the altitude percentage and get closer to defined altitude.

## V. DISCUSSION

UAVs are flying robots. Even though initially intended for military use, they currently generally utilized in different areas, from sporting games and putting out fires. In this part, a utilization introduced of UAVs on Agricultural businesses. A significant sort of UAV been introduced. However, with its capacity to float on spot and take-off and landing vertical, the multi-rotor UAV may appear to be appropriate for agriculture. Its restricted flight time is a significant impediment. The half-and-half fixed-wing-motor-rotor may be a superior fit. An itemized understanding of the uses of UAVs in crop creation and animal cultivating likewise introduced. UAVs have the benefit of improving and are even more effective when contrasted with people.

A few benefits of applying UAVs in Agriculture introduced, some of which incorporate restricted way requirements, efficiency, and reduced extensive labor. Nonetheless, there are various difficulties restricting UAVs, commonly the incorporated cost. UAVs that are appropriate for agriculture use are costly. Activity and support likewise include some significant downfalls. Thus, it is challenging to persuade ranchers and agriculture related parties to invest in UAVs for their business. One-time cost, battery impediments, security, and legitimate related issues are significant obstacles that should be scaled before UAVs can track down solid traction in agriculture. Following Fig. 8. shows a figure while the drone releases the seeds.

Fig. 8. Aerial Seeding through the drone

## VI. CONCLUSION

A UAV designed to spray seeds on the agricultural fields safely. The UAV is a transportation platform that can carry materials from one place to another. Pixhawk selected to control the drone, and the thrust calculations done to understand the total weight the frame holds. Therefore, it can be helpful in enhancing the stability of the drone since the drone should be properly stable in order to have proper control over the drone. Moreover, Raspberry Pi micro controller incorporated in this study as it has an inbuilt Wi-Fi module and USB ports. As USB ports are available, it is easy to connect a dongle for a high-speed internet connection to communicate with the web portal. In addition, the Raspberry Pi has a better processing speed that is useful in this project. Thus, this study proves that the use of drone technology is far more efficient in speed spraying over human labor due to the variable - speed. Therefore, in conclusion it is evident that drone technology creates an effective platform for harvesting without damaging seeds and by maintaining proper records of seed spraying paths.

## VII. FUTURE SCOPE

In this project, a prototype of UAV designed to be use in the agricultural field. It is required to improve the flying time of the UAV while increasing the payload by considering the throttle percentage formerly it will be useful for larger agricultural areas.

Also further studying about the aerodynamics in order to make a better body design and better propellers that will help to increase the flying time and make the drone fly under 6-sigma accuracy level. In addition, more precisely, control the drone via the website in real-time.

## REFERENCE

[1] "Croptracker - Drone Technology In Agriculture." https://www.croptracker.com/blog/drone-technology-in-agriculture.html (accessed Jul. 18, 2021).

[2] "Three Drones Work Together to Spray and Seed Fields." https://www.dtnpf.com/agriculture/web/ag/crops/article/2021/06/02/three-drones-work-together-spray (accessed Jul. 15, 2021).

[3] "Automated Plant Counts with Drones for Precision Agriculture." https://www.precisionhawk.com/blog/media/topic/understanding-your-aerial-data-plant-counting (accessed Jul. 17, 2021).

[4] "Introduction to Remotely Piloted Aircraft Systems (RPAS) - SKYbrary Aviation Safety." https://www.skybrary.aero/index.php/Introduction_to_Remotely_Piloted_Aircraft_Systems_(RPAS) (accessed Jul. 15, 2021).

[5] E. P. Fortes, "Seed Plant Drone for Reforestation," *J. Grad. Res. Creat. Scholarsh.*, vol. 2, no. 1, pp. 13–26, 2017.

[6] "Crop Spraying Drones For Sale | Agriculture Drones | HSE." https://hse-uav.com/product-category/commercial-drones/agriculture-and-crop-spraying-drones/ (accessed Jul. 15, 2021).

[7] B. Publications, "Agriculture Sciences," vol. 9, no. 7.

[8] S. C. Hassler and F. Baysal-Gurel, "Unmanned aircraft system (UAS) technology and applications in agriculture," *Agronomy*, vol. 9, no. 10, 2019, doi: 10.3390/agronomy9100618.

[9] L. Meier, P. Tanskanen, F. Fraundorfer, and M. Pollefeys, "PIXHAWK: A system for autonomous flight using onboard computer vision," *Proc. - IEEE Int. Conf. Robot. Autom.*, no. May, pp. 2992–2997, 2011, doi: 10.1109/ICRA.2011.5980229.

[10] A. Andrio, "Development of UAV technology in seed dropping for aerial revegetation practices in Indonesia," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 308, no. 1, 2019, doi: 10.1088/1755-1315/308/1/012051.

[11] S. C. Choi, N. M. Sung, J. H. Park, I. Y. Ahn, and J. Kim, "Enabling drone as a service: OneM2M-based UAV/drone management system," in *International Conference on Ubiquitous and Future Networks, ICUFN*, Jul. 2017, pp. 18–20, doi: 10.1109/ICUFN.2017.7993739.

[12] A. K. Saha *et al.*, "IOT-based drone for improvement of crop quality in agricultural field," in *2018 IEEE 8th Annual Computing and Communication Workshop and Conference, CCWC 2018*, Feb. 2018, vol. 2018-January, pp. 612–615, doi: 10.1109/CCWC.2018.8301662.

[13] "Generic Radiolink PIXHAWK Flight Controller FPV RC Drone FC 32 Bit with Power Module for Racing Drone/Multirotor/Quadcopter: Buy Online at Best Price in UAE - Amazon.ae." https://www.amazon.ae/Generic-Radiolink-Controller-Multirotor-Quadcopter/dp/B07L9GW8TT (accessed Apr. 29, 2021).

[14] "All About a Multirotor FPV Drone Frame | GetFPV Learn." https://www.getfpv.com/learn/new-to-fpv/all-about-multirotor-fpv-drone-frame/ (accessed Jul. 20, 2021).

[15] "Quick Drone Parts Overview Along With Handy DIY Tips - DroneZon." https://www.dronezon.com/learn-about-drones-quadcopters/drone-components-parts-overview-with-tips/ (accessed Jul. 20, 2021).

[16] "How to design a Propeller in Autodesk Fusion 360 | Fan Blade design - YouTube." https://www.youtube.com/watch?v=g_Wi_JSmKvI (accessed Oct. 21, 2020).

[17] "3D Racing Drone Frame Design in Autodesk Fusion 360 - YouTube." https://www.youtube.com/watch?v=hwo0R2_2_nI (accessed Oct. 31, 2020).

[18] "Find the Perfect Motor to Fit Your Drone | Tom's Guide Forum." https://forums.tomsguide.com/faq/find-the-perfect-motor-to-fit-your-drone.26313/ (accessed Jul. 20, 2021).

[19] "Twitter." https://uavsystemsinternational.com/blogs/drone-guides/best-drones-for-spreading-best-seed-fertilizer-spreader-drones (accessed Jun. 10, 2021).

[20] A. Dorzhigulov, B. Bissengaliuly, B. F. Spencer, J. Kim, and A. P. James, "ANFIS based quadrotor drone altitude control implementation on Raspberry Pi platform," *Analog Integr. Circuits Signal Process.*, vol. 95, no. 3, pp. 435–445, Jun. 2018, doi: 10.1007/s10470-018-1159-8.

[21] I. Brand, J. Roy, A. Ray, J. Oberlin, and S. Oberlix, "PiDrone: An Autonomous Educational Drone Using Raspberry Pi and Python," in *IEEE International Conference on Intelligent Robots and Systems*, Dec. 2018, pp. 5697–5703, doi: 10.1109/IROS.2018.8593943.

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index No: ES-23-76*

# Impact of ZnO NPs on *In Vitro* Germination and Growth Characteristics of Rice (*Oryza sativa*)

Sujani Perera (1st Author)
*Department of Applied Computing*
*Faculty of Computing and Technology*
*University of Kelaniya*
Peliyagoda, Sri Lanka
tosujani@gmail.com

Thilini Rupasinghe (2nd Author)
*Department of Applied Computing*
*Faculty of Computing and Technology*
*University of Kelaniya*
Peliyagoda, Sri Lanka
thilinir@kln.ac.lk

Lahiru Wijenayaka (3rd Author)
*Department of Chemistry*
*Faculty of Natural Sciences*
*Open University of Sri Lanka*
Nawala, Sri Lanka
lawj@ou.ac.lk

***Abstract* — A key feature of contemporary nanotechnology is the successful utilization of nanotechnological concepts in agriculture, for applications such as enhancement of the efficiency of crop production through improved seed germination & growth, smart fertilizers, smart pesticides etc. Effect of different nanoparticles (NPs) such as $TiO_2$, Ag, Si, Au, Cu, Zn and ZnO on seed germination have been studied in literature for crops such as canola, mung beans, onions, fenugreek and watermelon. Current study focuses on investigating the impact of ZnO NPs on the seed germination and growth of rice (*Oryza Sativa*). ZnO is a non-toxic NP and has the potential to boost the yield and growth of food crops. According to literature, ZnO NPs have been used to enhance the seed germination in different plants such as mung beans, peanut and black gram. Herein, ZnO NPs were synthesized by a wet chemical method by mixing ethanolic solutions of NaOH and Zn $(CH_3COO)_2{\cdot}2H_2O$.The morphology of the synthesized ZnO NPs was studied using Scanning Electron Microscopy (SEM) and synthesized ZnO NPs exhibited a spherical shape with a diameter ranging from 65 nm to 95 nm, with an average diameter of (73±2) nm. The impact of ZnO NPs on germination and growth of rice was studied under different NP concentrations (0-2000 mg/L) for Sudu samba and traditional rice varieties, Suwandel and Madathawalu. Further, the growth characteristics were investigated by measuring the root length and the shoot length of rice seeds. A significant enhancement of seed germination was observed in all three rice varieties after treating with ZnO NPs. Sudu Samba seeds showed a 7.3 % enhancement of seed germination (at 500 mg/L ZnO) while Suwandel and Madathawalu showed an enhancement of 20 % and 17% respectively. This can be attributed to the ability of ZnO NPs to penetrate through the cell wall and release targeting genes to specific cellular organelles to boost the cell division. As Zinc is an enzymatic component, it has the ability to influence secretion of indole acetic acid (IAA) which important to regulate plant growth and when the level of IAA is increased, it causes an effective response in seed germination. Interestingly, no significant difference in root length and shoot length was observed for Madathawalu while, a negative effect was observed for Sudu samba and Suwandel at 7 days. Further investigations with varying ZnO concentrations and other types of traditional Sri Lankan rice varieties are currently ongoing to gain more insight into the phenomena observed above.**

***Keywords — Zinc oxide nanoparticles, germination, seedling growth, traditional rice***

## I. INTRODUCTION

Nanotechnology is an interdisciplinary research area that deals with the atomic and molecular scale of materials between 1 and 100nm scale. Owing to the unique properties of materials at nanoscale, its utilization has been researched in almost every field [1]. In agriculture, concepts of nanotechnology has successfully been utilized for applications such as enhancement of the efficiency of crop production through improved seed germination & growth, smart fertilizers, smart pesticides, nanosensors and nanobionics [2]. Several germination studies have been conducted in literature using nanoparticles (NPs) such as $TiO_2$, Ag, Si, Au, Cu, Zn and ZnO to enhance seed germination and crop growth [2]. For example, $TiO_2$ NPs have been experimented on canola and broad beans, Ag NPs have been experimented on fenugreek, watermelon and jasmine rice, whereas Si NPs have been experimented on sunflower. Further, $Mg(OH)_2$ NPs have been tested on corn and ZnO NPs have been experimented on peanuts and mung beans [3]-[4]. According to these studies, an increase in germination up to 75% has been observed when canola seeds (*Brassica napus*) were embedded with $TiO_2$ NPs (2000 mg/L). Additionally, the incorporation of $TiO_2$ NPs has also resulted in larger radicle and plumule growth of canola seedling at 1200 mg $L^{-1}$, 1500 mg $L^{-1}$ [5]. Furthermore, an enhancement of germination and root lengths have been observed in $TiO_2$ treated broad bean (*Vicia faba*) seeds compared to the controlled seed [3]. Silver NPs on fenugreek (*Trigonella foenum-graecum*) has shown an enhanced germination percentage, where the highest germination (of 78% ) has been observed at a AgNP concentration of 20µg $mL^{-1}$ [6]. Acharya *etal* have investigated the seed germination, growth and yield of triploid and diploid watermelon (*Citrullus lanatus*) using AgNP, $AgNO_3$ and TNE (Turmeric Oil

Emulsion) and higher germination rates, lower mean germination times and higher final emergence percentages have been reported after treating with NPs [7]. Further, Jasmine rice (*Oryza sativa*) treated with Ag NPs have exhibited significantly higher germination percentages, vigor index and lengths of roots, shoots and seedling biomass compared to $AgNO_3$ primed seeds, while sunflower (*Helianthus annuus*) treated with Si NPs (0.2 and 0.4 mM) have shown a substantial reduction in days to 50% germination, mean germination time, improved root length, mean daily germination, seedling vigour index and final germination percentage [8], [9]. Also, corn ( *Zea mays*) treated with $Mg(OH)_2$ NPs have shown enhanced seed germination (100%) and seedling growth at 500 ppm [10], whereas tomato (*Solanum lycopersicum*) seeds treated with Ag NPs have shown significantly high germination rates and seedling growth compared to untreated seeds [11]. Interestingly, these seeds have shown significant improvement in germination at only a specific concentration. At higher or lower concentrations apart from that concentration had caused different negative impacts on the seed germination, further growth and crop yield as well [6]. For example, Hojjat *etal* have reported a negative effect of Ag NPs on the germination of fenugreek at higher concentrations [6], [10]. Therefore, maintaining the effective and optimum concentration levels is enormously important for better results and maintaining the ecological balance.

Zinc Oxide NPs can be named as one of the most widely used NPs in the current context, owing to their antimicrobial, antifungal, optical and electrical properties. More importantly, ZnO NPs have the potential to boost the yield and the growth of food crops [9]. For example, N.Jayarambabu *etal*, have reported a higher degree of germination (of 84.75%) of mung beans (*Vigna radiata*) in the presence of ZnO NPs, with an increase in the root length by 29.2%, shoot length by 9.36% and seedling length by 2.51 % at 125 ppm of ZnO NPs [12]. Further, T.N.V.K.V Prasad *etal* have reported a 100% seed germination on peanut, while K.Raja *etal* have observedma 4 % increase of seed germination, 16.3 % increase in root length, 18.1 % increase in shoot length and 25.2 % increase in seedling vigor, over control in black gram (*Vigna mungo*) when treated with ZnO nanoparticles at 1100 mg $kg^{-1}$ [4] [13].

## II. OBJECTIVES

The objective of the current study is to investigate the effects of ZnO NPs on the process of *in vitro* seed germination and seedling growth of rice (*Oryza sativa*), specifically focusing on traditional rice varieties such as Suwandel, Madathawalu, Kalu heenati and Pachchaperumal, which are considered to be rich in nutrients such as proteins, minerals and antioxidants, compared to other varieties.

## III. METHODOLOGY

### A. Materials

Zinc acetate dehydrate [Zn $(CH_3COO)2.2H_2O$] (Molecular Weight 219.50 $g.mol^{-1}$, Minimum Assay 99.5%, Sisco Research Laboratories Pvt Ltd, India), sodium hydroxide (NaOH) (Molecular weight 40 $g.mol^{-1}$, Minimum Assay 97%, S D Fine-Chem Limited, India) and methanol ($CH_3OH$) (Sigma Aldrich, USA) were used to synthesis the ZnO NPs from the wet chemical method. Seeds of rice (*Oryza sativa*), Sudu samba, Suwandel, Madathawalu were gathered from local farmers.

### B. Synthesis of ZnO NPs

The synthesis was adapted from "Germination and Growth Characteristics of Mung-bean Seeds (*Vigna radiata* L.) affected by Synthesized Zinc Oxide Nanoparticles [12]. ZnO NPs were synthesis as stated in literature via a chemical method. 0.1M zinc acetate was dissolved in methanol and stirred for 30 minutes at $50^0$C. 0.2M sodium hydroxide was dissolved in methanol with stirring and it was scaled up to 1:2 by adding 100ml methanol. Sodium hydroxide was added drop wise to the zinc acetate solution that was kept at $50^0$C. During the addition of sodium hydroxide solution, 50ml methanol was added to zinc acetate solution to prevent methanol evaporation, followed by another 50 ml at the end. A white milky solution was obtained and stirring continued for 80 minutes at $50^0$C. The solution was aged for 17 hrs at $27^0$ C temperature. After that the solution was centrifuged for 10 minutes for volume reduction. The volume reduction solution was washed three times using centrifuge. The sample was then placed in oven at $70^0$C for 90 minutes. The observed white powder was collected to a ceramic crucible. The crucible was placed in a muffle furnace and heat treated at $400^0$C for 2 hours. Finally, the powder was grinded using mortar and pestle.

### C. Characterization of ZnO NPs

The characterization of the ZnO NPs, presented herein, was performed at an accredited research institute, Sri Lanka Institute of Nanotechnology at Homagama, Sri Lanka. The morphology of particles was determined using Scanning Electron Microscopy (SEM) (Hitachi SU6600 Scanning Electron Microscope, Europe).

### D. Seed Preparation

The seeds were first checked for their viability by suspending them in distilled water. The seeds which are settled to the bottom were selected for further study. For in vitro cultures 50 seeds were tested for each sample. The seeds were soaked under different NP concentrations (0 – 1000 $mgL^{-1}$). The control (0 $mgL^{-1}$) seeds were treated with 10 mL distilled water. The seeds were kept in the prepared medium for two days and then kept on a wetted paper towel for a day. Sudu samba and traditional rice varieties, Suwandel and Madathawalu seeds were experimented in the research.

### E. Seed Germination Test

The seed germination rate {SGR- Equation (1)}, relative root growth {RRG - Equation (2)} and germination index {GI - Equation (3)} were calculated using the following equation [12].

$$\text{Seed Germination Rate} \quad SGR = \frac{SS}{SC} \times 100\% \qquad (1)$$

$$\text{Relative Root Growth} \quad RRG = \frac{RS}{RC} \times 100\% \qquad (2)$$

$$\text{Germination Index} \quad GI = \frac{SGR}{RRG} \times 100\% \qquad (3)$$

Where SS is the number of seed germinated in ZnO NPs sample, SC is number of seed germinated in control, RS is the average root length in ZnO NPs sample and RC is average root length in control.

### *F. Root Length and Shoot Length*

Root length was taken from the seminal root of the seed. Shoot length was measured from the beginning of the mesocotyl to the end of the second leaf. The root and shoot length were measured with the help of a thread and scale.

## IV. RESULTS AND DISCUSSION

### *A. Characterization of ZnO NPs*

Morphology of the synthesized ZnO NPs was studied using Scanning Electron Microscope (SEM) images (Fig. 1). According to the SEM analysis, synthesized ZnO NPs exhibited a spherical shaped with a diameter ranged from 65 nm to 95 nm, with an average diameter of (73±2) nm.

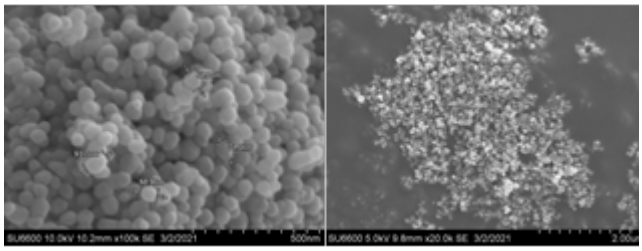


Fig. 1. Scanning electron microscopy (SEM) images of zinc oxide nanoparticles (ZnO NPs) at different scales of 500 nm, 2µm from left to right

### *B. Seed Germination*

Table 1 summarizes the seed germination percentages of ZnO NPs embedded Sudu samba, Suwandel and Madathawalu. As can be seen in Table 1 a significant enhancement in percent germination was observed in ZnO NPs treated samples compared to the control, at lower concentrations. For Sudu samba, the highest germination was observed at 500 mg/L, which was 88 %. This was an enhancement of 7.3% compared to the control which showed germination percentage of 82%. Interestingly, seeds treated with 1000 $mgL^{-1}$ ZnO NPs showed the least germination, and it was even lower than the control. A similar behavior was observed with Suwandel where the sample treated with 200 $mgL^{-1}$ ZnO NPs showed the highest germination of 96%, while the control showed a seed germination percentage of 70 (Fig. 2). The seeds which were treated with 1000 $mgL^{-1}$ and 2000 $mgL^{-1}$ ZnO NPs showed the least germination and the seeds treated with 500 mg/L exhibited a germination percentage of 84 %. Overall, the dynamics of seed germination were most intensive in the media with ZnO NPs at concentration of 200 mg/L. Further, Madathawalu showed the highest germination of 14% at the 500 $mgL^{-1}$ ZnO NPs while the control showed 12%. And this was a 17 % enhancement in the percent germination. Overall form these preliminary data, it can be stated that the elements of seed germination have generally escalated in the media with ZnO NPs at the convergence of difference concentration of ZnO NPs.

Table 1. . Seed germination percentage of different seeds types on different concentration of ZnO NPs at second day

| **Seed Type** | **ZnO NPs Concentration ($mgL^{-1}$)** | **Germination Percentage at 2nd day** |
|---|---|---|
| Sudu samba | 0 | 82 % |
| | 500 | 88 % |
| | 1000 | 78% |
| Suwandel | 0 | 70 % |
| | 200 | 96 % |
| | 500 | 84 % |
| | 1000 | 60 % |
| | 2000 | 66 % |
| Madathawalu | 0 | 12 % |
| | 500 | 14 % |

Fig. 2. Suwandel germination on 200 $mgL^{-1}$ ZnO NPs at Day 2

### *C. Seedling Growth*

The seedling growth was observed for a period of 7 days. As can be seen in Table 2 and 3 the control has exhibited the highest root length and shoot length in Sudu samba and Suwandel. Interestingly, no significant difference in root length and shoot length was observed for Madathawalu while, the above-mentioned negative effect was observed for Sudu samba and Suwandel at 7 days.

Table 2. Seedling growth characteristics of Sudu samba within 7 days

| **Concentration ($mgL^{-1}$)** | **Germination Percentage (%)** | **Root Length(mm)** | **Shoot Length(mm)** | **Relative Root Growth** | **Germination Index** |
|---|---|---|---|---|---|
| 0 | 92 | 12.1 ± 5.0 | 30.7±10.9 | 100.0 | 92.0 |
| 500 | 90 | 6.1 ± 2.3 | 29.8 ± 7.4 | 50.3 | 179.0 |
| 1000 | 84 | 7.4 ± 2.6 | 28.5±10.9 | 61.2 | 137.3 |

Table 3. Seedling growth characteristics of Suwandel within 7 days

| Concentration ($mgL^{-1}$) | Germination Percentage (%) | Root Length(mm) | Shoot Length(mm) | Relative Root Growth | Germination Index |
|---|---|---|---|---|---|
| 0 | 96 | 17.7 ± 8.6 | 38.9 ± 9.8 | 100.0 | 96.0 |
| 200 | 98 | 15.3 ± 6.7 | 37.0±10.4 | 86.7 | 113.1 |
| 500 | 100 | 10.3 ± 4.8 | 33.7±11.9 | 58.2 | 171.7 |
| 1000 | 96 | 10.9 ± 4.3 | 36.0± 9.7 | 61.7 | 155.5 |
| 2000 | 96 | 10.8 ± 4.1 | 35.0±11.2 | 61.0 | 157.3 |

Herein, a significantly positive impact on early seed germination was observed when rice seeds were treated with ZnO NPs. Interestingly, different concentrations of ZnO NPs have affected on different seeds types in different ways. The experiment results showed that ZnO NPs can be used to enhance the seed germination and early seedling growth of rice seeds. However, it is crucial to find the optimum concentration of ZnO NPs for individual rice varieties as there can be side effects on the growth of the seeds at higher concentration.

This observation of enhanced germination in the presence of ZnO NPs can be attributed to their ability to pass through the cell wall and release targeting genes to specific cellular organelles. Furthermore, NPs have the ability to stimulate or inhibit of seed germination or seedling growth, activation of genes involved in metabolism and induction of photosynthesis or reactive oxygen species. Specifically, ZnO NP has the potential to pass through the cell wall and release targeting genes to specific cellular organelles, enhance the cell division while zinc is an enzymatic component influencing the secretion indole acetic acid (IAA) which is a phytohormone(auxin) which significantly regulates plant growth. By increasing the level of IAA, zinc gives a positive response in seed germination.[1], [14]-[15]. Moreover, the early seed germination and seedling growth might have happened due to the water uptake by the ZnO NPs treated seeds and also it caused to reduce the adverse effects of heat, drought and salt stresses [1]. Additionally, the antifungal and antibacterial action of ZnO NPs can also be useful herein to control the spread of infections by a variety of plant pathogens and thereby enhancing the crop productivity.

Fig. 3. Abnormal seed germination due to the high concentration of ZnO NPs

Interestingly, at higher concentrations of ZnO NPs, some negative effects such as slow germination rate, lower seedling growth such as root growth, shoot growth and distorted root growth were observed (Fig.3). For example, Suwandel showed only a 66% of seed germination at 2000 mg/L while 200 $mgL^{-1}$ sample showed a 96% germination, whereas the control seeds showed 70% seed germination. According to literature, similar phenomena have been observed where different types of NPs have shown positive impacts on the seed germination of crops such as fenukgreek (*Trigonella foenum-graecum*), sunflower (*Helianthus annuus*), corn (*Zea mays*), watermelon (*Citrullus lanatus*), peanut (*Arachis hypogaea*), jasmine rice (*Oryza sativa*) and chickpea (*Cicer arietinum*) at lower concentrations. For example, S. S. Hojjat etal have shown that Ag NPs enhance the germination and seedling growth of fenugreek by 78 % at 20µg $mL^{-1}$ concentration. However, at higher concentrations, Ag NPs have shown adverse effects. The higher concentrations of AgNPs (>20µg $mL^{-1}$) strongly inhibited both the shoot and root growth (especially as moisture weight), with a more marked inhibition of the shoot growth than the root growth[6]. Also, M. Sedghi *etal* have observed slightly adverse effects such as reduction in the residual fresh and dry weight (respectively 0.36g and 0.090g) at $1g.L^{-1}$ with soybean (*Glycine max*) compared to the control as a consequence of increasing the concentration of ZnO NPs [16].

## V. CONCLUSION

In this study, ZnO NPs were synthesized via a wet chemical method and the impact of ZnO NPs on the germination and the growth of rice (*Oryza sativa*) was studied using different types of rice samples including traditional Sri Lankan rice spices such as Suwandel and Madathawalu. Synthesized ZnO NPs were characterized by Scanning Electron Microscope to determine the morphology and the size. And ZnO NPs were found to be spheres with a diameter in the range of 65nm to 95 nm. According to the results, the application of ZnO NPs had significantly enhanced the seed germination of rice where Sudu Samba seeds showed a 7.3 % enhancement of seed germination (at 500 mg/L ZnO) while Suwandel and Madathawalu showed an enhancement of 20 % and 17% respectively. This can be attributed to enhanced cell division supported by ZnO NPs, a critical nutrient in the cell division and growth in plants. Further, according to preliminary data, it was also observed that lower concentration of ZnO NPs leads to best germination percentage. Sudu samba seeds treated with 500 $mgL^{-1}$ showed 88% germination while seeds treated with 1000 mg/L resulted a germination percentage of 78 % at the second day. Suwandel seeds treated with 200 $mgL^{-1}$ showed 96% germination and at concentrations above 500 mg/L, percentages of germination were below 70 %, which was the germination percentage of the control at second day. According to preliminary data, it can be concluded that ZnO NPs have exhibited a positive impact on *the in vitro* germination of Rice (*Oryza sativa)* for Sudu samba, Suwandel and Madathawalu. However, it is significantly important to focus on the optimum and effective concentration of ZnO NPs for higher seed germination, as

well as for early plant growth which is essential in achieving higher crop productivity.

## REFERENCES

[1] A. Tymoszuk and J. Wojnarowicz, "Zinc oxide and zinc oxide nanoparticles impact on in vitro germination and seedling growth in allium cepa L.," *Materials (Basel).*, vol. 13, no. 12, pp. 1–16, 2020, doi: 10.3390/ma13122784.

[2] Y. Shang, M. Kamrul Hasan, G. J. Ahammed, M. Li, H. Yin, and J. Zhou, "Applications of nanotechnology in plant growth and crop protection: A review," *Molecules*, vol. 24, no. 14, 2019, doi: 10.3390/molecules24142558.

[3] M. Ruffini Castiglione, L. Giorgetti, L. Bellani, S. Muccifora, S. Bottega, and C. Spanò, "Root responses to different types of TiO2 nanoparticles and bulk counterpart in plant model system Vicia faba L.," *Environ. Exp. Bot.*, vol. 130, pp. 11–21, 2016, doi: 10.1016/j.envexpbot.2016.05.002.

[4] T. N. V. K. V. Prasad *et al.*, "Effect of nanoscale zinc oxide particles on the germination, growth and yield of peanut," *J. Plant Nutr.*, vol. 35, no. 6, pp. 905–927, 2012, doi: 10.1080/01904167.2012.663443.

[5] H. Mahmoodzadeh, "c r v i h o e f c r v i h o e f," vol. 3, no. 1, pp. 25–32, 2013.

[6] S. S. Hojjat and H. Hojjat, "Effect of Nano Silver on Seed Germination and Seedling Growth in Fenugreek Seed," *ETP Int. J. Food Eng.*, vol. 1, no. 2, pp. 106–110, 2015, doi: 10.18178/ijfe.1.2.106-110.

[7] P. Acharya, G. K. Jayaprakasha, K. M. Crosby, J. L. Jifon, and B. S. Patil, "Nanoparticle-Mediated Seed Priming Improves Germination, Growth, Yield, and Quality of Watermelons (Citrullus lanatus) at multi-locations in Texas," *Sci. Rep.*, vol. 10, no. 1, pp. 1–16, 2020, doi: 10.1038/s41598-020-61696-7.

[8] W. Mahakham, A. K. Sarmah, S. Maensiri, and P. Theerakulpisut, "Nanopriming technology for enhancing germination and starch metabolism of aged rice seeds using phytosynthesized silver nanoparticles," *Sci. Rep.*, vol. 7, no. 1, pp. 1–21, 2017, doi: 10.1038/s41598-017-08669-5.

[9] N. S. Mohsen Janmohammadi, "Botanica lithuanica," *Bot. Lith.*, vol. 18, no. 1, pp. 40–45, 2012.

[10] S. Shinde, P. Paralikar, A. P. Ingle, and M. Rai, "Promotion of seed germination and seedling growth of Zea mays by magnesium hydroxide nanoparticles synthesized by the filtrate from Aspergillus niger," *Arab. J. Chem.*, vol. 13, no. 1, pp. 3172–3182, 2020, doi: 10.1016/j.arabjc.2018.10.001.

[11] A. Noshad, C. Hetherington, and M. Iqbal, "Impact of AgNPs on seed germination and seedling growth: A focus study on its antibacterial potential against clavibacter michiganensis subsp. Michiganensis infection in solanum lycopersicum," *J. Nanomater.*, vol. 2019, 2019, doi: 10.1155/2019/6316094.

[12] N.Jayarambabu, B. S. Kumari, K. V. Rao, and Y. T. Prabhu, "Germination and Growth Characteristics of Mungbean Seeds ( Vigna radiata L . ) affected by Synthesized Zinc Oxide Nanoparticles," *Int. J. Curr. Eng. Technol.*, vol. 4, no. 5, pp. 3411–3416, 2014.

[13] K. Raja, R. Sowmya, R. Sudhagar, P. S. Moorthy, K. Govindaraju, and K. S. Subramanian, "Biogenic ZnO and Cu nanoparticles to improve seed germination quality in blackgram (Vigna mungo)," *Mater. Lett.*, vol. 235, pp. 164–167, 2019, doi: 10.1016/j.matlet.2018.10.038.

[14] L. Yin, B. P. Colman, B. M. McGill, J. P. Wright, and E. S. Bernhardt, "Effects of Silver Nanoparticle Exposure on Germination and Early Growth of Eleven Wetland Plants," *PLoS One*, vol. 7, no. 10, 2012, doi: 10.1371/journal.pone.0047674.

[15] A. Gorczyca, E. Pociecha, M. Kasprowicz, and M. Niemiec, "Effect of nanosilver in wheat seedlings and Fusarium culmorum culture systems," *Eur. J. Plant Pathol.*, vol. 142, no. 2, pp. 251–261, 2015, doi: 10.1007/s10658-015-0608-9.

[16] M. Sedghi, M. Hadi, and S. Toluie, "Effect of Nano Zinc Oxide on tSedghi, M., Hadi, M., & Toluie, S. (2013). Effect of Nano Zinc Oxide on the Germination Parameters of Soybean Seeds Under Drought Stress. Annals of West University of Timişoara, Ser. Biology, 16(2), 73–78.he Germination Param," *Ann. West Univ. Timişoara, ser. Biol.*, vol. 16, no. 2, pp. 73–78, 2013.

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index No: TL-24-30*

# Teacher Attitudes towards the Use of a Learning Management System for Teaching Performing Arts Undergraduates During the COVID-19 Pandemic

Sumudu Embogama (1st Author)
*Department of English Language Teaching*
*University of the Visual & Performing Arts*
Colombo, Sri Lanka
sumudu.e@vpa.ac.lk

***Abstract* — The sudden and unplanned shift from traditional face-to-face instruction to online modes due to the unprecedented increase in COVID-19 patients resulted in many challenges faced by both teachers and students. This situation became even more compounded for educational institutions that offered courses in performing arts subjects since these subjects demand the tangible presence of the teachers and students in a single physical environment. The primary purpose of this research was to find out whether educators teaching these subjects had a positive or negative attitude toward the use of a learning management system in a Sri Lankan University that offers degrees' programs in performing arts. The descriptive research design was adopted for this purpose, and data were gathered using a survey-type online questionnaire from a sample of academic staff from the University. Firstly, the survey intended to find out the respondents' attitudes towards using an LMS in general. Secondly, it considered how effective such a system is concerning teaching practical-oriented subjects such as dance and drama. Although the teachers had positive perceptions regarding the effectiveness of using an LMS during emergency remote teaching, they appeared to be uncertain whether it was an effective tool for teaching performing arts subjects. In addition, the findings indicate the need for the teachers to be provided with further training on using an LMS for teaching practical-oriented subjects. According to the findings of this study, there is a clear indication that an LMS has limitations when it comes to teaching subjects that demand the physical presence of teachers and students. A mechanized system cannot fully accommodate tangible interactions that are required for the development of practical knowledge of these subjects at the undergraduate level.**

***Keywords* — *emergency remote teaching, online learning, learning management system, performing arts***

## I. INTRODUCTION

A Learning Management System (LMS) or e-learning platform is a software system that includes a range of services to assist teachers with the management of their courses [1]. Technologies that facilitate the provision of courses over long distances are broadly termed "learning management systems" or "LMSs." According to Turnbull et al., LMSs are defined as web-based software platforms that provide an interactive online learning environment and automate the administration, organization, delivery, and reporting of educational content and learner outcomes [2]. Ortega & Arcos, defines them as " a computing device that groups several tools and ensures the educational lines across dedicated platforms to the ODL (open and distance learning). All conduits are preserved and expanded for the learner, tutor, coordinator and administrator within the e-learning platform" [3]. Given the above explanations of what an LMS is, there is a need also to understand that the system is not an all-inclusive, well-packed replacement to formal education even in crises such as the Covid-19 pandemic that prevented students physical access to their educational institutions.

As a result of the Covid19 pandemic, Universities have been forced to remove/limit physical access to students. As a result, University teachers were compelled to learn to use LMSs such as Moodle as well as online video conferencing tools such as Zoom. This was the mandate given by the Sri Lankan education authorities and the University administration to continue teaching using distance education methods such as synchronous and asynchronous teaching. Only a short time period was provided for adjusting to this new paradigm of teaching. Therefore, teachers have felt challenged by this 'new normal' system of education. However, they appear to make every attempt to adopt and instruct their students to the best of their capabilities given the unwarranted conditions of the pandemic.

Under these circumstances, performing arts teachers were introduced to teaching University-level performing arts courses in subject areas subjects such as dance, music,

theatre, and drama, which are heavily dependent on face-to-face teaching of complex practical aspects related to these subjects. Therefore, a for performing arts educators to adjust their teaching practices to suit a digitalized teaching environment is a challenge. The main questions that these teachers have repeatedly asked but found no comprehensive answer to is "How can one teach dancing, singing, playing instruments and acting, all of which essentially require one-on-one teaching and learning including human-to-human observation and whole-body involvement, be taught through the third-hand mediation of a screen?" Teaching online entails technical issues such as connection disruptions, unavailability of devices, lack of signals as well as practical issues such as not being able to see or hear the students, or check for the aptness of pitch, tune, movement or gesture. These are next to impossible to achieve with the limited range of the camera angle of a computer or a smartphone. Such concerns expressed by dance, music and drama teachers are investigated in this study.

## II. OBJECTIVES

Three primary objectives of this research are as follows:

1. To find out what kind of attitude performing arts teachers had towards the use of their University LMS for teaching purposes in general.

2. To discover how receptive these faculty members were to the possibility of teaching performing arts students' subjects such as dance, music, theatre and drama using the digitalized environment of an LMS.

3. To determine whether these teachers were willing and motivated to obtain further training on effectively using the LMS.

## III. METHODOLOGY

According to Brotherton [4], descriptive research describes what is known about an issue/context but with increased accuracy and precision. The survey was designed using the quantitative method of inquiry which was adopted to gather numerical data for analysis purposes. The survey was a standardized online questionnaire that used a five-point Likert scale to calculate mean scores for individual statements. Subsequently, mean scores for measuring the attitudes of the performing arts teaching staff were generated. The survey consists of three parts. First, the respondents' response towards the use of the University LMS for teaching purposes in general; secondly, their receptiveness to its use for teaching performing arts subjects; and finally, their motivation for using it were statistically measured. The survey indicated to what extent the respondents agreed or disagreed with a particular statement developed concerning the initial research questions that led to the study.

The purposive sampling method, which is a non-probability sampling technique, was selected as the sampling technique. This gave ease of access to the sample population and allowed her the flexibility of approaching study subjects without any undue constraints. Moreover, because the sample size was small, using a random sampling technique would have limited the number of research participants when gathering data. Although the total sample size was limited to 48 University teachers in performing arts, given the narrowed scope of the study, the findings can be considered as reasonably representative of others engaged in the same profession within the same context. To cross-validate the results, opinions were elicited from the participants with their perceptions of using the University LMS for teaching purposes stop.

The University in which the study was conducted was purposively selected to represent the Universities where performing arts subjects are taught. The selected University is the only one of its kind in Sri Lanka which is exclusively devoted to teaching performing arts subjects at undergraduate and postgraduate level. The lecturers and instructors who are teaching the undergraduate level performing arts degree courses were approached to collect the required data.

To collect data, the survey was developed in a digital format using a Google form. Subsequently, the link to the data collection instrument was shared among the staff members using WhatsApp and personal email addresses. This type of data collection instrument was regarded as the most appropriate since it provided the best form of inquiry, it was cost-effective as well as time effective, and precise. Moreover, the auto-tabulated facility provided when using Google forms facilitated accurate analysis with minimal room for human error during data entry as well as when the data were mechanically stored and tabulated.

The online administered, self-report questionnaire involved a series of statements on a 5-point likert scale whereby respondents can select their level of agreement to each assertion presented. Here, number 1 stood for "strongly disagree" and 5 stood for "strong agreement". The first 27 items in the questionnaire were in the form of statements for which the respondents selected their preferred choice from the 5-point Likert scale. The final item (28) was an open-ended question which was expected to gather secondary data in order to identify any themes that emerged about the faculty members' candid responses to the question: "Please explain why you think using the University LMS is effective/ineffective for teaching performing arts subjects to undergraduates".

The data were mechanically calculated for each statement to measure the level of agreement. The findings appeared in the form of a report that included representations of the tabulated data through the mode of pie charts. In addition, the data were transported into an Excel file for further analysis. Since the study mainly intended to describe the attitude of the teachers towards the use of a LMS, mean scores were generated for each statement. Consequently, the responses

given to the open-ended question were analyzed by placing each response as under a theme that matches it.

## IV. RESULTS AND DISCUSSION

The survey was categorized according to three sub-themes:

A. Faculty members' need for LMS training
B. Faculty members' attitude towards the use of the LMS for teaching and learning processes in general.
C. Faculty members' attitude towards the use of the LMS for teaching and learning performing arts subjects through the use of this system.

Therefore, the findings and discussion are presented under each category.

### *A. Faculty members' need for LMS training*

There were four statements that aimed to capture the respondents' requirement for further training on how effectively they can use the LMS for teaching purposes. The overall average score for the responses given to the statements indicated high levels of agreement for the need to obtain more LMS training with an average mean score of 4.02. Out of the relevant statements, a significantly high mean score was received (M = 4.3) for the statement "I'd like to get more training on how to use Learning Management System (LMS)".

Given the above findings, it becomes evident that despite the various training programs that the teacher has already been exposed to, they feel the need further training on how best to use the LMS. Also, this brings us to the understanding that they are yet to feel convinced that they are competent users of this system, and this call for help has to be considered seriously and as an urgency by the relevant authorities. Suppose the teachers are of the understanding that they require further training on using the LMS. In that case, this may be one main obstacle that gets in the way of them at least attempting to try to use it for teaching in this case performing arts subjects. Based on the open-ended inquiry, most of these teachers have stated that they have a basic understanding of the system but that it does not empower them to effectively teach practically-oriented subjects through it.

As such, trainers have to keep in mind that merely teaching the functions of the LMS will not suffice under these circumstances but rather that there should be a specific focus on how best to use the LMS for teaching practical subjects and what effective strategies can be utilized to mitigate issues that may arise. There are also various free and open source resources that the teachers can refer to update themselves on using information communication technologies for teaching and learning purposes. For example, the online open-source resource portal titled "Teaching in the context of COVID-19" describes general aspects of online education, as well as provides tailored resources, activities and assignments, and practical tips for educators.

### *B. Faculty members' attitude toward the use of the LMS for teaching and learning processes in general.*

An average score of 3.5, out of a total score of 5, was received when measuring the academics' expression of the agreement for using this type of mechanized system for pedagogical intent. The finding indicates positive agreement in terms of using the LMS for teaching any subject using this system with the highest mean scores gained for the assertions: "The LMS helps to send information to students quickly" (M=4.08), "The LMS helps teachers to organize their courses well" (M=3.79) and "I feel that the use of LMS develops the teaching process systematically" (M=3.79). Strong levels of agreement with the above statements testify the respondents' awareness of the main functions of an LMS, which is to organize courses systematically and to transfer information and educational material to students rapidly and with minimum hassle.

From this category, the least score of 2.87 was given to the notion that the LMS reduces the role of the teacher in the classroom and increases the role of the student. This moderate level of acceptance indicates that almost half of the sample did not feel as threatened by the presence of this technology as was expected. Of course, the fact that they did not disagree with the statement indicates that more than half the respondents acknowledge the above view, which means that they feel that teachers are gradually becoming redundant as a result of this system's recognizably rapid use in education.

Reassurance that an LMS is merely a carrier of information that cannot be sustained without the presence and mediation of educators has to be reminded constantly to teachers. It is recommended that the training given to the staff should include mostly attitudinal altering workshops that assists teachers to realize the benefits of an LMS instead of perceiving it as a potential threat to their careers as teachers. The findings demonstrate the misconception that a substantial number of educators believe that a digitized system can actually supersede the human presence of a teacher. This notion has to be demystified if the use of the LMS is to be perceived in an endorsing manner.

On the whole, though, we can note that as academic intellectuals who have been selected to teach University-level courses, these teachers are now considerably more aware of the benefits that an LMS has to offer to them as teachers and their students as active learners. Due to its apparent ability to grant autonomy to both these stakeholders, as well as the inevitable need to use an LMS under the present context, the participants of this study have a largely exhibited a positive attitude toward the adaptation of a LMS such as Moodle.

*C. Faculty members' attitude toward the use of the LMS for teaching and learning the performing arts subjects through the use of this system.*

Despite the presence of relatively high agreement levels to the use of the LMS for teaching and learning as a whole, the responses given to the statements capturing the faculty members' perceptions in relation to the use of it for teaching performing arts subjects is considerably lower. The total mean score obtained for this particular category of the survey is 2.87. Thus, it falls under the scale "neither agree nor disagree". This
particular finding effectively captures the dilemma faced by the performing arts teachers as they try to grapple with the new reality of the changes happening in the system of education as a result Covid 19 pandemic.

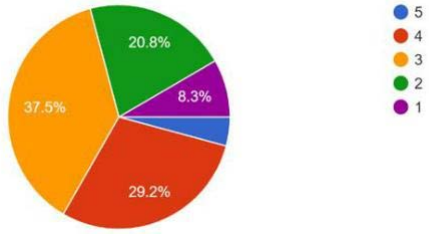



Fig. 1. Responses of the respondents regarding the use of the LMS to generate more positive changes in performing arts undergraduate courses chart.

Note: 1 is for "strongly disagree" and 5 is for "strongly agree")

Figure 1 exemplifies the fact that the majority of the study sample are sensing that the presence of an LMS can lead to more positive than negative changes in the performing arts courses. However, about 30% of the respondents disagree with this claim. The lowest mean score received under this category was to the statement; "The LMS increases flexibility in teaching performing arts subjects" (M=2.54), and this was followed by the statement "Using blended learning, including face to face instruction and the LMS, will offer better results than normal teaching practices in performing arts education" (M=2.6). When considering the range attributed to the mean score we can note the teacher's disagreement with the statement.

This can be noted in the figure below, which also indicates mainly an uncertain attitude from the respondents concerning how useful the LMS is for teaching performing arts subjects.

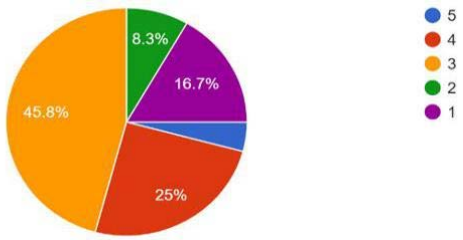



Fig. 2. Respondents level of agreement as to whether the LMS is useful for teaching performing arts subjects to undergraduates

It does not come as a surprise that the teachers have a negative attitude towards the LMS about its inability to be flexible enough to be adopted for teaching performing arts subjects. However, what is an even more noteworthy finding is that the majority of the sample are also hesitant in recommending the use of the blended mode of education for teaching these types of subjects. Given that these faculty members appear to be aware of the multiple benefits of an LMS, their resistance to offering performing arts degree courses using a blended mode comes as an surprise finding. This is because in spite of the restrictions that are present in this computerized system, using the blended mode seems to be possibly the most rational option.

The majority of the teachers (M=3.6) endorse the fact that the use of the LMS can, in fact, bring about more positive than negative changes when teaching these types of practically-oriented subjects. It means that although skeptical concerning the use of ICTs for teaching performing arts courses at the undergraduate level, teachers seem to be hopeful and optimistic about the potential benefits of an LMS for teaching these subjects to a certain extent.

## V. CONCLUSION

The present study's key finding is that, performing arts teachers in the University continue to request further training in using the LMS. Based on the levels of agreement given to the various assertions presented in the survey, the performing arts University teachers appear to be already aware of the benefits of an LMS; nevertheless, they want to gain more knowledge of how this system can accommodate the teaching and learning of practical subjects that invariably require a human touch. Also, the other key finding is the negative attitude and perceptions towards blended learning and its use for teaching practically-oriented subjects. This highlight the necessity for LMS training that does not merely look at the technical aspects of the system but also provide coaching in relation to its wider implications, benefits and how it can be effectively used for blended teaching of performing arts subjects.

Based on the findings, it is realized that when it comes to practical subjects such as dance, drama and music, teachers face many challenges due to the incapacity of the LMS to facilitate human interaction. This is required for developing the practical knowledge of these subjects at the undergraduate level. As such, performing arts educators should be conscious of the fact that teaching performing arts using the LMS can be a feasible option until circumstances dictate otherwise. Hence, as with any other pedagogical approach we select, we must focus on the learning objectives of our performing arts courses and use technologies to mediate our delivery by utilizing it both creatively and intelligently.

Given the inevitability of the change in pedagogy due to the emergence of the Covid-19 pandemic, which is beyond human control, the way forward appears to be to overcome

the obstacles and cultivate a sense of resilience through training and attitudinal change in teachers. This will enable them to move forward despite the volatile environment in which both students and teachers are placed during the pandemic. Given that the study population is not simply a group of teachers, as they are professional artists, we can argue that their capacity for being resilient, creative, adaptable, and flexible can help them surmount the obstacles of online education to a large extent.

## REFERENCES

[1] Mohammed, Ouadoud & Chkouri, Mohamed & Nejjari, Amel, Learning Management System and the Underlying Learning Theories: Towards a new Modeling of an LMS. International Journal of Information Technology, 2018.

[2] Turnbull, Darren & Chugh, Ritesh & Luck, Jo, Learning Management Systems: An Overview, 2019. Doi: 10.1007/978-3-319-60013-0_248-1.

[3] P. Ortega, F.Arcos,Blended Learning: an efficient response to some of the problems raised by youths at risk. INTED2008. International Technology, Education and Development Conference, 2009.

[4] B. Brotherton, Researching Hospitality and Tourism: a student guide, London: Sage, 2008.

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index No: TL-25-34*

# Review on Decision Support Systems used for Resource Allocation in Health Crises

Deshanjali Diyasena (1st Author)
*Software Engineering Teaching Unit*
*Faculty of Science*
*University of Kelaniya*
Kelaniya, Sri Lanka
deshanjalidiyasena@gmail.com

Nimasha Arambepola (2nd Author)
*Software Engineering Teaching Unit*
*Faculty of Science*
*University of Kelaniya*
Kelaniya, Sri Lanka
nimasha_2019@kln.ac.lk

Lankeshwara Munasinghe (3rd Author)
*Software Engineering Teaching Unit*
*Faculty of Science*
*University of Kelaniya*
Kelaniya, Sri Lanka
lankesh@kln.ac.lk

***Abstract*—A disaster or crisis can be stated as a serious disruption occurring for a certain period of time, which could cause loss of human lives, properties, and disrupt the day-to-day life of people. Managing such situations is always a challenge due to various reasons. Especially, allocating and providing resources to manage disaster situations to restore the normal life of people is the main challenge in a disaster situation. Having a proper mechanism for resource allocation could save thousands of human lives as well as properties. Modern smart technologies play a vital role in designing and developing solutions for efficient and effective resource allocation mechanisms. For example, the COVID-19 pandemic has forced people to work from home using digital platforms. Those digital platforms have been able to support people to do their routine work while maintaining social distancing which minimizes the spread of Covid-19. On the other hand, those digital platforms provide an easy and fast way for healthcare officials to reach infected patients to provide necessary treatments and care. Present research critically reviews the past research on managing resources in health crises particularly falls under pandemics and epidemics.**

***Keywords*—resource allocation systems, health crises**

## I. INTRODUCTION

Natural and man-made disasters commonly and frequently happen around the globe. Such disasters always pose threats to human life. For example, past disasters have claimed thousands of human lives and properties. In addition, the destruction they have caused to the environment is hard to recover. Especially, health disasters such as epidemics and pandemics could cause serious damages that negatively impacts human lives and the economies. For example, Cholera is responsible for 120-400 thousand deaths annually worldwide [1]. During the 2009 H1N1 flu pandemic, 203,000 deaths were reported around the globe [2]. The ongoing COVID-19 pandemic has caused 4 million deaths in 192 countries so far and declared a public health emergency by World Health Organization (WHO) [3]. Among them, a significant number of deaths have occurred due to lapses in health resources allocation. For example, in USA 15,571 deaths per month would have been prevented if the resources were effectively allocated [4], and in china it has been shown that a 10-fold per capita increase in the number of hospital beds, medical staff, and assisting staff would decrease the mortality rate by 0.393%, 0.24%, and 0.134% respectively [5]. It highlights the necessity to have a system to manage and allocate needful resources efficiently, which could save thousands of lives.

Disaster or crisis management is a broad domain that includes the phases of mitigation, preparedness, recovery, and response [6]. One of the key challenges in a disaster is that critical resources become scarce as the demand rises unexpectedly. For instance, human resources such as police, medical staff, rescue teams, and/or essential supplies like food, water, medicines/vaccines, equipment, etc. become scarce due to the high demand. A substantial number of researchers have studied and investigated this problem. However, it still remains a challenging task due to various reasons [1], [2], [6]-[18]. Resource allocation can be described as assigning resources considering various aspects of a crisis situation such as priorities, requirements (type(s) of resources and quantities needed), delivery options, and constraints such as ethical aspects [19]. For instance, in the current pandemic, the essential health resources are PCR test kits, personal protective equipment (PPE), vaccines, masks, hospital beds, ICUs equipment, medicines, etc [10]. WHO's declaration of pandemic triggered efforts to boost healthcare facilities and capacity especially in developed countries [20]. Unlike in developing countries, real-time communication technologies, control technologies, and detection technologies have been put together to provide a reliable and accurate medical resources supply system to fight against the health crisis in developed countries [8]. However, many developing countries such as Sri Lanka are suffering from lack of resources and effective resource allocation mechanisms which have made it hard to cope with the COVID-19 pandemic. The poor health infrastructures, lack of well-trained human resources, budget constraints, and poor resource management are the main contributors to this deprivation. Among those factors, poor resource management has been identified as a root cause for poor resource allocation in health crises. For example, quantifying available resources, surveilling the allocations, lack of social

awareness and the inability of taking timely decisions are the main issues with the current process. A substantial amount of research has proposed various solutions to mitigate the above problems but still optimizing resource allocation remains a challenging task due to poor resource management. It motivated us to investigate the existing systems to find their pros & cons.

## II. OBJECTIVES

Objectives of this study are as follows;

- Exploring the existing resource allocation systems in Health Crises
- Analyzing advantages and limitations of existing resource allocation systems
- Studying existing ethical frameworks and guidelines
- Identifying a necessity of a proactive resource allocation system dedicated to health crises

The rest of the paper is organized as follows. The methodology section presents the method used to select the related research followed by section IV which contains a critical review of the selected research and highlights the limitations & issues with the existing resource allocation systems. Finally, section V discusses the literature concisely and concludes this paper with the future directions of the research.

## III. METHODOLOGY

In a crisis situation, priority is given to save human lives. For example, at present, the whole world is racing towards finding a sustainable solution for COVID-19 to save and protect people from the deadly virus. While medical experts are working around the clock to find a medical solution, other healthcare workers are trying to mitigate the crisis with the available resources. As a consequence, the need for efficient and effective resource allocation mechanisms have been proposed based on past research. Studying and investigating the existing research provides a better insight into their advantages, effectiveness, applicability, and limitations. Therefore, the present research critically reviews some of the related research published in the recent past. This review was carried out as qualitative analysis in two stages. First, the relevant research papers were collected from different sources such as IEEE Xplore, Science Direct, ResearchGate, Google scholar, and random search on the Internet using Google search engine. In addition, the most cited papers in the reference lists were also added to the collection. We filtered papers published in the past decade, and the Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [21] standard was used as the guide for filtering papers according to its excluding and including criteria. Apart from that, the details of the community healthcare infrastructure and the resource allocation process to mitigate the COVID-19 pandemic in Sri Lanka were collected through observations and the Ministry of Health publications.

In the second stage, the selected papers were critically reviewed based on qualitative aspects to identify the features, pros & cons, and limitations of existing resource allocation systems. Prior to this, a keyword-based clustering technique was used to identify the most prominent keywords in related studies because keywords reflect the areas that have been covered in a particular study. The main reference tool used was Mendeley, and Microsoft Excel was used as a support tool for maintaining the keywords of selected papers. Out of the selected research papers, only 25 research papers included keywords. Thus, those papers were considered for keyword-based clustering that contained altogether 118 keywords as in the word cloud shown in Figure 1. After pre-processing the collected keywords, the feature extraction was performed using the Term Frequency-Inverse Document Frequency (TF-IDF) vectorizer. Then the Affinity Propagation algorithm, which is an unsupervised Machine Learning (ML) algorithm available for clustering was used for identifying inner patterns in the keywords.


Fig. 1. Keyword word cloud of selected research papers

## IV. SMART DECISION SUPPORT SYSTEMS FOR RESOURCE ALLOCATION IN HEALTHCARE

Designing health-related information systems is always challenging because critical decisions are made based on the information and analytics produced by those systems. Therefore, the accuracy of the system is a crucial factor to be considered when designing such decision support systems. Apart from that, due to the sensitivity of health data, it is required to consider the ethical aspects when allocating resources.

### *A. Decision Support Systems (DSS)*

Resource related decision-making is a complex process especially during crises but the accuracy and the effectiveness of the decisions can be improved with the use of new technologies. Recent studies have proposed promising resource allocation systems with the aims of prioritization, utilization, scheduling, and dispatching resources effectively [6], [7]. Those systems consist of data collection frameworks, databases, and resource deployment guidelines. Besides, different algorithms have been used for victims clustering, resource allocation, and dispatching resources during disaster situations. To take effective and efficient resource allocation decisions it is required to have robust mechanisms to verify data collected from various sources using different data

collection methods. Among them, IoT devices such as sensors, RFID, and Google Maps API have been used in modern systems to collect real-time data. As a smart device, mobile phones play a vital role in modern data collection, and it enables people to be aware of the current situation of the crisis [9].

As this research is narrowing the scope to the health crises, it is worth reviewing the past research that has investigated decision support systems used in the health domain. Optimizing healthcare resource allocation, and distribution has been a key topic in past research. For instance, a platform has been built based on cyber-physical systems for optimizing the city medical resource distribution [8]. It is capable of training people for medical treatment, reducing imbalance and improving utilization of resources with the help of its intelligent and real-time network. A data-driven optimization approach has also been introduced specifically for a cholera outbreak [1]. It considers only vaccines and antibiotics as resources and practices multi-period allocation based on location-specific data of a meta-population. Unlike that, both human and material healthcare resources are assigned, monitored, predicted, and coordinated by the IMPRESS system along with resource mobilization [11]. Another promising way is introducing mathematical models as an aid for resource allocation. For example, patient prioritization and smart dynamic allocation models usage are possible in a patient-driven mobile medical information system [15], [22]. In order to improve decisions, some frameworks also have been designed based on a country's Geopolitical position that is capable of locally collecting, centrally aggregating data prior to being analyzed and communicated [23].

There is a surge in recent studies that focused COVID-19 pandemic as the case study for implementing resource allocation systems. For example, an optimal solution for ventilator distribution during pandemics was proposed as a mixed-integer model adapting an epidemic model called SEIAR (suspectable, exposed, infected, asymptomatic, recovered) for demand prediction [12]. The findings suggest that inter-regional pooling strategies are ensuring adequate provision of ventilators. Further, a time-varying linear optimization-based approach has been proposed for vaccine allocation during a pandemic using the same epidemic model [14]. This solution considers immunity uncertainty along with population density and infected ratio. Apart from that, it has introduced a vulnerability score for vaccine allocation along with the strategies for clustering regions for warehouse selection. A similar epidemic model is also used in the multi-criterion intelligent decision support system suggested by Aggarwal et al. [13]. Most importantly, the resource allocation systems need to be cost-effective in order to prevent extra expenses during an emergency like the COVID-19 pandemic. Thus, recent researchers have proposed a cloud-based smart resource allocation advisor which can identify low-cost high benefit allocations [10]. It has been utilized to analyze the severity of the pandemic with the help of different algorithms and also this system could detect resource demand, adjust allocation schema, learn from the knowledge about the patients as well. Moreover, it is visible that both qualitative and quantitative parameters have been considered when deciding the factors for flattening the growing curves of COVID-19 with the help of ML algorithms. There are possibilities to utilize ML models during an emergency to guide resource allocation-related decisions [24]. Those models have been used in areas like risk monitoring, analyzing disease transmission, and to find people who are less likely to complete the vaccination, etc. In order to analyze the disease, spread patterns, usage of social media platforms as the key source for ML algorithms is also noticeable. For instance, the Tweets have been used to predict influenza transmission where it provides a real-time accurate assessment and predicts the weekly percentage of patients and future influenza activities which help resource planning [2].

With the advancement of technology, researchers are focusing on simulation approaches when analyzing, developing, and testing resource allocation systems. For example, to evaluate the spread of the coronavirus pandemic a simulation-based multi-agent approach is suggested by Aboulaich et al. [25] that uses a compartmental model to identify individuals with the same behavior, and for forecasting the evolution of the disease against preventive measures. The evaluation of alternative healthcare resource allocation possibilities is also essential in emergency situations [17]. Even though a few of these approaches are not introduced particularly for public health emergencies, most of them contain the objective of levelling the resource utilization according to the severity of patients where it has a possible use case in COVID-19 pandemic. Furthermore, several studies have produced solutions through simulation by integrating optimization techniques as well. For example, in order to reduce the door-to-doctor time, an optimal approach was introduced for an emergency department that mainly focuses on allocating human resources using a meta-model-based simulation and an imperialist competitive algorithm for resource optimization within a limited budget [16]. A simulation environment is also suggested for optimal resource planning during the COVID-19 crisis [18]. In there, deep-learning and linear optimization methods have been used for making predictions in resource demand, deaths, and infected cases as well as resource exchange decisions.

With the current pandemic, public health officers face a number of difficulties when collecting close contacts' information for PCR testing. In order to mitigate those difficulties to some extent, a community-based system has been introduced that includes a risk status indication via a QR code by analyzing close contacts and virus hot zones based on GPS technology [9]. This system can also schedule and send notifications about infection testing to people while following the smart city concept and, it also has been integrated with private institutions along with the capability of continuous data visualization. Here, public agents deal

with the web UI and a dashboard and community through the mobile interface. Even though similar approaches were taken in other applications such as China's 'Health Code' and India's 'Aarogya Setu', they had effectively performed only in the initial stage of the pandemic. In most cases, attention on resource allocation is commonly shifted from individual patient-centered to population-centered during a pandemic. Therefore, visualization of the aggregated data is vital. Thus, some studies have used social media to visualize and understand how people are concern and react to the pandemic. For example, topic modelling, an unsupervised ML technique was used on the Tweets by Tao et al. [26] along with an emotion analysis. Another visualization has done on a dashboard according to the interaction with a speech interface which consists of multiple views and tested using the COVID-19 pandemic as the case study [27]. As a gist, we can say that many decision support systems have been proposed addressing health crises and especially for COVID-19 pandemic using various approaches including optimization techniques, simulation approaches, ML techniques, and IoT for allocating both specific or generic resource types.

### *B. Ethical Frameworks and Guidelines for Resource Allocation*

This section focuses on the proposed guidelines and frameworks that have been related to resource allocation morality during health crises. Legal and ethical issues are possible to raise during a health emergency [28]. Therefore, it is a must to have proper principles for resource allocation procedures that help to draw conclusions about transparency, sharing, balance, and communication under the influence of community obligations, community well-being, and good preparedness practice. It is worthwhile to have a rapid response core guidance list to ensure clear, collaborative, consistent, and context-sensitive decisions during a health emergency [29]. Likewise, for the COVID-19 health emergency, a set of recommendations was made considering ethical values such as maximum benefit, equality, and prioritization [30].

The prioritization is solely focused on a few studies. For example, a resource allocation based on multi-criteria dimensions of priorities is suggested by Angelis et al [31]. Most importantly, WHO has introduced ethics and priority settings for resource allocation during the pandemic [32]. In a health crisis, assigning equity weights on vulnerable groups for triaging people such as dementia patients could be considered as a fairer approach [33], [34]. It also argues that the cost-effective analysis is better on resource allocation. Moreover, indicators such as counts, ratios, percentages, proportions are very important in the surveillance of communicable disease [35]. For instance, a field model has been proposed with a set of potential indicators [36] which are community-based and focused on ensuring basic services are reached to the local. The percentages and ratios of elders, smokers, asthma/diabetes patients, medical officers are some of those indicators which help to make decisions. Aside from those, several studies have suggested resource-specific guidelines. For example, a study has recommended steps for allocating ICU equipment according to community engagement while considering moral reference [20]. In addition, a rapid review was also conducted for ICU resources during an infectious disease outbreak [37] along with the evaluation of the validity of the already existing triage tools and ethical frameworks. From the studies that are discussed in this section, ethical value considerations, guidelines & frameworks could be commonly seen with the main objective of maximizing the community benefit.

### *C. Limitations of Existing Resource Allocation Systems*

Despite their benefits, existing resource allocation systems have their own limitations. For example, some of these systems are specifically developed for a particular type of disaster and health crisis. Thus, the existing resource allocation systems proposed by previous researchers are difficult to customize for the COVID-19 pandemic situation in developing countries like Sri Lanka due to the following reasons.

- Healthcare budget constraints lead to extreme material resource limitations
- Poor healthcare infrastructure and differences in procedures
- A separate body for social status tracking and visualization
- Lack of hospital capacities and maintaining quarantine centers
- Secondary difficulties such as resource transportation to remote areas delay resource accessibility
- The inability to store and dispatch resources by establishing temporary facilities
- The limited availability of staff
- The community-dependent systems are not suitable because people of remote areas don't have smart devices or the technology literacy

The above conclusions are drawn by analyzing the existing literature where only a few have explicitly stated the difficulties [38] but in contrast, many have proposed systems for countries that have a high quality of life, stable economies, and technological advancements. Therefore, most of the systems have been proposed, while considering the continuous budget increases, rich healthcare infrastructures, multiple warehouse & emergency center establishments, transportation facilities, smart cities, high availability of healthcare resources [2], [6]-[9], [11], [12], [14], [16], [17], [23], [24], [30], [33]. Hence these systems would not be very much suitable for Sri Lanka.

According to statistics, as of the second week of November 2021, more than 549,500 reported cases, 13,972 deaths and, 523,503 recovered cases have been identified in Sri Lanka. In order to mitigate the further damages of this health crisis, a preparedness and response plan for COVID-

19 has been introduced by the Ministry of Health [39]. It includes resource management and priority measures such as contact tracing, quarantine, and isolation guidelines. Apart from that, a set of provisional guidelines on COVID-19 patients has also been introduced [40] which includes procedures about patient identification, monitoring, admissions, and treatments are stated especially for the health practitioners. Moreover, the way of assessing the severity of patients, level of risk by exposure, conducting PCR tests or rapid Antigen tests, taking the history of patients, usage of essential resources have been emphasized.

In countries like Sri Lanka, the Public Health Inspectors (PHI) and (Medical Officer of Health) MOH are directly dealing with society and, handle all the community-based procedures manually. There are about 60,000 people who live in a particular MOH area, and they receive resources from the regional medical supply division under the coordination of the PHIs. With the accelerated growth of the pandemic, there are a large number of problems occurred in the area of health resource allocation. As the result of extreme demands, there is a possibility that coronavirus disease would overwhelm the healthcare infrastructure in Sri Lanka. It has come to the level where the infected patients are also isolated at homes while only the severe cases are transferred to a hospital. With the skyrocketing COVID infected cases, the admissions of severe patients are also getting late day by day. In the current procedure, when there is a delay in conducting PCR tests and getting the test results, one should contact the PHI officer directly which is very difficult with the current rush. On the other hand, the vaccination campaigns are overcrowded without any prioritization. When a limited number of vaccines are received for a certain MOH area, insufficient vaccine utilization and improper tracking of people for the second dosage could be clearly observed. The disease-related indicators such as asymptomatic and symptomatic rates, case fatality rate, infected percentages by age that are important for MOH officers & high authorities in surveillance and for allocating resources are irregularly traced for each area. In addition, transporting patients, distributing food, and case monitoring are also not handled effectively. These unnecessary delays of healthcare resources may put the patient in a dangerous position and the community in hardships. Therefore, a decision support system for resource allocation is a necessity that combines data and decision logic to assist decision-makers to overcome the above difficulties. Providing timely and accurate decisions regarding resource controlling/coordinating, distributing and monitoring as well as building community awareness should be the main objectives of a system along with automating the tasks handled by the PHI officers in each MOH area from the patient identification to conducting PCR tests for close contacts. In such system, data visualization could be used to visualize the current resource allocations in that particular area which helps to increase social awareness as well. Likewise, the external parties who receive updates would be able to make decisions easily with the use of resource allocation systems.

## V. RESULTS AND DISCUSSION

The analysis of relevant existing systems in support of making resource-related decisions during a pandemic like health crisis have been discussed throughout this study. The keyword-based clustering provided greater insights into the literature during the analysis and, it supported the analysis process by showing up the key areas that previous researchers have focused on. It generated 14 clusters highlighting the recurrent areas as shown in Table I. Some of the dominant keywords identified were related to the scope such as COVID-19, coronavirus, resource allocation, disaster planning, pandemic, critical care, public health, and healthcare. As per the remaining categories of the keywords, several reasoning can be made by analyzing relevant literature. The resource allocation guidelines, policies, legal and medical frameworks are generally established as common components or proactive plans of disaster management [29]-[31], [33], [36]. The resource-related decisions should be taken by considering the maximum community benefit, ethics, and vulnerable groups. In several studies, triage has been considered when allocating resources according to the severity levels of the patients during an emergency [20], [28], [37]. The selection, adaption of these frameworks, and dealing with uncertainties of situations must be more focused based on the health crisis. The simulation and optimization can be also identified as prominent keywords [8], [12], [14], [16]-[18], [25]. The simulation approaches were commonly used in the areas of pandemic spread analysis, events generation & representation, and resource allocation with the help of GIS (Geographic Information System) where some are embedded with allocation optimization techniques as well.

Moreover, the use of ML techniques for the predictions of required resources, pandemic transmission, contact tracing, infected and suspected cases can be commonly seen [7], [13], [18]. When studying the literature, one could identify that the resource allocation is hugely addressed in the response phase of a disaster and there were several interrelated problems attached such as transportation, scheduling, and facility locations. The resource allocation-related decision support systems, health information, and healthcare systems were proposed in numerous studies [9]-[14]. Both disaster-specific and resource-specific systems are mostly focusing on the COVID-19 pandemic. However, the technology has enabled the community to collaborate with public agents during a crisis. Emerging technologies such as Big Data, IoT, facial/voice recognition, and Artificial intelligence have a dominant role in those smart systems.

Table 1. Clustered keywords

| Label | Keyword | Label | Keyword |
|---|---|---|---|
| 0 | COVID-19 | 7 | Pandemic |
| 1 | Medical ethics | 8 | Triage |
| 2 | Resource allocation | 9 | Critical care |
| 3 | Machine learning | 10 | Decision support |

| | | | |
|---|---|---|---|
| 4 | Optimization | 11 | Healthcare systems |
| 5 | Coronavirus | 12 | Public health |
| 6 | Disaster planning | 13 | Simulation |

In this study, different kinds of resource allocation systems were discussed pertaining to the health crises. A detailed overview of the current stand of resource allocation systems was explicitly presented as the main objective of this study. The keyword clustering and PRISMA standard were used to investigate and filter previous studies, in a systematic manner. The allocation of both the critical/non-critical and human/material resources was considered in this study.

## REFERENCES

[1] M. Du, A. Sai, and N. Kong, "A data-driven optimization approach for multi-period resource allocation in cholera outbreak control," *Eur. J. Oper. Res.*, no. xxxx, 2020, doi: 10.1016/j.ejor.2020.09.052

[2] K. Lee, A. Agrawal, and A. Choudhary, "Forecasting Influenza Levels Using Real-Time Social Media Streams," *Proc. - 2017 IEEE Int. Conf. Healthc. Informatics, ICHI 2017*, pp. 409–414, 2017, doi: 10.1109/ICHI.2017.68.

[3] B .News, Coronavirus declared global health emergency by WHO, 2020 (accessed June 6, 2021). [Online]. Available: https://www.bbc.com/news/world-51318246.

[4] A. T. Janke, H. Mei, C. Rothenberg, R. D. Becher, Z. Lin, and A. K. Venkatesh, "Analysis of Hospital Resource Availability and COVID-19 Mortality Across the United States," *J. Hosp. Med.*, vol. 16, no. 4, pp. 211–214, 2021, doi: 10.12788/jhm.3539.

[5] L. Xie, H. Yang, X. Zheng, Y. Wu, X. Lin, and Z. Shen, "Medical resources and coronavirus disease (COVID-19) mortality rate: Evidence and implications from Hubei province in China," *PLoS One*, vol. 16, no. 1 January, pp. 1–13, 2021, doi: 10.1371/journal.pone.0244867.

[6] M. Choksi and M. A. Zaveri, "Multiobjective Based Resource Allocation and Scheduling for Postdisaster Management Using IoT," *Wirel. Commun. Mob. Comput.*, vol. 2019, 2019, doi: 10.1155/2019/6185806.

[7] R. Kondaveti and A. Ganz, "Decision support system for resource allocation in disaster management," *Proc. 31st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. Eng. Futur. Biomed. EMBC 2009*, pp. 3425–3428, 2009, doi: 10.1109/IEMBS.2009.5332498.

[8] Z. Liu, Z. Wang, Y. Ren, Q. Feng, D. Fan, and Z. Zuo, "A City Medical Resources Distribution Optimization Platform Based on Cyber Physical Systems (CPS)," *Proc. - 2018 Int. Conf. Sensing, Diagnostics, Progn. Control. SDPC 2018*, pp. 269–273, 2019, doi: 10.1109/SDPC.2018.8664843.

[9] L. A. do Araujo, Adelson de ; Garcia, Israel ; Cacho, Nelio; Nascimento, "A platform for citizen cooperation during the COVID-19 pandemic in RN, Brazil," vol. 1, no. 12, p. 56, 2020.

[10] O. Cosgun and A. Umar, "Smart resource allocation advisor in the cloud for COVID-19 and other pandemics," *Proc. - 2020 IEEE Cloud Summit, Cloud Summit 2020*, pp. 124–130, 2020, doi: 10.1109/IEEECloudSummit48914.2020.00025.

[11] F. Cibella, S. Panunzi, V. Cusimano, and A. De Gaetano, "Decision support for medical disasters: Evaluation of the IMPRESS system in the live Palermo demo," *Int. J. Disaster Risk Reduct.*, vol. 50, p. 101695, 2020, doi: 10.1016/j.ijdrr.2020.101695.

[12] P. Vxsso, F. Lq, K. V Vwhpv, and D. Wkh, "Optimization of mechanical ventilator distribution in response to pandemics similar to COVID-19," 2020.

[13] L. Aggarwal, P. Goswami, and S. Sachdeva, "Multi-criterion Intelligent Decision Support system for COVID-19," *Appl. Soft Comput.*, vol. 101, p. 107056, 2021, doi: 10.1016/j.asoc.2020.107056.

[14] S. Roy, R. Dutta, and P. Ghosh, "Optimal time-varying vaccine allocation amid pandemics with uncertain immunity ratios," *IEEE Access*, vol. 9, pp. 15110–15121, 2021, doi: 10.1109/ACCESS.2021.3053268.

[15] C. K. Lo, H. C. Chen, P. Y. Lee, M. C. Ku, L. Ogiela, and C. H. Chuang, "Smart dynamic resource allocation model for patient-driven mobile medical information system using C4.5 algorithm," *J. Electron. Sci. Technol.*, vol. 17, no. 3, pp. 231–241, 2019, doi: 10.11989/JEST.1674-862X.71018117.

[16] M. Yousefi and M. Yousefi, "Human resource allocation in an emergency department: A metamodel-based simulation optimization," *Kybernetes*, vol. 49, no. 3, pp. 779–796, 2019, doi: 10.1108/K-12-2018-0675.

[17] L. Bedoya-Valencia and E. Kirac, "Evaluating alternative resource allocation in an emergency department using discrete event simulation," *Simulation*, vol. 92, no. 12, pp. 1041–1051, 2016, doi: 10.1177/0037549716673150.

[18] N. Petrovic, "Simulation Environment for Optimal Resource Planning during COVID-19 Crisis," *2020 55th Int. Sci. Conf. Information, Commun. Energy Syst. Technol. ICEST 2020 - Proc.*, pp. 23–26, 2020, doi: 10.1109/ICEST49890.2020.9232908.

[19] J. P. Leider, D. Debruin, N. Reynolds, A. Koch, and J. Seaberg, "Ethical guidance for disaster response, specifically around crisis standards of care: A systematic review," *Am. J. Public Health*, vol. 107, no. 9, pp. e1–e9, 2017, doi: 10.2105/AJPH.2017.303882.

[20] M. R. Ehmann *et al.*, "Operational Recommendations for Scarce Resource Allocation in a Public Health Crisis," *Chest*, vol. 159, no. 3, pp. 1076–1083, 2021, doi: 10.1016/j.chest.2020.09.246.

[21] A. D. D Liberati A, Tetzlaff J, "Preferred reporting items for systematicreviews and meta-analyses: The prisma statement,"The PRISMA Group(2009), 07 2009.

[22] I. Sung and T. Lee, "Optimal allocation of emergency medical resources in a mass casualty incident: Patient prioritization by column generation," *Eur. J. Oper. Res.*, vol. 252, no. 2, pp. 623–634, 2016, doi: 10.1016/j.ejor.2016.01.028.

[23] A. M. Abd-Alrahman and L. Ekenberg, "Modelling health information during catastrophic events - A disaster management system for Sudan," *2017 IST-Africa Week Conf. IST-Africa 2017*, pp. 1–9, 2017, doi: 10.23919/ISTAFRICA.2017.8102390.

[24] A. Bhansali and V. Jain, "Using Machine Learning for Healthcare Resource Allocation in COVID-19: Opportunities and Challenges for LMICs | Center For Global Development," no. Ml, pp. 1–9, 2020, [Online]. Available: https://www.cgdev.org/blog/using-machine-learning-healthcare-resource-allocation-covid-19-opportunities-and-challenges.

[25] R. Aboulaich, K. Bensaid, S. Chabbar, and J. El Karkri, "Mathematical modeling and multi-agents approach for the evolution of the Coronavirus pandemic," *2020 IEEE Int. Conf. Technol. Manag. Oper. Decis. ICTMOD 2020*, pp. 1–6, 2020, doi: 10.1109/ICTMOD49425.2020.9380596.

[26] G. Tao, Y. Miao, and S. Ng, "COVID-19 Topic Modeling and Visualization," *Proc. Int. Conf. Inf. Vis.*, vol. 2020-Septe, no. Iv, pp. 734–739, 2020, doi: 10.1109/IV51561.2020.00129.

[27] I. Chowdhury, A. Moeid, E. Hoque, M. A. Kabir, M. S. Hossain, and M. M. Islam, "Designing and Evaluating Multimodal Interactions for Facilitating Visual Analysis with Dashboards," *IEEE Access*, vol. 9, pp. 60–71, 2021, doi: 10.1109/ACCESS.2020.3046623.

[28] D. J. Barnett, H. A. Taylor, J. G. Hodge, and J. M. Links, "Resource allocation on the frontlines of public health preparedness and response: Report of a summit on legal and ethical issues," *Public Health Rep.*, vol. 124, no. 2, pp. 295–303, 2009, doi: 10.1177/003335490912400218.

[29] L. Frank, T. Concannon, and K. Patel, "Health Care Resource Allocation Decisionmaking During a Pandemic," *Heal. Care Resour. Alloc. Decis. Dur. a Pandemic*, 2020, doi: 10.7249/rra326-1.

[30] E. J. Emanuel *et al.*, "Fair Allocation of Scarce Medical Resources in the Time of Covid-19," *N. Engl. J. Med.*, vol. 382, no. 21, pp. 2049–2055, 2020, doi: 10.1056/nejmsb2005114.

[31] A. Angelis, P. Kanavos, and G. Montibeller, "Resource Allocation and Priority Setting in Health Care: A Multi-criteria Decision Analysis Problem of Value?," *Glob. Policy*, vol. 8, no. March, pp. 76–83, 2017, doi: 10.1111/1758-5899.12387.

[32] World Health Organization, "Ethics and COVID-19 : resource allocation and priority-setting," *World Heal. Organ.*, pp. 1–5, 2020, [Online]. Available: https://www.who.int/ethics/publications/ethics-covid-19-resource-allocation.pdf?ua=1.

[33] C. Roadevin and H. Hill, "How can we decide a fair allocation of healthcare resources during a pandemic?," *J. Med. Ethics*, pp. 1–4, 2021, doi: 10.1136/medethics-2020-106815.

[34] E. E. Smith *et al.*, "Pandemic dementia scarce resource allocation," *Can. Geriatr. J.*, vol. 23, no. 3, pp. 260–262, 2020, doi: 10.5770/CGJ.23.457.

[35] WHO, "Communicable disease surveillance and response systems. A guide to Planning," *World Health*, p. 33, 2006.

[36] M. A. S. Jane S. Durch, Linda A. Bailey, "Improving Health in the Community: A Role for Performance Monitoring," in *Health Care Resource Allocation*, no. 1997, 2021, pp. 1–17.

[37] K. M. Fiest *et al.*, "Allocation of intensive care resources during an infectious disease outbreak: a rapid review to inform practice," *BMC Med.*, vol. 18, no. 1, pp. 1–17, 2020, doi: 10.1186/s12916-020-01871-9.

[38] W. H. Organisation, "Strengthening health systems : Strengthening health systems," *Heal. (San Fr.*, 2004.

[39] S. L. Ministry of Health, "Sri Lanka Preparedness & Response Plan COVID-19," *Minist. Heal. Indig. Med. Serv.*, no. April, 2020, [Online]. Available: http://www.health.gov.lk/moh_final/english/others.php?spid=7.

[40] Ministry of Health Sri Lanka, "Provisional Clinical Practice Guidelines on COVID-19 suspected and confirmed patients," *Minist. Heal. Sri Lanka*, no. March, p. 46, 2020, [Online]. Available: http://www.epid.gov.lk/web/images/pdf/Circulars/Corona_virus/covid-19 cpg _ version 4.pdf.

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index No: TL-26-36*

# Heart Disease Prediction Using Machine Learning Techniques: A Comparative Analysis

L. Gamage (1st Author)
*Department of Software Engineering*
*University of Kelaniya*
Kelaniya, Sri Lanka
*Postgraduate Institute of Science*
*University of Peradeniya*
Peradeniya, Sri Lanka
lgama191@kln.ac.lk

***Abstract* — In today's world, heart disease is one of the leading causes of death. In clinical data analysis, predicting heart disease is a difficult task. Machine Learning (ML) helps assist with the decision-making and prediction of large volumes of data generated by the healthcare industry. The main goal of this study is to find the best performance model and compare machine learning algorithms for predicting heart disease. This work applies supervised machine learning algorithms, namely Logistic Regression, Support Vector Machine, K-Nearest Neighbor, and Random Forest, to the Cleveland Heart Disease dataset to predict heart disease. Our experimental analysis using preprocessing steps and model hyperparameter tuning, Logistic Regression, Support Vector Machine, K- Nearest Neighbor and Random Forest achieved 90.16%, 86.89%,86.89%, and 85.25%, accuracies respectively. As a result, Logistic Regression classification outperforms other machine learning algorithms in predicting heart disease.**

***Keywords* — *Machine Learning, Heart Disease Prediction, Logistic Regression, Support Vector Machine, K- Nearest Neighbor, Random Forest***

## I. INTRODUCTION

The heart is the most important organ in the human body. It is at the center of the human circulatory system, a network of blood vessels that delivers blood to every part of the human body. Blood carries oxygen and other essential nutrients that all body organs need to stay healthy and work properly. If it fails to function correctly, then the brain and various other organs will stop working, and within a few minutes, a person can die. In today's world, people are getting very busy in their lives and work so much that they do not have time to take care of themselves. Due to this hectic lifestyle, they experience stress, anxiety, depression, and many more conditions. These factors contribute to people being ill and having severe diseases. There are many diseases such as cancer, diabetes, heart diseases, tuberculosis, Etc., which lead to the death of people each year. Due to a hectic lifestyle and not having good food habits, heart-related diseases increase among people more than any other disease. Also, some factors, diseases, and habits influence increasing the risk of having heart diseases. They are Age, Family History, Diabetes, High blood pressure, high cholesterol, being obese, depression, hypertension, smoking, or too much drinking and physical inactivity. So, taking care of the heart as well as other organs becomes a challenge in this era. Apart from the factors, there are different types of heart disease, such as Angina, Arrhythmia Fibrillation, Congenital heart disease, Coronary artery disease, Myocardial infarction, and Heart failure. To reduce the heart disease caused death, it is mandatory to early detection of heart disease and treatment on time. So, the Importance of Prediction of heart-related diseases can significantly impact the medical field and people's lives. Now day's heart disease is the main reason for deaths throughout the world. Many people die from heart disease than from any other disease. According to the World Health Organization, for the past 20 years, heart disease has been the leading cause of death at the global level[1]. The number of deaths from heart disease increased by more than 2 million since 2000, to nearly 9 million in 2019. Heart disease currently accounts for 16% of total deaths from all causes. Even though heart disease is the riskiest disease globally, people are unaware of the risks and symptoms of heart disease. Therefore, the prediction of heart disease is a significant concern of humankind. The foremost challenge facing healthcare organizations is the provision of quality services at a reasonable cost. Quality service indicates diagnosing patients correctly and administering effective treatments. Most of the clinical decisions are made based on the Doctor's perception and practice rather than the hidden knowledge on the patient database. This practice may lead to mistakes and extreme medical expenses, which affect the quality of health care services. It is frequently difficult for medical practitioners to predict heart diseases as it requires experience and knowledge, which is very difficult to accomplish. In that case, machine learning techniques help the healthcare industry and the professionals in diagnosing heart diseases based on patients' data.

In the medical domain, many medical datasets help make compelling predictions. These data can be exploited using

machine learning techniques to extract hidden information from hidden patterns of datasets. Moreover, this extractive data will help to predict the medical diagnosis. The collected medical data are massive in size, and it can be noisy. This data which is too complicated for the human mind to understand can be easily explored using machine learning techniques. The future predictions will also help the doctors to diagnose the disease by discovering previous dataset patterns and taking the right step to treat the patients. It will save humankind and increase the quality of health care services. Several research studies have been conducted on predicting heart disease. They used several machine learning techniques to predict heart disease and achieved different results using different methods. [4]Their work on "Cardiovascular Disease Forecast using Machine Learning Paradigms" used machine learning algorithms and achieved 86.25% accuracy from the logistic regression model, giving the best accuracy among all four models. [9] Their work on" Heart Disease Prediction Using Machine Learning Algorithms" summarized the recent research with comparative results that have been done on heart disease prediction and also make analytical conclusions. Their experimental results show that the Decision Tree Classifier algorithm has the most precise and significant result compared to the others algorithm. Also, they mentioned that the predictions that are carried out with large datasets make better accuracy.

This study is for analyzing and finding an appropriate model for the heart disease data. Machine learning models are used to predict heart disease while conducting a comparative analysis on models. In this work, supervised machine learning algorithms, namely Logistic regression (LR), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and Random Forest (RF), are used to predict heart disease and compare their performance.

## II. OBJECTIVES

The goal of this study is to find out the best model that predicts whether the patient has heart disease. This model will provide important insights to doctors who can then adapt their diagnosis and treatment per-patient basis. This study is to be done on supervised machine learning classification.

The research objectives (ROs) of the research work are stated in the following.

**RO1:** To predict heart disease using selected four Supervised machine learning techniques and analyze the models; Logistic regression, Support Vector Machine, K-nearest neighbor, and Random Forest.

**RO2:** To find the best model that predicts heart disease more accurately among all the models.

**RO3:** To conduct a comparative analysis of machine learning models based on their performances.

## III. METHODOLOGY

This section illustrates various resources and approaches that used in this study. Primarily, the description of the dataset is provided to understand how to work on it, followed by the preprocessing steps involved. Finally, the methodologies used for the experiment in this study are discussed.

In this section, the research materials and methodologies are presented and discussed in brief.

### *A. Dataset Description*

In this work-study, the Cleveland Heart Disease dataset [4] has been collected from the UCI machine learning repository that has been used for both training and testing purposes. The dataset is a collection of medical analytical reports with values for 76 attributes and 303 rows, but this work considers a feature subset of 14 numerical valued attributes. The output level has two classes, where 0 represents not having heart disease, and one represents having heart disease. The information on the heart disease dataset is given

**Table**, where the attribute name and description are presented.

### *B. Data Preprocessing*

In this step, data preprocessing is applied to identify the missing values, process the noisy, incomplete, irreverent, and inconsistent values, and remove some attributes' redundancy. Then separation, feature scaling, and normalization are performed to find the standard format of data. After data preparation, the dataset is divided into a training set (80% of data) and a test set (20% of data).

Table 1. Attributes of the dataset and their description

| Attribute Name | Attribute Description |
|---|---|
| age | age in years |
| sex | 1=Male,0=Female |
| cp | Chest pain type.<br>0: asymptomatic,<br>1: atypical angina,<br>2: non-anginal pain,<br>3: typical angina |
| trestbos | Resting blood pressure (in mmHg) |
| chol | Serum Cholesterol (in mg/dl) |
| fbs | Fasting blood sugar> 120 mg/dl.1 = true; 0 = false |
| restecg | Resting Electrocardiography results.<br>0: showing probable or definite left ventricular hy pertrophy by Estes' criteria,<br>1: normal,<br>2: having ST-T wave abnormality<br>(T wave inversions and/or ST elevation or depres sion of > 0.05 mV) |
| thalach | Maximum heart rate achieved |
| exang | Exercise-induced angina ;(1 = yes; 0 = no) |

| Attribute Name | Attribute Description |
|---|---|
| oldpeak | Depression induced by exercise relative to rest |
| slope | The slope of the peak exercise segment.<br>0: down sloping.<br>1: flat.<br>2: upsloping |
| ca | Number of major vessels colour by fluoroscopy that ranges between 0 and 3 |
| thal | Results of the blood flow were observed via the r adioactive dye.<br>0=Normal (Thallium test.)<br>1=fixed defect<br>2=reversible defect |
| target | 0 = No heart disease,<br>1 = heart disease |

### *C. Model Generation*

In this stage, machine learning algorithms are applied to the training set to develop different classification models. Models are developed using four machine learning algorithms named LR, KNN, SVM and RF classifiers. Then using these four generated models, the test set is classified and evaluated the performance.

In machine learning, hyperparameter tuning is one of the most significant research issues. If the hyperparameters are tuned or optimized, then it is considered that the machine learning algorithms can give better performance. Models are developed with the help of grid search, and cross-validation approaches, hyperparameters are optimized and tuned. It considers cross-validation to guide the performance metrics. Grid search is an exhaustive search that can exercise to compute the optimal values of hyperparameters. It can build a model that generates every parameter combination and stores each combination of the model. The efforts and resources can be saved using this search. Then with the tuned parameters, the LR, KNN, SVM and RF classifier models are generated. After the generation of the classification model, the test set is applied to the proposed model with the tune hyperparameter and evaluated the test set's performance.

The flowchart of the model is shown in Fig. 1 below. It summarizes the steps for our proposed method.

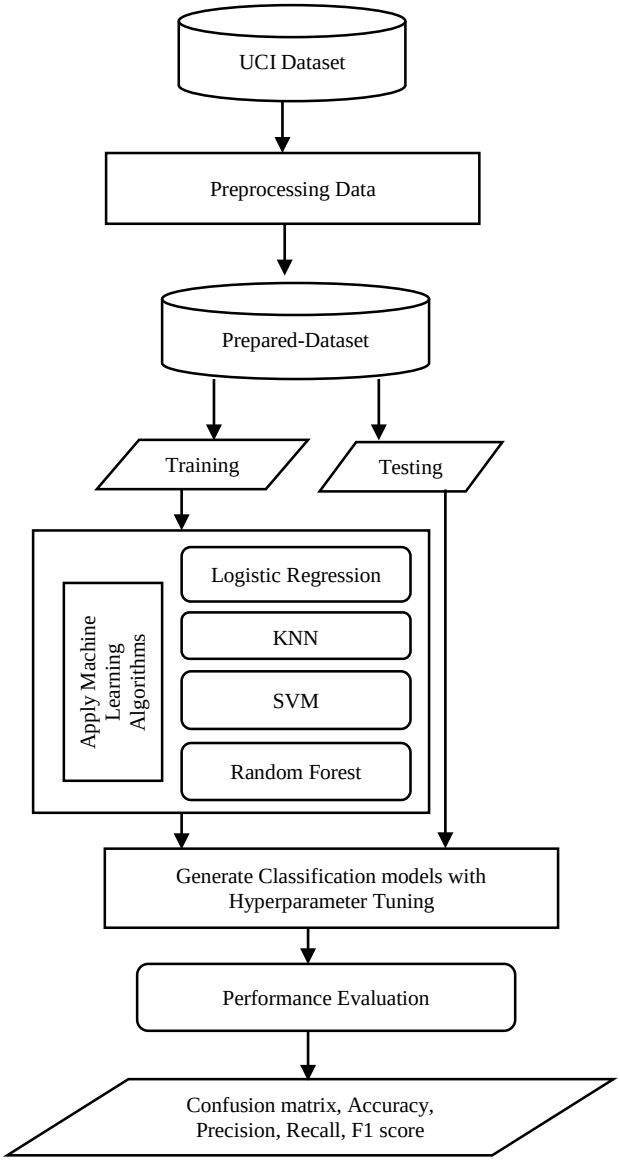



Fig. 1. Flowchart of proposed work

### *D. Machine Learning Algorithms*

This section explains the supervised algorithms of Machine learning that are used in this work. In the model generation process, Logistic regression (LR), K nearest neighbor (KNN), Support vector machine (SVM) and Random Forest (RF) classifiers are used as machine learning algorithms.

#### *1) Logistics Regression*

Logistic regression [5] is a supervised learning machine used for regression and classification problems. Logistic regression uses probability to predict the classification of categorical data. It is mainly used for a binary classification problem and a logistic function to predict a binary dependent variable [6].

#### *2) KNN*

K-Nearest Neighbor (KNN) [2] is also a supervised classification algorithm. It predicts the target class based on how similar that data is from other provided training data labels to the model. KNN is used widely in the machine learning classification problem. It is simple to understand and generates a non-parametric model that is applied to practical problems. It is a lazy learner or instance-based learner, which depend on the distance. It works well but does not learn any classification rule[5].

*3) SVM*

Support Vector Machine [6] is a prevalent supervised machine learning technique (having a pre-defined target variable) that can be used as a classifier and a predictor. For classification, it finds a hyper-plane in the feature space that differentiates between the classes. In SVM, a hyper-plane is created margin as wide as possible to separate different types of data or keep similar data of one type at one side and similar data of another type of data at another side of the margin [5].

*4) Random Forest*

Random forest (RF) [2]is a supervised machine learning algorithm. As the name predicts, it is a forest of randomly generated decision trees. It uses an approach bagging, where various learning models are combined to improve the overall results. To perform the bagging operation, it produces manifold decision trees and synthesizes them together to obtain a refined result. It is one of the finest machine learning algorithms. It uses a random subset of features by splitting a node to get the best feature that contributes the most to building the model.

## *E. Performance Evaluation*

Performance evaluation of the proposed work is done based on the following measures.

*1) Confusion Matrix*

Confusion Matrix [2] is a matrix that is used to evaluate the performance of a model. The four terms associated with the confusion matrix which is used to determine the performance matrices are:

True Positive (TP): An outcome when the model correctly predicts the positive class

True Negative (TN): An outcome when the model correctly predicts the negative class

False Positive (FP): An outcome when the model incorrectly predicts the positive class

False Negative (FN): An outcome when the model incorrectly predicts the negative class

In this step, models evaluate the performance of the training set and test set and find the confusion matrix. Then the performance metrics of these two models have been calculated and assessed in terms of accuracy, precision, recall, and F1 Score with the help of the confusion matrix. The mathematical expressions of accuracy, precision, recall and F1-Score are shown in Equation (1), Equation (2), Equation (3) and Equation (4) respectively.

*2) Accuracy*

Accuracy is the ratio of the number of correct predictions given by the model to the total number of instances[2].

$$\text{Accuracy} = (TP+TN) / (TP+TN+FP+FN) \quad (1)$$

*3) Precision*

Precision in this work measures the proportion of individuals predicted to be at risk of developing heart disease and had a risk of developing heart disease.

$$\text{Precision} = TP / (TP+FP) \quad (2)$$

*4) Recall*

Recall, in this work, measures the proportion of individuals that were at risk of developing heart disease and were predicted by the algorithm to be at risk of developing heart disease.

$$\text{Recall} = TP / (TP+FN) \quad (3)$$

*5) F1-Score*

F1 Score is the harmonic mean of precision and recall.

$$\text{F1 score} = 2* (\text{Recall}*\text{Precision}) / (\text{Recall} + \text{Precision}) \quad (4)$$

# IV. RESULTS AND DISCUSSION

## *A. Results of Data Preprocessing*

The heart disease dataset consists of 303 samples with 14 attributes, where 138 are healthy (0) instances and 165 instances (1) having heart disease. In the preprocessing step, the statistical operations have been performed to identify and remove the missing values and to find the maximum, minimum, mean, and standard deviation of each feature set.

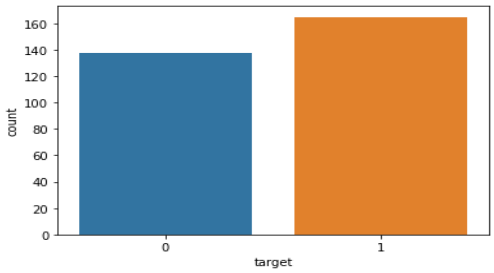



Fig.2.Count of target Variables

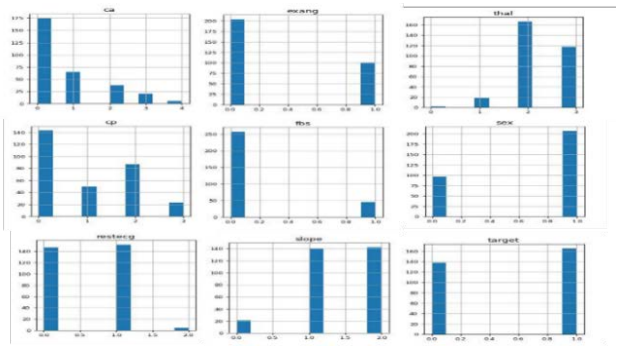


Fig.**3**. Histogram of categorical valued attributes

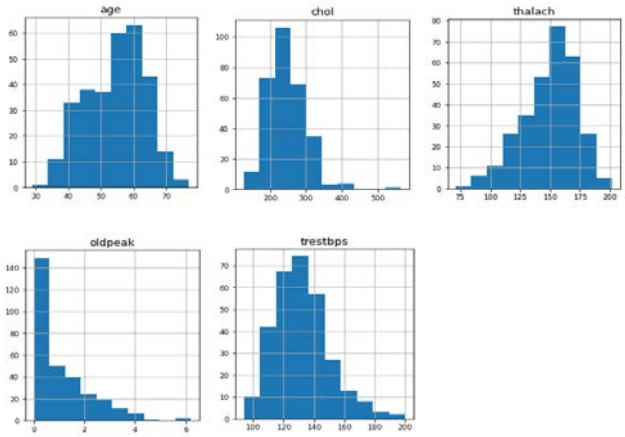



Fig. 4. Histogram of continuous-valued attributes

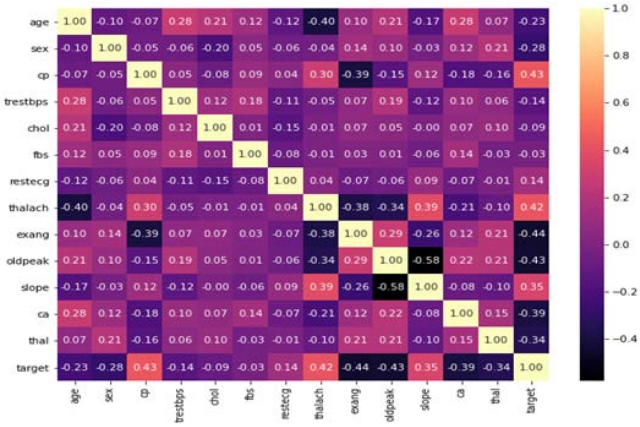



Fig. 5. The heat map for correlation features of the heart disease dataset

Consequently, several non-heart disease and heart disease patients are illustrated in **Fig. 2;** they are 138 and 165. This further confirms that the two classes are almost balanced, and it is good to proceed with data preprocessing. Then the histogram of categorical and continuous features has been plotted for easy and better understanding. The histogram plots are presented for the pattern and frequency distribution of categorical and continuous measurements of data. The distribution of each feature value is shown in Fig. and Fig. **4** as a histogram plot. It can help to identify the trend and patterns of data to understand the distribution of features.

The histogram plot of continuous variables shows how each feature and label is distributed along with different ranges, which further confirms the need for scaling data. The discrete plots in figure 5 show that each of these is categorical variables, and it confirms the need of converting them into numerical variables before training models.

Fig. 5 represents the heat map, which describes the co-relation among the features of the heart disease dataset. Here, different colors have been used to represent the values on the two-dimensional surface. It's clear that no one feature has a particularly strong relationship with the target variable, and categorical-valued attributes are more concentrated than continuous-valued attributes. The heat map of the heart disease dataset shows the hierarchical clustering and a general view of numeric data. After investigating the dataset, the categorical valued attributes have been converted into dummy attributes. Then, centering and scaling operations have been performed to standardize each feature by computing the relevant statistics on the dataset. The resultant dataset has been divided into a training set and a test set, with an 80% and 20% split.

## *B. Experiment Results of the Classifiers*

### *1) Performance Evaluation of the Classifiers with Default Hyper-Parameters*

In this experiment, the machine learning algorithms are applied with the default parameters. Table 2 shows the result of this system. In the training phase, Logistic Regression is fitted and executed the model with parameters of C=1 and solver= 'liblinear' and found 86.78%,86.33%,90.22% and 88.23% of accuracy, precision, recall, and F1 score respectively. The test set is predicted on this LR model and provides 90.16%, 93.30%,87.50% and 90.32% of accuracy, precision, recall, and F1 score, respectively. Again, in the training phase, KNN is fitted and executed the model with the parameters of no. of neighbor=7 and weights= 'uniform' and found 82.23%, 82.60%, 85.71% and 84.13% of accuracy, precision, recall, and F1 score respectively. The test set is predicted on this KNN model and provides 86.89%,90.00%,84.37%and 87.09%, of accuracy, precision, recall, and F1-score respectively

In another training model, SVM is fitted and executed with the parameters of C= 1.0, gamma= 0.01, and 'rbf' kernel, and found 82.23%,82.14%,86.46% and 84.25% of accuracy, precision, recall, and F1 score respectively. The test set is predicted on this SVM model and provides 83.61%, 89.28%, 78.13% and 83.33% of accuracy, precision, recall, and F1 score, respectively. In the last phase of training, Random Forest is fitted and executed with the parameter estimators = 100, min_samples_leaf=1, min_samples_split=2 and found 100%, 100%, 100%, and 100% of accuracy, precision, recall, and F1 score respectively. The test set is predicted on this Random Forest model and provides 81.97%,86.20%,78.12% and 81.97% of accuracy, precision, recall, and F1-score, respectively.

Table 2. Performance of evaluation and comparison of classification models on the training set and the test set

| Algorithm | Training Dataset | | | | Test Dataset | | | |
|---|---|---|---|---|---|---|---|---|
| | Accuracy (%) | Precision (%) | Recall (%) | F1- score (%) | Accuracy (%) | Precision (%) | Recall (%) | F1- score (%) |
| LR | 86.7 | 86.3 | 90.2 | 88.2 | 90.1 | 93.3 | 87.5 | 90.3 |
| KNN | 82.2 | 82.6 | 85.7 | 84.1 | 86.8 | 90.0 | 84.3 | 87.0 |
| SVM | 82.2 | 82.1 | 86.4 | 84.2 | 83.6 | 89.2 | 78.1 | 83.3 |
| RF | 100 | 100 | 100 | 100 | 81.9 | 86.2 | 78.1 | 81.9 |

*2) Performance Evaluation of the Classifiers with Tuned Hyper-Parameters*

In this experiment, the Grid search is used to find the optimal hyperparameters. After tuning the hyperparameters, the classification models are generated. Table 3 shows the result of the proposed system. In the training phase, Logistics Regression is fitted and executed with the tuned hyperparameters of C=1.08, penalty =' l2' and solver= 'liblinear' and found 86.78%, 86.33%, 90.22% and 88.23% of accuracy, precision, recall, and F1 score respectively. The test set is predicted on this LR model and provides 90.16%,93.33%,87.50% and 90.32% of accuracy, precision, recall, and F1 score, respectively. Again, in the training phase, KNN is fitted and executed with the tuned hyperparameters of no. of neighbor=5 and weights=' uniform', metric: 'manhattan' and found 88.43%, 87.76%, 91.72% and 89.70% of accuracy, precision, recall, and F1 score respectively. The test set is predicted on this KNN model and provides 86.89%,87.50%,87.50%, and 87.50% of accuracy, precision, recall, and F1 -core, respectively.

In another training model, SVM is fitted and executed with the tuned hyperparameters of C= 2.0, gamma= 0.1 and RBF kernel, and found 92.98%,92.03%,95.48% and 93.72% of accuracy, precision, recall, and F1 score respectively. The test set is predicted on this SVM model and provides 86.89%, 90.00%,84.37% and 84.37% accuracy, precision, recall, and F1 score, respectively. In the last phase of training, RF is fitted and executed with the tuned hyperparameters of 1000 no of estimators, two minimum samples of the leaf, and five minimum samples the split, maximum depth = 50, maximum features = square, and found 96.69%, 94.96%,99.24% and 97.05% of accuracy, precision, recall, and F1- score respectively. The test set is predicted on this RF model and provides 85.25%,89.65%,81.25% and 85.24% of accuracy, precision, recall, and F1-score, respectively.

Table 3. Performance of evaluation and comparison of classification models with a hyperparameter tuning approach on the training set and the test set

| Algorithm | Training Dataset | | | | Test Dataset | | | |
|---|---|---|---|---|---|---|---|---|
| | Accuracy (%) | Precision (%) | Recall (%) | F1- score (%) | Accuracy (%) | Precision (%) | Recall (%) | F1- score (%) |
| LR | 86.7 | 86.3 | 90.2 | 88.2 | 90.1 | 93.3 | 87.5 | 90.3 |
| KNN | 88.4 | 87.7 | 91.7 | 89.7 | 86.8 | 87.5 | 87.5 | 87.5 |
| SVM | 92.9 | 92.0 | 95.5 | 93.7 | 86.8 | 90.0 | 84.3 | 84.3 |
| RF | 96.6 | 94.9 | 99.2 | 97.0 | 85.2 | 89.6 | 81.2 | 85.2 |

*3) Performance Comparison of the Classifiers with default and tuned Hyper- Parameters*

Table **4** represent the performance comparison between without and with the hyperparameters tuning approach of four machine learning algorithms.

Table 4. Accuracy comparison

| Algorithm | Accuracy (%) of Training Dataset | | Accuracy (%) of Test Dataset | |
|---|---|---|---|---|
| | Without parameter tuning | With Hyperparameter tuning | Without parameter tuning | With Hyperparameter tuning |
| LR | 86.7 | 86.7 | 90.1 | 90.1 |
| KNN | 82.2 | 88.4 | 86.8 | 86.8 |
| SVM | 82.2 | 92.9 | 83.6 | 86.8 |
| RF | 100.0 | 96.6 | 81.9 | 85.2 |

These comparisons show that the models with tuned hyperparameters provide better accuracy results. It also noticed that the LR and KNN accuracies remain the same while SVM and Random Forest have increased their test accuracy.

## V. CONCLUSION

Heart disease is one of the significant deaths anywhere in the world. Early detection of heart diseases will increase the survival rate; hence this work-study is intended to predict whether the patient has heart disease or not with the help of clinical data, which will assist the diagnosis process. However, machine learning techniques are helpful to predict the output from existing data. The results of this study confirm the application of machine learning algorithms in the prediction and early detection of heart disease. To our best understanding, the models built with tuned hyperparameters exhibit better accuracy than the models with default hyperparameters. The prediction accuracy of our proposed models reaches 90.16 % in heart disease detection using Logistic Regression, 86.89% in Support Vector Machine classifier,86.89% using KNN classifier and 85.25% using Random Forest Classifier. The experimental results show that the Logistic Regression algorithm predicts heart disease with the highest performance measures in terms of accuracy of 90.16% and other evaluating metrics.

The future aspect of this study will be to implement the model with the feature selection approach. Furthermore, new feature selection methods can be developed to get a broader perception of the significant features to increase the performance of heart disease prediction.

## REFERENCES

[1] "WHO reveals leading causes of death and disability worldwide: 2000-2019." https://www.who.int/news/item/09-12-2020-who-reveals-leading-causes-of-death-and-disability-worldwide-2000-2019 (accessed May 28, 2021).

[2] "UCI Machine Learning Repository: Heart Disease Data Set." https://archive.ics.uci.edu/ml/datasets/Heart+Disease (accessed May 31, 2021).

[3] R. Katarya and S. K. Meena, "Machine Learning Techniques for Heart Disease Prediction: A Comparative Study and Analysis," *Health and Technology*, vol. 11, no. 1, pp. 87–97, Jan. 2021, doi: 10.1007/s12553-020-00505-7.

[4] S. Islam, N. Jahan, and M. E. Khatun, "Cardiovascular Disease Forecast using Machine Learning Paradigms," in *Proceedings of the 4th International Conference on Computing Methodologies and Communication, ICCMC 2020*, Mar. 2020, pp. 487–490. doi: 10.1109/ICCMC48092.2020.ICCMC-00091.

[5] E. Kabir Hashi and M. Shahid Uz Zaman, "Developing a Hyperparameter Tuning Based Machine Learning Approach of Heart Disease Prediction," *Journal of Applied Science & Process Engineering*, vol. 7, no. 2, 2020.

[6] V. v. Ramalingam, A. Dandapath, and M. Karthik Raja, "Heart disease prediction using machine learning techniques: A survey," *International Journal of Engineering and Technology(UAE)*, vol. 7, no. 2.8 Special Issue 8, pp. 684–687, 2018, doi: 10.14419/ijet.v7i2.8.10557.

[7] S. Mohan, C. Thirumalai, and G. Srivastava, "Effective heart disease prediction using hybrid machine learning techniques," *IEEE Access*, vol. 7, pp. 81542–81554, 2019, doi: 10.1109/ACCESS.2019.2923707.

[8] Galgotias University. School of Computing Science and Engineering, Institute of Electrical and Electronics Engineers. Uttar Pradesh Section, and Institute of Electrical and Electronics Engineers, *2018 4th International Conference on Computing Communication and Automation (ICCCA)*.

[9] V. Rachapudi, S. S. Vaddi, R. R. Karumuri, and S. Sripurapu, "Heart disease prediction using machine learning algorithms," *International Journal of Recent Technology and Engineering*, vol. 7, no. 5, pp. 805–809, Feb. 2019, doi: 10.32628/cseit206421.

[10] S. Anitha and N. Sridevi, "HEART DISEASE PREDICTION USING DATA MINING TECHNIQUES," 2019. [Online]. Available: www.ijaconline.com,

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index No: TL-27-41*

# Road Accident Severity Prediction in Mauritius using Supervised Machine Learning Algorithms

Jameel Ahmad Sowdagur (1st Author)
*University of Technology Mauritius (UTM)*
*Pointe-aux-Sables*
Republic of Mauritius.
jsowdagur@umail.utm.ac.mu

B. Tawheeda B. Rozbully-Sowdagur (2nd Author)
*University of Technology Mauritius (UTM)*
*Pointe-aux-Sables*
Republic of Mauritius.
brozbullysowdagur@umail.utm.ac.mu

Geerish Suddul (3rd Author)
*University of Technology Mauritius (UTM)*
*Pointe-aux-Sables*
Republic of Mauritius.
g.suddul@umail.utm.ac.mu

***Abstract* — Road accidents with high severities are a major concern worldwide, imposing serious problems to the socio-economic development. Several techniques exist to analyse road traffic accidents to improve road safety performance. Machine learning and data mining which are novel approaches are proposed in this study to predict accident severity. Support Vector Machine (SVM), Gradient Boosting (GB), Logistic Regression (LR), Random Forest (RF) and Naïve Bayes (NB) were applied to perform effective data analysis for informed decisions using Python programming language. The gradient boosting outperformed all the other models in predicting the severity outcomes, yielding an overall accuracy of 83.2% and an AUC of 83.9%.**

***Keywords* — *Data mining, machine learning, accident severity***

## I. INTRODUCTION

Road accidents comprise significant public health and development threats. Each year, 1.35 million deaths occur due to road traffic accidents worldwide and in many countries, it accounts for 3% of the Gross Domestic Product (GDP) [1]. Among the victims, more than half are young adults aged between 15 and 44. An increase in fatal road accidents is anticipated by the World Health Organization (WHO) if road safety is not addressed [1]. A National Observatory for road safety in Mauritius for the period 2016-2025 has been set up by the concerned Ministry with one of its objectives of curbing the number of road fatalities to 50% by the year 2025 in line with the target set by the WHO [2]. However, figures published by Statistics Mauritius indicates 21.4% increase in road accidents from the year 2013 to 2020 resulting in a rise of the fatality index from 3.8 to 4.5 during the same period. This increasing trend has made it a necessity to handle the problem in a more scientific approach.

This paper presents road accident severity prediction using a machine learning approach and evaluation of the classification models to identify the most performant one among support vector machine, gradient boosting, logistic regression, random forest and naïve Bayes.

The remaining part of the paper is structured as follows: Section II will cover the literature review on the features related to injury accidents and machine learning.

Approaches to predict accident severity. Section III will address the data collection procedures. Section IV will deal with the data preprocessing. The different machine learning algorithms and their performance metrics will be highlighted in the methodology in section V. The results obtained will be discussed in section VI and ultimately, the paper will be concluded in section VII together with future work

## II. LITERATURE REVIEW

Data mining is a diagnostic operation that integrates artificial intelligence, statistics, and machine learning which focuses on pattern detection, prediction, and forecasting [3]. Further, [4] added that machine learning is effective in road traffic injury predictions and can help to mitigate road accidents. The literature review highlights some contributing features identified and machine learning techniques used in this particular field.

### *A. Features contributing to injury accidents*

According to [5], the multiple factors affecting road crashes are human causes, weather circumstances, road designs, traffic characteristics and vehicle conditions. In the study of [6], it was concluded that the use of data mining can determine and forecast leading factors amidst human, vehicle and environment. The study of [7] revealed that junction type, road type, location, signposting, the hour of the day, license type, driver's age, day of the week and vehicle type are all significantly related to injury severity.

### *B. Machine learning techniques to model road accidents*

The utilization of machine learning classifiers provides alternatives to traditional data mining techniques for generating higher results and accuracy, as highlighted in [8]. The study of [9] compared machine learning algorithms to predict the severity of motorcycle crashes. J48 decision tree classifier, random forest and instance-based learning with parameter "k" (IBk) were used in modelling the severity outcome. These models were validated employing the

technique of 10-fold cross-validation. Comparisons were made with each other and with a statistical model namely the multinomial logit model
(MNLM). Their experimental results revealed that the prediction performances of the machine learning algorithms were better than MNLM. The random forest showed its superiority with the experimental data for its optimization and extrapolation capability. In [10], the authors employed J48, rule induction (PART), naïve Bayes and multilayer perceptron (MLP) which showed similar conclusions from the algorithms apart from the naïve Bayes classifier which exhibited less accuracy.

A comparative study in [11] was made with four models for crash severity prediction using multinomial logit (MNL), nearest neighbor classification (NNC), SVM and RF. The results demonstrated that NNC performed satisfactorily overall and in high severity crashes. NNC's best achievement was since this method does not require a distributional assumption of the data. The limitation in this study was the forced removal of part of the probably paramount predictors, like speed and impact type due to the presence of unavailable information in them. The study of [12] employed gradient boosted, decision tree (DT) and RF to identify hazard factors and injury severity prediction of drivers. The gradient boosted had the highest accuracy of 73.3%. Similarly, in [13], six algorithms namely LR, DT, SVM, neural network (NN), RF and extreme gradient boosting (XGBoost) were compared in predicting injury severity levels. The XGBoost outperformed the other algorithms with an accuracy of 74.4%, followed by the RF (73.8%).

In [14], to predict injury severity crashes, a comparison of the SVM with the ordered probit (OP) model was made. It was deduced that the SVM model had improved predicting power (48.8%) for severity predictions over the OP model (44.0%). As limitations, the authors suggested that apart from the basic radial basis function (RBF), different kernel functions can be experimented with to enhance the results of the models. Another study comparing SVM models with polynomial and Gaussian RBF kernels was carried out by [15] to enquire about driver's injury severity. It was found that the SVM model gave feasible prediction achievement and the polynomial kernel surpassed the Gaussian RBF kernel. The authors stated that SVM algorithms are a common non-parametric classification technique that has been extensively used in the transportation field, but are yet somewhat new in the road accident analysis field. The polynomial SVM classifier worked best on the majority of instances. It was also confirmed that transforming multi-categorical variables into numerical ones is an effective way to enhance the ability of the classification model. The limitation was however the small sample sizes for each type of injury analyzed.

## III. DATA COLLECTION

Road accidents data including information on accidents circumstances, vehicles and casualties collected by the local authority in Mauritius was used.

The dataset initially comprised 12,523 instances and 49 variables, including the target variable "Accident_Severity". The quality of the data was assessed by investigating if the type of the variables was correct and the presence of any missing data. The data preparation phase consisted of removing unwanted and conflicting variables, handling missing data, treatment of class variables, renaming of factors, binning, dimension reduction/feature selection, association test and handling of imbalance class.

### *A. Excluded variables*

Unwanted variables which don't add intrinsic value to the dataset and were not influential to the target variable were removed. Some variables were: "Accident key", "Vehicle reference number" among others. The response from the variables "Casualty injury" or "Driver injury" or both depending on whichever is higher, determine the overall accident severity type. These two conflicting variables were removed after the strength of their associations were tested using the chi-square test.

### *B. Handling missing data*

Associated entries containing missing values were not dropped, as this would have resulted in considerable loss of information. The missing values were replaced by the next highest code to be further labelled by "Unknown" for the categorical variables and only drivers above 15 years were considered.

### *C. Treatment of class variables and renaming of factors*

The variables were converted to their appropriate classes and renamed accordingly. The numerical variables were converted to an integer and the categorical variables to factors.

### *D. Binning*

The "Time" variable was converted into a factor category containing three levels: 'Night', 'Morning' and 'Afternoon'. The variables "Casualty age" and "Driver age" were binned in the range of five years.

### *E. Feature selection*

Feature selection is primarily based on the exclusion of non-informative or irrelevant predictors from the model and is also a dimensionality reduction technique. This technique helps in improving the performance of some predictive modelling and computational time in the case of large datasets [16]. After one hot encoding of the categorical variables to be used for the SVM, the most important features were selected according to their *k* highest score.

### *F. Handling of imbalance class*

As highlighted in [17], machine learning algorithms are influenced by a high imbalance class between dominant and classes of the minority, which may result in favour of the majority (negative) class during prediction. The authors further stated that ratios of 75:25 and 65:35 are classified as slight imbalance, whereas the ratio of 90:10 is regarded as

moderately imbalance. Fig. 1 shows that the dataset under study had two minority classes (fatal and serious) suggesting an imbalanced dataset.

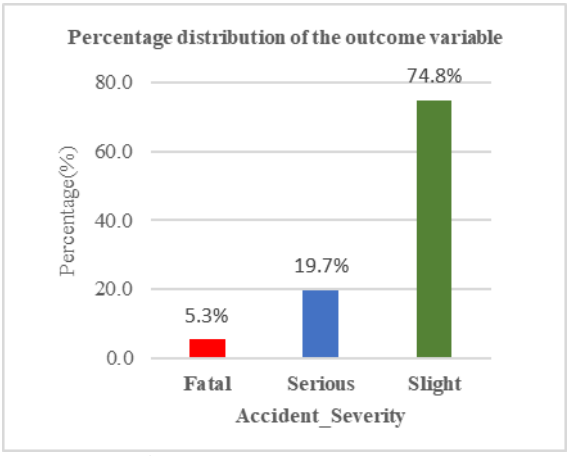



Fig. 1. Distribution of the target variable.

As such, these classes were grouped to form the positive class KSI (Killed and serious injury) which later for classification purposes in model prediction will be allocated a value of "1". The negative class or majority slight class will be assigned a value of "0". The result obtained after merging the two minority classes is shown in Fig. 2.

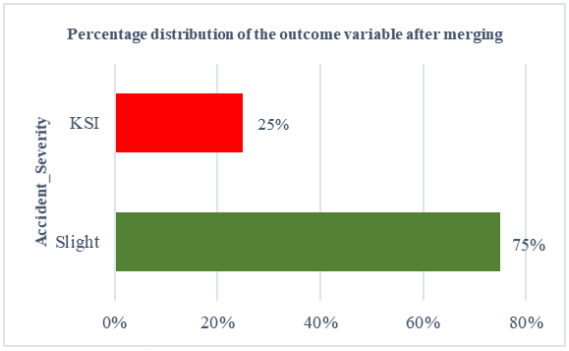



Fig. 2. Distribution of the merged classes.

A slightly imbalanced dataset was obtained with a ratio of 25:75, from which the classification models were built.

### G. *Final dataset*

We now have a clean dataset with 12,111 observations and 30 variables, including the outcome.

## IV. METHODOLOGY

In this work, the experimentation was based on five machine learning algorithms namely support vector machine, gradient boosting, logistic regression, random forest and naïve Bayes. Different split ratios (75:25, 80:20 and 90:10) were employed for the training and test set respectively. SVM requires numerical inputs. Hence, the categorical variables were one-hot encoded. The numerical variables were scaled using the min-max normalization. To assess the performance of each model, the confusion matrix was considered.

### A. *Logistic regression*

Logistic regression utilizes the power of regression to classify using maximum likelihood to fit a sigmoid curve on the target variable distribution. In this study, the logit is the natural logarithm of the odds or the likelihood ratio that the response is '1' (KSI) opposing to '0' (Slight). The probability '$p$' of a KSI accident is given by:

$$Y = logit = ln\ ln\left(\frac{p}{1-p}\right) = \beta X. \quad (1)$$

Where $'Y'$ is the dependent variable (Accident_Severity; $Y$=1, if every severity is KSI and $Y$=0 if the severity is Slight), '$\beta$' is a vector of parameters to be calculated and '$X$' is a vector of the predictors.

### B. *Naïve Bayes*

The naive Bayes classifier assumes that the existence (or lack) of one class feature (i.e., attribute) has no bearing on the presence (or absence) of any other feature. The model uses trained data to compute the probability of each class and the conditional probability of each class given each '$x$' value. It is effective for a large range of complicated problems [18]. The study of [19] concluded naïve Bayes to be more accurate than decision tree algorithm in predicting the severity of an accident.

### C. *Decision tree*

The decision tree can be employed to solve problems of regression and classification. ID3, classification and regression trees algorithm, J48, alternating decision tree (ADTree) form part of the decision tree algorithms [20]. It can be used to visually and explicitly represent decisions and decision-making based on the lowest Gini index and highest information gain.

### D. *Random Forest*

It is an ensemble of many decision trees and can be used in both classification and regression. RF is viewed as an improved method of the decision tree [21]. Being non-parametric, RF does not require any formal distributional assumption. It can handle many predictor variables and missing data [22].

### E. *Support vector machine*

This machine learning model was initially used in training data where it could be separated without errors and later was extended to non-separable training data [23]. It is another supervised algorithm used in the analysis of both classification and regression. In this study, we experimented with the linear SVM.

### F. *Gradient boosting*

The Gradient Boosting method works by constructing a series of successive decision trees, each of which seeks to

improve on the one before it by optimizing errors. It is a stage-wise additive model, in which each weak learner is added one at a time, previous weak trees are left unmodified and the model is trained by iteratively improving each tree [12].

### G. *Confusion matrix*

The classifying ability of the learned models was assessed as per the confusion matrix in Fig. 3.

| | | Actual values | |
|---|---|---|---|
| | | KSI | Slight |
| Predicted values | KSI | TP | FP |
| | Slight | FN | TN |

Fig. 3. (2x2) confusion matrix.

True Positive (TP) refers to the predicted positive outcome by the model that corresponds to the positive real value. True Negative (TN) is the predicted negative result that fits the actual value that was negative. False Positive (FP) is also termed as Type 1 error. It is the predicted positive value by the model that was incorrectly predicted as the real value was negative. False Negative (FN) also referred to as Type 2 error is the forecasted negative value by the model that was wrongly predicted as the real value was positive.

In the context of this study, to gauge the performance of the algorithms, the accuracy and area under the receiver operating characteristic (ROC) curve were considered. Accuracy is a measure of the actual positives plus the actual negatives that are correctly classified in the test data. The higher is the area under the curve, the better is the classifier.

## V. RESULTS

Table 1 provides a holistic view of the metrics used to gauge the performance of the different classification models with their numerical results.

Table 1. Accuracy of classifiers by different split ratios.

| **Classifier algorithm** | **Ratio of training set: test set** | | |
|---|---|---|---|
| | *75:25* | *80:20* | *90:10* |
| SVM (linear) | 76.8% | 77.0% | 77.8% |
| Gradient boosting | 82.9% | 83.0% | 83.2% |
| Logistic regression | 76.8% | 77.0% | 77.7% |
| Random forest | 78.7% | 78.5% | 77.4% |
| Naïve Bayes | 67.7% | 67.1% | 66.7% |

It is observed from Table 1 that the gradient boosting has the best performance irrespective of the ratio used when compared with the other classifiers. However, its performance was the best for the split ratio of 90:10 yielding an accuracy of 83.2%, which concur with the results of [12] and [13]. The naïve Bayes was the only algorithm with an accuracy of less than 75% and was therefore not considered for other metrics calculation. This result is in contrast with the findings of [19]. The ability of the other remaining classifiers was appraised in terms of the AUC as shown in Fig. 5 which further confirms the superiority of the gradient boosting with an AUC value of 0.839. Although the SVM and the LR had better accuracy than the RF, the latter had a better AUC value of 0.687 as compared to 0.652 for the other two classifiers.

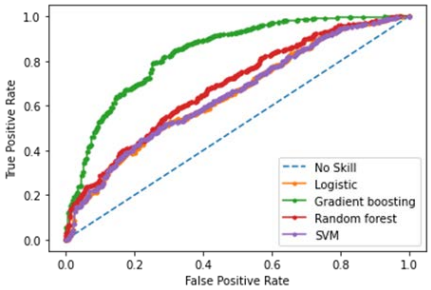



Fig. 5. Area under curve of classifiers.

## VI. CONCLUSION AND FUTURE WORK

The primary objective of this study was to apply several accident severity prediction models using machine learning. The dataset was preprocessed to deal with all possible issues that affect the performance of machine learning algorithms. The different algorithms were made to learn from the training set and their performance was assessed on the test set. The two minority classes were merged to deal with the imbalanced problem. The classifications were made according to a 2x2 table through a confusion matrix and a model's performance was established based on accuracy and the area under curve. The gradient boosting with an accuracy of 83.2% and an AUC of 83.9% showed better performance in terms of classifying the minority and majority classes. As an improvement to this work, it is proposed that the k-fold cross validation be carried out as a means to detect any over/underfitting. Other models like the artificial neural network (ANN), SVM with different kernels and other deep learning models may be implemented and compared.

## REFERENCES

[1] "Road traffic injuries," Who.int. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries, Accessed on: Jul 21, 2021.

[2] *Govmu.org*. [Online]. Available: http://www.govmu.org/French/News/Pages/Création-d'un-Observatoire-national-pour-la-sécurité-routière-à-Maurice.aspx, Accessed on: Jul 22, 2021.

[3] F. M. Nafie Ali and A. A. Mohamed Hamed, "Usage Apriori and clustering algorithms in WEKA tools to mining dataset of traffic accidents," *J. Inf. Telecommun.*, vol. 2, no. 3, pp. 231–245, 2018.

[4] D. Donchenko, N. Sadovnikova, and D. Parygin, "Prediction of road accidents' severity on Russian roads using machine learning techniques," in *Lecture Notes in Mechanical Engineering*, Cham: Springer International Publishing, 2020, pp. 1493–1501.

[5] M. Vilaça, E. Macedo, and M. C. Coelho, "A rare event modelling approach to assess injury severity risk of vulnerable road users," *Safety (Basel)*, vol. 5, no. 2, p. 29, 2019.

[6] S. Ramya, S. K. Reshma, V. D. Manogna, Y. S. Saroja and G. S. Gandhi, "Accident Severity Prediction Using Data Mining Methods," *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, vol. 5, no. 2, 2019.

[7] L. Eboli, C. Forciniti, and G. Mazzulla, "Factors influencing accident severity: an analysis by road accident type," *Transp. res. procedia*, vol. 47, pp. 449–456, 2020.

[8] A. Venkat, M. Gokulnath, V. K. P. Guru, S. T. Irish and D. Ranjani, "Machine learning based analysis for road accident prediction," *International Journal of Emerging Technology and Innovative Engineering*, vol. 6, no. 2, pp. 31-37, 2020.

[9] L. Wahab and H. Jiang, "A comparative study on machine learning based algorithms for prediction of motorcycle crash severity," *PLoS One*, vol. 14, no. 4, p. e0214966, 2019.

[10] M. Taamneh, S. Alkheder, and S. Taamneh, "Data-mining techniques for traffic accident modeling and prediction in the United Arab Emirates," *J. Transport. Safety Security*, vol. 9, no. 2, pp. 146–166, 2017.

[11] A. Iranitalab and A. Khattak, "Comparison of four statistical and machine learning methods for crash severity prediction," *Accid. Anal. Prev.*, vol. 108, pp. 27–36, 2017.

[12] S. Elyassami, Y. Hamid, and T. Habuza, "Road crashes analysis and prediction using gradient boosted and random forest trees," in *2020 6th IEEE Congress on Information Science and Technology (CiSt)*, 2020, pp. 520–525.

[13] M. K. Nour, A. Naseer, B. Alkazemi, and M. Abid, "Road traffic accidents injury data analytics," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 12, 2020.

[14] Z. Li, P. Liu, W. Wang, and C. Xu, "Using support vector machine models for crash injury severity analysis," *Accid. Anal. Prev.*, vol. 45, pp. 478–486, 2012.

[15] C. Chen, G. Zhang, Z. Qian, R. A. Tarefder, and Z. Tian, "Investigating driver injury severity patterns in rollover crashes using support vector machine models," *Accid. Anal. Prev.*, vol. 90, pp. 128–139, 2016.

[16] D. Bhalla, "Feature selection: Select important variables with Boruta package," *Listendata.com*. [Online]. Available: https://www.listendata.com/2017/05/feature-selection-boruta-package.html, Accessed on: Aug 13, 2021.

[17] T. Hasanin, T. M. Khoshgoftaar, J. L. Leevy, and N. Seliya, "Examining characteristics of predictive models with imbalanced big data," *J. Big Data*, vol. 6, no. 1, 2019.

[18] G. Shobha and S. Rangaswamy, "Machine Learning," in *Handbook of Statistics*, vol. 38, V. N. Gudivada and C. R. Rao, Eds. Elsevier, 2018, pp. 197–228.

[19] V. M. Ramachandiran, P. N. Kailash Babu, and R. Manikandan, "Prediction of Road Accidents Severity using various algorithms," Acadpubl.eu. [Online]. Available: https://acadpubl.eu/hub/2018-119-12/articles/6/1546.pdf, Accessed on: Aug 29, 2021.

[20] A. K. Sharma and S. Sahni, "A Comparative Study of Classification Algorithms for Spam Email Data Analysis," *International Journal on Computer Science and Engineering*, vol. 3, no. 5, pp. 1890-1895, May 2011.

[21] J. Wu, Z. Gao, and C. Hu, "An empirical study on several classification algorithms and their improvements," in *Advances in Computation and Intelligence*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 276–286.

[22] S. Richmond, "Algorithms exposed: Random forest," Org.au. https://bccvl.org.au/algorithms-exposed-random-forest/, Accessed on: Jul 30, 2021.

[23] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index No: TL-28-54*

# Open Educational Resources (OER) usage in learning: Perspectives of undergraduate students, University of Vocational Technology, Sri Lanka

M.Thenabadu (1st Author)
Department of Agriculture and Food Technology
*University of Vocational Technology*
Ratmalana,Sri Lanka
mthenabadu@uovt.ac.lk

H.A. Seneviratne (2nd Author)
Department of Multimedia and Web Technology
*University of Vocational Technology*,
Ratmalana Sri Lanka
helawikum@gmail.com

***Abstract*- Open Educational Resources has gained momentum during the past decade as a way of sharing educational resources with the learning community. This initiative has been very influential in changing the learning culture of higher education worldwide. However, there is a dearth of literature on OER usage, particularly among higher education students in developing countries, despite the fact that, higher education students in developing countries are portrayed as the primary beneficiaries of such initiatives. The aim of this study was to examine the awareness and perceived barriers of using OERs at the University of Vocational Technology. Data were gathered using a survey questionnaire from undergraduate students representing ICT and Food Technology degree programs (n=150). The findings revealed that a significant proportion of students were having less awareness of OER. It was observed that students face many barriers in using OER such as search techniques, content and environmental issues. According to the study, university and faculty members should take the lead in practice and dissemination of the concept of OER among students in order to encourage adoption of this valuable initiative.**

***Keywords — Open Educational Resources (OER), student perception, barriers in OER***

## I. INTRODUCTION

The emergence of Open Educational Resources (OER) is one of the most significant educational innovations in recent developments of open education[9][10]. OER is a relatively new concept in the field of Open Education, defined by UNESCO as "teaching, learning, and research materials in any medium, digital or otherwise, that reside in the public domain or have been released under an open license that permits no-cost access, use, adaptation, and redistribution by others with no or limited restrictions."[11] The recent definition proposed by the Organization for Economic Cooperation and Development (OECD) is "digitized materials offered freely and openly for educators, students, and self-learners to use and reuse for teaching, learning, and research." [10]

OER is originated from the common and widespread practice of creating and sharing learning resources where MIT's Open Courseware (OCW)project is frequently cited as the first example of OER. Since then, OCW has grown rapidly in both developed and developing countries [1][3] In 2017, UNESCO hosted the second World OER Congress, which resulted in OER action plan [12]. This policy recommended OER focus to its member countries as an enabler of open education towards innovative education and pedagogical strategy. The goal is to increase stakeholder capacity to create, use, adapt, and redistribute open educational resources (OER); develop supportive OER policy; encourage the development of inclusive and equitable quality OER; foster the development of OER sustainability models; and facilitate international cooperation on OER.

OER is still a relatively new concept in the Sri Lankan education system. In 2017 on par with the introduction of sustainability bill in line with United Nations' Sustainable Development Goals, adoption of OER for teaching and learning is promoted as sustainable education practice. Open University of Sri Lanka had been pioneering in introducing OER concept to the country , where several studies have been conducted to promote OER adoption by practitioners[5] Aaccording to Karunanayaka, Fernando & De Silva [6] practitioners' awareness of OER was extremely low, but their readiness and motivation to adopt the concept was seen quite high.

Previous research has emphasized the importance of establishing close collaboration between researchers and practitioners in awareness-raising and capacity-building initiatives, as well as the importance of designing appropriate

experiences in a systematic manner [6] Existing literature, however, fails to provide sufficient evidence on the utilization of OCR among student groups in Sri Lanka. Therefore, aim of this research is to investigate the awareness and perception of using OER among undergraduate students of Sri Lanka.

## II. OBJECTIVES

The objective of this research is to investigate the awareness and perception of using OER among undergraduate students of the University of Vocational Technology, Sri Lanka.

## III. METHODOLOGY

The target population was undergraduate students in University of Vocational Technology, Sri Lanka. Student groups of two degree programs of the University of Vocational Technology were selected as the sample. A total of 150 students were selected from two degree programs as per the convenient sampling method. The survey was done using online methods by posting questions using Google web form.

To identify undergraduate students' awareness and perceptions on OER survey instrument was developed based on the OCED's definition on OER (2007) and adapted version of survey items developed by Hu, E. et al [4]

The survey instrument included eight items. Demographic data were gathered in the first part and OER usage was investigated using five items indicating, OER usage experience, Where OER information received, purpose of OER usage , Frequency of using OER, Frequently used OER content, and Barriers /Issues in using OER is queried using open-ended question and focus group discussion .The data obtained were analyzed using descriptive statistics.

## IV. RESULTS AND DISCUSSION

Among the 150 students who participated for the survey 82( 55%) were male students and 68 (45%) were female, 80 (53%) were from ICT degree and 70 ( 47%) were from Food Technology degree,26 (17%) were first years students, 45 (30%) were second year students and 79 (53%) were from third year students. The table 01 shows the details of participants' distribution in terms of demographical variables of gender, Degree and academic year.

Table 1: Respondent demographics ( n=150)

| Personal characteristics | Sample | Percent (%) |
|---|---|---|
| **Gender** | | |
| Male | 82 | 55% |
| Female | 45 | 30% |
| **Degree Programme** | | |
| B.Tech ICT | 80 | 53% |
| B.Tech FT | 70 | 47% |
| **Academic Year** | | |
| Year 01 | 26 | 17% |
| Year 02 | 45 | 30% |
| Year 03 | 79 | 53% |

Awareness and OER usage patterns of undergraduates are summarized in Table 2. As per the survey results, only 29% of respondents had experience in using OER.and the majority of respondents had no experience in OER usage. (71%) Among the 29% of the students who have OER experience, indicated they got to know about OER via other students ( 50%) ,by search engines (27%) and though Faculty members (23%).With regard to the purpose of using OER 68% students indicated that they used OER for their personal needs , 11% used to get content of other disciplines ,9% used OER to view and listen to other scholars' work.

In terms of frequency of OER usage , 59% indicated that the frequency vary with the requirement. 27% used OER weekly . only small percentage of students (5%) used OER daily. With respect to mostly used OER content, it was noted that video was the most popular with 57%, while 23% listed text discussions 9% and other content 11%.

Table 2 : OER usage of undergraduate students

| OER usage status | Number | Percent(%) |
|---|---|---|
| **OER usage experience** | | |
| Have OER usage experience | 44 | 29% |
| Have no OER usage experience | 106 | 71% |
| **Channel to get to know OER** | | |
| Search engine | 12 | 27% |
| Other students | 22 | 50% |
| Faculty members | 10 | 23% |
| **Purpose of using OER** | | |
| To assist personal learning | 30 | 68% |
| To get to know content in other disciplines | 5 | 11% |
| To view other scholars' work | 4 | 9% |
| **Frequency of using OER** | | |
| Monthly | 4 | 9% |
| weekday | 12 | 27% |
| daily | 2 | 5% |
| As per the requirement | 26 | 59% |
| **Most frequently used OER content** | | |
| Video | 25 | 57% |
| Text | 10 | 23% |
| Online messaging and discussions | 4 | 9% |
| Other contents | 5 | 11% |

**Perceived barriers to OER usage:**

Students perceived several barriers to using OERs, including a lack of understanding of what OERs are and their usability. Many students have stated that they have used some form of OER without fully comprehending its benefits. Faculty members have not encouraged students to use OER for learning and information on OER are not readily available to the undergraduate students

Some students were resistant to learning through online means. Those who are aware of OERs claim that there are few repositories that provide high-quality learning materials and are difficult to find such. There are only a few subjects and disciplines covered, and not sufficient learning material present. Also there was no suitable platform to communicate and interact on the websites of OER.

Based on the analysis of students' perceived barriers to using OERs, it is strongly recommended that university should promote the OER initiative to raise student awareness. The advisory role of faculty members, in particular, is highly recommended. Many students had no idea how to look for OER in their subject discipline. To narrow this gap, providing an OER database to undergraduates is a good option.

## V. CONCLUSION

According to the findings of this study, undergraduate students' awareness and use of open educational resources (OER) at University of Vocational Technology with reference to the two degree programs studied are at a bare minimum. Faculty members' contribution towards making students aware of OER and have taught students how to use them, were also less according to the findings. University should implement special initiatives to introduce OER to their students in order to broaden the influence of OER and effectively advance quality education.

Furthermore, the findings of this study should inform university authorities about potential barriers to using OER and help them develop strategies to overcome these barriers in the context of the institutional environment. As suggested in focus group discussions, awareness sessions, faculty member involvement, creating OER database may help in minimizing potential barriers of using OER among students.

Because the scope was limited to an exploratory study of two degree programs at one university, the findings cannot be generalized. A discipline-specific instrument, which was not considered in this study, may also reflect more authentically and effectively on students' actual perceptions of OER usage.

## REFERENCES

[1] D. E. Atkins, J. S. Brown, and A. L. Hammond, A review of the open educational resources (OER) movement: Achievements, challenges, and new opportunities, vol. 164. Creative common Mountain View, 2007.

[2] P. Arinto, C. Hodgkinson-Williams, and H. Trotter, "OER and OEP in the Global South: Implications and recommendations for social inclusion." African Minds, International Development Research Centre & Research on Open …, 2017.

[3] C. Hockings, P. Brett, and M. Terentjevs, "Making a difference—inclusive learning and teaching in higher education through open educational resources," Distance education, vol. 33, no. 2, pp. 237–252, 2012 "Full article: Open Educational Resources: reviewing initiatives and issues." https://www.tandfonline.com/doi/full/10.1080/02680510802625443 (accessed Dec. 07, 2021).

[4] E. Hu, Y. Li, J. Li, and W.-H. Huang, "Open educational resources (OER) usage and barriers: a study from Zhejiang University, China," *Educational Technology Research and Development*, vol. 63, no. 6, pp. 957–974, 2015

[5] S. P. Karunanayaka and S. Naidu, "Capacity building of academic staff in the integration of ICT and OER in teacher education programs at the Open University of Sri Lanka," 2013.

[6] S. P. Karunanayaka and S. Naidu, "Impact of integrating OER in teacher education at the Open University of Sri Lanka," International Development Research Centre (IDRC), 2017.

[7] J. Knox, "Five critiques of the open educational resources movement," *Teaching in Higher Education*, vol. 18, no. 8, pp. 821–832, 2013.

[8] D. A. Nyberg, The Philosophy of Open Education (International Library of the Philosophy of Education Volume 15). Routledge, 2010.

[9] C. M. Stracke, "Quality frameworks and learning design for open education," *The International Review of Research in Open and Distributed Learning*, vol. 20, no. 2, 2019.

[10] "Giving Knowledge for Free: The Emergence of Open Educational Resources - OECD." https://www.oecd.org/education/ceri/givingknowledgeforfreetheemergenceofopeneducationalresources.htm (accessed Dec. 07, 2021).

[11] https://plus.google.com/+UNESCO, "The Paris OER Declaration 2012," UNESCO, Jul. 20, 2017. https://en.unesco.org/oer/paris-declaration (accessed Dec. 07, 2021).

[12] "ljubljana_oer_action_plan_2017.pdf." Accessed: Dec. 07, 2021. [Online]. Available: https://en.unesco.org/sites/default/files/ljubljana_oer_action_plan_2017.pdf

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index No: TL-29-55*

# M-LMS Adoption Intention; Empirical Evidence from Postgraduates in Developing Context

Isuru Chandradasa (1st Author)
Department of Human Resource Management
Faculty of Management and Finance
*University of Ruhuna*
Matara, Sri Lanka
isuru@mgt.ruh.ac.lk

***Abstract*-Mobile learning is a widely used teaching and learning method and a Mobile learning management system is a part of the mobile learning process. During the pandemic period, the demand for mobile learning raised rapidly and is continuing to rise. After the pandemic if academic institutions hope to continue this trend, it is important to know about students' intention to use this Mobile learning management system in the future. Without knowing this continuing this application in the future is a risk for the organization as well as a negative factor for users' satisfaction. The evidence from past literature in this regard is lacking. Therefore, the objective of this study is to examine the continued intention of Sri Lankan postgraduate's mobile learning management system use in the future. The variables that might influence their continued intention of using this Mobile learning management system are identified from previous literature and based on that conceptual framework and hypothesis of the study developed. This study used quantitative techniques and data collected from 119 postgraduate students of one of the reputed state university in Sri Lanka by employing a convenience sampling technique. The data was analyzed using SPSS 25 statistical software package. The results of the study showed that direct utilization, ease of use, mobility value, academic relevance, and university management support will have a significant impact on the post-graduates' intent on the mobile learning management system in their future academic pursuits. This study provides theoretical and practical insights for researchers and practitioners about M-LMS usage and future implications. The data included postgraduates only from one state university which is the major limitation of this study.**

***Keywords - Continuous intention of adopt M-LMS, M-LMS, Mobile learning (M-Learning)***

## I. INTRODUCTION

The rapid development of technologies influences every field of business and transforms into many convenient forms that ensure higher satisfaction of its users (Han & Shin, 2016). Mobile learning is one of the dynamically growing trends in the educational field and the expansion of these digital technologies enables continuous development (Han & Shin, 2016). In the past two years from 2020 to 2021, the world suffering from the COVID-19 virus, and alternative mechanisms to physical teaching to continue academic activities without interruption from the pandemic are used.

Developing countries are not ready to fully accept that challenge compared with developed countries that are abundant in necessary infrastructure and other necessities to face this new circumstance, (Gao, Krogstie, & Siau, 2014). In Sri due to the Covid situation all physical academic activities were temporarily stopped, and educational institutions started teaching using alternative mechanisms. Mobile learning was one of them that is popular in developing contexts. Sharma and Kitchens (2004) defined mobile learning as a new style of learning enabled by mobile devices, which include ubiquitous communication technologies and sophisticated user interfaces. With the advent of mobile learning, students may now experience a teaching atmosphere on their mobile devices (Brantes Ferreira, Zanela Klein, Freitas, & Schlemmer, 2013). Many innovative mobile services entered the university educational systems due to the pandemic situation (Qashou, 2021). Moreover, most of the past studies emphasized that using mobile learning was related to higher exam scores (Han and Shin, 2016).

In the advancement of mobile learning, numerous applications including the existing learning management system (LMS) needed to improve as compatible to bring a better user experience (Nguyen, Barton, & Nguyen, 2015). This advanced form of the learning management system which provides access via mobile devices is known as a mobile learning management system (M-LMS) (Brantes Ferreira, Zanela Klein, Freitas, & Schlemmer, 2013). M-LMS is a type of mobile learning platform that allows academics and students to access lessons on their mobile devices at any location and at any time (Nguyen et al., 2015).

Currently, educational institutions are gradually reopening for physical education but continue the use of mobile learning. Thus, this trend will continue. Most of the businesses promote work from home concept due to the advantages that are

incorporated, and, in education, it is similar. (Brantes et al., 2013). If this tendency persists in the future, it is important to be aware of the continuing intentions of the practices that use it.

Literature regarding the use M-LMS in the education sector in the past is limited. The availability of an M-LMS does not guarantee that students will utilize it in their daily academic activities continuously. Students' impressions of the system may be varying by different factors (Brantes et al., 2013). Some students were unaware of the system's potential utility (Nguyen, Barton, & Nguyen, 2015).
Further, due to the lack of evidence on university students continued use of M-LMS this study tries to fill this empirical knowledge gap.

The relevance of this m-LMS is especially important for online university students who are relatively old, part-time, and off-campus students (Qashou, 2021). According to Dahlstrom and Bichsel's (2014) research, older students are more likely to indicate that they learned best when a course was online, but younger students were more likely to report that they learned best when a course was completely online. The object of this study is to examine the factors that influence postgraduate students for continued use of M-LMS.

## II. LITERATURE REVIEW

### A. *M -LMS*

Users can access learning materials via mobile learning management systems on portable wireless devices such as smartphones and tablets (Qashou, 2021).

### B. *The continuous intention of using M-LMS*

Technology usage refers to either the amount of effort exerted in interacting with a particular technological system (Fitzgerald, 1993). From this perspective, continuous use of technology refers to a person's future desire, expectation, or goal to employ presently in use technology or system (Brantes et al., 2013). M-LMS is a technology-enabled system, and the current study focuses on the continuous intention of use- M-LMS for their future learning activities. According to Ajzen and Fishbein, 1980 continuous intention is a measure of a person's propensity to keep on using technology or system. Several theoretical models have emphasized the significance of behavioral intention as the most important predictor of human behavior in the continuous use of technology (Lee & Rao, 2009). In the context of the present study, the intention was to assess if the postgraduates, who are using M-LMS for their academic activities, would be willing to continue in the future.

### C. *Theories of study variables*

Many prior studies used or expanded the Technology acceptance model (TAM) to investigate the adoption intention of mobile learning. Akman & Turhan (2017) used an extended TAM to investigate students' acceptance of online learning systems in higher education. Recent studies discovered that perceived usefulness and perceived ease of use were significant determinants of users' attitudes towards the use of online educational systems (Mehta, Morris, Swinnerton, & Homer, 2019). We construct an expanded TAM model to examine the adoption of M-LMS based on the facts supporting the TAM and its success in integrating other variables to better understand aspects linked to continuous intention to use M-LMS. In addition to TAM's core constructs, three major predictors, Perceived Mobility Value (Huang, 2007), Academic Relevance (Venkatesh & Davis, 2000), and University Management Support (McGill, Klobas & Renzi, 2014) are extracted from prior studies to explore students' continuous intention to use M-LMS within a university education setting. Based on that figure 1 conceptual framework was developed.

### D. *Conceptualization*

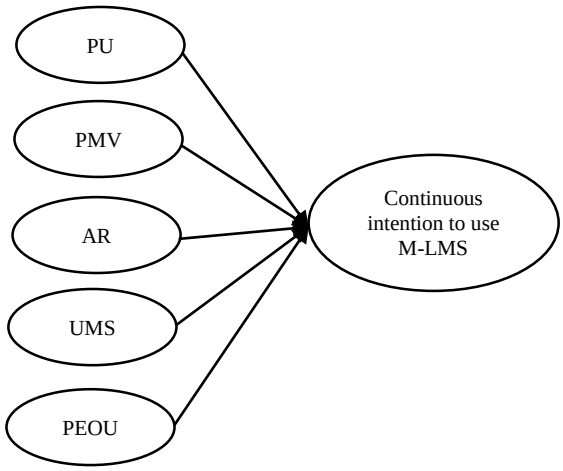



Fig. 1. Conceptual framework Source: Author constructed, 2021

### E. *Hypothesis development*

TAM is a user acceptance model that focuses on a system's perceived usefulness and perceived ease of use. The degree to which a person feels that using a given system will improve his or her job performance is referred to as perceived usefulness (Davis, 1989) and perceived ease of use refers to the degree to which a person feels that utilizing a certain system would be devoid of effort (Davis, 1989). These two variables had been tested in several empirical studies and found a positive association between these two variables with technology adoption intention (Salloum, Alhamad, Al-Emran, Monem & Shaalan, 2019). Further, it has proven as largely reliable in predicting user acceptance of new information technologies Pavlou, (2003); Gefen, (2003). Based on those following hypotheses were derived.

*H1: Perceive usefulness has a positive influence on continuous intention to use M-LMS.*

*H2: Perceive ease of use has a positive influence on continuous intention to use M-LMS.*

The extent to which users are aware of the flexibility of mobile services and systems is characterized as the perceived mobility (Rahmi, Birgoren,&Aktepe, 2018). Previous studies identified the positive effect of perceived mobility value on technology adoption and use for mobile learning (Huang et al., 2007). Based on that following hypothesis was derived.

*H3: Perceived mobility value has a positive influence on continuous intention to use M-LMS.*

Academic relevance refers to the general significance of M-LMS in university education. (Venter, van Rensburg, & Davis, 2012). Venkatesh & Davis (2000) discovered that general relevance had a favorable effect on technology adoption and use. As a result, this study contends that academic relevance in general influences the intention of continuous use of M-LMS. As a result, the following hypothesis was derived.

*H4: Academic relevance has a positive influence on continuous intention to use M-LMS.*

University management support refers to the service provided by the administration and authorities of the university for the successful implementation of technology for students. Iqbal & Bhatti (2017) indicated the significance of university management support for mobile learning efforts in earlier studies (Saroia & Gao, 2019). Moreover, Venkatesh (2000) discovered that university management support had a significant impact on technology adoption and use. According to Barker, Krull, and Mallinson (2005), university support workers played a significant part in the day-to-day support and maintenance of their learning institution's mobile learning infrastructure. Based on that following hypothesis was derived.

*H5: University management support has a positive influence on continuous intention to use M-LMS.*

## III. METHODOLOGY

A descriptive research design was employed that allows assessing the associations between the variables described in the model. After reviewing the literature, five independent variables were identified and a conceptual framework for the study was developed using the five hypotheses.

This study focuses on investigating the intention to continuous use of M-LMS in the future of the Sri Lankan postgraduates. Hence, the unit of analysis was the individual. The theoretical population of this study is postgraduate students at state universities in Sri Lanka. Due to the practical limitation and complexities, postgraduate students who are reading for Master of Business Administration(MBA) in one reputed state university in Colombo district of Sri Lanka was selected as the study population. Designed questionnaires were distributed among 150 postgraduates who are reading for their MBA. By employing a convince sampling technique 119 responses were gathered. The constructs of the research model were measured using previously validated instruments. The all the constructs used a five-point Likert scale where respondents marked their agreement scaling from strongly disagree (1) to strongly agree (5). Gathered data was analyzed by using SPSS 25 version of the software package.

## IV. DATA ANALYSIS

First, the normality was checked and the results indicated that the data set is normally distributed. Frequency analysis was used to describe sample composition. Consequently, the internal consistency of the constructs was tested using Cronbach alpha and all constructs met the threshold value of 0.6 or above. Further, correlations among variables were tested using Pearson correlation and found no multi collinearity issues among the predictors as all the correlations among independent variables were reported below 0.44. The hypothesis was tested by using regression analysis results.

### *A. Sample Composition*

Table 1. Sample composition

| Variable | Category | Frequency | Percentage (%) |
|---|---|---|---|
| Age group | 20-25 | 17 | 14 |
| | 26-30 | 66 | 56 |
| | 31-35 | 32 | 27 |
| | More than 36 | 4 | 3 |
| Gender | Male | 56 | 47 |
| | Female | 63 | 53 |
| The period that uses M-LMS | 1-3 months | 42 | 35 |
| | 4 months to 6 months | 18 | 15 |
| | 1 year | 51 | 43 |
| | More than 1 year | 8 | 7 |

Source: Survey Data, 2021

### *B. Hypothesis testing*

The Adjusted R Square value amounts to 0.721. Thus, the regression model explains 72% of the variance in the intention to continuously use M-LMS with the five independent variables specified in the research model and the ANOVA test confirmed that the regression model is statistically significant (F = 76.689, P= 0.000)

Based on the observed results in table 2, indicated that perceived usefulness, perceived ease of use, perceived mobility value, academic relevance, and university management support have a positive influence on postgraduate students' continued intention to use M-LMS in the future. In sum, this study confirms the results of previous studies (Saroia & Gao, (2019); Ooi, Hew & Lee, (2018). Supporting H1, and H2 Perceived usefulness (β=0.444, sig =.001). and Perceived

ease of use (β=0.388, sig =.002) had a significant positive influence on postgraduate's continuo intention to use M-LMS.

Table 2. Regression results

| | **Unstandardized Coefficients** | | | |
|---|---|---|---|---|
| | **B** | **Std. Error** | **t** | **Sig.** |
| Perceived usefulness | .444 | .091 | 4.926 | .001 |
| Perceived mobility value | .904 | .168 | 5.368 | .000 |
| Academic relevance | .199 | .052 | 3.692 | .004 |
| University Management support | .807 | .168 | 5.368 | .000 |
| Perceived ease of use | .388 | .137 | 3.068 | .002 |
| Adjusted R Square | .721 | | | |
| ANOVA | F = 76.689, (P= 0.000) | | | |

Source: Survey Data, 2021

Moreover, this study found empirical support for the H3, which claims that perceived mobility value (b =0.904 p = 0.000) and academic relevance (b =0.199 p = 0.004) are having a significant positive influence on determining postgraduates continued intention to use M-LMS. The study also found that university management support (β=0.807, sig =.000) positively influenced on postgraduate's continuo intention to use M-LMS.

## V. RESULTS AND DISCUSSION

The observed results of the study indicate that the five variables selected has a significant positive influence on continuous intention of use M-LMS. Mainly postgraduates feel that using M-LMS use improve his or her academic performance and if they feel that utilizing M-LMS would be devoid of effort they intend to use M-LMS in the future. Moreover, when M-LMS consisted in flexibility benefit and characteristics postgraduates' future intention will positively derive that for continued use of M-LMS in future. Further significance of using M-LMS will increase comparatively in academic activities it also becomes a reason to influence of Use M-LMS in continuously. Furthermore, study results reveal that if the administration and authorities of the university have positive support for their postgraduates regarding M-LMS it will positively influence the postgraduate's continuous intention of using M-LMS in the future.

From a theoretical standpoint, this research is regarded as valuable for academic research since it expands and improves knowledge of mobile learning uptake. This study contributed to the existing literature on the adoption of M-LMS. This study illustrates the robustness and explanatory power of the proposed TAM framework in predicting users' intention to adopt an M-LMS in a higher education setting.

The findings enable the understanding of how university postgraduates' students will engage with and realize the usage of an M-LMS. They provide insight into creating and deploying of mobile technologies for educational purposes. From the postgraduates' perspective perceived ease of use, perceived usefulness, academic relevance, perceived mobility value, and university management support in M-LMS are the crucial aspects to ensuring continuous intention of use M-LMS in the future. Therefore, university administration authorities must consider these aspects because these affect the continued use of the system. Further this study provides insights to developers and service providers of mobile learning apps such as M-LMS to identify the important characteristics that would play a significant role in postgraduate users of M-LMS.

We identified the following study limitations while proposing future research objectives in this work. Due to time and budgetary restrictions, the sample was limited to one state university respondent. Bigger sample size would enhance statistical power and provide more robust results (Hair et al., 2010). Future research with a larger sample size covering postgraduates as well as undergraduates at other institutions is also recommended. This study applied a cross-sectional design, with data collected at a single moment in time. Students' continuous intention is considered a psychological concept that requires longitudinal empirical investigations to obtain a thorough understanding. Future studies using a longitudinal study methodology would substantially contribute to the literature. Compared to gathering responses to a questionnaire alternative technique, such as interviews, might allow for a more in-depth study of intention and its causes The study used only four factors based on TAM and empirical findings. Additional variables specified in other theories might have an impact on the continued intention of using M-LMS.

## REFERENCES

[1] Adel Ali, R., & Rafie Mohd Arshad, M. (2018). Empirical analysis on factors impacting on intention to use m-learning in basic education in Egypt. *International Review of Research in Open and Distributed Learning*, *19*(2).

[2] Akman, I., & Turhan, C. (2017). User acceptance of social learning systems in higher education: an application of the extended Technology Acceptance Model. *Innovations in Education and Teaching International*, *54*(3), 229-237.

[3] Alshahrani, H. A., & Walker, D. A. (2017). Validity, reliability, predictors, moderation: The UTAUT model revisited. *General Linear Model Journal*, *43*(2), 23-34.

[4] Alshurideh, M., Al Kurdi, B., Salloum, S. A., Arpaci, I., & Al-Emran, M. (2020). Predicting the actual use of m-learning systems: a comparative approach using PLS-SEM and machine learning algorithms. *Interactive Learning Environments*, 1-15.

[5] Barker, A., Krull, G., & Mallinson, B. (2005, October). A proposed theoretical model form-learning adoption in developing countries. In *Proceedings of mLearn* (Vol. 2005, p. 4th).
[6] Berge, Z. L., & Muilenburg, L. (Eds.). (2013). *Handbook of mobile learning*. Routledge.
[7] Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 319-340.
[8] El-Hussein, M. O. M., & Cronje, J. C. (2010). Defining mobile learning in the higher education landscape. *Journal of Educational Technology & Society*, *13*(3), 12-21.
[9] Essel, H. B., Nunoo, F. K. N., Tachie-Menson, A., & Amankwa, J. O. (2018). Higher Education Students' Ownership and Usage of Smart Phones and Tablets: The Case of Kwame Nkrumah University of Science and Technology (KNUST). *International Journal of Educational Technology*, *5*(1), 20-28.
[10] Ferreira, J. B., Klein, A. Z., Freitas, A., & Schlemmer, E. (2013). Mobile learning: definition, uses and challenges. In *Increasing student engagement and retention using mobile applications: Smartphones, Skype and texting technologies*. Emerald Group Publishing Limited.
[11] Fishbein, M., Jaccard, J., Davidson, A. R., Ajzen, I., & Loken, B. (1980). Predicting and understanding family planning behaviors. In *Understanding attitudes and predicting social behavior*. Prentice Hall.
[12] Fitzgerald, L. F., & Cortina, L. M. (2018). Sexual harassment in work organizations: A view from the 21st century.
[13] Gao, S., Krogstie, J., & Siau, K. (2014). Adoption of mobile information services: An empirical study. *Mobile Information Systems*, *10*(2), 147-171.
[14] Gefen, D., Karahanna, E., & Straub, D. W. (2003). Trust and TAM in online shopping: An integrated model. *MIS quarterly*, 51-90.
[15] Gondim, S. M. G., Techio, E. M., Paranhos, J., Moreira, T., Brantes, C., Sobrinho, J. B., & Santana, V. (2013). Imigração e trabalho: um estudo sobre identidade social, emoções e discriminação contra estrangeiros. *Revista Psicologia em Pesquisa*, *7*(2).
[16] Han, I., & Shin, W. S. (2016). The use of a mobile learning management system and academic achievement of online students. *Computers & Education*, *102*, 79-89.
[17] Herrador-Alcaide, T. C., Hernández-Solís, M., & Hontoria, J. F. (2020). Online learning tools in the era of m-learning: Utility and attitudes in accounting college students. *Sustainability*, *12*(12), 5171.
[18] Huang, S., & Czech, M. P. (2007). The GLUT4 glucose transporter. *Cell metabolism*, *5*(4), 237-252.
[19] Jaldemark, J., Hrastinski, S., Olofsson, A. D., & Öberg, L. M. (2018). Editorial introduction: Collaborative learning enhanced by mobile technologies. *British Journal of Educational Technology*, *49*(2), 201-206.
[20] Kukulska-Hulme, A. (2012). How should the higher education workforce adapt to advancements in technology for teaching and learning?. *The Internet and Higher Education*, *15*(4), 247-254.
[21] Kurdi, B., Alshurideh, M., & Alnaser, A. (2020). The impact of employee satisfaction on customer satisfaction: Theoretical and empirical underpinning. *Management Science Letters*, *10*(15), 3561-3570.
[22] Liao, C. H., Tsou, C. W., & Huang, M. F. (2007). Factors influencing the usage of 3G mobile services in Taiwan. *Online Information Review*.
[23] McGill, T. J., Klobas, J. E., & Renzi, S. (2014). Critical success factors for the continuation of e-learning initiatives. *The Internet and Higher Education*, *22*, 24-36.
[24] Mehta, A., Morris, N. P., Swinnerton, B., & Homer, M. (2019). The influence of values on E-learning adoption. *Computers & Education*, *141*, 103617.
[25] Nam, G. H., Gim, J. A., Mishra, A., Ahn, K., Kim, S., Kim, D. H., ... & Kim, H. S. (2018). Identification of transposable elements fused in the exonic region of the olive flounder genome. *Genes & genomics*, *40*(7), 707-713.
[26] Nguyen, L., Barton, S. M., & Nguyen, L. T. (2015). iP ads in higher education—Hype and hope. *British Journal of Educational Technology*, *46*(1), 190-203.
[27] Pavlou, P. A. (2003). Consumer acceptance of electronic commerce: Integrating trust and risk with the technology acceptance model. *International journal of electronic commerce*, *7*(3), 101-134.
[28] Qashou, A. (2021). Influencing factors in M-learning adoption in higher education. *Education and information technologies*, *26*(2), 1755-1785.
[29] Rahmi, B. A. K. I., Birgoren, B., & Aktepe, A. (2018). A meta analysis of factors affecting perceived usefulness and perceived ease of use in the adoption of e-learning systems. *Turkish Online Journal of Distance Education*, *19*(4), 4-42.
[30] Saroia, A. I., & Gao, S. (2019). Investigating university students' intention to use mobile learning management systems in Sweden. *Innovations in Education and Teaching International*, *56*(5), 569-580.
[31] Sharma, D., & Jain, S. (2020). Carcinogenic risk from exposure to PM2. 5 bound polycyclic aromatic hydrocarbons in rural settings. *Ecotoxicology and environmental safety*, *190*, 110135.
[32] Shishov, S. E., Yu, S., Lyakhova, N. B., Pivneva, S., Kapustina, D. M., & Arkatov, P. A. (2021). Digitalization Policy Influence: Implementation of Mobile Learning in the University Educational Process. *Management*.
[33] Shoukat, S., Bhatti, H. N., Iqbal, M., & Noreen, S. (2017). Mango stone biocomposite preparation and application for crystal violet adsorption: a mechanistic study. *Microporous and Mesoporous Materials*, *239*, 180-189.
[34] Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management science*, *46*(2), 186-204.
[35] Venter, P., van Rensburg, M. J., & Davis, A. (2012). Drivers of learning management system use in a South African open and distance learning institution. *Australasian Journal of Educational Technology*, *28*(2).

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index No: TL-30-56*

# A Needs Analysis for English for Specific Purposes Online Course for Accounting: A Tertiary Level Sri Lankan Study

Imani Randhuli (1st Author)
*Department of Computing*
*Informatics Institute of Technology (IIT)*
Colombo, Sri Lanka
irandhuli94@gmail.com

***Abstract* — Curriculum designing has been considering English for Specific Purposes pertaining to many areas of study to be delivered online. Accounting is yet another area of study which comes under Management studies. The present study addresses the need to develop a curriculum to meet the needs of the undergraduates in the field, making them ready for the world of work. The present research intends to investigate the subjective needs, present needs and target needs of the sample. The sample consists of fifty undergraduates at the Department of Accountancy, University of Kelaniya. In order to validate the industry needs, six industry professionals with more than seven years of experience are interviewed. The interviews are audio recorded and transcribed verbatim. The data gathered are analyzed mainly qualitatively however student questionnaire's closed ended items in the Google Form are analyzed qualitatively using SPSS. The findings reveal that students are least confident about their speaking skills and expecting to improve both speaking and writing skills. According to the professionals, speaking, writing and reading are equally demanded at workplaces. Business emails, letters, presentations, conversations are said to be the most recurrent instances of English use. The methods of learning preferred are group work, pair work, grammar exercises, vocabulary exercises and individual work. Since the course is delivered online, the item which inquires about online learning gets positive responses from the majority although challenges are explained in the open-ended item that followed. However, in order to cater to the specific needs of Accounting students, both perceptions of students themselves and industry professionals are taken into account to create a more effective course with relevant content.**

***Keywords* — *needs analysis for online English course, English for specific purposes, English for Accounting (online course)***

## I. INTRODUCTION

The needs analysis is conducted to design an English course- "English for Accounting". The target audience is level one (first year) undergraduates of the Department of Accountancy, Faculty of Management Studies, University of Kelaniya. The author has been teaching the students for one Academic year. Currently, the duration of the course is two years and it is a Certificate course which is now a compulsory course unit in their first two academic years. Until half of the second semester, I, and the lecturers of all the departments of Faculty of Management Studies, have been using self-made materials out of a general textbook. However, we are now using a revised and a subject specific text material designed for all five departments of the Faculty of Management Studies which is more effective.

## II. OBJECTIVES

The objective of conducting the current needs analysis for the Department of Accountancy is to investigate their own needs in the field of Accounting which could be different from that of any in the faculty along with the method of delivery being changed to "online" due to the pandemic. The study also attempts to investigate the role played by technology in the delivery of the course and discuss effective implementation methods. The present study is conducted to analyse the needs in order to design an effective course for the audience. Saragih [1] cites, setting up a new course means "a skillful blending of what is already known about language teaching and learning with the new elements that a group of learners bring to the classroom; their own needs, wants, attitudes, knowledge of the world". Hence, the research extends to investigate the needs of the undergraduates of the Department of Accountancy.

The research questions are as follows;

i. How do students perceive their current proficiency?

ii. What are the language needs in the industry of Accountancy?

ii. What are students' perceptions on the methods to be employed in the classroom?

## III. METHODOLOGY

The purpose of the needs analysis is to investigate the present needs of the undergraduates of the Department of Accountancy in order to design a course- English for Accountancy under ESP- English for Specific Purposes. The sample consists the students and the employees in the field. Fifty first undergraduates who are studying at the Department of Accountancy following either studying the Bachelor of Business Management (Special) Degree in accounting or Bachelor of Business Management Honors in Auditing and Forensic Accounting and six professionals in the field of accounting are selected. Purposive sampling is utilized in choosing the student sample whereas the employees are chosen according to convenient sampling technique. Hutchinson and Waters [2] discover learners' needs and necessities and lacks and wants by a variety of means: "by testing, by questioning and interviewing, by recalling previous performance, by consulting employers, by collecting data such as textbooks and manuals and by investigating the situations where the learners will need to use the language [3] Hence, the present study also includes both parties; employees apart from students, in order to identify industry needs better. The study follows a mixed approach. The students are given a questionnaire comprising 21 items, a Google form with multiple choice questions and three open ended questions. The questionnaire is sectioned into four parts which elicited their Background, Current Proficiency, Strengths and Weaknesses and Subject Areas of Use in the Field. On the other hand, a short Google questionnaire and short interviews are conducted to gather data. In fact, interviewing is considered an effective method of collecting more in-depth and emergent data [4] and qualitative research designs tend to work with a relatively small number of cases. The semi-structured interviews were audio-recorded and transcribed.

## IV. RESULTS AND DISCUSSION

This section discusses the results and findings of the study. The quantitative data will be analyzed followed by qualitative data. Apart from student questionnaires, interviews with professionals are transcribed. As implemented in the ESP (English for Specific Purposes) study of Alsamadani [5]. The interview was used to support and/or explain the results obtained from the questionnaire.

### *A. Background- Subjective Needs Analysis*

The participants of the Google questionnaire are fifty level I undergraduates at the department of Accountancy, University of Kelaniya where 96.2% belongs to the age range of 20-23. There are students from all provinces, however, the majority are from the Western Province. 96.2% of the participants state that their first language is *Sinhala*. The profile of the professionals is also equally important. All of them are from Western province. Four of the six participants have experience for 11 years or more while the other two participants have experience in the field for 7-10 years. This is evidence that they have been in the field for a considerably long period that they are able to share the experience over the years. The sample contains two Accountants, a Manager, an Assistant Manager and a Chief Accountant and a Deputy Auditor General.

As for subjective needs and present situation analysis, item 06 in the student questionnaire inquires about whether they have learnt any courses beyond school learning. The results indicate that the majority which accounts for 78% have learnt English in school.

As illustrated in Fig. 1, the majority of the students which accounts for 76% of the sample have learnt English in schools whereas few have gone beyond and either followed a certificate course, diploma or a private class.

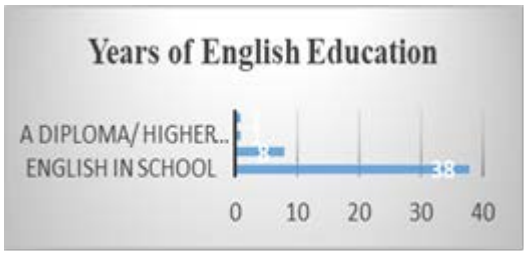



Fig. 1. Years of English Education

### *B. The Need for English Skills in the field of Accounting (Present Situation Analysis)*

Table 1. Students Self-evaluation of English Proficiency

| category | Very Weak | Weak | Fair | Very Good |
|---|---|---|---|---|
| Grammar | 2 | 23 | 24 | 1 |
| Vocabulary | 6 | 22 | 22 | 0 |
| Pronunciation | 1 | 20 | 28 | 1 |
| Speaking | 5 | 26 | 21 | 0 |
| Reading | 0 | 7 | 37 | 6 |
| Listening | 1 | 15 | 31 | 3 |
| Writing | 1 | 15 | 31 | 3 |

One obvious advantage of needs analysis is that, by identifying them learners in general, needs analysis will identify the weaknesses and strengths of the skills that they may use in potential business or academic contexts) [5]. In Sri Lanka, there is the gap between the industry needs and existing English skills

In order to investigate into that, the present level of proficiency in English, according to the students' perceptions is taken into account. The rated their proficiency in a rating scale of 5, 5 being the highest. As for overall English proficiency, none of the participants rated 5= highest. 25 accounting for 50% of the sample rated their proficiency as "average", 36.5% as "weak", and 1 student as "very weak". Hence, there is the implication that the majority of the students are not very confident about their current knowledge of knowledge of English. However, an interesting finding is, the Professionals rated their overall proficiency, 2 of them as "very good", another 2 as "Excellent" and the remaining 2 as "Good".

As illustrated in Table 1, at a glance it is evident that some skills need more attention and that the group is multi-level

everyway that there is a need to cater to the range of different proficiencies.

*C. Skills that needs Improvement (Lack Analysis)*

Item 7 of the questionnaire inquires whether the learners are more confident in speaking or writing. 73.1% of the participants are more confident in writing than speaking where 26.9% are more confident in speaking than in writing. Item 8 observes which productive skill they perceive that needs improvement.

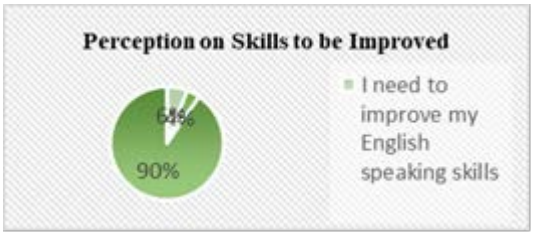



Fig. 2. Perceptions on Skills to be improved

*D. Tasks and Activities to be covered (Target Situation Analysis)*

According to Miyake and Tremaco [6], Target situation analysis refers to the tasks and activities in which English is used. Target needs, as defined by Hutchinson and Waters [2], refer to "what the learner needs to do in the target situation." Their necessities, lacks and wants [5] have to be analyzed.

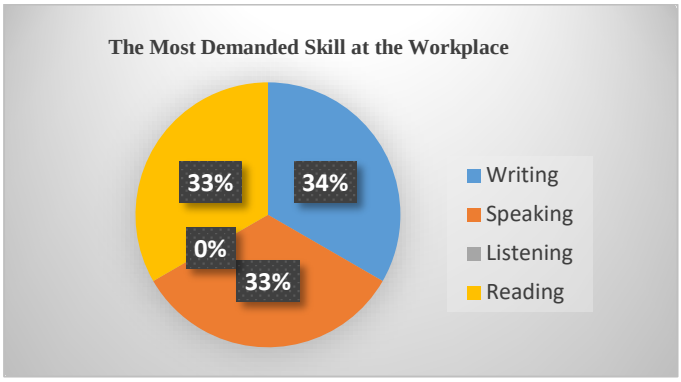



Fig. 3. The Most Demanded Skill at the Workplace from the Perspective of Accounting Professionals

Fig. 3 illustrates writing, speaking and reading as the most demanded skills at the workplace as rated by the professionals of accounting in the field. Notably, the skills are equally demanded. Parallel to the findings of the study by Alsamadani [5], writing skills probably out of the scant attention given to it, are perceived by as one of the most important targeted needs. However, as shown in Table one, there are many students who perceive they are not very good in those skills which is substantial and needs attention.

Communication at the workplace is said to be demanded at several stages. The semi-structured interviews with the professionals revealed that there are challenges at workplaces unless the skills are developed.

A communicative approach or a focus of communication needs to be integrated in the syllabus. In such an approach, 'while the teaching of structures, vocabulary and pronunciation are not neglected, learners are generally given the opportunity to communicate in speech and writing, with an emphasis on fluency and the primacy of communication'. In fact, this could benefit the undergraduates of Accounting.

As described by the participants (Table 2), there are instances where English is required at the workplace. This is further investigated in two ways; the instances of English use and activities that would benefit.

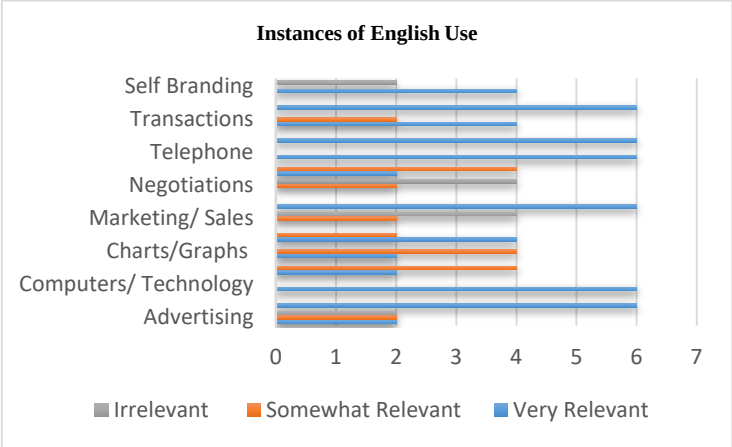



Fig. 4. Instances of English Use- Industry Professionals' Perspective

As illustrated in Fig. 4, the importance of several topics in curriculum designing can be influenced by the data collected. There are several topics and areas that are said to be of more relevance than the rest, moreover they need focus. Writing and speaking skills are demanded in various situations and contexts.

Apart from that, item 19, open ended question of the student questionnaire inquires any other topics that is of interest which they think would benefit their career in the future. The very same question was asked from the industry professionals.

### *E. Learning Needs Analysis*

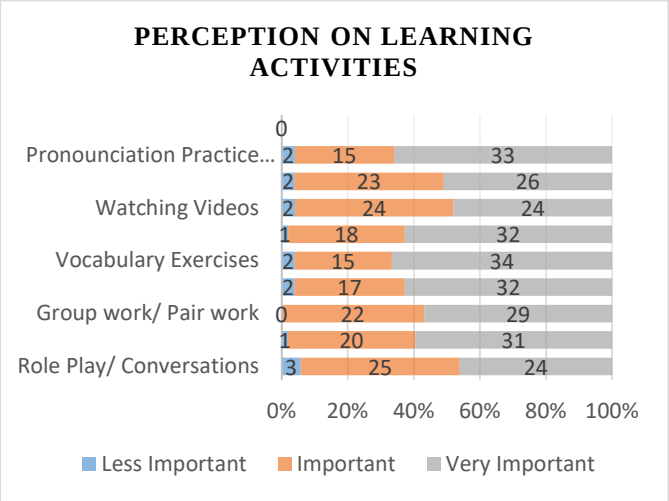



Fig. 5. Student Perceptions on Learning Activities

Miyake and Tremarco [6] define learning needs as language learning information about the effective ways of learning the skills and language. This includes learning techniques and preferences. Thus, an item was included in the questionnaire in order to investigate their learning preferences.

Fig. 5 illustrates the student preferences for various types of activities, how important they think it would be to practice them in the course. Since they are the stakeholders it is very important to take their opinion in this regard into account. Pronunciation practice activities, vocabulary exercises, grammar exercises and presentations are perceived as being of highest importance.

Furthermore, since grammar teaching and standard pronunciation have been subjected to debates in second language contexts, it is beneficial to investigate how grammar and pronunciation practice which are identified as two of the most important areas that needs practice (Fig. 3). A question was posed at the industry professionals at the semi-structured interview to which they pinpointed that grammar is important in mainly writing due to the professionalism that it entails.

### *F. Means Analysis*

Means analysis deals with the information about the environment in which the course will be run [6]**.** The intended course will be delivered either in an online environment or blended learning situation. It is important to know whether the stakeholders are able and willing to adapt to the environment in which the course will be delivered.

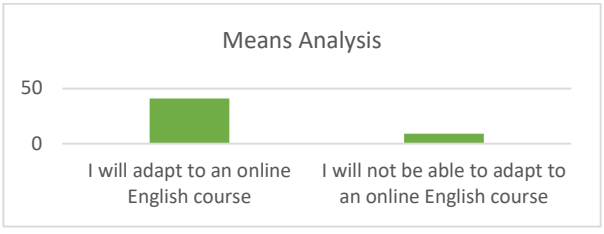



Fig. 6. Perception on Platform where the Course is delivered

Another item was included in the questionnaire in order to gather their overall perception about learning English online. However, some of them had concerns which can be related to “interaction”. The lessons, therefore, needs to be designed in such a way that it attempts to address the means, the platform in which the course will be delivered, to some extent.

As illustrated in Table 1, at a glance it is evident that some skills need more attention and that the group is multi-level everyway that there is a need to cater to the range of different proficiencies.

### *G. Skills that needs Improvement (Lack Analysis)*

Item 7 of the questionnaire inquires whether the learners are more confident in speaking or writing. 73.1% of the participants are more confident in writing than speaking where 26.9% are more confident in speaking than in writing. Item 8 observes which productive skill they perceive that needs improvement.

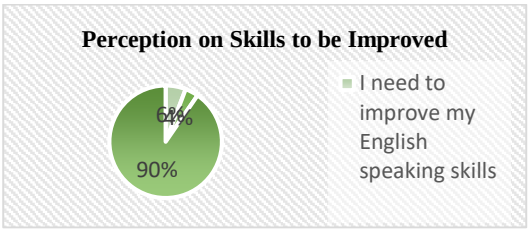



Fig.7. Perceptions on Skills to be improved

### *H. Tasks and Activities to be covered (Target Situation Analysis)*

According to Miyake and Tremaco [6], Target situation analysis refers to the tasks and activities in which English is used. Target needs, as defined by Hutchinson and Waters [2], refer to “what the learner needs to do in the target situation.” Their necessities, lacks and wants [5] have to be analyzed.

Fig. 8 illustrates writing, speaking and reading as the most demanded skills at the workplace as rated by the professionals of accounting in the field. Notably, the skills are equally demanded. Parallel to the findings of the study by Alsamadani [5], writing skills probably out of the scant attention given to it, are perceived by as one of the most important targeted needs. However, as shown in Table one, there are many

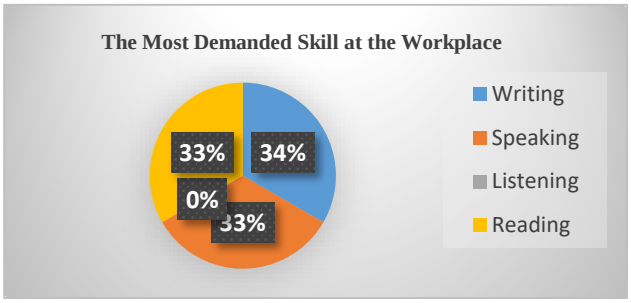



Fig. 8. The Most Demanded Skill at the Workplace from the Perspective of Accounting Professionals

students who perceive they are not very good in those skills which is substantial and needs attention.

Communication at the workplace is said to be demanded at several stages. The semi-structured interviews with the professionals revealed that there are challenges at workplaces unless the skills are developed.

A communicative approach or a focus of communication needs to be integrated in the syllabus. In such an approach, 'while the teaching of structures, vocabulary and pronunciation are not neglected, learners are generally given the opportunity to communicate in speech and writing, with an emphasis on fluency and the primacy of communication'. In fact, this could benefit the undergraduates of Accounting.

As described by the participants (Table 2), there are instances where English is required at the workplace. This is further investigated in two ways; the instances of English use and activities that would benefit.

As illustrated in Fig. 8, the importance of several topics in curriculum designing can be influenced by the data collected. There are several topics and areas that are said to be of more relevance than the rest, moreover they need focus. Writing and speaking skills are demanded in various situations and contexts.

Apart from that, item 19, open ended question of the student questionnaire inquires any other topics that is of interest which they think would benefit their career in the future. The very same question was asked from the industry professionals.

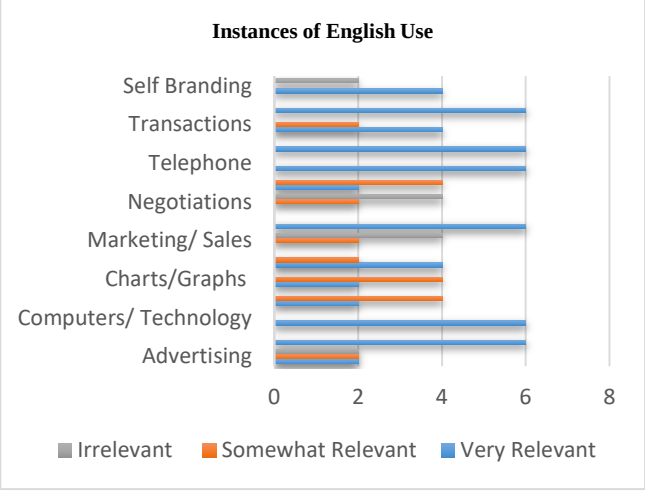



Fig. 9. Instances of English Use- Industry Professionals' Perspective

## I. *Learning Needs Analysis*

Miyake and Tremarco [6] define learning needs as language learning information about the effective ways of learning the skills and language. This includes learning techniques and preferences. Thus, an item was included in the questionnaire in order to investigate their learning preferences.

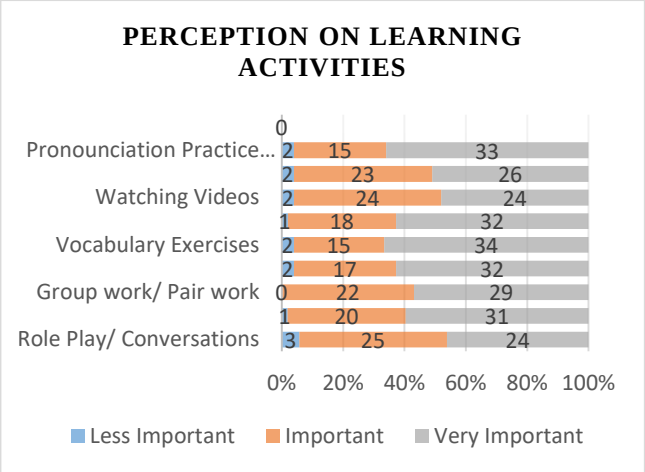



Fig. 10. Student Perceptions on Learning Activities

Fig. 10 illustrates the student preferences for various types of activities, how important they think it would be to practice them in the course. Since they are the stakeholders it is very important to take their opinion in this regard into account. Pronunciation practice activities, vocabulary exercises, grammar exercises and presentations are perceived as being of highest importance.

Furthermore, since grammar teaching and standard pronunciation have been subjected to debates in second language contexts, it is beneficial to investigate how grammar and pronunciation practice which are identified as two of the most important areas that needs practice (Fig 8). A question was posed at the industry professionals at the semi-structured interview to which they pinpointed that grammar is important in mainly writing due to the professionalism that it entails.

## J. *Means Analysis*

Means analysis deals with the information about the environment in which the course will be run [6]. The intended course will be delivered either in an online environment or blended learning situation. It is important to know whether the stakeholders are able and willing to adapt to the environment in which the course will be delivered.

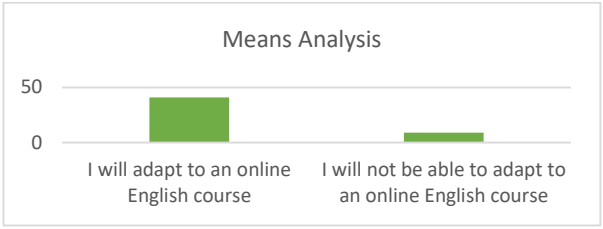



Fig. 11. Perception on Platform where the Course is delivered

Another item was included in the questionnaire in order to gather their overall perception about learning English online. However, some of them had concerns which can be related to "interaction". The lessons, therefore, needs to be

designed in such a way that it attempts to address the means, the platform in which the course will be delivered, to some extent.

## V. CONCLUSION

The research found that the majority of the stakeholders of the course have followed English in school. They are more confident in writing skills than speaking. Majority of the sample rated that they are "not proficient" in speaking and good in "reading". This correlates with the findings of Rahmatunisa and Augustiana [8] needs analysis for Indonesian Accounting students where the sample indicated lower proficiency for "speaking". As for productive skills, however, the present sample perceives that they need to improve both writing and speaking skills. The most demanded skills in industry environment are speaking, reading and writing. This corresponds to Alsamadani's [5] finding that writing skill, although neglected, is very important in workplaces.

The needs analysis also focuses on the English use in the workplace (see Fig. 8). The findings indicate that email writing, business documents and letters, interviews, money and finance, computers and technology and transactions are very relevant topics.

Learning needs analysis have provided insight into the type of activities that would benefit the target learners. The participants highly prefer grammar activities, vocabulary activities, pronunciation related activities, presentations, group work/pair work etc. Similar findings are indicated in Rahmatunisa and Augustiana [8] research where the sample prefers "group learning". The course can be designed in a way to in-cooperate a range of activities. Furthermore, although the students have a positive attitude on online learning, where the course will be delivered, the answers to the open ended question suggested that there are several challenges pertaining to interaction and speech practice that needs attention in curriculum design.

Especially in ESP courses, the contexts need to be taken into account since more focus is placed on the utility value of language. The present study attempts to discuss the present needs, subjective needs, target needs, lacks analysis and means analysis of a Needs Analysis research to design an English course for Accounting. Further research can be conducted with a different sample (2nd year/ 3rd year) if an extended course needs to be designed. Due to the pandemic in the time of conducting student data were gathered using a Google form. Probably, interviews with some students would have provided more insight since they will be less restricted. Further research can implement different instruments of data collection and analysis.

In a teaching perspective, a curriculum, although is a guidance of teaching could be and needs to be altered to different environments. As Nation and Macalister [3] states, even where a course uses a prescribed textbook, has an externally imposed syllabus, or leads up to an externally set test, there are still plenty of opportunities for negotiation, particularly as to how activities are carried out and how internal assessment is done. Hence, individual needs of the particular group of learners are non-negligible in every stage of teaching, from curriculum design to lesson planning.

## REFERENCES

[1] Saragih, E. (2014). Designing ESP materials for nursing students based on needs analysis. International journal of Linguistics, 6(4), 59.

[2] Alsamadani, H. A. (2017). Needs Analysis in ESP Context: Saudi Engineering Students as a Case Study. Advances in Language and Literary Studies, 8(6), 58-68.

[3] Brown, J. D. (2009) Foreign and second language needs analysis, In M. H. Long & C. J. Doughty (Eds) The Handbook of Language Teaching, Oxford: Blackwell Publishing Ltd. DOI: https://doi.org/10.1002/9781444315783.ch16

[4] Hutchinson, T., & Waters, A. (1987). English for specific purposes. Cambridge: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511733031

[5] Miyake, M., & Tremarco, J. (2005). Needs analysis for nursing students utilizing questionnaires and interviews. Kawasaki journal of medical welfare, 11(1), 23-34.

[6] Nation, I.S.P., & Macalister, J. (2009). Language curriculum design. New York: Routledge.

[7] Nunan, D. (2000). The learner-centred curriculum. Cambridge: Cambridge University Press.

[8] Rahmatunisa, W., & Agustiana, V. (2018). Needs Analysis of 'English for Accounting'in Tertiary Education. ELT in Focus, 1(2), 52-64.

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index No: TL-31-62*

# Relationship between Demographic Factors and the Misuse of Recreational Drugs among the Students of the Faculty of Science of the University of Kelaniya

Thanuja Dewanmini (1st Author)
*Department of Statistics and Computer Science, Faculty of Science*
*University of Kelaniya*
Paliyagoda, Sri Lanka

Harshana Munasinghe (2nd Author)
*Department of Statistics and Computer Science*
*Faculty of Science*
*University of Kelaniya*
Paliyagoda, Sri Lanka
harshanam@kln.ac.lk

***Abstract* — Drug abuse has been identified as a global health issue. It can be identified as one of the root causes of destroying peoples' lives and souls. This has a huge impact on their academic performance as well as the productivity of their lives. This study has explored the factors affecting the misuse of recreational drugs among university students. The main objective of the study was to identify the relationship between demographic factors and the misuse of recreational drugs among the students of the Faculty of Science of the University of Kelaniya. Data were collected through an online questionnaire survey. Snowball sampling was used as the sampling technique. A descriptive analysis was performed and a binary logistic regression, Support Vector Machine (SVM), and Probabilistic Neural Network (PNN) were used to evaluate the best statistical model to predict drug usage. Participants included 220 students from the Faculty of Science of the University of Kelaniya. The descriptive analysis showed that male drug usage was higher than female drug usage. Also, the drug consumption of the third-year students was higher than the other students. Students whose parents were illiterate showed a higher value of drug consumption than the other students. Also, the drug usage among the students who lived in the hostel was significantly higher than the others. It was also revealed that the most commonly used drug is alcohol. Among the fitted models, SVM Non-Random spilt model showed the highest accuracy (93.1818%) in predicting drug usage. Based on the results, gender, religion, year of study, involvement of a part-time job, participation in the sports activities, financial support from bursary or mahapola, mother's education level, father's occupation, ethnicity, marital status of the parents were identified as the associated factors of the drug usage among the students.**

***Keywords* — *Alcohol, Drug abuse, Drug usage***

## I. INTRODUCTION

Human beings have had a deep connection to their social environment since past time, and have therefore maintained an intimate, cooperative, and long-term relationship with one another. As can be seen throughout history, people adapted and coordinated themselves to achieve their common goals, and this generally resulted in positive and desired outcomes. Globalization has fundamentally altered the essence of traditional society, transforming it into a tangle of contradictions unlike anything seen before in any conventional society. These changes and complexities have resulted in, and continue to result in, a variety of social issues among citizens. Teenage pregnancy, drug addiction, poverty, domestic violence, child abuse, trafficking, traffic accidents, divorce, and crime are only a few examples of today's social issues, many of which have negative implications for individuals and communities. Drug abuse has recently emerged as one of the world's most significant social issues. This problem hasn't an age limit. In every age limit, people are addicted to this matter. Substance abuse among youth is a huge problem. University students are facing this problem at that moment. The majority of drug users will argue that using drugs is a personal matter and that their actions affect only them and not others. Everyone knows, however, that it is not as easy as the drug addicts say. It is a well-known fact that drug addiction influences not only drug users, but also their families, culture, and world. Drug abuse is not a straightforward issue that just affects a few people. It is a multi-faceted, all-encompassing social phenomenon that affects almost every country in the world. Drug misuse and addiction are not new occurrences. Millions of people all over the world have lost control of their lives as a result of drug addiction. In particular, illegal drug use has increased dramatically across the world in recent decades. Many people have begun to use drugs without realizing the harmful effects that these substances can have on their own lives and the lives of their dependents. In the origin, People only used drugs on rare occasions at first, and they were capable of handling themselves in unusual circumstances. However, the scene has gradually changed as drugs have taken hold of people's

lives and made them believe they could not survive without them. Drug abuse has enslaved a large portion of the population in many countries around the world today, which is tragic. Drug addiction has now been added to the list of social problems, and it is particularly noticeable among the poor, and working population of Sri Lanka.[1]

This research has explored the Relationship between Demographic Factors and Misuse of Recreational Drugs among the Students of the Faculty of Science of the University of Kelaniya. The demographic factors of the students were considered for the analysis. Namely, Gender, Ethnicity, Religion, Accommodation place, Marital status of parents, Life status of parents, Father's Occupation, Father's Education Level, Mother's Education Level, Scholarship status, Engagement of the part-time jobs, Engagement in sports activities, Accommodation place and year of study at university. These factors were identified by studying the relevant research papers. These variables were used to find the determinants of drug usage and to develop a statistical model to predict drug usage

## II. OBJECTIVES

The main objective of this study was to identify the relationship between demographic factors and misuse of recreational drugs among university students. Also, this study helped to identify the present situation of drug addiction among the Students of the Faculty of Science of the University of Kelaniya, Which year university students more use drugs, Investigate the drug abuse of students who participates in sports, Investigate the drug abuse of students who do a part time job, Develop a statistical model to predict the drug usage, Identify the types of drug currently being used by university students, their involvement in multiple drug misuse and drugs combinations, To determine the frequency of drug abuse in a sample of a university student. And also, the demographic variables were used to find the determinants of the drug usage and to develop a statistical model to predict the drug usage.

## III. METHODOLOGY

In this study gender, ethnicity, religion, marital status of parents, life status of parents, father education, mother education, father occupation, accommodation, year, the status of the scholarships (Mahapola/bursary), engagement in a part-time job, and engagement in sports were used to find out the influence on the drug usage of the university students. All the outputs of the research are obtained with the use of Minitab, SPSS, and MATLAB software. A sample of 220 students who are studying at the Faculty of Science of the University of Kelaniya was used in this research. The sample was obtained with the use of the Snowball Sampling procedure. A descriptive analysis was carried out to check the composition of the sample. Correlation analysis was conducted using the Chi-square test in order to find out the most influential factors of drug usage. After checking the association between the independent variables and the dependent variable, the following variables were identified as they were having a significant association with the response variable. Gender, Religion, Father's Occupation, Year of study, Engagement in a part-time job, Engagement in sports activities. These variables were used to fit the binary logistic regression model in order to predict drug usage. Since the VIF for all independent variable were less than 10 multicollinearity was identified among the independent variables. Support Vector Machine (SVM) random split and nonrandom split data divide method and Probabilistic Neural Network (PNN) random split and nonrandom split data divide methods were used to evaluate the best statistical model to predict drug usage. The polynomial kernel function has been chosen as the optimal kernel function in the SVM. 0.1 was the optimal spread parameter in the PNN. Model diagnostic analysis was carried out to check the accuracy of all the fitted models.

## IV. RESULTS AND DISCUSSION

According to the fitted binary logistic regression model, male drug usage is 2.008 higher than the females. For students who are doing a part-time job their drug consumption value is higher at 1.564 than the students who are not doing a part-time job. Drug consumption of the students who are engaged in Sports is 0.493 times higher than a student who is not doing sports. Third-year students' drug consumption and second-year students' drug composition are higher than the fourth-year students' drug composition. As a value, it is 1.289 times and 0.066 times respectively. Third-year students' drug consumption is 1.494 times higher than first-year students. Drug usage of Hindu students is 2.946 times higher than the drug usage of Muslim students and 1.411 higher than Buddhists. Christian students' drug consumption is 2.079 times higher than muslin students' drug consumption. The students whose father works in the private sector have used drugs 1.328 times higher than students whose father is Unemployed (base category).

According to the Table 1, we can conclude that the Gender, Part-time job variable (Do you do any part-time job?), Year 3, Hinduism (Religion), variables are significant in the fitted model. Since they are positively correlated, they have a positive association with students' drug usage after controlling other variables.

Based on the results of binary logistic analysis Gender, part-time job, Year 3, Hinduism (Religion) variables are the only influenced factor for drug usage from the final model with the accuracy of 79.1% correctly classified.

Table 1: Variable in the Equation Table

**Variables in the Equation**

| | | B | S.E. | Wald | df | Sig. | Exp(B) |
|---|---|---|---|---|---|---|---|
| Step 1[a] | G1 | 2.008 | .463 | 18.808 | 1 | .000 | 7.451 |
| | R1 | 1.535 | .851 | 3.251 | 1 | .071 | 4.641 |
| | R2 | 2.079 | 1.126 | 3.410 | 1 | .065 | 7.996 |
| | R3 | 2.946 | 1.152 | 6.537 | 1 | .011 | 19.031 |
| | P1 | 1.564 | .476 | 10.812 | 1 | .001 | 4.780 |
| | S1 | .493 | .387 | 1.625 | 1 | .202 | 1.637 |
| | F1 | -.271 | .945 | .082 | 1 | .775 | .763 |
| | F2 | 1.328 | .934 | 2.020 | 1 | .155 | 3.772 |
| | F3 | 1.095 | .941 | 1.353 | 1 | .245 | 2.988 |
| | F4 | .323 | .926 | .122 | 1 | .727 | 1.381 |
| | Y1 | -.205 | .609 | .114 | 1 | .736 | .814 |
| | Y2 | .066 | .587 | .013 | 1 | .911 | 1.068 |
| | Y3 | 1.289 | .581 | 4.926 | 1 | .026 | 3.630 |
| | Constant | -4.922 | 1.411 | 12.162 | 1 | .000 | .007 |

a. Variable(s) entered on step 1: G1, R1, R2, R3, P1, S1, F1, F2, F3, F4, Y1, Y2, Y3.

Then fitted the support vector machine to the data set. First, we divide the data set-in two-part training set and testing set. In the analysis below the two sets are divided as 80% training set and 20% testing set. Artificial Neural network modeling has been carried out using MATLAB mathematical software. To measure the performance using a classification matrix. First, we divide the whole data set nonrandom split. This data set has fourteen variables. Firstly, label the demographic variable Gender(G), Ethnicity(E), Religion(R), Marital Status of Parents (MS), Life Status of Parents (LS), Father Education (FE), Mother Education(ME), Father Occupation (FO), Accommodation(A), Bursary/Mahapola (BM), Part-time job(P), Sports(S), Year(Y). Then remove one by one variable and checked the accuracy. Then checked the most accurate input combination by removing one by one variable out of fourteen variables. Father Occupation, Part time job, Mother Education, Religion, Marital Status, Sport, Year, Gender, Ethnicity and student have a bursary or mahapola variables were included most accuracy input combination model. Then did the parameter adjustment of this model. There are many Kernel Functions that can be used in support vector machine. Polynomial, Gaussian, Rbf (Radial Basis Function), Linear kernel are the most commonly used Kernel Function. When the Kernel Function change while all other factors are Keep Constant. Gaussian Kernel Function and Rbf Kernel Function have the almost same performance in this SVM. All the other kernel function performance was lesser than the polynomial kernel function performance. Therefore, the polynomial Kernel function has been chosen as the kernel function in the SVM. After changing the kernel function in the SVM final model has included the polynomial kernel function and Father Occupation, Part time job, Mother Education, Religion, Marital Status, Sport, Year, Gender, Ethnicity and student have a bursary or mahapola input variables.

Table 2. Overall performance of SVM Non-Random Split model

| Accuracy | Misclassification | Precision | Recall | F-measure |
|---|---|---|---|---|
| 93.1818 | 6.82 | 0.89 | 0.94 | 0.91 |


Fig. 1. Classification Chart

1-Not Use Drug, 2- Drug Use

This is the classification matrix of the optimum model Out of the training data set, the model predicts the 25 students are not using the drug. They are actually not drug use students. And also, the model predicts that 16 students are used drug. They are actually drug used students. The model predicts the two students are not used drugs. But actually, they have used drugs. And also, the model predicts the one student who used a drug. But actually, this student not used the drug

Then checked the SVM model by using the random split data dividen method. This method firstly identified the most accurate input combination. Removed one by one variable in the full model and checked the accuracy. Then we identified the less accurate variable in the full model. Accommodation(A), Life Status of Parents (LS), Ethnicity(E), Father Occupation (FO), Marital Status Parents (MS) variable have the lowest association in this SVM random split data dividen model. Then removed the lowest association variables combination. Checked the highest accuracy model. Father Occupation (FO), Marital Status of Parents (MS), Religion(R), Mother Education (ME), Sport(S), Student who have Bursary or Mahapola (BM), Part-time job(P), Gender(G), Year(Y) Variables were included highest accuracy random split support vector machine model. Then did the parameter adjustment of this model. There are many Kernel Functions that can be used in support vector machine. Polynomial, Gaussian, Rbf (Radial Basis Function), Linear kernel are the most commonly used Kernel Function. When the Kernel Function change while all other factors are Keep Constant. Gaussian Kernel Function and Rbf Kernel Function have the almost same performance in this SVM. All the other kernel function performances are lesser than the polynomial kernel function performance. Therefore, the polynomial Kernel function has been chosen as the kernel function in the SVM. After changing the kernel function in the SVM final model has included the polynomial kernel function and Father Occupation, Part time job, Mother Education, Religion, Marital Status, Sport, Year, Gender

and student have a bursary or mahapola input variables. The below table shows the performance statistic of the fitted model.

Table 3. Overall performance of SVM Random Split model

| Accuracy | Misclassification | Precision | Recall | F-measure |
|---|---|---|---|---|
| 77.95 | 22.05 | 0.82 | 0.82 | 0.82 |

Then fitted the Probabilistic Neural Network model to the data set. First, we divide the data set-in two-part training set and testing set. In the analysis below the two sets are divided as 80% training set and 20% testing set. Artificial Neural network modeling has been carried out using MATLAB mathematical software. To measure the performance using a classification matrix.

This data set have fourteen variables. In this section to identify what are optimum input combination. Firstly, label the demographic variable Gender(G), Ethnicity(E), Religion(R), Marital Status of Parents (MS), Life Status of Parents (LS), Father Education (FE), Mother Education (ME), Father Occupation (FO), Accommodation(A), Bursary/Mahapola (BM), Part-time job(P), Sports(S), Year(Y)these inputs were presented to the network in different combinations and the optimum combination of inputs. Performance measures were used to identify the best input combination. All the other factors and parameters in the network were kept constant during the training process. Marital Status of Parents, Sports, Gender, Religion, Life Status of Parents, Student have Bursary or Mahapola, Year these inputs variables were identified highest accuracy model. PNN has only one parameter adjustment. It is the spread parameter. After getting the optimum input combination model then we changed the spread parameter. The spread parameter should be near zero. Because The network performs the function of the closest neighbor classifier. As the spread widens, the designed network considers many nearby design vectors.

Table 4. Performance of different Spread Parameter in PNN For Non-Random Split

| Spread | Accuracy | Misclassification | Precision | Recall | F-measure |
|---|---|---|---|---|---|
| 0.1 | 90.91 | 9.09 | 0.93 | 0.82 | 0.88 |
| 0.2 | 90.91 | 9.09 | 0.88 | 0.88 | 0.88 |
| 0.3 | 90.91 | 9.09 | 0.88 | 0.88 | 0.88 |
| 0.4 | 90.91 | 9.09 | 0.88 | 0.88 | 0.88 |
| 0.5 | 90.91 | 9.09 | 0.88 | 0.88 | 0.88 |

According to the above table, we can conclude that the optimal model is the first model. This model spread parameter 0.1 and it has the highest Precision than the other model. All the models have the same accuracy and misclassification rate. After increasing the spread parameter to 0.1 all the performance values are the same.

Changing the spread parameter in the PNN final model has included the polynomial 0.1 spread parameter and Mother Education, Religion, Life Status of Parents, Sport, Year, Gender and student have a bursary or mahapola input variables. The below table shows the performance statistic of the fitted model.

Table 5. Overall performance of PNN Non-Random Split model

| Accuracy | Misclassification | Precision | Recall | F-measure |
|---|---|---|---|---|
| 90.9091 | 9.09 | 0.93 | 0.82 | 0.88 |

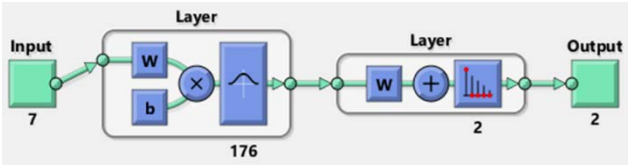



Fig. 2. Architecture of the PNN for Random Split data set

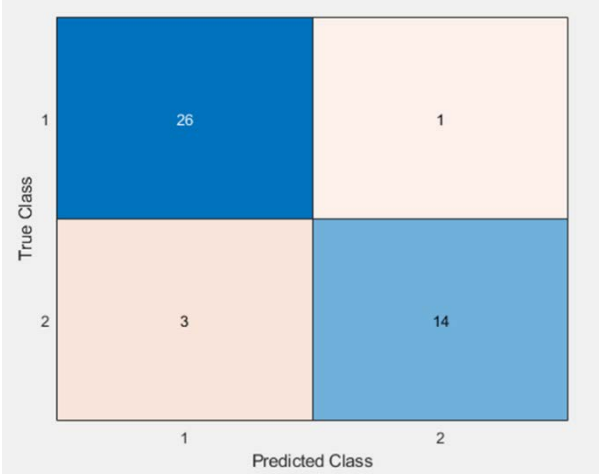



Fig.3. Confusion Chart
1-Not Use Drug, 2- Drug Use

This is the classification matrix of the optimum model Out of the training data set, the model predicts the 26 students are not using the drugs. They are remove actually not drug use students. And also, the model predicts that 14 students are used drugs. They are actually drug used students. The model predicts the three students are not used drug. But actually, they have used drugs. And also, the model predicts the one student who used a drugs. But actually, this student not used the drug.

Then fitted the PNN model by using the random data split method. This method firstly identified the most accurate input combination. Removed one by one variable

in the full model and checked the accuracy. Accommodation(A), Part time job(P), Year(Y), Sport(S), Religion(R), Marital Status of Parents (MS), Gender(G) and Year(Y) Variables were included highest accuracy model. This model includes only eight variables out of fourteen variables.

Table 6. Performance of different Spread Parameter in PNN For Random Split

| Spread | AA | AM | AP | AR | AF |
|---|---|---|---|---|---|
| 0.1 | 76.82 | 23.18 | 0.82 | 0.80 | 0.81 |
| 0.2 | 76.82 | 23.18 | 0.83 | 0.79 | 0.81 |
| 0.3 | 73.18 | 26.82 | 0.81 | 0.76 | 0.78 |
| 0.4 | 73.18 | 26.82 | 0.80 | 0.78 | 0.78 |

According to the above table, we can conclude that the optimum spread parameter 0.1. This model has high Average Accuracy (AA), Average Misclassification (AM), Average Precision (AP), Average Recall (AR)and Average F-measure (AF).

Changing the spread parameter in the PNN final model has included the polynomial 0.1 spread parameter and Accommodation, Religion, Marital Status of Parents, Sport, Year, Gender, Part-time job and Farther Education input variables. The below table shows the performance statistic of the fitted model.

Table 7. Overall performance of PNN Random Split model

| Accuracy | Misclassification | Precision | Recall | F-measure |
|---|---|---|---|---|
| 76.82 | 23.18 | 0.82 | 0.80 | 0.81 |

According to the above four models, we can conclude that the Non-Random split method model accuracy is higher than the Random split method. support vector machine non-random split method accuracy (93.1818%) is higher than the SVM random split (77.9545%), probabilistic neural network random split (90.9091%) and non-random split method (76.8182%). According to the SVM random and non-random split method, both two models include the same dependent variable there is father occupation, part-time job, mother education, religion, marital status, sports, year, gender and student have a bursary or mahapola. But SVM non-random split has included one extra variable it is student ethnicity. And also, according to the probabilistic neural network random split and non-random split method, both models have the same dependent variable there are religion, sport, year and gender. Then we consider all four models have the same dependent variable there are religion, sport, year, gender variables. Finally, we conclude that the most suitable and accurate model support vector machine (SVM) non-random split data dividend method. It has high accuracy and a low misclassification rate. It is correctly classified than the other model, a student who have used drug and a student who has not used.

Table 8. Model Comparison

| Model | Accuracy |
|---|---|
| Binary Logistic Model | 79.1% |
| Support Vector Machine Random Split Model | 77.95% |
| Support Vector Machine Non-Random Split Model | 93.18% |
| Probabilistic Neural Network Random Split | 76.83% |
| Probabilistic Neural Network Non-Random Split | 90.91% |

According to the Binary Logistic, SVM and PNN non-random split statistical model, SVM Non-Random spilt model is the highest accuracy (93.18%) model to classify the drug usage. Based on the result, gender, religion, year, does the student do any part-time job? Do the student participants in the sports activities? does the student have a bursary or Mahapola? Mother's Education Level, Father's Occupation, Ethnicity, Marital Status of Parents were in influenced factors for student drug usage.

## V. CONCLUSION

In this research, I have studied the Relationship between Demographic Factors and Misuse of Recreational Drugs among the Students of the Faculty of Science of the University of Kelaniya. The demographic factors of the students were considered for the analysis. Namely, gender, ethnicity, religion, accommodation place, marital status of parents, life status of parents, father's Occupation, Father's Education Level, Mother's Education Level, Does the student has a bursary or Mahapola? Does the student do any part-time job? Does the student participants in the sports activities? Accommodation place and year of study at university. This is a survey-based study. Data were collected by using a questionnaire. The Snowball sampling technique was used to collect the data.220 students were respondents to this questionnaire. Out of 88-drug use respondents, 13.6 % female respondents and 55.4 % male respondents, according to that result, male drug usage is more than female drug usage. Burger students have used more drug than the other students and according to their religion. Hinduism students have used drug higher than the other students. The student whose parents are Divorced has a higher value of drug consumption than the other students and the students whose parents are Both Dead have a higher value of drug consumption than the other students. According to these two results, we can conclude that parent life status and marital status highly impact student drug consumption. furthermore, Students whose parents are Illiterate have a higher value of drug consumption than other students. we can conclude that parents' education status is an important factor in student drug consumption. The students whose father works in the private sector have used drugs higher than the other students. The students who live in the hostel have used drugs higher than boarding and home and students who have bursary used drugs higher

than those who don't have bursary and Mahapola. And also, the students who are doing a part-time job and students who have participated in sports used drugs higher those who are not doing a part-time job and have not participated in a sport. Third-year students' drug consumption is higher than the other year student. The majority of the students do not use any drug (75.9%). And most commonly used drug is alcohol.

Binary Logistic regression fitted model male drug usage is 2.008 higher than the females. For students who are doing a part-time job their drug consumption value is higher at 1.564 than the students who are not doing a part-time job. Drug consumption of the students who are engaged in Sports is 0.493 times higher than a student who is not doing sports. Third-year students' drug consumption and second-year students' drug composition are higher than the fourth-year students' drug composition. As a value, it is 1.289 times and 0.066 times respectively. Third-year students' drug consumption is 1.494 times higher than first-year students.

According to the fitted Binary Logistic regression model, Support Vector Machine and Probabilistic Neural Network non-random split and random split statistical model, SVM Non-Radom spilt model is the highest accuracy (93.18%) model to correctly classified the drug usage. In this model use the polynomial kernel function. Out of the training data set, the model predicts the 25 students are not using the drug. They are remove actually not drug use students. And also, the model predicts the 16 students are using the drug. They are actually drug used students. The model predicts the two students are not used drugs. But actually, they have used drugs. And also, the model predicts the one student who used a drug. But actually, this student not used the drug. Based on the result, gender, religion, year, student engage part-time job, student participants in the sports activities, student have a bursary or Mahapola, Mother's Education Level, Father's Occupation, Ethnicity, Marital Status of Parents were in influenced factors for student drug usage.

This study is a sensitive case study because of that, I have faced a huge problem in the data collection part. Lots of time was spent on the data collecting part. This research is limited to the Faculty of Science, University of Kelaniya. This study is does not apply to students outside of this designation. This research can be extended to the whole University system to draw more accurate results.

## REFERENCES

[1] T. H. Bennett and K. R. Holloway, "Drug use among college and university students: Findings from a national survey," *J. Subst. Use*, vol. 20, no. 1, pp. 50–55, 2015, doi: 10.3109/14659891.2013.878762.

[2] B. Kasundu, M. Mutua Mutiso, P. Sutter Chebet, and P. Wanjiku Mwirigi, "Factors Contributing To Drug Abuse among the Youth in Kenya: A Case of Bamburi Location," *Elixir Soc. Sci.*, vol. 46, no. 16, pp. 8259–8267, 2012, [Online]. Available: http://www.elixirpublishers.com/articles/1350378627_46 (2012) 8259-8267.pdf.

[3] S. Senanayake *et al.*, "Smoking, Alcohol Consumption, and Illegal Substance Abuse among Adolescents in Sri Lanka: Results from Sri Lankan Global School-Based Health Survey 2016," *Adv. Public Heal.*, vol. 2018, pp. 1–7, 2018, doi: 10.1155/2018/9724176.

[4] N. Geramian *et al.*, "Development of a questionnaire to assess drug abuse among high school students of Isfahan province, Iran: An action research," *Int. J. Prev. Med.*, vol. 5, no. 14, p. 146, 2014, doi: 10.4103/2008-7802.157678.

[5] K. Holloway and T. Bennett, "Characteristics and correlates of drug use and misuse among university students in Wales: a survey of seven universities," *Addict. Res. Theory*, vol. 26, no. 1, pp. 11–19, 2018, doi: 10.1080/16066359.2017.1309031.

[6] M. Meray, F. Ahmed, and A. Rania, "Study of Drug Abuse among University Students in Sohag , Egypt," *Journal of Forensic,Toxicology and Medicolegal Analysis*, vol. 1, no. 2, pp. 19–23, 2016.

[7] O. Rezahosseini, A. Roohbakhsh, V. Tavakolian, and S. Assar, "Drug Abuse among University Students of Rafsanjan, Iran.," *Iran. J. psychiatry Behav. Sci.*, vol. 8, no. 2, pp. 81–5, 2014, [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/25053961%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4105608.

[8] M. Zaman *et al.*, "Drug abuse among the students," *Pakistan J. Pharm. Res.*, vol. 1, no. 1, p. 41, 2015, doi: 10.22200/pjpr.2015141-47.

[9] T. Evgeniou and M. Pontil, "Support vector machines: Theory and applications," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 2049 LNAI, no. May, pp. 249–257, 2001, doi: 10.1007/3-540-44673-7_12.

[10] T. H. misuse among university students in the U. I. for preventiongh Bennett and K. R. Holloway, "Drug misuse among university students in the UK: Implications for prevention," *Subst. Use Misuse*, vol. 49, no. 4, pp. 448–455, 2014, doi: 10.3109/10826084.2013.846378.

[11] F. Seguel, G. Santander, and M. L. Ramos, "Consumo de drogas y factores demográficos asociados en estudiantes de primer año de una universidad Chilena," *Acta Paul Enferm*, vol. 25, no. 2, pp. 33–39, 2012.

[12] T. Osman *et al.*, "Epidemiology of Substance Use among University Students in Sudan," *J. Addict.*, vol. 2016, pp. 1–8, 2016, doi: 10.1155/2016/2476164.

[13] I. L. M. Mahir and T. M. F. Wazeema, "Social Aspects of Drug Addiction in Sri Lanka," *J. Polit. Law*, vol. 13, no. 2, p. 54, 2020, doi: 10.5539/jpl.v13n2p54.

[14] R. Atkinson, F. Flint, and N. Gilbert, "Snowball Research Strategies," *Dep. Sociol. Univ. Surrey Guildf. GU2 5XH United Kingdom.*, no. 33, p. null-null, 2001.

[15] T. M. Franke, T. Ho, and C. A. Christie, "The Chi-Square Test: Often Used and More Often Misinterpreted," *Am. J. Eval.*, vol. 33, no. 3, pp. 448–458, 2012, doi: 10.1177/1098214011426594.

[16] K. A. Pituch and J. P. Stevens, "Binary Logistic Regression," *Appl. Multivar. Stat. Soc. Sci.*, pp. 454–490, 2020, doi: 10.4324/9781315814919-16.

[17] F. Ancona, A. M. Colla, S. Rovetta, and R. Zunino, "Implementing probabilistic neural networks," *Neural Comput. Appl.*, vol. 5, no. 3, pp. 152–159, 1997, doi: 10.1007/BF01413860.

[18] J. Sujatha and S. P. Rajagopalan, "Performance evaluation of machine learning algorithms in the classification of parkinson disease using voice attributes," *Int. J. Appl. Eng. Res.*, vol. 12, no. 21, pp. 10669–10675, 2017.

[19] D. P. Ranjan, N. Namita, and R. M. Chaturvedi, "A study of socio-demographic factors contributing to the habit of drug abuse in the urban slum community of Mumbai," *Biomed. Res.*, vol. 21, no. 3, pp. 277–284, 2010.

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index No: TL-32-64*

# Use of Early Semester Student Feedback for Enhancing Effectivness of Teaching and Learning

Indrajith D. Nissanka (1st Author)
*Faculty of Engineering*
*University of Moratuwa*
Katubedda, Sri Lanka
nissankai@uom.lk

Gayani K. Nandasiri (2nd Author)
*Faculty of Engineering*
*University of Moratuwa*
Katubedda, Sri Lanka
gayanin@uom.lk

***Abstract* — Student feedback in basic engineering modules is important, as the module involves the application of theory into practice. The feedback is used to assess the teaching and learning process at the end of the semester, which is the current practice. This mainly focuses on summative assessment through quantitative scores, where feedback is addressed in subsequent academic years. The early semester feedback can be used to improve the teaching and learning process within the semester itself. It can be designed as formative feedback, focusing on meaningful improvement of the teaching and learning. Hence, it is explored whether early semester feedback can be applied for enhancing the effectiveness of teaching and learning process. In this study, early semester feedback was obtained for the Basic Thermal Sciences module from a selected sample size of 122 students representing two engineering disciplines of the same semester of study. The feedback was collected in two stages of the semester using both paper and Moodle based online questioners. The feedback survey was designed in two sections: the first section provided for quantitative evaluation using rating questions while the second included open-ended questions to obtain qualitative feedback. Survey results depicted students' self-assessments on their learning and the suggestions for improving the teaching and learning process. The feedback provided diagnostic information on the key changes to be adopted in teaching, that resulted in improved student engagement and performance. Almost 90% of the students responded that their interest was valued, and they felt inclusive in the class while 80% of students were of the impression that class materials are relevant to their professional practice. Also, the subsequent assessment has shown a 10% increase in the average marks for group assignments. It was evident that the students were appreciative of taking the early semester feedback, and it helped to improve the inclusiveness of the student's requirements into the module.**

***Keywords —Early semester feedback, teaching and learning, Engineering education***

## I. INTRODUCTION

Engineering education over the past decades has adopted several pedagogical practices on enhancing the quality of the programmes. The lack of literature in the methods of assessing effective teaching alarms the educational practitioners; even though teaching is the most important role of the faculty members [1]. Currently, student evaluation of teaching in the end semester is considered as the only metric used to evaluate the effectiveness of the teacher's delivery method as well as the student learning [2]. The literature suggests four purposes of collecting feedback: diagnostic feedback to teachers about the effectiveness of their teaching, a measure of teaching effectiveness to be used in administrative decision making, information for the students in selecting course modules, and finally as an outcome or process description for use in research on teaching and learning. Although the student feedback on teaching has proven valid and reliable [3] [4] end semester feedback stands out to be a form of a summative assessment through quantitative scores. Hence, the instructor/ lecturer may use the feedback received to improve teaching and learning in the subsequent years. This form of feedback is not beneficial for students who provided the feedback; as well as teachers if they might not teach the same course next year. The end semester feedback is mainly used for administrative decision making and feedback forms are standardized over different disciplines, thus only providing general information about the module [5]. Hence, the end semester feedback is not the most effective method in evaluating the teaching and learning process [6].

In contrast, early semester feedback provides an opportunity to improve the teaching and learning process within the same semester. It provides the students to raise their concerns within the module, indicating what is and what is not working, while time remains for the adjustments required by the teacher. Hence, it provides a broader and deeper view into one's teaching methods and behaviours than it is likely to gain otherwise. In fact, a critically reflective teacher would cultivate a heightened awareness of their own teaching from as many different perspectives as possible [7]. Thus, early semester feedback is designed as formative feedback focusing on meaningful improvement of teaching and learning during the semester. Such feedback is specific, timely, corrective and positively framed [8] that it motivates the teacher to adopt required changes to their practice, while the student feels

involved. The literature suggests that a disconnect often occurs for first-year students, as their focus tends toward general study skills rather than the development of domain specific or course specific knowledge [9]. This along with less frequent use of feedback creates a gap between what students struggle to learn and how to improve their learning in Engineering education. In fact, this is often observed in first-year undergraduates of the Engineering faculties, where students struggle to improve their learning with regards to the course learning outcomes. The use of early semester feedback could be one way to bridge the gap, as it is considered as an effective dialogue with students. The teacher is required to listen, react, and make changes during the semester for a better learning experience. This constructive dialog with the students makes them feel they are valuable in the learning community thus, they are motivated by the lecturers' concern for their learning [10]. Hence, it improves the overall student learning experience and automatically the ratings of the end semester evaluations will improve.

The early semester feedback can be customized to obtain feedback for an overall module or on certain teaching activities. The lecturer has the flexibility to control the timing, questions, and analysis of these informal surveys, which ensures the best information is collected to contribute to student learning. There are several methods of obtaining early semester feedback depending on the class size and the required information [11]. Classroom assessment quality circles: where teachers regularly meet with the small group of students to get their feedback on the course could be used for large classes, while for small to midsize classes group instructional feedback technique could be adopted: where the external person (other than the teacher) is arranged to interview the students based on prepared open-end questions.

The teacher designed surveys, which is the most common method, could be used for any size class. In designing a feedback survey, it is important to include questions directed to diagnose teaching/ learning aspects that could be addressed during the semester and the aspects that the lecturer is willing to make a change.

An interactive discussion with students is critical in basic Engineering modules, as it requires them to apply the theoretical concepts into practice. The proactive implementation of the changes to address the student feedback is important in improving student satisfaction in their learning in Engineering modules [12].

## II. OBJETIVES

This study aims to explore the effectiveness of the use of early semester feedback within a basic Engineering module to enhance the student learning experience, using a teacher designed survey. The key objectives of the study are to: identify the key aspects of early semester feedback, apply early semester feedback techniques in basic engineering modules, evaluate the effectiveness of early semester feedback in improving the teaching and learning experience in Engineering education.

## III. METHODOLOGY

In this study, the early semester feedback strategy was implemented for the Basic Thermal Sciences module, which is a compulsory fundamental engineering module offered for both Civil Engineering and Earth Resources Engineering students for the same semester which runs through a total of 14 weeks period. The module had a total of 175 registered students from both engineering disciplines. The module discusses fundamental theoretical concepts, and it is expected from students to identify applications related to their discipline to apply the theories into practice. Thus, it is important that the teacher maintains a continuous dialogue with the students to motivate them while understanding their learning requirements. During this study student feedback was obtained in two stages during the semester: during the 3rd and 8th weeks. The first stage of the feedback was obtained specifically for an in-class group activity, where the students were allowed to provide feedback as a group in a written paper later uploaded to Moodle. This group assessment was focused on fundamentals of Engineering Thermodynamics where students had to self-study a given journal article and submit answers for six questions provided. These questions were designed such that it encourages group discussions and increase the active learning during the classroom. Total time allocated for the group assessment was 30 minutes. The group feedback strategy was implemented to avoid students' reluctance to provide feedback and to allow them to reflect on their learning experience with peers. It was expected that this would impact the student response, while providing a collective reflection on the group activity. The students were divided into 16 groups, and after completing the activity they were asked to complete the feedback forms. The feedback forms included three open ended questions: two questions about the group activity and one question on the learning experience in the class. The obtained qualitative feedback was analyzed to get any common suggestions while avoiding the outliers; based on the analysis it was identified that most of the students suggested to allocate more time for discussions. Hence, more time was allocated for small group discussions to promote active engagement and active learning. Except for this there were no other suggestions provided by the students to improve the learning experience in the class.

In the second stage, it was decided to get individual feedback from the students on their learning experience of the course module. The common method used for student feedback on teaching and learning is the questionnaire-based survey with rating questions. In recent years with the online teaching platforms and learning management systems online based teaching evaluations became much popular. Online teaching evaluation saves time, and cost while providing quick analysis and reporting compared to the traditional paper-based information collection [13]. The online evaluations provide the opportunity for students to complete the questionnaire anytime without the influence of the teacher, which allows the student to provide a better reflection on the learning experience in the class [14] However, the requirement to have access to computer and internet is one of the drawbacks of the online feedback [15], whereas low student response rate, biasness of the student's response due

to unwillingness to participate in the online survey were also highlighted by the previous research [16]. However, the advantages of an online feedback survey are significant than the disadvantages, thus, a Moodle based online feedback survey was selected to adapt to the second stage of collecting the feedback.

Table 1. Different teaching and learning aspects covered in the feedback survey

| Category | Question |
|---|---|
| Effectiveness of Teaching | The lecturer explains the material clearly. |
| | The lecturer indicates important points to remember |
| | The lecturer shows genuine interest in students |
| | The lecturer explains the thinking behind statements and theories. |
| | The lecturer seems well-prepared and knowledgeable on the subject. |
| Class Engagement | The lecturer effectively encourages students to ask questions and give answers |
| | The lecturer effectively directs and stimulates discussion. |
| Student Learning experience | The lecturer adjusts the pace of the class to the students' level of understanding |
| | The lecture is effective, overall, in helping me learn |
| | Class materials are adequate to learn, and the lecturer provides enough guidance to learn |
| | The lecturer stimulates my interest in the class material |
| Student's interest in the module | I enjoyed participating in this class |
| | In this class, my learning focuses on issues that interest me and is important for my professional practice. |
| | I enjoy participating in this class online. |

The online feedback was obtained during the 8th week of the semester, which allowed the students to have sufficient time to interact with the teacher, while there is enough time remaining in the semester to address the student comments/concerns.

The feedback survey was designed in two sections: the first section focused on quantitative evaluation using 14 rating questions while the second section included two open-ended questions to obtain qualitative feedback. The rating questions were focused on covering four aspects: effectiveness of teaching, student class engagement, student learning experience in the class, and student's interest in the module, as given in Table 1.

The survey was opened for the students to respond at the beginning of the week and a student could provide their feedback anytime during the same week. There was about a 70% student response rate for the survey, while the provided quantitative questions were analyzed using a grade point scale (ref). A corresponding weighted average grade point for each question was calculated, and the weighted average grade point for four aspects of teaching and learning was also calculated. To evaluate the effectiveness of the early semester feedback, the student marks for two group assessments with one conducted before and the other conducted after the feedback was compared. The second assessment was designed in the similar manner to the first, where students have to discuss among their groups to answer seven questions designed in the assessment. Furthermore, the end semester feedback of the current year as compared with the previous year to identify the effects of early semester feedback.

## IV. RESULTS AND DISCUSSION

Feedback survey results of the stage 1 depicted students' self-assessments on their learning and the suggestions for improving the teaching and learning process of the group assessment. Out of the 16 groups, two groups have opted not to submit the feedback while all the other groups have provided the feedback. The student response statistics for the three open-ended questions are given in Table 1 and the student suggestions are depicted in Fig. 1. In the second stage of the feedback survey, 122 students (almost 70% of the registered students) have responded and the results of the weighted averages grade points of each question are given in Table 1 and the weighted averages of the students' responses for different aspects of teaching and learning are depicted in Fig. 2.

The feedback collected for the group activity in 3rd week of the semester depicted in Table 2 indicates that students were keen on providing feedback on the group activity, where almost 85% of the student's groups provided their reflection on the activity. However, it was clear from the limited responses received for the last question (below 30%) that the students had very little idea on providing overall comments on the module. Also, the received limited responses were also not providing any meaningful information to improve the teaching and learning of the module. This suggested that the students did not have enough interaction in the class to provide overall feedback on their learning experience. Thus, it could be concluded that the time of obtaining the early semester feedback could affect the results of the survey and it should be implemented only after students had enough time to experience the module.

This was further evident from the good student responses received for the second feedback obtained during the 8th week of the semester. Almost 70% of the registered students have responded to this voluntary survey and out of which 90% of them have provided relevant comments to the qualitative questions. The feedback provided diagnostic information on the key changes to be adopted in teaching, that resulted in improved student engagement and performance. One of the key suggestions adapted was to conduct more tutorial sessions, since there were no timetable hours allocated as tutorials, as a first step one lecture was redesigned in tutorial mode to enhance the application of theoretical concepts of thermodynamics in real life scenarios. As can be seen from Table 2, almost 90% of the students responded that their interest was valued, and they felt inclusive in the class while 80% of students were of the impression that class materials are relevant to their professional practice. According to the results

depicted in Figure 2, it was clear that both teaching effectiveness and class engagement/inclusiveness have excellent grade points (9 out of 10), suggesting the students felt inclusive in the class and they have appreciated the teaching pedagogy. However, because this class was conducted completely online student learning has been a little bit hindered and students' interest has clearly declined. This was also evident from the comparatively low average grade point received for the last question: "I enjoyed participating this class online" in Table 3 as well as the student response to the qualitative questions, where 70% of the respondents highlighted the importance of having physical interaction in the classroom. Amid all these difficulties, it was clear from the student responses that obtaining their views has increased their inclusiveness in the classroom. Since the classroom activities were designed in discussion mode it improved the student interaction during the lecture time. This was evident in the second assessment as well as the participation in the second feedback survey. Furthermore, subsequent analyzis of the student grades for two group assessments: one conducted before the feedback survey and the other conducted after the survey, has shown that the average student grades have increased by 10%. This confirmed that the early semester feedback has improved the student performance in the class. Also, a comparison of end semester feedback received from the students for this module with the previous year showed that the student response rate for the end semester feedback has increased by 15% showing that the students have felt they were inclusive in the teaching & learning process.

Table 2. Student response statistics of the feedback on group learning activity

| Question | No responses | % |
|---|---|---|
| What do you most like about this activity? | 12 | 85.7 |
| What aspects would you like to change to make this activity more useful for your learning? | 9 | 64.3 |
| What suggestions would you give to improve the learning experience in this class? | 4 | 28.6 |

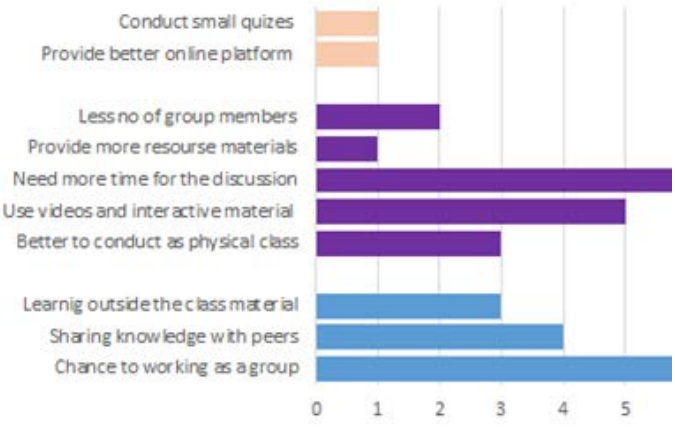



Fig. 1. Student responses for the feedback survey for the group activity

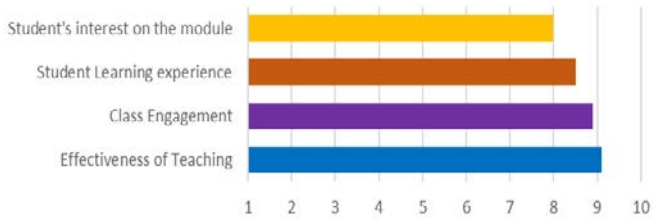



Fig. 2. Student response for different teaching and learning aspects

## V. CONCLUSION

This study was conducted to evaluate the effectiveness in This study clearly showed that obtaining early semester feedback and having continuous dialogue with students is important in Engineering education and it resulted in improved class interaction and student performance in the module assessments. Also, it is important to plan ahead on when to obtain the feedback and on what aspects the feedback is required, which ultimately impacts the effectiveness of the feedback process. using early semester feedback to improve the student class engagements in basic Engineering Modules. A teacher designed survey conducted as paper based and the online mode was used to collect the early semester feedback in two stages of the semester.

The collected feedback included both qualitative and quantitative information, which were analyzed to evaluate the student engagement during the semester. The results of the study showed that 90% of the respondents felt their interests were valued and the survey has provided diagnostic information on key changes to teaching which resulted in an improved learning experience in the class. This contributed towards increased grades in the subsequent assessments which proves early semester feedback positively affects student learning, Also, the results showed that correct timing of the feedback is important for success, where the feedback should be collected only after the students had enough time to interact with the teacher and experience the module.

Table 3. Student response details of the survey conducted in 8th week of the semester and grade point values

| Teaching & Learning aspect | Respondents = 122 | Never (2) | score | Rarely (4) | score | Occasionaly (6) | score | Frequetly (8) | score | Very Frequetly (10) | score | Total Marks | Wt.Avg (10 point-scale) | Wt. Avg (10 point scale) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Effectiveness of Teaching | The lecturer explains the material clearly. | 0 | 0 | 1 | 4 | 6 | 36 | 42 | 336 | 73 | 730 | 1106 | 9.1 | 9.1 |
| | The lecturer indicates important points to remember | 0 | 0 | 2 | 8 | 7 | 42 | 47 | 376 | 66 | 660 | 1086 | 8.9 | |
| | The lecturer shows genuine interest in students | 1 | 2 | 1 | 4 | 7 | 42 | 34 | 272 | 79 | 790 | 1110 | 9.1 | |
| | The lecturer explains thinking behind statements and theories. | 0 | 0 | 1 | 4 | 8 | 48 | 44 | 352 | 69 | 690 | 1094 | 9.0 | |
| | The lecturer seems well-prepared and knowledgeable on the subject. | 1 | 2 | 0 | 0 | 0 | 0 | 15 | 120 | 106 | 1060 | 1182 | 9.7 | |
| Class Engagement | Lecturer effectively encourages students to ask questions and give answers | 1 | 2 | 2 | 8 | 9 | 54 | 19 | 152 | 91 | 910 | 1126 | 9.2 | 8.9 |
| | The lecturer effectively directs and stimulates discussion. | 0 | 0 | 2 | 8 | 11 | 66 | 58 | 464 | 51 | 510 | 1048 | 8.6 | |
| Student Learning experience | The lecturer adjusts the pace of the class to the students' level of understanding | 0 | 0 | 3 | 12 | 16 | 96 | 50 | 400 | 52 | 520 | 1028 | 8.4 | 8.6 |
| | Class materials are adequate to learn and lecturer provide enough guidance to learn | 2 | 4 | 2 | 8 | 12 | 72 | 44 | 352 | 62 | 620 | 1056 | 8.7 | |
| | The Lecturer stimulates my interest in the class material | 1 | 2 | 2 | 8 | 18 | 108 | 52 | 416 | 49 | 490 | 1024 | 8.4 | |
| | The lecture is effective, overall, in helping me learn | 0 | 0 | 3 | 12 | 11 | 66 | 40 | 320 | 68 | 680 | 1078 | 8.8 | |
| Student's interest on the module | I enjoyed participating in this class | 1 | 2 | 3 | 12 | 22 | 132 | 51 | 408 | 45 | 450 | 1004 | 8.2 | 8.0 |
| | interest me and is important for my professional practice. | 0 | 0 | 3 | 12 | 19 | 114 | 60 | 480 | 40 | 400 | 1006 | 8.2 | |
| | I enjoy participating in this class online. | 7 | 14 | 8 | 32 | 29 | 174 | 40 | 320 | 38 | 380 | 920 | 7.5 | |

## REFERENCES

[1] K. C. Dee, "Student perceptions of high course workloads are not associated with poor student evaluations of instructor performance," *Journal of Engineering Education*, vol. 96, no. 1, pp. 69–78, 2007, doi: 10.1002/j.2168-9830.2007.tb00916.x.

H. Murzi, "Using continuous feedback as an alternative form of students' evaluation of teaching," in *Proceedings of the 8th Research in Engineering Education Symposium, REES 2019 - Making Connections*, 2019, pp. 160–169.

[2] P. A. Cohen, "Student Ratings of Instruction and Student Achievement: A Meta-analysis of Multisection Validity Studies," *Review of Educational Research*, vol. 51, no. 3, pp. 281–309, 1981, doi: 10.3102/00346543051003281.

[3] H. W. Marsh, "Students' evaluations of University teaching: Research findings, methodological issues, and directions for future research," *International Journal of Educational Research*, vol. 11, no. 3, pp. 253–388, 1987, doi: 10.1016/0883-0355(87)90001-2.

[4] Frick, T.W., Chadha, R., Watson, C. and Zlatkovska, E., "Improving course evaluations to improve instruction and complex learning in higher education," *Educational Technology Research and Development*, vol. 58, no. 2, pp. 115–136, 2010.

[5] R. Stark, P., Freishtat, "An Evaluation of Course Evaluations," *Science Open Research*, 2014, doi: https://doi.org/10.14293/S2199-1006.1.SOR-EDU.AOFRQA.v1.

[6] S. D. Brookfield, *Becoming a critically reflective teacher*. San Francisco: John Wiley & Sons, 1995.

[7] P. Gormally, C., Evans, M., & Brickman, "Feedback about teaching in higher ed: Neglected opportunities to promote change.," *CBE—Life Sciences Education*, vol. 13, no. 2, pp. 187–199, 2014.

[8] H. A. Diefes-Dux, "Student reflections on standards-based graded assignments," in *Proceedings - Frontiers in Education Conference, FIE*, 2016, vol. 2016-Novem. doi: 10.1109/FIE.2016.7757445.

[9] S. E. Hurney, C., Harris, N., Bates Prins, S. and Kruck, "The impact of a learner-centered, mid-semester course evaluation on students," *The Journal of Faculty Development*, vol. 28, no. 3, pp. 55–62, 2014.

[10] N. V. N. Chism, T. A. Angelo, and K. P. Cross, *Classroom Assessment Techniques: A Handbook for College Teachers.*, 2nd ed., vol. 66, no. 1. Jossey-Bass Publishers, 1995. doi: 10.2307/2943957.

[11] N. K. Mandal, "Importance of student feedback in improving mechanical engineering courses," *International Journal of Mechanical Engineering Education*, vol. 47, no. 3, pp. 227–245, 2019, doi: 10.1177/0306419018759800.

[12] T. H. Eng, A. F. Ibrahim, and N. E. Shamsuddin, "Students' Perception: Student Feedback Online (SuFO) in Higher Education," *Procedia - Social and Behavioral Sciences*, vol. 167, pp. 109–116, 2015, doi: 10.1016/j.sbspro.2014.12.651.

[13] B. Ravelli, "Anonymous Online Teaching Assessments: Preliminary Findings," in *American Association for Higher Education*, pp. 3–12. [Online]. Available: http://eric.ed.gov/?id=ED445069

[14] H. M. Anderson, J. Cain, and E. Bird, "Online student course evaluations: Review of literature and a pilot study," *American Journal of Pharmaceutical Education*, vol. 69, no. 1, pp. 34–43, 2005, doi: 10.5688/aj690105.

[15] M. Lovric, "Traditional and Web-Based Course Evaluations-Comparison of Their Response Rates and Efficiency," 2006.

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index No: TL-33-66*

# Regularization Risk Factors of Suicide in Sri Lanka for Machine Learning

D. M. A. U. Delpitiya (1st Author)
*Department of Physical Science,*
*Faculty of Applied Science,*
*University of Vavuniya*
Vavuniya, Sri Lanka
achinidel@gmail.com

Prabha M. Kumarage(2nd Author)
*Department of Software Engineering,*
*Faculty of Computing and Technology,*
*University of Kelaniya*
Dalugama, Sri Lanka
prabhami_2019@kln.ac.lk

B. Yogarajah(3rd Author)
*Department of Physical Science,*
*Faculty of Applied Science,*
*University of Vavuniya*
Vavuniya, Sri Lanka
yoganbala@yahoo.com

Nagulan Ratnarajah (4th Author)
*Department of Physical Science,*
*Faculty of Applied Science,*
*University of Vavuniya*
Vavuniya,Sri Lanka
rnagulan@univ.jfn.ac.lk

***Abstract*— Indication to World Health Organization, suicide is a major world public health concern that is in the top twenty leading causes of death worldwide. Sri Lanka is a country that has the highest suicidal rates in the globe. The comprehensive study about risk factors for suicide is important because we can prevent or treat the recognized most risky categories of people. The emergence of big data concepts with machine learning techniques introduced a resurgence in predicting models using risk factors for suicide. Regularization is one of the most decisive components in the statistical machine learning process and this technique is used to reduce the error on the training dataset and prevent over-fitting. Comprehensive regularization approaches are presented here to select significant risk factors for age-specific suicide in Sri Lanka and build unique predictive models. The Least Absolute Shrinkage and Selection Operator (LASSO) approach presents regularization along with the feature selection to improve the prediction precision. The dataset collected for the study is rooted in the Sri Lankan people and the factors used for the analysis are, suicide person's gender, lived place, education level, mode of suicide, job, reason, suicide time, previous attempts, and marital status. Further, the riskiest age category of the people, who has exposure to suicide, is identified. Multiple linear regression and Ridge regression were used to evaluate the performance of LASSO. The selected most relevant factors with regularization to predict age-specific suicide prove the effectiveness of the proposed regularization approaches.**

***Keywords* - *Suicide, LASSO, Ridge, Machine Learning***

## I. INTRODUCTION

Suicide means someone ending their own life and it is a way for people to escape pain or suffering. Mental disorders, including depression, bipolar disorder, autism, schizophrenia, personality disorders, anxiety disorders, substance abuse including alcoholism and the use of benzodiazepines, stress, such as from financial difficulties, relationship problems such as breakups, or bullying and there are many more reasons for suicide. According to the World Health Organization (WHO) statistics, suicide is a major world public health concern [1]. It is among the top twenty leading causes of death all around the world, with more deaths due to suicide than to malaria, breast cancer, war, and homicide. There are close to 800 000 people who die by suicide every year [1]. Today, Sri Lanka has recorded one of the highest rates of suicide within the world.

Suicide prevention in developing countries is an essential social and public health objective. For this protection, the most significant risk factors should be identified and build a predictive model for counsellors, and health and human service workers. Many studies [2]-[5] on risk factors for suicide were published in developing countries based on socio-demographic, clinical, and environmental/situational factors. Studies of Sri Lankan suicidal data-based risk factors analysis were presented by the research community, such as studies on risk factors for acute pesticide poisoning [6] and demographic risk factors in pesticide-related suicides [7]. Jeanne Marecek [8] has presented the social ecology of young women's suicide in Sri Lanka and discussed the implications of the cultural, ecological, and psychological factors. Lakmali et al [9] identify the major factors affecting suicide in Sri Lanka using Quasi Poisson and negative binomial regression models. However, these traditional statistical approaches to the prediction of risk factors for suicide attempts have limited accuracy and scale of risk detection [10].

Recently, the advancement of computer power, the large amount of data, and more importantly development of more advanced machine learning algorithms such as deep learning [11], are utilized to analyze any complete data and predict models. Machine learning-based techniques, such as Support Vector Machines (SVM), Artificial Neural Networks (ANN) and deep learning-based algorithms such as Recurrent Neural Network (RNN), provide systems with the ability to self-learn and improve from practice. This process of learning begins with observations or data, such as examples, direct experience, or instruction and considers patterns of those data. Then it can make better decisions eventually based on the examples that we have given. The main desire of this concept is to allow computers to self-learning without human aid and adjust actions appropriately. Supervised learning is one of the machine learning tasks, which is a learning model, built to make predictions, given an unforeseen input instance. Supervised learning algorithms take a well-known input dataset and its known responses to the data of output to learn the regression or classification model. These learning algorithms then train a model to generate a prediction for the response to a new dataset. Supervised learning uses classification algorithms and regression techniques to develop predictive models.

The goal of machine learning is to model the pattern and ignore the noise. A machine learning algorithm is trying to fit the noise in addition to the pattern, it is called over-fitting. The model gets a low accuracy if it is over-fitted. Thus, the avoiding over-fitting issues are one of the main phases of training a machine-learning model. Over-fitting occurs when the model is tracking heavily to capture the noise in the training dataset. Most commonly, cross-validation is used for avoiding over-fitting that guides in examining the error over a dataset, and in determining which variables are working best for the model. Regularization is a form of regression, which regularizes or shrinks the coefficient estimates close to zero. Particularly, this approach discourages learning a more complex or flexible model to avoid the risk of over-fitting. LASSO is a regression analysis method that performs both variable selection and regularization to enhance the prediction accuracy and interpretability of the statistical model [10]. Some of the studies in the literature focused on predicting the risk of suicide attempts through machine learning without using proper regularization approaches [12]-[14].

## II. OBJECTIVES

This research paper aims to detect the significant variables, the risk factors for suicide, to create a unique predictive model for each age group to use in the machine learning purpose. eal world dataset of suicides collected from the Sri Lankan general hospitals is used to detect the risk factors for suicide according to age. The LASSO model is used to apply the suicide dataset for variable selection, regularization and prediction. In addition, the riskiest age group for the suicide attempt is identified in this study. The results were evaluated using a comparison with multiple regression analysis and Ridge regression with LASSO. The LASSO model provides excellent prediction precision, guides to extend the model intelligibility by removing negligible factors which do not relate to the responding variable and reduces over-fitting. The initial and important part of the machine learning process is accomplished here to choose the risk factors and build prediction models.

## III. METHODOLOGY

### *A. Suicide Dataset*

Mainly the suicide dataset was collected from the Forensic Medical Office, District General Hospital-Nawalapitiya in February 2020. We have conducted an interview with the forensic medical officer and gathered the dataset and got approval to use the dataset for research purposes. Furthermore, we have collected more data from Teaching Hospital-Peradeniya by an interview with the forensic medical officer of the hospital. The dataset consists of 120 data that is containing the details of persons, who have committed suicide from February 2016 to February 2020 in the Nawalapitiya and Kandy districts.

Originally, the dataset contains 14 risk factors of individual information. Data pre-processing was made to ensure that the collected data is correct, consistent, and usable by identifying any errors in the data. The 10 risk factors were extracted based on the pre-processing step (eliminating incomplete, noisy and inconsistent data) and the suggestion from the medical officer, which make machine learning algorithms work well, better represent the underlying problem to the predictive models and improve the accuracy of the model. It contains the data of the suicide person's age, gender, lived place, education level, mode of suicide, job, reason, suicided time, marital status, and the details about if there are previous attempts to suicide. By using the responding variable as age categories, we have done our analysis to get the most risk factors of suicide for using in the machine learning approach and to identify the riskiest age category for committing suicide. Here, we have decided to gather all the suicide persons' data of all age ranges.

### *B. Variables Selection Using Regression Models*

The variable (factor) selection method is used to select the most significant variables from the above exposure variables to predict the responding variable. Thus, we have to use the multiple linear regression, Ridge, and LASSO methods with making use of the R software [15]. However, in the first step, we have to analyze the dataset in a better understandable way.

*1) Categorization of Dataset:* The data was properly organized in the first phase using Microsoft Excel 2016 as a perfect dataset and it was categorized as the responding variable and the exposure variables. Then correlations between the exposure variables were computed.

*2) Implementation of Multiple Linear Regression Model*: At the second phase, a multiple linear regression model was built from the training dataset and an ANOVA table was revealed. This model is the best-illustrated model that consists of the highly correlated independent variables [16].

By using this model, we can inspect the exposure variables which have high significance and low significance for the model. We can observe the most significant exposure variables for use in the model from the ANOVA table. Moreover, by using the ANOVA table, we can select the factors for the finalized model. To use this regression model, the responding variable should be normally distributed, should be a linear relationship among the responding variable and the exposure variables, and those variables should not be correlated, mean should be zero and variance should be constant.

In addition, to check the multicollinearity of the training dataset, we calculated the Variance Inflation Factors (VIF) values. Examining these values, we can decide the multicollinearity exists or not in the model.

*3) Implementation of Ridge and LASSO Models:* The Ridge and LASSO regression models were used to avoid confusion and determine the factors, which are contained in the finalized model because those methods can be used for efficient variable selection. In the Ridge and LASSO regressions, correlated exposure variables were used and there were no assumptions to use with these models. For the Ridge and LASSO, implementations have been done by the following procedures.

First, two different models were obtained for the Ridge and LASSO regression models with different λ values. Here the λ is the tuning parameter and this λ values control the strength of these models with great importance. The most affecting factor, negatively and positively affected factors and the low significant factors were obtained by the number of graphs from those Ridge and LASSO models. After that, by doing cross-validation for the models, we got the most convenient λ values. Then we selected the minimum λ values and according to that λ values obtained the most significant exposure variables to use in the finalized models of Ridge and LASSO. The most appropriate models were constructed to predict the suicide causes using the most significant factors selected by the Ridge and LASSO models.

We computed the adjacent R-squared for all the variables and therefore the most significant factors independently for examining the effectiveness of the chosen factors. We researched the medical and psychological background for each significant variable to justify the correctness. Furthermore, we checked the correctness of the models by testing new data.

Finally, we compared all the built models by use of the Residual Sum of Squares (RSS). Initially, we approximately calculated the λ values based on the training dataset by double cross-validation methods using the best λ values. Then we applied these estimated λ values in the test dataset. In addition, coefficients were estimated, and the residual sum of squares (RSS) was calculated for each model.

## IV. RESULTS AND DISCUSSION

### *A. Variables Selection Using Regression Models*

The 10 risk factors were extracted from the raw suicide dataset for 120 committed suicide persons. The responding variable of this study is Age, and the Age categorized into 4 groups, 1-15 years, 16-30 years, 31-50 years and 51-70 years. The selected nine exposure variables are,

1. *Gn* : Gender of the suicided person
2. *Lp* : Division where he/she lives
3. *El* : Education level
4. *Mod* : Which type of the suicide
5. *Jb* : Occupation
6. *Rsn* : Reason of suicide
7. *Time* : Time of suicide
8. *PrAt* : Number of previous attempts to suicide
9. *Ms* : Marital status (single/married/divorced)

We have selected 110 samples among the suicide dataset as the training data and to assess the correctness of the regression model, the last 10 samples remained as testing data.

*1) Categorization of the Suicide Dataset:*

Table 1 illustrates the correlations among all response and exposure variables. It shows that correlations between variables are powerful. Some factors are positively dependent on other factors and some factors are negatively correlated with others. For illustration, the responding variable *Age* shows high dependency with the exposure variables, *Jb*, and *Rsn*. In addition, *Age* negatively depends on the variables *Gn*, *Time*, and *Ms* and positively depends on all other variables. Therefore, according to the Table 1, we can examine, the higher number of factors in this dataset are multi-correlated covariance variables.

Table 1. Correlations "Suicide Dataset"

| | Age | Gn | Lp | El | Mod | Jb | Rsn | Time | PrAt | Ms |
|---|---|---|---|---|---|---|---|---|---|---|
| **Age** | 1 | | | | | | | | | |
| **Gn** | -0.129 | 1 | | | | | | | | |
| **Lp** | 0.183 | -0.048 | 1 | | | | | | | |
| **El** | 0.164 | 0.056 | -0.001 | 1 | | | | | | |
| **Mod** | 0.201 | 0.076 | -0.015 | -0.010 | 1 | | | | | |
| **Jb** | 0.589 | 0.020 | -0.087 | 0.052 | -0.119 | 1 | | | | |
| **Rsn** | 0.052 | 0.079 | 0.047 | 0.218 | -0.024 | -0.042 | 1 | | | |
| **Time** | -0.102 | 0.004 | -0.075 | -0.036 | -0.065 | -0.030 | 0.213 | 1 | | |
| **PrAt** | 0.107 | -0.130 | -0.132 | 0.112 | -0.080 | 0.217 | 0.072 | -0.125 | 1 | |
| **Ms** | -0.341 | 0.005 | -0.301 | -0.028 | -0.005 | -0.246 | 0.091 | -0.099 | 0.134 | 1 |

*2) Implementation of Multiple Linear Regression Model:* We built an ANOVA table (Table 2) for examining which are the most significant for the model and which are less significant exposure variables for the model.

Analyzing Table 2, the results show that, according to the p-values, *Job* and *Mod* have high significance. Therefore, we can predict these exposure variables will be involved in the finalized model for prediction. We cannot actually wrap up these factors are contained in the finalized model due that it is not guaranteed whether those variables are optimal or not. There is a feasibility for the not optimal, because the variables *PrAt*, *Time*, and *Rsn* are not significant factors and should not be included in the finalized model

Table 2. ANOVA "Suicide Dataset"

| **Responding Variable: *Age*** | | | | | |
|---|---|---|---|---|---|
| | **Df** | **Sum Sq** | **Mean Sq** | **F Value** | **Pr (>F)** |
| **Gn** | 1 | 1.718 | 1.718 | 3.6323 | 0.059541 |
| **Lp** | 1 | 3. 254 | 3.254 | 6.8785 | 0.010088 |
| **El** | 1 | 3. 032 | 3.032 | 6.4094 | 0.012909 |
| **Mod** | 1 | 4. 814 | 4.814 | 10.1779 | 0.001899 |
| **Jb** | 1 | 40.654 | 40.654 | 85.9428 | 3.932e-15 |
| **Rsn** | 1 | 0.328 | 0.328 | 0.6928 | 0.407182 |
| **Time** | 1 | 0.313 | 0.313 | 0.6621 | 0.417749 |
| **PrAt** | 1 | 0.054 | 0.054 | 0.1152 | 0.735060 |
| **Ms** | 1 | 1.884 | 1.884 | 3.9826 | 0.048692 |
| **Residuals** | 100 | 47.303 | 0.473 | | |

Table 3 illustrates the computed Variance Inflation Factors (VIF) to check the multicollinearity of the training dataset. By examining Table 3, we can see some of these values are rather than high such as *VIFMs*, *VIFJob* and *VIFLp*. According to the table, we can conclude multicollinearity exists in this model.

Table 3. VIF values for "Suicide Dataset"

| **Variable** | **VIF** |
|---|---|
| Gn | 1.040016 |
| Lp | 1.182521 |
| El | 1.075680 |
| Mod | 1.032261 |

| | |
|---|---|
| Jb | 1.190100 |
| Rsn | 1.155741 |
| Time | 1.125189 |
| PrAt | 1.157541 |
| Ms | 1.282413 |

*3) Ridge Regression:* For obtaining the most significant exposure variables, we used the Ridge regression method with its functions to predict the responding variable *Age* by using all the exposure variables in the suicide dataset.

Fig. 1 shows a graphical representation of different values of log λ versus coefficients of each exposure variable. Each color line illustrates the exposure variables and their expansion in the model. From this graph, we can identify the entered position of each exposure variable to the model and the influenced scope of the responding variable.

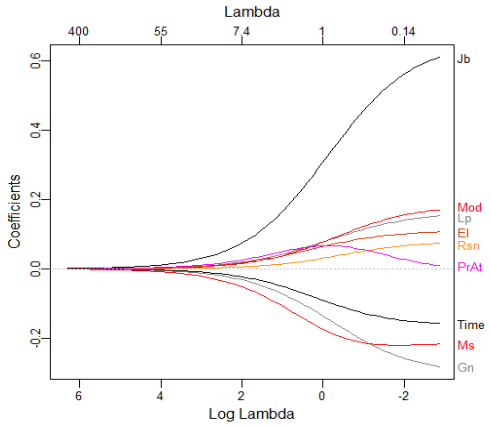



Fig. 1. Glmnet graph for exposure variables - Ridge regression

Fig. 1 reveals that the most influenced exposure variable is *Jb* in the model, it steadily and positively affected the responding variable *Age*. *Ms* is the second important variable and that negatively affected the responding variable *Age* and entered the model lately, but it also affected the trend of *Jb* after entered in the model. *Gn is* also one of the most significant exposure variables because it also enters the model later and it affects the trend of *Ms*. Furthermore, we can select *Mod* and *Lp* as other important factors by examining their trends. All other variables in the model can be discarded from the finalized model because those are less significant.

Fig. 2 shows the plot that the green lines illustrate more significant factors that negatively influence the responding variable, while red lines illustrate the ones that influence it positively.

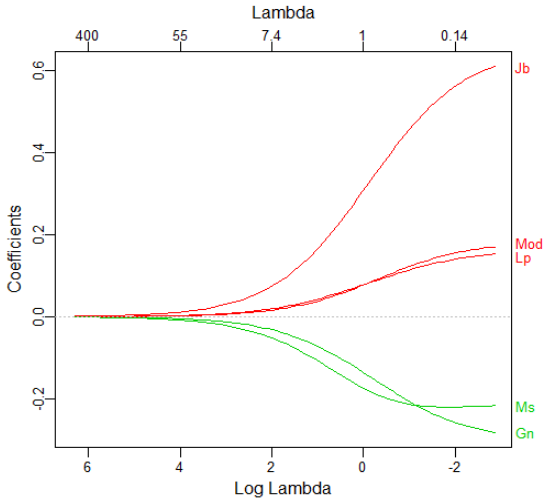



Fig. 2. Glmnet graph for significant variables - Ridge regression

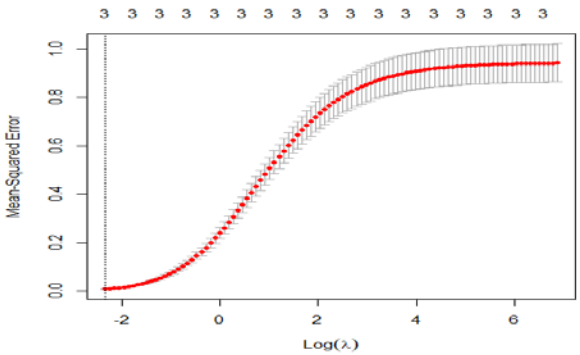



Fig. 3. Cross-validation – Ridge regression (nfold = 5)

As the next step, we used cross-validation to find the most suitable value for the λ. It is useful for controlling the strength of the penalty. For that, we considered $\lambda_{min}$ that gives minimum mean cross-validated error and $\lambda_{1se}$, the standard error of the minimum. In this study, these two values are the same. Therefore, both two values are represented by the vertical dotted line in Fig. 2. Moreover, we can select the value for tuning parameter λ that better fits to the situation. In this analysis, $\lambda_{min}$ is not noticeable due to that Fig. 3 plot illustrates an exponential trend. According to the plot, we can select factors as likely, the most significant variables from both the values of λ from the variables shown in Fig. 2. Thus, we have for $\lambda_{min}$ and $\lambda_{1se}$ as 0.09693229 and the most significant variables that we can contain in the finalized model are *Jb*, *Lp*, *Mod*, *Ms*, and *Gn*.

*4) LASSO Regression*: We obtained the most significant factors which are contained in the finalized model using LASSO regression.

Fig. 4 depicts the graphical representation of the different values of log λ versus coefficients of each exposure variable. Colored lines illustrate the exposure variables and their expansion into the model like Ridge regression.

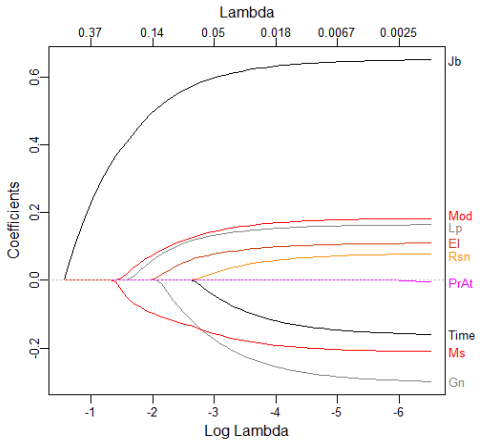



Fig. 4. Glmnet graph for exposure variables - LASSO

By examining Fig. 4, the most influenced exposure variable is revealed, which is *Jb*, and it steadily and positively affected the responding variable *Age*. We decided *Ms*, *Gn*, *Mod*, and *Lp* as the other important variables as in Ridge regression by looking at their trends and all the other variables in the model discarded from the finalized model because those are less significant.

Fig. 5 plot shows that the green lines determine more significant attributes that negatively influence the responding variable, while red lines determine the ones that influence it positively. Using the cross-validation into the LASSO, we found the most suitable value for the λ, like the Ridge regression. The $\lambda_{min}$, minimum mean cross-validated error and $\lambda_{1se}$, the standard error of the minimum is the same for the obtained LASSO values. Therefore, we represented both two values as a vertical dotted line in Fig. 5. In addition, we selected the value for the tuning parameter λ for the better fit problem.

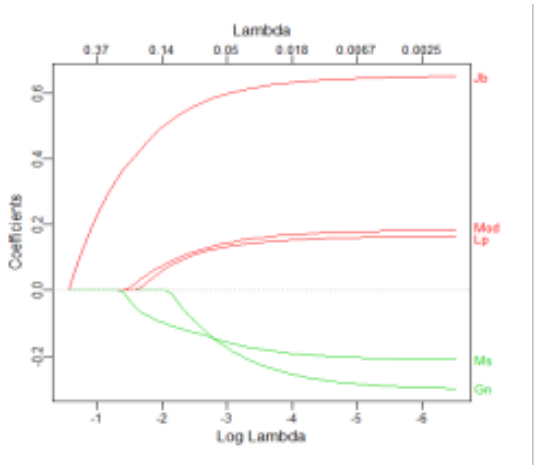



Fig. 5. Glmnet graph for most significant variables – LASSO

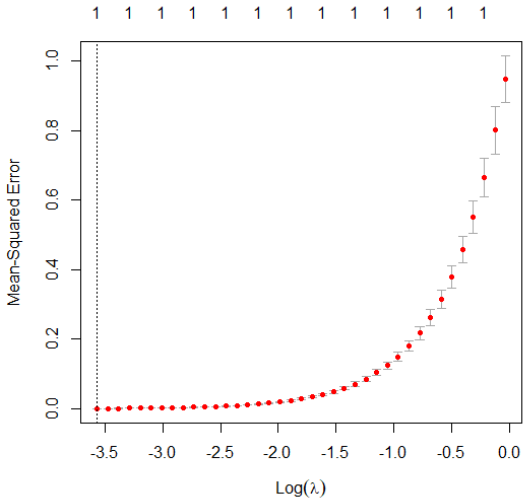



Fig. 6. Cross-validation (nfold=5)

In this analysis, $\lambda_{min}$ is not visible because Fig. 6 plot illustrates an exponential trend. Moreover, the factors were selected as likely, which are the most significant exposure variables according to both the values of λ. From that, we have the value for $\lambda_{min}$ and $\lambda_{1se}$ as 0.02825628. Thus, the most significant factors included in the finalized model are *Jb*, *Lp*, *Mod*, *Ms*, and *Gn* and those variables were selected according to the optimized $\lambda_{min}$ and $\lambda_{1se}$ values.

### *B. Adjusted R Squared*

The calculated adjusted R squared value is 0.5011 for multiple linear regression with all the nine exposure variables. After the selection of the most significant exposure variables *Jb*, *Lp*, *Mod*, *Gn*, and *Ms*, the adjusted R squared value was obtained as 0.4883. Thus, we proved that the factors we have selected using the Ridge and LASSO for the finalized model are highly significant risk factors for the responding variable *Age*.

### C. Comparing coefficients and RSS

The linear, Ridge, and LASSO models were compared by using the Residual Sum of Squares (RSS). Initially, we approximated the λ v8alues with respect to the training dataset by using two cross-validation methods of $\lambda_{min}$ and $\lambda_{1se}$. After that, we applied these estimated λ values for training and testing datasets separately to calculate the estimated coefficients and residual sum of squares for each model.

For those two cross-validation methods, we have got the same values as for LASSO the $\lambda_{min}$ as 0.02825628 and $\lambda_{1se}$ as 0.02825628 and for Ridge Regression the $\lambda_{min}$ as 0.09693229 and $\lambda_{1se}$ as 0.09693229. Therefore, we selected $\lambda_{min}$ from both to estimate the coefficients and RSS.

According to the results, we proved that the coefficient estimates get closer to zero. In LASSO, some of the non-zero coefficients changed to zero. Then, we applied the estimated λ values for the testing dataset to get new estimated coefficient values. From those results, most of the coefficient values in the linear model changed as negative values and most values are close to zero. And also in LASSO, the number of zero coefficient values in the test dataset is greater than the number of coefficient values in the training dataset.

Table 4. RSS coefficient values for train and test data

| RSS | Linear model | Ridge$\lambda_{best}$ | Lasso$\lambda_{best}$ |
|---|---|---|---|
| $RSS_{TRAIN}$ | 47.34059 | 47.8663 | 48.08071 |
| $RSS_{TEST}$ | 148.577 | 138.3825 | 138.4715 |

The RSS values for the training and testing datasets are shown in Table 4. According to that, the linear regression model predictably achieves the smallest RSS value on the training dataset. The RSS for Ridge and LASSO are again typically greater when λ is chosen using the "one-standard error" rule. Interestingly, the errors on the test dataset are ordered unexpectedly; the minimum is achieved by Ridge at λ = ridge.bestlam, and the second is LASSO at λ = lasso.bestlam while linear regression takes just the third position. Here, Ridge and LASSO are acting better than linear regression on the testing dataset.

### D. Most Risky Age Group

As a sub-finding from the dataset, we have selected the most vulnerable age group for suicide. In this study, we used four categories of age groups, 1-15 years, 16-30 years, 31-50 years, and 51-70 years. For the test dataset, we have randomly selected sets of two exposure variables from our finalized model variables and used the Analysis of Covariance (ANCOVA) method for analyzing.

Table 5. ANCOVA for age groups

| Responding variable is Gn | | | | |
|---|---|---|---|---|
| | **Estimate** | **Std. Error** | **t value** | **Pr(>\|t\|)** |
| **(Intercept)** | 1.380912 | 0.121496 | 11.366 | <2e-16 |
| **Lp** | -0.003278 | 0.041676 | -0.079 | 0.9374 |
| **Age1** | 0.091522 | 0.069983 | 1.308 | 0.1936 |
| **Age2** | -0.146364 | 0.064107 | -2.283 | 0.0243 |
| **Age3** | 0.087537 | 0.082822 | 1.057 | 0.2928 |

Table 5 shows the ANCOVA table for the variables *Gn* and *Lp*. We can see that group *Age2* (16-30 years) is the significant age group to attempt suicide because the *p*-value of *Age2* is less than a significant value of 0.05.

According to the ANCOVA values, the other sets of random significant variables also provided the most significant age category as *Age2*.

## V. CONCLUSION

In this study, comprehensive regularization approaches have been presented to select significant risk factors for suicide in Sri Lanka and an important part of the machine learning process was accomplished. Ridge and LASSO regressions for the machine learning approach are explained with the regularization task. To demonstrate the regularization with those two approaches, we utilized the Sri Lankan real time suicide dataset. We have concluded that the most related factors with regularization to predict the suicide ages by Ridge and LASSO are suicided person's gender, lived place, mode of suicide, occupation, and marital status. We proved the effectiveness of these most significant factors by computing the adjusted R-squared values. In addition, use of the coefficients of each and the residuals sum of squares (RSS), we have proven that Ridge is most suitable, and then LASSO for reach for the regularization. We have obtained that the riskiest age category for committing suicides is the young adults age between 16 - 30 years. The results of this study showed that Ridge and LASSO can use to build a unique model for predicting the suicide ages more accurately in machine learning approaches.

This study provides not only a platform for the machine learning approaches to further investigation of risk factors of suicide in Sri Lanka but also it will help for the activities of suicide prevention in Sri Lanka. Moreover, in the medical field researchers, counsellors, and doctors give advice or treatment to the most risk category of people.

## REFERENCES

[1] World Health Organization, *Suicide in the World: Global Health Estimates*, 2019.
[2] Lakshmi Vijayakumar Sujit John, Jane Pirkis, Harvey Whiteford Suicide in Developing Countries (2): Risk Factors, Crisis. 2005; 26(3):112-9.
[3] Lakshmi Vijayakumar1 Suicide prevention: the urgent need in developing countries, World Psychiatry. 2004 Oct; 3(3): 158–159.
[4] Murad M Khan Suicide Prevention and Developing Countries, Journal of the Royal Society of Medicine, Volume: 98 issue: 10, page(s): 459-463, 2005.
[5] Michael RPhillips Gonghuan Yang Yanping Zhang et al., Risk factors for suicide in China: a national case-control psychological autopsy study, The Lancet, Volume 360, Issue 9347, 30 November 2002, Pages 1728-1736.
[6] Wim Van Der Hoek, Flemming Konradsen Risk factors for acute pesticide poisoning in Sri Lanka, Tropical medicine and international health, Volume10,issue 6 June 2005, Pages 589-596
[7] E B R Desapriya, P Joshi, G Han, F Rajabali Demographic risk factors in pesticide related suicides in Sri Lanka, Injury prevention, Volume 10, Issue 2, 2004.
[8] Jeanne Marecek Young Women's Suicide in Sri Lanka: Cultural, Ecological, and Psychological Factors, Asian Journal of Counselling, 2006, Vol. 13 No. 1, 63–92

[9] Lakmali, S. M. M., & Nawarathna, L. S. (2019). Identifying and Predicting Major Factors Affecting the Suicides in Sri Lanka. Asian Journal of Probability and Statistics, 2(3), 1-7. https://doi.org/10.9734/ajpas/2018/v2i328785

[10] V. Fonti., "Feature Selection Using Lasso", Vrije Universiteit Amsterdam,2017

[11] M. M. Najafabadi1, F. Villanustre, T. M. Khoshgoftaar1, N. Seliya1, R. Wald1 and E. Muharemagic, "Deep learning applications and challenges in big data analytics", Najafabadi et al. Journal of Big Data, 2:1, DOI 10.1186/s40537-014-0007-7, 2015.

[12] L.E. Melkumova, S.Ya. Shatskikh," Comparin Ridge and Lasso Estimators for Data Analysis", Elsevier 2017: P. 746-755.

[13] Gregory D. S. "Regularization Techniques for Multicollinierity: Lasso, Ridge, and Elastic Nets," Henry M Jackson Foundation, 2018.

[14] (2018) Linear and Ridge Regression with R, [Online]. Available: https://www.pluralsight.com/

[15] Imran Amin, "Prediction of Suicide Causes in India Using MachineLearning", Shaheed Zulfikar Ali Bhutto Institute of Science and Technology, 2017.

[16] Prabha M. Kumarage, B. Yogarajah, and Nagulan Ratnarajah. "Efficient Feature Selection for Prediction of Diabetic Using LASSO", *19th International Conference on Advances in ICT for Emerging Regions (ICTer)*, 2019..

[17] (2016)Adolescent suicide behaviours in 32 low- and middle – income countries homepage on World Health Organization. [Online]. Available: https://www.who.int/bulletin/volumes/94/5/15-163295/en/

[18] Jana Vorlickova, "Least Absolute Shrinkage and Selection Operator Method", Faculty of Social Science, Charles University, 2017.

[19] Thevaraja, Mayooran & Gabriel, Mathew, "Comparing Linear, Ridge and LASSO Regressions". 10.13140/RG.2.2.30269.77282, 2018.

[20] Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani," An Introduction to Statistical Learning with Applications in R", 2013.

[21] Comparing Multiple Means in R homepage on DATANOVIA. [Online]. Available:https://www.datanovia.com/en/lessons/ancovain .

[22] Tibshirani, R., Regression Shrinkage and Selection via the Lasso. J. R. Stat. Soc. Ser. B (Methodological), 58, 267-288, 1996.

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index No: TL-34-79*

# UML System Model to Implement Authentic Learning in the 21st Century

Manuja Wickramasinghe (1st Author)
*Department of Software Engineering*
*(Faculty of Computing and Technology)*
*University of Kealniya*
Kelaniya,Srilanka
manujaw_2019@kln.ac.lk

Gamini Wijayarathna (2nd Author)
*Department of Software Engineering*
*(Faculty of Computing and Technology)*
*University of Kealniya*
Kelaniya,Srilanka
gamini@kln.ac.lk

Kasun Ferenando (3rd Author)
Department of Applied Computing
*(Faculty of Computing and Technology)*
*University of Kealniya*
Kelaniya,Srilanka
kasunf @kln.ac.lk

***Abstract* — The Pandemic situation demanded all universities to transform towards online teaching and learning as an alternative to face-to-face study through Technology Enabled Learning (TEL). In most developing countries, universities use Learning Management Systems (LMS) such as Moodle and Blackboard to facilitate online education. The LMS is used mainly to facilitate staff, administration, and students sharing module outlines, specifications, lecture schedules, lesson plans, assignment generation and submission, announcements, and generating assessment reports. However, it has been observed that many inefficiencies exist in the teaching and learning approaches as there is less focus on learner autonomy. Dynamic changes in the world and drastic transformations demand continuous learning and innovative thinking. Authentic learning is an approach that focuses on real-world situations to gain new knowledge and skills in a context rather than listening to lectures and memorizing information. In authentication learning, enabling collaborative learning has been observed as a practical approach to developing critical thinking, reflective thinking, and enhancing creativity. Sri Lankan culture is inherited with authentic learning as our cultural events, traditions, and values encourage living in harmony and learning from each other. This paper proposes a new system as a web portal to ensure that the learner can effectively gain the required knowledge, skills, and attitudes to face complex real-world situations, thus arriving at practical solutions to overcome contemporary issues. The proposed system focuses particularly on distance learning programs, as those could be advanced by the adoption of a model that can be used to guide the design of online learning environments focused on elements of authentic learning. In addition to presenting the authentic task model and its theoretical functionalities, the authors have implanted Bloom's Taxonomy Theory to ensure the quality and effectiveness of the system model. Researchers have incorporated use case diagrams and activity diagrams to exemplify the UML model.**

***Keywords- Authentic Learning, Online Education, UML***

## I. INTRODUCTION

Education is the process of imparting knowledge, values, skills and attitudes, which can be beneficial to an individual, and learning is the process of adopting knowledge, values and skills. Sustaining a nation's development is inextricably linked to the country's level of education.
Thus, the government's focus on educating their citizens is a key focus area as competing in the global workforce is a key concern. The world needs people with the skills to grow in the 21st century, and employers look for potential innovativeness and creativity as key criteria. Combined with technologies that continue to develop rapidly, the result is a world that needs continuous learning and development.

Thus, this education and learning are eventually considered one of society's fundamental demands. Technological advancements are a must to facilitate an authentic learning approach for distance learning programs.

Creativity, critical thinking, collaboration, and communication are the 4Cs that represent the skills in demand in the 21st century. They are largely excluded from our educational system and thus do not appear on most examinations. Sri Lanka ranked 106 out of 139 countries in the world for creativity.

The World Exports of Creative Goods in 2015 totaled the massive US $ 509 billion. According to Global Creativity, Sri Lanka could only capture 0.04 per cent of this revenue.
Index [2]

The learning environment should compromise high-speed internet connection and bandwidth, provision of multimedia information, asynchronous and synchronous communication and social networking tools to support teamwork, intelligent tutoring systems, virtual laboratories, feedback mechanisms, and mobile devices for accessing and inputting data during field-based investigations.

Research on distance education has been conducted for many years and continues today to investigate the use of online technologies to enhance and support learner activity and engagement.

The proposed system will contribute towards overcoming the existing inefficiencies in the traditional paper-based evaluation systems and will focus more on learning through real-life scenarios. One key to success appears to lie in the design of learning environments that effectively use the communications capabilities of technologies that can connect learners in meaningful ways [3].

## II. OBJECTIVES

The main goal of this software is to implement authentic learning, and it was designed in a way that a student self-evaluates. The student's profile will indicate that learning achievement level aligns with bloom's taxonomy framework. The software focuses on incorporating game-based learning to evaluate the learner where the levels will be defined based on bloom's taxonomy framework, remember, understand, apply, analyse, evaluate and create. Artificial intelligence will enable game-based learning and help adsorb complex phenomena in an exciting mechanism, ensuring the quality of education the goal through a web-based software application.

## III. IMPLEMENTATION OF AUTHENTIC LEARNING ELEMENTS IN THE SYSTEM

The researchers have implemented the authentic learning elements to map the tools in the pre-system as a new model. The following table (Table 1-Learning design mapped against the elements of authentic learning/components) illustrates how the elements of authentic learning have been addressed in the learning design of living and learning with technology, highlighting the characteristics of authentic activities.

The proposed system is designed based on the systematic investigations of authentic learning by examining successful, in-depth online authentic learning practices, including in-class complex tasks and evaluation approaches.

It was identified that the complexity of the task led to bringing out innovative solutions and their application to real-life situations. Exploring working examples and critical [1] analysis led to effective experiences for both the learner and the facilitator. The following table illustrates how the elements of authentic learning have been addressed in the learning design of Living and learning with technology. Table 1 shows the learning design mapped against the elements of authentic learning as components of the system.

The researchers have implemented the authentic learning elements to map the tools in the pre-system as a new model. The following table (Table 1-Learning design mapped against the elements of authentic learning/components) illustrates how the elements of authentic learning have been addressed in the learning design of living and learning with technology, highlighting the characteristics of authentic activities.

The proposed system is designed based on the systematic investigations of authentic learning by examining successful, in-depth online authentic learning practices, including in-class complex tasks and evaluation approaches.

It was identified that the complexity of the task led to bringing out innovative solutions and their application to real-life situations. Exploring working examples and critical [1] analysis led to effective experiences for both the learner and the facilitator. The following table illustrates how the elements of authentic learning have been addressed in the learning design of Living and learning with technology. Table 1 shows the learning design mapped against the elements of authentic learning as components of the system.

Table 1. Learning design mapped against the nine elements of authentic learning.

| **Authentic Learning Elements** | **Characteristics of Authentic Activities** |
|---|---|
| Authentic Context | • A design to preserve the complexity of real-world settings.<br>• It gives learnings a sense of purpose and motivation.<br>• In the context of a real-world situation, ideas can be explored in depth. |
| Authentic Activates | • It is defined by clear goals and relevance to real-world situations.<br>• By authorizing knowledge production rather than preproduction, it is necessary to prioritize knowledge production over preproduction.<br>• Students must decide how they will make decisions about how they decide they will complete the task. |
| Access to Expert Performances | • Provide students with access to expert thinking and processes.<br>• Students should see how an expert think and behaves and how they move around in the environment and what they do when performing the task.<br>• Sharing of narrative stories. |
| Multiple Perspectives | • Students who are capable of utilizing and exploring issues from various perspectives and capable of enhancing a variety of learning resources and materials, not just a single textbook. |
| Collaboration | • Provide joint problem solving and social support.<br>• The learning environment can provide it by arranging students into teams or pairs rather than as individuals.<br>• Encourage through technology task address to groups rather than to an individual. |
| Reflection | • The learning environment provides an opportunity for the students to think about, reflect and discuss choices.<br>• Students compare their thoughts and ideas to express, teachers, students and other students. |

| | |
|---|---|
| Articulation | • The learning environment provides opportunities for students to express themselves verbally and visually about their gowns and understanding.<br>• This allows students to present their work in front of a public, defending skill sets and ideas. |
| Coaching and Scaffolding | • In an authentic learning environment, the teacher's scaffolding role means no attempt to transmit knowledge. Instead, the teacher's role is to be supportive rather than a detective.<br>• Teachers are not always in charge of the activities; sometimes, more capable partners can assist.<br>• A teacher, guide, or helper is on hand to provide contextualized support. |
| Authentic Assessment | • Teachers and facilitators should be able to integrate the assessment with the task. This means that what they assess is generally a very polished kind of product, and usually, it has taken significant time and effort to collaborate with other students. It is important to align the tasks with the assessment so that the teacher can seamlessly integrate them. |

The following table illustrates how the elements of authentic learning have been addressed in main activities within the system, the relationship with each as component widgets. Authentic tasks form the core of an authentic learning environment. The completion of these tasks is what the students put most of their time and effort into during the unit [4] [5]

Table 2. Shows the Main Widgets in the system mapped against is each task and the relationship between authentic learning.

| **Widgets of the System** | **Task Description** |
|---|---|
| Webinars | • Opportunities to learn from others through access to various levels of expertise Lectures and expert guest lectures Optional peer review Open education and custom-made internet resources. |
| Discussion forum | • Teachers can create discussion boards to encourage students to think critically about their unit work and interact with each other's ideas.<br>• The teacher must first create one or more forums before users can start message threads. A forum is where participants discuss a topic or a group of related topics. |
| Collaborative Learning | • Real-time video conferencing tool with file sharing, apps, and virtual whiteboard. discussion. |
| Quiz-<br><br>(Mapping with the bloom's taxonomy rubrics remembering, understanding, applying, analyzing, evaluating) | • Based on the curriculum built in the software, the student is progressively challenged to remember, understand, apply it, and analyze by themselves.<br>• Find and create gamified quizzes and interactive lessons to keep any learner interested and involved. |
| Storyboarding- The Projects are close to students' own life. | • Each student expresses their ideas, experience, and interests, thus enabling creative technics. |
| Flash cards | • Contain exciting facts about the subjects being studied; students can add images, words, phrases, or numbers to test themselves on the subject being studied quickly.<br>• If learners prefer, they can search our database for an image that matches the card. |
| Online Video | • Learners have unlimited access to the system's high- quality, up-to-date, and engaging video tutorials. Automatically suggest videos based on the unit they are currently studying.<br>• Video courses are presented in bite-size chunks. It is simple to locate quick answers to specific questions. |
| Games-<br>Action games eg:Doom | • Creativity and innovative thinking-By brainstorming new game ideas.<br>• An opportunity to push students' skills further and demonstrate students, newfound skills as developers and creative thinkers.<br>• Categories students' thinking skills, ranging from recalling and evaluating students by themself in an interactive way. |
| My Library | • The system's libraries provide a variety of services and facilities to students and staff.<br>• Connecting and collaborating with peers and supervisors via technology and experimenting with new platforms and applications.<br>• The service is designed to provide just-in-time assistance when users are unsure what to do or where to click to resolve to learn technology issues as they arise and get back to work. |
| Turnitin | • Maintaining academic integrity.<br>• Providing instructors with software that streamlines |

| | |
|---|---|
| | manual grading processes will allow them to devote more time to teaching.<br>• Help students develop their original thinking skills by providing high-quality, actionable feedback that is easy to incorporate into their teachers' existing workflows. |

The interface comprises two leading portals. One is a student, and the other is a staff. The system enables easy access to both students and teachers and allows customization. Fig. 1 shows the staff portal, and fig. 2 and fig.3 show the student preview of the system.
The overall images illustrate the interface of the system. The system has different widgets different to a Moodle as it enables the learner to self-evaluate themselves through game-based learning aligned with Bloom's taxonomy framework.
The learner will be receiving badges and star points which will indicate the level of achievement. This way, they will evaluate their current level and the level they need to reach along with the tasks they need to complete.

The unique feature of the proposed system as it enables the learning more interesting, fun and enjoyable activity in a way that they also interact with others when it comes to game playing. Depending on the subject context, there will be games designed to take part as teams. This learning approach will also create so that those failures face in interactive tools and games and the successes are informative. The learner will learn from the losses or poor points and be motivated to excel in their learning journey.

When this approach is applied in an educational context, learners can be presented with challenging problems, requiring participation from multiple players around realistic scientific concepts. The system also enables a personalized experience sharing their ideas, expertise and interests through the storyboard widget, as storytelling is a way of bridging the generation gap and forming the foundation to understand the future better. The author has observed authentic learning as a practical approach to effective life-long learning.

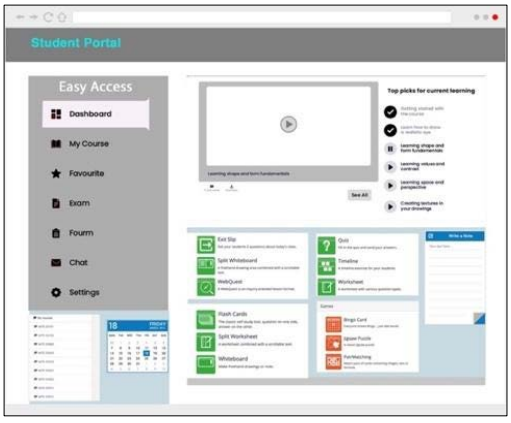


Fig. 1 Interface of the system Stuff Portal

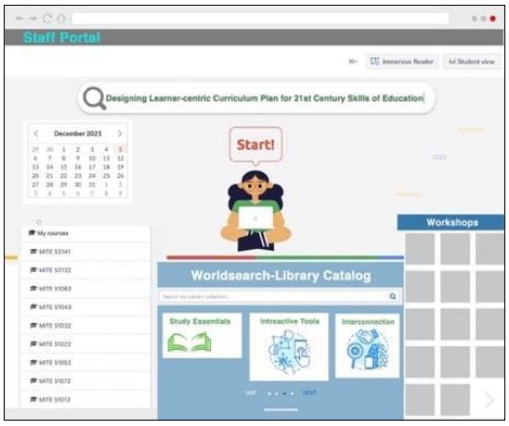


Fig. 2 Interface of the system Student Portal

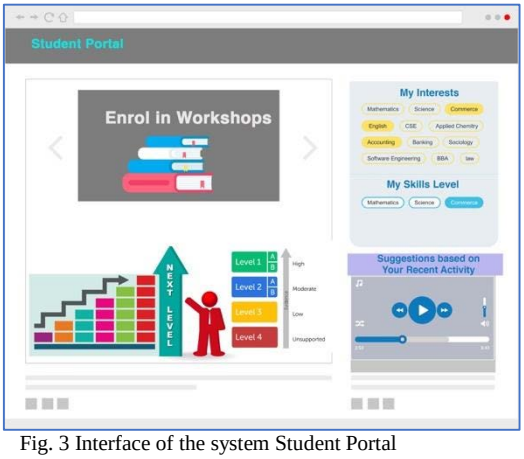


Fig. 3 Interface of the system Student Portal

This interface has easy access as the widget system has forums, chat options, links with open educational resources, whiteboard management, and different widgets as education components, including quizzes, timelines, worksheets. Worksheets will keep track of activities and enable the learner to evaluate themselves through refractive analysis. The facilitator will only guide the students. The widgets also include top ticks for current learning. This widget will pop up links to different sources and references aligned with the learner's interests and will motivate and engage learners in the interested study arena.

There will be action-based games such as Doom. Action games and These strategy games will enable the learner to self-evaluate and assess his or her levels. In Bloom's Taxonomy, the evaluation level is where students judge the value of ideas, items, materials, and more. Evaluation is the final level of Bloom's taxonomy pyramid. At this level, students are expected to bring in all they have learned to make informed and sound material evaluations. In the meantime, they can evaluate their mapping to Bloom's taxonomy frame. Fig.3 image Also, exercise-based games such as drawer mind mapping exercises will assist the learner in thinking creatively and selecting effective.

## IV. SYSTEM FUNCTIONAL VIEW

UML stands for Unified Modelling Language, which is diagrams capable of accurately describing a system from various perspectives. Each view consists of diagrams illustrating various aspects of the system [6]. Within this system, the functional and dynamic views of the system are presented and explained. In this paper, the author will discuss how two distinct types of UML diagrams can be used to depict the functional view of a system. Use Case diagrams and activity diagrams created based on the system's predefined functional requirements.

### *A. Use Case Diagram*

The use cases diagram the system's functional requirements. The use case illustrates the interactions between the system's various components. In the use case diagram, the actor represents a system component.
It reflects the functions of the user or application required to interact or communicate with the primary use case of the system. The proposed system has two main actors: teachers and students who interact directly with the system. Widgets and their functions are components of our system.

Use Case Diagrams are created to ensure that each actor has the appropriate access to the relevant components.
To create the Use Case Diagram of the defined system in the paper. The functional requirements for the system were determined, as shown in
Fig. 4. Use Case diagrams were developed to show learner and staff activities' significant processes.

Use Case diagram helps model the system's functional requirements and features the flow of events [7]. It is essential to mention the main scenario, alternative scenario and actors involved with the use case when writing the Use Case description [8], [9].
As in the system, one of the mains follows can be described as the system prompting the teacher to log in with the proper credentials. Then the system validates the ID and password. after that teacher can successfully access the staff portal of the system and access the functions for staff.

### *B. Activity Diagram*

Activity Diagrams describe how activities are coordinated to provide a service at different levels of abstraction. Typically, an event needs to be achieved by some operations, mainly where the operation is intended to perform several different things that require coordination, or how the events in a single-use case relate to one another, most use cases where activities may overlap and require coordination. Activity Diagrams describe how activities are coordinated to provide a service at different levels of abstraction. [10]

Activity diagrams, the objects needed to execute the functional requirements are identified. Furthermore, the responses taking place during the execution are clarified. can conclude that activity diagrams describe the system's functionality from a sequential aspect [11]

In this system, fig.5. shows the important processes of creating a quiz within the system. The user initiates the system by entering their login credentials then the system validates the user's ID and password. Then the based on that system verifies the user enter into the portal. According to the potential, the user has different system features. Furthermore, process notification management and updating notification process was defined in fig. 6 as activity diagrams.


Fig. 4 Use Case Diagram of main Actros's and system response

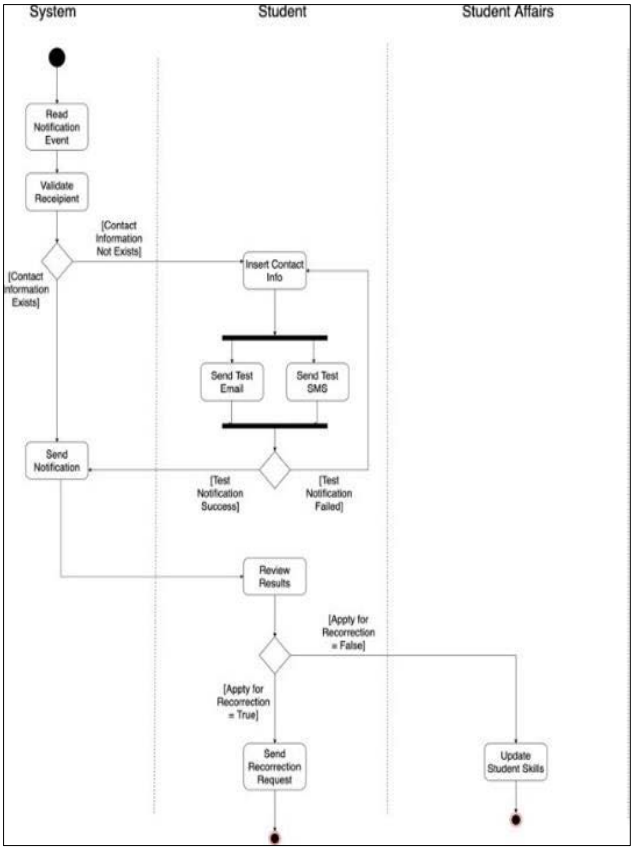



Fig. 5 Activity Diagram of Notification Management

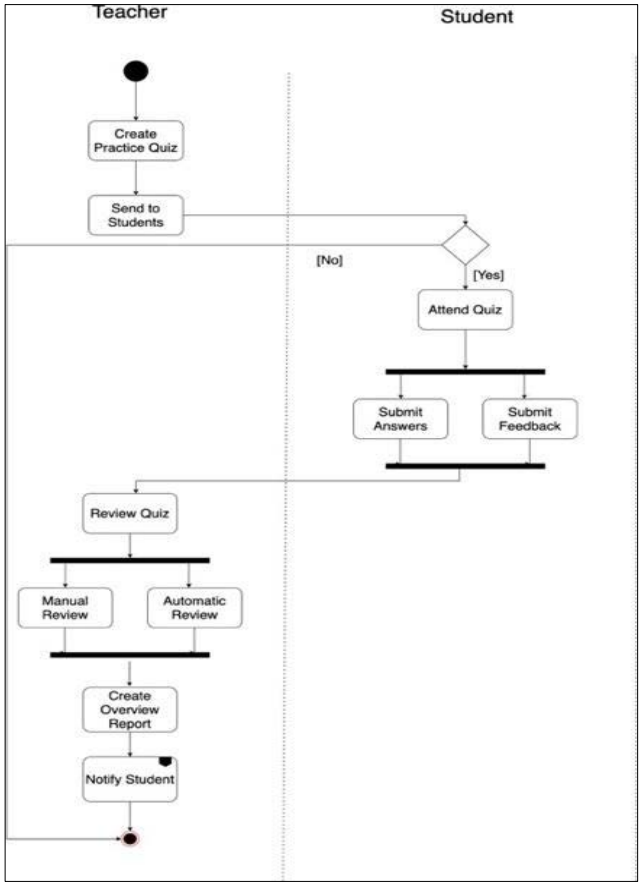



Fig. 6. Activity diagram of Quiz Crating Pro

## V. CONCLUSION

The proposed system is observed to significantly contribute to enhancing the teaching and learning approaches, overcoming the inefficiencies prevailing in the contemporary education system. Implementing the proposed system will bring positive change, transforming the educational system, enabling new thinking and new ways of learning aligned with the drastic dynamic technological changes resulting in the world. Thus, bringing innovation and creative lifelong learning is possible. Learning with Technology has attempted to prepare teachers for rapidly changing technological advancements.

The approach has been observed through literature to enhance students' knowledge and skills and promote self-directed learning and 21st- century learning skills. Students formulate Technology in a real scenario. This proposed system is an expert system typically composed of at least three primary components: inference engine, the knowledge base, and the User interface.

As the proposed system has micro components, the following can be identified as future directions. Software development life cycle (SDLC) is a series of phases the software will be realized and developed from an understanding of the concept. The iterative Incremental Model can be identified as the model going to align for the proposed system. Incremental development model helps to develop the software architecture and processes, overcoming the drawbacks of the waterfall model. [12]

In the interactive incremental model, as the initial phase, a partial implementation of the proposed system will be constructed as a deliverable phase and phase to phase increased functionality will be incorporated. Based on the actors' feedback, who are the key users of the system, if there are any defects or errors from the prior delivery, they will be fixed, and the working system will be delivered. This process will be repeated as a cycle of phases until the entire system development is completed. These repetitive processes are referred to as interactions, and at the end of each scenario, a system increment is delivered. Iterative, incremental development means improvements are made ongoing, so the result is likely to be delivered on time and of higher quality.

## REFERENCES

[1] Lombardi, Marilyn, Authentic Learning for the 21st Century: An Overview. (2007).
[2] D. L. Liyanage, “Authentic learning teaching model,” https://island.lk/authentic-learning-teaching-model/, 2021.
[3] J. Groff, “TECHNOLOGY-RICH INNOVATIVE LEARNING ENVIRONMENTS,” 2013.

[4] J. H. T. R. R. Oliver, "Authentic Tasks Online: A synergy among learner, task, and technology," ECU PUBLICATIONS PRE. 2011.

[5] T. C. R. R. O. an Herrington, A Guide to Authentic e-Learning, Routledge, 2010.

[6] A. Mohammed, "ML Modeling of Online Public Bus Reservation System in Egypt," 2020.

[7] a. Tonella, "Reverse engineering of object oriented code," Proceedings of the 27th international conference on Software engineering, 2005.

[8] A. R. Mohammed and S. S. Kassem, "E-Learning System Model For University Education Using UML," Sixth International Conference on e-Learning (econf),, no. doi: 10.1109/econf51404.2020.9385482., pp. pp. 35-39, 2020.

[9] Object Management Group, "Unified Modeling Language," omg, 22 02 2013. [Online]. Available: https://www.omg.org/spec/UML/About-UML/. [Accessed 2021].

[10] V. P. Online, "Visual Paradigm Online," 2021. [Online]. Available: https://www.visual-paradigm.com/guide/uml-unified-modeling-language/what-is-activity-diagram/.

[11] R. C. P. L. P. L. L. Li, "An automated approach to generating usage scenarios from UML activity diagrams," software Engineering Conference, 2005. APSEC '05. 12th Asia-Pacific, 2005.

[12] M. Sami, "Software development life cycle models and methodologies - Mohamed Sami," *Melsatar.blog*, 15-Mar-2012. [Online]. development-life-cycle-models-and-methodologies/.

INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
*Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka*
*18th December 2021*

*Index No: TL-35-80*

# Model for Integration of Technology in Authentic Education -An interpretation of a literature Review

Bhagya Manjaree HS(1st Author)
*Department of Chemistry*
*University of Colombo*
Colombo, Sri Lanka
2020s286@stu.cmb.ac.lk

Laalitha S.I. Liyanage (2nd Author)
*Department of Applied Computing*
*University of Kelaniya*
Kelaniya, Sri Lanka
laalitha@kln.ac.lk

Thilini P. Rupasinghe (3rd Author)
*Department of Applied Computing*
*University of Kelaniya*
Kelaniya, Sri Lanka
thilinir@kln.ac.lk

Aashani Tillekaratne (4th Author)
*Department of Chemistry*
*University of Colombo*
Colombo, Sri Lanka
taashani@sci.cmb.ac.lk

KMN de Silva (5th Author)
*Department of Chemistry*
*University of Colombo*
Colombo, Sri Lanka
kmnd@chem.cmb.ac.lk

***Abstract*— This literature review is conducted to identify, appraise and synthesize empirical evidence of a filtered list of recent literature regarding methods in which technology could be integrated to facilitate authentic learning pedagogy. A protocol was developed to carry out a search for screening[1]. iDiscover search engine of University of Cambridge library was used for selection and filtering of the articles for their appropriateness. Critical appraisal was performed and data was extracted to map, conceptualize and synthesize the proposed tripod model for integration of technology in authentic education. This model depicts the findings in three zones namely, foundational layer, operational layer and the stage which is the platform for authentic education. Understanding the landscape of the tripod model for integration of technology in authentic education could be quite decisive in selecting the best-fit technological tool. This article argues about how technological interventions could enhance the outcomes of authentic education and the need of an appropriate pedagogical strategy to align such interventions to the elements of authentic education.**

***Keywords—authentic education, technology integration***

## I. INTRODUCTION

Critical proclamations are made by education reformers on the irrelevance of schooling to the real world at the beginning of the 19th century. Their claims were related to the formal and abstract education that has minimal application to everyday life leaving students unprepared to utilize theoretical/conceptual knowledge to overcome everyday challenges. There is a necessary requirement to connect these two worlds in order to make school education relevant for the real-world [2]. Two persistent maladies that make school education irrelevant to the real world are debated extensively. Firstly, it is about student work not allowing them to use their minds and having no intrinsic meaning or value to students beyond achieving success in examinations. Secondly, it is about the student activities which were not interesting and relevant enough for adequate student engagement [3] . In other words, the lessons were focusing exclusively on theoretical components than the real-world practical components.

Enabling a student to explain what is authentically happening around him/her could perhaps be the most important purpose of education [4]. John Dewey who was a renowned and influential pragmatist, progressivist, educator, philosopher, and social reformer believed that school should be representative of the social environment and that students learn best in a natural social setting. Further he believed that students were all unique learners. Therefore, traditional classroom would not be developmentally appropriate for young learners [4]. Hence, if the classroom setting mimicked the real world, it would enable students to make connections between knowledge and skills learned at school to the challenges faced in the real world. This approach would provide a sense of purpose for learning in school and enhance student ownership of learning.

Dewey with Hans Freudenthal from The Netherlands put forth a bottom-up ‘authentic pedagogical process’ named ‘re-invention’ where relevant real-world contexts and personnel were made to be part of the teaching and learning process in the classroom. This would then make the classroom learning experience similar to ‘real-life’ and thereby expose the students directly to various real-world social situations [2].

Direct instruction knowledge is largely transferred to students as `inert knowledge’ [5]. Students act as passive listeners without much opportunity for the application of knowledge. According to Vygotsky, knowledge is considered to be developed and used in a social and physical environment to which it inseparably binds rather than being an abstract entity [6]. The structure, content and coherence of the

concepts are determined by the respective situation or context [7]. The concept of learning in situational contexts or 'Situated learning' came to limelight amidst an environment of such an understanding [8].

Experiential learning and authentic contexts are strongly inter-connected aspects in education. In the early nineteen hundred learning through real life contexts, learning by doing, learning through projects, learning through problem solving were considered as key pedagogies of experiential learning. However, by the end of 19th century these pedagogies were aligned with the three criteria of authentic learning which are inductive approach for solving problems using the knowledge learnt, student's active participation and application of knowledge beyond classroom contexts to real-life settings [9].

### *A. Constructivism and Authentic Education*

In constructivism knowledge was considered as a direct product of a learner's activities. It was not merely the result of dissemination. Rather, knowledge was constructed by the learner through relating new knowledge to already existing cognitive structures in the learner's mind. The learning of knowledge is done by actively participating in different social contexts and experiencing the realities of those contexts. In constructivism the student has different options for learning independently[2]. Authentic education displays many similarities to constructivism. It is an active process in which the student interprets and negotiates new information by himself in contrast to the traditional setup of teacher-centered education.

Authentic education recognizes that student learning depends upon past experiences or pre-knowledge of students. Therefore, teachers are expected to facilitate requirements of the learners arising from their past experiences, by being a coach, guide or a mentor directing students through meaningful learning. Facilitation is enacted through collaborative learning which is recognized as 'cognitive apprenticeship'[10]. This enables the student to develop higher-order thinking skills beyond simple reproduction of information. Learners will be motivated to interpret, evaluate, compare and contrast information to come up with innovative solutions to real-world problems without limiting themselves to mere recovery of information. Learning in the constructivist paradigm is more powerful when students can draw meaningful connections between their classroom learning and personal experiences in the real-world [11]. In this paradigm the learning takes place while immersed in the real world, whereas in conventional direct instruction the real-world experience comes much later at the end of schooling.

### *B. Characteristics of Authentic Education*

'Authentic Education' concept integrates core principles of learning into one pedagogy. Authentic education does not have a single definition. However, few widely accepted frameworks explain the core concepts of the pedagogy. Roelofs argues that any education setting has a degree of authenticity according to the characteristics of the frameworks that are considered for identifying authenticity. [2].

### *C. Herrington's 9 Principles of Authentic Learning*

Among numerous strategies put forth for authentic education, the framework by Jan Herrington is well-accepted [12]. Based on constructivist approach Herrington has put forth a model for authentic learning. It specifies nine critical characteristics that can be identified in authentic learning environments [13]. Further research has proposed effective instructional guidelines for the aforementioned situated learning framework for creating authentic learning environments. These 42 guidelines put forth by Herrington under each element is summarized in Table 1. 'Context' is the teaching and learning process using a realistic scenario which leads to meaningful/purposeful learning. Factors such as motivation and empathy could be argued to be automatically generated in such a situation. 'Authentic activity' is expected to mimic the complexity of a real-life situation as well as open-ended to have more than one solution. The activity should keep the student engaged for a considerably long period of time as per Herrington's model. 'Expert thinking and modelling' are where she suggests streamlining of thinking processes via the use of real-time expert tools. Facilitating communication and reflection through multiple perspectives obtained from peers, experts and teachers through collaborative learning encompass three elements of this framework. Articulating the findings in verbal or in written medium is another aspect in Herrington's framework. Coaching and scaffolding renders a supportive aspect that is missing from the conventional 'didactic model' of teaching. Finally, assessment is expected to be integrated within the task itself through the development of polished products which gives the learning process a 'value outside of the learning environment'.

Table 1. Checklist of guidelines for the instructional design of a learning environment which enables the situated elements to be operationalized [14],[15]

| **Element of Authentic learning – Herrington's model** | **Guidelines for design and implementation of learning environment** |
|---|---|
| 1. Provide authentic context that reflect the way the knowledge will be used in real-life<br>*-Contexualizing education by situating the instructions and event of teaching and learning within a realistic scenario to provide meaning for learning.* | A situated learning environment should provide:<br>-a physical environment which reflects the way the knowledge will ultimately be used<br>-a design to preserve the complexity of the real-life setting with 'rich situational affordances'<br>-a large number of resources to enable sustained examination from a number of different perspectives<br>- a design which makes no attempt to fragment or simplify the environment |
| 2. Provide authentic activities<br>*-Complex, open-ended, realistic task requiring* | -activities which have real-world relevance<br>-ill-defined activities |

| Element of Authentic learning – Herrington's model | Guidelines for design and implementation of learning environment |
|---|---|
| *long time to complete and not simple tasks that can be completed quickly.* | -an opportunity for students to define the tasks and sub-tasks required to complete the activity<br>-a sustained period of time for investigation<br>-the opportunity to collaborate<br>-tasks which can be integrated across subject areas |
| 3. Provide access to expert performances and the modelling of processes<br>*-Modelling expertise;distributed expertise to be tapped using technological tools.* | -access to expert thinking and modelling processes<br>-access to learners in various levels of expertise<br>-opportunity for the sharing of narratives and stories<br>-access to the social periphery or the observation of real-life episodes as they occur |
| 4. Provide multiple roles and perspectives<br>*-Ensuring that teaching and learning is not relying on a single source.* | -different perspectives on the topics from various points of view<br>-the opportunity to express different points of view through collaboration<br>-the opportunity to criss-cross the learning environment by providing more than one investigation within a resource sufficiently rich to sustain repeated examination |
| 5. Support collaborative construction of knowledge<br>*-Provide opportunities for students to collaborate.* | -tasks which are addressed to a group rather than an individual<br>-classroom organization into pairs or small groups<br>-appropriate incentive structure for whole group achievement |
| 6. Promote reflection to enable abstractions to be formed<br>*-Provide oppotunities for students to reflect both in task and on task* [15]. | -authentic context and task<br>-the facility for students to return to any element of the program if desired, and to act upon reflection<br>-the opportunity for learners to compare themselves with experts<br>-the opportunity for learners to compare themselves with other learners in varying stages of accomplishment<br>-collaborative groupings of students to enable reflection with aware attention |
| 7. Promote articulation to enable tacit knowledge to be made explicit<br>*-Provide oppotunities for students articulate their growing understanding in speech and in writing when overall task and activities are completed.* | -a complex task incorporating inherent, as opposed to constructed, opportunities to articulate<br>-collaborative, groups to enable social then individual understanding<br>-public presentation of argument to enable articulation and defense of learning |
| 8. Provide coaching by the teacher at critical times, and scaffolding and fading of teacher support<br>*-Provide scaffolding and coaching in the learning environment through less didactic approach through a more supporting role* [15]. | -a complex, open-ended learning environment<br>-no attempt to provide intrinsic scaffolding and coaching<br>-collaborative learning, where more able partners can assist with scaffolding and coaching<br>-recommendations that the teacher implementing the program is available for coaching and scaffolding assistance for a significant portion of the period of use |
| 9. Provide for integrated assessment of learning within the tasks<br>*-Using authentic assessments to assess products with a value outside the learning environment, in a realistic setup ratherthan implementing separate tests. Realworld clients can be used in this case and genunine polished products of students could be used in their working lives.* | -fidelity of context<br>-the opportunity for students to be effective performers with acquired knowledge, and to craft polished, performances or products<br>-significant student time and effort in collaboration with others<br>-complex, ill structured challenges that require judgement, and a full array of tasks<br>-the assessment to be seamlessly integrated with the activity<br>-multiple indicators of learning<br>-validity and reliability with appropriate criteria for scoring varied products |

[1]

## II. METHOD

Literature review was done through an electronic search of articles. All selected articles were from peer-reviewed journals. The main search terms used were Authentic education and technology. Peer reviewed articles written in English published between 2010 and 2021 for which full text is available online were used for the search. The initial search resulted in 307 journal articles, 02 reviews and 01 book chapter. Out of which 30 most relevant items from the initial search results were chosen based on where they have used technology within the teaching and learning setup, in order to support the elements of Herrington's framework. For example, context designing, assessments, collaboration, communication etc. Those were further filtered based on technological tool terms such as 'Audio, Video, Computer, Digital, Mobile, 3D, Simulation, Computational, Modelling, Cloud sharing, Data storage, Data base management, Web, Artificial intelligence, Virtual reality, Internet, Learning management systems, e-portfolio, Blog and Social media'. This resulted in a comprehensive analysis of 09 studies as listed below in Table 2.

### *Overview*

Table 2. Summary of literature for technology integration in different educational setups

| Description of Study (Research Design/Subject/ Source/Analysis) | *Remarks (Findings/Technology tool/s used)* |
|---|---|
| -Qualitative data from student comments on educational blog, interviews & classroom observations<br>-History<br>-USA School students<br>-Rubric based comment analysis-Thematic<br>[16] | (a) Students engaged in analysis while working in the blog environment<br>(b) Upon cultural experiences, they were able to better use their prior knowledge<br>(c) a variety ofaffordances related to blogging encouraged and supported students as they completed their work<br>(d) Blogging activities were constrained by the limits of students' literacy and subject specific skills, and the limits of technology.<br>- *Web, Social networking blogging tool Web 2.0* |
| -Survey and in-depth interviewing<br>-Wide array of subjects<br>-265 higher educators; South African university<br>-Statistical Analysis | -The highest levels of authenticity were found for element authentic context and task, wherein lowest was found for articulation.<br>-Moderate correlation identified between levels of authenticity and the |

| Description of Study (Research Design/Subject/ Source/Analysis) | Remarks (Findings/Technology tool/s used) |
|---|---|
| [17] | role played by emerging technologies in achieving the authenticity, showing a potentially symbiotic relationship between them.<br>*-Digital media videos/mobile phones/web/online tutoring platforms/blogs/Channels such as Wikis and Google docs* |
| - Design based research-written material, artifacts collection, interviewing and evaluation report<br>-09-University educators; Australian university<br>-Thematic analysis<br>[18] | -Professional learning community (PLC) would enrich understanding of teaching with mobile technologies and would enhance teaching.<br>-Enriched engagement with m-learning may be promoted by the establishment of a PLC<br>*-Mobile Technology-Podcast/audio/texting/blogging/use of social media* |
| -Review<br>-Special education<br>[19] | - The use of online digital portfolios has satisfied both the need to evaluate teacher candidates' performance in special education settings and encourage deeper reflection through the use of interactive digital technologies.<br>-Evidence for learning important instructional technologies by implementing them in authentic settings.<br>-The author suggests the implementation of digital portfolios may reinforce best practices in special education.<br>*-Phone, text, video hangout, Video, pictures, and graphs, Google Doc, Online chat functions, Learning Management Systems* |
| -Design based action research<br>-Language studies<br>-School students; San Diego<br>[20] | - Effective participation in peer review and constructive response<br>-Provision of assistance in setting up networking flatforms rather than setting up them on their own<br>-Students could authentically see the value learning<br>*- Facebook, Twitter, Myspace, video chats via Skype, e-portfolio application called Mahara-social networking and friending* |
| -Action research<br>-Climate change education<br>-Primary school learners; Greece<br>- Self-reflective analysis<br>[21] | -Successful application of the ExConTra learning paradigm-Experiential, Constructivist, Transformative-Changing, Reflecting and Acting<br>-Cross-thematic and interdisciplinary curricular approach,<br>-Learning activities that are long-term, interdisciplinary, student centered and integrated with real-world issues and practices in which students planned, implemented and evaluated projects with real-world applications beyond the classroom<br>*-Web-based learning environment-hypermedia technology, open-source learning technologies, ICT tools-concept maps (Text2Mindmap), spreadsheets (Zoho Sheet), presentations (280slides), paint tools (Pixlr), word processing (Zoho Writer), Venn diagrams (classtools.net), largely elicited from the Web, Drupal – an open-source content management system (CMS) similar to platforms like Joomla and Moodle* |
| -Quantitative survey<br>-Nursing, education, liberal arts and sciences<br>-Southwestern university students in the United States -28 online courses at in size from 8 to 25 enrolled students with a total survey response of 392<br>-Structural equation modeling (SEM) approach<br>[22] | -Student-centered learning in the online setting vs online learner satisfaction was studied.<br>-Learner relevance, active learning, authentic learning, learner autonomy, and computer technology competence has proven to predict students' perceived satisfaction with online courses and web-based distance education at a statistically significant level.<br>*-Online learning environment* |
| -Quasi experiment-pre-test and post-test questionnaire<br>-Health education<br>-Elementary school in northern Taiwan-Classes of fourth graders, 52 students comprised of 28 boys and 24 girls in the experimental group and 52 students comprised of 29 boys and 23 girls in the control group<br>-Statistical analysis<br>[23] | -Digital game-based learning, contextual learning, contextual decision-making, learning achievement, learning motivation, problem-solving ability<br>-Improved the students' learning motivation, but also their learning achievement and problem-solving competences.<br>-Significant two-way interaction suggested that the contextual game-based learning approach benefited the higher motivation students more than the lower motivation ones in terms of the advanced knowledge, showing the importance and potential of applying contextual games to health education activities.<br>*-Multimedia learning-digital game-based learning* |
| -Classical Structural Equation Modeling (SEM)<br>-Quantitative research approach<br>-Information Technology<br>-308 undergraduate students from a public university in Turkey<br>- Statistical analysis<br>[24] | -Students' information management (i.e., retrieve, store, share, and apply) practices were found to be significantly associated with their attitudes, which were in return significantly associated with the behavioral intentions.<br>-Employing the Mobile cloud computing (MCC) services for personal information management should be supported and encouraged in the higher education by designing authentic learning environments and by scaffolding the students in using such services.<br>*-Mobile cloud computing (MCC), electronic databases (i.e., Scopus, WOS, and Google Scholar), MCC services (i.e., OneDrive, Dropbox, and Google Drive)* |

## III. FINDINGS

According to the literature survey, it was evident that technology had been used in various forms and levels in educational setups. They range from classes and learning environments in kindergarten to tertiary and higher

educational levels. However, very few studies were found investigating all characteristics put forth in Jan Herrington's framework of authentic teaching and learning environment. Out of the 42 guidelines put forth by Jan Herrington in achieving the 09 elements, only one or few are considered in most of the referred studies [14]. Many studies used technology for various teaching and learning activities. In most of the studies technology was widely exploited for basic communication. However, communication among the main stakeholders of authentic education namely students, peers, senior students, facilitators/teachers, experts/outside world and public are not adequately considered. Although, technology could have been a vital part of facilitation of collaboration as an outcome was not adequately discussed. Technology was often used in creating various pedagogical contexts and content. However, it was rarely used for instruction or scaffolding purposes in authentic education setups. Furthermore, use of technology in design and conduct of assessments (both formative and summative) and providing timely feedback also might need further research. Considering the findings of the referred individual studies it is evident that the success of most of the implementations within these studies have depended upon three common factors. They are;

1. **R**esources
2. **E**fficacy[25],[26]
3. **K**nowledge

Access for physical resources such as hardware and software and access for professional development programs for knowledge and skill upgrade are considered as vital pre-requisites enabling teachers to select and apply the best-fit technique/s as per the situation [16]. Efficacy could be self-efficacy as well as organizational support[27]. Individuals who take the initiative of integrating technological tools to enhance their teaching and learning could be encouraged and motivated to uplift their self-efficacy to use technology in a personalized context, provided that the technological tools they use have favorable system efficacies[28]. It could be argued that, all three factors mentioned above would require support from the existing system/organization as well in addition to the individual acceptances; to result sustainable and collective change at the ground level. Therefore, these three factors (**REK**) could be regarded as the key pillars upon which the success of integrating technology in any educational setup would start its implementation process. In the proposed model REK appear at the foundation layer. Thereby, the resulting change could be made a sustainable one, if continuous update and upgrade occurs within the system especially in terms of physical resources, knowledge and skills. Further, efficacy could be assured via various positive reinforcements such as provision of motivational setups with appropriate recognition, reward for novelty and etc.

Aforementioned interpretation could be graphically represented by a 'Tripod' model as illustrated in Figure 01. The three 'pod legs' consist the foundation layer upon which two 'pod base' layers operate with varying degrees of freedom corresponding to their variation, usage and importance in an educational setup. These two operational layers are connected through the 'central column' technology in various ways and hoist the 'stage' which is the environment where 'Authentic Education' could take place. Above the foundation layer the two levels of the operational layer which encompass Context, Content, Communication, **C**reativity, **C**onnectivity and **C**ollaboration (the 6C's) appear in varying strengths.

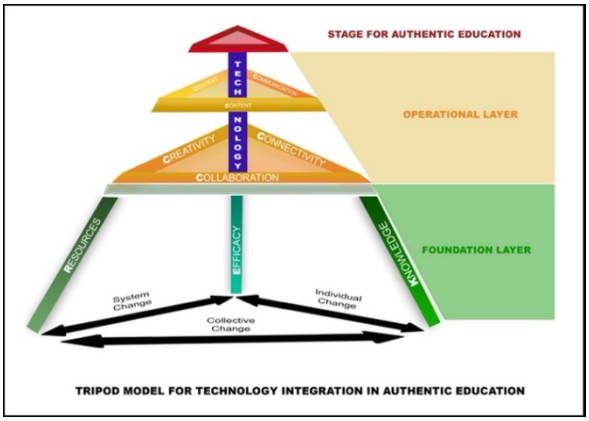



Fig.1. Proposing model for integration of technology within authentic education

In most studies **C**ontext, **C**ontent and **C**ommunication appear to be the factors with the highest diversity where most educators prefer to employ technology [16],[17]. This could be due its large scope considering subject variations, content variations and other cultural, social and demographical variabilities. Compared to context, content and communication there is limited diversity in collaboration, connectivity and creativity. It is difficult to justify which one is of more importance than the other. Since there cannot be education without context, content and communication in any pedagogy, it could be considered as a highly evolving layer subjected to increased diversity. **C**reativity, **C**onnectivity and **C**ollaboration stand comparatively selective in educational setups, therefore could be argued that these exist at two different levels in the operational framework. Considering the function of these elements these could be presented as the operational layer within which various technological tools could be integrated to appropriate pedagogical strategies. The two different levels which encompass the 6C's would construct the operational layer together in the model proposed.

## IV. CONCLUSION

This paper has summarized evidence for achieving authentic educational components through use of technological tools. A wide array of technological tools used in different educational setups for various subject disciplines were published in the literature. This shows the proven

general use of technological tools in aspects of authentic education. Some studies have intentionally used technology to implement authentic learning whereas some have discussed about the vital components of such a setup separately without mentioning or referring to any authentic education framework. High variation of technology was observed in context, content and communication. However, it can be argued that technology is not adequately integrated for peer, expert, teacher and senior student interactions. High creativity is required for the design of educational activities in the authentic learning paradigm. However, very limited studies show evidence explaining the effect of technology in successfully enabling connectivity, collaboration and creativity. Apart from those aspects, individual qualities and attitudes such as motivation, student achievement, engagement, performance is also found to be enhanced through technology integrated teaching and learning setups. However, it can be argued that these are always underpinning the latter with or without the researcher's acknowledgement in most of the studies. Therefore, further studies via qualitative approaches could assist in understanding more about the aforementioned individual qualities and attitudes and their intra and inter relationships along with the limitations and challenges connected to such implementations in educational setups. In addition, such studies could further be able to explain/expand the foundational layer and its components in the technology integration model proposed. Training is required for both knowledge and skill upgrade regarding educational technologies and design of creative educational experiences. Professional development programs aiming teachers revisiting the methods of access to resources, initiations for change of attitudes, beliefs and self-efficacies could be the most vital pillars to ensure continuous stability. This could also require a system change in terms of changes in policies and practices within a system encouraging technology integrated individual teaching approaches.

Considering such concerns, technology could not possibly be a 'bolted-on' aspect in an authentic educational setup. The current model promotes 'technology enabled learning' rather than its simple integration and will have to be coupled with creative pedagogies and strategies further explaining how technology could be best integrated within the 6Cs of operational layer proposed within the model.[29]

## REFERENCES

[1] MSKTC, "A Guide for Developing a Protocol for Conducting Literature Reviews," *Handb. Pediatr. Retin. OCT Eye-Brain Connect.*, no. 90, p. 7, 2019, [Online]. Available: http://training.cochrane.org/resource/writing-protocol.

[2] E. Roelofs and J. Terwel, "Constructivism and authentic pedagogy: State of the art and recent developments in the Dutch national curriculum in secondary education," *J. Curric. Stud.*, vol. 31, no. 2, pp. 201–227, 1999, doi: 10.1080/002202799183232.

[3] Fred M. Newmann and Gary G. Wehlage, "Educational Leadership:Authentic Learning:Five Standards of Authentic Instruction," *Educ. Leadersh. ASCD*, vol. 50, no. 7, pp. 8–12, 1993, [Online]. Available: http://www.ascd.org/publications/educational-leadership/apr93/vol50/num07/Five-Standards-of-Authentic-Instruction.aspx.

[4] M. Williams, "John Dewey in the 21st Century," *J. Inq. Action Educ.*, vol. 9, no. 1, pp. 91–102, 2017.

[5] R. S. Grabinger and J. C. Dunlap, "Rich environments for active learning: a definition," *Alt-J*, vol. 3, no. 2, pp. 5–34, 1995, doi: 10.1080/0968776950030202.

[6] L. S. Vygotsky, *Mind in Society The Development of Higher Psychological Processes*. 1978.

[7] J. S. Brown, A. Collins, and P. Duguid, "Situated Cognition and the Culture of Learning," *Educ. Res.*, vol. 18, no. 1, pp. 32–42, 1989, doi: 10.3102/0013189X018001032.

[8] P. M. Miller, G. A. and GIildea, "0046_SA1987_HowChildrenLearnWords.pdf." .

[9] N. A. Knobloch, "Is Experiential Learning Authentic?," *J. Agric. Educ.*, vol. 44, no. 4, pp. 22–34, 2003, doi: 10.5032/jae.2003.04022.

[10] G. Kou, "No Title*עלון הנוטע*," *ענף הקיווי: תמונת מצב*, vol. 66, pp. 37–39, 2012.

[11] F. Newmann, H. Marks, and A. Gamoran, "Standards That Boost Student Performance," *Issues Restruct. Sch.*, 1995.

[12] A. Herrington and J. Herrington, "What is an authentic learning environment?," in *Authentic Learning Environments in Higher Education*, no. January, 2005, pp. 1–13.

[13] J. Herrington and R. Oliver, "Critical Characteristics of Situated Learning : Implications for the Instructional Design of Multimedia," *Learn. with Technol.*, vol. withtechno, pp. 253–262, 1991, [Online]. Available: http://www.konstruktivismus.uni-koeln.de/didaktik/situierteslernen/herrington.pdf.

[14] J. Herrington and R. Oliver, "An instructional design framework for authentic learning environments," *Educ. Technol. Res. Dev.*, vol. 48, no. 3, pp. 23–48, 2000, doi: 10.1007/BF02319856.

[15] M. Shaughnessy, "Q & A with Ed Tech Leaders: Interview with Harold Stolovitch.," *Educ. Technol.*, vol. 56, no. 5, pp. 49–52, 2016.

[16] B. M. Groome, J. Rankin, and J. Wheary, "The Importance of Support & Coll aboration in Solving the Crisis of STEM Teacher Retention An Issue Brief *," 2012.

[17] V. Bozalek *et al.*, "The use of emerging technologies for authentic learning: A South African study in higher education," *Br. J. Educ. Technol.*, vol. 44, no. 4, pp. 629–638, 2013, doi: 10.1111/bjet.12046.

[18] S. Schuck, P. Aubusson, M. Kearney, and K. Burden, "Mobilising teacher education: A study of a professional learning community," *Teach. Dev.*, vol. 17, no. 1, pp. 1–18, 2013, doi: 10.1080/13664530.2012.752671.

[19] C. Wyatt-Smith, C. Alexander, D. Fishburn, and P. McMahon, "Standards of practice to standards of evidence: developing assessment capable teachers," *Assess. Educ. Princ. Policy Pract.*, vol. 24, no. 2, pp. 250–270, 2017, doi: 10.1080/0969594X.2016.1228603.

[20] A. Teng, "Writing Teachers Should Comment on Facebook Walls," *Voices from Middle*, vol. 19, no. 4, pp. 34–38, 2012, [Online]. Available: https://proxy.library.kent.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ976269&site=ehost-live%0Ahttp://www.ncte.org/journals/vm/issues/v19-4.

[21] V. Makrakis, D. Gkotzos, and N. Larios, "ICT-enabled climate change education and children's rights," *J. Teach. Educ. Sustain.*, vol. 14, no. 2, pp. 89–110, 2012, doi: 10.2478/v10099-012-0011-y.

[22] F. Ke and D. Kwak, "Constructs of student-centered online learning on learning satisfaction of a diverse online student body: A structural equation modeling approach," *J. Educ. Comput. Res.*, vol. 48, no. 1, pp. 97–122, 2013, doi: 10.2190/EC.48.1.e.

[23] H. Y. Sung, G. J. Hwang, and Y. F. Yen, "Development of a contextual decision-making game for improving students' learning

performance in a health education course," *Comput. Educ.*, vol. 82, pp. 179–190, 2015, doi: 10.1016/j.compedu.2014.11.012.

[24] I. Arpaci, "A hybrid modeling approach for predicting the educational use of mobile cloud computing services in higher education," *Comput. Human Behav.*, vol. 90, no. January 2018, pp. 181–187, 2019, doi: 10.1016/j.chb.2018.09.005.

[25] G. Rodríguez-Abitia, S. Martínez-Pérez, M. S. Ramirez-Montoya, and E. Lopez-Caudana, "Digital gap in universities and challenges for quality education: A diagnostic study in Mexico and Spain," *Sustain.*, vol. 12, no. 21, pp. 1–14, 2020, doi: 10.3390/su12219069.

[26] J. W. Richardson, "Challenges of adopting the use of technology in less developed countries: The case of Cambodia," *Comp. Educ. Rev.*, vol. 55, no. 1, pp. 8–29, 2011, doi: 10.1086/656430.

[27] M. Igbaria and J. Iivari, "The effects of self-efficacy on computer usage," *Omega*, vol. 23, no. 6, pp. 587–605, 1995, doi: 10.1016/0305-0483(95)00035-6.

[28] D. Jia, A. Bhatti, and S. Nahavandi, "The impact of self-efficacy and perceived system efficacy on effectiveness of virtual training systems," *Behav. Inf. Technol.*, vol. 33, no. 1, pp. 16–35, 2014, doi: 10.1080/0144929X.2012.681067.

[29] P. A. Ertmer and A. Ottenbreit-Leftwich, "Removing obstacles to the pedagogical changes required by Jonassen's vision of authentic technology-enabled learning," *Comput. Educ.*, vol. 64, pp. 175–182, 2013, doi: 10.1016/j.compedu.2012.10.008.

ISSN 2756-9160

# 6TH INTERNATIONAL CONFERENCE IN ADVANCES OF TECHNOLOGY AND COMPUTING 2021,

## 18th - 20th December 2021

ICATC

https://conf.kln.ac.lk/