ComURS

Faculty of Computing
Sabaragamuwa University of Sri Lanka

ComURS 2025

COMPUTING UNDERGRADUATE RESEARCH SYMPOSIUM

Data-Driven Approaches to Global Sustainability

ABSTRACTS

19th of February 2025

Faculty of Computing

Sabaragamuwa University of Sri Lanka

Abstracts of the ComURS2025 Computing Undergraduate Research Symposium 2025

Published by : Faculty of Computing,
Sabaragamuwa University of Sri Lanka

ISBN : ISBN 978-624-5727-57-5

Citation **:** Abstracts of the ComURS2025 Computing Undergraduate Research Symposium 2025,
Faculty of Computing,
Sabaragamuwa University of Sri Lanka.

Available from : Sabaragamuwa University of Sri Lanka.
P.O. Box 02, Belihuloya 70140, Sri Lanka.
Tel: +94 (045) 22 80015
Email: info@comurs.sab.ac.lk
Web: https://www.comurs.sab.ac.lk/

ComURS2025 Computing Undergraduate Research Symposium 2025

Organized by

Faculty of Computing
Sabaragamuwa University of Sri Lanka

ORGANIZING COMMITTEE

Symposium Chair
Mr. HMKT Gunawardhane

Symposium Secretary
Mrs. S Adeeba

Symposium Treasurer
Mr. PGP Kumara

Technical Programme Committee Chair
Mr. GACA Herath

Logistic Chair
Mrs. RMKK Wijerathna

Web Master
Mr. WGS Lakshitha

Information Systems Track
Dr. LS Lekamge (Chair)
Mrs. KGL Chathumini (Co-Chair)

Software Engineering Track
Prof. S Vasanthapriyan (Chair)
Mrs. N Ravikumar (Co-Chair)

Data Science Track
Prof. BTGS Kumara (Chair)
Mrs. S Wasalthilake (Co-Chair)

Open Track
Dr. UAP Ishanka (Chair)
Ms. PMAK Wijerathne (Co-Chair)

Publication Chair
Mrs. JDT Erandi

Publicity Chair
Mrs. PKDK Kaushalya

ADVISORY PANEL

Prof. S Vasanthapriyan	Professor, Sabaragamuwa University of Sri Lanka
Prof. BTGS Kumara	Professor, Sabaragamuwa University of Sri Lanka
Dr. LS Lekamge	Senior Lecturer, Sabaragamuwa University of Sri Lanka
Dr. UAP Ishanka	Senior Lecturer, Sabaragamuwa University of Sri Lanka

EDITORIAL BOARD

PANEL OF REVIEWERS

Professor BTGS Kumara	Faculty of Computing, Sabaragamuwa University of Sri Lanka
Dr. RAHM Rupasinghe	Senior Lecturer, Faculty of Social Sciences and Languages, Sabaragamuwa University of Sri Lanka
Mr. Samadhi Rathnayake	Lecturer, Faculty of Computing, SLIIT
Mr. Kanagasabai Thiruthanigesan	Lecturer, Faculty of Computing, SLIIT Northern University
Ms. Pipuni Wijesiri	Lecturer, Faculty of Computing, SLIIT
Ms. Chathuni Thilakarathne	Lecturer, SLIIT
Mr. Charith Dissanayake	Lecturer, Gampaha Wickramarachchi University
Mr. Janitha Rajakaruna	Lecturer, SLIATE
Mr. Aruna Withanage	Lecturer, Uwa Wellassa University of Sri Lanka
Ms. PBSN Ariyathilake	Lecturer (Probationary), Faculty of Geomatics, Sabaragamuwa University of Sri Lanka
Mr. Dimuth Pathirana	Lecturer (Probationary), Faculty Applied Sciences, Uwa Wellasa University
Mr. LL Gihan Chathuranga	Lecturer (Probationary), Faculty of Science, University of Ruhuna

Mr. M Janotheepan	Lecturer (Probationary), Faculty of Applied Sciences, Uwa Wellasa University
Mr. S Prashanth	Lecturer (Probationary), Faculty of Applied Sciences, Sabaragamuwa University of Sri Lanka
Mr. Thilina Gunarathna	Lecturer (Probationary), Faculty of Applied Sciences, University of Uwa Wellassa
Mr. Umesh Ranassoriya	Lecturer (Probationary), Faculty of Applied Sciences, Uwa Wellasa University
Ms. Akila Brahmana	Lecturer (Probationary), Faculty of Technology, University of Ruhuna
Ms. Buddika Gayashani	Lecturer (Probationary), Faculty of Technology, University of Ruhuna
Ms. Ayesha Chathurangi	Lecturer (Probationary), Faculty of Applied Sciences, Uwa Wellasa University
Ms. P Rukshani	Lecturer (Probationary), Faculty of Technology Studies, University of Vavuniya
Ms. Tharmini Kiriharan	Lecturer (Probationary), Faculty of Applied Sciences, Eastern University of Sri Lanka
Ms. Ashansa Wijerathne	Lecturer (Probationary), Faculty of Computing, Sabaragamuwa University of Sri Lanka
Mrs. W.V.S.K. Wasalthilaka	Lecturer (Probationary), Faculty of Computing, Sabaragamuwa University of Sri Lanka

Mr. WMCJT Kithulwatta	Lecturer (Probationary), Faculty of Technological Studies, Uwa Wellassa University of Sri Lanka
Mrs. KGL Chathumini	Lecturer (Probationary), Faculty of Computing, Sabaragamuwa University of Sri Lanka
Ms. UP Kudagamage	Lecturer (Probationary), Faculty of Computing, Sabaragamuwa University of Sri Lanka
Mrs. RMKK Wijerathne	Lecturer (Probationary), Faculty of Computing, Sabaragamuwa University of Sri Lanka
Mrs. WTS Somaweera	Lecturer (Probationary), Faculty of Computing, Sabaragamuwa University of Sri Lanka
Mrs. J.D.T. Erandi	Lecturer (Probationary), Faculty of Computing, Sabaragamuwa University of Sri Lanka
Ms. Pirunthavi Sivakumar	Lecturer (Probationary), Faculty of Technological Studies, University of Vavuniya
Mr. Pramuka Weerasinghe	Lecturer (Probationary), Faculty of Science, University of Peradeniya
Mrs. Roshani Wijesuriya	Lecturer (Probationary), University of Moratuwa
Mrs. Venuja Niranjan	Lecturer (Probationary), Faculty of Technological Studies, University of Vavuniya
Mr. Vijayakanthan G Ganesalingam	Lecturer (Probationary), Faculty of Applied Science, University of Vavuniya

Ms. Hiruni Peiris	Lecturer (Probationary), University of Ruhuna
Mrs. Tharika Weerakoon	Lecturer (Probationary), Faculty of Science, University of Ruhuna
Ms. Thilini Bakmeedeniya	Lecturer (Probationary), Faculty of Technological Studies, Uwa Wellassa University
Mrs. Tharmini K Kiriharan	Lecturer (Probationary), Faculty of Applied Science, Eastern university of Sri Lanka
Eng. Tiroshan Madushanka	Lecturer (Probationary), Faculty of Science, University of Kelaniya
Mathangi Krishnathasan	Lecturer (Probationary), University of Colombo School of Computing
Mrs. A.M Chandrika Malkanthi	Lecturer (Probationary), Faculty of Applied Science, Sabaragamuwa University
Mrs. PKDK Kaushalya	Lecturer (Temporary), Faculty of Computing, Sabaragamuwa University of Sri Lanka
Mrs. Rukshika R Premathilaka	Lecturer (Temporary), Faculty of Technological Studies, Uva Wellassa University
Mr. Susiri P Dias	Assistant Lecturer (Temporary), Faculty of Applied Science, University of Sabaragamuwa
Ms. Niroshi Alwis	University of Ruhuna
Ms. Malki S Jayawardhana	Lecturer (Temporary), Faculty of Technology, Rajarata University of Sri Lanka

Mrs. K Ranasinghe	MSc (Reading), University of Aizu, Japan
Mr. J Charles	MSc (Reading), Friedrich-Alexander-University Erlangen-Nuremberg (FAU)
Ms. Theshani Marambe	MSc (Reading), Finland
Ms. Anparasy Sivaanpu	PhD Student, University of Albetra
Mr. P. Vigneshwaran	PhD Student, Australia
S. Prasanth	PhD Student, Canada

TABLE OF CONTENTS

ORGANIZING COMMITTEE ii
ADVISORY PANEL iii
EDITORIAL BOARD iv
PANEL OF REVIEWERS v
TABLE OF CONTENTS x
MESSAGE FROM THE VICE-CHANCELLOR xi
MESSAGE FROM THE DEAN xii
MESSAGE FROM THE CHAIR xii
ABSTRACTS 1
Information Systems Track 3
Software Engineering Track 16
Data Science Track 28
Open Track 46
AUTHOR INDEX 56

MESSAGE FROM THE VICE-CHANCELLOR

Sabaragamuwa University of Sri Lanka

As the Vice Chancellor of Sabaragamuwa University of Sri Lanka, it is with great pleasure that I extend my warm greetings to all participants of the Computing Undergraduate Research Symposium (ComURS2025) which is organized by the Faculty of Computing, themed on "Data-Driven Approaches to Global Sustainability". This theme is a timely and forward-thinking initiative that highlights the important role of computing and data science in tackling today's most urgent challenges.

I firmly believe that universities are not merely repositories of knowledge but catalysts for change. Forums like ComURS empower our undergraduates to contribute meaningfully to shaping a sustainable future, fostering innovative thinking and a passion for discovery. Additionally, it provides an exceptional platform for students to showcase their research and exchange ideas with peers and experts from diverse fields.

I take this opportunity to commend the organizing committee for their professionalism and dedication in hosting the second edition of this symposium and also like to extend my gratitude to all contributors, including researchers, track chairs, and stakeholders, for their invaluable support in making this symposium a success. It is truly inspiring to see the commitment of our undergraduates and academic staff in fostering a culture of research and innovation. Together, we can position Sabaragamuwa University of Sri Lanka as a leader in addressing global challenges through research and innovation.

I wish ComURS 2025 every success and hope it inspires our students to excel in research while making meaningful contributions to sustainability.

Professor M. Sunil Shantha
Vice-Chancellor
Sabaragamuwa University of Sri Lanka

MESSAGE FROM THE DEAN

Faculty of Computing
Sabaragamuwa University of Sri Lanka

As the Dean of the Faculty of Computing at Sabaragamuwa University of Sri Lanka, it is my privilege to contribute this message to the Abstract Proceedings of the Computing Undergraduate Research Symposium (ComURS2025), organized by our esteemed Faculty. The 2025 symposium, themed "Data-Driven Approaches to Global Sustainability", highlights the transformative power of data and technology in addressing pressing global challenges.

I am delighted to announce the second edition of this symposium, which has grown into a significant platform for our undergraduates to showcase their innovative research and projects. ComURS2025 focuses on four major areas within Computing—Information Systems, Software Engineering, Data Science, and Open Computing. It provides students with a remarkable opportunity to present their findings, engage in thought-provoking discussions, and connect with experts in the field.

Since 2017, the Faculty of Computing has made remarkable strides in fostering a strong undergraduate research culture**.** Through dedicated mentorship, state-of-the-art research facilities, and collaborative industry partnerships, we have empowered students to pursue high-impact research that contributes to academia and industry alike. Our faculty members and students have actively engaged in peer-reviewed research publications, international and national conferences, and reputed journals**,** reflecting our commitment to academic excellence.

The faculty has played a pivotal role in organizing and participating in prestigious computing conferences**,** fostering an environment where students can learn, network, and present their research on global platforms. These efforts have not only enhanced our students' research capabilities but also positioned them as emerging scholars and innovators in the field of computing.

Our commitment to academic excellence has been recognized through numerous awards and achievements**,** including best paper awards, innovation grants, and research fellowships that acknowledge the outstanding contributions of our students and faculty members. These accomplishments serve as a testament to our unwavering

dedication to developing a culture of research, innovation, and technological advancement.

I extend my heartfelt gratitude to the Vice-Chancellor of Sabaragamuwa University of Sri Lanka, our Evaluators, Heads of Departments, faculty members, and all contributors for their invaluable support in nurturing a dynamic research ecosystem**.** Special appreciation goes to the organizing committee of ComURS2025 for their tireless dedication in ensuring the success of this symposium.

Finally, I wish all participants of ComURS2025 the very best as they present their pioneering research. May this symposium continue to inspire groundbreaking ideas and contribute to the advancement of computing and global sustainability.

Professor S Vasanthapriyan
Dean, Faculty of Computing
Sabaragamuwa University of Sri Lanka

MESSAGE FROM THE CHAIR

Computing Undergraduate Research Symposium
Faculty of Computing
Sabaragamuwa University of Sri Lanka

On behalf of the Organizing Committee, I am honored to present this message for the Abstract Proceedings of the Computing Undergraduate Research Symposium (ComURS2025), organized by the Faculty of Computing, Sabaragamuwa University of Sri Lanka. This year, ComURS marks its second edition under the theme "Data-Driven Approaches to Global Sustainability," highlighting the role of computing and data science in addressing global challenges.

ComURS2025 offers an excellent platform for our undergraduates to present their innovative research, exchange ideas, and receive feedback from academics and industry leaders. The symposium encourages intellectual collaboration, enabling students to apply their knowledge to real-world problems and fostering academic and professional growth. The symposium features presentations across four tracks: Information Systems, Software Engineering, Data Science, and Open. The "Best Presenter" awards will recognize exceptional contributions from each track.

I would like to express my heartfelt gratitude to the Vice-Chancellor of Sabaragamuwa University of Sri Lanka, the Dean of the Faculty of Computing, Heads of Departments, and the entire academic and non-academic staff for their support. My gratitude also goes to the Advisory Panel, Supervisors, Abstract Reviewers, Session Chairs, Poster Evaluators, and the volunteers for their hard work in making this event a success.

Finally, I appreciate the students for their passion and dedication. May this symposium inspire meaningful discussions, impactful research, and motivate all to contribute to a sustainable future. Wishing everyone an enriching experience at ComURS2025.

Mr. HMKT Gunawardhane
Chair
ComURS2025

ABSTRACTS
OF THE
ComURS2025 COMPUTING UNDERGRADUATE RESEARCH SYMPOSIUM 2025

Technical Sessions of the ComURS2025 are organized as follows;

- Abstracts of the **Undergraduate Poster Presentation Session** – (pp. 04–55)
  - Undergraduate Poster Presentations – Information Systems Track
  - Undergraduate Poster Presentations – Software Engineering Track
  - Undergraduate Poster Presentations – Data Science Track
  - Undergraduate Poster Presentations – Open Track

Abstracts

Information Systems Track

Artificial Intelligence Adoption in the Apparel Supply Chain in Sri Lanka: Strategies for Small and Medium-sized Enterprises

Prabodhani L.H.K.N.*, Wijerathna R.M.K.K. and Lekamge L.S.

Department of Computing & Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka
**lhknprabodhani@std.appsc.sab.ac.lk*

Artificial Intelligence (AI) is rapidly transforming supply chain operations across industries, driving efficiency, cost reduction, and sustainability improvements. In Sri Lanka, large-scale apparel companies have successfully integrated AI into their supply chain processes, but small and medium-sized enterprises (SMEs) face significant challenges in adopting AI. This study focused on identifying these challenges in SMEs, particularly key financial, technological, and organizational barriers, by referencing AI practices in large-scale apparel companies. Data were collected through surveys, receiving responses from 156 participants representing 50 SMEs and 280 participants from 12 large-scale apparel companies. Analytical techniques, including descriptive statistics, qualitative analysis, cross-analysis, and graphical representation, were employed to analyze the data. The findings showed that AI adoption increases operational efficiency and competitiveness in the apparel sector, with technologies such as Machine Learning, Predictive Analytics, Robotic Process Automation (RPA), Internet of Things (IoT), and Computer Vision while revealing the barriers for SMEs, such as high initial investment costs, ongoing maintenance costs, limited technical expertise, and resistance to change. Based on these results, a strategic framework with actionable strategies was proposed to guide SMEs in overcoming these barriers and integrating AI technologies effectively into their supply chain operations. This research contributes to the broader AI integration in apparel supply chains, specifically in SMEs in developing economies like Sri Lanka.

Keywords: *AI Adoption Barriers, Apparel Supply Chain, Artificial Intelligence, SMEs, Strategic Framework*

Usage of Digital Twin Technology for Sustainable Development in Sri Lanka's Construction Sector

Geekiyanage C.L.*, Wijerathna R.M.K.K. and Lekamge L.S.

Department of Computing & Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka
**clgeekiyanage@std.appsc.sab.ac.lk*

Digital Twin (DT) technology is transforming the global construction industry, offering real-time monitoring, better decision-making, and enhanced project outcomes. However, DT is primarily used in the design and planning phase, with limited integration across the entire construction lifecycle. This research explored the applications, benefits, and barriers to adopting DT technology in all phases of the construction lifecycle in the Sri Lankan context. Data was collected through a survey that received 254 responses from professionals, including engineers, construction managers, architects, consultants, and site supervisors, representing 25 construction companies in Sri Lanka. Qualitative techniques, descriptive statistics, cross-analysis, and graphical representation were used to analyze the data. The findings showed that DT is widely used in the design and planning phase (56.7% of respondents), with minimal usage in the construction (31.9%) and operational phases (30.3%) and no usage observed in the end-of-life phase within Sri Lanka's construction industry. Key barriers identified include a lack of awareness about DT technology (37.4% of respondents), insufficient technical expertise (32.7%), organizational resistance to change (26.8%), and high cost of implementation (25.2%). Based on these findings, the study proposed a framework of practical strategies to address these challenges and encourage the broader adoption of DT technology. These strategies contribute to sustainable development in the construction industry by using DT technology across all phases of the construction lifecycle, with a specific focus on the Sri Lankan context.

Keywords: *Adoption Barriers, Construction Lifecycle, Digital Twin Technology, Sri Lanka Construction Industry, Sustainable Development Strategies*

Enhancing User Engagement in Website Hero Sections: A Comparative Analysis of 2D and 3D Graphics Using Kansei Engineering

Sandaruwan R.M.T.* and Somaweera W.T.S.

Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka
** rmtsandaruwan@std.appsc.sab.ac.lk*

Hero sections are a critical component of websites responsible for capturing user attention and shaping overall engagement. Despite the increasing adoption of advanced graphics in web design, the effectiveness of two-dimensional (2D) versus three-dimensional (3D) graphics within hero sections has not been extensively studied. Therefore, this research tries to find out the implications of 2D and 3D graphics in hero sections concerning user engagement and emotional response. The study will be guided by the following research question: What are the differences in emotional responses and user engagement between 2D and 3D graphics in website hero sections? Quantitative analysis provides empirical data to guide UI/UX designers. A preliminary survey was conducted with 50 participants, 66% between 23 to 25 years old, and 52% male to identify emotional triggers and design the hero sections accordingly. Using these insights, 20 hero section designs were created as 10 in 2D and 10 in 3D. A follow-up survey was conducted with 100 participants, 69% between 23 and 25 years old, and 58% male to assessed engagement, emotional response, attractiveness, and memorability of these designs. The findings signify a strong preference for 3D graphics with 97% of participants found them more engaging, 91% more visually pleasing, and 94% more memorable compared to 2D graphics. Participants noted that 3D's visual attractiveness fosters greater emotional attachment, while 2D graphics, though functional, lacked dynamism. This study successfully evaluates the effectiveness of 2D versus 3D graphics and provides empirical evidence that 3D enhances user experience. These findings lay the groundwork for future research on complex visual elements in digital interfaces and offer practical insights for UI/UX designers aiming to create emotionally engaging, user-centered web experiences.

Keywords: *2D Graphics, 3D Graphics, Kansei Engineering, User Engagement, Website Hero Section*

Deep Learning Approach for Classifying AI-generated and Human-written Sinhala Answers

Ranathunga R.A.D.K[1*], Rupasingha R.A.H.M.[2] and Kumara B.T.G.S.[3]

[1]Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka
[2]Department of Economics and Statistics, Faculty of Social Sciences and Languages, Sabaragamuwa University of Sri Lanka, Sri Lanka
[3]Department of Data Science, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka
[*]radkranathunga @std.appsc.sab.ac.lk

With increased use of AI on answering academic questions, a real concern arises on cases of cheating especially where languages such as Sinhala are hardly supported by effective detection systems. As a result, it becomes difficult to distinguish content written by AI from human written content, which compromises and distorts fair evaluations and originality within education. To address this issue, this study introduces a deep learning model to differentiate AI-generated and human-written Sinhala answers. A proposed model is presented that enables the recognition of real Sinhala answers written by humans and those generated by AI. This help to prevent cheating in academic settings while overcoming shortages of resources for the Sinhala language study. A stepwise methodology is used, with data gathering at the first stage, which includes 1000 questions from academic areas of history, science, business studies and Buddhism with human-provided answers, and artificially intelligent responses. The text pre-processing like stemming, tokenization and elimination of stop-words are imposed on the data. Term frequency-inverse document frequency transforms the textual data into numerical forms that can be fed to actual learning algorithms. Then, the two algorithms were used such as Artificial Neural Networks (ANN) and the Long Short-Term Memory (LSTM). Based on the results LSTM with 86% accuracy out performs the accuracy of the ANN, therefore can be conclude to LSTM is better than the ANN. As well as the recall, F1-score and error values better in LSTM. Different hyperparameters and percentage split used for the evaluations. Data collection and computer issues are few challenges faced during this research. This research provides a solution for cheat detection in low frequent languages. It forms the basis for subsequent work that aims to detect content in AI underrepresented languages; further work will use cosine similarity to explore the relationship between lecturer and AI responses.

Keywords: *AI-generated, Classification, Deep learning, Human-written, Sinhala language*

Influence of Demographic Factors on Emotion-Driven User-Centered Design in Mobile Applications

Dias J.B.R.N. [*], Lekamge L.S. and Kaushalya P.K.D.K.

Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka
[*]jbrndias@std.appsc.sab.ac.lk

This study explores the impact of demographic characteristics including age, gender and culture on affective user centered design of mobile applications. Emotion-based design seeks to elicit a particular emotion so as to improve user satisfaction, interaction, and commitment. However, in the current literature, these demographic attributes are still investigated singly, and there is comparatively little knowledge on the ways in which these factors may interact to impact the emotion engagement in app design. The research proposed here will fill this gap by exploring the manner in which demographic factors influence emotional reactions toward specific design features such as color, spatial organization, fonts, and controls. The research goals are twofold, namely, to understand specific patterns of design preferences depending on the demographic characteristics of users and to assess overall levels of emotional engagement among the sample; finally, given the identified patterns of design preferences and levels of engagement, to formulate references for designing affective apps for users of different demographic backgrounds. This study applies a combination of qualitative and quantitative research methods for data collection. The participant selection seeks diversity among users of different age groups starting from young adults up to older users together with participants from diverse genders and cultural backgrounds for broad applicability in the research results. The collection of qualitative data will occur through thorough interviews that enable participants to share their emotional reactions to mobile app interface features. The surveys implemented to study user satisfaction and interface emotional connection will generate quantitative measurements for this research. The project will produce specific recommendations for middle-aged user interface design and enhance the development of interactive user interfaces suitable for different audience demographics.

Keywords: *Emotional engagement, demographic diversity, design optimization, user interaction.*

Consumer Preferences of Gamification Elements in Sri Lankan E-Commerce Websites: A Study of Preferences and Effectiveness

Duleesha G.R.*, Kaushalya P.K.D.K. and Lekamge L.S.

Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka
**duleeshagamalathge@gmail.com*

Gamification has become a tool used by most e-commerce websites to enhance engagement, motivation, and satisfaction through game techniques, such as Points Systems, Badges/Achievements, Leaderboards, Spin the Wheel, Quizzes/Challenges, Loyalty Rewards, and Referrals. This research aims to explore consumer preferences for these seven gamification elements in Sri Lankan retail, telecommunications, banking, and transportation e-commerce platforms, focusing on their effectiveness and ethicality. Qualitative data were gathered through surveys of over 100 Sri Lankan e-commerce users to gain further insights. The collected survey data were analyzed using SPSS. Descriptive statistics were employed to summarize demographic data and basic response trends. Correlation analysis was used to explore relationships between gamification features and consumer behavior. Regression analysis assessed the predictive power of specific gamification elements on user engagement and satisfaction. Administrative interviews and focus group discussions provide quantitative data to understand user engagement, perceptions, and ethical attitudes. Moreover, case studies of Sri Lankan e-commerce platforms illustrate the impact of gamification solutions. The results show that Sri Lankan users are mostly engaged with retail websites. When considering the gamification elements, they mostly prefer point systems, loyalty rewards, and spin-the-wheel promotions respectively. Considering the ethical concerns, results show a higher percentage of users feel that some gamification features are manipulative or deceptive while most users experience misleading promotions within gamifications. The results are anticipated to provide Sri Lankan e-commerce organizations with guidelines to customize and enhance gamification while steering clear of unethical practices. This study helps to bridge the knowledge gap regarding gamification's cultural and economic significance in emerging markets, which can inform future innovations.

Keywords: *Consumer Preference, Customer Satisfaction, e-Commerce, Ethical Marketing, Gamification*

Image-Based Skin Cancer Detection Using Machine Learning with Ayurvedic Treatment Recommendations

Abeygunawardhana D.K. [*], Lekamge L.S. and Wijerathna R.M.K.K.

Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka
[*]dkabeygunawardhana@std.appsc.sab.ac.lk

Early diagnosis and accurate identification of skin cancer together with dermatological diseases represent major worldwide health issues that require precise medical assessment for proper therapeutic interventions. A diagnostic model development project focuses on skin cancer identification by examining skin lesion images with high precision. The research uses MobileNetV2 as an advanced deep learning architecture to analyse image-based diagnostic models that detect different types of skin cancer. The study maintains uniqueness through its unification of machine learning techniques for skin cancer recognition with Ayurvedic remedy guidance which stems from Sri Lankan traditional medical practices. The research leverages a dermatology dataset which includes diverse 27200 labelled images of skin lesions under 10 skin diseases to support training data representation. The classification performance evaluation of MobileNetV2 outperforms ResNet50 and other models by achieving better accuracy and speed in model convergence. Experimental findings indicate that MobileNetV2 delivers 64.54% accuracy performance beyond CNN but demonstrates similar results to ResNet50 at 62.97% accuracy. The study connects the developed model with Ayurveda treatment suggestions using a rule-based system to develop an expanded method for complete skin cancer healthcare.

Keywords: *AI in Healthcare, Ayurvedic Medicine, Deep Learning, Medical Image Analysis, Skin Cancer Detection*

Analyzing the Impact of Audio Features on Popularity in Global Music Consumption using Machine Learning: A Spotify Case Study

Wanniarachchi W.A.S.Y[1*] and Nirubikaa R[2]

[1]*Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka*
[2]*Department of Software Engineering, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka*
[*] *wasywanniarachchi@std.appsc.sab.ac.lk*

The rapid growth of music streaming platforms such as Spotify has transformed how the audience consume music globally. This research analyses the influence of audio features on the song popularity in global music consumption using machine learning utilizing a Spotify dataset of around 2000 records. The audio attributes used for analyzation are acousticness, danceability, energy, instrumentalness, liveness, loudness, speechiness, valence, and tempo of the audio track along with the track popularity metrics, genres and release dates. The research is aimed to identify patterns and factors that contribute to a song's success. The research significance is that its potential to bridge the gap between the data-driven insights and global music consumption. Logistic Regression, Random Forest and Artificial Neural Networks are employed to predict song popularity and evaluate the importance of individual features driving audience engagement. The methodology involves data gathering, preprocessing, feature engineering, algorithm selection, model training and model evaluation. Data preprocessing phase involves handling missing values, normalizing numerical features and encoding categorical variables to ensure consistency. Feature engineering includes selecting key audio attributes, creating new derived features and applying dimensionality reduction techniques to enhance model performance. For each algorithm, precision, recall, F1-score and accuracy values were computed. Accuracy values for Logistic Regression, Random Forest and ANN are 93%, 92% and 90% respectively. The findings offer insights into how specific audio characteristics influence listener preferences and provide a model which supports the artists, producers, and streaming platforms to optimize song releases for high popularity ratings. The model incorporates global trends by analyzing playlist genres and listener engagement to reveal how audio features influence music consumption. This research contributes to the growing music industry and highlights the potential of machine learning to reveal patterns in global music consumption. The results help stakeholders make informed decisions in music production, curation, and marketing.

Keywords: *Audio Attributes, Global Music Consumption, Machine Learning, Popularity*

Exploring the Correlation Between Social Media Activity & Anxiety Levels Among Students: A case study for undergraduates in Sabaragamuwa University of Sri Lanka

Udeshika P.K.M.[1*] and Erandi J.D.T.[2]

[1]*Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka*
[2]*Department of Data Science, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka*
[*]*pkmudeshika@std.appsc.sab.ac.lk*

Social media use by university students raises concerns about mental health, particularly anxiety. While few studies in South Asia address this issue comprehensively, behavioral aspects are often neglected. This study explores the correlation between social media use and anxiety among undergraduates at Sabaragamuwa University of Sri Lanka, focusing on anxiety-inducing usage patterns such as social media checking frequency, usage during study periods, impact on sleep, Fear of Missing Out (FOMO), comparison with others, prioritization of social media, and discomfort when unable to access social media. A sample of 380 undergraduates, stratified by academic year, was selected from a population of 7,022. A pilot survey of 100 students validated the questionnaire. Social media use, the independent variable, was assessed alongside anxiety levels (minimal, mild, moderate, and severe) using the General Anxiety Disorder-7 (GAD-7) scale. Reliability analysis confirmed the dataset's validity (Cronbach's Alpha = 0.878). Descriptive statistics, Pearson, and Spearman correlations revealed significant relationships. A weak positive association was found between social media checking frequency and anxiety levels ($r = 0.501$, $p = 0.028$). Checking social media upon waking ($r = 0.559$, $p = 0.004$) and its impact on sleep ($r = 0.580$, $p = 0.002$) were associated with greater anxiety. Social media usage during study periods also correlated positively with anxiety ($r = 0.539$, $p = 0.008$). Descriptive statistics showed 47.5% of students had minimal anxiety, 34.2% mild, and 18.3% moderate. The highest anxiety level (42.5%) was observed in students using social media for 2–4 hours daily, while the lowest was in users with under 1 hour. This study demonstrates that frequent checking, use during study, and sleep disruption due to social media increase anxiety. These insights provide compelling evidence for interventions promoting healthier social media habits to mitigate mental health impacts among students.

Keywords: *Anxiety, GAD-7 Scale, Mental Health, Social Media, University Students*

Identification of the Factors Affected to User Satisfaction in Sri Lankan E-commerce Websites

Ananda K.V.A.I.[1*], Chathumini K.G.L.[1] and Erandi J.D.T.[2]

[1]Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka.
[2]Department of Data Science, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka.
[*]kvaiananda@std.appsc.sab.ac.lk

The rapid expansion of e-commerce in Sri Lanka has transformed consumer behavior, yet many local platforms struggle to meet user expectations. Understanding the key factors influencing user satisfaction is crucial for improving online shopping experiences and fostering business growth. This study investigates user satisfaction with Sri Lankan e-commerce websites through a structured online survey, collecting 706 responses from online shoppers. The research examines seven key constructs: website usability, product attributes, payment methods and security, delivery efficiency, customer service, cultural and localization factors, and consumer protection and trust. A quantitative methodology was employed, including reliability and correlation analysis. Reliability analysis yielded a Cronbach's alpha of 0.85, confirming high internal consistency among the constructs. Pearson correlation analysis showed that all independent variables significantly influenced user satisfaction (p-value < 0.001). Among these, website usability (r = 0.73) was the most influential factor, followed by product attributes (r = 0.57) and consumer protection and trust (r = 0.55). The findings also highlight the lesser yet notable influence of cultural and localization factors (r = 0.43), emphasizing the importance of adapting e-commerce platforms to local consumer preferences. Future work will further analyze these relationships through exploratory factor analysis and regression modeling to quantify the impact of each factor. Additionally, predictive models will be developed to identify the best determinants of user satisfaction, ensuring robust and data-driven recommendations for Sri Lankan e-commerce businesses. This study contributes to both academic literature and practical strategies, providing actionable insights to enhance website usability, product diversity, payment security, and consumer trust. Sri Lankan e-commerce platforms can enhance satisfaction, loyalty, and competitiveness by addressing this key user experience factors.

Keywords: *Consumer Trust, E-commerce, Online Shopping, User Satisfaction, Website Usability*

Enhancing Financial Retrieval Augmented Generation with Knowledge Graphs for Better Contextual Insights

Ranasinghe R.M.H.K.[1*], Herath G.A.C.A.[1] and Kumara B.T.G.S.[2]

[1]Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka
[2]Department of Data Science, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka
[*]hashan.ranasinghe98@gmail.com

In today's rapidly evolving financial domain, effective information retrieval and analysis are crucial for making well-informed decisions. This study focuses on constructing a domain-specific knowledge graph (KG) tailored for financial data retrieval, aiming to overcome challenges such as the time-intensive nature of traditional processes, hallucination in large language models (LLMs), and knowledge cut-off limitations. The related works have explored improving the efficiency of RAG pipelines in finance by enhancing retriever accuracy, reducing hallucinations, and increasing context relevance through techniques such as fine-tuning and re-ranking. However, none have focused on employing a KG to achieve these objectives. In developing the finance domain-specific KG, an LLM is employed to extract entities and relationships from unstructured financial news articles. The extracted data was structured into a graph format, stored in a Neo4j database, and was queried using Cypher queries. This approach increased the potential to recover related information while maintaining scalability and a stable accuracy level. The resulting KG offered a concise and insightful overview of the market, highlighting key entities and their interconnections. The accuracy of the KG is evaluated by comparing its extracted entities and relationships with validated financial articles. The KG comprises 858 nodes and 1,624 edges, with a density of 0.0022 and an average node degree of 3.79. A clustering coefficient of 0.1447 indicates moderate interconnectedness, while 12 node types and 47 relationship types ensure diversity and a relationship type entropy of 4.2073 reflects its complexity. While future work will focus on the seamless integration of the developed KG with RAG systems to enhance LLMs' contextual reasoning and response accuracy, the current findings already demonstrate that KGs significantly improve semantic search by reducing ambiguity and enhancing the quality of retrieved financial information.

Keywords: *Knowledge Graphs, RAG, LLM, Financial Data, Contextual Insights*

Integrating Federated Learning with Blockchain for Secure and Efficient Vehicle-to-Vehicle (V2V) Communication

Kumara A.V.A.N. P.[1*], Vigneshwaran P.[2] and Jayasena K.P.N.[3]

[1]*Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka*
[2]*Faculty of Science and Technology, Charles Darwin University Casuarina, Australia*
[3]*Postdoctoral Research Fellow, Technische University Dresden, Germany*
[*]*avanpkumara@std.appsc.sab.ac.lk*

The future of autonomous vehicles depends on secure, efficient, and trustworthy Vehicle-to-Vehicle (V2V) communication systems. Traditional centralized architectures face critical limitations, including cybersecurity risks, data privacy vulnerabilities, and high latency, which hinder real-time decision-making in dynamic automotive environments. To address these challenges, this work proposes a decentralized framework integrating Federated Learning (FL) with blockchain technology**,** designed specifically for privacy-preserving, low-latency V2V communication. A custom deep neural network (DNN) contains two hidden layers accommodating 64 followed by 32 neurons while including batch normalization features and dropout regularization for local vehicle training. The FL paradigm allows vehicles to train the model collaboratively without exposing raw data while simultaneously minimizing their network utilization. The global model aggregation combines weighted averaging techniques that utilize local dataset sizes to achieve 88.89% accuracy during synthetic V2V dataset collision prediction tasks. Real-time responsiveness is achieved through this framework which cuts communication delays by 30% as compared to centralized systems by using optimized parameter exchange along with parallel processing. Blockchain integration addresses trust and security gaps: model updates are hashed using SHA-256 and immutably stored via a lightweight Practical Byzantine Fault Tolerance (PBFT) consensus mechanism, enabling auditability while maintaining low overhead. Batch processing of transactions further mitigates blockchain's inherent latency, ensuring compatibility with time-sensitive V2V operations. Future work will deploy a permissioned blockchain network using smart contracts to automate model validation and mitigate adversarial attacks. The framework will be tested on real-world vehicular datasets to evaluate scalability in urban traffic scenarios. This work advances intelligent transportation systems by enabling secure, decentralized collaboration between vehicles, directly enhancing road safety and traffic efficiency.

Keywords: *Federated Learning, Blockchain, V2V Communication, Privacy Preserving, Intelligent Transportation Systems*

Software Engineering Track

Prediction of Sprint Delivery Capability in Agile Software Development Using Machine Learning

Lorshan N.[1*], Wijerathna P.M.A.K.[2] and Kumara B.T.G.S.[3]

[1] *Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka*
[2] *Department of Software Engineering, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka*
[3] *Department of Data Science, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka*
[*] *lorshannarayanasamy@gmail.com*

Scrum plays the dominant framework in agile software development. It delivers the project in an iterative and incremental manner under main concept called sprint. Tracking whether a sprint is capable of delivering its issues on duration is called sprint delivery capability which helps to enhance sprint planning and deliver the software on time. Previous works implemented prediction models using Machine Learning (ML) approach to forecast sprint delivery, but it lacked to capture temporal, hidden patterns and non-linear relationships. To tackle these problems, this research used ensemble learning approach by leveraging the strengths of different ML, Deep Learning (DL) and ensemble algorithms with enhanced feature aggregation and feature selection to fill the gap to improve the prediction of sprint delivery capability in different project progression levels. The data were collected from the JIRA open-source project, consisting of 1873 iterations and 10852 issues in tabular format at 30%, 50%, and 80% project progression levels, was utilized with data preprocessing to clean the data. Feature engineering enabled to improve the performance and interpretability of the model including, feature aggregation and feature selection. The feature aggregation was performed by leveraging statistical and clustering approaches on the features of the issue table to derive new features and merge them with the iteration table. The hybrid feature selection extracted the key features influencing the sprint delivery capability using statistical and ML techniques. The author developed the ensemble models using ensemble methods such as, bagging, boosting and stacking that helped to increase the prediction accuracy. The results demonstrated that 80% progression level provided more accuracy in all the models and Categorical Boosting (CatBoost) algorithm gained highest accuracy of 96.8%. The prediction made on different project progression levels supported to get more insights about sprint performance, identify potential risks and efficiently allocate resources throughout the project lifecycle.

Keywords: *Sprint Delivery Capability, Ensemble Learning, Feature Aggregation, Hybrid Feature Selection*

Evaluating Cost-Effective Consensus Algorithms to Enhance Blockchain Scalability in Low-Resource Education Systems

Udeshika W.A.[1*] and Wasalthilaka W.V.S.K.[2]

[1]Department of Computing & Information Systems, Faculty of Computing, Sabaragamuwa University, Sri Lanka
[2] Department of Software Engineering, Faculty of Computing, Sabaragamuwa University, Sri Lanka
**ayomiudeshika19@gmail.com*

The issues such as the fake credentials, delays in verification, and vulnerable centralized data have increased the demand for a safe and tamper-proof academic credentialing system. Blockchain is a great solution, but conventional consensus algorithms involve resource-intensive traditional consensus algorithms, which generally require high computational power and energy consumption, making it unsuitable for low-resource settings. This research investigates cost-effective consensus algorithms to enhance blockchain scalability while ensuring security and efficiency. The methodology consists of preparation of a dataset of blockchain-based academic records, data preprocessing, and extraction of key performance metrics such as TPS(Transaction Per Second),and block latency. It evaluated the existing performance and scalability issues in academic systems and consensus algorithms and identified Proof of Stake, Delegated Proof of Stake, Proof of Authority, and Practical Byzantine Fault Tolerance as suitable consensus algorithms for low-resource settings. Because they are comparatively low resource intensive, cost-effective and fast. A private Ethereum blockchain environment had been emulated using Ganache, where an analysis of consensus algorithms considering different parameter. PoA achieved the highest TP with minimal latency and other relevant parameters also should be evaluated to get the more accurate decision. As compared to the real-world blockchain implementations, higher throughput was generated in the simulated environments due to the absence of consensus overhead; therefore, extended real-world validation is required. The study contributes to developing scalable and cost-effective blockchain solutions for academic credentialing in low-resource settings, providing insight into the performance-security-affordability trade-offs.

Keywords: *Academic Systems, Blockchain, Consensus Algorithms, Low-Resource Systems, Scalability*

Machine learning based detection of software vulnerabilities in C code

Junoj S.[1*], Wijeratne P.M.A.K.[2] and Kumara B.T.G.S.[3]

[1]Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka
[2]Department of Software Engineering, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka
[3]Department of Data Science, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka
[*]Junojsiva1998@gmail.com

Software security vulnerability detection is an integral part of creating secure and reliable software. C programming is used extensively in system-level and embedded applications for efficiency and direct control over hardware resources without inherent security features, thus being especially vulnerable to common categories of attacks, including buffer overflow and null pointer dereferences. Classic security vulnerability detection methods based on manual code reviews and static analysis tools are unable to discover complicated security bugs. Given the shortcomings of previous approaches, this research work presents a machine learning-based approach to automate the detection and classification of vulnerabilities in C code. Datasets were gathered from Kaggle and IEEE DataPort, consisting of real-world samples of C code that are quite varied. Feature extraction was performed with the Word2Vec model, which is more powerful than traditional frequency-based methods for capturing semantic and contextual relationships in code. Various machine learning and deep learning models have been explored in the research: Logistic Regression (LR), Support Vector Machine (SVM), Decision Tree (DT), Random Forest (RF), Convolutional Neural Network (CNN), and Long-Term Short Memory (LSTM). Later on, a hybrid CNN-LSTM model is suggested for better results. These models were then developed, trained, validated, and tested using the 80-10-10 split, evaluated based on the accuracy, precision, recall, and F1-score. These results show that the Decision Tree model had the highest accuracy of 93.46% in vulnerability detection, while the hybrid CNN-LSTM model performed best in classification with an accuracy of 94.55%. These results prove that machine learning significantly enhances software vulnerability detection compared to traditional methods. The study further elucidates how these models can be integrated into real-world software development workflows and improve automated security assessments. Future studies should compare this approach to state-of-the-art vulnerability detection frameworks in order to further tune the machine learning-based security solution.

Keywords: *Convolutional Neural Network, Vulnerability detection, vulnerability classification, Machine learning, Deep learning*

Enhancing Software Quality through Comparative Analysis of Machine Learning Techniques for Test Case Prioritization using Object-Oriented Metrics

Kumari M.K.M.N.[1*] and Wasalthilaka W.V.S.K.[2]

[1]Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka.
[2]Department of Software Engineering, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka.
[*]nadeeshamayuri15@gmail.com

Software quality plays an important role in software engineering by maintaining system reliability and ensuring efficiency. Achieving high-quality software mainly depends on software testing. Regression testing is important to this process, but it is very time-consuming and resource-intensive. Test case prioritization (TCP) techniques can optimize this process. It reduces test time and optimizes resource usage. Conventional TCP mechanisms, like coverage-based and risk-based prioritization, have limited ways of handling software structures. This study compared different machine learning algorithms like Decision Tree, Random Forest, Neural Networks, K-Nearest Neighbor, and Logistic Regression to identify the best technique for TCP using object-oriented metrics like Coupling Between Objects (CBO), Weighted Methods per Class (WMC), Depth of Inheritance Tree (DIT), Number of Children (NOC), and Lack of Cohesion in Methods (LCOM). Used a dataset from Zenodo, which includes 232,468 observations and 53 attributes for this research. Each observation in the dataset represents a software file, class, or method. The dataset is divided into training and testing sets using data preprocessing and feature selection methods. Models are trained using the training dataset and tested using the testing dataset. Then, the trained models were evaluated using performance metrics like accuracy, precision, recall, and F1-score. Finally, each model is compared using evaluated results to get the best-performing model. The decision tree outperformed others in TCP due to its ability to manage decision boundaries with minimal overfitting. It achieved 71% accuracy, reducing the execution time for testing by 32.5% and improving the detection of errors by 15.8% over the traditional methods. This result highlighted their huge potential to increase regression testing efficiency and software quality.

Keywords: *Machine Learning, Object-Oriented Metrics, Software Quality, Test Case Prioritization*

Classifying Code Quality in Java Open Source Software Projects using Machine Learning and Contribution Metrics.

Welagedara H.T.[1*] and Wasalthilaka W.V.S.K.[2]

[1]*Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka.*
[2]*Department of Software Engineering, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka.*
[*]*hashiniwel@gmail.com*

Geographically dispersed volunteer teams can achieve collaborative and transparent processes with Open-Source Software Development (OSSD). While it outperforms traditional methodologies, challenges remain in preserving code quality, managing third-party dependencies leading to compatibility issues, and inconsistencies in developer contributions that can lead to code redundancies. Java as the foundation of software development, has fostered numerous open-source projects, enhancing research dependability. This study proposes a machine learning model that classifies code quality in Java-based open-source software projects by analyzing code contributions. Popular machine learning techniques used for software quality prediction, such as Regression, Decision Trees, Random Forest, Support Vector Machine and Bayesian Learning are used, as well as established software quality metrics to measure the developer's contribution using source code such as Lines of Code (LOC), Coupling Between Objects (CBO), Response for a Class (RFC), Weighted Methods per Class(WMC), Lack of Cohesion in Methods (LCOM), Depth of Inheritance Tree (DIT) and Number of Children (NOC). The proposed model was evaluated using a dataset, containing over 200,000 observations and 53 software metrics extracted from open-source projects. Performance was measured using Precision, Accuracy, Recall and F1-Score. Decision Tree and Random Forest currently show the highest model accuracy, with 58%. Neural networks weren't included due to their high computational cost and limited interpretability. This analysis enhances Java OSSD projects by accurately evaluating code contributions, ensuring reliability and sustainability. Refining code review, prioritizing refactoring, and leveraging the best ML approach to predict code quality can strengthen development processes and advance OSSD efficiency.

Keywords: *Code Contributions, Code Quality, Machine Learning, Open-Source Software Development, Software Quality Metrics*

Emotion Driven Prediction of Conversational Thread Depth

Kumara H.K.M.H.V.[1*], Weerakkodi S.[2] and Kumara B.T.G.S.[3]

[1]*Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka*
[2]*Colombo Regional Centre, The Open University of Sri Lanka, Sri Lanka*
[3]*Department of Data Science, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka*
[*]*hkmhvkumara@std.appsc.sab.ac.lk*

While chatbots have become enhanced in human-computer interaction, they hardly provide quality and interesting interactions. This study emphasizes on conversational thread depth prediction, which quantifies the ability of a conversational thread to stay productive and valuable with a text-only emotion-based flow. Conversational thread depth is determined using the emotional features taken from textual data, then used to enhance chatbot responses for their appropriateness to the context or sensitivity to the emotions being expressed. The model developed under this study makes use of feature extraction through the use of Keras and Word2Vec, and implemented using Long Short-Term Memory (LSTM) with attention mechanism. This enables the model to detect more important affective features of text, thus making the model better suited to predicting the depth of conversation. Given experimental findings, it is noted that the text-only scenario is rather efficient, with training accuracy equal to 80% and validation accuracy increased to 76%. This means that the model can generalize well given the limitations of text-based nature of the data but some of the issues such as the problem of over fitting and accommodating for variety in patterns of human conversational exchanges. Chatbot engagement and automated customer service are the real-world uses in this study. Although this study is limited to the text-only mode, the current basis suggested for subsequent additions and improvements. Future studies will use audio-visual information to extend the range of rankings and broaden the emotions and context levels.

Keywords: *Chatbot Interaction, Conversational AI, Deep Learning, Emotion Recognition, Thread Depth.*

Software Defect Prediction in Agile Environments using Deep Learning

Pathirana P.U.N.[1*] and Wasalthilaka W.V.S.K.[2]

[1]Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka.
[2]Department of Software Engineering, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka.
[*]uthpalanp@gmail.com

Software defect prediction is a critical part of development, identifying mismatches between expected and actual outcomes as detected by developers or end users. The main purpose of agile defect prediction is identifying defects in timely manner. Because it helps teams to prioritize tasks, allocate the resources such as time, effort and personnel to the riskiest parts of the code, enhance the sprint planning and improve the software quality by supporting continuous development. But the iterative and fast paced nature of agile environments raises few challenges for the defect prediction such as handling code changes, managing limited development time and addressing the dynamic and evolving project requirements. So, there is a notable gap related to the research studies of agile defect prediction. Traditional defect prediction methods frequently struggle to identify dynamic and complex data patterns. This study proposes a solution by using deep learning models so agile teams can use them to detect defects early, optimize testing and improve real-time software quality of the projects. For the research, Convolutional Neural Network (CNN), Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM) and Deep Belief Network (DBN) models are used which can identify complex data patterns and relationships by extracting the meaningful features automatically. Jira defect dataset was used and cleaned using data pre-processing and feature selection techniques, and the processed dataset was divided into training and testing sets. The trained models were evaluated using metrics like accuracy, precision, recall and f1-score. The study exposes LSTM outperforms other models with 76% accuracy, handling long-term dependencies and processing historical defect logs. The findings of the research emphasize the effectiveness of using deep learning models in agile software defect prediction for high-quality, reliable real-world agile development practices.

Keywords: *Agile, Deep Learning, Software Defect Prediction, Software Development*

AI-Driven Time-Series Forecasting in JIRA to Enhance Risk Mitigation and Predictive Insights for Agile Project Management

Samarakoon S.M.T.D.[*] and Dangalla R.L.

Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka.
[*] *thakshilad2018@gmail.com*

Agile project management prioritizes adoptability and collaboration, continues to face significant challenges in risk management, dynamic requirements, resource allocation, and project delays. A major concern is the inconsistency in meeting project deadlines, which results in inefficiencies and cost overruns. This research proposes an AI-driven theoretical framework for time-series forecasting, specifically tailored for JIRA, aiming to mitigate delays and enhance project predictability. The study employs a structured approach, commencing with data collection through surveys directed at project managers, team leaders, developers, QA engineers, business analysts, and UI/UX designers to identify critical risks inherent in agile environments. A comprehensive risk analysis is conducted to evaluate prevalent project challenges, thereby guiding the selection of an appropriate AI model for forecasting. The ARIMA model is developed and integrated with the JIRA API, which facilitates automated risk predictions and supports proactive decision-making. The results derived from this integration contribute to the refinement of the proposed theoretical framework, which is further validated by experts in the field. By utilizing predictive analytics, particularly through the ARIMA model, the framework analyses historical project data, forecasts delays in sprints, identifies resource bottlenecks, and anticipates trends in task completion. These AI-powered insights enable project managers to proactively manage risks, enhancing both sprint planning and resource distribution. Additionally, the study evaluates the framework's applicability beyond JIRA, enhancing its relevance to a wider array of agile tools. Findings indicate that AI-driven forecasting enhances the predictability of sprints, reduces bottlenecks, and improves decision-making processes within agile methodologies. This research contributes to the evolution of agile practices by offering a scalable, AI-driven approach to risk mitigation, adaptable to various project management contexts. The proposed framework seamlessly integrates traditional agile methodologies with AI-driven automation, offering practical recommendations for incorporating AI-based forecasting into modern software development.

Keywords: *AI-Driven Forecasting, Risk Mitigation, JIRA, ARIMA Model, Sprint Planning.*

Explainable Artificial Intelligence Approach for Agile Story Point Estimation and Issue Type Prediction

De Silva H.M.C.J.[1*], Wijerathna P.M.A.K.[2] and Kumara B.T.G.S.[3]

[1]*Department of Computing & Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka*
[2]*Department of Software Engineering, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka*
[3]*Department of Data Science, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka*
[*]*hmcjdesilva@std.appsc.sab.ac.lk*

Agile Software Development (ASD) relies on accurate story point estimation and issue type prediction for effective sprint planning and resource allocation. Despite advancements in Machine Learning (ML), existing approaches lack interpretability and have not yet addressed both tasks together. This study fills this gap by proposing two specialized deep learning models: a regression model for story point estimation (0-20) and a classification model for the prediction of the issue type as bug, story, or task. Both models were trained on a subset of the TAWOS dataset with 65,427 issue reports. Issue titles and descriptions were combined and used as input. Then these textual inputs were converted into vector representations using Word2Vec embeddings to achieve feature extraction. To address the class imbalance in the classification task, dynamic augmentation using BERT-based contextual substitutions was applied. Bidirectional Long Short-Term Memory (BiLSTM) networks were selected after evaluating several other models, including Random Forest, XGBoost, Support Vector Machine, and Logistic Regression. Compared to other traditional ML models, BiLSTM demonstrated better performance. To enhance the interpretability of the models, we incorporated Local Interpretable Model-agnostic Explanations (LIME). This method provides transparency by offering insights into which specific words most influenced the predictions for each issue, both for the regression and classification tasks. The regression model achieved a mean absolute error of 2.11 and a root mean square error of 3.46. The classification model achieved 84% accuracy, with F1 scores of 0.85 for bugs, 0.87 for stories, and 0.71 for tasks. For future work, we propose integrating these models into a unified model. This research fills a gap in ASD practices by introducing an explainable approach that aligns with agile industry norms and fosters trust among practitioners.

Keywords: *BiLSTM, Effort estimation, Interpretability, Issue report*

Automating Software Documentation: Leveraging AI, NLP, and DocOps for Real-Time Integration in Agile Development

Wijethunga W.A.T.G.R.[*], Dayarathna K.O.K.S., Perera R.D.M., Fernando M.F.P.S.L. and Ravikumar N.

Department of Software Engineering, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka
[*]tharushigrw@gmail.com

The increasing complexity of modern software systems highlights the critical role of documentation in ensuring maintainability, knowledge transfer, and onboarding efficiency. However, documentation is often deprioritized due to its time-consuming nature, leading to incomplete records. This research presents an AI-powered framework that automates software documentation using Artificial Intelligence (AI), Natural Language Processing (NLP), and Documentation Operations (DocOps) to ensure accuracy and relevance updates. The system is trained on a predefined template, learning document structure, formatting, and content organization. Once trained, it continuously generates documentation following the given template, dynamically adapting to changes. The framework integrates advanced NLP techniques, large language models (LLMs), and abstract syntax trees (ASTs) to extract meaningful insights from **co**de, user interactions, and system log**s.** These components were chosen over traditional rule-based methods due to their superior ability to analyze context and generate structured documentation. Seamlessly embedded within CI/CD pipelines and agile workflows, the system ensures that documentation evolves alongside software iterations, automatically detecting gaps, personalizing content based on user roles, and securing proprietary data. It is designed to be scalable and developer-friendly**,** enabling easy integration into diverse software development environments, with applications beyond software engineering, such as healthcare and finance. To validate its effectiveness, a prototype was tested with software engineers, technical writers, and project managers**,** ensuring relevance across different roles. The evaluation, based on documentation accuracy, content personalization, usability, and efficiency, demonstrated improved consistency in documentation, better alignment with project updates, and reduced cognitive load for developers by minimizing manual documentation efforts**.** User feedback highlighted the system's intuitive content organization, ability to generate role-specific documentation, and seamless adaptation to evolving project requirements**.** This AI-driven approach not only automates repetitive tasks but also enhances collaboration, reduces technical debt, and ensures that documentation remains accurate, accessible, and adaptable to evolving software requirements, transforming industry best practices.

Keywords: *Artificial Intelligence (AI), Natural Language Processing (NLP), Documentation Operations (DocOps), Continuous Integration/Deployment (CI/CD), Software Documentation, Scalability, Developer Experience*.

Adaptive Behavioral Analysis Systems (ABAS) for Detecting URL Phishing in LLM-Based Environments

Wanigasooriya W.A.D.A.U.* and Dangalla R.L.

Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka
**anubwabt414@gmail.com*

Large Language Models (LLMs) have transformed various industries by enhancing natural language processing applications. However, their capabilities also introduce cybersecurity vulnerabilities, particularly regarding phishing attacks. Cybercriminals are using LLMs to create highly deceptive phishing URLs, making it increasingly challenging for existing detection systems to keep pace with evolving attack strategies. To address these challenges, this study introduces an Adaptive Behavioral Analysis System (ABAS) specifically designed to detect phishing URLs in LLM-based environments. ABAS combines behavioral analytics with URL feature extraction and employs preprocessing techniques such as cleaning, normalization, and tokenization to identify meaningful patterns in URLs. The model is trained and validated on a dataset of 50,000 legitimate and phishing URLs, ensuring its adaptability to real-world phishing threats. Experimental evaluations show that ABAS achieves an accuracy of 96.4%, outperforming current phishing detection systems. The results highlight ABAS's capacity to dynamically adapt to evolving phishing tactics, providing a robust and efficient defense mechanism against LLM-generated phishing threats. This research not only uncovers vulnerabilities in LLM-based phishing attacks but also contributes to the development of adaptive and resilient cybersecurity frameworks. Future work will focus on further enhancing ABAS by integrating real-time detection and continuous learning capabilities.

Keywords: *Adaptive Behavioral Analysis System, Cybersecurity, Large Language Models, Phishing Detection, URL Analysis*

Data Science Track

Enhanced U-Net Architecture for Analyzing Sinhala Document Layouts and Styles

Hulathdoowage S.K.D.[1*] and Kumara B.T.G.S.[2]

[1]Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka
[2]Department of Data Science, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka
[*] skdilukshi@gmail.com

Document layout analysis is a process of identifying and segmenting various elements within a document. However, accurate digitization and information extraction from documents require effective analysis of complex layouts, particularly in documents with diverse elements such as titles, images, paragraphs, tables, and mathematical expressions. Additionally, comprehensive layout analysis alone is insufficient for document digitization; style analysis plays a critical role in preserving structural and typographical integrity, which is essential for accurate text recognition in Sinhala script. This research proposes an enhanced U-Net architecture for the semantic segmentation of Sinhala document layouts and font styles. A dataset of 600 manually annotated Sinhala document images with 27 labels, including Title Level 1, Title Level 2, Title Level 3, Paragraph, Table, Image, Text Bold, and Text Italic, was used. Furthermore, it improves optical character recognition performance by element-wise integration of optical character recognition technologies, ensuring improved text extraction accuracy. The initial convolutional layers of the U-Net encoder were integrated with a vision transformer block. The input image was divided into patches, which were flattened and processed by the vision transformer block with adaptive positional embedding. The accuracy, precision, recall, and F1-score for the test dataset were 79.27%, 71.12%, 69.85%, and 70.48%, respectively. These modifications enabled the model to capture long-range dependencies and global context in the input images, potentially improving feature extraction. Compared to conventional U-Net models, this approach demonstrated superior segmentation accuracy, particularly in complex document structures. Finally, this study contributes to Sinhala document digitization by providing a comprehensive framework for layout and style analysis, enhancing OCR performance, and offering adaptability for multilingual document processing in real-world applications such as automated archiving and digital library systems.

Keywords: *Document layout analysis, Optical character recognition, Semantic segmentation, U-net architecture, Vision transformers*

Knowledge-Graph Base Document Chunking and Improve Retrieval's Relevancy of RAG Applications with Recursive Counter Pointing Agent

Perera W.P.K.I.[1*], Vigneshwaran P.[2] and Charles J.[3]

[1]*Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka*
[2]*Faculty of Science and Technology, Charles Darwin University Casuarina, Australia*
[3]*School of business, University of Southern Queensland Springfield, Australia*
[*]*kishari1998@gmail.com*

Retrieval Augmented Generation (RAG) has emerged as a transformative approach to address key limitations of Large Language Models (LLMs), particularly regarding hallucination, private data access, and contextual accuracy. Both traditional and current RAG systems, while showing promising results, continue to face significant challenges in maintaining semantic accuracy, optimizing retrieval performance, and adapting to complex knowledge structures. Recent implementations like HyDE (Hypothetical Document Embeddings) and Self-RAG have made progress in addressing these issues, but fundamental limitations persist in handling dynamic knowledge structures and maintaining contextual relevance. In this study we introduced an enhanced RAG framework with three novel components: (1) a flexible schema-less knowledge graph that dynamically adapts to new information, (2) an iterative retrieval verification system with integrated web search capabilities that validates extracted information, and (3) an adaptive query refinement mechanism that progressively improves search accuracy. Our approach integrates a schema-less knowledge graph as the primary retrieval source, enhanced by an agent-based architecture where reviewer and query reformer agents work iteratively. These agents dynamically validate and refine queries, while leveraging web sources as a supplementary knowledge base to compensate for temporal gaps in the primary source. This dual-source, agent-driven architecture enables robust handling of complex queries while maintaining up-to-date contextual relevance. We evaluate our framework using both small-context (≤3,000 tokens, Conversational Question Answering (CoQA) and Document Question Answering (DoQA)) and large-context (≥8,000 tokens, INSCIT and TopiOCQA) datasets following Nvidia's benchmark standards. On CoQA, our system achieves higher recall (0.52 vs 0.47) and precision (0.34 vs 0.29) compared to baseline RAG, while maintaining competitive average precision (0.45 vs 0.47) and perfect reciprocal rank. These improvements demonstrate particular effectiveness in complex, high-dimensional contexts, with practical applications in medical record analysis, financial compliance documentation, legal case retrieval, and real-time customer support systems where accuracy and context preservation are crucial.

Keywords: *Retrieval Augmented Generation (RAG), Large Language Model (LLM), Hallucination, Knowledge Graph, Information retrieval*

Sentiment Analysis of Tourist Reviews for enhancing Sri Lanka's Tourism Industry

Mathusan S.[1*], Adeeba S.[1] and Kumara B.T.G.S.[2]

[1]Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka
[2]Department of Data Science, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka
[*]smathusan@std.appsc.sab.ac.lk

Sri Lanka's tourism industry generates a vast volume of tourist reviews often, providing valuable insights into traveller experiences across various attractions. However, manually analysing these reviews is time consuming and challenging due to their large volume. The research questions are, how can sentiment analysis techniques effectively classify tourist reviews, which Machine Learning (ML) models perform best in sentiment classification and how can the findings be practically implemented to enhance Sri Lanka's tourism sector. The objectives of this research are to develop a robust sentiment analysis framework for classifying tourist reviews and predict the new reviews whether positive or negative or neural, evaluate the performance of various ML and deep learning models, and provide actionable recommendations for stakeholders to improve tourist experiences and promote sustainable tourism growth. Data was collected from TripAdvisor using a web scraping tool. Sentiment labelling is done using the Vader technique to categorize reviews as positive, negative, or neutral. The review was pre-processed and converted into numerical formats using methods like TF-IDF, Bag of Words, and Word2Vec. Several ML and deep learning model including Random Forest, Naive Bayes, Decision Tree, Logistic Regression, Artificial Neural Networks (ANN), and Long Short-Term Memory Networks (LSTM) were used for classification. An ensemble approach combining Random Forest, Naive Bayes, Logistic Regression and Decision Tree achieved the highest accuracy as 89.5%, outperforming individual models. The study aimed to develop a robust sentiment analysis framework for classifying tourist reviews as objective1 and evaluate the performance of various models as objective2. The findings enable stakeholders to effectively identify key sentiment factors, address negative factors, and implement strategies to enhance the tourist experience. For instance, tourism boards can use insights to improve service quality, while businesses can tailor offerings to meet traveller preferences.

Keywords: *Tourist Reviews, Sri Lanka Tourism, Ensemble model*

Automated Performance Analysis in Tennis Using YOLOv8 for Object Detection and Trajectory Tracking

Fernando W.D.P.[1*], Adeeba S.[1] and Kumara B.T.G.S.[2]

[1]*Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka*
[2]*Department of Data Science, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka*
[*] *wdpfernando@std.appsc.sab.ac.lk*

Tennis ranks among the most popular sports worldwide, with over 89 million participants in 2021. Success in tennis requires technical skill, strategy, mental resilience, and peak physical conditioning. Traditional performance analysis relies on manual observation, which is time consuming and requires specialized equipment. Calculating shot speeds and player movements manually is inefficient. This research introduces an automated approach using object detection to quickly and accurately extract performance metrics, reducing manual effort and analysis time. The study collected tennis match footage from YouTube, using a dataset of 578 images from various court angles to detect the tennis ball. Videos were standardized through pre-processing, and annotation was performed using Roboflow to label players and the ball. A separate dataset was used to mark court key points for performance calculations. The YOLOv8 model was trained to detect players, the ball, and key points on the court. YOLOv8 balances speed and accuracy, with an improved anchor-free detection mechanism for better generalization across datasets. It also supports multiple vision tasks, making it a versatile choice. The model successfully drew bounding boxes, tracked trajectories, and identified key points for precise measurements. Frame by frame analysis determined player and shot speeds, providing insights into match intensity and efficiency. Visual inspection confirmed accurate player identification and key point mapping. This research minimizes manual effort in performance analysis and delivers actionable data. The system enables coaches to track player performance, identify weaknesses, and tailor training programs. Players can analyse matches, assess shot efficiency, and refine strategies based on objective data.

Keywords: *Object Detection, Player and Ball Tracking, Sports Data Analytics, Tennis Performance Analysis, YOLO Framework*

Machine Learning Approach for Predicting the Career Paths of IT Undergraduates

Rukais S.L.M.[1*], Kumara B.T.G.S.[2] and Adeeba S.[1]

[1]Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka
[2]Department of Data Science, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka
[*]slmrukais@std.appsc.sab.ac.lk

Career trajectories of Information Technology (IT) undergraduates are crucial for aligning academic programs with industry demands and supporting students' professional aspirations. Rapid changes in the IT industry present both opportunities and challenges for new graduates. Despite the availability of information, many IT undergraduates struggle to secure jobs that match their aspirations and abilities. Traditional career development frameworks often fail to provide accurate predictions due to their inability to address industry-specific requirements, technical skills, and soft skill demands, leading to graduate unemployment and skills imbalances. To address these challenges, this research proposes a machine learning (ML)-based framework for predicting suitable career paths for Sri Lankan IT undergraduates by analysing a combination of these features to tailor career predictions to individual profiles. The study focuses on six career fields: Software Engineering, UI/UX Design, Quality Assurance, Business Analysis/Project Management, Data Science/Artificial Intelligence, and Networking/DevOps/System Administration. Through an extensive literature survey and expert opinions, 49 attributes were identified, including technical skills, soft skills, academic performance, and internship experiences. A dataset of 520 IT professionals was collected, pre-processed, and used to train and evaluate eight ML models: XGBoost, Artificial Neural Network, Decision Tree, Random Forest, SVM, Gradient Boosting, AdaBoost, and an Ensemble model. The Ensemble model, combining Random Forest, Gradient Boosting, and SVC in a stacking approach, achieved the best performance with 91.39% accuracy, 92.47% precision, 91.39% recall, and a 91.48% F1-score. This framework seamlessly integrates into career development systems, such as university counselling platforms and job portals, by enhancing traditional methods with data-driven insights. It highlights the impact of key features, ensuring accessibility for educators, policymakers, and students. By bridging the gap between academia and industry, this research empowers IT undergraduates to make informed career decisions, enhancing job satisfaction and industry alignment.

Keywords: *Career path, Ensemble model, IT Undergraduates, Predictive Modeling*

Automated Fish Freshness Classification Using CNN: A Scalable Solution for Sri Lankan Coastal Regions

Peries R.F.S.[1*], Adeeba S.[1] and Kumara B.T.G.S.[2]

[1]Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka
[2]Department of Data Science, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka
[*]rfsperies@std.appsc.sab.ac.lk

Fish is a crucial component in sustaining human nutrition, serving as a major source of protein throughout history. Ensuring fish freshness is the key determinant of fish quality, consumer health and market value globally. In Sri Lanka, within traditional fish freshness assessment methods, manual sensory evaluations are often subjective and inaccurate due to humanized measurements. Chemical sensors-based evaluation methods are reliable but remain inaccessible to small-scale fishermen and local consumers, due to higher costs and resource limitations. To address this gap, this study employs Convolutional Neural Networks (CNNs) to automate fish freshness classification. Image samples were collected from Mannar coastal regions including Gulf and Island of Mannar and Vankalai Lagoon. The study comprised obtaining diverse dataset of 400 images (200 fresh and 200 non-fresh) per region of interest (ROI) including whole fish, fish eye and fish gill, total 1,200 images (600 fresh and 600 non-fresh). Each sample was manually categorized as fresh or non-fresh while capturing pictures. A singular dataset was created by combining all the images. Then data augmentation increased the dataset to 6500 images (3250 fresh and 3250 non-fresh). Then sequenced preprocessing techniques were applied including resizing, labelling and splitting dataset into training and testing sets in 7:3 ratio. A basic custom CNN model was developed with three convolutional layers with max-polling, dropout and dense layers. Adam optimizer was utilized for training with early stopping to prevent overfitting. The model achieved an impressive testing accuracy 94% along with excellent 94% precision, recall and F1-scores. The confusion matrix and precision-recall curve further validated the model's effectiveness. By providing a scalable, cost-effective solution for fish freshness classification, outperforming traditional methods and holding significant potential to develop mobile applications. Future work will explore larger datasets and real-time implementation in fish markets.

Keywords: *Convolutional Neural Networks, Fish Freshness Classification, Freshness Assessment*

Identifying Duplicate GitHub Issues in Open-Source Repositories Using Deep Learning

Dharmadasa T.K.R.S.[1*], Rupasingha R.A.H.M.[2] and Kumara B.T.G.S.[3]

[1]*Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka*
[2]*Department of Economics and Statistics, Faculty of Social Sciences and Languages, Sabaragamuwa University of Sri Lanka, Sri Lanka*
[3]*Department of Data Science, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka*
[*]*tkrsdharmadasa@std.appsc.sab.ac.lk*

GitHub is a popular platform that is used to maintain software repositories, where users can publish bugs, feature requests, and questions in the form of GitHub issues. Due to the uncoordinated nature of the open-source repositories, which are hosted publicly, there is a huge probability of creating duplicate GitHub issues that may lead to redundant efforts. After manually identifying a duplicate GitHub issue, the standard practice to mark that as a duplicate is to add the corresponding duplicate tag and move it to the closed-issue section. To overcome this manual detection, the study proposes an automated solution that classifies issues as duplicates or non-duplicates using a combination of pre-engineered features and deep learning algorithms. The proposed methodology extracts over 5000 duplicate and non-duplicate GitHub issues using the GitHub Application Programming Interface. Creating the dataset was challenging due to differences in issue reports. After pre-processing the data set by lowercasing, stop word removal, lemmatizing, and tokenizing, it is converted into a machine-trainable format. A feature vector is then constructed using various feature extraction methods such as Term Frequency - Inverse Document Frequency, Word2Vec, SBERT and semantic similarity metrics such as cosine similarity, along with additional features derived from the data. Then the created feature vector serves as an input for different deep learning models including Long Short-Term Memory (LSTM), Convolutional Neural Network, Artificial Neural Network, and Recurrent Neural Network. After evaluating the performance of each algorithm, the LSTM model demonstrated superior results achieving 88% classification accuracy along with the highest precision, recall, f-measure, and lowest error rates compared to other models showcasing its temporal pattern recognition ability. With the proposed approach, duplicate GitHub issues can be automatically detected, reducing manual effort and preventing repositories from accumulating similar issues while allowing for comprehensive discussions on existing ones.

Keywords: *Deep learning, Duplicate detection, GitHub Issues*

Personalized Education: Leveraging LLMs for Dynamic Content Generation and Creating Contextual Quizzes to Enhance Learning Outcomes

Wanigasekara W.M.P.S.B.[1*], Gunawardhana H.M.K.T.[1] and Kumara B.T.G.S.[2]

[1]*Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka*
[2]*Department of Data Science, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka*
[*]*shavinpasan99@gmail.com*

The research study developed an adaptive learning system based on Large Language Models (LLMs) and Retrieval-Augmented Generation (RAG) technology to deliver customized educational content. The system differentiates from traditional LLM educational software by accepting complete lecture materials which ensures quiz responses and feedback match the specific content of the current course. The application retrieves dynamic relevant content from lecture slides to provide focused structured learning beyond standardized pre-trained responses. Pinecone serves as a vector database for semantic content retrieval and open provides GPT for natural language generation from the system architecture. The educational materials undergo SentenceTransformers processing to create semantic embeddings that enable both precise content retrieval as well as contextual adjustments. Specific course materials determine the alignment of quizzes and feedback through this method instead of using pre-existing knowledge as a basis. The first phase concentrates on implementing the system for a single subject to provide detailed refinement of retrieval techniques and quiz adjustment. The system's functionality verifies that it obtains lecture-specific content while producing structured quiz questions that create context-based feedback responses. The evaluation of the system's learning effectiveness through quantitative methods remains an unaddressed task. The forthcoming research tasks will analyse new retrieval approaches alongside optimized content selection methods and different retriever comparison to boost accuracy and adaptability rates. This research integrates RAG retrieval alongside LLM-based generation with direct course materials to demonstrate improved learning results over static and generic educational tools utilizing LLMs. Organized assessments will determine how well the approach improves student engagement while enhancing both understanding and customized learning efficiency.

Keywords: *Adaptive Learning, Large Language Models, Personalized Education, Quiz Generation, Retrieval-Augmented Generation*

Machine Learning Approach of Lung Cancer Prediction using Multi Machine Learning Models

Karunakaran L.[*], Ratnayaka R.M.T.K. and Malkanthi C.

Department of Physical Sciences and Technology, Faculty of Applied Sciences, Sabaragamuwa University of Sri Lanka, Sri Lanka
[*]Klluxshi99@gmail.com

Lung cancer is still one of the leading causes of cancer deaths worldwide. Cancer detection is a complex and challenging process; however, when identified at an incipient stage, it is amenable to curative interventions. Machine learning is one of the most promising artificial intelligence methods widely used in oncological diagnosis and detection of early stages of disease. Thus, this study assesses the application of machine learning to predict lung cancer for symptom-based diagnosis. We introduce a novel approach by evaluating deep learning methods, whereas previous research primarily relied on traditional machine learning models. The numerical dataset in the Kaggle repository was preprocessed to ensure the quality of inputs and the holdout method was used to evaluate the model performance. Various implemented like Logistic Regression, Decision Trees, Gradient Boosting, SVM, ANN, Random Forest, XGBoost, Linear Regression, Naive Bayes, and LSTM. The models were evaluated using Accuracy, Sensitivity, Specificity, and ROC-AUC matrices. The SVM model outperformed all the other models with an accuracy value of 96.98% followed by ANN (96.82%) and LSTM (96.52 %) models. SVM recorded higher sensitivity compared to ANN (98.48%) and LSTM (98.99%). However, SVM and LSTM recorded a lower specificity value of 74.24%, whereas ANN recorded a highest value of 81.81%. The ROC-AUC was highest for both LSTM and ANN (99.32%) while SVM resulted in 98.58%. These results show that machine learning algorithms can classify lung cancer with acceptable accuracy which opens the way toward the improvement of clinical diagnosis. This study emphasizes the usage of machine learning models like SVM in clinical practice to improve the detection rate of early lung cancer as a binary classification task. Advanced machine learning algorithms can be further finetuned and coupled with different cross-validation methods to check the suitability to detect lung cancers in the future.

Keywords: *Lung cancer, Machine learning, Numerical data, Predictive analysis, SVM*

Enhancing Mobile Banking Apps in Sri Lanka: UI or UX Classification on Negative Reviews

Nirosana K.[1*], Prasanth S.[2] and Ishanka U.A.P.[3]

[1]Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka
[2]Memorial University of Newfoundland and Labrador, Canada
[3]Department of Data Science, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka
[*]nirosanakarunanithy@gmail.com

The widespread adoption of mobile banking applications, initially accelerated by the COVID-19 pandemic, has permanently transformed financial transaction patterns. This study analyzed user reviews of popular mobile banking apps in Sri Lanka to evaluate user interface (UI) and user experience (UX) satisfaction. Using topic modelling techniques, we extracted and ranked significant UI and UX-related keywords from Google play store reviews respective to the banking application to identify prevalent user issues. The primary objective was to determine either UI or UX improvements were critical for future updates, improving user satisfaction and engagement. Classification algorithms including K-Nearest Neighbors (KNN), Random Forest, Extreme Gradient Boosting (XGBoost), Categorical Boosting (CatBoost), and Artificial Neural Networks (ANN) were used to classify UI and UX negative reviews based on topic-related keywords. The ANN model achieved 89% accuracy in classification, revealing that UI-related issues dominated negative feedback for a major government banking app. Future work will extend this analysis to other major Sri Lankan banking apps to provide comprehensive insights into specific UI and UX factors affecting user satisfaction.

Keywords: *Mobile Banking, Topic Modelling, User Interface (UI), User Experience (UX)*

Deep Learning Approach for Painting Authentication: Differentiating AI-Generated and Human-Drawn Paintings

Warnakulasooriya A.I.[1*], Rupasingha R.A.H.M.[2] and Kumara B.T.G.S.[3]

[1]Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka
[2]Department of Economics and Statistics, Faculty of Social Sciences and Languages, Sabaragamuwa University of Sri Lanka, Sri Lanka
[3]Department of Data Science, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka
[*]aiwarnakulasooriya@std.appsc.sab.ac.lk

The arrival of Generative Artificial Intelligence and Diffusion Models revolutionised the artwork creation process with the ability to generate hyper-realistic paintings based on a textual description. Consequently, it's crucial to distinguish synthetic paintings from human-drawn paintings (HDPs) for museums, art galleries, and art auctions to protect the social, cultural, and monetary value of art and the artists. To fill the absence of specifically detecting AI-generated paintings (AIGPs) from HDPs, this study proposes a deep-learning-based approach utilising a Convolutional Neural Network (CNN) and an Artificial Neural Network (ANN). A diverse and balanced dataset of 3000 paintings across 10 different art styles was captured from the AI-ArtBench dataset. The AIGPs were generated with the equal use of Latent Diffusion and Standard Diffusion models. The ANN and CNN models were built and methodically fine-tuned using hyperparameters. The implemented CNN model achieved 89% classification accuracy at the training dataset size of 30% while the ANN model reached 76% of optimum accuracy at the same training dataset size. The CNN's superior performance over the ANN was exceptionally evident in detecting discrete artistic patterns, making CNN more suitable for complex painting classification tasks. Furthermore, evaluation metrics such as precision, recall, F1-score, etc., were analysed for both models and compared. Moreover, the key visual features the model relied on classifying paintings were identified by examining the resulting heatmaps of the paintings from the Gradient Class Activation Maps (Grad-CAM) function. Although the study encountered limitations in processing high-resolution images with the available run-time environments of the Google Colab virtual machine, the recommended approach contributes to computer vision advancements and art authentication enhancing the automated art analysis capabilities. In future works, the models' performance across different art styles and the key features specified for each art theme will be analysed.

Keywords: *AI-Generated Paintings, ANN, Classification, CNN, Human-Drawn Paintings*

Improving the Performance of Machine Learning Classifiers for Imbalanced Multiclass Reputation Analysis

Nugekotuwa C.K.[1*] and Ishanka U.A. P.[2]

[1]*Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka*
[2]*Department of Data Science, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka*
[*]*chamodnugekotuwa@gmail.com*

Class imbalance in multiclass datasets remains a persistent challenge, often leading to biased models with reduced generalization, particularly affecting the minority class most. This research proposes a hybrid approach to addressing class imbalance by integrating data-level techniques, including custom oversampling and noise cleaning, with algorithm-level techniques such as cost-sensitive self-paced learning and a deep neural network architecture. Custom oversampling balances class distributions by generating synthetic samples for minority classes, while noise cleaning identifies and removes outliers to enhance data quality. Cost-sensitive self-paced learning assigns dynamic weights to samples based on their difficulty and rarity, enabling the model to focus on underrepresented classes while mitigating overfitting. The dataset comprises 11,276 mobile application reviews used for brand reputation analysis, exhibiting a highly imbalanced distribution. The dataset consists of 4,981 negative reviews (44.16%), 3,658 neutral reviews (32.45%), and 2,637 positive reviews (23.39%), making it an ideal benchmark for evaluating this approach. The dataset is trained using a multi-layer perceptron feedforward network, incorporating progressive neuron reduction, batch normalization, and dropout layers to improve learning efficiency and prevent overfitting. The proposed method extends the baseline model, Random Oversampling with Neighborhood Cleaning Rule (ROS-NCL), which effectively balances datasets but does not optimize the decision boundary for minority classes. Experimental results show that the proposed method outperforms the baseline model, improving accuracy from 91.62% to 94.47%, while achieving class-wise precision scores of 95.1% and 91.3% for the two minority classes and 96.2% for the majority class. The results confirm that this hybrid approach is able to perform well on multiclass-imbalanced datasets and has strong implications for real-world applications, particularly in brand reputation monitoring, customer sentiment analysis, and business decision-making, where accurate classification of opinions is crucial for maintaining the competitive advantage.

Keywords: *Class Imbalance, Cost-Sensitive Learning, Noise Cleaning, Oversampling, Reputation Analysis)*

Adaptive Feature-Weighted Ensemble Learning for Chronic Kidney Disease Classification

Sooriyage H.D.[1*], Ishanka U.A.P.[2] and Madushani U.A.K.[3]

[1]Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka
[2]Department of Data Science, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka
[3]Kahawatta Base Hospital, Gampola, Sri Lanka
[*]hashansooriyage@gmail.com

This research proposes a feature-weighted ensemble learning model to achieve optimal Chronic Kidney Disease (CKD) classification. The number of people affected with CKD continues to grow rapidly because this condition has become a major health challenge affecting more than 850 million individuals worldwide. The accurate identification of diseases remains fundamental to managing the escalating healthcare system challenges. The proposed study develops an improved high-performance diagnostic model by combining state-of-the-art machine learning algorithms with ensemble methods to enhance diagnostic precision. To detect important clinical markers, the proposed approach includes Recursive Feature Elimination (RFE) and Least Absolute Shrinkage and Selection Operator (LASSO) along with other robust feature selection algorithms. According to the outcome, Hemoglobin and Packed Cell Volume demonstrated higher clinical importance than Serum Creatinine and Specific Gravity. Random Forest, XGBoost, CatBoost, and LightGBM classifiers were used, combined with adaptive weights to enhance the model's performance.The UCI CKD dataset was used for this research, and it originally contained 24 features, but preprocessing together with feature selection reduced it to seven features, which produced an accurate model with 99.75% accuracy and 0.50% standard deviation. The model delivers performance that exceeds standard diagnostic methods to generate a dependable and efficient decision-making tool with explanatory capabilities for clinical use. The model demonstrates wide applicability across healthcare systems, which enables it to revolutionize CKD diagnosis and treatment, particularly in resource-limited healthcare environments.This feature-weighted ensemble learning methodology not only enhances the diagnosis of CKD but also demonstrates that data-driven methods are critical to achieving sustainable development in healthcare systems worldwide.

Keywords: *Chronic Kidney Disease, Ensemble Learning, Feature Selection, Machine Learning, Classification*

Enhancing Flutter App Development: Addressing Configuration and Compatibility Bugs Using Large Language Models

Rajaguru R.M.S.N.[1*], Chathumini K.G.L.[1], Kumara B.T.G.S.[2] and Wijerathna P.M.A.K.[3]

[1]*Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka*
[2]*Department of Data Science, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka*
[3]*Department of Software Engineering, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka*
[*] *supunnilakshana1999@gmail.com*

In today's world, there are cross-platform frameworks available, like Flutter, that allow the developer to develop an app from a single codebase. However, debugging issues and compatibility with different devices and dependencies are still the greatest challenges for Flutter developers because there are continuous framework changes and divergent platform-specific behaviours. To overcome these challenges, the research adopts the use of advanced LLMs, including GPT-4o, Claude Sonnet 3.5, and Gemini 2.0 Flash, combined with RAG capabilities. The collection of data from multiple sources like Stack Overflow, GitHub, and Flutter documentation initiated the study. The dataset was cleaned and preprocessed using markdown/HTML stripping, deduplication, code-text separation, and case normalization. To kick off the evaluation a fixed set of prompt templates to 350 Stack Overflow questions across all the LLMs were applied. The assessment consisted of two approaches: expert validation of 25 key questions and assessing high-scored Stack Overflow answers using cosine similarity and prediction correlation against the developed framework and standalone LLM. Through this assessment, it was able to gauge the models' efficacy in finding configuration bugs and addressing them. The RAG pipeline, designed with a dual-model embedding approach (separate models for text and code) in a vector database, achieved a higher similarity score of 0.8371. Retrieval performance was further improved through re-ranking methods. Early evidence hints at the enhanced accuracy/precision of a RAG-enhanced LLM when that LLM is a standalone model for better Flutter configuration. This paper presents a modular framework to integrate RAG-enhanced LLMs for bug detection and fixing. The framework requires much lesser debugging efforts along with more reliability of the app. Additionally, it can also adopt the fast-evolving Flutter framework. By filling a significant gap in cross-platform development literature, it helps advance AI-assisted debugging and improve the development workflows of Flutter apps.

Keywords: *Compatibility Issues, Configuration Bugs, Flutter, Large Language Models, Retrieval-Augmented Generation*

A Hybrid Deep Learning Model for Improving Tea Leaf Disease Detection: Overcoming Challenges in Sri Lanka's Tea Industry.

Rathnayaka R.M.D.N.[1*], Chathumini K.G.L.[1] and Kumara B.T.G.S.[2]

[1]*Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka*
[2]*Department of Data Science, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka*
[*]*dinithinayanamali98@gmail.com*

In the Sri Lankan tea industry, tea leaf diseases affect yield and quality. Therefore, this study proposes a hybrid deep learning model that combines MobileNet, VGG16, and ResNet50 to improve the accuracy, reliability, and robustness in detecting tea leaf disease. Each model is unique in that MobileNet contributes with its lightweight and computationally efficient nature, VGG16 extracts deep features hierarchically, and ResNet50 relies on residual connections to improve learning. However, each of them suffers from common problems in capturing complex features effectively, a higher computational load, and a longer training time while working with large or diverse datasets. Thus, the hybrid model will be able to achieve maximum performance in classification by overcoming the limitations of each model through strategic combinations of the architectures. 5,000 images of tea leaves were collected from Kaggle, labeled, and preprocessed by resizing, normalizing, and augmenting to improve the generalization and avoid overfitting. The experimental performance of individual models, as shown by the results, revealed that MobileNet classified tea leaf diseases with an accuracy of 83.9%, ResNet50 with 96.30%, and VGG16 with 93.63%. These models are to be integrated into an organized hybrid framework through ensemble learning methods, which are intended to provide optimized robustness, reduce prediction errors, and improve model interpretability for applications in agriculture. This study fills an important gap in the literature by investigating how the hybrid approach outperforms both individual models and demonstrates that the synergistic use of multiple architectures enhances disease detection accuracy and consistency. Such an approach may revolutionize large-scale agricultural disease monitoring by reducing dependency on manual inspections, which are often prone to human error. In the future, fine-tuning of the hybrid model will be done through the implementation of fine-tuning strategies, optimization of hyperparameters, and testing on real-world deployments to enhance its feasibility for tea plantations.

Keywords: *Agricultural Disease Diagnosis, Hybrid Deep Learning Models, Model Integration, Sri Lankan Tea Industry, Tea Leaf Disease Detection*

Enhancing Cinnamon Plants Disease Detection Using Advanced Machine Learning Techniques

Senevirathne M.B.L.M.[*] and Herath G.A.C.A.

Department of Computing and Information Systems, Sabaragamuwa University of Sri Lanka, Sri Lanka.
[*]mblmsenevirathne@std.appsc.sab.ac.lk

Cinnamon is a vital crop in Sri Lanka. However, some diseases can cause serious damage to the cinnamon plant. Existing methods for detecting diseases in cinnamon plantations are constrained by limited expertise and the time-intensive nature of manual testing. Because of that, there is a critical need for an efficient, comprehensive, and accurate disease identification system. The lack of a reliable and automated solution for detecting cinnamon diseases is the primary research problem. This study aims to develop an advanced deep-learning model capable of detecting and classifying six types of cinnamon diseases (Rough Bark, Stripe Canker, Leaf Blight, Black Sooty Mold, Leaf Gall Forming Louse, Leaf Gall Forming Mites), using a dataset of 1,200 images collected through field visits and Kaggle. The research explores various deep-learning models for disease classification. Baseline models starting with DenseNet, followed by VGG16 and InceptionV3, demonstrated test accuracies of 46.78%, 73.39%, and 73%, respectively**.** Further improvements were achieved using transfer learning, where EfficientNetB4 achieved the highest accuracy of 92%, followed by InceptionV3 (90%), ResNet50 (89%), and VGG16 (87.5%)**.** The ultimate goal of this research is to develop an AI-powered mobile application that will enable farmers to quickly and accurately identify cinnamon diseases, facilitate timely interventions, enable effective crop management, and minimize agricultural losses. The system is designed to be user-friendly, accessible, and deployable in real-world farming environments, ensuring practical benefits to the agricultural community.

Keywords: *Cinnamon Diseases Detection, CNN, Deep Learning, Transfer Learning*

Triage Categorization in Emergency Departments Using Ensemble Learning: Health Care

Thasitha T.[1*] and Kumara B.T.G.S.[2]

[1]Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka
[2]Department of Data Science, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka
[*]thasithathangarasa@gmail.com

Emergency department (ED) triage is a critical process that determines patient treatment priority based on medical urgency. While EDs worldwide struggle with overcrowding and resource constraints, these challenges are particularly acute in Sri Lanka, where inconsistent manual triage decisions and limited healthcare infrastructure can compromise patient care. This study develops an advanced triage prediction system using ensemble learning techniques to address these challenges. Through expert consultation, we identified key clinical indicators including vital signs, heart rate, blood pressure, respiratory rate, and oxygen saturation. Our approach implements a Stacking Classifier that combines Logistic Regression, Support Vector Machine (SVM), and Light Gradient Boosting Machine (LightGBM) algorithms, with a tuned LightGBM model serving as the final estimator. This ensemble method achieved 90.88% accuracy in predicting triage categories, offering a robust solution for optimizing patient prioritization and resource allocation in emergency settings.

Keywords: *Emergency Departments, Ensemble Learning, Health Care, Machine Learning, Triage Categorization.*

Open Track

Comparative Analysis of Arduino UNO R3 and ESP32 Microcontrollers: A Multi-Sensor Data Acquisition and Automation Perspective

Hosan K.S.[1], Perera W.A.H.V.[2], Priyadarshana H.V.V.[1,3], Wijesekara R.J.M.D.D.P.[2,3], Koswattage K.R.[*1,3]

[1]*Department of Engineering Technology, Faculty of Technology, Sabaragamuwa University of Sri Lanka, Sri Lanka.*
[2]*Department of Biosystems Technology, Faculty of Technology, Sabaragamuwa University of Sri Lanka, Sri Lanka.*
[3]*Centre for Nanodevices Fabrication and Characterization, Faculty of Technology, Sabaragamuwa University of Sri Lanka, Sri Lanka.*
[*]*koswattagekr@apppsc.sab.ac.lk*

In modern scientific experiments, researchers widely use development boards with microcontrollers to operate their experimental-level test beds and systems. This study was conducted based on a multi-sensor experimental setup incorporating eight KY028 thermistor temperature sensors, a DHT11 humidity sensor, and a YF-S201 flow rate sensor operating as a data acquisition system integrated with a cooling system controller. Initially, the experimental setup was developed using an Arduino UNO R3 development board and the same system was operated using an ESP32 microcontroller to conduct a comparative analysis between two microcontrollers. During the operation of the system, the Arduino UNO R3 demonstrated continuous irregularities in data acquisition when operating with multi-sensor systems, especially when the temperature sensor data were not accurate, which affected the data acquisition system's reliability. Comparatively, with ESP 32 controller, there were not any value-based issues, and were able to collect the temperature data with higher accuracy than the Arduino UNO R3 controller. The ESP32's enhanced computational power and efficient data-handling capabilities made it more suitable for this application. In the results analyzing process mainly identified the more than 5 to 6 Celsius temperature difference in temperature sensors and one sensor showed more than 10 Celsius abnormal temperature difference through Arduino UNO R3 microcontroller. Flowrate sensor and humidity sensor showed the same values in both microcontrollers. The results demonstrate that the ESP32 microcontroller outperforms the Arduino UNO R3 in multi-sensor data acquisition and automated control for sensitive systems. Furthermore, through this experiment these controllers represented the operational performance under complex operational conditions. This study verifies the ESP32's reliability in complex operational environments requiring precise monitoring and automation through multi-sensor inputs. Through this experimental study clearly abled to identify usability and capability of the above-mentioned microcontrollers for experimental setups.

Keywords:*Arduino-UNO-R3, ESP32, Data-acquisition, Microcontrollers, Temperature-sensing*

Comprehensive Study on User-Centric Approaches to Preventing Social Engineering Attacks

Pathirana E.P.K.H.[*] and Abeysinghe D.V.D.S.

Department of Computer Science, Faculty of Computing, General Sir John Kotelawala Defence University, Sri Lanka
[*]39-bcs-0016@kdu.ac.lk

Social engineering attacks are a critical threat to organizational security because they exploit human psychological vulnerabilities. Most users are generally unprepared to detect and mitigate the impact of such attacks despite the availability of technical safeguards. It highlights a gap in current practice and prevention strategies. This study employs a mixed-methods approach, which combines a literature review with primary data collected through questionnaires and interviews conducted with 75 participants from diverse professional backgrounds in Sri Lanka, selected through purposive sampling to ensure a representative sample. The term mixed-methods indicates both analysis of open-ended and closed-ended questions through qualitative and quantitative methods. The results showed that 78.7% of respondents were aware of the existing social engineering attacks, but most were less confident in identifying such an attack, with only 25.3% very confident in identifying them. Behavioral factors such as cognitive biases (trust, fear, and urgency) and overconfidence especially gained recognition as one of the key critical factors influencing vulnerability. Users make substandard choices even with knowledge of potential dangers because biases cloud their judgment. On the other hand, real-time simulations and personalized interactive training tools have been identified as more effective for improving user readiness than traditional training methods. These findings have identified the need for user-centered cybersecurity education that integrates psychological and technological measures as a means of better positioning users against the threats of social engineering attacks. Further research should focus on developing such tools and the expansion of adaptive training programs for a wide range of user groups.

Keywords: *Adaptive Cybersecurity Education, Interactive Security Tools, Social Engineering Attacks, User-Centric Security*

A Systematic Review of Latency Reduction Techniques in Smart Home IoT Networks

Vinnath W.L.V.[*] and Abeysinghe D.V.D.S.

Department of Computer Science, Faculty of Computing,
General Sir John Kotelawala Defence University, Sri Lanka
[*]39-bcs-0005@kdu.ac.lk

The Internet of Things (IoT) is improving automation, security, and energy management by converting traditional living spaces into smart settings. However, latency issues pose significant challenges, affecting the seamless operation of smart home IoT networks. This systematic review explores latency reduction techniques in smart home IoT systems, focusing on their impact on network performance, security, and energy efficiency. The study evaluates research from conference proceedings and peer-reviewed publications between 2010 and 2024. It considers three primary categories of solutions: Fog and edge computing frameworks, which enable localized data processing to reduce transmission delays, Machine learning-based optimization methods, which dynamically adapt to network conditions in real-time and Blockchain-enabled hybrid systems, which enhance security while mitigating latency. The review discusses the trade-offs associated with these methods, including computational overhead, resource consumption, and security benefits. Additionally, it examines the types of IoT devices commonly deployed in smart homes, such as smart sensors, connected appliances, and security systems, providing insights into their unique latency challenges. Findings indicate that no single approach provides a universally optimal solution, emphasizing the need for context-aware, hybrid strategies. The review concludes that integrating fog computing for decentralized data processing with machine learning for adaptive task scheduling offers a more effective approach to mitigating latency. Future research should focus on refining hybrid frameworks to enhance scalability, adaptability, and overall efficiency in evolving smart home IoT ecosystems.

Keywords: *Smart home IoT, Latency reduction, Fog computing, Edge computing, Machine learning*

Assistive Navigation for the Visually Impaired: Real-Time Crosswalk Detection Using Satellite Images and Machine Learning

Thawindi W.G.M. [1*], Senarathna W.G.H.R.[1], Appuhamilage D.K.M.A.S.N.D.K.M. [1], Mendis W.M.H.[1] and Herath W.B.P.N.[2]

[1] *Department of Physical Sciences, Faculty of Applied Sciences, Rajarata University of Sri Lanka, Sri Lanka*
[2] *Department of Computing, Faculty of Applied Sciences, Rajarata University of Sri Lanka, Sri Lanka*
**monishathawindi1999@gmail.com*

Visually impaired individuals face many challenges when navigating urban areas, particularly when they are crossing streets. Existing navigation tools lack real-time detection of crosswalks, and depend on immediate surroundings, posing challenges to safe and independent navigation. This study addresses this limitation by presenting a wearable that utilizes machine learning to detect crosswalks in real-time from satellite imagery. This approach with the use of Google Maps satellite imagery ensures accurate and timely crosswalk detection, unlike the proximity-based navigation systems that rely on local sensors. A dataset of 1,740 satellite images annotated for binary classification of crosswalks and non-crosswalks instances was trained using the YOLOv5s object detection model. To enhance detection accuracy, the model was fine-tuned using hyperparameter optimization, including batch size, learning rate, and the number of epochs. The precision of the trained model was 95.24%, recall 98.45%, and mean average precision (mAP@0.5) 99.71%. The trained model, along with a Raspberry Pi 4 and GPS receiver, gives audio notifications at 5 m from a crosswalk, enabling the user's safety and independence. Simultaneously, the dependence on internet connectivity to obtain satellite images, and the image quality being hampered due to low resolutions can be named as the limitations of the study. The contribution of this study lies in bridging the gap between existing assistive technologies and the need for real-time, reliable crosswalk detection. This research offers a novel, energy-efficient, and cost-effective solution for visually impaired individuals navigating urban environments by integrating machine learning with satellite imagery. Future directions will expand the dataset by including additional images, optimizing model efficiency, and enhancing adaptability across diverse environmental conditions. Unlike traditional navigation aids, which lack real-time crosswalk detection, this system provides an innovative way to ensure safety and independence, especially in urban areas.

Keywords: *Crosswalk detection, Google Maps satellite imagery, Real-time navigation, visually impaired assistance*

A Modification of Hungarian Method to solve One Dimensional Bin Packing Problems

Weerasinghe T.D.T. *, Ekanayake E.M.U.S.B.

Department of Physical Sciences, Faculty of Applied Sciences, Rajarata University of Sri Lanka, Sri Lanka
**dananjithathsarani1@gmail.com*

In bin packing problems, items of various sizes must be packed into a limited number of bins or containers, each with a fixed capacity. There are several kinds of BPP such as one-dimensional (1D), Two-dimensional (2D), three-dimensional (3D). While First Fit and Best Fit method is faster and simpler it does not ensure that optimal solution and can result in smaller spaces in the bins that could have been occupied by larger objects. This research introduces a new method for solving the 1D BPP inspired by the row reduction method from Hungarian method. To minimize the bin usage, carefully analyzes item placement by applying the concept of row reduction. Existing methods often produce suboptimal solutions, resulting in wasted space within bins. This method aims to achieve optimal packing, minimizing waste. First, arrange items in ascending order of size as Row 1 and descending order of size as Row 2. Subtract the smallest value in Row 2 from all values, recording results in Row 3. Identify items in Row 1 corresponding to zeroes in Row 3 and allocate them to the first bin. Select the smallest non-zero value in Row 3, subtract it from all values, and update as Row 4. Repeat the allocation process for remaining zeroes in Row 4, ensuring bin capacity is not exceeded. This new approach to 1D BPP has been thoroughly tested using examples taken from research papers Continue above steps until all items are assigned to bins. This suggested method presents a new approach to the 1D BPP that consistently produces optimal solutions that perform better than existing methods such as the First Fit and Best Fit approaches. This study provides a foundation for researching more developments for bin packing problems such as 2D and 3D.

Keywords: *Bin Packing, Hungarian method, Optimal solution, First Fit*

Machine Learning-based Quality Detection of Cinnamon from Outer Bark Images

Samarajeewa S.G.M.D., Balasooriya W.B.I.S., Rasanjalee W.G. R.*, Jeewanthi J.G.A., Ravihara M.P.N. and Herath W.B.P.N.

Department of Computing, Rajarata University of Sri Lanka, Sri Lanka
**ridmirasanjalee63@gmail.com*

Cinnamon is a vital component of Sri Lanka's agriculture, internationally recognized for its flavor, aroma, and medicinal properties. Despite its significance, traditional cinnamon quality grading relies on manual visual inspections, which are subjective, time-consuming, and inconsistent. Also, misclassification of cinnamon bark's highest obtainable grade can lead to improper grading decisions and suboptimal processing, reducing its potential value. Previous research has focused primarily on species identification or post-processing quality assessment, leaving a gap in grading cinnamon bark before processing. This study introduces a machine learning model to detect the highest grade obtainable from cinnamon bark, ensuring its optimal utilization. A four-stage deep learning pipeline: Data Collection, Preprocessing, Segmentation, and Classification was developed. A dataset of over 1,500 high-resolution cinnamon bark images, captured in a 4:3 aspect ratio using phone cameras under controlled lighting conditions, was prepared. These images were labeled into four predefined quality classes: Extra Special High Quality, High Quality, Medium Quality, and Low Quality, by industrial experts. Class imbalance was addressed using augmentation techniques, blurriness detection, image resizing, normalization, and CLAHE (Contrast Limited Adaptive Histogram Equalization) to enhance clarity and texture. Segmentation was performed using the U-Net architecture integrated with attention and residual blocks, achieving an IoU (Intersection over Union) of 0.87 and a Dice Coefficient of 0.90, outperforming basic models such as thresholding and edge detection. Classification was carried out using a fine-tuned ResNet101 CNN with transfer learning, achieving an accuracy of 92%, recall of 0.89, and an F1-score of 0.89. The model was integrated into a mobile application, allowing farmers to submit bark images for instant grading. Model's reliance on controlled lighting and lack of generalization to other types of cinnamon bark, remain as limitations of the study. Nevertheless, this research offers an efficient, scalable solution for cinnamon quality detection, reducing human error, improving grading accuracy, and paving the way for broader AI adoption in the cinnamon industry.

Keywords: *Cinnamon Quality Detection, Image Segmentation, Deep Learning, ResNet101, U-Net, CLAHE*

Automated vs. Manual: Investigation of Human Error Associated with Time-related Measurements

Lakmal H.D.S.M.*, Ranasingha K.M.G.G.I.U., Gunasinghe P.A.K.G. and Lakshan K.A.S.

Department of Physics, University of Ruhuna, Sri Lanka
** savirumalinda@gmail.com*

In this study, we have studied the human error associated with time-related measurements. The free fall of an object is one of the most significant time-related phenomena studied in physics. This study compares manual time measurements using a stopwatch with analyzed time derived from a recorded video of the free fall of a small ball to evaluate the precision of human timing during the experiment. In this experiment, a small ball was dropped from a height of one meter. The free fall experiment was repeated three times by four participants, resulting in a total of twelve trials using a smartphone stopwatch. The time analyzed from the video was considered the reference value for the experiment. Video analysis was performed using an image processing method implemented in Python. The average time measured using the stopwatch was (0.40 ± 0.01) seconds, while the free fall time obtained from the video was 0.4667 seconds. Relative to the reference value, the results from the stopwatch showed a percentage error of 14.2918%. For this phenomenon, the value of time calculated using the equation of motion was 0.4516 seconds. Compared to this value, the result from the video showed an error percentage of 3.3437%. The time measured using the stopwatch was examined to evaluate the accuracy and consistency of human measurements. Contributing factors to human reaction time errors include variability in participant's eye levels, the reaction time when starting and stopping the stopwatch, and difficulty stopping the stopwatch precisely when the ball hit the ground. Environmental air resistance further affected time deviations. This experiment provides insight into the unpredictability introduced by human error in manual timing techniques during physics investigations. The results highlight the need for precise instruments to improve reliability and accuracy in experimental physics.

Keywords: *Free Fall, Human Error, Image Processing*

Assessing the Role of Internet Connectivity in IT Project Success and Stakeholder Satisfaction in Sri Lanka: A Comprehensive Perspective

Perera M. A. M. D.[1*] and Nirubika R.[2]

[1]*Department of Computing and Information Systems, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka*
[2]*Department of Software Engineering, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka*
[*]*mamdperera@std.appsc.sab.ac.lk*

Reliable internet connectivity is crucial for effective project execution, seamless communication, and stakeholder satisfaction in Sri Lanka's evolving IT sector. However, persistent challenges such as inconsistent quality, regional disparities, and frequent disruptions hinder project timelines, budgets, and stakeholder trust. This study examines these challenges using a mixed-methods approach, integrating survey data and qualitative insights from IT professionals, project managers, and end-users. The quantitative analysis revealed that while over 50% of respondents rated connectivity as "Good" or "Very Good," 14% rated it as "Poor," highlighting notable regional inconsistencies. Qualitative findings emphasized that frequent disruptions significantly impact productivity, leading to delays and cost overruns. Furthermore, 67% of stakeholders acknowledged the positive correlation between improved connectivity and project success, while over 80% supported incorporating connectivity metrics into Key Performance Indicators (KPIs). Unlike previous studies that primarily assessed general internet accessibility, this research specifically evaluates its implications for IT project success and stakeholder engagement. Taking a requirement engineering perspective, the study proposes actionable strategies, including infrastructure enhancements, regulatory reforms, and improved stakeholder collaboration, to bridge connectivity gaps. The findings provide a roadmap to optimize project outcomes, enhance communication, and strengthen Sri Lanka's IT sector's competitiveness on a global scale.

Keywords: *Internet Connectivity, IT Project Success, Stakeholder Engagement, Digital Infrastructure, Requirement Engineering*

Enhancing Cloud Computing Performance Through the Integration of Weighted LRU Caching and Dynamic Task Scheduling

Randeni R.M.S.D.[1*] and Jayasena K.P.N.[2]

[1] *Department of Computing and Information System, Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka*
[2] *Postdoctoral Research Fellow, Technische University Dresden, Germany*
[*] *sashikadulaj20643@gmail.com*

The scalability of cloud services comes with performance limitations due to resource contention and inefficient task scheduling. This leading to increased latency, reduced throughput, poor resource utilization, higher operational costs, and degraded Quality of Service. This study presents an innovative approach by integrating dynamic task scheduling with Weighted Least Recently Used (WLRU) caching to address these issues. WLRU caching improves cache efficiency by prioritizing critical data, such as task execution details, virtual machine (VM) metrics, and VM failure rates at the data center level, reducing latency and enhancing system performance. The dynamic scheduling algorithm adapts to workload variability, optimizing resource utilization and task execution. The methodology integrates WLRU caching with the Heterogeneous Earliest Finish Time (HEFT) scheduling algorithm, which has not been considered in prior studies. Simulations were conducted using CloudSim, focusing on task completion time as the key performance metric across two test scenarios. The integrated approach outperforms traditional HEFT without caching, improving task completion time by 17.78% for small workloads (3 VMs with 1–5 GB RAM, 10 cloudlets with 4,000–20,000 million instructions) and 15.23% for large workloads (10 VMs with 1–5 GB RAM, 100 cloudlets with 4,000–20,000 million instructions). By storing detailed task execution data and VM metrics, this approach significantly decreases latency and improves resource usage. The study's findings underscore the potential of combining caching and scheduling techniques to optimize cloud performance, offering a unique solution for addressing the limitations of current models and contributing to more efficient, scalable, and sustainable cloud systems. Future studies can incorporate multiple performance metrics, such as cache hit and miss ratio and power usage.

Keywords: *Cache-Aware Task Scheduling, Cloud Computing, Dynamic Task Scheduling, HEFT Scheduling, WLRU Caching*

AUTHOR INDEX

Abeygunawardhana D.K.	35
Abeysinghe D.V.D.S.	8, 14
Adeeba S.	21, 22, 23, 25
Ananda K.V.A.I.	57
Appuhamilage D.K.M.A.S.N.D.K.M.	15
Balasooriya W.B.I.S.	19
Charles J.	13
Chathumini K.G.L.	57, 65, 66
Dangalla R.L.	62, 76
Dayarathna K.O.K.S.	73
De Silva H.M.C.J.	64
Dharmadasa T.K.R.S.	28
Dias J.B.R.N.	33
Duleesha G.R.	34
Ekanayake E.M.U.S.B.	17
Erandi J.D.T.	51, 57
Fernando M.F.P.S.L.	73
Fernando W.D.P.	22
Geekiyanage C.L.	27,
Gunasinghe P.A.K.G.	20
Gunawardhana H.M.K.T.	41
Herath G.A.C.A.	69, 72
Herath W.B.P.N.	15, 19
Hosan K.S.	4
Hulathdoowage S.K.D.	7
Ishanka U.A. P.	45, 61, 63
Jayasena K.P.N.	70
Jeewanthi J.G.A.	19
Junoj S.	47
Karunakaran L.	44
Kaushalya P.K.D.K	33, 34
Koswattage K.R.	4
Kumara B.T.G.S.	6, 7, 21, 22, 23, 25, 28, 31, 41, 47, 54, 56, 64, 65, 66, 69, 81
Kumara H.K.M.H.V.	54
Kumari M.K.M.N.	48

Lakmal H.D.S.M.	20
Lakshan K.A.S.	20
Lekamge L.S.	26, 27, 33, 34, 35
Lorshan N.	6
Madushani U.A.K.	63
Malkanthi C.	44
Mathusan S.	21
Mendis W.M.H.	15
Nirosana K.	45
Nirubika R.	42, 43, 73
Nugekotuwa C.K.	61
Pathirana E.P.K.H.	8
Pathirana P.U.N.	60
Perera M.A.M.D.	43
Perera R.D.M.	73
Perera W.A.H.V.	4
Perera W.P.K.I.	13
Peries R.F.S.	25
Prabodhani L.H.K.N.	26
Prasanth S.	45
Priyadarshana H.V.V.	4
Rajaguru R.M.S.N.	65
Ranasingha K.M.G.G.I.U.	20
Ranasinghe R.M.H.K.	69
Ranathunga R.A.D.K	31
Randeni R.M.S.D.	70
Rasanjalee W.G. R.	19
Rathnayaka R.M.D.N.	66
Ratnayaka R.M.T.K.	44
Ravihara M.P.N.	19
Rukais S.L.M.	23
Rupasingha R.A.H.M.	28, 31, 56
Samarajeewa S.G.M.D.	19
Samarakoon S.M.T.D.	62
Sandaruwan R.M.T.	29
Senarathna W.G.H.R.	15
Senevirathne M.B.L.M.	72
Somaweera W.T.S.	29

Sooriyage H.D.	63
Thasitha T.	81
Thawindi W.G.M.	15
Udeshika P.K.M.	51
Udeshika W.A.	32
Vigneshwaran P.	13
Vinnath W.L.V.	14
Wanigasekara W.M.P.S.B.	41
Wanigasooriya W.A.D.A.U.	76
Wanniarachchi W.A.S.Y	42
Warnakulasooriya A.I.	56
Wasalthilaka W.V.S.K.	32, 48, 49, 60
Weerakkodi S.	54
Weerasinghe T.D.T.	17
Welagedara H.T.	49
Wijerathna P.M.A.K.	6, 47, 64, 65
Wijerathna R.M.K.K	26, 27, 35
Wijesekara R.J.M.D.D.P.	4
Wijethunga W.A.T.G.R.	73

Best wishes for the ComURS2025 Computing Undergraduate Research Symposium 2025