



MINISTRY OF ECONOMY  
DEPARTMENT OF STATISTICS MALAYSIA

The Second International Conference on  
Data Science and Emerging Technologies

# DaSET 2023

TOWARDS GREEN ARTIFICIAL INTELLIGENCE AND SUSTAINABLE COMPUTING



PROGRAMME & ABSTRACTS



4-5 December 2023  
Virtual Conference  
<https://www.icdaset.com>



IN COLLABORATION WITH:



## **PROGRAMME & ABSTRACTS BOOK**

**International Conference on Data Science and Emerging Technologies  
(DaSET 2023)**

*Towards Green Artificial Intelligence and Sustainable Computing*

Copyright © 2023 by UNITAR International University, Kelana Jaya, Petaling Jaya, Selangor, Malaysia

All rights reserved. Personal use of this material is permitted. However, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or any means, electronic, mechanical, photocopying, recording or otherwise, without prior permission, in writing, from the publisher.

## **TABLE OF CONTENTS**

About UNITAR International University	4
Message from the Vice-Chancellor	6
Message from the Conference Chairs	8
Introduction to DaSET 2023	10
Keynote Speaker 1: Professor Dr Muhammad Khuram Khan	11
Keynote Speaker 2: Mr Raja Segaran	14
Keynote Speaker 3: Professor Dr Naomie Salim	16
Keynote Speaker 4: Dato' Dr. Amirudin Abdul Wahab	18
Keynote Speaker 5: Professor Dr Seifedine Kadry	20
Keynote Speaker 6: Ms Puteri Anis Aneeza binti Zakaria	22
Keynote Speaker 7: Associate Professor Dr Chin Kim On	24
Keynote Speaker 8: Professor Dr Schahram Dustdar	26
<b>Conference Schedule DaSET 2023</b>	<b>28</b>
Paper Presentations Schedule	35
Proceeding Abstracts	45
Track 1: Artificial Intelligence	45
Track 2: Machine & Deep Learning	54
Track 3: Statistical Learning	70
Track 4: Health & Industrial Applications	84
Conference Organization	95
Reviewers	98
Acknowledgements	100

## **About UNITAR International University**



UNITAR International University (UNITAR) was established in 1997 as the first virtual university in Southeast Asia and among the earliest private universities in the country. It is the first institution in Asia to be awarded a QS 5-Star Rating for the Online Learning category. In alignment with the United Nation's Sustainable Development Goals, UNITAR's holistic and innovative learning ecosystem provides a comprehensive programme offering for all socio-economic levels whilst developing highly employable and entrepreneurial global citizens.

It aims to go beyond just providing the learning journey for students in acquiring relevant skill sets, but through their C.A.R.E. — Collaborative, Adaptive, Reflective, and Entrepreneurial education methodology, UNITAR nurtures its students in developing self-awareness of their talents and abilities, as well inculcating in them the need for social responsibility whilst moulding them to be better citizens of the nation.

UNITAR's programmes are offered in 13 locations, including its main campus in urban Kelana Jaya. It has a total enrolment of over 7,000 students spread within 40 homegrown programmes ranging from the Foundation, Diploma and Bachelor's levels to the postgraduate Master's and Doctorate's. To meet the current needs of the industries, UNITAR also provides over 50 professional certification programmes. With APEL, working professionals can continue their education journey with the option to reduce their programme duration.

UNITAR takes pride in having committed academics housed in various faculties and schools. The academic verticals comprise of the Faculty of Business, Faculty of Education & Humanities, the School of Information Technology (SIT) and the School of Media, Arts & Design (SMARd). Since its inception, UNITAR has embarked on an

innovative learning methodology where technology is featured significantly, but never aimed to replace human relationships. Its pursuit to offer competitive quality education in both conventional and online teaching modes continues till today, meeting the needs of 21<sup>st</sup> Century education.

<b>Business</b>				
<b>Certificate</b> <ul style="list-style-type: none"> <li>• Business Studies 🌐</li> </ul>	<b>Diploma</b> <ul style="list-style-type: none"> <li>• Business Administration 🌐 🇲🇾</li> <li>• Logistics Management 🌐</li> </ul>	<b>Bachelor's Degree</b> <ul style="list-style-type: none"> <li>• Business Administration 🌐</li> <li>• Fintech 🌐</li> </ul>	<b>Master's Degree</b> <ul style="list-style-type: none"> <li>• Business Administration 🌐</li> </ul>	<b>Doctorate</b> <ul style="list-style-type: none"> <li>• Business Administration 🌐</li> <li>• Philosophy (PhD) in Business Administration</li> <li>• Philosophy (PhD) in Management</li> </ul>
<b>Accounting</b>				
<b>Diploma</b> <ul style="list-style-type: none"> <li>• Accounting 🌐</li> </ul>	<b>Bachelor's Degree</b> <ul style="list-style-type: none"> <li>• Accounting 🌐</li> </ul>			
<b>Information Technology</b>				
<b>Foundation</b> <ul style="list-style-type: none"> <li>• Information Technology 🌐</li> </ul>	<b>Diploma</b> <ul style="list-style-type: none"> <li>• Information Technology 🌐</li> </ul>	<b>Bachelor's Degree</b> <ul style="list-style-type: none"> <li>• Information Technology 🌐</li> </ul>	<b>Master's Degree</b> <ul style="list-style-type: none"> <li>• Information Technology 🌐</li> </ul>	<b>Doctorate</b> <ul style="list-style-type: none"> <li>• Philosophy (PhD) in information Technology</li> </ul>
<b>Education &amp; Humanities</b>				
<b>Diploma</b> <ul style="list-style-type: none"> <li>• Early Childhood Education 🌐 🇲🇾</li> <li>• Psychology 🌐</li> </ul>	<b>Bachelor's Degree</b> <ul style="list-style-type: none"> <li>• Education 🌐</li> <li>• Education (UPSI-UNITAR Dual-Degree)</li> <li>• Early Childhood Education 🌐</li> <li>• Psychology 🌐</li> </ul>	<b>Post-Graduate Diploma</b> <ul style="list-style-type: none"> <li>• Education 🌐</li> </ul>	<b>Master's Degree</b> <ul style="list-style-type: none"> <li>• Early Childhood Education 🌐</li> <li>• Educational Leadership &amp; Management 🌐</li> <li>• Education (Teaching English as a Second Language) 🌐</li> <li>• Instructional Technology 🌐</li> </ul>	<b>Doctorate</b> <ul style="list-style-type: none"> <li>• Philosophy (PhD) in Education</li> </ul>
<b>Art &amp; Design</b>				
<b>Certificate</b> <ul style="list-style-type: none"> <li>• Art and Design</li> </ul>	<b>Diploma</b> <ul style="list-style-type: none"> <li>• Animation Design</li> <li>• Fashion Design</li> <li>• Interior Design</li> </ul>	<b>Bachelor's Degree</b> <ul style="list-style-type: none"> <li>• Animation Design with Game Art</li> <li>• Fashion Design with Marketing</li> </ul>	<div style="border: 1px solid black; border-radius: 15px; padding: 5px; display: inline-block;"> <span>🌐 Online mode available</span>  <span>🇲🇾 Offered in Bahasa Melayu</span> </div>	

## **Message from the Vice-Chancellor**



It is indeed a great pleasure for me to welcome all respected guests, esteemed keynote speakers, paper presenters, and participants to the Second International Conference on Data Science and Emerging Technologies 2023 (DaSET 2023). We are very honoured to have YBhg. Dato' Sri Dr. Mohd Uzir Mahidin, Chief Statistician of Malaysia to officiate the opening of DaSET2023.

I am also very proud to welcome Professor Dhiya Al-Jumeily, General and Founding Chair of DaSET conference who is visiting us at UNITAR International University. Professor Al-Jumeily is also our Visiting Professor for the School of Information Technology, and we are grateful for his support in establishing international collaborations for academic and research excellence.

On behalf of UNITAR, I would like to express my gratitude to renowned international and keynote speakers for being on board with us on the journey of knowledge sharing on the application of data science and emerging technologies for environment sustainability, and education, economic, and social development.

International conferences are important for knowledge sharing and building strategic partnerships between academics and industries. Conferences are the catalyst to encourage a more inclusive dialogue, plus knowledge and skills transfer on a global scale. The role of universities is to cultivate creative and innovative thinkers who can provide solutions to solve social, economic, and environmental issues. In addition, universities play an important role in ensuring an industry-informed curriculum to produce graduates who will fill the talent gap in the industry.

With the emphasis on IR4.0 skills, industries play an important role to share on technological advancements, especially in the area of data science and emerging technologies. Topics like Artificial Intelligence, Internet of Things, Cloud Computing,

Blockchain, Cyber-Security, Computer Vision, and Augmented Virtual Reality provide innovative and progressive developments in many areas including healthcare, manufacturing, agriculture, transport, business, finance, and education- to name a few.

Smart partnerships with industries will support the university's mission to empower graduates with the right skills to push forward the digital transformation agenda in the 21<sup>st</sup> century for society and economic growth. The theme of DaSET 2023 is: Towards Green Artificial Intelligence and Sustainable Computing. Green AI refers to applications of artificial intelligence technology that address environmental issues. Green AI promotes circular economy practices and aims to reduce waste and energy consumption.

It is essential that researchers from various disciplines develop their capabilities in research related to data science, AI, and emerging technologies. I am optimistic that this two-day event will provide opportunities for multi-disciplinary collaborations, which will accelerate impactful research outputs.

On that note, I would like to thank and congratulate the committees from UNITAR International University, Liverpool John Moores University and the Department of Statistics Malaysia for their efforts in organizing this conference. I am also happy with the great support from our local and international university partners: UK Malaysia University Consortium, Universiti Teknologi MARA, Universiti Teknologi Malaysia, CerDaS (Centre for Data Science), Universiti Teknologi PETRONAS, Institut Teknologi Sepuluh Nopember, Indonesia; Chulalongkorn University, Thailand; Charles Sturt University, Australia; Institut Teknologi Bandung, Indonesia; National Defence University of Malaysia and Data Analytics and Collaborative Computing Group, University of Macau, China. We also appreciate the strong support from our industry partners: Microsoft Malaysia, Malaysia Digital Economy Corporation (MDEC), Cybersecurity Malaysia, and Statworks (M) Sdn Bhd.

I wish everyone a rewarding and fulfilling experience at DaSET 2023. May you move forward with strategic collaborations and bring forth innovative ideas and impactful research for the benefit of the global community.

**Professor Emeritus Tan Sri Dato' Sri Ir. Dr. Sahol Hamid Bin Abu Bakar FASc**  
**Vice-Chancellor**  
**UNITAR International University**

## **Message from the Conference Chairs**



On behalf of UNITAR International University, Liverpool John Moores University, and the Department of Statistics Malaysia, we would like to extend a warm welcome to all our honorable guests, esteemed keynote speakers, and participants of the Second International Conference on Data Science and Emerging Technologies 2023 (DaSET 2023) with the theme “Towards Green Artificial Intelligence and Sustainable Solutions.”

DaSET2023 aims to provide a platform for academics and industry practitioners to share recent trends in data science and emerging technologies research and to establish collaborations toward impactful, innovative research for community development, business success, and economic prosperity. We are very grateful for the support from UNITAR’s university partners: UK Malaysia University Consortium, Universiti Teknologi MARA, Universiti Teknologi Malaysia, CerDaS (Centre for Data Science), Universiti Teknologi PETRONAS, Institut Teknologi Sepuluh Nopember, Indonesia; Chulalongkorn University, Thailand; Charles Sturt University, Australia; Institut Teknologi Bandung, Indonesia; National Defence University of Malaysia and Data Analytics and Collaborative Computing Group, University of Macau, China. We also appreciate the strong support from Microsoft Malaysia, Malaysia Digital Economy Corporation (MDEC), Cybersecurity Malaysia, and Statworks (M) Sdn Bhd.

We are very honored to have Dato’ Sri Dr Mohd Uzir Mahidin, Chief Statistician of Malaysia to officiate the opening of DaSET2023. We are proud to have eight distinguished international and local keynote speakers: Professor Dr Muhammad Khurram Khan, King Saud University, Saudi Arabia; Prof. Dr. Schahram Dustdar, Tu Wien, Vienna University of Technology, Austria; Professor Dr Seifedine Kadry, Noroff Education AS, Norway; Professor Dr. Naomie Salim, Universiti Teknologi Malaysia, Assoc. Prof. Dr Chin Kim On, Universiti Malaysia Sabah, Dato’ Dr. Amirudin Abdul Wahab, Cybersecurity Malaysia, Ms Puteri Anis Aneeza binti Zakaria, Statworks Group, and Mr Raja Segaran, MDEC, Malaysia.

We would like to thank Professor Emeritus Tan Sri Dato' Sri Ir Dr. Sahol Hamid Bin Abu Bakar, Vice Chancellor of UNITAR International University, for his outstanding leadership, advice, and support of local and international academic activities to foster collaborations that lead to the exchange of knowledge and skills for research with impactful outcomes for social and economic prosperity. We are also grateful to Mr. Puvan, CEO of UNITAR, and the corporate staff for their ongoing assistance and support.

From a total of 80 submitted papers, 43 were selected after a rigorous review process for oral presentation. The authors and presenters for these 43 papers represented 10 different countries. We would like to thank Professor Dr. Michael W. Berry, University of Tennessee, for his contribution as editor of DaSET 2023 proceedings which will be published by Springer in the Lecture Notes in Data Engineering and Communication series. We gratefully acknowledge the tremendous support from all the technical reviewers who generously sacrificed their time to review the papers. We also thank the Series Editor, Springer, Lecture Notes on Data Engineering and Communications Technologies), for the opportunity to organize this guest-edited volume. Finally, we are grateful to Mr. Aninda Bose (Senior Publishing Editor, Springer India Pvt. Ltd.) and Mr. Radhakrishnan Madhavamani for the excellent collaboration, patience, and help during the preparation of this volume. We are confident that the publication will provide insightful information to researchers, practitioners, and graduate students in the areas of data science, artificial intelligence, and emerging technologies which are important in the digital information era.

Our heartfelt thanks go to our dedicated DaSET 2023 committees, who have worked tirelessly to ensure a successful conference. We also thank all those who have contributed directly or indirectly to the success of DaSET 2023. Lastly, we wish everyone a fruitful discussion and sharing during DaSET 2023. May new innovative ideas, friendships, and collaborations be established from this conference.

**Professor Dr Yap Bee Wah & Professor Dr Dhiya Al-Jumeily, OBE  
General and Founding Conference Chairs**

## **Introduction to DaSET 2023**

DaSET2023 aims to bring together leading industry practitioners, academic scientists, researchers and research scholars to exchange and share their experiences and research results on all aspects of Data Science and Emerging Technologies. It also provides a premier interdisciplinary platform for researchers, practitioners, and educators to present and discuss the most recent innovations, trends, and concerns as well as practical challenges encountered, and solutions adopted in the fields of data science and emerging technologies. One of the main challenges of Industry 4.0 is data science knowledge and skills for various domains such as business, services, education, healthcare and manufacturing. This conference provides the platform for industry practitioners and academia to come together, share and learn from each other and form effective collaborations to accelerate talent development for enterprises and governments.

DaSET 2023 is organized by UNITAR International University, Liverpool John Moores University, UK and Department of Statistics (DOSM) and supported by UK Malaysia University Consortium, Universiti Teknologi MARA, Universiti Teknologi Malaysia, CerDaS (Centre for Data Science), Universiti Teknologi PETRONAS, Institut Teknologi Sepuluh Nopember, Indonesia; Chulalongkorn University, Thailand; Charles Sturt University, Australia; Institut Teknologi Bandung, Indonesia; National Defence University of Malaysia; Data Analytics and Collaborative Computing Group, University of Macau, China, Microsoft Malaysia, Malaysia Digital Economy Corporation (MDEC), Cybersecurity Malaysia, and Statworks (M) Sdn Bhd.

The theme of the conference is 'Towards Green Artificial Intelligence and Sustainable Computing.' DaSET 2023 has invited renowned international and local keynote speakers who are academia or practitioners to share their knowledge and experience in the applications of data science in various disciplines. This conference aims to create rigorous international and local university-university and university-industry and collaborations for knowledge transfer and talent development in the area of data science.

All accepted and registered papers of DaSET2023 will be published in the renowned series of "**Lecture Notes on Data Engineering and Communications Technologies**" <https://www.springer.com/series/15362> by Springer Verlag.

Authors of selected papers can submit extended papers to the **Journal of Information and Communication Technology (JICT)** or **Indonesian Journal of Electrical Engineering and Computer Science**. There are also opportunities to publish extended papers in a **Special Issue of Pertanika Journal of Science and Technology**.

## **Keynote Speaker 1: Professor Dr Muhammad Khuram Khan**

**King Saud University, Saudi Arabia**



### **Security and Privacy in Connected and Autonomous Vehicles: State-of-the-art and Future Directions**

#### **Abstract**

Connected and autonomous vehicles (CAVs) are a collection of smart components such as sensors, Lidars, Electronic Control Units (ECUs), and cameras programmed through voluminous lines of code. There is growing interest in CAVs in global markets, as they hold the potential to create a more sustainable future. They are generating increased revenues and experiencing significant market penetration, thereby enabling disruptive innovations in the automotive industry. With the booming technological sophistication of today's vehicles as well as their hyper-connectivity, cybersecurity becomes more critical than ever. In this speech, we will focus on recent technological advancements and developments in CAVs. Moreover, we will explore security and privacy issues and challenges that could exacerbate the risks of data breaches and cyberattacks in CAVs. Lastly, we will highlight some emerging cybersecurity trends in CAVs along with future research directions.

#### **Biography**

Muhammad Khuram Khan is a global thought leader and influencer in cybersecurity. He is a Professor of Cybersecurity at the Center of Excellence in Information Assurance, King Saud University, Saudi Arabia. He is one of the founding members of CoEIA and has served as Manager R&D from March 2009 until March 2012. He, along with his team, developed and successfully managed the cybersecurity research program at CoEIA, which turned it into one of the best

centers of excellence in the region. He is the founder and CEO of the "Global Foundation for Cyber Studies and Research", an independent and non-partisan cybersecurity think-tank in Washington, D.C., USA. He has contributed to cyber policy work for the G20 (Saudi and Italian Presidencies) in shaping a safer cyberspace for children, protecting the masses and vulnerable populations in cyberspace, and empowering and enabling women in the cybersecurity profession. In 2019, he played an instrumental role as a cybersecurity subject expert for a USD 6 million series B investment in a South Korean startup, "SecuLetter," which has received a corporate valuation of over USD 40 million (2020) with a pre-IPO valuation of USD 100 million (2022).

He is the Editor-in-Chief of a well-reputed international journal, 'Telecommunication Systems', published by Springer-Nature for over 27 years, with a recent impact factor of 2.5 (JCR 2023). He is also the Editor-in-Chief of Cyber Insights Magazine. Furthermore, he is the area editor (Cybersecurity) of IEEE Communications Surveys & Tutorials, which has a 35.6 impact factor in JCR 2023. Moreover, he has been on the editorial board of several journals, including IEEE Transactions on Information Forensics and Security, IEEE Communications Magazine, IEEE Transactions on Intelligent Vehicles, IEEE Internet of Things Journal, IEEE Transactions on Consumer Electronics, Journal of Network & Computer Applications (Elsevier), IEEE Access, IEEE Consumer Electronics Magazine, PLOS ONE, Electronic Commerce Research, IET Wireless Sensor Systems, Journal of Information Hiding and Multimedia Signal Processing, and International Journal of Biometrics, etc. He has also played the role of guest editor for several international journals published by IEEE, Springer, Wiley, Elsevier, and Wiley. Besides, he is one of the organizing chairs of more than 5 dozen international conferences and a member of the technical committees of more than 10 dozen international conferences. In addition, he is an active reviewer of several international journals as well as research foundations in Switzerland, Italy, Saudi Arabia, and the Czech Republic.

Prof. Khurram is an honorary professor at the IIIRC, Shenzhen Graduate School, China. He is also an adjunct professor at Taylor's University, Malaysia, and Fujian University of Technology, China. He secured an outstanding leadership award at the 2009 IEEE International Conference on Networks and Systems Security in Australia. He has been included in the Marquis Who's Who in the World 2010 edition. In addition, at the AIT international Conference in Japan in June 2010, he received a certificate of appreciation for outstanding contributions in "Biometrics and Information Security Research". He has been awarded a Gold Medal for the 'Best Invention and Innovation Award' at the 10th Malaysian Technology Expo 2011, Malaysia. Moreover, in April 2013, his invention won a Bronze Medal at the

"41st International Exhibition of Inventions" in Geneva, Switzerland. In addition, he was awarded the best paper awards from the Journal of Network and Computer Applications (Elsevier) and the IEEE Systems Journal. He was the recipient of the King Saud University Award for Scientific Excellence (Research Productivity) in May 2015. He was also a recipient of the King Saud University Award for Scientific Excellence (Inventions, Innovations, and Technology Licensing) in May 2016.

He has published over 450 research papers and is the inventor of 10 US/PCT patents. He has edited nine books/proceedings published by Springer-Verlag, Taylor & Francis, and IEEE. He has secured several national and international competitive research grants totalling over USD 4 million in the domain of cybersecurity. He has played a leading role in developing the "BS Cybersecurity Degree Program" and the "Higher Diploma in Cybersecurity" at King Saud University.

His research areas of interest are cybersecurity, IoT security, automotive cybersecurity, digital authentication, multimedia security, cyber policy, and technological innovation management. He is a fellow of the Institution of Engineering & Technology (UK), a fellow of the British Computer Society (UK), a fellow of the FTRA (Korea), a senior member of the IEEE, and a senior member of the IACSIT (Singapore). He is also the Vice Chair of the IEEE Communications Society Saudi Chapter. He has been a distinguished lecturer of the IEEE (CTSoC) from 2019–2021. He is currently working as a chair of the task group (TG) of the International Telecommunication Union (ITU) on child online protection in the Metaverse. His research has been cited over 26,500 times with an h-index of 87 (Google Scholar) .

His detailed profile is available at <http://www.professorkhurram.com>

**Keynote Speaker 2: Mr Raja Segaran**

**Director, Head of Strategy, Research & Insights  
Malaysia Digital Economy Corporation (MDEC)**



**The Digital Economy and Its impact to Society, Business and  
Government: A Malaysia's Perspective**

**Abstract**

The digital economy, a transformative force powered by technological advancements, is reshaping the fabric of society, business, and government. In this dynamic landscape, the profound impact of digitization is evident across various facets of our interconnected world.

**Society**

The digital economy is a catalyst for societal evolution, fostering connectivity and changing the way individuals interact, communicate, and access information. With the ubiquity of smartphones and high-speed internet, communities are more connected than ever. This connectivity has facilitated the rise of digital communities, enabling individuals to engage in social, educational, and economic activities across borders. The digital economy necessitates a concerted effort to bridge the digital divide, ensuring that the benefits of technology are inclusive and accessible to all.

**Business**

In the business realm, the digital economy is a driving force behind innovation, efficiency, and global competitiveness. Companies embracing digital transformation are optimizing operations, leveraging data analytics, and adopting artificial intelligence to enhance decision-making. E-commerce has revolutionized

the way businesses reach consumers, breaking down geographical barriers and creating new avenues for growth. Yet, this transformation brings challenges such as cybersecurity threats and the need for upskilling the workforce to navigate the evolving technological landscape.

### **Government**

Governments worldwide are grappling with the complexities of governing in the digital age. The digital economy demands agile regulatory frameworks to foster innovation while safeguarding consumer rights and data privacy. Smart governance, enabled by digital technologies, enhances public service delivery, streamlines administrative processes, and promotes transparency. However, it also raises questions about cybersecurity, digital rights, and the ethical use of emerging technologies. Governments are navigating this terrain, seeking to strike a balance between encouraging innovation and safeguarding the public interest.

In summary, the digital economy is a transformative force that permeates every aspect of modern life. Its impact on society, business, and government is profound, presenting opportunities for progress and posing challenges that require careful navigation. As we navigate this digital frontier, a collaborative, inclusive, and forward-looking approach is essential to ensure that the benefits of the digital economy are realized by all.

### **Biography**

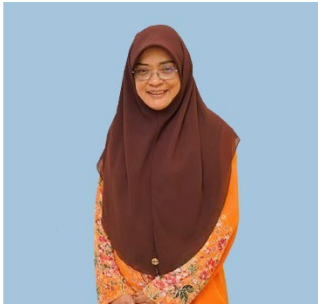
Raja is currently the Director of Strategy, Research & Insights in Malaysia Digital Economy Corporation (MDEC), an agency tasked by the Government to develop and lead the country's digital economy. He has led the development of various national macro-strategies and roadmaps for the country's digital economy and sits in several national-level platforms, committees and task forces to help the country to shape strategies and policies to deliver impactful outcomes.

Raja is an Adjunct Professor in UNITAR International University for Digital Strategies to support the university's goal to bring in industry perspectives in developing future-ready talents. He is also a Fellow with the Malaysian Institute of Management (FMIM).

At the international front, Raja is the Global Chair for the Digital Economy & Financial Inclusion (DEFI) Sub-Committee under the IEEE's Digital Resilience Industry Connection Program. Prior to joining MDEC, Raja was a Director with Ernst & Young Advisory in the Kuala Lumpur office as one of the leaders in the Transformation domain. He has led the set-up and management of several medium to large transformation offices for various multinational and local clients in the private sector.

### **Keynote Speaker 3: Professor Dr Naomie Salim**

**Universiti Teknologi Malaysia**



### **Mining the Malaysian Social Media: Challenges and Solutions**

#### **Abstract**

Social media is widely used as a means of interactions among virtual communities as more and more people openly create, share, and/or exchange information and ideas in Facebook, Twitter, Instagram, LinkedIn, YouTube and online forums. With the increasing amount of data available on social media channels, valuable insights into how people feel about products, person, issues, initiatives or services can be obtained if the social media is monitored for mentions of brand, competitor brands, initiatives, issues, persons and related keywords to help organizations make smarter, data-driven decisions. However, many of the currently available social media listening tools cannot be effectively used for monitoring the Malaysian social media. Malaysian use a lot of mixed languages such as Manglish. We also have our own Malay short forms, neologism, sarcasm and idioms. There are also issues of other noises and spams in the Malaysian social media before it can be mined accurately. Many of the current rule-based approaches, machine learning-based approaches and their hybrids fails to accurately mine the Malaysian social media because of these issues. In this talk, we will discuss the approaches that can be used to handle the issues, such as using our own valence shifter, Malaysian lexicons, and sarcasm handler. Finally, we will discuss how deep learning, data fusion and Generative Pre-trained Transformer tools can be used to mine the Malaysian social media.

## **Biography**

Dr Naomie is a professor at the Faculty of Computing UTM. She obtained her PhD (Computational Informatics) from University of Sheffield. She is currently the Director of UTM Big Data Centre. She works on design of new algorithms to improve the effectiveness of searching and mining of new knowledge from various types of data. She has been involved in 57 research projects with a total amount of nearly RM5 million. She has published more than 400 papers. She has been awarded as 2020 Malaysia Top Research Scientist by Akademi Sains Malaysia. Some of her other awards are PECIPTA 2011 Gold Medal award for her cross-language semantic plagiarism detection system, the I-inova 2010 Gold Medal award for her Islamic Ontology-based Quran search engine, BioInnovation 2011 Bronze Award for UTMChem Workbench Molecular Database System, iPhex Gold Medal Award for innovation in teaching and learning, UTM 2011 Best Research Award and UTM 2014 Best Research Award. Her innovation in big data, data warehousing and mobile applications has benefited industry such as MAMPU, Johor Corporations, Malaysia Marine and Heavy Engineering Ptd. Ltd., Malaysia Higher Education Ministry and Companies Commission of Malaysia (SSM). 48 PhD students have graduated under her supervision. She was a fellow of the Japan Society for Promotion of Science and is the Head of the Malaysian Big Data Research Consortium.

## **Keynote Speaker 4: Dato' Dr. Amirudin Abdul Wahab**

**CEO, Cybersecurity Malaysia**



### **Cyber Treats in the Age of Emerging Tech**

#### **Abstract**

In today's digitally connected world, the importance of cybersecurity cannot be underestimated. The convergence of various technologies adds to the complexity of cyberspace. The future will definitely involve a lot more technologies. Artificial Intelligence, the Internet of Things, cloud services, and 5G technology are among many technologies that have emerged today. These advancements provide opportunities for various users to explore while also creating opportunities for cybercriminals to exploit. As ever-evolving cyber threats are increasing, they pose serious challenges to individuals, organisations, and even nations. However, these threats hide untapped opportunities for innovation, collaboration, and advances in cybersecurity. This session will examine the shift in the cybersecurity landscape and of using cybersecurity as a mere defensive measure.

#### **Biography**

DATO' Ts. DR HAJI AMIRUDIN ABDUL WAHAB is the Chief Executive Officer of CyberSecurity Malaysia, a cybersecurity specialist and technical agency that monitors e-sovereignty of the country.

He has more than 30 years of ICT working experience in telecom and IT sector in both public and private sectors. A PhD holder from University of Queensland, Australia, Master's in Business Administration (MBA) from University of Duquibue, Iowa, USA, Master's in information technology, National University of Malaysia

(UKM) and also a Bachelor of Science Engineering in Electrical Engineering from University of Michigan, Ann Arbor, USA.

Dato' Ts. Dr. Amir became the first Malaysian to chair World Trustmark Alliance (2014 - 2015); received Outstanding Chief Information/Security Officer at the 10th ASEAN CIO | CSO Summit and Awards 2014 in Vietnam and bestowed Distinguished Alumni Award Winner 2021 by Australian Alumni Awards. On 28 June 2022, Dato' Ts. Dr Amir was conferred 2022 World Digital Economy Distinguished Leader Award by the KSI Strategic Institute for Asia Pacific. In 22 Oct 2022, he has been awarded Blockchain Icon of the Year 2022.

Dato' Ts. Dr. Amir is an Adjunct Professor at various institutions of higher learning in Malaysia. He received numerous awards and recognition and the most recent is the Honorary Doctorate in Information Technology (Cybersecurity) from Malaysia Technical University Melaka (UTeM).

Dato' Ts. Dr. Amir has built a strong reputation for developing effective, unique strategies, incubating new models and designs, and maintains a well-deserved national and international reputation where he has been invited as a keynote speaker and panelist in various local and international conferences.

**Keynote Speaker 5: Professor Dr Seifedine Kadry**

**Noroff Education AS, Norway**



**Computer Vision based Plants Phenotyping: Challenges, Opportunities,  
and Perspectives**

**Abstract**

The increasing demand for food production due to the growing population is raising the need for more productive plant environments. The genetic behavior of plant traits remains different in different growing environments. However, it is tedious and impossible to manually look after the individual plant component traits. Plant breeders need computer vision-based plant monitoring systems to analyze a more productive, suitable environment, quantitative analysis, geometric analysis, and yield rate analysis. Many of the data collection methods according to the needs have been used by plant breeders. In the presented review, most of them are discussed with their corresponding challenges and limitations.

Furthermore, the traditional approaches of segmentation and classifications of plant phenotyping are also discussed. The data limitation problems and their currently adapted solutions in the computer vision aspect are highlighted that somehow solve the problem but are not genuine. The available data sets and current issues are enlightened. The presented study covers the data collection to classification tasks.

## **Biography**

He has a bachelor's degree in 1999 from Lebanese University, an MS degree in 2002 from Reims University (France) and EPFL (Lausanne), Ph.D. in 2007 from Blaise Pascal University (France), an HDR degree in 2017 from Rouen University (France). His research currently focuses on Data Science, education using technology, system prognostics, stochastic systems, and applied mathematics. He is an ABET program evaluator for computing, Data Science, and Engineering Tech. He is a full professor of data science at Noroff University College, Norway. He is an IET Fellow.

**Keynote Speaker 6: Ms Puteri Anis Aneeza binti Zakaria**

**Senior Technical Manager, Statworks Group**



**Leveraging XLSTAT for Data Science: A Comprehensive Overview and Practical Applications**

**Abstract**

XLSTAT, a statistical analysis software seamlessly integrated into Microsoft Excel, has gained popularity for its accessibility and user-friendly interface. XLSTAT serves as a valuable bridge between traditional statistical analyses and the expansive field of data science. The key applications of XLSTAT in data science include descriptive statistics, data visualization, hypothesis testing, regression analysis, multivariate analysis, and time series analysis. The software's capabilities in these domains enable users to gain insights into data patterns, relationships, and trends. Its integration with Excel and a rich set of statistical tools makes it accessible to a wide range of users, particularly those accustomed to Excel's interface.

This presentation provides an overview of how XLSTAT can be a valuable tool in data science and offers a range of statistical and analytical functionalities that cater to users who prefer a familiar Excel environment.

## **Biography**

Puteri Anis Aneza binti Zakaria is a skilled and dynamic professional trainer with a passion for empowering individuals and teams to reach their best. With over ten years of experience in education and analytic software, she has led many successful workshops, seminars, and corporate training sessions. Puteri Anis Aneza is known for her engaging and innovative style, creating environments that support learning and success. Her dedication and proven track record make her a trusted expert in the field of professional training.

**Keynote Speaker 7: Associate Professor Dr Chin Kim On**

**Universiti Malaysia Sabah**



**Support Vector Machines for Efficient Identification and Classification of Orangutan Nests**

**Abstract**

The Bornean orangutan population is critically endangered, and efficient methods for identifying and classifying their nests are needed to monitor their populations and track their health. Manual identification and classification are time-consuming and require extensive knowledge of orangutan behavior and their natural habitats. We propose a novel approach to classifying orangutan nests using Support Vector Machines (SVMs). SVMs are a powerful machine learning algorithm that has been shown to be effective in image classification tasks. Our method incorporates a comprehensive image pre-processing and enhancement pipeline, along with a rigorous training and testing procedure. Experimental results indicate that our method can achieve remarkable accuracy in classifying orangutan nests, with an achieved accuracy of 79.90%, an F1-score of 79.87%, a precision of 79.91%, and a recall of 79.82%. These findings underscore the efficacy of our proposed method and highlight the potential of SVMs for this critical conservation task. The proposed method has significant implications for orangutan conservation efforts, enabling researchers and conservationists to efficiently identify and monitor orangutan populations. By automating the nest classification process, it becomes possible to analyze large-scale datasets and gain valuable insights into orangutan behavior, habitat preferences, and population dynamics. In conclusion, our study highlights the effectiveness of SVMs in classifying orangutan nests and contributes to the ongoing efforts in preserving these critically endangered species. The developed methodology offers a

promising solution for wildlife conservationists and researchers working towards the preservation of Bornean orangutans.

## **Biography**

Kim On, Chin is an Associate Professor at the Faculty of Computing and Informatics, Universiti Malaysia Sabah. His research interests encompass evolutionary computing, machine learning, image processing, Internet of Things (IoT), and biometric security systems. He has successfully secured 30 research grants and 14 community grants, amounting to almost RM 5 million in total. He has authored or co-authored about 130 publications, including journal articles, book chapters, and conference proceedings. As a Senior Member of IEEE, Kim On has actively engaged in over 80 community-based activities. He has developed web applications for teaching and learning using Google Tools and Google Data Studio, and conducted training sessions on Computational Thinking, Arduino, 3D printing, and Python programming. He has also undertaken consultancy projects involving the management of Digital Maker Hubs for Science, Technology, Engineering, Arts, and Mathematics (STEAM), the development of an e-summon system utilizing image processing and neural networks for campus surveillance, and plastic object classification through machine learning-based classifiers for onshore plastic waste detection using near-infrared hyperspectral images. Kim On has more than 18 years of teaching experience, and is highly skilled in delivering professional training courses and giving invited talks. His expertise and contributions in the field of computer science and informatics are significant, making him a valuable asset to the academic community. He is also one of the appointed MBOT and MQA Assessors for computer science programme accreditation.

## **Keynote Speaker 8: Professor Dr Schahram Dustdar**

**TU Wien, University Technology Vienna**



### **Building Intelligent Sustainable Internet-based Ecosystems**

#### **Abstract**

Modern distributed systems also deal with uncertain scenarios, where environments, infrastructures, and applications are widely diverse. In the scope of IoT-Edge-Fog-Cloud computing, leveraging these neuroscience-inspired principles and mechanisms could aid in building more flexible solutions able to generalize over different environments. A captivating set of hypotheses from the field of neuroscience suggests that human and animal brain mechanisms result from few powerful principles. If proved to be accurate, these assumptions could open a deep understanding of the way humans and animals manage to cope with the unpredictability of events and imagination.

#### **Biography**

Schahram Dustdar is a Full Professor of Computer Science at the TU Wien, heading the Research Division of Distributed Systems, Austria. He holds several honorary positions: University of California (USC) Los Angeles; Monash University in Melbourne, Shanghai University, Macquarie University in Sydney, University Pompeu Fabra, Barcelona, Spain. From Dec 2016 until Jan 2017 he was a Visiting Professor at the University of Sevilla, Spain and from January until June 2017 he was a Visiting Professor at UC Berkeley, USA.

From 1999 – 2007 he worked as the co-founder and chief scientist of Caramba Labs Software AG in Vienna (acquired by ProjectNetWorld AG), a venture capital co-funded software company focused on software for collaborative processes in teams. He is co-founder of edorer.com (an EdTech company based in the US) and co-founder and chief scientist of Sinoaus.net, a Nanjing, China based R&D organization focusing on IoT and Edge Intelligence.

He serves as Editor-in-Chief of Computing (Springer). Dustdar is recipient of multiple awards: IEEE TCSVC Outstanding Leadership Award (2018), IEEE TCSC Award for Excellence in Scalable Computing (2019), ACM Distinguished Scientist (2009), ACM Distinguished Speaker (2021), IBM Faculty Award (2012). He is an elected member of the Academia Europaea: The Academy of Europe, as well as an IEEE Fellow(2016) and an Asia-Pacific Artificial Intelligence Association (AAIA) Fellow (2021) and the AAIA president (since 2021).



## **PROGRAMME SCHEDULE**

### **International Conference on Data Science and Emerging Technologies (DaSET 2023)**

*Towards Green Artificial Intelligence and Sustainable Computing*

### **Conference Schedule DaSET 2023 - The International Conference on Data Science and Emerging Technologies**

#### **Virtual Links:**

- [Main Room](#)
- [Room 1](#)
- [Room 2](#)
- [Room 3](#)

**Day 1: Monday, 4<sup>th</sup> December 2023**

Time	Activities
8:00 am -9:00 am MYT	Online Registration
9:00 am -9:05 am MYT	( <u>Main Room</u> ) National Anthem & Doa Recital
9:05 am -9:10 am MYT	<b>Welcoming Remarks</b>  <b>Professor Dr Yap Bee Wah &amp; Professor Dr Dhiya Al-Jumeily, OBE</b>  Conference Chairs UNITAR International University, Malaysia & Liverpool John Moores University, UK
9:10 am -9:20 am MYT	<b>Opening Speech by:</b>  <b>Dato' Sri Dr. Mohd Uzir Mahidin</b> Chief Statistician, Department of Statistics Malaysia
9:20 am - 9:25 am MYT	Montage DaSET2023
9:25 am -9:30 am MYT	Online Photography Session
9:30 am -9:45 am MYT	<b>Special Address by:</b>  <b>Professor Witold Pedrycz</b> Professor, Faculty of Engineering - Electrical & Computer Engineering Department, University of Alberta, Canada
9:45 am -10:30 am MYT	<b>Keynote Speaker 1</b>  <b>Professor Dr Muhammad Khuram Khan</b> King Saud University, Saudi Arabia  <i>Keynote Title: Security and Privacy in Connected and Autonomous Vehicles: State-of-the-art and Future Directions</i>  <b>Moderator: Dr Aitizaz Ali</b> UNITAR International University, Malaysia

10:30 am -10:45 am MYT	Break (play video- UNITAR Video/Song)
10:45 am -11:30 am MYT	<p><b>Keynote Speaker 2</b></p> <p><b>Mr Raja Segaran</b> Director, Head of Strategy, Research &amp; Insights Malaysia Digital Economy Corporation (MDEC) &amp; Adjunct Professor UNITAR International University, Malaysia</p> <p><i>Keynote Title: The Digital Economy and Its Impact to Society, Business and Government: A Malaysia's Perspective</i></p> <p><b>Moderator: Dr Esther Loo</b> Head, Organisational Design and Effectiveness Malaysia Airlines Bhd &amp; Adjunct Professor UNITAR International University, Malaysia</p>
11:30 am -12:15 pm MYT	<p><b>Keynote Speaker 3</b></p> <p><b>Professor Dr Naomie Salim</b> Universiti Teknologi Malaysia</p> <p><i>Keynote Title: Mining the Malaysian Social Media: Challenges and Solutions</i></p> <p><b>Moderator: Dr Abdulaziz Yahya Al-Nahari</b> UNITAR International University, Malaysia</p>
12:15 pm - 1:30 pm MYT	<p><b><u>Paper Presentations (Room 1)</u></b></p> <p>Track 1a: Artificial Intelligence (5 papers) Each paper 15 minutes</p> <p><b>Moderator: Dr Azlin Ahmad</b> Universiti Teknologi MARA, Malaysia</p>
	<p><b><u>Paper Presentations (Room 2)</u></b></p> <p>Track 2a: Machine &amp; Deep Learning (5 papers) Each paper 15 minutes</p> <p><b>Moderator: Ts. Dr Syarifah Bahiyah Rahayu Binti Syed Mansoor</b> Universiti Pertahanan National Malaysia, Malaysia</p>

1:30 pm -2:30 pm MYT	Lunch Break
2:30 pm -3:15 pm MYT	<p><b>Keynote Speaker 4</b></p> <p><b>Dato' Dr. Amirudin Abdul Wahab</b> CEO, Cybersecurity Malaysia &amp; Adjunct Professor UNITAR International University, Malaysia</p> <p><i>Keynote Title: Cyber Treats in the Age of Emerging Tech</i></p> <p><b>Moderator: Dr Tahir Mehmood</b> UNITAR International University, Malaysia</p>
3:15 pm - 4:30 pm MYT	<p><b><u>Paper Presentations (Room 1)</u></b></p> <p>Track 1b: Artificial Intelligence (3 papers) Each paper 15 minutes</p> <p><b>Moderator: Assoc. Prof. Dr Syerina Azlin Mohd Nasir</b> Universiti Teknologi MARA, Malaysia</p>
	<p><b><u>Paper Presentations (Room 2)</u></b></p> <p>Track 2b: Machine &amp; Deep Learning (5 papers) Each paper 15 minutes</p> <p><b>Moderator: Ts. Dr Syarifah Bahiyah Rahayu Binti Syed Mansoor</b> Universiti Pertahanan National Malaysia, Malaysia</p>
	<p><b><u>Paper Presentations (Room 3)</u></b></p> <p>Track 3a: Statistical Learning (5 papers) Each paper 15 minutes</p> <p><b>Moderator: Assoc. Prof. Dr Wan Fairos Wan Yaacob</b> Universiti Teknologi MARA, Malaysia</p>
4:30 pm - 5:15 pm MYT	<p><b>Keynote Speaker 5</b></p> <p><b>Professor Dr Seifedine Kadry</b> Noroff Education AS, Norway</p>

	<p><i>Keynote Title: Computer Vision based Plants Phenotyping: Challenges, Opportunities, and Perspectives</i></p> <p><b>Moderator: Professor Dr Heri Kuswanto</b> Institut Teknologi Sepuluh Nopember, Indonesia</p>
5:15 pm - 5:20 pm MYT	Closing of Day 1

**Day 2: Tuesday, 5th December 2023**

Time	Activities
08:00 am- 09:00 am MYT	Online Registration
09:00 am - 09:15 am MYT	Welcoming Remarks
09:15 am - 10:00 am MYT	<p><b>Keynote Speaker 6</b></p> <p><b>Ms Puteri Anis Aneeza binti Zakaria</b> Senior Technical Manager, Statworks Group</p> <p><i>Keynote Title: Leveraging XLSTAT for Data Science: A Comprehensive Overview and Practical Applications</i></p> <p><b>Moderator: Dr Iznora Aini Binti Zolkifly</b> UNITAR International University, Malaysia</p>
10:00 am – 10:45 am MYT	<p><b>Keynote Speaker 7</b></p> <p><b>Associate Professor Dr Chin Kim On</b> Universiti Malaysia Sabah</p> <p><i>Keynote Title: Support Vector Machines for Efficient Identification and Classification of Orangutan Nests</i></p> <p><b>Moderator: Assoc. Prof. Dr Wan Fairos Wan Yaacob</b> Universiti Teknologi MARA, Malaysia</p>
10:45 am – 11:00 am MYT	<b>Break</b>
11:00 am -12:15 pm MYT	<p><b><u>Paper Presentations (Room 1)</u></b></p> <p>Track 4a: Health &amp; Industrial Applications (5 papers) Each paper 15 minutes</p> <p><b>Moderator: Dr Umar Farooq Khattak</b> UNITAR International University, Malaysia</p>
	<p><b><u>Paper Presentations (Room 2)</u></b></p> <p>Track 2c: Machine &amp; Deep Learning (4 papers) Each paper 15 minutes</p> <p><b>Moderator: Assoc. Prof. Dr Farhad Nadi</b> UNITAR International University, Malaysia</p>

	<p><b><u>Paper Presentations (Room 3)</u></b></p> <p>Track 3b: Statistical Learning (5 papers) Each paper 15 minutes</p> <p><b>Moderator: Professor Dr Sapto Wahyu Indratno</b> Institut Teknologi Bandung, Indonesia</p>
12:15 pm - 02:15 pm MYT	Lunch Break (play video)
02:15 pm - 03:15 pm MYT	<p><b><u>Paper Presentations (Room 1)</u></b></p> <p>Track 4b: Health &amp; Industrial Applications (4 papers) Each paper 15 minutes</p> <p><b>Moderator: Dr Wasiq Khan</b> Liverpool John Moores University, UK</p>
	<p><b><u>Paper Presentations (Room 3)</u></b></p> <p>Track 3c: Statistical Learning (2 papers) Each paper 15 minutes</p> <p><b>Moderator: Dr Sulaf Assi</b> Liverpool John Moores University, UK</p>
03:15 pm - 03:30 pm MYT	<b>Break</b>
03:30 pm - 04:15 pm MYT	<p><b>Keynote Speaker 8</b></p> <p><b>Professor Dr Schraham Dustdar</b> Tu Wien, Vienna University of Technology</p> <p><i>Keynote Title: Building intelligent sustainable Internet-based ecosystems</i></p> <p><b>Moderator: Dr Hadi Naghavipour</b> UNITAR International University, Malaysia</p>
4:15 pm - 4:45 pm MYT	<p>Best Paper Awards Closing Ceremony</p>

## Paper Presentations Schedule

<b>Track 1a: Artificial Intelligence</b>				
Venue: <a href="#">Room 1</a>				
Date: 4 <sup>th</sup> December 2023 (Monday)				
Session 1: 12:15 pm – 1:30 pm MYT				
Session Chair: Dr Azlin Ahmad, Universiti Teknologi MARA, Malaysia				
NO	TIME	ID	TITLE	AUTHORS
1.	12.15 pm – 12.30 pm MYT	1570950538	A Comparative Study of Lemmatization Approaches for Rojak Language	Liu Jun Yoon, Xuan Yi Tan , Khai Yin Lim, Chi Wee Tan, Ling Ern Cheng and Jenny Tan
2.	12.30 pm – 12.45 pm MYT	1570952139	Multi-Aspect Extraction in Indonesian Reviews through Multi-Label Classification using Pre-trained BERT Models	Nur Hayatin, Suraya Alias, Lai Po Hung and Yuliana Setiowati
3.	12.45 pm – 1.00 pm MYT	1570953952	Artificial Intelligence (AI) Empowered Sign Language Recognition Using Hybrid Neural Networks	Ambar Saxena, Nailya Sultanova, Jamila Mustafina and Noor Lees Ismail
4.	1.00 p.m – 1.15 pm MYT	1570959215	The Performance of GPT-3.5 in Summarizing Scientific and News Articles	Sabkat Arshad, Muhammad Yaqoob and Tahir Mehmood
5.	1.15 p.m – 1.30 pm MYT	1570948456	Wound Stage Recognition using YOLOv5	Clair Abela and Frankie Inguanez

<b>Track 2a: Machine &amp; Deep Learning</b>				
Venue: <a href="#">Room 2</a> Date: 4th December 2023 (Monday) Session 1: 12:15 pm – 1:30 pm MYT Session Chair: Ts. Dr Syarifah Bahiyah Rahayu Binti Syed Mansoor, Universiti Pertahanan National Malaysia, Malaysia				
NO	TIME	ID	TITLE	AUTHORS
1.	12.15 pm – 12.30 pm MYT	1570954006	Fraudulent e-Commerce Website Detection Using Convolutional Neural Network based on Image Features	Nurfazrina Mohd Zamry, Anazida Zainal, Eric Khoo, Mohamad Nizam Kassim and Zanariah Zainudin
2.	12.30 pm – 12.45 pm MYT	1570959213	Unlocking Security: Advanced Ransomware Prediction and Classification with Artificial Neural Networks	Saaman Nadeem, Tahir Mehmood and Muhammad Yaqoob
3.	12.45 pm – 1.00 pm MYT	1570948608	Leveraging Gamification for Engaged Learning in Online Teaching and Learning Experiences	Norshahriah Abdul Wahab, A'tifah Hanim Rosli, Syarifah Bahiyah Rahayu Syed Mansoor, Norazliana Akmal Jamaludin and Siti Hajar Adam
4.	1.00 p.m – 1.15 pm MYT	1570952240	Sentiment Analysis Using Large Language Models: A Case Study of GPT-3.5	Farhad Nadi, Hadi Naghavipour, Tahir Mehmood, Alliesya Binti Azman, Jeetha A/P Nagantheran, Kezia Sim Kui Ting, Nor Muhammad Ilman Bin Nor Adnan, Roshene A/P Sivarajan, Suita A/P Veerah and Romi Fadillah Rahmat

5.	1.15 p.m – 1.30 pm MYT	1570953458	Telecom Customer Experience Analysis using Sentiment Analysis and Natural Language Processing – Comparative Study	Ahmed Mohamed Abdou Ahmed, Raghad Al_Shabandar, Abdulaziz Al-Nahari and A.H. Mohammed
----	---------------------------	------------	---	---

**Track 1b: Artificial Intelligence**

Venue: [Room 1](#)

Date: 4th December 2023 (Monday)

Session 2: 3:15 pm – 4:30 pm MYT

Session Chair: Associate Professor Dr Syerina Azlin Mohd Nasir, Universiti Teknologi MARA, Malaysia

NO	TIME	ID	TITLE	AUTHORS
1.	3.15 pm – 3.30 pm MYT	1570951781	Harvest Palm Tree Based on Detection through 2D LiDAR Sensor Using Power Equation	Luqman Hakim Bin Yusof, Abdulaziz Yahya Yahya Al-Nahari, Danny Ngo Lung Yao and Normaiza Mohamad
2.	3.30 pm – 3.45 pm MYT	1570954165	Enhancing Security Surveillance through Business Intelligence with NVIDIA DeepStream	Vishal Pednekar, Nidhi Shettigar, Sayli Tawhare
3.	3.45 pm – 4.00 pm MYT	1570955846	Fuzzified Hybrid Metaheuristics for QoS-aware Service Composition	Hadi Naghavipour, Farhad Nadi, and Ali Aitizaz

**Track 2b: Machine & Deep Learning**

Venue: [Room 2](#)

Date: 4th December 2023 (Monday)

Session 2: 3:15 pm – 4:30 pm MYT

Session Chair: Ts. Dr Syarifah Bahiyah Rahayu Binti Syed Mansoor, Universiti Pertahanan National Malaysia, Malaysia

NO	TIME	ID	TITLE	AUTHORS
1.	3.15 pm – 3.30 pm MYT	1570953467	Efficient Time Series of Smoothing and Auto-regressive Forecasting Models for Predicting Police Officer Fatalities in the USA	Danush Nagappan, Manoj Jayabalan, Ahmad Alanezi, Farhad Nadi, Thomas Coombs
2.	3.30 pm – 3.45 pm MYT	1570953476	Multimodal Emotion Recognition using Attention-based Model with language, Audio, and Video Modalities	Disha Sharma and Manoj Jayabalan, Nailya Sultanova, Jamila Mustafina and Danny Ngo Lung Yao
3.	3.45 pm – 4.00 pm MYT	1570953964	Comparative Analysis of Emotion Recognition using Large-Language Models and Conventional Machine Learning	Mangu Soujanya Rao, Thomas Coombs, Normaiza Binti Mohamad, Vinay Kumar, Manoj Jayabalan
4.	4.00 pm – 4.15 pm MYT	1570955167	The Impact of Clustering-Based Sequential Multivariate Outliers Detection in Handling Missing Values	Mety Agustini, Kartika Fithriasari and Dedy Dwi Prastyo
5.	4.15 pm – 4.30 pm MYT	1570956190	Sarcasm Detection in Newspaper Headlines	Vishnu Sai Reddy Chilpuri, Saaman Nadeem and Tahir Mehmood

**Track 3a: Statistical Learning**

Venue: [Room 3](#)

Date: 4th December 2023 (Monday)

Session 2: 3:15 pm – 4:30 pm MYT

Session Chair: Associate Professor Dr Wan Fairos Wan Yaacob, Universiti Teknologi MARA, Malaysia

NO	TIME	ID	TITLE	AUTHORS
1.	3.15 pm – 3.30 pm MYT	1570943390	A Case Study via Bayesian Network: Investigating Factors Influencing Student Academic Performance in Online Teaching and Learning During Covid-19 Pandemic	Zheng Ning Looi, Poh Choo Song, Huai Tein Lim and Sing Yan Looi
2.	3.30 pm – 3.45 pm MYT	1570949397	Harnessing the XG-Boost Ensemble for Intelligent Prediction and Identification of Factors with a High Impact on Air Quality: A Case Study of Urban Areas in Jakarta Province, Indonesia	Wahyu Wibowo, Harun Al Azies, Susi A. Wilujeng and Shuzlina Abdul-Rahman
3.	3.45 pm – 4.00 pm MYT	1570949756	Modeling Earthquake Catalog in Sumatra by Space-Time Epidemic-Type Aftershock Sequences Model: Combining Davidon Fletcher Powell and Stochastic Declustering Algorithms	Christopher Andreas, Achmad Choiruddin and Dedy Dwi Prastyo
4.	4.00 pm – 4.15 pm MYT	1570949888	Small Area Estimation of Mean Years of Schooling Under Time Series and Cross-sectional Models	Reny Ari Noviyanti, Setiawan and Agnes Tuti Rumiasi
5.	4.15 pm – 4.30 pm MYT	1570950075	Probabilistic Seismic Hazard Analysis for Sulawesi-Maluku Region of Indonesia using the Space-Time Epidemic Type Aftershock Sequence Model	Sonia Faradilla, Achmad Choiruddin and Bambang Widjanarko Otok

<b>Track 4a: Health &amp; Industrial Applications</b>				
Venue: <a href="#">Room 1</a> Date: 5 <sup>th</sup> December 2023 (Tuesday) Session 3: 11:00 am – 12.15 pm MYT Session Chair: Dr Umar Farooq Khattak, UNITAR International University, Malaysia				
NO	TIME	ID	TITLE	AUTHORS
1.	11.00 am – 11.15 am MYT	1570956304	W@RK: ATTENDANCE APPLICATION FRAMEWORK USING BLOCKCHAIN TECHNOLOGY	Putra Roskhairul Fitri Kaha, Syarifah Bahiyah Rahayu, Afiqah M. Azahari, Mohd Hazali Mohamed Halip and K. Venkatesan
2.	11.15 am – 11.30 am MYT	1570950111	Exploring the Impact of COVID-19 on Individuals' Mental Health through Cluster Analysis	Azlin Ahmad, Siti Nabilah Mohd Abdul Hakim Amir, Ezzatul Akmal Kamaruzaman and Abdulaziz Al-Nahari
3.	11.30 am – 11.45 am MYT	1570953200	Enhancing Data Collection Efficiency through Multithreaded Web Scraping Approach	Dahlan Nariman
4.	11.45 am – 12.00 pm MYT	1570958670	Tracking High Potential Transmission Risk Spots of Infectious Disease Using Spatial Social Network Analysis and Visualisation (SSNAV) Techniques	Ilham Abdul Jalil and Abdul Rauf Abdul Rasam
5.	12.00 pm – 12.15 pm MYT	1570950594	Breast Cancer Classification: A Transfer Learning Approach Using EfficientNets	Samiya QANOUNE, Hassan AMMOR and Zouhair GUENNOUN

**Track 2c: Machine & Deep Learning**

Venue: [Room 2](#)

Date: 5<sup>th</sup> December 2023 (Tuesday)

Session 3: 11:00 am – 12.15 pm MYT

Session Chair: Associate Professor Dr Farhad Nadi, UNITAR International University, Malaysia

NO	TIME	ID	TITLE	AUTHORS
1.	11.00 am – 11.15 am MYT	1570956272	Transformer-Based Named Entity Recognition Model - Tamil Language	Karthi Dhayalan, Nailya Sultanova, Jamila Mustafina and Paridah Daud
2.	11.15 am – 11.30 am MYT	1570956523	A Comparative Study of Methods for Topic Modelling in News Articles	Swapna D Rajan, Thomas Coombs, Manoj Jayabalan and Noor Azma Ismail
3.	11.30 am – 11.45 am MYT	1570956622	Application of Deep Learning Algorithms to Terahertz Images for Detection of Concealed Objects	Soumen Sardar, Sulaf Assi, Iznora Aini Zolkifly, Manoj Jayabalan, Manea Alsaleem, Ammar H. Mohammed and Dhiya Al-Jumeily OBE
4.	11.45 am – 12.00 pm MYT	1570958001	Multivariate Comparative Analysis of Statistical and Deep Learning Models for Prediction Hardware Failure	Saurabh Gupta, Raghad Alshabandar Chong Kim Loy, and Ammar H. Mohammed

**Track 3b: Statistical Learning**

Venue: [Room 3](#)

Date: 5<sup>th</sup> December 2023 (Tuesday)

Session 3: 11:00 am – 12.15 pm MYT

Session Chair: Professor Dr Supto Wahyu Indratno, Institut Teknologi Bandung, Indonesia

NO	TIME	ID	TITLE	AUTHORS
1.	11.00 am – 11.15 am MYT	1570950211	Application of Time Series Regression, Double Seasonal ARIMA, and Long Short-Term Memory for Short-Term Electricity Load Forecasting	Hafez Afghan and Hidayatul Khusna
2.	11.15 am – 11.30 am MYT	1570950306	A Bayesian Network for Classifying and Predicting Ship Collision	Iis Dewi Ratih, Ketut Buda Artana, Heri Kuswanto, Emmy Pratiwi and Muhammad Farhan Nuari
3.	11.30 am – 11.45 am MYT	1570950508	Outlier Detection in Simultaneous Equations with Panel Data	Suci Ismadyaliana, Setiawan and Jerry Dwi Trijoyo Purnomo
4.	11.45 am – 12.00 pm MYT	1570952489	Fast Maximum Likelihood Estimation of Large Data Spatial Error Model: Matrix Exponential Spatial Specification Approach	Marsono, Setiawan and Heri Kuswanto
5.	12.00 pm – 12.15 pm MYT	1570953393	Assessing Departmental Efficiency at Sepuluh Nopember Institute of Technology: A Comparative Study Using Classical and Advanced Data Envelopment Analysis	Zakiatul Wildani, M Naufal Mubarik, Sri Pingit Wulandari, Lucia Aridinanti and Muhammad Alifian Nuriman

**Track 4b: Health & Industrial Applications**

Venue: [Room 1](#)

Date: 5<sup>th</sup> December 2023 (Tuesday)

Session 4: 02:15 pm – 03.15 pm MYT

Session Chair: Dr Wasiq Khan, Liverpool John Moores University, UK

NO	TIME	ID	TITLE	AUTHORS
1.	02.15 pm – 02.30 pm MYT	1570955852	Evaluation Of Machine Learning Algorithms For Early Prediction Of Liver Disease	Sushmitha Geddam, Sulaf Assi, Hadi Naghavipour, Manoj Jayabalan, Abdullah Al-Hamid and Dhiya Al-Jumeily OBE
2.	02.30 pm – 02.45 pm MYT	1570957416	An agricultural information recommendation method based on matrix decomposition knowledge graph algorithm	Ruipeng Tang, Narendra Kumar Aridas and Mohamad Sofian Abu Talip
3.	02.45 pm – 03.00 pm MYT	1570947738	Performance of the Auxiliary Information Based Hybrid EWMA Chart with Fast Initial Response	Peh Sang Ng, Huai Tein Lim, Wai Chung Yeong and Sajal Saha
4.	03.00 pm – 03.15 pm MYT	1570958440	Evaluation of Machine Learning Models for Breast Cancer Detection in Microarray Gene Expression Profiles	Mohammad Nasir Abdullah and Yap Bee Wah

**Track 3c: Statistical Learning**

Venue: [Room 3](#)

Date: 5<sup>th</sup> December 2023 (Tuesday)

Session 4: 02:15 pm – 03.15 pm MYT

Session Chair: Dr Sulaf Assi, Liverpool John Moores University, UK

NO	TIME	ID	TITLE	AUTHORS
1.	02.15 pm – 02.30 pm MYT	1570953406	Multivariate Adaptive Fuzzy Clustering Means Regression Splines model using Generalized Cross-Validation (GCV) on stunting cases in Southeast Sulawesi.	Mira Meilisa, Bambang Widjanarko Otok and Jerry Dwi Trijoyo Purnomo
2.	02.30 pm – 02.45 pm MYT	1570955599	Statistical Inferences for Multivariate Generalized Gamma Regression Model	Hasbi Yasin, Purhadi and Achmad Choiruddin

## **Proceeding Abstracts**

### **Track 1: Artificial Intelligence**

<b>NO</b>	<b>ID</b>	<b>TITLE</b>	<b>AUTHORS</b>
1.	1570950538	A Comparative Study of Lemmatization Approaches for Rojak Language	Liu Jun Yoon, Xuan Yi Tan , Khai Yin Lim, Chi Wee Tan, Ling Ern Cheng and Jenny Tan
2.	1570952139	Multi-Aspect Extraction in Indonesian Reviews through Multi-Label Classification using Pre-trained BERT Models	Nur Hayatin, Suraya Alias, Lai Po Hung and Yuliana Setiowati
3.	1570953952	Artificial Intelligence (AI) Empowered Sign Language Recognition Using Hybrid Neural Networks	Ambar Saxena, Nailya Sultanova, Jamila Mustafina and Noor Lees Ismail
4.	1570959215	The Performance of GPT-3.5 in Summarizing Scientific and News Articles	Sabkat Arshad, Muhammad Yaqoob and Tahir Mehmood
5.	1570948456	Wound Stage Recognition using YOLOv5	Clair Abela and Frankie Inguanez
6.	1570951781	Harvest Palm Tree Based on Detection through 2D LiDAR Sensor Using Power Equation	Luqman Hakim Bin Yusof, Abdulaziz Yahya Yahya Al-Nahari, Danny Ngo Lung Yao and Normaiza Mohamad
7.	1570954165	Enhancing Security Surveillance through Business Intelligence with NVIDIA DeepStream	Vishal Pednekar, Nidhi Shettigar, Sayli Tawhare
8.	1570955846	Fuzzified Hybrid Metaheuristics for QoS-aware Service Composition	Hadi Naghavipour, Farhad Nadi, and Ali Aitizaz

## **A Comparative Study of Lemmatization Approaches for Rojak Language**

Liu Jun Yoon<sup>1</sup>, Xuan Yi Tan<sup>2</sup>, Khai Yin Lim<sup>3</sup>, Chi Wee Tan<sup>4</sup>, Ling Ern Cheng<sup>5</sup>, Jenny Tan<sup>6</sup>

<sup>1,2,3,5,6</sup> Department of Computing and Information Technology, Tunku Abdul Rahman University of Management and Technology, Penang Branch

<sup>4</sup> Faculty of Computing and Information Technology, Tunku Abdul Rahman University of Management and Technology, Kuala Lumpur Campus

oonlj-pm20@student.tarc.edu.my, tanxy-pm20@student.tarc.edu.my, limky@tarc.edu.my, chiwee@tarc.edu.my, chengle-pm20@student.tarc.edu.my, tanj-pm20@student.tarc.edu.my

**Abstract.** Lemmatization is an important preprocessing step in most Natural Language Processing (NLP) applications where it extracts a valid and linguistically meaningful lemma from an inflectional word. This allows different inflected forms of a word to be grouped into a common root which is the base-form or dictionary-form of a word, known as lemma. Due to the rapid spread of code-mixing languages like the Rojak language that mixes English with Malay, a lemmatizer capable of lemmatizing the language is needed for NLP applications involving this language. Thus, this work proposes a Rojak language lemmatization approach that is able to handle both languages without requiring users to input texts in different language separately. Various methods including rule-based, corpus-based, machine learning, and deep learning-based were experimented and compared using the (English Web Treebank) EWT and Indonesian GSD corpora from the Universal Dependencies (UD) framework. Besides, the effect of POS tags on the performance of lemmatizers was also evaluated based on the accuracy of the train and test sets. From the experiments conducted, the corpus-based approach produced the best results with 99.90% and 92.27% test set accuracy for Malay and English, respectively, whereas the deep learning-based with POS tag approach produced the worst results of 79.78% and 91.15%.

**Keywords:** Lemmatization, Rojak Language, Natural Language Processing (NLP), POS Tag.

## **Multi-Aspect Extraction in Indonesian Reviews through Multi-Label Classification using Pre-trained BERT Models**

Nur Hayatin<sup>1,2</sup>, Suraya Alias<sup>2</sup>, Lai Po Hung<sup>2</sup>, Yuliana Setiowati<sup>3</sup>

<sup>1</sup> Informatics Department, University of Muhammadiyah Malang, Indonesia

<sup>2</sup> Computing and Informatics Faculty University Malaysia Sabah, Kota Kinabalu, Malaysia

<sup>3</sup> Informatics and Computing Department, Politeknik Elektronika Negeri Surabaya, Indonesia

suealias@ums.edu.my

**Abstract.** Aspect extraction automatically identifies and categorizes specific aspects mentioned in the text to enable fine-grained opinion analysis in sentiment analysis. While previous research has successfully extracted aspects, they often focused on a single aspect per review, overlooking the presence of multiple aspects within sentences. This limitation has affected capturing a complete user opinion, thus posing a challenge despite the complexity of computations and annotations in supervised learning. In this study, we address the task of extracting multiple aspects from Indonesian reviews based on multi-label classification using Bidirectional Encoder Representations from Transformers (BERT) by implementing pre-trained models. BERT is chosen due to the ability to capture contextual information through bidirectional encoder mechanisms and the capacity to catch complex word and sentence relationships. In the experiment, we conducted the tests with various Indonesian pre-trained BERT models to enhance the performance of multi-aspect extraction on Indonesian hotel reviews. Our findings indicate that Indonesian-BERT-1.5G pre-trained model can improve the classifier performance and achieve an impressive F1-score of 0.84 with the lowest loss of 0.3029 for Indonesian reviews data.

**Keywords:** aspect extraction, BERT, user reviews, Indonesian pre-trained model, deep learning.

## **Artificial Intelligence (AI) Empowered Sign Language Recognition Using Hybrid Neural Networks**

Ambar Saxena<sup>1</sup>, Nailya Sultanova and Jamila Mustafina<sup>2</sup>, Noor Lees Ismail<sup>3</sup>

<sup>1</sup> Liverpool John Moores University, United Kingdom

<sup>1</sup>a.saxena@ljmu.stud.ac.uk

<sup>2</sup> Kazan Federal University, Russia

NaRNurutdinova@kpfu.ru, DNMustafina@kpfu.ru

<sup>3</sup> UNITAR International University, Malaysia

lees@unitar.my

**Abstract.** Hand gestures serve as the primary means of communication in sign languages, which are composed of visual gestures made up of hands, faces, and other bodily motions. Although sign language has become more common in recent years, communicating with sign language speakers or signers remains difficult for non-sign language speakers. There has been promising progress in the disciplines of motion and gesture detection utilizing Artificial Intelligent techniques because of recent advances in deep learning and computer vision. The deep learning network makes full use of the advantages of time series classification provided by the recurrent neural network model as well as the feature extraction capabilities of convolutional neural networks to achieve more accurate recognition. High precision, scalability, and robustness, on the other hand, remain significant issues in sign language recognition research. The purpose of this research is to examine hybrid neural network to improve the accuracy and robustness of sign language recognition. This research proposes a sign language recognition system using an ensemble of convolutional neural network (CNN) models followed by a long short-term memory (LSTM) model. The proposed system is designed to recognize hand gestures and interpret sign language, with a focus on American Sign Language (ASL) and sign digit dataset. Research evaluated the performance of the proposed system using the same ASL dataset and achieved an accuracy of 99.3%. We compared the performance of the proposed model with standalone CNN models and found that the proposed hybrid model outperformed standalone models.

**Keywords:** Artificial Intelligence, Sign language recognition, Neural network, Convolutional Neural Network, Long short-term memory.

## **The Performance of GPT-3.5 in Summarizing Scientific and News Articles**

Sabkat Arshad<sup>1</sup>, Muhammad Yaqoob<sup>1\*</sup> and Tahir Mehmood<sup>2</sup>

<sup>1</sup>Department of Computer Science  
University of Hertfordshire, Hatfield, UK

<sup>2</sup>School of Information Technology  
UNITAR International University, Selangor, Malaysia

\*Corresponding authors: m.yaqoob3@hertzs.ac.uk,  
tahir.mehmood@unitar.my

**Abstract.** In the age of information, we are overwhelmed with large amounts of data. The quest to know more in less time has increased the need for efficient text summarization models that convert information into precise summaries such that essential details are not overlooked. Recently, GPT-3.5 has demonstrated impressive performance in text completion, generation, and question answering. However, its effectiveness in generating concise and coherent summaries for scientific articles and news reports remains under-explored. This work evaluates the performance of GPT-3.5 in summarizing scientific research articles and news data. Scientific articles were collected from arXiv STEM dataset, whereas news articles were sampled from the CNN/DailyMail dataset. Using the GPT-3.5 OpenAI API, the pre-trained model is prompted to generate summaries of the scientific and news articles. In the next step, the ROUGE score is computed for the generated summaries against the reference summaries to analyse the performance of the model. Our results show that GPT-3.5 performs slightly better in summarizing scientific articles as compared to news articles with an average ROUGE score of 0.35 and 0.31, respectively. Moreover, in agreement with the literature, we show that the ROUGE is not the best measure for evaluating text similarity as it heavily relies on similar vocabulary rather than semantics.

**Keywords:** Text Summarization · Scientific Papers · ChatGPT · Natural Language Processing · Large Language Model.

## **Wound Stage Recognition using YOLOv5**

Clair Abela and Frankie Inguanez\*

Institute of Information & Communication Technology, Malta College of Arts Science & Technology, Corradino Hill, Paola PLA9032, Malta.

\*frankie.inguanez@mcast.edu.mt,  
clair.abela.f32840@mcast.edu.mt

**Abstract.** Pressure injuries or pressure sores are lesions that develop over the bony prominences of the body. It is noted from professional nurses and from first-hand experience, that such wounds are often wrongly classified, making the healing process difficult and painful for the patient, when it does not need to be the case. This study aims to research the use of modern computer vision and artificial intelligence techniques to aid in the classification of pressure ulcers to a degree superior to that achieved by humans. The YOLOv5 algorithm was trained in various combinations of three data sets. Single-phase and multi-phase architecture models were developed and trained for the different data set configurations to identify the ideal scenario of classification for the different stages of pressure ulcers. A single-phase architecture model trained using the public data sets and enhanced with the novel data set improves on the current state of the art. While the single-phase architecture was deemed the better option for the wound stage classification throughout this study, both architectures had their own strong points. The multi-phase architecture was found to be better suited for smaller data sets, while the single-phase architecture was better suited for larger ones. It was also found that throughout both architectures, the combination of 3 data sets is varied enough, however, less is not. This study is further strengthened through the input of different viewpoints that complement the lead researchers work.

**Keywords:** Pressure Injury, Staging, Machine learning, YOLO.

## **Harvest Palm Tree Based on Detection through 2D LiDAR Sensor Using Power Equation**

Luqman Hakim Bin Yusof<sup>1</sup>, Abdulaziz Yahya Yahya Al-Nahari<sup>2</sup>, Danny Ngo Lung Yao<sup>3\*</sup> and Normaiza Mohamad<sup>4</sup>

<sup>1,2,3,4</sup> School of Information Technology, UNITAR International University, 47301 Petaling Jaya, Selangor, Malaysia

danny.ngo@unitar.my

**Abstract.** The palm oil industry in Malaysia is encountering challenges that need innovative and technology-driven solutions. This study focuses on palm tree detection where 2D LiDAR sensors are utilized to collect data like distance and reflection strength. Through analysis, the gathered data is compared with an array of trend lines to ascertain the optimal data relationship. Among the equations considered, including linear, logarithmic, polynomial, and power equations, the power equation emerges to fit more for the detection algorithm. The chosen equation is integrated into the ESP32 firmware. The evaluation of the algorithm's efficacy transpires through its accuracy in identifying palm trees. The algorithm exhibits an accuracy rate of 98%, attesting to its proficiency in discerning palm trees within plantations.

**Keywords:** Power Equation, Palm Tree Detection, 2D LiDAR Sensor, ESP32 firmware.

## **Enhancing Security Surveillance through Business Intelligence with NVIDIA DeepStream**

Vishal Pednekar<sup>1</sup>, Nidhi Shettigar<sup>1</sup>, Sayli Tawhare<sup>1</sup>

<sup>1</sup> Department of Computer Engineering, K.C. College of Engineering and Management  
Studies and Research, Thane, India

shettigarnidhi813@gmail.com

**Abstract.** Security systems, whether in public or private establishments, frequently encounter challenges in delivering tailored, insightful analytics and real-time security alerts. In the contemporary business landscape, grasping customer behaviour and enriching their experience is necessary. Business analytics emerges as the linchpin for well-informed decision-making across financial aspects, daily operations, and beyond. Concurrently, upholding premises security remains of utmost importance. To address these requirements, we introduce URSA - an economical application seamlessly integrated into existing surveillance systems. URSA harmoniously merges business analytics with heightened security measures, leveraging NVIDIA DeepStream technology to analyze data acquired from surveillance systems. The customized DeepStream pipeline within URSA amalgamates diverse deep learning models, including PeopleNet, FaceNet, and ST-GCN, to deliver targeted outcomes. Our evaluation of URSA's performance, conducted with video footage from bustling areas like college campuses and railway stations, underscores its exceptional accuracy in handling challenging real-world scenarios.

**Keywords:** Analytics, artificial intelligence, deep learning, machine learning, real-time, surveillance.

## **Fuzzified Hybrid Metaheuristics for QoS-aware Service Composition**

Hadi Naghavi-pour, Farhad Nadi, and Ali Aitizaz

School of Information Technology,  
UNITAR International University, Petaling Jaya, Malaysia

email@hadi-naghavi-pour.com

**Abstract.** Service composition has been a centric issue in various distributed computational paradigms arising from the necessity of the composition of services to deliver more complicated computing tasks. In many cases, service composition was the process of selecting the optimal set of services from a repository of available services, which led to a multi-objective combinatorial optimization problem falling in the category of NP-hard. As a result, seeking optimal solutions in a minimum time budget has been a continuous quest in the research agenda. Meta-heuristics and hybrid metaheuristics have been the primary avenues to address this problem. Nonetheless, metaheuristics suffer from randomness, slow or premature convergence and stochastic behavior. A plethora of hybridization methods leveraging operator modification is evident in existing literature. This paper proposed a fuzzy hybrid method of mathematical techniques using fuzzy linguistics as a search operator to accelerate convergence and minimize the stochastic behaviors of metaheuristics to achieve a stable search mechanism.

**Keywords:** Fuzzy Metaheuristics · Hybrid Metaheuristics · QoS Aware Service Composition.

## **Track 2: Machine & Deep Learning**

<b>NO</b>	<b>ID</b>	<b>TITLE</b>	<b>AUTHORS</b>
1.	1570954006	Fraudulent e-Commerce Website Detection Using Convolutional Neural Network based on Image Features	Nurfazrina Mohd Zamry, Anazida Zainal, Eric Khoo, Mohamad Nizam Kassim and Zanariah Zainudin
2.	1570959213	Unlocking Security: Advanced Ransomware Prediction and Classification with Artificial Neural Networks	Saaman Nadeem, Tahir Mehmood and Muhammad Yaqoob
3.	1570948608	Leveraging Gamification for Engaged Learning in Online Teaching and Learning Experiences	Norshahriah Abdul Wahab, A'tifah Hanim Rosli, Syarifah Bahiyah Rahayu Syed Mansoor, Norazliana Akmal Jamaludin and Siti Hajar Adam
4.	1570952240	Sentiment Analysis Using Large Language Models: A Case Study of GPT-3.5	Farhad Nadi, Hadi Naghavipour, Tahir Mehmood, Alliesya Binti Azman, Jeetha A/P Nagantheran, Kezia Sim Kui Ting, Nor Muhammad Ilman Bin Nor Adnan, Roshene A/P Sivarajan, Suita A/P Veerah and Romi Fadillah Rahmat
5.	1570953458	Telecom Customer Experience Analysis using Sentiment Analysis and Natural Language Processing – Comparative Study	Ahmed Mohamed Abdou Ahmed, Raghad Al_Shabandar, Abdulaziz Al-Nahari and A.H. Mohammed
6.	1570953467	Efficient Time Series of Smoothing and Auto-regressive Forecasting Models for Predicting Police Officer Fatalities in the USA	Danush Nagappan, Manoj Jayabalan, Ahmad Alanezi, Farhad Nadi and Thomas Coombs
7.	1570953476	Multimodal Emotion Recognition using Attention-based Model with language, Audio, and Video Modalities	Disha Sharma and Manoj Jayabalan, Nailiya Sultanova, Jamila Mustafina and Danny Ngo Lung Yao

8.	1570953964	Comparative Analysis of Emotion Recognition using Large-Language Models and Conventional Machine Learning	Mangu Soujanya Rao, Thomas Coombs, Normaiza Binti Mohamad, Vinay Kumar and Manoj Jayabalan
9.	1570955167	The Impact of Clustering-Based Sequential Multivariate Outliers Detection in Handling Missing Values	Mety Agustini, Kartika Fithriasari and Dedy Dwi Prastyo
10.	1570956190	Sarcasm Detection in Newspaper Headlines	Vishnu Sai Reddy Chilpuri, Saaman Nadeem and Tahir Mehmood
11.	1570956272	Transformer-Based Named Entity Recognition Model - Tamil Language	Karthi Dhayalan, Nailya Sultanova, Jamila Mustafina and Paridah Daud
12.	1570956523	A Comparative Study of Methods for Topic Modelling in News Articles	Swapna D Rajan, Thomas Coombs, Manoj Jayabalan and Noor Azma Ismail
13.	1570956622	Application of Deep Learning Algorithms to Terahertz Images for Detection of Concealed Objects	Soumen Sardar, Sulaf Assi, Iznora Aini Zolkifly, Manoj Jayabalan, Manea Alsaleem, Ammar H. Mohammed and Dhiya Al-Jumeily OBE
14.	1570958001	Multivariate Comparative Analysis of Statistical and Deep Learning Models for Prediction Hardware Failure	Saurabh Gupta, Raghad Alshabandar, Chong Kim Loy and Ammar H. Mohammed

## **Fraudulent e-Commerce Website Detection Using Convolutional Neural Network based on Image Features**

Nurfazrina Mohd Zamry<sup>1</sup>, Anazida Zainal<sup>1</sup>, Eric Khoo<sup>1</sup>, Mohamad Nizam Kassim<sup>2</sup>  
and Zanariah Zainudin<sup>3</sup>

<sup>1</sup>Cyber Threat Intelligence Lab, Information Assurance and Security  
Research Group, Faculty of Computing 81310 UTM Johor Bahru, Johor, Malaysia

<sup>2</sup>Bahagian Penyelidikan Strategik, National Anti-Financial Crime Center, Putrajaya,  
Malaysia

<sup>3</sup>Faculty of Information and Communication Technology  
(FICT), Universiti Tunku Abdul Rahman, Jalan Universiti, Bandar Barat, 31900 Kampar,  
Perak, Malaysia

nurfazrina.mz@utm.my, anazida@utm.my,  
erichooi1995@gmail.com,  
mnizam.kassim@gmail.com and zanariah@utar.edu.my

**Abstract.** Due to the growth of e-Commerce's website, attackers have motivated to intrude these e-commerce websites. These illegitimate websites, also known as fraudulent websites, exploited the internet to trick people into committing fraud or engaging in malicious cyberattacks, which cost billions of dollars of loss. Numerous research and methodologies have been developed to recognize fraudulent Internet websites, yet none of them has been able to offer a feasible solution to bring an end to these illegal acts. Moreover, as the number of fraudulent webpages increases, so does the sophistication of the threat. Due to this issue, this paper presents a method for detecting fraudulent websites that is based on image of the website that may contain some patterns that show it is fraudulent. In this paper, Convolutional Neural Network (CNN) classification is used to classify the fraud and legal website image utilized. A total of 530 website images were collected with 272 fraud website images and 258 fraud website images. The websites are partitioned into 80% (370 images) samples as training set, 10% (80 images) samples as testing set and the rest 10% (80 images) samples as validation set. Three CNN optimizers namely including Adam, RMSprop (Root Mean Square) and SGD (Stochastic Gradient Descent) were tested for training and validation accuracy and loss in different epoch. The results show that proposed fraudulent website detection model accuracy of 78.25 % using Adam optimizer for 120 epochs.

**Keywords:** Fraudulent website, Convolutional Neural Network, HTML image

## **Unlocking Security: Advanced Ransomware Prediction and Classification with Artificial Neural Networks**

Saaman Nadeem<sup>1</sup>, Tahir Mehmood<sup>2</sup>, and Muhammad Yaqoob<sup>3</sup>

<sup>1</sup>Department of Computer Science  
University of Management and Technology, Lahore, Pakistan  
saamannadeem350@gmail.com

<sup>2</sup>School of Information Technology  
UNITAR International University, Selangor, Malaysia  
tahir.mehmood@unitar.my

<sup>3</sup>Department of Computer Science  
University of Hertfordshire, Hatfield, UK  
m.yaqoob3@herts.ac.uk

**Abstract.** With the increase in the development of technology, the threat of “Ransomware” has also increased especially towards organizations. Ransomware is a malicious software that encrypts all the user’s data or system and demands a ransom payment for decryption. Despite various machine learning approaches proposed for ransomware detection, they often fail to identify those threats accurately in time, thus leading to data loss and victimization. This research introduces a novel framework, primarily based on static analysis of ransomware and predicting the presence of ransomware on users’ systems by monitoring a defined set of ransomware activities. In this study, we used the Resilient Information Systems Security (RISS) ransomware dataset, encompassing 582 ransomware samples from 11 distinct families and 982 instances of goodware. We proposed a generic neural network framework for the identification of ransomware and compared the performance of Artificial Neural Networks (ANN) and Deep Neural Networks (DNN) in terms of accurately classifying ransomware and goodware. The suggested framework secured an accuracy of 98.56% with ANNs, and achieved a slightly better performance (99.06%) when ANN was replaced with DNN. Our results showed that a basic ANN can achieve performance comparable to that of a DNN for ransomware detection. In future work, we plan to evaluate the performance of the proposed framework in a real-time setting.

**Keywords:** Ransomware, Artificial Neural Network, Deep Neural Network, Malware, Machine Learning

## **Leveraging Gamification for Engaged Learning in Online Teaching and Learning Experiences**

Norshahriah Abdul Wahab<sup>1</sup>, A'tifah Hanim Rosli<sup>1</sup>, Syarifah Bahiyah Rahayu Syed Mansoor<sup>1</sup>, Norazliana Akmal Jamaludin<sup>1</sup> and Siti Hajar Adam<sup>2</sup>

<sup>1</sup>Computer Science Department, Faculty of Defence Science & Technology

<sup>2</sup>Faculty of Medicine and Defence Health

shahriah@upnm.edu.my

**Abstract.** In the rapidly evolving landscape of education, the integration of technology and evolving student dynamics have ushered in a new era of Teaching and Learning (TnL). This article explores the transformative power of gamification in online education, shedding light on its capacity to ignite motivation, enhance engagement, and facilitate meaningful learning experiences. Traditional education paradigms have expanded, requiring educators to guide students in exploring knowledge and essential skills in digital environments. Gamification emerges as a potent solution in this context, seamlessly blending educational content with gaming elements, creating stimulating and rewarding learning journeys. Gamification, rooted in the fusion of game design principles and non-game contexts, has risen to prominence. It has been recognized for its potential to revolutionize engagement and behavior transformation, with over 70% of Fortune Global 2000 organizations adopting it by 2014. This article explores the burgeoning realm of gamification, its rapid evolution, and its role as an innovation catalyst in the teaching and learning landscape. A study conducted with 30 Computer Science students at the National Defence University of Malaysia reveals the impact of gamification on student engagement. Results indicate that the gamified approach effectively encourages active participation in online discussions and engages students' passion for learning. These findings underscore gamification's potential to create dynamic and interactive learning environments, aligning with its growing popularity and transformative potential. As the digital generation shapes education, gamification offers an innovative approach to fostering effective learning experiences that cater to evolving preferences. Gamification represents a revolutionary method bridging education and gaming, promoting motivation, active involvement, and improved learning outcomes. This study prepares the way for a new era of enhanced and interactive learning experiences, equipping educators to harness the potential of gamification and adapt education to the evolving needs of contemporary learners.

**Keywords:** Gamification, Teaching and Learning, Student Engagement, Learning Experiences, Education Technology

## **Sentiment Analysis Using Large Language Models: A Case Study of GPT-3.5**

Farhad Nadi<sup>1</sup>, Hadi Naghavipour<sup>1</sup>, Tahir Mehmood<sup>1</sup>, Alliesya Binti Azman<sup>1</sup>, Jeetha A/P Nagantheran<sup>1</sup>, Kezia Sim Kui Ting<sup>1</sup>, Nor Muhammad Ilman Bin Nor Adnan<sup>1</sup>, Roshene A/P Sivarajan<sup>1</sup>, Suita A/P Veerah<sup>1</sup>, Romi Fadillah Rahmat<sup>2</sup>

<sup>1</sup>School of Information Technology, UNITAR International University, Kelana Jaya 47301, Malaysia

<sup>2</sup>Department of Information Technology, Universitas Sumatera Utara, Indonesia

farhad.nadi@unitar.my;farhad.nadi@gmail.com

**Abstract.** Sentiment analysis, which utilizes machine learning, natural language processing, and computational linguistics, has been developed to comprehend the emotions and viewpoints of individuals on social media platforms. This area of research has undergone extensive investigation, leading to the integration of diverse algorithms and techniques specifically tailored to this objective. A re-cent breakthrough in the field of general AI is exemplified by Large Language Models, which have surpassed numerous research domains as quintessential achievements. This study examines the performance of GPT-3.5 in sentiment analysis. We evaluate its performance on established benchmark datasets as well as a collection of reviews extracted from Google Maps. The findings indicate that the Large Language Models can outperform traditional methods in the literature, and they do not require pre-processing, unlike traditional methods.

**Keywords:** Sentiment Analysis, GPT3.5, Large Language Models, Artificial Intelligence.

## **Telecom Customer Experience Analysis using Sentiment Analysis and Natural Language Processing – Comparative Study**

Ahmed Mohamed Abdou Ahmed<sup>1</sup>, Raghad Al\_Shabandar<sup>2</sup>, Abdulaziz Al-Nahari<sup>3</sup>,  
and A.H. Mohammed<sup>2</sup>

<sup>1</sup>Independent Researcher, India

<sup>2</sup>Iraqi Prime Minister's Office, Iraq

<sup>3</sup>School of Information Technology, UNITAR International University, Petaling Jaya,  
Selangor, Malaysia

abdulaziz.yahya@unitar.my

**Abstract.** In today's competitive telecom landscape, service providers are increasingly seeking real-time customer experience analysis, prompt responses to customer feedback, and the ability to effectively promote new services. To achieve these goals, telecom operators are embracing digitalization initiatives that encompass the entire customer journey, broadly divided into three phases: engaging, using, and evaluating. Recent advancements in natural language processing (NLP) and sentiment analysis (SA) techniques have empowered telecom service providers to rapidly analyze and categorize millions of customer tweets, gaining valuable insights into service perceptions and user satisfaction. With a significant presence of telecom service providers in Arab countries, where customers frequently share their service experiences through Arabic tweets, the need for specialized NLP and SA techniques that can effectively handle Arabic language data becomes paramount. This study focuses on Arabic language processing and sentiment analysis to support one of the Middle East's largest telecom service providers in analyzing and enhancing customer experience. The study successfully applied BERTopic, a topic modeling technique, to Arabic telecom-related text, generating six distinct clusters for 50% of the analyzed tweets. Support vector machine (SVM) outperformed XGBoost as the machine learning classifier when combined with the BERT-base model, achieving an F1-score of 0.71 compared to XGBoost's 0.65. The fine-tuned MARBERT model demonstrated superior performance in text classification compared to machine learning algorithms, achieving an F1-score of 0.8571, a 3% improvement over the best-performing machine learning classifier.

**Keywords:** Latent Dirichlet Allocation (LDA), MARBERT, Arabic Natural Language; Sentiment Analysis.

## **Efficient Time Series of Smoothing and Auto-regressive Forecasting Models for Predicting Police Officer Fatalities in the USA**

Danush Nagappan<sup>1</sup>, Manoj Jayabalan<sup>1</sup>, Ahmad Alanezi<sup>2</sup>, Farhad Nadi<sup>3</sup>, Thomas Coombs<sup>4\*</sup>

<sup>1</sup>Liverpool John Moores University, Liverpool, United Kingdom

<sup>2</sup>Saad Alabdullah Police Academy, Kuwait

<sup>3</sup>School of Information Technology, UNITAR International University, Petaling Jaya, Malaysia

<sup>4</sup>British American Tobacco, Southampton, United Kingdom

coombstom25@gmail.com

**Abstract.** With police deaths rising, predicting the number of police deaths is now of significant importance, and is necessary to take precautions to prevent deaths from affecting the police force, public, and government from an associated reduction in police numbers. The aim of this research is to compare different time series forecasting models and find the most efficient model in predicting police deaths occurring in the USA. The dataset used in this study consisted of details of police officers who had died in service on duty in the USA. A total of 26,269 records between 3rd January 1791 and 3rd June 2022. The dataset was obtained through the Kaggle website data repository. A total of four smoothing models and four auto-regressive models were used and compared in this research. The smoothing models had better RMSE and MAPE scores, with HWES being the best-performing model. In summary, the HWES model performed the best on the USA police deaths dataset by producing less error compared to the other smoothing and auto-regressive models used in this research.

**Keywords:** Time series, Forecast, HWES, Police Officer Fatalities

## **Multimodal Emotion Recognition using Attention-based Model with language, Audio, and Video Modalities**

Disha Sharma and Manoj Jayabalan<sup>1</sup>, Nailiya Sultanova and Jamila Mustafina<sup>2</sup>,  
Danny Ngo Lung Yao<sup>3</sup>

<sup>1</sup>Liverpool John Moores University (LJMU), United Kingdom  
d.sharma@ljmu.stud.ac.uk, m.jayabalan@ljmu.ac.uk

<sup>2</sup>Kazan Federal University, Russia  
NaRNurutdinova@kpfu.ru, DNMustafina@kpfu.ru

<sup>3</sup>UNITAR International University, Malaysia  
danny.ngo@unitar.my

**Abstract.** Multimodal emotion identification is becoming increasingly important in human-computer interaction due to the amount of emotional information in human communication. Multimodal emotion recognition is the technique of simultaneously considering several modalities to boost accuracy and robustness. As emotion identification studies become more vital to human-computer interactions, automatic emotion detection systems become increasingly necessary. However, a lack of data presents a problem for multimodal emotion identification. To address this issue, we suggest employing transfer learning, which uses pretrained models such as RoBERTa and attention-based mechanisms such as Self Attention to extract relevant features from multiple modalities and Multi-head Attention to fuse data across modalities. The aim of this paper is to provide a strategy for reliably forecasting emotions in audio, visual and text by merging and complementing aspects traditionally handled by humans with those typically handled by deep learning. During the study three popular multimodal emotion recognition datasets, IEMOCAP, CMU-MOSI, and CMU-MOSEI, are analyzed and ranked based on their quality. This study will help in constructing the network with the right amount of focus placed on each feature modality by creating an architecture that efficiently combines textual characteristics retrieved from RoBERTa with other modality-based features. A model better than BERT is introduced as part of this work that helps to improve the performance.

**Keywords:** Emotion identification, Multimodal Emotion Recognition, Self-Attention, Multi-head Attention.

## **Comparative Analysis of Emotion Recognition using Large-Language Models and Conventional Machine Learning**

Mangu Soujanya Rao<sup>1</sup>, Thomas Coombs<sup>2</sup>, Normaiza Binti Mohamad<sup>3</sup>, Vinay Kumar<sup>1</sup>,  
Manoj Jayabalan<sup>1</sup>

<sup>1</sup> Liverpool John Moores University, Liverpool, United Kingdom

<sup>2</sup> British American Tobacco, Southampton, United Kingdom

<sup>3</sup> School of Information Technology, UNITAR International University, Petaling Jaya,  
Malaysia

coombstom25@gmail.com

**Abstract.** Emotions are significant aspects of human existence and influence interaction between individuals and groups, influencing how we think and behave. In this research we aim to use conventional and neural network models to identify emotions from textual data and compare which performed best. The Go Emotions dataset contained 27 different emotions across 58,000 samples. The approach involves modelling the Conventional machine learning models and the Neural Network based models and comparing the results over test dataset and choosing the best model. Upon comparing the classification reports for the conventional and neural network-based models on the Ekman taxonomy, conventional machine learning algorithms were outperformed by neural network-based models which gained almost 10% more than conventional models. Conventional models averaged the values around 50% of macro average F1-Score except for the KNN classifier which performed poorly getting the macro average F1-Score of 21%. BERT classifier with Ekman taxonomy including neutral emotion had a macro average precision of 55% and a sensitivity of 68%. This classifier also outperformed the macro-average F1-score by 106 61%. While the ROBERTA classifier had a macro-average precision of 65%, the recall, or sensitivity, was found to be 53%. This study clearly states that Neural Network based models outperformed conventional models. Our study proposed BERT Model which achieved a macro average F1-Score of 0.50 across Go Emotion taxonomy.

**Keywords:** Neural network, Emotion, Machine learning.

## **The Impact of Clustering-Based Sequential Multivariate Outliers Detection in Handling Missing Values**

Mety Agustini<sup>1,2</sup> Kartika Fithriasari<sup>1</sup> Dedy Dwi Prastyo<sup>1</sup>

<sup>1</sup>Department of Statistics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia  
kartika\_f@statistika.its.ac.id, dedy-dp@statistika.its.ac.id

<sup>2</sup>Badan Pusat Statistik Provinsi Kepulauan Bangka Belitung, Pangkalpinang, Indonesia  
mety.assahid@bps.go.id

**Abstract.** The presence of missing values is a common issue that frequently leads to incomplete data in a wide range of research. They diminish the accessibility of the dataset that can be utilized and degrade the statistical power of the analysis. A significant focus in numerous studies has been directed towards the methods of missing value imputation. In cases where the dataset includes outliers, the imputation of missing values might be incorrect or significantly deviate from the actual values. One of the challenges that impacts the quality of data is the handling of missing values and outliers simultaneously. Several studies removed outliers before imputing missing values or deleted observations with missing values before detecting outliers. The removal approach leads to a lack of information included within the data. Other researchers integrate clustering methods into the process of missing value imputation to mitigate the impact of outliers and data variations, thereby enhancing the accuracy of the imputation model. This paper proposes a new clustering-based sequential multivariate outlier detection (SMOD) method to effectively handle incomplete data within outliers. The method is applied to an official economic statistics dataset that involves outliers and performs a missing value rate scenario of about 50 percent. In comparison to a well-known and widely used clustering technique, i.e., model-based clustering (MBC), the proposed method works well in missing value imputation.

**Keywords:** Clustering, Missing Value, Outlier Detection, Multiple Imputation.

## **Sarcasm Detection in Newspaper Headlines**

Vishnu Sai Reddy Chilpuri<sup>1</sup>, Saaman Nadeem<sup>2</sup>, Tahir Mehmood<sup>3</sup>,  
and Muhammad Yaqoob<sup>1</sup>

<sup>1</sup>Department of Computer Science  
University of Hertfordshire, Hatfield, UK  
{vc20aaw, m.yaqoob3}@herts.ac.uk

<sup>2</sup>Department of Computer Science  
University of Management and Technology, Lahore, Pakistan  
saamannadeem350@gmail.com

<sup>3</sup>School of Information Technology  
UNITAR International University, Selangor, Malaysia  
tahir.mehmood@unitar.my

**Abstract.** Language is an essential medium for human communication. It allows us to convey information, express our ideas, and give instructions to others. The rise of sarcasm can be attributed to the increasing number of negative comments and expressions posted on social networks such as Twitter, Facebook, and newspapers. Due to the use of positive vocabulary in sarcastic comments, it is hard to detect sarcasm in news reports. Sarcasm is intentionally used in news reports to grab the readers' attention. Unfortunately, many people find it hard to identify the ironic tone of the headlines and may pass incorrect information. This work focuses on detecting sarcasm in newspaper headlines and investigates the performance of four machine learning algorithms (Logistic Regression, Naive Bayes, Decision Tree, and Random Forest) and one deep learning model BiLSTM (Bi-directional Long Short-Term Memory) for sarcasm detection in news headlines. We demonstrate that regardless of the machine learning model, the application of vectorization technique i.e., BoW (Bag of Words) and TF-IDF (Term Frequency - Inverse Document Frequency) has minimal influence on the ability to detect sarcasm in news headlines. We also show that the performance of the three machine learning algorithms (Logistic Regression, Random Forest, and Decision Tree) remains stable across two tokenization techniques (Unigram or Bigram) except Naive Bayes which secured a high precision with Unigram analysis. We further found that BiLSTM is the most preferred model for sarcasm detection in news headlines.

**Keywords:** Sarcasm Detection · Newspaper Headlines · Natural Language Processing · Machine Learning · Deep Learning.

## **Transformer-Based Named Entity Recognition Model - Tamil Language**

Karthi Dhayalan<sup>1</sup>, Nailya Sultanova and Jamila Mustafina<sup>2</sup>, Paridah Daud<sup>3</sup>

<sup>1</sup> Liverpool John Moores University, United Kingdom  
karthi.dhayalan@gmail.com

<sup>2</sup> Kazan Federal University, Russia  
NaRNurutdinova@kpfu.ru, DNMustafina@kpfu.ru

<sup>3</sup> UNITAR International University, Malaysia  
paridah69@unitar.my

**Abstract.** This work presents different Transformer based approaches to the Named Entity Recognition problem in the Tamil language. The WikiANN-ta dataset is used for training, validating and testing the model. Person, Organization and Location tags are annotated in the dataset. Context features of Tamil languages are extracted by the BERT pre-trained models. In this task, 5 significant BERT variations are explored, 1) mBERT (Multilingual BERT), 2) XLM-RoBERTa, 3) Indic BERT, 4) MuRIL-base & 5) +MuRiL large. Context features extracted by mBERT/XLM-R/Indic BERT/ MuRIL-base/ MuRiL large are sent to the feedforward layer with SoftMax to identify named entities. Two models out of these 5 BERT based models achieve state-of-the-art results in the test dataset of WikiANN-ta 1) NER model with MuRIL-base, 2) NER model with MuRIL-large. MuRIL-base is proposed for better runtime and MuRiL-large is proposed for better accuracy.

**Keywords:** Transformer, Tamil language, Named Entity Recognition, Machine learning.

## **A Comparative Study of Methods for Topic Modelling in News Articles**

Swapna D Rajan<sup>1</sup>, Thomas Coombs<sup>2</sup>, Manoj Jayabalan<sup>3</sup> and Noor Azma Ismail<sup>4\*</sup>

<sup>1</sup>Independent researcher, India

<sup>2</sup>British American Tobacco, Southampton, United Kingdom

<sup>3</sup>Liverpool John Moores University, Liverpool, United Kingdom

<sup>4</sup>UNITAR International University, Selangor, Malaysia

azma1706@unitar.my, coombstom25@gmail.com

**Abstract.** The past few decades have seen an increase in textual data and influence from the news media. With the rise in available data, especially in regard to textual data from news media, it is imperative to quickly categorise news topics. In this research, the primary aim is to suggest a method for automatically identifying news topics in articles. The dataset used in this research was the News Category published on Kaggle and comprised of 210,294 headlines and abstracts from HuffPost between 2012 to 2022. The dataset consisted of a total of 42 categories and 6 columns. Traditional modelling techniques did not perform well in comparison to Top2Vec, NMF or BERTopic. This research confirms the efficacy of Top2Vec and BERTopic, followed by NMF, LDA, and LSA for analysing news category data from a human-interpretation perspective. Though BERTopic was able to deduce 1145 topics from the data, it could not chunk unwanted words like “to”, “say”, “for” which do not add any value to the topic semantics. In summary, TF-IDF proved to be the best feature extraction technique and Top2Vec the best topic modelling technique.

**Keywords:** Topic modelling, machine learning, Top2Vec, BERTopic

## **Application of Deep Learning Algorithms to Terahertz Images for Detection of Concealed Objects**

Soumen Sardar<sup>1</sup>, Sulaf Assi<sup>1</sup>, Iznora Aini Zolkify<sup>2</sup>, Manoj Jayabalan<sup>1</sup>, Manea  
Alsalem<sup>3</sup>, Ammar H. Mohammed<sup>4</sup>, and Dhiya Al-Jumeily OBE<sup>1,\*</sup>

<sup>1</sup> Liverpool John Moores University, UK

<sup>2</sup> School of Information Technology, UNITAR International University, Petaling Jaya,  
Selangor, Malaysia

<sup>3</sup>Khabash General Hospital, Saudi Arabia

<sup>4</sup>Iraqi Prime Minister Office, Iraq

d.aljumeily@ieee.org\_

**Abstract.** Safety of the public at large venues is of utmost importance and therefore it is important to detect threats encountered due to concealed objects especially on human bodies. Terahertz imaging have gained popularity over the last few years due to its ability to detect concealed objects inside fabric without harm to humans or invasion to their privacy. However, terahertz images suffer from poor resolution and low signal to noise ratio. Deep learning has shown high performance in classifying images lately especially feed forward neural networks. Therefore, this study utilised deep learning for detection of concealed objects on human bodies using the public Active Terahertz Imaging (ATZ) dataset that contained images of 11 tiny objects. An end-to-end framework was applied and involved image enhancement using wavelet filtering, locating object using skip attention generative adversarial networks (SAGAN) and Forward-Forward Mixed Convolution with Gaussian Affinity Network (FMCGNet). The results showed that wavelet filtering did not affect the performance of the deep learning models. The SAGAN showed performance accuracy of 68.4% and that was related to the insufficient training of the model that required high computational power. However, the accuracy of classifying images with anomalies was high when FMCGNet was applied and was featured in the true positive rate of 91.9% though it was applied to imbalanced dataset. In summary, the proposed end-to-end framework successfully identified concealed objects on human bodies in an efficient way. Future work involves adopting this approach to different types of imaging datasets to further understand its strengths and limitations.

**Keywords:** Safety, Terahertz, imaging, deep learning, SAGAN, FMCGNet.

## **Multivariate Comparative Analysis of Statistical and Deep Learning Models for Prediction Hardware Failure**

Saurabh Gupta<sup>1</sup>, Raghad Alshabandar<sup>2</sup>, Chong Kim Loy<sup>3</sup>, and Ammar H. Mohammed<sup>2</sup>

<sup>1</sup>Independent Researcher, India

<sup>2</sup>Iraqi Prime Minister's Office, Iraq

<sup>3</sup>UNITAR International University, Petaling Jaya, Selangor, Malaysia

Chongkl@unitar.my, Saurabhguptacse@gmail.com

**Abstract.** The 21st century is witnessing a transformative shift towards digitalization, where various services are undergoing a paradigm change from traditional to digital platforms. This digital revolution, evident in the transition from commerce to e-commerce and agriculture to agro-tech, necessitates the development of robust and sustainable systems to support these applications. While cloud providers like Amazon, Azure, and Google Cloud offer Infrastructure as a Service (IaaS) solutions, many organizations still prefer on-site hardware maintenance. The existing studies on system failure prediction have primarily focused on either machine learning techniques like Random Forest and Naive Bayes or deep learning neural networks like RNN, LSTM, and CNN. However, there has been no comprehensive comparison of prediction accuracy between statistical models and deep learning models. This research aims to identify the algorithm that yields the highest prediction accuracy while considering hardware resource utilization, including CPU, RAM, ROM, and network usage. The study utilizes multivariate time series analysis to predict system failures based on the "BitsBrain" dataset. The models evaluated include ARIMA, auto-regression, SARIMAX, exponential smoothing, LSTM, and bi-LSTM. The results demonstrate that the bi-LSTM model outperforms the other models, achieving the lowest mean absolute error (MAE) and the highest R2 score. Conversely, the exponential smoothing model exhibits the poorest performance. Additionally, a comparison of actual versus predicted plots reveals that the bi-LSTM model generates more accurate predictions. These findings suggest that the bi-LSTM model can serve as a valuable tool for forecasting system failures and enhancing system reliability. Its superior prediction accuracy, coupled with relatively efficient resource utilization, makes it a compelling choice for organizations seeking to optimize their system performance and minimize downtime.

**Keywords:** Auto Regressive Integrated Moving Average (ARIMA), Long Short-Term Memory (LSTM).

### **Track 3: Statistical Learning**

<b>NO</b>	<b>ID</b>	<b>TITLE</b>	<b>AUTHORS</b>
1.	1570943390	A Case Study via Bayesian Network: Investigating Factors Influencing Student Academic Performance in Online Teaching and Learning During Covid-19 Pandemic	Zheng Ning Looi, Poh Choo Song, Huai Tein Lim and Sing Yan Looi
2.	1570949397	Harnessing the XG-Boost Ensemble for Intelligent Prediction and Identification of Factors with a High Impact on Air Quality: A Case Study of Urban Areas in Jakarta Province, Indonesia	Wahyu Wibowo, Harun Al Azies, Susi A. Wilujeng and Shuzlina Abdul-Rahman
3.	1570949756	Modeling Earthquake Catalog in Sumatra by Space-Time Epidemic-Type Aftershock Sequences Model: Combining Davidon Fletcher Powell and Stochastic Declustering Algorithms	Christopher Andreas, Achmad Choiruddin and Dedy Dwi Prastyo
4.	1570949888	Small Area Estimation of Mean Years of Schooling Under Time Series and Cross-sectional Models	Reny Ari Noviyanti, Setiawan and Agnes Tuti Rumiati
5.	1570950075	Probabilistic Seismic Hazard Analysis for Sulawesi-Maluku Region of Indonesia using the Space-Time Epidemic Type Aftershock Sequence Model	Sonia Faradilla, Achmad Choiruddin and Bambang Widjanarko Otok
6.	1570950211	Application of Time Series Regression, Double Seasonal ARIMA, and Long Short-Term Memory for Short-Term Electricity Load Forecasting	Hafez Afghan and Hidayatul Khusna

7.	1570950306	A Bayesian Network for Classifying and Predicting Ship Collision	Iis Dewi Ratih, Ketut Buda Artana, Heri Kuswanto, Emmy Pratiwi and Muhammad Farhan Nuari
8.	1570950508	Outlier Detection in Simultaneous Equations with Panel Data	Suci Ismadyaliana, Setiawan and Jerry Dwi Trijoyo Purnomo
9.	1570952489	Fast Maximum Likelihood Estimation of Large Data Spatial Error Model: Matrix Exponential Spatial Specification Approach	Marsono, Setiawan and Heri Kuswanto
10.	1570953393	Assessing Departmental Efficiency at Sepuluh Nopember Institute of Technology: A Comparative Study Using Classical and Advanced Data Envelopment Analysis	Zakiatul Wildani, M Naufal Mubarak, Sri Pingit Wulandari, Lucia Aridinanti and Muhammad Alifian Nuriman
11.	1570953406	Multivariate Adaptive Fuzzy Clustering Means Regression Splines model using Generalized Cross-Validation (GCV) on stunting cases in Southeast Sulawesi.	Mira Meilisa, Bambang Widjanarko Otok and Jerry Dwi Trijoyo Purnomo
12.	1570955599	Statistical Inferences for Multivariate Generalized Gamma Regression Model	Hasbi Yasin, Purhadi and Achmad Choiruddin

## **A Case Study via Bayesian Network: Investigating Factors Influencing Student Academic Performance in Online Teaching and Learning During Covid-19 Pandemic**

Zheng Ning Looi<sup>1</sup>, Poh Choo Song<sup>2</sup>, Huai Tein Lim<sup>3</sup>, Sing Yan Looi<sup>4</sup>

<sup>1234</sup> Department of Physical and Mathematical Science,  
Faculty of Science, University Tunku Abdul Rahman.

<sup>2</sup> Corresponding author

<sup>1</sup>zhengninglooi@gmail.com

<sup>2</sup>songpc@utar.edu.my

<sup>3</sup>limhuait@utar.edu.my

<sup>4</sup>looisyy@utar.edu.my

**Abstract.** COVID-19 pandemic has an impact on numerous Malaysia industries, particularly the education sector. Due to the changes in class delivery during the pandemic, students' academic performance has been impacted, for instance, decrease of CPGA. Hence, we would like to determine the variables that could influence academic performance of the undergraduate students in Malaysia during COVID-19 pandemic, whereby educators may be able to help the students to adapt the changes. In this study, 27 items categorized into ten factors are investigated, and Bayesian Network (BN) is employed to discover the determinants and forecast the students' academic performances. The accuracy of prediction between BN and Support Vector Machines (SVM) model is compared. The findings indicate a significant relationship between students' academic achievement and the factors including the type of learning device, faculty, attendance, age, self-efficacy, and teaching technique. Additionally, the MAPE for BN is 5% less than that of SVR in comparison to the actual results. This study assists higher education institutions to effectively strengthen the significant variables to raise students' academic performance, especially for OTL.

**Keywords:** Bayesian Network, support vector machine, academic performance

## **Harnessing the XG-Boost Ensemble for Intelligent Prediction and Identification of Factors with a High Impact on Air Quality: A Case Study of Urban Areas in Jakarta Province, Indonesia**

Wahyu Wibowo<sup>1</sup> and Harun Al Azies<sup>2,3</sup>, Susi A. Wilujeng<sup>4</sup>, Shuzlina Abdul-Rahman<sup>5</sup>

<sup>1</sup>Department of Business Statistics, Faculty of Vocational Studies,  
Institut Teknologi Sepuluh Nopember, 60111, Surabaya, Indonesia

<sup>2</sup>Study Program in Informatics Engineering, Faculty of Computer Science,  
Universitas Dian Nuswantoro, 50131, Semarang, Indonesia

<sup>3</sup>Research Center for Materials Informatics, Faculty of Computer Science,  
Universitas Dian Nuswantoro, 50131, Semarang, Indonesia

<sup>4</sup>Department of Environmental Engineering, Faculty of Civil, Planning, and Geo  
Engineering,

Institut Teknologi Sepuluh Nopember, 60111, Surabaya, Indonesia

<sup>5</sup>College of Computing, Informatics and Mathematics, Universiti Teknologi MARA, 40450  
Shah Alam, Selangor, Malaysia

wahyu\_w@statistika.its.ac.id

**Abstract.** This article aims to develop an accurate air quality prediction model to handle Jakarta's air pollution challenges. In this study, data from air quality monitoring stations' conventional air pollution indexes were employed. In the research phase, data is explored, SMOTE is used to manage imbalances, and XG-Boost is used to develop a model with the best parameters. The evaluation stage shows the model's ability to predict air quality. With an accuracy rate of 99.516%, an F1 score of 99.528%, and a recall rate of 99.509%, the results were very astounding. These performance indicators show the model's exceptional ability to classify and predict air quality levels. Furthermore, this study investigates the significance of various variables in predicting air quality. A thorough evaluation of measures such as weight, gain, total gain, and cover indicators reveals the significance of numerous aspects. Even while SO<sub>2</sub> helps predict air quality, the prevalence of PM<sub>2.5</sub> on several measures reveals a significant influence. This study contributes to a better understanding of the complicated dynamics of air quality prediction by employing advanced analytical approaches and accurate models. This knowledge is useful in developing targeted solutions to address air pollution issues and promote healthier urban environments.

**Keywords:** Air Quality Prediction, XGBoost Algorithm, Jakarta Air Pollution, Predictive modeling

# **Modeling Earthquake Catalog in Sumatra by Space-Time Epidemic-Type Aftershock Sequences Model: Combining Davidon Fletcher Powell and Stochastic Declustering Algorithms**

Christopher Andreas<sup>1</sup>, Achmad Choiruddin<sup>2</sup>, and Dedy Dwi Prastyo<sup>3</sup>

<sup>1,2,3</sup> Department of Statistics, Faculty of Science and Data Analytics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

choiruddin@its.ac.id

**Abstract.** Earthquake hazard mapping has an important role as a mitigation measure for earthquake disasters. In Indonesia, especially in the Sumatra, the risk of major earthquakes is high due to its geographical location. In this study, we conduct statistical analysis of earthquake catalog data (coordinates, time occurrences, and magnitudes) and develop earthquake hazard map in the Sumatra during 2005-2023 using the Space-Time Epidemic-Type Aftershock Sequences (Space-Time ETAS) model. We perform computationally feasible method for parameter estimation combining Davidon Fletcher Powell and Stochastic De-clustering algorithms. To design the study and obtain the best model, we define 4 different models exploring the use of complementary events. Of the 4 de-signed models, we found that model with short period of complementary events performs best. In addition, we also found that aftershock generation mechanism in the Sumatra is mainly dominated by the magnitudes of mainshocks (rather than locations or time occurrences of mainshocks). The resulting hazard map also shows the mainshocks dominate the aftershocks, which is in accordance with Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) records with re-spect to the earthquake-prone areas in Sumatra.

**Keywords:** Complementary Events, Disaster Risk Reduction, Earthquake Hazard Mapping, Space-Time ETAS.

## **Small Area Estimation of Mean Years of Schooling Under Time Series and Cross-sectional Models**

Reny Ari Noviyanti<sup>1,2</sup>, Setiawan<sup>1(\*)</sup>, Agnes Tuti Rumiati<sup>1</sup>

<sup>1</sup> Institut Teknologi Sepuluh Nopember (ITS), Surabaya 60111, Indonesia  
setiawan@statistika.its.ac.id

<sup>2</sup> BPS, Statistics of Sumatera Utara Province, Medan 20123, Indonesia

**Abstract.** Small area estimation develops within the framework of time series and cross-sectional models. The restricted estimation maximum likelihood method was used to obtain the empirical best linear unbiased prediction for small areas and its mean squared error estimators. The model focuses on applying statistical models that permit borrowing strength over area and time. The process uses regularly conducted survey data, where the areas of interest are observed repeatedly under a predetermined scheme. The time series and cross-sectional models was able to capture heterogeneity across area and time, so it can be used to enhance sample size effectiveness, thereby minimizing mean squared error and producing a more accurate estimation. The simulation results show that the degree of correlation parameters significantly affects the efficiency of the model. The application to estimate mean years of schooling at the sub-district level in Langkat Regency, North Sumatra, Indonesia, for the period of 2018–2021 showed that the time correlation coefficient was 0.3758, the variance of the area random effect was 1.1125, and the variance of the area-time random effect was 0.3241. The estimations derived from time series and cross-sectional models had a lower mean squared error than those obtained from the Fay-Herriot models and direct estimation.

**Keywords:** Small area estimation, empirical best linear unbiased prediction, mean squared error, mean years of schooling

## **Probabilistic Seismic Hazard Analysis for Sulawesi-Maluku Region of Indonesia using the Space-Time Epidemic Type Aftershock Sequence Model**

Sonia Faradilla<sup>1</sup>, Achmad Choiruddin<sup>1</sup>, Bambang Widjanarko Otok<sup>1</sup>

<sup>1</sup> Department of Statistics, Institut Teknologi Sepuluh Nopember, Kampus ITS-Sukolilo, Surabaya 60111, Indonesia

choiruddin@its.ac.id

**Abstract.** Sulawesi and Maluku are the only regions in Indonesia where three major plates (Eurasian, Indo-Australian and Pacific) meet, resulting in the most seismically active region in Indonesia. It is important to perform probabilistic seismic hazard analysis to understand the seismic activity in this region, study aftershock sequences, and provide comprehensive seismic hazard maps. In this study, we consider spatio-temporal analysis and perform space-time epidemic-type aftershock sequencing (space-time ETAS) to model earthquake catalog data in Sulawesi-Maluku, Indonesia. The method combines Davidon-Fletcher-Powell (DFP) algorithm and stochastic declustering algorithm. We conclude that the model has good performance and can be used to construct seismic hazard maps based on the generated model. The results of the seismic map provide information on the total rate of earthquake events occurring around the Maluku Sea, with the clustering of earthquake events spread almost throughout Sulawesi and Maluku.

**Keywords:** Climate Change Management, Davidon-Fletcher-Powell Algorithm, Earthquake Catalog, Space-Time ETAS Model, Stochastic Declustering.

## **Application of Time Series Regression, Double Seasonal ARIMA, and Long Short-Term Memory for Short-Term Electricity Load Forecasting**

Hafez Afghan and Hidayatul Khusna

Department of Statistics, Sepuluh Nopember Institute of Technology,  
Arief Rahman Hakim, Surabaya 60111, Indonesia

0621194000018@student.its.ac.id, hidayatul@its.ac.id

**Abstract.** Electricity load must be accurately estimated since electricity is non-storable. If electricity is generated more than customer's demand, it will be wasted and the power of generator should be lowered. Meanwhile, if electricity is generated less than customer's demand, it may cause power outage and must undergone backup-plant operation. This study compares three forecasting methods such as Time Series Regression (TSR), Double Seasonal Autoregressive Integrated Moving Average (DSARIMA) and Long-Short Term Memory (LSTM) to predict electricity load. These three methods can be applied to forecast electricity load which has double seasonal pattern. Each method has an ability to capture the data pattern. The observation data is half-hourly recorded electricity load of East Java Province in Mega Watt (MW) units from January 1<sup>st</sup> 2020 to January 31<sup>st</sup> 2023. This study obtains the out-sample sMAPE (symmetric Mean Absolute Percentage Error) of those methods as many as 0.6436%, 0.5504% and 0.9713% respectively. From these findings, DSARIMA  $([2, 10, 11, 12, 15, 16, 17, 18, 19, 20, 21, 22, 23], 1, [1, 2, 3, 7, 8, 30, 34, 35, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48]) (0, 1, 1)^{48} (0, 1, 1)^{336}$  is apparent to be the best forecasting model for its lowest out-sample sMAPE value. This model gained a white noise residual, but it does not have a normal distribution due to outliers.

**Keywords:** Double Seasonal ARIMA, Electricity Load, Long Short-Term Memory, Time Series Regression.

## **A Bayesian Network for Classifying and Predicting Ship Collision**

Iis Dewi Ratih<sup>1</sup>, Ketut Buda Artana<sup>2</sup>, Heri Kuswanto<sup>3</sup>, Emmy Pratiwi<sup>4</sup>, Muhammad Farhan Nuari<sup>5</sup>

<sup>1,2,3,4,5</sup> Institut Teknologi Sepuluh Nopember, ITS, Indonesia

kb.artana@gmail.com

**Abstract.** Ship collisions are the type of accident with the highest percentage of investigations, making them the type of accident with a high variation in causes. Additionally, ship collisions pose a serious threat because they occur between two different vessels, resulting in material losses and loss of life. This condition makes ship collisions a serious problem that requires efforts to minimize prevention and adjust existing conditions. This study aims to model the causes of ship collisions in Indonesia to determine the probability of a ship experiencing a collision or a near miss. The modeling will be conducted using the Bayesian Network method. The Bayesian Network model is based on the factors that cause ship collisions, relying on past incidents and written reports from National Transportation Safety Committee (NTSC) investigations and judgments from the Maritime Court. The purpose of this study is to identify the factors that cause ship collisions, determine the probability of a ship experiencing a collision, and identify the factors that contribute the most to the probability of ship collisions in Indonesia through sensitivity analysis. The results obtained from the model, with a 70% weight for training data, show that the probability of a ship experiencing a collision during a dangerous condition is 63%, with an accuracy and sensitivity of 93.75% and 100% respectively. According to the model, the factors with the greatest influence are "Crew Competence," "Decision Making," "Maneuverability," and "Ship Communication."

**Keywords:** Bayesian Networks, Probability, Sensitivity Analysis, Ship Collision

## **Outlier Detection in Simultaneous Equations with Panel Data**

Suci Ismadyaliana<sup>1,2</sup>, Setiawan<sup>1(\*)</sup>, and Jerry Dwi Trijoyo Purnomo<sup>1</sup>

<sup>1</sup> Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia  
setiawan@statistika.its.ac.id

<sup>2</sup> Badan Pusat Statistik, Jakarta, Indonesia

**Abstract.** An outlier in a set of data is an observation (or subset of observation) that appears to be inconsistent with the remainder of that set of data. Outliers can have a marked effect on any type of empirical analysis. Therefore, outlier detection is very important in a statistical analysis. Outlier detection in regression model used residual to determine the anomaly score. This research used internally studentized residual (ISR) to detect outliers in panel data model and simultaneous equation with panel data. The application of data in this study is data for 10 countries that are members of the ASEAN-China Free Trade Area (ACFTA) during the period 2007–2021. The result showed that outlier observations will vary depending on the type of modeling. Each model will obtain certain residuals, so that the observations detected as outliers will also be different. Outlier detection results in the simultaneous equations show that the outlier observations are a subset of the independent equation outliers that compose them.

**Keywords:** ACFTA, internally studentized residual, outlier detection, panel data, simultaneous equation

## **Fast Maximum Likelihood Estimation of Large Data Spatial Error Model: Matrix Exponential Spatial Specification Approach**

Marsono<sup>1</sup>, Setiawan<sup>1</sup>, and Heri Kuswanto<sup>1</sup>

<sup>1</sup> Department of Statistics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

setiawan@statistika.its.ac.id

**Abstract.** The model of spatial dependence between regions caused by unknown factors or errors is known as spatial error model (SEM). In its application to large datasets, SEM experiences several difficulties in parameter estimation, such as non-linear optimization, theoretical complexity, and computation. One way to overcome these difficulties is by using matrix exponential spatial specification (MESS). The purpose of this research is to find another alternative in modeling data containing spatial autocorrelation error as a substitute for SEM. Labeled MESS(0,1) is an alternative model to replace SEM. Based on the properties possessed by MESS, the MESS(0,1) model is analytically simpler than SEM when using maximum likelihood estimation (MLE). Based on the simulation study results, in addition to being faster in the estimation process, the estimation results obtained from MESS(0,1) have a smaller root mean square error (RMSE) value compared to the SEM model. By looking at these advantages, the MESS(0,1) model is applied to model real data. The real data used is the GRDP of districts and cities in Java Island in the construction sector. This is in line with the massive infrastructure development as a government program. The predictor variables used and considered influential on the GRDP of the construction sector are domestic investment, foreign investment, labor, and wages. The modeling results show that all predictor variables have a significant effect on GRDP in the construction sector. In addition, there are factors of GRDP from other regions, causing inter-regional dependence or spatial dependence.

**Keywords:** SEM, MESS(0,1), RMSE, and MLE.

## **Assessing Departmental Efficiency at Sepuluh Nopember Institute of Technology: A Comparative Study Using Classical and Advanced Data Envelopment Analysis**

Zakiatul Wildani<sup>1\*</sup>, M Naufal Mubarik<sup>1</sup>, Sri Pingit Wulandari<sup>1</sup>, Lucia Aridinanti<sup>1</sup>,  
Muhammad Alifian Nuriman<sup>1</sup>

<sup>1</sup>Department of Business Statistics, Vocational Faculty, Sepuluh Nopember Institute of Technology, Surabaya, Indonesia

zakia@its.ac.id

**Abstract.** As the significance of sustainable development continues to grow, particularly in the field of education, it becomes imperative to evaluate the effective utilization of resources by universities. Efficiency plays a pivotal role in enabling organizations to optimize their resource allocation and achieve optimal outcomes. In this study, we employ Data Envelopment Analysis (DEA) to assess the efficiency of 38 departments at Sepuluh Nopember Institute of Technology Surabaya or ITS. We examine three distinct approaches: Teaching and Learning, Research and Development, and an overarching overall performance assessment. Our analysis employs both classical and advanced models, incorporating four input and five output variables. The findings consistently indicate that the Department of Instrumentation Engineering, excels in efficiency, regardless of the chosen approach or DEA model. This underscores its role as an exemplar for less efficient departments, offering valuable insights for improvement and aiding university administrators in setting ambitious objectives. Notably, a substantial portion of ITS's departments were identified as inefficient for research and development scheme. Therefore, it is imperative for the university to concentrate on enhancing departmental productivity through initiatives such as workshops and training programs.

**Keywords:** Data Envelopment Analysis, Departmental Efficiency, Efficiency, Higher Education

## **Multivariate Adaptive Fuzzy Clustering Means Regression Splines model using Generalized Cross-Validation (GCV) on stunting cases in Southeast Sulawesi.**

Mira Meilisa<sup>1,2</sup>, Bambang Widjanarko Otok<sup>1</sup>, and Jerry Dwi Trijoyo Purnomo<sup>1</sup>

<sup>1</sup>Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia  
bambang\_wo@statistika.its.ac.id, jerry@statistika@its.ac.id

<sup>2</sup>Muhammadiyah University of Sumatera Barat, Sumatera Barat, Indonesia  
mirameilisa@umsb.ac.id

**Abstract.** Regression modeling cannot be solved with a parametric approach if there is no information about the shape of the function and there is no clear pattern of relationship between the response variable and the predictor variables. One approach that can be done is to use a nonparametric regression approach. Multivariate adaptive regression splines (MARS) are one of the nonparametric regression models that can accommodate additive effects and interaction effects between predictor variables in data modeling. MARS is a model capable of handling high-dimensional data, for sample sizes of 50-1000 and 3-20 predictor variables. This study uses the MARS model with fuzzy clustering means (FCM) which is then called the multivariate adaptive fuzzy clustering means regression splines model (MAFCMRS). Creating a continuous model at knots by paying attention to aspects of heterogeneity. The results of the cluster validity test using the validity of the Xie and Beni index produced 5 clusters which will be formed into the multivariate adaptive fuzzy clustering means regression splines model. This model was used to model the prevalence of stunting in Southeast Sulawesi province using 222 district observation units. The results obtained from the 4 predictor variables used are number pregnant women with chronic energy deficiency (CED), number of low birth weight of the baby, number of neonate visits, number of health centers, all of which affected the prevalence of stunting at the 5% significance level.

**Keywords:** MARS, FCM, MAFCMRS, GCV, Stunting

## **Statistical Inferences for Multivariate Generalized Gamma Regression Model**

Hasbi Yasin<sup>1,2</sup>, Purhadi<sup>1</sup> and Achmad Choiruddin<sup>1</sup>

<sup>1</sup> Department of Statistics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

<sup>2</sup> Department of Statistics, Universitas Diponegoro, Semarang, Indonesia

purhadi@statistika.its.ac.id

**Abstract.** Generalized Gamma (GG) distribution serves as a widely applied statistical tool, particularly suitable for scenarios where data distribution skews positively and lacks symmetry. In many real-world situations, multiple factors can simultaneously influence various outcomes. This article introduces the Multivariate Generalized Gamma Regression (MGGR) model, tailored for data adhering to a Multivariate Generalized Gamma (MGG) distribution. Parameter estimation in MGGR relies on the Maximum Likelihood Estimation (MLE) technique, further optimized with the Berndt-Hall-Hausman (BHHH) algorithm to enhance precision. To assess the model's significance, we deploy the Maximum Likelihood Ratio Test (MLRT) and conduct partial testing using the Wald test. Rigorous validation through simulations demonstrates the MGGR model's adeptness in parameter estimation, exhibiting minimal bias. To underscore its practicality, we apply the MGGR model to a real-world case study. Specifically, we employ it to analyze three education indicators spanning 2017 to 2021 in Central Java, Indonesia. Our findings highlight the efficacy of multivariate modeling over its univariate counterpart, revealing a more logical approach to data analysis. In summary, this research underscores the robustness of the MGGR model in parameter estimation and highlights the benefits of embracing multivariate modeling for comprehensive data insights.

**Keywords:** Educational Indicator, Generalized Gamma Distribution, Maximum Likelihood Estimation, Multivariate Generalized Gamma Regression

## **Track 4: Health & Industrial Applications**

<b>NO</b>	<b>ID</b>	<b>TITLE</b>	<b>AUTHORS</b>
1.	1570956304	W@RK: ATTENDANCE APPLICATION FRAMEWORK USING BLOCKCHAIN TECHNOLOGY	Putra Roskhairul Fitri Kaha, Syarifah Bahiyah Rahayu, Afiqah M. Azahari, Mohd Hazali Mohamed Halip and K. Venkatesan
2.	1570950111	Exploring the Impact of COVID-19 on Individuals' Mental Health through Cluster Analysis	Azlin Ahmad, Siti Nabilah Mohd Abdul Hakim Amir, Ezzatul Akmal Kamaruzaman and Abdulaziz Al-Nahari
3.	1570953200	Enhancing Data Collection Efficiency through Multithreaded Web Scraping Approach	Dahlan Nariman
4.	1570958670	Tracking High Potential Transmission Risk Spots of Infectious Disease Using Spatial Social Network Analysis and Visualisation (SSNAV) Techniques	Ilham Abdul Jalil and Abdul Rauf Abdul Rasam
5.	1570950594	Breast Cancer Classification: A Transfer Learning Approach Using EfficientNets	Samiya QANOUNE, Hassan AMMOR and Zouhair GUENNOUN
6.	1570955852	Evaluation Of Machine Learning Algorithms For Early Prediction Of Liver Disease	Sushmitha Geddam, Sulaf Assi, Hadi Naghavipour, Manoj Jayabalan, Abdullah Al-Hamid and Dhiya Al-Jumeily OBE
7.	1570957416	An agricultural information recommendation method based on matrix decomposition knowledge graph algorithm	Ruipeng Tang, Narendra Kumar Aridas and Mohamad Sofian Abu Talip
8.	1570947738	Performance of the Auxiliary Information Based Hybrid EWMA Chart with Fast Initial Response	Peh Sang Ng, Huai Tein Lim, Wai Chung Yeong and Sajal Saha

9.	1570958440	Evaluation of Machine Learning Models for Breast Cancer Detection in Microarray Gene Expression Profiles	Mohammad Nasir Abdullah and Yap Bee Wah
----	------------	--	---

## **W@RK: ATTENDANCE APPLICATION FRAMEWORK USING BLOCKCHAIN TECHNOLOGY**

Putra Roskhairul Fitri Kaha<sup>1</sup>, Syarifah Bahiyah Rahayu<sup>1,2\*</sup> Afiqah M. Azahari<sup>3</sup>, Mohd Hazali Mohamed Halip<sup>1</sup>, K. Venkatesan<sup>1</sup>

<sup>1</sup> Faculty of Defense Science and Technology, National Defence University of Malaysia (UPNM), Malaysia.

<sup>2</sup> Cyber Security and Digital Industrial Revolution Centre, National Defence University of Malaysia (UPNM), Malaysia.

<sup>3</sup> Campus Sophia Tech, 450 Route des Chappes, 06410 Biot, France

\*SyarifahBahiyah@upnm.edu.my

**Abstract.** Post-COVID has inadvertently caused a widespread shift towards remote working for many employees across the globe. This new normal has necessitated a focus on employee's well-being, making it a top priority for companies. However, many businesses are concerned about the potential for abuse or remote working arrangements and the risk of data security breaches due to various security concerns. As a result, there is a need to develop a new attendance application that can facilitate a shift from a work-from-office setup to a work-from-anywhere arrangement. This paper introduces a contactless attendance, W@RK, embedded with a face recognition feature and location-based for recording employee attendance using blockchain technology. The aim is to monitor employee attendance and empower employees to work from anywhere. The objective is to develop an alternative contactless attendance application. Agile software development methodology is adopted to develop this attendance application. A facial recognition feature is used to scan the face of employees to authenticate themselves through verification services. While the location is detected using geo-fencing, the blockchain technology record and verify the attendance of the employee. The facial data is recorded on a blockchain, along with a timestamp and other relevant information such as location. The findings show the proposed framework has a promising capability to record and verify attendance securely and efficiently. As the use of blockchain technology continues to grow, it is likely more organizations will adopt facial attendance systems to track the attendance of employees, students, and other individuals at events and meetings.

**Keywords:** Attendance, location-based, face recognition, blockchain, agile.

## **Exploring the Impact of COVID-19 on Individuals' Mental Health through Cluster Analysis**

Azlin Ahmad<sup>1</sup>, Siti Nabilah Mohd Abdul Hakim Amir<sup>1</sup>, Ezzatul Akmal Kamaruzaman<sup>1</sup>, Abdulaziz Al-Nahari<sup>2</sup>

<sup>1</sup> School of Computing, College of Computing, Informatics and Mathematics  
Universiti Teknologi MARA, Shah Alam, Malaysia

<sup>2</sup> School of Information Technology, UNITAR International University, Petaling Jaya,  
Selangor, Malaysia

<sup>1</sup>azlin@tmsk.uitm.edu.my, <sup>1</sup>sitinabilahnorhisham@gmail.com, <sup>1</sup>ezzatul@tmsk.uitm.edu.my,  
abdulaziz.yahya@unitar.my

**Abstract.** The COVID-19 pandemic has had a significant impact on mental health, resulting in anxiety and other issues among many individuals due to the lockdowns implemented to curb its spread. With the world moving towards 2030, it is crucial to reduce premature mortality from non-communicable diseases through prevention and treatment. Sustainable Development Goal (SDG) 3 emphasizes prioritizing mental health and well-being to address the increasing burden of mental health issues. The study utilized text clustering through the K-Means algorithm to gain a better understanding of the mental health issues people are facing. The Term Frequency-Inverse Document Frequency (TF-IDF) was used to determine each word's weight after extracting tweets from Twitter and pre-processing the data. The K-Means clustering algorithm was then employed in the data, which revealed that the clusters could be classified into three categories of mental health: 'stress,' 'depression,' and 'pressure.' It was found that using three clusters provided more dependable outcomes since clusters with more than three tended to have overlapping mental health conditions. This study sheds light on the mental health problems that people face during the COVID-19 pandemic, which can help guide efforts to support those in need. Moreover, it would be more beneficial to incorporate Bahasa Malaysia in future research since there has yet to be much exploration done on this language despite it being Malaysia's official language. By adopting a holistic approach and prioritizing mental health, we can work towards ensuring a healthier and happier future for everyone.

**Keywords:** Clustering, K-Means Clustering, Mental Health, Term Frequency-Inverse Document Frequency.

## **Enhancing Data Collection Efficiency through Multithreaded Web Scraping Approach**

<sup>1</sup>Dahlan Nariman

<sup>1</sup> College of Sustainability and Tourism, Ritsumeikan Asia Pacific University (APU),  
1-1 Jyumonjibaru Beppu-shi, Oita, Japan

dahlan@apu.ac.jp

**Abstract.** With the penetration of new technologies, there is rapid growth of internet users and data. The Internet provides a variety of human-created facts and data sources such as online customers reviews. The data commonly consists of a vast array of disjointed structured and unstructured data that are difficult to collect by physical means and problematic to use in mechanical processes. In addition, it is required to use a simple tool with simple user interface to make everyone easier to collect huge numbers of customers reviews and store it in specific database. We are developing a tool to collect hotels customers reviews data called MULARS (Multi-Languages Reviews Scrapers) for collecting large-scale of multi-language customers reviews. In the current system, we faced several challenges while collecting huge number of disjoint unstructured of online reviews data, such as speed of data collection, intensive computation capacity and reliability of data extraction. In this paper, we investigate multi-thread web scraping architecture possibilities to enhance performances and efficiency of reviews data collection. Our solution tries to address both scraping efficiency and feasibility for big data analyses and applications. This paper is started by describing the existing online data collection methods and the problems statement. Then, the proposed system will be explained, and the results of experiment of data collection performances will be presented. Finally, the paper discusses the improvement by adopting the multi-threading, the tool advantages, and future possibilities of the tool applications on the different areas.

**Keywords:** Big Data, Customers Reviews, Data Collection, Multi-thread.

## **Tracking High Potential Transmission Risk Spots of Infectious Disease Using Spatial Social Network Analysis and Visualisation (SSNAV) Techniques**

Ilham Abdul Jalil<sup>1</sup> and Abdul Rauf Abdul Rasam<sup>2\*</sup>

<sup>1,2</sup> School of Geomatics and Natural Resources, College of Built Environment, Universiti Teknologi MARA, Shah Alam, Selangor, MALAYSIA

ilhamabdjilil@gmail.com, \*rauf@uitm.edu.my

**Abstract.** The risk map for infectious disease shows the importance of the Geographical Information System (GIS) and spatial social network analysis and visualisation (SSNAV) as a preparedness and response tool to strengthen the capacity for assessing health risks. The current mapping method still needs to be revised to detect the potential risk areas of the disease due to the need for more dynamic spatial and social elements, especially in identifying human mobility effects in detecting missing tuberculosis (TB) cases. This study has combined GIS-MCDM and SSNAV techniques to evaluate whether this combination will enhance TB's general existing disease hotspot mapping in Klang, Selangor. The social network structure of selected TB cases in Klang as actors (nodes) and human mobility (home-workplace) data as edges have been used to investigate social network mobility structures, analyse the relationships among the nodes and study their edges regarding their network centrality. The main finding has revealed that the higher the node's centrality in the network structure, the higher the chance the node influences the TB spread in the whole network after comparing the network graph results with the GIS mapping technique. Combining these techniques increases the existing mapping capabilities towards enhancing the understanding of how diseases move through the population and creating a reliable potential risk map in Malaysia.

**Keywords:** SSNAV, GIS-MCDM, Risk Disease Map, TB Hotspots, Infectious Disease.

## **Breast Cancer Classification: A Transfer Learning Approach Using EfficientNets**

Samiya QANOUNE<sup>1</sup>, Hassan AMMOR<sup>2</sup> and Zouhair GUENNOUN<sup>3</sup>

<sup>1,2,3</sup> Smart Communication Research Team, Mohammadia Engineering School,  
Mohammed V University, Rabat, MOROCCO.

samiyaqanoune@research.emi.ma

**Abstract.** In recent years, Deep Learning has emerged as a powerful tool in the early detection of breast cancer, enabling timely initiation of treatment. However, existing approaches often grapple with extended training/testing durations and a notable incidence of false positives. In response, this paper introduces a novel model leveraging transfer learning and deep learning features' extraction techniques. By adapting a pre-trained CNN model to address the specific problem at hand, our proposed model attains an exceptional accuracy rate of 99.66%. Notably, it surpasses the accuracy and efficiency of conventional methods, marking a significant advancement in breast cancer detection.

**Keywords:** Breast cancer , Artificial intelligence, Deep learning, Transfer learning, EfficientNets.

## **Evaluation Of Machine Learning Algorithms For Early Prediction Of Liver Disease**

Sushmitha Geddam<sup>1</sup>, Sulaf Assi<sup>1</sup>, Hadi Naghavipour<sup>2</sup>, Manoj Jayabalan<sup>1</sup>, Abdullah Al-Hamid<sup>3</sup> and Dhiya Al-Jumeily OBE<sup>1,\*</sup>

<sup>1</sup>Liverpool John Moores University, UK

<sup>2</sup> School of Information Technology, UNITAR International University, Petaling Jaya, Selangor, Malaysia

<sup>3</sup>College of Clinical Pharmacy, King Faisal University, Saudi Arabia

d.aljumeily@ieee.org

**Abstract.** Liver disease contributes to major morbidities and mortalities globally. Around 2 million people die annually worldwide from liver diseases. Diagnosing liver disease is usually made by measuring levels of biomarkers that could be enzymes linked to liver function tests and/or proteins. However, not often the levels of these biomarkers change in cases of liver diseases. The majority of diagnosis based on biomarkers occur when the liver is partially or fully damaged. On the other hand, early diagnosis of the disease can prevent further complication, decrease the burden on healthcare systems and saves lives. Subsequently, this research evaluated machine learning algorithms for early prediction of liver disease. These algorithms included logistic regression, decision tree, Random Forest, adaptive boosting, extreme gradient boost, support vector machine and Naïve Bayes, and were applied to a dataset of patients with and without liver diseases after data pre-processing. Metrics for evaluation included accuracy, precision, recall, AUC-ROC and F1 score. Based on these metrics and after hyperparameter tuning, support vector machine, random forest, adaptive gradient boost and extreme gradient boost showed to be the best performing models. However, the performance of these models did not exclude overfitting that could be related to the low sample size of the data set. Future work involves applying these algorithms to a larger sample size of patients and more features.

**Keywords:** Liver disease, Diagnosis, Liver Function Tests, Machine Learning Algorithms, Evaluation.

## **An agricultural information recommendation method based on matrix decomposition knowledge graph algorithm**

Ruipeng Tang<sup>1\*</sup>, Narendra Kumar Aridas<sup>1</sup> and Mohamad Sofian Abu Talip<sup>1</sup>

<sup>1</sup> Department of Electrical Engineering, Faculty of Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia;

22057874@siswa.um.edu.my(RT);narendra.k@um.edu.my(NKA);sofian\_abutalip@um.edu.my(MSAT);

<sup>2</sup> Correspondence: 22057874@siswa.um.edu.my; Tel.: (+60)11-3339-1886

**Abstract.** With the promotion of the Internet of Things, big data, and other technologies in agricultural production, farmers are accustomed to searching for crop cultivation information through internet. However, the traditional way of information search has problems such as time-consuming, inefficient, and inaccurate, so farmers urgently need an efficient, fast, and accurate information recommendation method. So this study proposes an agricultural information recommendation model based on the MDKG algorithm (matrix decomposition knowledge graph algorithm). It introduces the matrix decomposition algorithm based on the knowledge graph, which can constructs and trains the interaction matrix between users and agricultural information. It also solves the decomposed interaction matrix, and obtains the correlation score between users and agricultural information through calculation. Experimental results show that the MDKG algorithm is better than the DKN and RippleNet algorithms in AUC and ACC of information flow clicks and Recall and Precision of cold start recommendations, which indicates that the algorithm can deeply mine user preference characteristics to improve click-through rates and perform better in cold start scenarios. It can better alleviate the negative impact of data sparseness, which improves the efficiency of farmers to obtain agricultural information.

**Keywords:** Knowledge graph, Matrix decomposition algorithm, Agricultural information recommendation, Personalized recommendation, Information retrieval

## **Performance of the Auxiliary Information Based Hybrid EWMA Chart with Fast Initial Response**

Peh Sang Ng<sup>1</sup>, Huai Tein Lim<sup>2</sup>, Wai Chung Yeong<sup>3</sup>, and Sajal Saha<sup>4</sup>

<sup>1,2</sup> Department of Physical and Mathematical Science, Faculty of Science,  
Universiti Tunku Abdul Rahman, 31900 Kampar, Perak, Malaysia

<sup>3</sup> School of Mathematical Sciences, Sunway University, 47500 Petaling Jaya, Malaysia

<sup>4</sup> Department of Mathematics, International University of Business Agriculture  
and Technology, Dhaka, Bangladesh

<sup>1</sup>psng@utar.edu.my

<sup>2</sup>limhuait@utar.edu.my

<sup>3</sup>waichungy@sunway.edu.my

<sup>4</sup>sajal.saha@iubat.edu

**Abstract.** An integration of the fast initial response (FIR) feature into a control chart helps in increasing the shift detection speed for the process with start-up quality problem or inefficient corrective measures towards an occurrence of off-target process. Due to the salient feature of the FIR feature, this study investigates the performance of the FIR auxiliary information-based hybrid EWMA (denoted as the AIB-HEW-FIR) chart by using the regression estimator. To evaluate the efficiency of the AIB-HEW-FIR chart in triggering off-target process, the performance indicators of the Average Run Length (ARL) and Expected ARL (EARL) are used, and the values of the ARL and EARL are computed by using the Monte Carlo simulation. The proposed chart is then evaluated and compared with the auxiliary information-based hybrid EWMA (AIB-HEW) chart without the FIR feature. Results show that the AIB-HEW-FIR chart is superior to the AIB-HEW chart in detecting infrequent changes in the process mean.

**Keywords:** AIB-HEW, Fast Initial Response, Auxiliary Information, Average Run Length, Expected Average Run Length, Off-target Process.

## **Evaluation of Machine Learning Models for Breast Cancer Detection in Microarray Gene Expression Profiles**

Mohammad Nasir Abdullah<sup>1</sup> and Yap Bee Wah<sup>2</sup>

<sup>1</sup> School of Mathematical Sciences, College of Computing, Informatics and Mathematics, Universiti Teknologi MARA, Tapah Campus, 35400 Tapah, Perak, Malaysia.  
nasir916@uitm.edu.my

<sup>2</sup> School of Information Technology, UNITAR Graduate School, UNITAR International University, Tierra Crest, Jalan SS6/3, 47301 Petaling Jaya, Selangor, Malaysia.  
bee.wah@unitar.my

**Abstract.** Breast cancer (BC) is a leading global health challenge, with survival rate varying significantly across regions due to socioeconomic disparities and healthcare accessibility. This research seeks to identify the most efficient machine learning (ML) classifier for precise BC classification using gene expression data. Utilizing the CuMiDa database's microarray BC dataset, which includes 35,983 gene biomarkers from 146 breast adenocarcinoma patients and 143 normal subjects, the study employed R-programming for data pre-processing and feature selection. The Boruta algorithm pinpointed 214 key biomarkers, and the dataset was subsequently balanced using the SMOTE technique. Among the seven ML classifiers assessed, the support vector machine (SVM) showcased superior performance metrics such as sensitivity, specificity, and accuracy, while naïve Bayes (NB) underperformed. A thorough examination of the BC dataset revealed that SVM is the premier ML classifier, highlighting its potential for enhancing BC predictive modeling.

**Keywords:** Boruta algorithm, Breast Cancer, CuMiDA, Machine Learning, and Support Vector Machine.

## **Conference Organization**

### **Patron**

Professor Emeritus Tan Sri Dato' Sri Ir Dr Sahol Hamid Bin Abu Bakar,  
Vice-Chancellor, UNITAR International University, Malaysia

### **Honorary Chair**

Prof. Dr. Witold Pedrycz, University of Alberta, Canada  
Dato' Sri Dr Uzir Mohd Uzir Mahidin, Department of Statistics, Malaysia

### **General and Founding Chairs**

Prof. Dr. Yap Bee Wah, UNITAR International University, Malaysia  
Prof. Dr. Dhiya Al-Jumeily, OBE, Liverpool John Moores University, UK

### **International Advisory Committee**

Prof. Dr. Mario Koeppen, Kyushu Institute of Technology, Japan  
Prof. Dr. Heri Kuswanto, Institute Teknologi Sepuluh Nopember  
Prof. Dr. Mohammed Bennamoun, University of Western Australia, Australia  
Prof. Dr. Chidchanok Lursinsap, Chulalongkorn University, Thailand

### **Proceeding Editors**

Prof. Dr. Yap Bee Wah, UNITAR International University, Malaysia  
Prof. Dr. Dhiya Al-Jumeily, OBE, Liverpool John Moores University, UK  
Prof. Dr. Michael W. Berry, University of Tennessee, USA

### **Finance and Registration Committee**

Chong Kim Loy, UNITAR International University, Malaysia  
Normaiza Binti Mohamad, UNITAR International University, Malaysia  
Simranpreet Kaur Hansaram, UNITAR International University, Malaysia

### **Secretary**

Sharifah Nurul Aina binti Sayed Burhanudin, UNITAR International University,  
Malaysia

### **Technical Programme Committee**

Abdulaziz Al-Nahari, UNITAR International University, Malaysia (Chair)  
Azlin Ahmad, Universiti Teknologi MARA, Malaysia (Co-Chair)  
Aitizaz Ali, UNITAR International University, Malaysia  
Hadi Naghavipour, UNITAR International University, Malaysia  
Jamila Mustafina, Kazan Federal University, Russia  
Sulaf Assi, Liverpool John Moores University, UK  
Umar Farooq Khattak, UNITAR International University, Malaysia

Tahir Mehmood, UNITAR International University, Malaysia  
Wasiq Khan, Liverpool John Moores University, UK  
Abir Jaafar Hussain, Liverpool John Moores University, UK  
Anazida Binti Zainal, Universiti Teknologi Malaysia, Malaysia  
Wan Fairos Wan Yaacob, Universiti Teknologi MARA, Malaysia  
Syerina Azlin Md Nasir, Universiti Teknologi MARA, Malaysia  
Mohd Fadzil Hassan, Universiti Teknologi PETRONAS, Malaysia  
Sapto Wahyu Indratno, Institut Teknologi Bandung, Indonesia  
Farhad Nadi, UNITAR International University, Malaysia  
Syarifah Bahiyah Rahayu Binti Syed Mansoor, Universiti Pertahanan Nasional  
Malaysia, Malaysia

### **Programme Book Committee**

Noor Lees Ismail, UNITAR International University, Malaysia  
Iznora Aini Binti Zolkifyl, UNITAR International University, Malaysia  
Rohaizah Abd Latif, UNITAR International University, Malaysia  
Jan Lunn, Liverpool John Moores University, UK

### **IT Committee**

Danny Ngo Lung Yao, UNITAR International University, Malaysia  
Noor Azma Binti Ismail, UNITAR International University, Malaysia  
Paridah Binti Daud, UNITAR International University, Malaysia  
Intan Izzatul Fariza Binti Rosli, UNITAR International University, Malaysia  
Ker Boon Chin, UNITAR International University, Malaysia  
Muhamad Hadri Bin Mohd Hassan, UNITAR International University, Malaysia  
Mohd Aiman Mohd Sani, UNITAR International University, Malaysia  
Ahmad Ruzaini bin Rahim, UNITAR International University, Malaysia

### **Logistics Committee**

Mohd Farizudin Mohammad Fauzi, UNITAR International University, Malaysia  
Muhammad Shahrir Aizat Muhammad Shuhaili, UNITAR International University,  
Malaysia  
Roziyana Binti Bahrudin, UNITAR International University, Malaysia  
Nurlisnawati Binti Mohd Hijazi, UNITAR International University, Malaysia  
Syed Munir Barakbah Bin Syed Faozi Barakbah, UNITAR International University,  
Malaysia  
Mohd Amar Bin Mohd Mokhtar, UNITAR International University, Malaysia  
Farah Nazurah Binti Zainal, UNITAR International University, Malaysia

### **Publicity & Strategy Committee**

Hadi Naghavipour, UNITAR International University, Malaysia (Chair)  
Wan Zakiyatussariroh Wan Husin, Universiti Teknologi MARA, Malaysia  
Bander Ali Al-Rimy, Universiti Teknologi Malaysia, Malaysia

Izzatdin Abdul Aziz, Universiti Teknologi PETRONAS, Malaysia  
Norshahriah Binti Abdul Wahab, Universiti Pertahanan Nasional Malaysia, Malaysia

**Corporate Committee**

Stella Chua Ching Yee, UNITAR International University, Malaysia  
Marc Kevin Natusch, UNITAR International University, Malaysia  
Mohamad Shah Andrew Ibrahim, UNITAR International University, Malaysia  
Nikki Poh Li Yi, UNITAR International University, Malaysia

**Industry Committee**

Badrie Abdullah, UNITAR International University, Malaysia  
Visnuvarthen A/ L Sakayam, UNITAR International University, Malaysia

**International Scientific Committee**

Adel Al-Jumaily, Charles Sturt University, Australia (Chair)  
Naomi Bt Salim, Universiti Teknologi Malaysia, Malaysia  
Ku Ruhana Ku Mahmud, Universiti Utara Malaysia, Malaysia  
Siddhivinayak A. Kulkarni, MIT-World Peace University, Pune, India  
Simon Fong, University of Macau, China  
Richard Millham, Durban University of Technology, South Africa  
Layth Sliman, Efrei, Paris, France

## **Reviewers**

Dr. Abdullah Al Hamid, Saudi Ministry of Health, Saudi Arabia  
Dr. Adam Shariff Adli Aminuddin, Universiti Malaysia Pahang, Malaysia  
Dr. Aitizaz Ali, UNITAR International University, Malaysia  
Dr. Arunava Roy, The University of Memphis, USA  
Dr. Noor Asiah Ramli, Universiti Teknologi MARA, Malaysia  
Dr. Azlin Ahmad, Universiti Teknologi MARA, Malaysia  
Dr. Bander Ali Al-Rimy, Universiti Teknologi Malaysia, Malaysia  
Chong Kim Loy, UNITAR International University, Malaysia  
Dr. Danny Ngo Lung Yao, UNITAR International University, Malaysia  
Dr. Dittaya Wanvarie, Chulalongkorn University, Thailand  
Prof. Ensar Gul, Maltepe University, Turkey  
Assoc Prof. Dr. Faiz Ahmed Elfaki, Qatar University, Qatar  
Dr. Feras Zen Alden, Cihan University Erbil, Iraq  
Dr. Hizir Sofyan, Syiah Kuala University, Indonesia  
Dr. Yun Huoy Choo, Universiti Teknikal Malaysia Melaka, Malaysia  
Dr. Nur Intan Raihana Ruhaiyem, Universiti Sains Malaysia, Malaysia  
Dr. Iznora Aini Zolkifly, UNITAR International University, Malaysia  
Assoc Prof. Dr. Izzatdin Abdul Aziz, Universiti Teknologi PETRONAS, Malaysia  
Prof. Jamila Mustafina, Kazan Federal University, Russia  
Dr. Kholoud Baba, Beirut Arab University, Lebanon  
Assoc Prof. Dr. Chin Kim On, Universiti Malaysia Sabah, Malaysia  
Dr. You Beng Koh, Universiti Malaya, Malaysia  
Assoc Prof. Dr. Kok-Haur Ng, University of Malaya, Malaysia  
Prof. Chidchanok Lursinsap, Chulalongkorn University, Thailand  
Dr. Siow Hoo Leong, Universiti Teknologi MARA, Malaysia  
Dr. Mahasen El Dimassi, University of Engineering, Lebanon  
Dr. Muhammad Yaqoob, University of Hertfordshire, United Kingdom  
Dr. Muhammad Babar, Prince Sultan University, Saudi Arabia  
Assoc Prof. Dr. Mohd Fadzil Hassan, Universiti Teknologi PETRONAS, Malaysia  
Dr. Michael Loong Peng Tan, Universiti Teknologi Malaysia, Malaysia  
Assoc Prof. Dr. Mohamed Imran Mohamed Ariff, Universiti Teknologi MARA,  
Malaysia  
Dr. Mohd Zamri Osman, Universiti Teknologi Malaysia, Malaysia  
Dr. Najwa Hayaati Mohd Alwi, Universiti Sains Islam Malaysia, Malaysia  
Dr. Mohammad Nasir Abdullah, Universiti Teknologi MARA, Malaysia  
Dr. Muhammad Noryanti, Universiti Malaysia Pahang Al-Sultan Abdullah, Malaysia  
Dr. Nuhman Ul Haq, COMSATS University Islamabad, Abbottabad Campus Pakistan  
Dr. Nurain Ibrahim, Universiti Teknologi MARA, Malaysia  
Dr. Nurfazrina Mohd Zamry, Universiti Teknologi Malaysia, Malaysia  
Dr. Pakawan Pugsee, Chulalongkorn University, Thailand  
Muhammad Mazhar Rathore, University of New Brunswick, Canada

Prof. Richard Millham, Durban University of Technology, South Africa  
Dr. Ruhaila Maskat, Universiti Teknologi MARA, Malaysia  
Dr. Sulaf Assi, Liverpool John Moores University, United Kingdom  
Dr. Sakhinah Bakar, Universiti Pertahanan Nasional Malaysia, Malaysia  
Prof. Dr. Sapto Wahyu Indratno, Institut Teknologi Bandung, Indonesia  
Norkhushaini Awang, Universiti Teknologi MARA, Malaysia  
Dr. Sameer Khan, King Abdullah University of Science and Technology, Saudi Arabia  
Dr. Syed Hamad, Hazara University, Malaysia  
Assoc. Prof. Dr. Syerina Azlin Md Nasir, Universiti Teknologi MARA, Malaysia  
Dr. Umar Farooq Khattak, UNITAR International University, Malaysia  
Dr. Umara Urooj, Universiti Teknologi Malaysia, Malaysia  
Dr. Uswah Khairuddin, Universiti Teknologi Malaysia, Malaysia  
Dr. Wasiq Khan, Liverpool John Moores University, United Kingdom  
Dr. Wan Zakiatussariroh Wan Husin, Universiti Teknologi MARA, Malaysia  
Assoc Prof. Dr. Wan Fairos Wan Yaacob, Universiti Teknologi MARA, Malaysia  
Dr. Waseem Alromema, Taibah University, Saudi Arabia  
Dr. Yazan AL-Khassawneh, Zarqa University, Jordan  
Dr. Abdulrazak Yahya, University Malaysia Sarawak, Malaysia  
Dr. Siti Zanariah Satari, Universiti Malaysia Pahang, Malaysia

## Acknowledgements

Organised by



In collaboration with

