The 11th International Conference on Computer Science and Computational Intelligence (ICCSCI)

Program Book ICCSCI 2026

18th - 19th June 2026

BINUS @ Medan



Co-host:

TABLE OF CONTENTS

01

Welcoming Remark
Chairman ICCSCI 2026

02

Welcoming Remark - Dean
School of Computer Science
BINUS University

03

Welcoming Remark - Vice Rector
Research and Technology Transfer
BINUS University

05

Welcoming Remark - Rector
BINUS University

06

Keynote Abstract
Jae-Hyeon Park, Ph.D.
Bina Nusantara University, Indonesia

TABLE OF CONTENTS

07

Keynote Abstract

Prof. Poltak Sihombing, Ph.D.

Universitas Sumatera Utara (USU), Indonesia

08

Keynote Abstract

Prastudy Fauzi, Ph.D.

Nanyang Technological University, Singapore

09

General Rundown

12

Parallel Session Rundown and Paper Presentation Abstract

Welcoming Remark
Chairman ICCSCI 2025

Dear esteemed participants, colleagues, and guests,

On behalf of the organizing committee, it is my great honor and privilege to warmly welcome all of you to ICCSCI 2026. The 11th International Conference on Computer Science and Computational Intelligence (ICCSCI) .

We are deeply honored by the presence of our university leaders, distinguished guests, and international participants who have joined us in this important academic gathering. Your presence reflects our shared commitment to advancing knowledge, innovation, and global collaboration in the rapidly evolving digital era.

The theme of ICCSCI 2026, “Fostering Human-Centered and Inclusive Innovation for a Digitally Empowered Society,” highlights the importance of ensuring that technological advancement remains focused on humanity, accessibility, inclusivity, and societal well-being. As Artificial Intelligence, computational intelligence, and digital technologies continue to transform our lives, we must ensure that innovation not only becomes more advanced, but also more ethical, equitable, and beneficial for all communities.

We are especially honored to host this milestone edition of ICCSCI in Medan, a dynamic gateway city in North Sumatra known for its cultural diversity, historical richness, and vibrant culinary landscape. Beyond the exchange of innovative ideas and scholarly discussions, we hope you will also enjoy the unique charm of the city and the genuine hospitality of its people.



We are proud to share that ICCSCI 2026 received an outstanding total of 891 paper submissions. Following a thorough and rigorous peer-review process, 151 submissions were accepted, and an overall acceptance rate of 16.6%. The accepted proceedings will be published in Procedia Computer Science, a highly regarded Q2 journal with strong impact, with which we have proudly collaborated for over a decade. We extend our sincere gratitude to all participants, reviewers, and contributors for their invaluable efforts in making this conference a success.

We sincerely hope that your time in Medan, whether in person or through our hybrid platform, will be memorable, enriching, and intellectually rewarding. Thank you for being part of ICCSCI 2026.

Warm regards,

Dr. Mochammad Haldi Widianto, ST., MT.
Chairman ICCSCI 2026

Welcoming Remark

Dean – School of Computer Science
BINUS University



Dear esteemed participants, colleagues, and guests,

On behalf of the School of Computer Science, Bina Nusantara University, it is my great pleasure to welcome you to the 11th International Conference on Computer Science and Computational Intelligence (ICCSCI 2026), held at BINUS @Medan, Indonesia. We are honored to gather once again distinguished researchers, academics, industry professionals, and students from around the world to share ideas, present innovations, and strengthen collaborations in the fields of Computer Science and Computational Intelligence.

On behalf of the School of Computer Science, Bina Nusantara University, it is my great pleasure to welcome you to the 11th International Conference on Computer Science and Computational Intelligence (ICCSCI 2026), held at BINUS @Medan, Indonesia. We are honored to gather once again distinguished researchers, academics, industry professionals, and students from around the world to share ideas, present innovations, and strengthen collaborations in the fields of Computer Science and Computational Intelligence.

It is my distinct pleasure to welcome our esteemed keynote speakers:

- Prof. Poltak Sihombing
- Jae-Hyeon Park, Ph.D.
- Prastudy Fauzi, Ph.D.

Their expertise and contributions to the field will undoubtedly inspire meaningful discussions and valuable insights throughout the conference.

Over the years, ICCSCI has evolved into a respected international platform for knowledge exchange and academic networking.. As we enter the 11th edition of ICCSCI, we are proud to witness the continued growth of this conference and its contribution to advancing research, innovation, and global academic cooperation.

In today's rapidly evolving digital era, collaboration between academia and industry has become increasingly important. Through ICCSCI 2026, we hope to encourage innovative ideas, impactful research, and sustainable technological solutions that can contribute positively to society.

I would like to express our sincere appreciation to all keynote speakers, authors, reviewers, session chairs, participants, and organizing committee members for their dedication and support in making this conference possible. My special gratitude also goes to Bina Nusantara University and all partners who continuously support ICCSCI throughout the years. To all participants, I wish you a productive, inspiring, and enjoyable conference experience. May ICCSCI 2026 become a valuable opportunity to build new collaborations, exchange knowledge, and create lasting academic connections.

Thank you, and I look forward to welcoming you again at ICCSCI 2026.

Sincerely yours,

Prof. Derwin Suhartono
Dean – School of Computer Science
Bina Nusantara University, Jakarta, Indonesia

Welcoming Remark

Vice Rector - Research and Technology Transfer **BINUS University**



Dear esteemed participants, colleagues, and guests,

To the honorable:

- Dr. Nelly, Rector of BINUS University,
- Dr. Derwin Suhartono, Dean of School of Computer Science,
- Dr. Tri Juniarty, Campus Director of BINUS@Medan,

Vice Rectors, Directors, Deans, and esteemed Professors,

Our distinguished keynote speakers:

- Prof. Poltak Sihombing
- Jae-Hyeon Park, Ph.D.
- Prastudy Fauzi, Ph.D.

Distinguished guests, ladies, and gentlemen,

Good morning. Let us praise God the Almighty for His merciful blessings that allow us to gather here in high spirits for The 11th International Conference on Computer Science and Computational Intelligence (ICCSCI) 2026, a proud part of the BINUS Joint International Conference (BJIC) 2026.

This year, ICCSCI centers around a profound conviction: "Fostering Human-Centered and Inclusive Innovation for a Digitally Empowered Society."

We live in an era where AI and advanced data analytics are no longer just laboratory breakthroughs—they are active agents of change. However, technology without a human-centric soul can widen the gap of inequality. That is why our theme today is critical. We are not just discussing faster algorithms or massive datasets; we are talking about how computational intelligence can alleviate poverty, revolutionize inclusive education, transform healthcare, and serve as an indispensable ally in climate action. We build technology by humanity, for humanity.

As scientists and educators, our tech innovations must stand on a bedrock of truth. Therefore, in line with our deep commitment to academic integrity, ICCSCI 2026 strongly advocates for Open Data and Open Contributorship. We strongly encourage authors to share their datasets through public repositories and to clearly outline author contributions using the CRediT (Contributor Roles Taxonomy) format.

By embracing these practices, we ensure that our research is transparent, reproducible, and verifiable. This is the ultimate manifestation of our BINUS SPIRIT values—where Integrity meets Innovation and Teamwork to foster true scientific progress.

Ladies and gentlemen,

It is highly fitting that we host this conference at BINUS@Medan. North Sumatra has historically been a vibrant crossroads of cultures, trade, and ideas along the Malacca Strait. Today, BINUS@Medan translates that heritage into the digital age.

As a dynamic campus in this western hub of Indonesia, BINUS@Medan is strategically positioned to nurture localized, high-impact innovations.

Welcoming Remark

Vice Rector - Research and Technology Transfer **BINUS University**



The campus serves as a fertile ground for nurturing digital talent and technopreneurship, bridging global computer science advancements with the unique socio-economic landscape of Sumatra. Through its modern ecosystem, BINUS@Medan ensures that the spark of innovation we discuss today will directly empower local and global communities alike.

We are honored to have prominent scholars and professionals from diverse backgrounds, including our colleagues joining us from Germany and Indonesia, to bridge global perspectives. This collaboration reflects BINUS University's bold commitment as a world-class university to generate high-impact research toward our BINUS Vision 2035—to continuously foster and empower society.

Before I conclude, I would like to express my deepest gratitude to Dr. Mochammad Haldi Widianto as the Conference Chair, along with the entire organizing committee, for their tireless dedication. My appreciation also goes to all presenters, reviewers, track directors, participants, and sponsors who have made this event possible.

I wish you all an inspiring, collaborative, and impactful conference.

Thank you very much. Horas!

Medan, North Sumatra, 18 June 2026

Prof. Dr. Juneman Abraham, S.Psi., M.Si.
Vice Rector – Research & Technology Transfer, BINUS University



Welcoming Remark
Rector - **BINUS University**

Dear esteemed participants, colleagues, and guests,

Distinguished Keynote Speakers,

- ·Prastudy Fauzi, Ph.D, Nanyang Technological University, Singapore,
- Prof. Poltak Sihombing, Universitas Sumatera Utara (USU), Indonesia,
- Jae-Hyeon Park, Ph.D., Bina Nusantara University, Indonesia,

Fellow Researchers, Academics, and Conference Participants,Ladies and Gentlemen,

Good morning, and a warm welcome to BINUS @Medan, Delipark Mall, Medan, North Sumatera.

It is a great honour and privilege for me, on behalf of BINUS University, to welcome you to the 11th International Conference on Computer Science and Computational Intelligence, ICCSCI 2026. For more than ten years, ICCSCI has served as an international forum, bringing together scholars, researchers, and professionals from around the world to share ideas, advance knowledge, and build meaningful collaboration in Computer Science, Computational Intelligence, and Information Technology. Initiated by the School of Computer Science at BINUS University, this conference reflects our long-standing commitment to research excellence and to knowledge that creates real impact for society.

This year's theme, "Fostering Human-Centered and Inclusive Innovation for a Digitally Empowered Society," speaks directly to the values that BINUS University holds at its core. As we mark 45 years of BINUS's journey as a knowledge institution, this theme reaffirms our belief that technology must be built around human dignity, accessibility, and the shared aspiration for a more equitable world. This is the spirit that guides BINUS's vision: "A World-class university fostering and empowering the society in building and serving the nation." Through platforms like ICCSCI, we aim to contribute not only to academic knowledge, but to the broader mission of building a digitally inclusive and globally connected society.



This year, we are privileged to welcome three distinguished keynote speakers whose expertise covers some of the most important areas where technology and society meet today: privacy-preserving systems in AI and blockchain, human-centered governance of IoT and robotics, and AI-driven approaches to chemical detection and signal analysis. Their insights will deepen our collective understanding and open the kind of meaningful conversations that move both science and society forward.

To all paper contributors, peer reviewers, organizing committee members, co-host institutions, and participants who have traveled from near and far to be here: your dedication and scholarly commitment are what make ICCSCI a conference of true international standing. We are deeply grateful for each of you.

I wish everyone two productive days of rich discussion, strong collaboration, and new connections that extend well beyond this conference. May this 11th edition of ICCSCI mark a milestone, not only in the history of this conference, but in the development of technology that truly serves all of humanity.

Regards,
Rector of BINUS University

Keynote **Abstract**



Jae-Hyeon Park, Ph.D.

Blina Nusantara University, Indonesia

Abstract

Integrated Approaches to Chemical Detection: Spectroscopy, Signal Analysis, and AI

This seminar presents advanced chemical detection technologies using hyperspectral spectroscopy, signal processing, and artificial intelligence. It introduces long-range FTIR and Raman spectroscopy systems for non-contact detection of hazardous chemical agents.

The seminar covers hyperspectral signal processing techniques, including noise reduction, spectral correction, feature extraction, and background removal, along with detection algorithms such as CC, ASD, OSP, NMF, SVM, and DNN. It also highlights gas visualization, concentration reconstruction, and AI-based detection methods for real-time chemical monitoring.

This seminar is intended for students, researchers, and professionals interested in spectroscopy, AI, signal processing, and advanced chemical sensing systems.

Keynote **Abstract**



Prof. Poltak Sihombing, Ph.D.

Universitas Sumatera Utara (USU), Indonesia

Abstract

Super Smart Society for All: Human-Centered Governance of IoT, AI, and Robotics

Society 5.0 is the concept of a "super-intelligent society" first proposed by Japanese government in 2016. Essentially, this concept places humans as the primary component, not technology. The convergence of Internet of Things (IoT), Artificial Intelligence (AI), and robotics constitutes the technological foundation of Society 5.0, that aims to integrate cyberspace and physical space to resolve social challenges while advancing economic development. This paper examines human-centered governance frameworks necessary to ensure that these technologies serve all members of society. Through comparative analysis of the Organization for Economic Co-operation and Development (OECD) AI Principles, the European Union AI Act, and Japan's Society 5.0 policy architecture, we identify core tensions between innovation, risk mitigation, and inclusion. We propose a layered implementation model addressing governance challenges across device, data, algorithmic, and societal layers. The analysis concludes that effective governance requires moving beyond entity-based personhood toward functional accountability mechanisms that prioritize human dignity, autonomy, and equitable access. The findings contribute to ongoing policy discourse on establishing trustworthy and inclusive socio-technical systems. We also recommend prioritizing human dignity in every decision in the era of the Super-Intelligent Society. The convergence of IoT, AI, and robotics technologies is within our grasp, but the responsibility rests with us, and it's all human-centered.

Keynote **Abstract**



Prastudy Fauzi, Ph.D.

Nanyang Technological University, Singapore

Abstract

Privacy-Preserving Technologies in Practice: Securing AI, Blockchain, and Electronic Voting Systems

This presentation explores five notable privacy-preserving technologies: differential privacy, secure multiparty computation, fully homomorphic encryption, zero-knowledge proofs, and trusted execution environment. We will also discuss other interesting technologies, such as private set intersection and searchable symmetric encryption.

These technologies are used to preserve individuals' privacy while allowing maximum value and insight to be extracted from available data. In particular, we will look at how these technologies can be used to enhance security & privacy in machine learning, blockchain, and electronic voting applications.

General **Rundown**

Zoom link for online participant :

https://binus.zoom.us/j/93422345876
Meeting ID: 934 2234 5876

Day 1 : Thursday, June 18th 2026

Time	Activity	Room
08.00-09.00	Registration	Xperience Space, 1st Fl
09.00-09.05	Welcoming Ceremony	Xperience Space, 1st Fl
09.05-09.10	Opening Speech : Chairman ICCSCI 2026	Xperience Space, 1st Fl
09.10-09.15	Opening Speech : Dean - School of Computer Science, BINUS University	Xperience Space, 1st Fl
09.15-09.20	Opening Speech : Binus @ Medan Campus Director	Xperience Space, 1st Fl
09.20-09.25	Opening Speech : Vice Rector - Research & Technology Transfer	Xperience Space, 1st Fl
09.25-09.30	Opening Speech : Rector BINUS University	Xperience Space, 1st Fl
09.30-09.35	Performance : Stamanara	Xperience Space, 1st Fl
09.40-10.10	Keynote Speaker - Prastudy Fauzi, Ph.D "Privacy Preserving Technologies in Practice: Securing AI, Blockchain and Electronic Voting Systems"	Xperience Space, 1st Fl
10.10-10.40	Keynote Speaker - Prof. Poltak Sihombing, Ph.D "Super Smart Society for All: Human-Centered Governance of IoT, AI, and Robotics"	Xperience Space, 1st Fl
10.40-10.55	Q&A Session	Xperience Space, 1st Fl
10.55-11.05	Coffee Break	Xperience Space, 1st Fl

General **Rundown**

Zoom link for online participant :

https://binus.zoom.us/j/93422345876
Meeting ID: 934 2234 5876

Day 1 : Thursday, June 18th 2026

Time	Activity	Room
11.05-11.35	Keynote Speaker - Jae Hyeon Park, Ph.D. "Integrated Approaches to Chemical Detection: Spectroscopy, Signal Analysis, and AI"	Xperience Space, 1st Fl
11.35-11.50	Q&A Session	Xperience Space, 1st Fl
11.50-12.00	Photo Session	Xperience Space, 1st Fl
12.00-13.00	Lunch	Lunch Room
13.00-15.00	Parallel Session 1 - Paper Presentation	Onsite (Xperience Room) and Online (Zoom Link attached)
15.00-15.30	Coffee Break	LKC 2nd Fl
15.30-17.30	Parallel Session 2 - Paper Presentation	Onsite (Xperience space) and Online (Zoom Link attached)
17.30-17.45	Closing Day 1	Xperience Space, 1st Fl

General **Rundown**

Zoom link for online participant :

https://binus.zoom.us/j/93422345876
Meeting ID: 934 2234 5876

Day 2 : Friday, June 19th 2026

Time	Activity	Room
09.00-09.30	Re-Registration	Xperience Space, 1st Fl
09.30-10.00	Seminar : Jae Hyeon Park, Ph.D. "Autonomous control of robots using AI - Jaehyon Park"	Xperience Space, 1st Fl
10.00-11.45	Parallel Session 3 - Paper Presentation	R.205 and Online (Zoom Link attached)
10.30-11.45	Workshop : Dr Adli "AI Workflow Automation with ImpresivAI"	Xperience Space, 1st Fl
11.45-13.00	Lunch	Xperience Space, 1st Fl
13.00-15.00	Parallel Session 4 - Paper Presentation	Onsite and Online (Zoom Link attached)
13.00-15.00	Video Reels Competition Presentation	R. 206
15.00-15.30	Coffee Break	LKC 2nd Fl
15.30-16.30	Closing Ceremony and Awarding Session	Xperience Space, 1st Fl
16.30-16.45	Closing Day 2	Xperience Space, 1st Fl

Parallel **Session Rundown**

&

Paper **Presentation Abstract**

Onsite Paralel Session 1 Rundown - BINUS @ Medan

Parallel Session 1				
Room		**Xperience Room - 1st Floor, BINUS @ Medan Main Campus**		
Session Chair		**Indra Hartarto Tambunan, Ph.D.**		
Day / Time		**Thursday, June 18th 2026 / 13.00-15.00 WIB (GMT +7)**		
No	**Paper ID**	**Author**	**Presenter**	**Title**
1	498	Hernawati Samosir (Institut Teknologi Del)* Samuel Sibuea (Institut Teknologi Del) Yoseplin Hutauruk (Institut Teknologi Del) Novrael Marbun (Institut Teknologi Del) Michael Joseph Christian Situ (Institut Teknologi Del)	Hernawati Susanti Samosir	Mobile Application for Strawberry Plant Disease Classification Using CNN-Based Leaf Image Analysis
2	928	Jayalakshmi Vaithiyanathan (Vels Institute of Science Technology and Advanced Studies,Chennai)*	Dr. Jayalakshmi V	SAD-HSA: A Secure Framework for Detecting Malicious Client Updates in Federated Learning through Statistical Anomaly Detection and Hybrid Aggregation
3	84	Sutarman Sutarman (Universitas Sumatera Utara)* Maya Silvi Lydia (Universitas Sumatera Utara) Syahriol Sitorus (Universitas Sumatera Utara) Tito Afwi Senda (Universitas Sumatera Utara) Adli Abdillah Nababan (Bina Nusantara University)	Sutarman	A Unified Benchmark of Interpretable Regression Models under Nonlinear Embeddings for High-Dimensional Mixed-Type Intelligent Systems

4	274	Linh Thuy Thi Pham (Can Tho University of technology) Hat Nguyen Le (Can Tho University) Thai T. Vo Nguyen (FPT University) Bao Gia Tran (FPT University) Ngoc Huynh Pham (Can Tho University) Thanh-Hai Nguyen (Can Tho University)*	Bao Gia Tran	Identification of Key Microbial Biomarkers for Colorectal Disease Prediction through Univariate Feature Selection
5	684	Arlinta Christy Barus (Institut Teknologi Del) Humasak Simanjuntak (Institut Teknologi Del) Lusye Triksi Pasaribu (Institut Teknologi Del) Hedrin S. Sitorus (Institut Teknologi Del) Greyssenly Ester Lya Simanjuntak (Institut Teknologi Del) Samuel Situmeang (Institut Teknologi Del)*	Arlinta Christy Barus	Architectural Refactoring and Container-Based Orchestration for DiTenun Platform

Onsite Paralel Session 3 Rundown - BINUS @ Medan

Parallel Session 3				
Room		**Xperience Room - 1st Floor, BINUS @ Medan Main Campus**		
Session Chair		**Rosni Lumbantoruan, M.ISD., Ph.D.**		
Day / Time		**Friday, June 19th 2026 / 10.00-11.45 WIB (GMT +7)**		
No	**Paper ID**	**Author**	**Presenter**	**Title**
1	685	Samuel Situmeang (Institut Teknologi Del)* Arlinta Christy Barus (Institut Teknologi Del) Humasak Simanjuntak (Institut Teknologi Del) Rachel Riris Rotua Damanik (Institut Teknologi Del) Reinhard Hottua Simbolon (Institut Teknologi Del) Reinhard Hottua Simbolon (Institut Teknologi Del)	Samuel Situmeang	Automated Coloring of Ulos Motif Designs via the UlosColor-GAN
2	430	Arnold Aguilar (De La Salle University)*	Arnold Aguilar	Target-Oriented Robust Optimization for Blockchain Governance Implementation Risk Management in the Philippines
3	504	Ranty Deviana Siahaan (Institut Teknologi Del)* Benhard Yudha (Institut Teknologi Del) Archico Darius Simpar Sembiring (Institut Teknologi Del) Rafael A. C. Manurung (Institut Teknologi Del)	Ranty Deviana Siahaan	User-Centered Gamification in E-Learning: A Preference-Based Personalization Approach using DPE and BWS

4	418	Cynthia Nababan (Institut Teknologi Del)* Fanny Clara Sinaga (Institut Teknologi Del) Yennisa Marianti Napitupulu (Institut Teknologi Del) Exaudi Siregar (Institut Teknologi Del)	Cynthia Deborah Nababan	Integrating Computational Thinking into Project-Based Learning: A Collaboration Model and Minimum Design Standard for Computer Science Students
5	677	Romindo Romindo (Universitas Pelita Harapan)* Kevin Bastian Sirait (Universitas Pelita Harapan) Jefri Junifer Pangaribuan (Bina Nusantara University) Valentino Riffan (Universitas Pelita Harapan) Chailine Garcia Wijaya (Universitas Pelita Harapan)	Kevin Bastian Sirait	Sentiment Analysis of Toxic YouTube Comments Using Naive Bayes and Confusion Matrix

Parallel Session 1 Rundown

No	Paper ID	Online		
		Room 1	https://binus.zoom.us/j/94617129923	
		Username	**946 1712 9923**	
		Day / Time	**Thursday, June 18th 2026 / 13.00-15.00 WIB (GMT +7)**	
		Session Chair	**Dr. R. Priya Anand**	
		Author	**Presenter Name**	**Title**
1	152	Jerry Wongso (BINUS University)* Adrian Tiongardi (BINUS University) Jeffrey Ali (BINUS University) Rezki Yunanda (BINUS University)	Jerry Wongso	Image-Based Tea Leaf Disease Classification Using SingleBackbone EfficientNet-B4 with Multi-Stage Gated Fusion
2	96	Grace Hadhinata (Binus University)* Sidharta Sidharta (Binus University)	Grace Hadhinata	Cost-Sensitive Heterogeneous Ensemble Learning for E-commerce Customer Churn Prediction with Explainable AI
3	106	Catherine Winanda (Bina Nusantara University)* Andreas Wiranata (Bina Nusantara University) Azani Cempaka Sari (Bina Nusantara University)	Catherine Olivia Winanda	Phishing Website Detection Using a Stacking Ensemble Classifier of Random Forest, LightGBM, and CatBoost
4	115	M. Raffa Sutanto (BINUS University)* Radityo Rasendrya Putra Wibowo (BINUS University) Nadhifa Risya Aqilah (BINUS University) Nur Damayanti (BINUS University)	Radityo Rasendrya P. W.	AI-Driven Personalization in Smart Urban Digital Ecosystems: Examining Brand Awareness and Consumer Behavioral Response among Gen Z in Jabodetabek Metropolitan Area

5	143	Thi Hong Yen Nguyen (Vinh Long University of Technology Education) Thanh-Hai Nguyen (Can Tho University)* Minh Thanh Ta (Vinh Long University of Technology Education) Loc Vinh Cu (Can Tho University)	Thi Hong Yen Nguyen	Enhancing Image Steganography using GAN with Multi-Scale Channel Attention and Adaptive Instance Normalization
6	147	Jason Sulistyawan (Bina Nusantara University)* Mahmud Isnan (Bina Nusantara University) Bens Pardamean (Bina Nusantara University)	Jason Sebastian Sulistyawan	Machine Learning Analysis of Longitudinal Biomarker Trajectories for Post-Stroke Recovery Prediction

No	**Paper ID**	**Online**		
		Room 2	https://binus.zoom.us/j/99897105452	
		Username	**998 9710 5452**	
		Day / Time	**Thursday, June 18th 2026 / 13.00-15.00 WIB (GMT +7)**	
		Session Chair	**Dr. Hady Pranoto**	
		Author	**Presenter Name**	**Title**
1	157	Nhan Luu (Can Tho University)* Nhan Luu (Can Tho University)	Nhan Trong Luu	Quantifying How Training Gradient Sparsity Affect Spiking Neural Network Accuracy And Robustness
2	180	Angelyn Gunawan (Bina Nusantara University)* Wina Permana Sari (Bina Nusantara University)	Angelyn Gunawan	Comparison of Stacked Ensemble Machine Learning for In-Season Maize Yield Forecasting in Ukraine Using Satellite and Climate Data

3	189	Andrea Stevens Karnyoto (Bina Nusantara University)* Claudia Rachel Wijaya (Bina Nusantara University) Harsawardana Harsawardana (Institut Pertanian Stiper) Wiwin Dyah Ully Parwati (Institut Pertanian Stiper) Muchammad Ali Nur Secha (Institut Pertanian Stiper) Taufik Yusuf (Institut Pertanian Stiper) Yasin Addafi (Institut Pertanian Stiper) Najia Al Umri (Bina Nusantara University) Bens Pardamean (Bina Nusantara University)	Andrea Stevens Karnyoto	Transfer Learning for Melon Plants Leaf Rust Symptoms Classification
4	192	Eddy Tungadi (Politeknik Negeri Ujung Pandang)* M. Ashar Alhakim (Politeknik Negeri Ujung Pandang) Tantri Indrabulan (Politeknik Negeri Ujung Pandang) Andrea Stevens Karnyoto (Bina Nusantara University) Bens Pardamean (Bina Nusantara University) Najia Al Umri (Bina Nusantara University)	Eddy Tungadi	Developing AI-Based Solutions for Indonesian Freshwater Fish Breeding: The Case of Tilapia Seed Calculation
5	200	Yanhong Gong (Guangdong University of Science and Technology) Anton Louise De Ocampo (Batangas State University)*	Yanhong Gong	Reliability Analysis of Landing Gear Retraction and Extension Actuators Based on Wavelet Threshold Denoising and Gaussian Process Regression Algorithm

6	231	Rafael Yudianto (Bina Nusantara University) Rojali Rojali (Bina Nusantara University) Franz Junior (Bina Nusantara University)*	Rafael Yudianto	Enhancing Wordlist-based DGA Detection via N-Gram Feature Fusion and Linguistic Statistical Analysis in Ensemble Learning

No	Paper ID	Online		
		Room 3	https://binus.zoom.us/j/95209485541	
		Username	**952 0948 5541**	
		Day / Time	**Thursday, June 18th 2026 / 13.00-15.00 WIB (GMT +7)**	
		Session Chair	**Dr. Jayalakshmi Vaithiyanathan**	
		Author	**Presenter Name**	**Title**
1	212	IRMA SYARLINA CHE ILIAS (UNIVERSITI KUALA LUMPUR)* Muslihah Wook (National Defence University of Malaysia) Norazimah Awang (UNIVERSITI KUALA LUMPUR) Muhammad Ezra Ismail (UNIVERSITI KUALA LUMPUR)	IRMA SYARLINA CHE ILIAS	A Decade of Augmented Reality Integration in Malaysian TVET: Adoption, Practices, and Challenges
2	296	Bustami Bustami (Universitas Syiah Kuala)* Yuwaldi Away (Universitas Syiah Kuala) Husni Husin (Universitas Syiah Kuala) Asri Asri (Universitas Malikussaleh) Muhammad Fikry (Universitas Malikussaleh)	Bustami	Automated Personal Protective Equipment Detection Using Convolutional Neural Networks for Laboratory Safety Monitoring
3	302	Daniel Wahyudi (Binus University)* Andry Chowanda (Binus University)	Daniel Kevin Jhon Feko Wahyudi	Constructing a Performance Report System as a Profiling Tool to Measure Game Object Spawning Performance on Unreal Engine 5

4	543	Rudy Hartanto (Bina Nusantara University)* Francesco Sebastian (Bina Nusantara University) Anderies Anderies (Bina Nusantara University)	Rudy Hartanto	Adaptive SNR-Aware Zero-DCE For Domain Generalizable Low-Light Image Enhancement (LLIE)
5	658	Anastasia Putri Hartanti (BINUS University)* Gregorius Natanael Elwirehardja (BINUS University)	Anastasia Putri Hartanti	Detecting Indonesian Illegal Online Gambling Site Advertisements in Social Media Images: A Comparative Study
6	758	Daniel Kurniawan (Bina Nusantara University)* Puti Andam Suri (Bina Nusantara University) Sani Muhammad Isa (Bina Nusantara University)	Daniel Kurniawan	Application of AI for Recipe Recommendations Using Images of Food Ingredients

No	**Paper ID**	**Online**		
		Room 4	https://binus.zoom.us/j/98217380023	
		Username	**982 1738 0023**	
		Day / Time	**Thursday, June 18th 2026 / 13.00-15.00 WIB (GMT +7)**	
		Session Chair	**Dr. Kalaivani K.**	
		Author	**Presenter Name**	**Title**
1	817	Ercent Tannius (BINUS)* Jos Hwen Sim (BINUS) Nicholas Lee (BINUS) Jurike V. Moniaga (BINUS) Setiawan Joddy (BINUS)	Nicholas Lee	Segmentation of Liver and Tumor Regions in CT Scan Images Using U-Net

2	926	Balaji VISTAS (Vels Institute of Science Technology and Advanced Studies)*	S.Saravana Kumar	Adaptive AI-Based Models for Detecting and Mitigating Evolving Threats in Cloud-Hosted E-Commerce Platforms
4	937	Parameswari Ramachandran (Vels Institute of Science,Technology and Advanced Studies)* Jyothilakshmi K N (Vels Institute of Science,Technology and Advanced Studies)	Jyothilakshmi KN	Grey wolf–Firefly algorithm-based feature selection with Support Vector Machine for Pulmonary Tuberculosis detection from Chest X-ray images
5	938	Akila A (Vels Institute of Science, Technology and Advanced Studies)* Priya Dharshini V (Vels Institute of Science, Technology and Advanced Studies)	V. Priyadharshini	AI-Based Automated Data Quality Management System for Noise Detection in Numerical Data Pipelines
6	948	Nikitha V (VISTAS)*	Nikitha.V	Integrated Hybrid CNN–LSTM Framework for Multi-Omics-Based Prediction of Salinity-Responsive Genes in Plants

No	**Paper ID**	**Online**		
		Room 5	https://binus.zoom.us/j/94900232408	
		Username	**949 0023 2408**	
		Day / Time	**Thursday, June 18th 2026 / 13.00-15.00 WIB (GMT +7)**	
		Session Chair	**Dr. Aditya Kurniawan**	
		Author	**Presenter Name**	**Title**

1	31	Andrew Colin (Bina Nusantara University) Jason Tandiari (Bina Nusantara University)* Yehezkiel Ferdinand (Bina Nusantara University) Aditya Kurniawan (Bina Nusantara University) Franz Junior (Bina Nusantara University)	Franz Adeta Junior	Integration of Decentralized Identity (DID) and Digital Signature for Blockchain-Based Digital Copyright Storage
2	191	Andrea Stevens Karnyoto (Bina Nusantara University)* Meiton Boru (Universitas Nusa Cendana) Dony Martinus Sihotang (Universitas Nusa Cendana) Nelci Dessi Rumlaklak (Universitas Nusa Cendana) Bertha Selviana Djahi (Universitas Nusa Cendana) Najia Al Umri (Bina Nusantara University) Bens Pardamean (Bina Nusantara University)	Meiton Boru	Mapping Stunting Prevalence: A Digital Geospatial Approach for Sustainable Health Monitoring in Nusa Tenggara Timur
3	273	Benedictus Lukito (Universitas Bina Nusantara)* Jason Loderick (Universitas Bina Nusantara) Rezki Yunanda (Universitas Bina Nusantara)	Rezki Yunanda	Data-Driven Volcanic Disaster Mitigation in Indonesia
4	415	Marvelous Nathanael (Bina Nusantara University)* Murniati Murniati (Bina Nusantara University)	Marvelous Nathanael	A Boosting MSME Growth with AI-Powered Social Media Analytics: A New Era for Guerrilla Marketing

5	597	Jordan Jordan (Binus University)* Rojali Rojali (Binus University)	Jordan	Cluster-Based Labeling and Feature Reduction for Money Laundering Detection Using Hybrid Models

No	Paper ID	Online		
		Room 6	https://binus.zoom.us/j/93955351000	
		Username	**939 5535 1000**	
		Day / Time	**Thursday, June 18th 2026 / 13.00-15.00 WIB (GMT +7)**	
		Session Chair	**Dr. P Sujatha**	
		Author	**Presenter Name**	**Title**
1	503	Stanislaus Wirianto (Bina Nusantara University) Sesilia Dewi (Bina Nusantara University) Louisa Halim (Bina Nusantara University) Nunung Qomariyah (Bina Nusantara University) Feri Setiawan (Bina Nusantara University)*	Sesilia Nadine Puspa Dewi	Hormonal Imbalances: Analyzing its Role on the Menstrual Cycle in Women
2	505	Alexander Gunawan (Binus University)* Roger Julianto Angryawan (Binus University) Fabian Surya Pramudya (Binus University)	Fabian Surya Pramudya	Indo-UnCRtainTS: Sentinel-1 and Sentinel-2 Fusion for Cloud Removal in Indonesian Urban Areas
3	507	Alexander Gunawan (Binus University)* KELVIN ALEXANDER BONG (Binus University) JASON PANGESTU (Binus University)	Kelvin Alexander Bong	Developer Discourse on Generative AI Programming Tools on Reddit: A Computational Grounded Theory Analysis
4	511	Alexander Gunawan (Binus University)* Kasandika Andariefli (Binus University) Vincentius Dewanto (Binus University) Jennifer Patricia (Binus University)	Kasandika Andariefli	Refining Spatial Detail in Sentinel-2 Wildfire Segmentation through a Data-Driven Post-Processing Pipeline

5	513	Alexander Gunawan (Binus University)* Wilson Gregory Pribadi (Binus University)	Wilson Gregory Pribadi	Machine Learning for Indonesian Stock Market Index Forecasting: The Role of Macroeconomic Indicators
6	520	Alexander Gunawan (Binus University)* Teuku Rusydi Khairi (Binus University) Wawan Hermawan (Padjajaran University)	Teuku Rusydi Khairi	Interpretable Machine Learning for Credit Rating Classification: Evidence from Indonesian Companies

Parallel Session 2 Rundown

No	Paper ID	Online		
		Room 1	https://binus.zoom.us/j/94617129923	
		Username	**946 1712 9923**	
		Day / Time	**Thursday, June 18th 2026 / 15.30-17.30 WIB (GMT +7)**	
		Session Chair	**Fiqri Ramadhan Tambunan, S.Kom., M.Kom**	
		Author	**Presenter Name**	**Title**
1	249	Idayati Mazlan (Universiti Teknologi MARA)*	Idayati Mazlan	Personalized Mental Health Interventions: A Comparative Study of Collaborative Filtering Approaches Using DASS-21 Data
2	252	Louis Ricardo (Binus University)* Jhoner Lou (Binus University) Bakti Jabar (Binus University)	Louis Ricardo	Hybrid SARIMA-LSTM Architecture for XAG/USD Price Forecasting based on Historical Trends and Technical Indicators
3	271	Ferre Deka Santosa (Bina Nusantara University)* Raymond Willen Kristopo (Bina Nusantara University) Dennise Adrianto (Bina Nusantara University)	Ferre Deka Santosa	Comparative Sentiment Analysis of the Terra (LUNA) and FTX Crises Using BERTweet, DistilBERT, and FinBERT
4	277	Nicholas Huang (Universitas Bina Nusantara)* Rezki Yunanda (Universitas Bina Nusantara)	Nicholas Huang	Enhancing Occupational Safety and Health Information Retrieval Using a RAG-Powered Question Answering System on NIOSH Guideline Documents
5	284	Immanuela Puspasari Saputro (Bina Nusantara University)* Muhammad Rafiq Fajar (Bina Nusantara University)	Muhammad Rafiq Fajar	Web Based AI for Running Warm-Up and Cool-Down Correction Using Blaze Pose and Joint-Angle Analysis

6	292	Muhammad Rasyak (Bina Nusantara University)* Mahmud Isnan (Bina Nusantara University) Bens Pardamean (Bina Nusantara University)	Muhammad Rezki Rasyak	A Comparative Study of Deep Learning and Machine Learning Approaches for Viral Transcriptomic Classification and Alcohol-Associated Hepatocellular Carcinoma
7	315	Shanaa Salsabil Erawan (Bina Nusantara University)* Najwa Zukhrufa (Bina Nusantara University) Celine Yovela (Bina Nusantara University) Azani Cempaka Sari (Bina Nusantara University) Nicholaus Hendrik Jeremy (Bina Nusantara University)	Celine Yovela	Predicting Jakarta's Air Quality Index Through SMOTE-Augmented Machine Learning: A Multi-Model Comparative Study on Imbalanced ISPU Data

No	**Paper ID**	**Online**		
		Room 2	https://binus.zoom.us/j/99897105452	
		Username	**998 9710 5452**	
		Day / Time	**Thursday, June 18th 2026 / 15.30-17.30 WIB (GMT +7)**	
		Session Chair	**Arya Krisna Putra, S.Kom., M.Kom**	
		Author	**Presenter Name**	**Title**
1	635	CORNELIUS Joe (Bina Nusantara University) Joel Christopher (Bina Nusantara University) Sulistyo Triutomo (Bina Nusantara University)* Rani Puspita (Bina Nusantara University)	Joel Christopher Purnomo	Data Mining for Discovering Consumer Behavioral Patterns in E-Commerce to Inform Marketing Strategies

2	638	Jasmine Chandra (Bina Nusantara University)* Meilani Meilani (Bina Nusantara University)	Jasmine Fidelia Chandra	Algorithmic Curation and Local Content Discoverability: A Survey-Based Assessment of Netflix's Recommendation Interface in Indonesia
3	662	Luthfillah Basya (Bina Nusantara University)* Murniati Murniati (Bina Nusantara University)	Luthfillah Basya	The Impact of Social Media Analytics Capability on Entrepreneurial Performance in the Digital Era: The Role of Engagement and Trust
4	722	Aisyah At Thahirah (Bina Nusantara University)* Suharjito - (Bina Nusantara University)	Aisyah Nur Fajri At Thahirah	Optimized Hybrid Neural-Statistical Architecture for Time-Series Forecasting: An Adaptive Multicategory Evaluation
5	725	Seline Loewel (Bina Nusantara University) Gareth Aurelius Harrison (Bina Nusantara University) Alleina Felisa Widjaya (Bina Nusantara University) Belvareia Reva Utamandarya (Bina Nusantara University) Benedictus Rahardjo (Bina Nusantara University)* Evaristus Didik Madyatmadja (Bina Nusantara University)	Gareth Aurelius Harrison	Public Emotion Analysis on Social Media and Short-horizon Stock Market and Rupiah Exchange Rate Movements A Case Study of the 17+8 Protests
6	726	Winston Khogres (Bina Nusantara University) Seline Loewel (Bina Nusantara University) Benedictus Rahardjo (Bina Nusantara University)*	Winston Khogres	Quantifying Collective Intelligence as a Volatility-moderating Signal in Short-horizon Digital Asset Markets

		Online	
		Room 3	https://binus.zoom.us/j/95209485541

No	Paper ID	Username	952 0948 5541	
		Day / Time	Thursday, June 18th 2026 / 15.30-17.30 WIB (GMT +7)	
		Session Chair	Nikita Masaling, S.Kom., M.Kom	
		Author	Presenter Name	Title
1	338	Laurentia Alyssa Castilani (BINUS University)* Muhammad Zarlis (BINUS University)	Laurentia Alyssa Castilani	The Role of Gemma 3 Fine-Tuning in Classifying Indonesian Fake News: A Large Language Model Approach
2	339	Kezya Tanudiredja (Bina Nusantara University)* Yanfi Yanfi (Bina Nusantara University)	Kezya Madeline Tanudiredja	Sales Trend Forecasting System for Snack Distribution using Extreme Gradient Boosting Algorithm
3	355	Richard Kurniawan (Binus)* Evan Somangkey (Binus) Felix Giovan (Binus) Eko Setyo (Binus)	Evan Somangkey	No Time to Waste: A Comparative Analysis of Deep Learning Models for Dynamic Garbage Classification
4	356	Emmanuel Widhiarto (Bina Nusantara University)* Abba Suganda Girsang (Bina Nusantara University)	Emmanuel Daniel Widhiarto	Long-Context Contrastive Adaptation for Automated Essay Scoring: Mitigating Information Blindness with BGE-M3
5	371	Fakhri Aulia (Bina Nusantara University) Emny Yossy (Bina Nusantara University)*	Fakhri Robi Aulia	Improving Implicit Feedback Recommendation with Matrix Factorization Model WARP Loss and SVD Prior Loss
6	383	Angelyn Gunawan (Bina Nusantara University)* Albert Verasius Dian Sano (Bina Nusantara University)	Angelyn Gunawan	Deep Transfer Learning for Predicting Corn Yield in Ukraine Using Remote Sensing and Climate Reanalysis Data
7	412	Stephanie Makmur (BINUS)* Grisella Zega (BINUS) Richie Hartono (BINUS) Bakti Jabar (BINUS)	Richie Hartono	Development of Rule-Weighted Ensemble Method for Masking Personally Identifiable Information in Financial Document Using Regex and NER

No	Paper ID	Online		
		Room 4	https://binus.zoom.us/j/98217380023	
		Username	**982 1738 0023**	
		Day / Time	**Thursday, June 18th 2026 / 15.30-17.30 WIB (GMT +7)**	
		Session Chair	**Anang Prasetyo, S.Kom, M.Kom.**	
		Author	**Presenter Name**	**Title**
1	730	Sheila Hadi (BINUS)* Fabian Surya Pramudya (BINUS) Muhammad Fikri Hasani (BINUS)	Sheila Atindriya Salsabila Hadi	Fuzzy TOPSIS for Digital-Readiness MSME Selection: A Multi-Criteria Decision-Making Method Employing Ministry of Manpower TKM Program Administrative Data
2	787	Anna Gracia Halim (Bina Nusantara University) Valerie Angelica Dharmo (Bina Nusantara University) Nadya Mulyarini (Bina Nusantara University) Benedictus Rahardjo (Bina Nusantara University)*	Anna Gracia Halim	Forecasting Indonesia's cashless transition: a SARIMA and MIDAS approach with cross-country benchmarking
3	822	Rowell Hernandez (Batangas State Univeristy)*	AUNG CHAN NYEIN	Application of Self-Attention-based Imputation for Time Series on Sparse Morbidity Data for Deep Learning Forecasting
4	825	Rosni Lumbantoruan (Institut Teknologi Del)* Katrina Tambunan (Institut Teknologi Del)	Katrina Arlyanti Tambunan	Out of Distribution Detection in Structured Data: A Systematic Review with Empahis on Autoencoder-Based Gatekeeper Methods
5	839	Rowell Hernandez (Batangas State Univeristy)*	Thein Kyaw Lwin	A Systematic Ablation of CBAM Depth in SqueezeNet for optimizing Cross-Domain Chickpea Disease Detection

6	210	Safitri Aprillia Putri (Bina Nusantara)* Viona Putri Salim (Bina Nusantara) Jevon Jeremy (Bina Nusantara)	Viona Putri Salim	Augmented Reality Card Game as a Learning Medium for Acehnese Cultural and Geographical Education
7	217	Jeremy Dewanto (Bina Nusantara University)* Kelvin Kelvin (Bina Nusantara University) Galih Pratama (Bina Nusantara University)	Jeremy Dewanto	Assessing Human Adaptability Under Time and Environmental Constraints Through Time-loop Mechanic in the Game “30 Seconds”

No	**Paper ID**	**Online**		
		Room 5	https://binus.zoom.us/j/94900232408	
		Username	**949 0023 2408**	
		Day / Time	**Thursday, June 18th 2026 / 15.30-17.30 WIB (GMT +7)**	
		Session Chair	**Samson Ndruru, S.Kom., M.Kom**	
		Author	**Presenter Name**	**Title**
1	264	Andhika Suryanto (Bina Nusantara University)* Reynata Praja Oktavian (Bina Nusantara University) Mirza Ramadhani (Bina Nusantara University)	ANDHIKA SURYANTO	Enhancing Cultural Awareness Through a Rhythm-Based Serious Game: A Study Using Indonesian Traditional Song
2	322	Crescentia Prasetya (Binus University)* Muhammad Fajar (Binus University)	Crescentia Karen Prasetya	THE EFFECT OF SOUND ON FEAR LEVELS IN THE 3D HORROR GAME FAMINE
3	365	Stephanie Vania Suwardi (Binus University)* Safitri Aprillia Putria (Binus University) Natalia Limantara (Binus University)	Stephanie Vania Suwardi	UI/UX DESIGN OF THE "EDUPHANY" APPLICATION FOR ENHANCING THE QUALITY OF ELEMENTARY SCHOOL MATHEMATICS OLYMPIAD EDUCATION IN INDONESIA

5	460	Raphael Ponda'ag (Bina Nusantara) Andrew Hanjaya (Bina Nusantara)* Dimas Ramdhan (Bina Nusantara)	Andrew Julian Hanjaya; Muhammad Raphael Ponda'ag	Evaluating a Serious Rhythm RPG Game for Enhancing Decision Making Through Mental Workload Measurement
6	779	Dr Muruganandham S K (Rajarajeswari college of engineering)*	Dr Muruganandham S K	CRISPR-AF: A Bio-Inspired Adaptive Firewall with Self-Curating Signature Memory for Network Intrusion Detection1

No	**Paper ID**	**Online**		
		Room 6	https://binus.zoom.us/j/93955351000	
		Username	**939 5535 1000**	
		Day / Time	**Thursday, June 18th 2026 / 15.30-17.30 WIB (GMT +7)**	
		Session Chair	**Dr. Dany Eka**	
		Author	**Presenter Name**	**Title**
1	930	Revathi S (VELS UNIVERSITY)*	Revathi S	Hybrid CNN-BiLSTM Based Secure Predictive Maintenance Framework for Accurate Fault Detection in IIoT Systems
2	936	Van Long Nguyen Huu (Can Tho University)* An Tran Cong (Can Tho University)	Van Long Nguyen Huu	Dual-Emotion Mining for Fake News Detection on Social Media
3	940	Rosniza Roslan (Universiti Teknologi MARA)* NURUL ANIS NAJIHAH AZMAN (Universiti Teknologi MARA) AHMAD HAFIZUDIN HAKIM DAWAM (Universiti Teknologi MARA) Noor Hasimah Ibrahim Teo (Universiti Teknologi MARA)	AHMAD HAFIZUDIN HAKIM DAWAM	Mobile Application for Real-Time Fresh Fruit Bunch Oil Palm Ripeness Classification

4	519	Cornelius Sarungu (Bina Nusantara University)* Ery Lestari Cahya Ningrum (Bina Nusantara University) Yosieka Wibawa (Bina Nusantara University) Zalsa Kusumaningtyas (Bina Nusantara University)	CORNELIUS MELLINO SARUNGU	Analysis of Factors Affecting The Use of Digital Libraries Through The Starry Application on The Reading Interest of Students at SMK Telkom Purwokerto

Parallel Session 3 Rundown

No	Paper ID	Online		
		Room 1	https://binus.zoom.us/j/97739221005	
		Username	**977 3922 1005**	
		Day / Time	**Friday, June 19th 2026 / 10.00-11.45 WIB (GMT +7)**	
		Session Chair	**Sidharta, S.Si., M.MT.**	
		Author	**Presenter Name**	**Title**
1	596	Muhammad Ferlanda (BINUS University)*	Muhammad Agrifian Ferlandaa	A Dual-Stream Deep Learning Architecture for Deforestation Land Cover Classification in Indonesia Using Optical and Infrared Satellite Imagery
2	600	Vincent Sebastian Gunawan (BINUS)* Christopher Alexander Chandra (BINUS) Said Achmad (BINUS)	Christopher Alexander Chandra	Robust Deepfake Image Detection Under Image Distortion Using Fine-Tuned CNNs and Stacking Ensembles
3	604	Christopher Purnomo (Binus University)* Melson Perawira (Binus University) Rani Puspita (Binus University)	Melson Mangala Perawira	A Genetic Algorithm-Based Optimization of Keyboard Layouts Using Indonesian Language
4	618	Devin Farrell (Bina Nusantara University)* Navizar Lubis (Bina Nusantara University) Stefanus Benhard (Bina Nusantara University) Derwin Suhartono (Bina Nusantara University)	Devin Farrell; Navizar Maestino Lubis	Fusion of Trajectory-Based and Behavioral Features for Session-Replay Bot Detection in Mouse Dynamics
5	622	Jordan Wijaya (Universitas Bina Nusantara)* Pualam Nusantara (Universitas Bina Nusantara)	Jordan Pascal Wijaya	Closing the Explainability Gap: Bridging ML and Deep Learning for Diabetes Prediction Using SHAP-Consensus Feature Selection

No	Paper ID	Online		
		Room 2	https://binus.zoom.us/j/97174313139	
		Username	**971 7431 3139**	
		Day / Time	**Friday, June 19th 2026 / 10.00-11.45 WIB (GMT +7)**	
		Session Chair	**Muhammad Rizki Nur Majiid, S.T., M.Kom**	
		Author	**Presenter Name**	**Title**
1	577	Adi Teguh Suprapto (Bina Nusantara University)* Panji Arisaputra (Bina Nusantara University)	Panji Arisaputra	Which Deep Learning Model Works Best for Emerging Market Stocks? Comparing LSTM and Transformer Models on 47 Indonesian Banks
2	579	Amrita Naik (Don Bosco College of Engineering)*	Dr Amrita Naik	Automated Interpretation of ECG to Predict Potential Cardiovascular Diseases
3	580	Elisabeth Nilasari (Sanata Dharma University)* Bambang Soelistijanto (Sanata Dharma University)	Elisabeth Nilasari	A Preliminary Study on Mitigating Upcoding in Indonesia's National Health Insurance using Multi-Agent Multi-Armed Bandits
4	583	Matthew Halim (BINUS University) Nathanael Widjaja (BINUS University)* Derwin Suhartono (BINUS University)	Matthew Nathanael Halim; Nathanael Hosea Widjaja	Gated Hybrid Architecture: Fusing Contextual Representations and Lexical Signals for Indonesian Hate Speech Classification

No	Paper ID	Online		
		Room 3	https://binus.zoom.us/j/99983659433	
		Username	**999 8365 9433**	
		Day / Time	**Friday, June 19th 2026 / 10.00-11.45 WIB (GMT +7)**	
		Session Chair	**Satria Santika, S.Stat., M.Kom**	
		Author	**Presenter Name**	**Title**

1	648	Chelsea Fakhira (Bina Nusantara University)* Latasha Ivy Khoe (Bina Nusantara University) Clarisse Xaviera Kurniaman (Bina Nusantara University) Dani Suandi (Bina Nusantara University)	Latasha Ivy Khoe	Cost sensitive ordinal classification for meningitis risk stratification considering clinical aspects to minimize false negatives
2	683	Mikhael Kevin Narendra Jayadharma (BINUS)* Jason Evansaputra Sudargo (BINUS) Prabowo Wahyu Sudarno (BINUS) Anang Prasetyo (BINUS) Fredy Purnomo (BINUS)	Mikhael Kevin Narendra Jayadharma	Explainability in Multi-Relational Opinion Spam Detection: The Trade-Offs of XAI-Guided Graph Pruning
3	689	Helena Sanjaya (Binus University)* Marvel Collin (Binus University) Bertrand Tjahyadi (Binus University) Derwin Suhartono (Binus University)	Helena Aurelia Sanjaya	Quantization Effects on Explainability in Indonesian Transformer-Based Sentiment Analysis
4	590	Daniel Amarta (Bina Nusantara University)* Nicholas Laksono (Bina Nusantara University) Stefanus Benhard (Bina Nusantara University) Derwin Suhartono (Bina Nusantara University)	Daniel Amarta	Improving the Accuracy of Facial Age Estimation Models for Social Media Platforms Using Hybrid CNN-Transformer Model

No	**Paper ID**	**Online**		
		Room 4	https://binus.zoom.us/j/94765940563	
		Username	**947 6594 0563**	
		Day / Time	**Friday, June 19th 2026 / 10.00-11.45 WIB (GMT +7)**	
		Session Chair	**Anderies, B.Eng., S.Kom., M.Kom**	
		Author	**Presenter Name**	**Title**

1	566	Jonathan Wijaya (BINUS University)* Suryadiputra Liawatimena (BINUS University)	Jonathan Wijaya	Evaluating YOLOv12 for Road Pothole Detection: A Cross-Dataset Performance Comparison
2	569	Suwandi The (Binus University)* Alif Tanoto (Binus University) Joshua Rawatan (Binus University) Feri Setiawan (Binus University) Nunung Qomariyah (Binus University)	Suwandi The	Improving YOLOv11 Model for Fabric defect detection using Knowledge Distillation approach
3	571	Leander Farrell Suryadi (BINUS) Kent Alber Fredson (BINUS)* Yithro Paulus Tjendra (BINUS) Abba Suganda Girsang (BINUS)	Leander Farrell Suryadi	Robust Detection of AI-Generated Text via Hybrid mBERT and Stylometric Feature Modeling
4	576	Pandu Dwi Luhur Pambudi (BINUS University)* Imelda Juliana BR. Sibarani (BINUS University) Cedric Sebastian (BINUS University)	Pandu Dwi Luhur Pambudi	Dynamic Pricing under Queueing Constraints in Electric Vehicle Charging Systems: A Comparative Analysis of Rule-Based and Reinforcement Learning Policies

No	**Paper ID**	**Online**		
		Room 5	https://binus.zoom.us/j/96616222142	
		Username	**966 1622 2142**	
		Day / Time	**Friday, June 19th 2026 / 10.00-11.45 WIB (GMT +7)**	
		Session Chair	**Abram Setyo , S.Kom., M.Cs**	
		Author	**Presenter Name**	**Title**
1	694	Elgan Kenlie (Universitas Bina Nusantara)* Anderies Anderies (Universitas Bina Nusantara)	Elgan Alviano Kenlie	Benchmarking LangChain4j Retrieval-Augmented Generation for Data Warehouse Querying on AdventureWorksDW2025

2	700	Tu-Nhu Le (Can Tho University) Van Long Nguyen Huu (Can Tho University) An Cong Tran (Can Tho University)*	An Cong Tran	A Personalized Wardrobe Management Application with Attention-Based Outfit Recommendation
3	732	Laura Marcella Pratama (Bina Nusantara University) Andreas Edward Putra Jatmiko (Bina Nusantara University) Simeon Yuda Prasetyo (Bina Nusantara University)*	Andreas Edward Putra Jatmiko	Voiced-Segment Demarcation and Translation-Invariant Scattering Networks for Parkinson's Disease Detection
4	764	Sitarani Safitri (Badan Riset dan Inovasi Nasional)* Akhmad Riqqi (Institut Teknologi Bandung) Albertus Deliar (Institut Teknologi Bandung)	Sitarani Safitri	A Probabilistic ECOC-SVM Framework for Multi-Class Land Suitability Assessment

No	**Paper ID**	**Online**		
		Room 6	https://binus.zoom.us/j/92946571906	
		Username	**929 4657 1906**	
		Day / Time	**Friday, June 19th 2026 / 10.00-11.45 WIB (GMT +7)**	
		Session Chair	**Andien Dwi Novika, S.Kom., M.Kom**	
		Author	**Presenter Name**	**Title**
1	780	Renaldy Fredyan (Bina Nusantara University)* Luis Frentzen Salim (National Taiwan University of Science and Technology) Khairunnisa (Universitas Muhammadiyah Bima)	Renaldy Fredyan	An Evaluation of Quantized Large Language Models and Entailment-Based Reformulation for Sentiment Classification

2	789	Shamna AR (Musaliar College of Engineering)*	Shamna AR	Attention-Constrained Mask-Aware Transformer For Intelligent Reconstruction Of Missing Biomedical semg Signals
3	819	Harry Santosa (Binus University) Garent Ecklesia (Binus University)* Derwin Suhartono (Binus University)	Harry Santosa	Evaluating Data Augmentation and Half-Precision Efficiency in Transformer-based Hate Speech Classification
4	872	Vittorio DinataVittorio Dinata (Bina Nusantara University)* Jefferson Louis (Bina Nusantara University) Wilbert Leung (Bina Nusantara University) Setiawan Joddy (Bina Nusantara University) Jurike Moniaga (Bina Nusantara University)	Jefferson Louis	Comparative Analysis of Feature Extraction Strategies for Atrial Fibrillation Detection Using ECG Signals and Machine Learning
5	923	Fathima Rumaiza SM (Vels Institute of Science, Technology & Advanced Studies)* Kamalakkannan S (Vels Institute of Science, Technology & Advanced Studies)	Fathima Rumaiza SM	A Multi-Objective Reinforcement Learning Framework for Antibiotic Selection Using Drug–Pathogen Interaction

Parallel Session 4 Rundown

No	Paper ID	Online		
		Room 1	https://binus.zoom.us/j/97739221005	
		Username	**977 3922 1005**	
		Day / Time	**Friday, June 19th 2026 / 13.00-15.00 WIB (GMT +7)**	
		Session Chair	**Said Achmad, S.Kom., M.Kom**	
		Author	**Presenter Name**	**Title**
1	431	Gilbert Rae (BINUS University)* Alberto Rusli (BINUS University) Rafael Kalyana (BINUS University) Michael Angelus (BINUS University)	Alberto Jecklyn Rusli	The Impact of AI Usage For Accounting Students Learning Process And Motivation
3	468	Violitta Yesmaya (BINUS)*	Violitta Yesmaya	Explainable Hybrid Machine Learning Model for Daily AQI Predicting in Jakarta
4	471	Karldritz Hanson (Bina Nusantara University)* Eikel Barus (Bina Nusantara University) Frederick Kuntjoro (Bina Nusantara University) Ivan Edbert (Bina Nusantara University)	Karldritz Farrel Hanson	Pareto-Based Evolutionary Reward Shaping for Deep Reinforcement Learning Traffic Control
5	472	Filbert Winch (Universitas Bina Nusantara)* Andry Chowanda (Universitas Bina Nusantara)	Filbert Christian Winch	A Comparative Study on EEG Classification using FFT for Emotion Analysis with GRU, BI-GRU, LSTM, BI-LSTM
6	474	Golda Salazar (Bina Nusantara University)* Adila Yuhanas (Bina Nusantara University) Karto Iskandar (Bina Nusantara University)	Golda Clarabella Salazar	Skin Tumor Classification using Convolutional Neural Network (CNN): Benign, Pre-Cancer, or Malignant
7	477	Reinardus Djafar (Binus University) Diandra Auwlia (Binus University) Abram Setyo (Binus University)*	Reinardus Djafar	Explainable Ensemble Learning for Water Potability Classification Using SHAP Based Feature Interpretation

No	Paper ID	Online		
		Room 2	https://binus.zoom.us/j/97174313139	
		Username	**971 7431 3139**	
		Day / Time	**Friday, June 19th 2026 / 13.00-15.00 WIB (GMT +7)**	
		Session Chair	**Muhamad Keenan Ario, S.Kom.,M.Kom**	
		Author	**Presenter Name**	**Title**
1	526	Pumudu Fernando (Informatics Institute of Technology)* Salinda Premadasa (Sabaragamuwa University of Sri Lanka)	Pumudu Fernando	Enhancing Fraction Learning in Primary Education through Curriculum-Aligned Game-Based Learning and Bloom's Digital Taxonomy
2	555	Cecilia Widayat (BINUS University)* Edi Purnomo Putra (BINUS University)	Cecilia Octavia Widayat	Integrating Formative Feedback and the Technology Acceptance Model: Enhancing Student Satisfaction Through Automated Grading in Gamification-Based Digital Learning Environments
3	612	Lana Aliyah Amran Amran (Universiti Teknologi Mara) Nurul Hidayah Mat Zain (Universiti Teknologi Mara)* Mohd Nabil Zulhemay (Universiti Teknologi Mara) Chiou Sheng Chew (Universiti Teknologi Mara) Anita Mohd Yasin (Universiti Teknologi Mara) Nor Azida Mohamed Noh (Universiti Teknologi Mara)	Lana Aliyah Mohd Amran	A Virtual Reality Game for Financial Literacy among Young Learners: An Empirical Study of User Engagement

4	631	Adi Teguh Suprapto (Bina Nusantara University)* Raizidan Rahman (Bina Nusantara University) Revi Arnan (Bina Nusantara University)	Adi Teguh Suprapto	The Role of Interactivity on Co-Developing, Influencing, Augmenting, and Mobilizing Behaviors: A Study of Roblox Mobile Gamers
5	674	Kayla Lukman (Binus University)* Jeslyn Djunaidy (Binus University) Muhamad Fajar (Binus University)	Kayla Cynthia Lukmana;Jeslyn Marianne Djunaidya	Santa-Pan Nusantara: Pilot Study on Educational Serious Game About Indonesian Local Food and Geography
6	687	Evlyn Manurung (Bina Nusantara University)* Muhammad Rachmadhani (Bina Nusantara University) Ayu Cahyani (Bina Nusantara University)	Evlyn Indah Manurung	Integrating Design Thinking and Technology Acceptance Model In Developing A Sustainable E-Commerce Prototype For Carbon Emission Reduction
7	782	Fadzlin Ahmadon (Universiti Teknologi MARA)* Muna Azuddin (International Islamic University Malaysia)	Muna Azuddin	Development and Validation of a Three-Layer Heuristic Evaluation Framework for Faith-Aligned Human-Computer Interaction

No	**Paper ID**	**Online**		
		Room 3	https://binus.zoom.us/j/99983659433	
		Username	**999 8365 9433**	
		Day / Time	**Friday, June 19th 2026 / 13.00-15.00 WIB (GMT +7)**	
		Session Chair	**Rilo Chandra, S.Stat., M.Kom**	
		Author	**Presenter Name**	**Title**
1	854	Vincent Wijaya (Binus University)* Evan Jonathan (Binus University) Radovieo Daffacetta (Binus University) Galih Pratama (Binus University)	Vincent Pho Wijaya	Increasing Awareness and Knowledge of Burnout Syndrome Through Sunside Cafe: A Serious Game Study

2	250	Vivian Alicia Santoso (Universitas Bina Nusantara)* Daniel Kurniawan Tjahjono (Universitas Bina Nusantara) Jovanda Carissa Joewono (Universitas Bina Nusantara)	Vivian Alicia Santoso	Determinants of Cybersecurity Awareness among Indonesian Generation Z: A PLS-SEM Approach
3	334	Dexter Krisnadhi (BINUS University)* Hosannia Hardiputri (BINUS University) Evotianus Evotianus (BINUS University) Muhammad Hasani (BINUS University)	Dexter Valerian Krisnadhi	Implementation of Retrieval-Augmented Generation on Indonesian AI-Driven Job Interview Practice Chatbot for Software Engineering
4	397	Riyanthi Sianturi (Institut Teknologi Del)*	Riyanthi Angrainy Sianturi	Software Reengineering of Village Administration Information System: Achieving Regulatory Compliance with Ministry of Home Affair Regulation No. 47 of 2016
5	416	Adhe Lingga Dewi (Binus University)* Anang Prasetyo (Binus University) Jayden Samuel Kurniawan (Binus University) Charles Nobel Wijaya Salim (Binus University)	Jayden Samuel Kurniawan, Charles Nobel Wijaya Salim	Regression-Based Theoretical Calibration of Multiple MQ Gas Sensors Using Datasheet Curves for IoT Air Quality Monitoring

No	**Paper ID**	**Online**		
		Room 4	https://binus.zoom.us/j/94765940563	
		Username	**947 6594 0563**	
		Day / Time	**Friday, June 19th 2026 / 13.00-15.00 WIB (GMT +7)**	
		Session Chair	**Galih Dea, S.Kom., M.Kom**	
		Author	**Presenter Name**	**Title**

1	51	Vincent Kwee (BINUS University) Michael Windarta (BINUS University) Vito Putra (BINUS University)*	Vincent Ferdinand Kwee; Michael Aurelio Windarta; Vito Tantra Putra	Memetic Algorithm Approach in Code Caves for Detection Evasion
2	74	Dedy Setiawan (Binus University)* Benfano Soewito (Binus University)	Dedy Setiawan	Integrating ISO 31000, ISO/IEC 27005, and NIST SP 800-30 for Risk Management in High-Classification Government Application
3	90	Priskila Tsunami (Bina Nusantara University)* Benfano Soewito (Bina Nusantara University) Ika Rachmawati (Bina Nusantara University)	Priskila Natalia Tsunami	Proactive Behavior-Based Ransomware Detection Using System Monitor Log Feature Extraction (Sysmon) And Random Forest Algorithm
4	153	Muhammad Fajar (Badan Pusat Statistik-Statistics Indonesia)* Rinda Nariswari (BINUS UNIVERSITY) Celine Ghefira Nurhaliza (Badan Pusat Statistik-Statistics Indonesia)	Rinda Nariswari	A New Alternative for Testing Brownian Motion in Time Series
5	234	Boby Siswanto (Bina Nusantara University)*	Boby Siswanto	Development of a Freshwater Fish Cultivation IoT System Based on Correlation Analysis to Identify the Most Influential Variables
6	247	Van Long Nguyen Huu (Can Tho University)* An Tran Cong (Can Tho University)	Van Long Nguyen Huu	An Explainable Web-Based Application for Rule-Driven Deep Learning Intrusion Detection System

No	**Paper ID**	**Online**	
		Room 5	https://binus.zoom.us/j/96616222142
		Username	**966 1622 2142**
		Day / Time	**Friday, June 19th 2026 / 13.00-15.00 WIB (GMT +7)**

		Session Chair	**Dr. Anita Rahayu**	
		Author	**Presenter Name**	**Title**
1	440	Uvindya Jayasundara (Sri Lanka Institute of Information Technology)* Yeshani Kumarage (Sri Lanka Institute of Information Technology) Hiruni Liyanagamage (Sri Lanka Institute of Information Technology) Kithmi Muthuweera (Sri Lanka Institute of Information Technology) Samadhi Rathnayake (Sri Lanka Institute of Information Technology) Amila Alexander (Sri Lanka Institute of Information Technology)	Yeshani Kumarage	Adaptive and Explainable Human-Robot Collaboration for Intelligent Sorting Systems Woody the Educational Robot
2	509	Alexander Gunawan (Binus University)* Alicia Taslim (Binus University) Junaedi Dede (Binus University) Muhamad Daffa Mennawi (Binus University)	Alexander Agung Santoso Gunawan	Fuzzy-PID Based Motion Control for a Mecanum-Wheeled Smart Trolley
3	518	Gintoro Gintoro (Universitas Bina Nusantara)*	Gintoro	Impact of Micro Services Architecture Design Patterns on Software Maintenance Cost
4	535	Yulius Prabowo (Univesitas Bina Nusantara)* Albert Wibisono (Universitas Atma Jaya Yogyakarta) Martinus Maslim (Universitas Atma Jaya Yogyakarta)	Yulius Denny	Absolute Zero Reasoning for Autonomous Penetration Testing: A Hybrid LLM–MCTS Framework with Action-Constrained Exploration
5	565	Nur Akmal Fahmi Nuri (FSKM, UiTM) Alya Geogiana Buja (FSKM, UiTM) Albin Lemuel Kushan (FSKM, UiTM)*	Nur Akmal Fahmi Nuri	Network performance of multiple VPN topologies under simulated DDoS attacks using GNS3

6	609	Muhammad Hariz Hakim Zakaria (FSKM, UiTM) Alya Geogiana Buja (FSKM, UiTM)* Raihah Aminuddin (FSKM, UiTM) Sulaiman Mahzan (FSKM, UiTM) Dziauddin Sharif (ACIS, UiTM) Anwar Farhan Zolkeplay (FSKM, UiTM)	Muhammad Hariz Hakim Zakaria	Blockchain-based Islamic family asset management system with a secured encrypted repository
7	614	Tengku Zafran Shah Tengku Iskandar Zulkarnain (FSKM, UiTM) Alya Geogiana Buja (FSKM, UiTM)* Mohd Hafifi Mohd Supir (FSKM, UiTM) Noor Afni Deraman (FSKM, UiTM) Nor Intan Shafini Nasaruddin (FSKM, UiTM) Nor Aimuni Md Rashid (FSKM, UiTM)	Tengku Zafran Shah Tengku Iskandar Zulkarnain	Development of Secure Credential Repository System Using Multifactor Authentication and Encryption

No	**Paper ID**	**Online**		
		Room 6	https://binus.zoom.us/j/92946571906	
		Username	**929 4657 1906**	
		Day / Time	**Friday, June 19th 2026 / 13.00-15.00 WIB (GMT +7)**	
		Session Chair	**Rico Halim, S.Kom., M.Kom**	
		Author	**Presenter Name**	**Title**

1	489	Darren Gunawan (Binus University International) Viona Djunaidy (Binus University International) Rafael Angelo Christianto (Binus University International) Nunung Nurul Qomariyah (Binus University International) Feri Setiawan (Binus University International)*	Viona Djunaidy	Advancing Emotion Recognition through Cross-Modal Fusion and Gender-Aware Learning
2	512	Slamet Riyanto (Bina Nusantara University)*	Slamet Riyanto	A Two-Step Machine Learning Approach for Clinker Free Lime Prediction toward Energy-Efficient Cement Kiln Operation
3	523	Vincent Ramaputra (Bina Nusantara University)* Setiawan Joddy (Bina Nusantara University)	Vincent Ramaputra	Evaluating the Impact of Quantization on Indonesian Question Answering Performance in Large Language Models
4	525	Andry Chowanda (Bina Nusantara University)* Mohamed Imran Bin Mohamed Ariff (UiTM Cawangan Perak)	Andry Chowanda	Social Media Based Depression Recognition Using BERT with Hyperparameters Optimisation
5	530	Prabowo Wahyu Sudarno (Bina Nusantara University)* Nur Sitha Afrilia (Bina Nusantara University) Kumara Pinasthika Dharaka (Bina Nusantara University) Nico Yonatan Wicaksana (Bina Nusantara University) Hermawan Honggo Widagdo (Bina Nusantara University)	Prabowo Wahyu Sudarno	Leveraging AI-Driven Data Analytics to Identify the Sanitation-Poverty Correlation via Comparative Clustering Analysis

6	563	Dhanushraaj Shanmuganandan (PES University)* BG Pranav (PES University) Suhas Chandan (PES University) Abhijit Shanbhag (PES University) Nitin Pujari (PES University)	Dhanush S	Controllable Identity Obfuscation via Multiple SVM Boundaries for Privacy-Preserving Face Dataset Generation

A Unified Benchmark of Interpretable Regression Models under Nonlinear Embeddings for High-Dimensional Mixed-Type Intelligent Systems

Sutarman Sutarman (Universitas Sumatera Utara)*
Maya Silvi Lydia (Universitas Sumatera Utara)
Syahriol Sitorus (Universitas Sumatera Utara)
Tito Afwi Senda (Universitas Sumatera Utara)
Adli Abdillah Nababan (Bina Nusantara University)

Abstract

Interpretable regression is essential in intelligent engineering systems that operate on high-dimensional data with mixed numerical, categorical, and ordinal features. Nonlinear dimensionality reduction is often used to address the curse of dimensionality, yet its influence on interpretable regression models remains insufficiently understood. This paper presents a unified and reproducible benchmarking protocol that decouples representation learning from regression modelling, enabling fair comparison across nonlinear latent spaces. Three embedding strategies—UMAP with Gower distance, standard autoencoders, and sparse orthogonal autoencoders—are evaluated in combination with interpretable regression models, including variants of Multivariate Adaptive Regression Splines (MARS) and Generalized Additive Models (GAM). Experiments on a real-world biomedical mixed-type dataset and multiple synthetic datasets show that embedding choice strongly affects downstream regression behaviour. GAM achieves lower prediction error on smooth latent representations, while MARS variants remain advantageous in computational efficiency and localized modelling.

Cost-Sensitive Heterogeneous Ensemble Learning for E-commerce Customer Churn Prediction with Explainable AI

Grace Hadhinata (Binus University)*
Sidharta Sidharta (Binus University)

Abstract

Customer loyalty is crucial for long-term business sustainability, making customer churn prediction a strategic priority for digital platforms to minimize revenue loss and improve customer retention. However, customer churn prediction remains challenging in the industry, particularly in e-commerce environments, due to imbalanced data distributions and unequal business costs associated with misclassification, especially false negative predictions. Although machine learning models have demonstrated strong predictive capabilities, most existing approaches focus on overall accuracy without considering the higher cost of misclassifying churn customers. Conventional models tend to bias predictions toward the majority non-churn class on imbalanced datasets, increasing false negative predictions that leading to higher customer reacquisition costs and potential revenue loss. This study proposes a cost-sensitive heterogeneous ensemble learning framework for predicting e-commerce customer churn. It integrates multiple classifiers with class weighted optimization to reduce false negative errors that could lead to increased revenue loss. To enhance transparency and support decision-making, explainable artificial intelligence (XAI) utilizing SHAP identifies key features that affect churn behavior. Experiments conducted on an e-commerce customer dataset demonstrate that the proposed framework significantly increases churn recall and outperforms the non-cost-sensitive model by effectively reducing the false negative (FN) rate while maintaining a high AUC of 0.9980. The XAI analysis provides actionable insights into customer habits and patterns, enabling the formulation of more efficient, targeted retention strategies. Overall, the results indicate that combining cost-sensitive ensemble learning with explainable AI yields an effective and interpretable solution for customer churn prediction, supporting more reliable and cost aware decision making for customer retention strategies, particularly in e-commerce.

Phishing Website Detection Using a Stacking Ensemble Classifier of Random Forest, LightGBM, and CatBoost

Catherine Winanda (Bina Nusantara University)*
Andreas Wiranata (Bina Nusantara University)
Azani Cempaka Sari (Bina Nusantara University)

Abstract

Phishing website attacks have become a complex and tough cyber threat to detect by conventional methods only. This work suggested a phishing website detection study, which is based on ensemble stacking model using Random Forest, LightGBM, CatBoost. The dataset used is a list of 84,773 URLs, both from the combination of Tranco for legitimate URLs and PhishTank for phishing URLs. Model performance was monitored using accuracy, precision, recall and F1 score metrics. In addition, an imbalanced approach of stress testing were used with class ratio 1:10 to simulate the real-world distribution of data, where legitimate class is very much more dominants than phishing class. The result of the experiment has shown that ensemble stacking model's accuracy rate is 96.3% and F1-score rate is 97%, and it shows stable and robust performance in unbalanced data conditions. Hence, these results indicate that the ensemble stacking approach is effective and can be applied to real-world website phishing detection systems.

AI-Driven Personalization in Smart Urban Digital Ecosystems: Examining Brand Awareness and Consumer Behavioral Response among Gen Z in Jabodetabek Metropolitan Area

M. Raffa Sutanto (BINUS University)*
Radityo Rasendrya Putra Wibowo (BINUS University)
Nadhifa Risya Aqilah (BINUS University)
Nur Damayanti (BINUS University)

Abstract

This study analyzes AI-driven personalization and social media marketing in a smart urban digital ecosystem and their relationships to brand awareness and purchase intention among Generation Z TikTok users living in the Jabodetabek area. In smart city environments, algorithmic platforms act as a data-dependent framework that shapes and modifies the consumption patterns of users. Using the Stimulus-Organism-Response (SOR) model, the study breaks down technological stimuli into advertising informativeness and perceived advertising personalization, leading to brand awareness as the organismic cognitive state, and to purchase intention as the behavioral response. Utilizing an online questionnaire, the study collected data from Gen Z (18+) TikTok users in the Jabodetabek metropolitan region, receiving 345 qualified responses, which were analyzed using the PLS-SEM method. The study found that social media advertising and AI-driven personalization positively impact brand awareness and purchase intention. Brand awareness is a strong indicator of purchase intention and advertising personalization. The structural model has a high degree of explanatory value and, in an AI-influenced urban platform context, explains 35.8% of the variation in brand awareness and 58.7% in purchase intention. The study revealed the impact of AI-driven advertising technology and the role of smart city infrastructure on the consumption behavior of urban residents. The study enhances the understanding of how social platform-embedded computational intelligence systems sustain urban digital economies and highlights the role of responsible and effective AI engagement strategies.

Enhancing Image Steganography using GAN with Multi-Scale Channel Attention and Adaptive Instance Normalization

Thi Hong Yen Nguyen (Vinh Long University of Technology Education)
Thanh-Hai Nguyen (Can Tho University)*
Minh Thanh Ta (Vinh Long University of Technology Education)
Loc Vinh Cu (Can Tho University)

Abstract

In today's digital landscape, safeguarding information is essential in fields such as finance, healthcare, and public services. Image steganography-embedding data within images-offers a practical solution for secure communication. This study compares two typical approaches: the traditional Least Significant Bit (LSB) technique and a deep learning–driven model based on Generative Adversarial Networks (GANs). LSB is straightforward and supports high payloads, but is easily exposed to visual artifacts and statistical attacks. In contrast, our GAN-based approach integrates a convolutional encoder–decoder–critic design with Multi-Scale Channel Attention (MSCA) and Adaptive Instance Normalization (AdaIN), allowing more subtle and resilient data embedding. Experiments show that although LSB performs well in capacity, the GAN model delivers superior visual quality, reconstruction fidelity, and robustness against steganalysis. These findings highlight the potential of deep learning methods as a strong and practical direction for secure information hiding in digital media.

Machine Learning Analysis of Longitudinal Biomarker Trajectories for Post-Stroke Recovery Prediction

Jason Sulistyawan (Bina Nusantara University)*
Mahmud Isnan (Bina Nusantara University)
Bens Pardamean (Bina Nusantara University)

Abstract

Predicting neurological recovery after acute ischemic stroke (AIS) remains a major clinical challenge. Although static admission biomarkers offer limited prognostic value, the longitudinal trajectory of inflammatory and hemostatic markers over the subacute period may provide richer predictive information. We analyzed a clinical trial dataset of 52 AIS patients who received oral DLBS1033 as adjunctive therapy. We took serial measurements of TNF-α, IL-6, D-Dimer, MMP-9, and carotid intima-media thickness (CIMT) on Days 0, 14, and 28. We engineered trajectory features, absolute change, rate of change, trapezoidal area under the curve (AUC), and variability per biomarker and combined them with baseline clinical data. Five machine learning classifiers were evaluated under repeated stratified 5-fold cross-validation across three feature sets of increasing temporal richness, plus a Spearman-guided top-7 feature selection step. Random Forest with the selected features achieved the best cross-validated AUC-ROC of 0.872 ± 0.122—a 16.6 percentage point improvement over static baseline approaches, with bootstrap confirmation of 0.962 (95 % CI: 0.909–0.994). Notably, 5 of the 7 selected features are trajectory derived, indicating that temporal dynamics provide additional prognostic information. The leading trajectory predictors were MMP-9 trajectory variance and 14-day delta, TNF-α AUC, and CIMT variance. Unsupervised clustering revealed two clinically distinct patient phenotypes that were separable by trajectory pattern,but not baseline scores. These findings demonstrate that longitudinal biomarker trajectories, when appropriately selected, substantially outperform static admission models for post-stroke recovery prediction.

Image-Based Tea Leaf Disease Classification Using SingleBackbone EfficientNet-B4 with Multi-Stage Gated Fusion

Jerry Wongso (BINUS University)*
Adrian Tiongardi (BINUS University)
Jeffrey Ali (BINUS University)
Rezki Yunanda (BINUS University)

Abstract

Tea leaf diseases threaten both the quantity and quality of tea production in Indonesia's plantation commodity, yet early prevention and accurate identification remain difficult due to manual inspection by field workers leading to a time-consuming and subjective process. Existing EfficientNet-based approaches mainly rely on the final backbone output, leaving rich discriminative information contained in intermediate layers not being utilized. To address this limitation, this study proposes an image-based tea leaf disease classification framework based on single-backbone EfficientNet-B4 enhanced with a Multi-Stage Gated Fusion mechanism. The model utilizes feature representations from multiple internal stages from S2 to S5, and splits into three variants such as, Two-Stage (S2 and S3), ThreeStage (S3 to S5), and Four-Stage (S2 to S5), classifying images into four target classes: Healthy, Algal Leaf Spot, Brown Blight, and Red Spot. Experiments were conducted on a merged publicly available dataset. The Four-Stage variant achieves the best performance with an accuracy of 97.23%, demonstrating that utilizing hierarchical multilevel representations within a single-backbone offers a more effective approach for tea leaf disease classification.

Quantifying How Training Gradient Sparsity Affect Spiking Neural Network Accuracy And Robustness

Nhan Luu (Can Tho University)*
Nhan Luu (Can Tho University)

Abstract

Spiking Neural Networks (SNNs) have recently received increasing attention in both computational neuroscience and artificial intelligence owing to their potential for energy-efficient computation and reduced memory requirements. Despite these advantages, improving adversarial robustness in SNNs (particularly for vision-based applications) remains an emerging and relatively underexplored research problem. Recent work has suggested that encouraging sparse gradients can act as a regularization mechanism to improve resistance against adversarial perturbations. In this study, we report an unexpected observation: under certain architectural configurations, SNNs inherently exhibit sparse gradients and can attain state-of-the-art adversarial defense performance without requiring any explicit regularization strategy. Further investigation reveals an inherent trade-off between robustness and generalization. Specifically, increased gradient sparsity enhances resistance to adversarial attacks but may reduce the model's generalization capability, whereas denser gradients tend to improve generalization while simultaneously increasing susceptibility to adversarial perturbations. These findings provide new perspectives on the role of gradient sparsity in the training dynamics of SNNs.

Early Stroke Prediction Using Hybrid Machine Learning Approach

Christopher Tan (Bina Nusantara University)
Egbert Hung (Bina Nusantara University)*
fondry cendranata (Bina Nusantara University)
Ika Dyah Agustia Rachmawati (Bina Nusantara University)

Abstract

Stroke is a life-threatening disease. Data collected by the WHO estimates that there are more than 12 million new strokes each year, and over 7 million people die from stroke annually. Several studies have been done to find the best method to predict stroke, with some even using physiological bio signals. However, the most effective way to predict stroke is by using laboratory data derived from blood tests; some research has already applied blood test results in their methodologies. In this study, four algorithms are used to find the best machine learning models and evaluate their capabilities to predict stroke: Random Forest, Gradient Boosting Machine, Logistic Regression, and Support Vector Machine. After testing, the results show that all four algorithms perform well, Random Forest and Gradient Boosting Machine have the best accuracy and Followed by LR and SVM. These percentages are influenced by the dataset containing far more non-stroke samples than stroke cases. The best results come from training and testing the algorithms with larger datasets. Ultimately, all four algorithms show strong potential in predicting stroke possibilities using blood test results.

Comparison of Stacked Ensemble Machine Learning for In-Season Maize Yield Forecasting in Ukraine Using Satellite and Climate Data

Angelyn Gunawan (Bina Nusantara University)*
Wina Permana Sari (Bina Nusantara University)

Abstract

Ukraine is the fourth-largest maize exporter in the world, yet the country is facing war which limits traditional ground-based crop surveys. Satellite-based analysis paired with machine learning may be the only viable alternative method; however, no in-depth study comparing stacking machine learning models and determining the earliest time yield can be forecast accurately exists specifically for Ukraine yet. This study develops and tests multiple stacking generalized ensemble models (Ridge regression + XGBoost, SVR + XGBoost and Ridge + SVR, with Ridge meta-learner) for annual in-season maize yield forecasting in Ukraine on oblast-level, and compares multiple time windows (March–May, March–June, March–July and March–August). Multi-source data comprise of remote sensing data, climate features and official national yield statistics from 2015–2024. Results show that stacking reduces RMSE and MAPE by up to 28.3% and 31.4%, respectively, relative to individual base models. The best performers were SVR + XGBoost in cross-validation (RMSE = 0.97 t/ha; MAPE = 12.9%) and Ridge + SVR in 2024 held-out data (RMSE = 0.64 t/ha; MAPE = 9.8%), and the most accurate forecasting time window is March–July, especially for oblasts in central and south-east Ukraine. All of the stacked models reached RMSE below 1 t/ha as early as three months before harvest by the end of May, proving that this satellite-based agricultural assessment system is capable of producing accurate early yield forecasts.

Transfer Learning for Melon Plants Leaf Rust Symptoms Classification

Andrea Stevens Karnyoto (Bina Nusantara University)*
Claudia Rachel Wijaya (Bina Nusantara University)
Harsawardana Harsawardana (Institut Pertanian Stiper)
Wiwin Dyah Ully Parwati (Institut Pertanian Stiper)
Muchammad Ali Nur Secha (Institut Pertanian Stiper)
Taufik Yusuf (Institut Pertanian Stiper)
Yasin Addafi (Institut Pertanian Stiper)
Najia Al Umri (Bina Nusantara University)
Bens Pardamean (Bina Nusantara University)

Abstract

In this research, we developed a method to detect leaf rust symptoms using transfer learning. Leaf rust is a plant disease caused by fungi from the Pucciniales group (formerly called Uredinales). Rust disease can impair plants' physical performance and reduce the quality of melon growth. It can also deplete starch reserves in the roots and branches. Eventually, the disease may kill the trees. We obtained the dataset by collecting leaf images from the Institut Pertanian Stiper (INSTIPER) greenhouse in Yogyakarta. We collected images of 81 healthy leaves and 82 leaves affected by rust symptoms. Each image was divided into nine parts using a 3x3 grid to increase the data size. After splitting, the number of healthy leaf images was 729, and those with rust symptoms were 738. We used K-fold cross-validation to split the dataset into 5 folds. Four parts were used for training and one for testing. Using K-fold provides a comprehensive performance metric. For the model, we used transfer learning models such as Xception, ResNet50V2, and EfficientNetV2S. Testing showed that EfficientNetV2S performed best. It achieved a Precision of 0.959, a Recall of 0.96, an F1-score of 0.9595, and an Accuracy of 0.958.

Developing AI-Based Solutions for Indonesian Freshwater Fish Breeding: The Case of Tilapia Seed Calculation

Eddy Tungadi (Politeknik Negeri Ujung Pandang)*
M. Ashar Alhakim (Politeknik Negeri Ujung Pandang)
Tantri Indrabulan (Politeknik Negeri Ujung Pandang)
Andrea Stevens Karnyoto (Bina Nusantara University)
Bens Pardamean (Bina Nusantara University)
Najia Al Umri (Bina Nusantara University)

Abstract

Tilapia (Oreochromis niloticus) is a widely consumed freshwater fish in Indonesia. Its cultivation is economically significant. Conventional methods for calculating fish seeds are time-consuming and prone to error. These issues present challenges for farmers. This study presents the development of an Android-based application for tilapia seed calculation. The application utilizes the YOLO (You Only Look Once) object detection algorithm. The dataset comprised 400 images, with 300 used for training and 100 for testing. Images were collected during the seeding process in Soppeng Regency. The trained YOLO model was converted to TFLite format and integrated into the application. Application functionality was evaluated using the Blackbox testing method. Results indicate that the application achieved 86% accuracy in tilapia seed calculation compared to manual data testing. The application offers a more efficient and accurate solution for tilapia farmers.

Reliability Analysis of Landing Gear Retraction and Extension Actuators Based on Wavelet Threshold Denoising and Gaussian Process Regression Algorithm

Yanhong Gong (Guangdong University of Science and Technology)
Anton Louise De Ocampo (Batangas State University)*

Abstract

The landing gear system is one of the key subsystems of an aircraft, directly affecting the safety of takeoff and landing. Due to the presence of noise interference and limited sample size in engineering data, this paper adopts the Monte Carlo (MC) method to conduct 1000 small-batch random samplings of random input parameters. The maximum stress of the landing gear retraction and extension actuator (LGREA) is obtained through the transient analysis module of ANSYS. The wavelet threshold denoising (WTD) method is used to denoise the 1000-sample output responses. The Gaussian process regression (GPR) model is trained by combining the constant kernel, radial basis function kernel RBF, and white noise kernel. Through large-scale sampling by the MC method, 1,000,000 stress means and variances predicted by the GPR model are obtained. Through statistical analysis, the reliability of the LGREA is calculated to be 99.92%. Compared with the reliability of 99.90% calculated by the finite element method, it not only ensures a certain prediction accuracy but also greatly improves the computational efficiency.

Enhancing Wordlist-based DGA Detection via N-Gram Feature Fusion and Linguistic Statistical Analysis in Ensemble Learning

Rafael Yudianto (Bina Nusantara University)
Rojali Rojali (Bina Nusantara University)
Franz Junior (Bina Nusantara University)*

Abstract

As Domain Generation Algorithms (DGA) evolve from random character series to wordlist-based formats that closely resemble actual human language phrases, the complexities of cybersecurity threats are also growing. It causes a notable decline in both traditional detection methods and transforms deep learning algorithms into something more complicated requiring high processing power and showing no explainable behavior. In this paper, we propose an efficient DGA detection method using a Feature Fusion strategy that combines N-Gram statistical features and linguistic statistical analysis as well as pronounceability score and vowel frequency/consonant frequency for the verification of obtained domains. The experiments used TRANCO and DGArchive an integrative dataset that contains 133,334 samples closing representations of network traffic in a ratio of 70:30. The results suggest that these 12 best features can enhance the F1 score significantly. The F1 score, which is used to see how well a model predicts Matsnu family passages, went from 8.81% in earlier tests of the LightGBM and XGBoost models to 80.53%. This research demonstrates that customized human feature engineering can surpass detection limitations in complex architectures, while simultaneously decreasing computational expenses and improving logical accountability.

Personalized Mental Health Interventions: A Comparative Study of Collaborative Filtering Approaches Using DASS-21 Data

Idayati Mazlan (Universiti Teknologi MARA)*

Abstract

Personalized mental health interventions provide essential benefits which enhance both the treatment of mental health cases and their overall support system. Recommender system techniques function as an effective method to provide personalized interventions because they utilize user characteristics and behavior patterns to identify suitable support solutions. The study evaluates three different methods which include a basic recommendation system and two versions of collaborative filtering which include profile matching. The evaluation uses data collected from the DASS-21 (Depression, Anxiety, and Stress Scale). The system evaluates method performance by measuring precision rates and recall rates and calculating F1-score results. The basic recommendation method serves as the baseline system while users experience better results through collaborative filtering which detects user similarities. The system performance improves through the addition of profile matching which evaluates extra user characteristics. The results demonstrate that the combination of collaborative filtering with profile matching achieves higher accuracy results than both baseline and standard collaborative filtering systems. The research demonstrates how recommender systems can develop personalized mental health support solutions which deliver effective treatment results.

Hybrid SARIMA-LSTM Architecture for XAG/USD Price Forecasting based on Historical Trends and Technical Indicators

Louis Ricardo (Binus University)*
Jhoner Lou (Binus University)
Bakti Jabar (Binus University)

Abstract

Forecasting silver price is still becomes one of a challenging problems because of its nature in financial econometrics as financial hedge and also industrial commodity. This results in not only linear seasonal trends, but also non-linear stochastic volatility. Statistical models such as SARIMA aren't capable of processing structural breaks and also deep learning models such as Long Short-term Memory (LSTM) also suffer from computational redundancy and overfitting when applied to non-stationary raw price data. This paper proposes a solution by combining both to be able to cover each of the weaknesses of both models, a two-stage hybrid model framework: SARIMA-Residual-LSTM approach. The SARIMA model will cover the linear trajectory and also seasonal components of daily silver prices in the 10-year period timeframe (from 2016-2026). The non-linear residuals will later be modeled using a multivariate LSTM network. The multivariate LSTM network will be integrated with a combination of technical indicators (RSI, MACD) and intermarket exogenous variables (Gold and DXY returns). This hybrid approach demonstrates an incredible result in outperforming both of the standalone models, achieving a Root Mean Square Error (RMSE) of 2.1633 compared to the SARIMA model with a score of 2.2161 and the standalone LSTM with an unstable result achieving 19.9570.

Comparative Sentiment Analysis of the Terra (LUNA) and FTX Crises Using BERTweet, DistilBERT, and FinBERT

Ferre Deka Santosa (Bina Nusantara University)*
Raymond Willen Kristopo (Bina Nusantara University)
Dennise Adrianto (Bina Nusantara University)

Abstract

To understand public opinion in the crypto market, sentiment analysis is crucial, especially during major events such as the collapse of Terra (LUNA) and the bankruptcy of FTX. This study conducts a comparative evaluation of three transformer models, which are: FinBERT, DistilBERT, and BERTweet, on Twitter tweets related to these events. This research uses a dataset consisting of 1,001 tweets that were manually labeled as negative, neutral, or positive. The results show that BERTweet achieved the highest accuracy at 82.79 percent, outperforming DistilBERT (78.81 percent) and FinBERT (76.16 percent). These results indicate that models trained using social media data process more informal and noisy text more effectively. Although FinBERT had the lowest validation score, domain mismatch limited its performance. Furthermore, this study shows that in handling informal discussions about cryptocurrencies, social media based models outperform models specialized for the financial domain. For improved performance, future research could include larger datasets and hybrid methods.

Enhancing Occupational Safety and Health Information Retrieval Using a RAG-Powered Question Answering System on NIOSH Guideline Documents

Nicholas Huang (Universitas Bina Nusantara)*
Rezki Yunanda (Universitas Bina Nusantara)

Abstract

Workplace accidents and illnesses cause human harm and economic loss. Practices of occupational safety and health are essential to minimize this problem, however the documents required for these practices are often lengthy, hard to navigate, and contain difficult terms. This research proposes a retrieval-augmented generation chatbot system on National Institute for Occupational Safety and Health guideline documents. The purpose of the system is to provide access to information required for practicing and understanding occupational safety and health. To help practitioners prioritize what actions to do, a feature by using a keyword-based approach to label obligation signal is used to produce deterministic results during indexing process. During answer generation, these labels are shown in the answer to help practitioners prioritize action. The system shows improvement in answer relevancy, context precision, and context recall with decrease in faithfulness when adding reranking stage. The optimal configuration for the system is baseline rag + reranker at K =10 with faithfulness score of 94.42%, answer relevancy score of 93.60%, context precision of 65.01% and context recall of 80.58%. This result shows the system's ability to retrieve information for answer generation and provide valuable advancements in guidelines RAG system.

Web Based AI for Running Warm-Up and Cool-Down Correction Using Blaze Pose and Joint-Angle Analysis

Immanuela Puspasari Saputro (Bina Nusantara University)*
Muhammad Rafiq Fajar (Bina Nusantara University)

Abstract

The increasing popularity of running among beginners is not always followed by adequate understanding of proper warm-up and cool-down techniques, which may increase injury risk. This study proposes a web-based AI application for real-time correction of running warm-up and cool-down movements using MediaPipe BlazePose and joint-angle analysis. The system captures user movements through a single RGB camera, extracts body landmarks, computes relevant joint angles, and evaluates movement correctness using rule-based criteria. Seven movements were implemented: Arm Circle, Quad Stretch, Lunges, High Knees, Frankenstein Walk, Stand Hamstring Stretch, and Shoulder Stretch. System evaluation involved seven participants who performed all movements in correct and incorrect conditions with five repetitions per condition, resulting in 490 movement-level trials. The results show an overall accuracy of 95.31% under a controlled testing protocol, with movement-level accuracy ranging from 90.00% to 100.00%. Shoulder Stretch achieved the best performance, while Frankenstein Walk showed the lowest accuracy due to its dynamic movement pattern. Preliminary user feedback indicated that the interface was understandable and responsive. These findings suggest that the system is feasible as an assistive tool for novice runners, although broader validation is still required.

A Comparative Study of Deep Learning and Machine Learning Approaches for Viral Transcriptomic Classification and Alcohol-Associated Hepatocellular Carcinoma

Muhammad Rasyak (Bina Nusantara University)*
Mahmud Isnan (Bina Nusantara University)
Bens Pardamean (Bina Nusantara University)

Abstract

Although distinct molecular profiles are found based on the etiology of hepatocellular carcinoma (HCC), classifying cases as viral or alcohol-based remains difficult due to the high-dimensionality nature of transcriptomic data. In this analysis, transcriptomic data from the TCGA GDC portal, which contains data for 134 viral-associated HCC samples and 119 alcohol-associated HCC, were used to evaluate classification. After eliminating overlapping samples found in both groups, 190 samples were contained in the final dataset — including a total of 112 viral-associated samples and 78 alcohol-associated samples. Following processing an initial collection of 34,686 genes, 3,000 of the most variable genes were selected for an autoencoder to perform dimensionality reduction to 128 latent dimensions ($R^2 = 0.918$). Classifiers including CNN, Random Forest, Support Vector Machine, and Logistic Regression were evaluated using features created through the dimensionality-reduced autoencoder. Among the classification methods used, Random Forest produced the best results (AUC-ROC = 0.817, accuracy = 73.7%) and outperformed the CNN (AUC-ROC = 0.747, accuracy = 68.4%) methods. Through gene-latent feature correlation analysis, the following genes are representative of the classifiers: BRCA1, PRKCZ, UBE2I and ATG16L1, all of which have been associated with pathways in DNA repair, cell signaling, and autophagy. Machine learning techniques that utilize classical machine-learning techniques and effective dimensionality reductions can produce results that are superior to deep-learning techniques when dealing with limited samples of transcriptomic data. The approach proposed in this study provides a foundation for defining an etiology-based classification system for HCC.

Predicting Jakarta's Air Quality Index Through SMOTE-Augmented Machine Learning: A Multi-Model Comparative Study on Imbalanced ISPU Data

Shanaa Salsabil Erawan (Bina Nusantara University)*
Najwa Zukhrufa (Bina Nusantara University)
Celine Yovela (Bina Nusantara University)
Azani Cempaka Sari (Bina Nusantara University)
Nicholaus Hendrik Jeremy (Bina Nusantara University)

Abstract

Air pollution in Jakarta has become one of the most crucial health issue and is also one of the most polluted city in the world.This research presents a comparative analysis of 5 machine learning models, such as Random Forest, XGBoost, Decision Tree,CatBoost, and LightGBM to predict Jakarta's Air Pollution Standard Index (ISPU) category. The categories are divided into four groups: 'Good', 'Moderate', 'Unhealthy', and 'Very Unhealthy'. The dataset with 5,531 records covered daily ISPU standards from 2010 to 2025 based on five train stations in Jakarta. Two experimental scenarios are evaluated to get the best result. First, a baseline without resampling and tuning. Second, an optimized framework that uses SMOTE application and Random Grid Search. The results show that Random Forest has the highest performance based on accuracy (95.03%) and macro F1 score (95.55%). SMOTE produces divergent effects across models: Random Forest showed consistent improvement across both accuracy and macro F1-score after resampling, while Decision Tree improved slightly in accuracy but declined in F1-score, and gradient boosting models (XGBoost, CatBoost, and LightGBM) experienced declines in both metrics despite gains in minority class recall for the Good class. Feature importance analysis identifies O3 and PM10 as the most influential pollutants. These findings provide a reproducible benchmark for ISPU classification and demonstrate the importance of algorithm-specific resampling strategies for imbalanced air quality datasets.

The Role of Gemma 3 Fine-Tuning in Classifying Indonesian Fake News: A Large Language Model Approach

Laurentia Alyssa Castilani (BINUS University)*
Muhammad Zarlis (BINUS University)

Abstract

As technology evolves, information is produced within a short period of time and in large quantities, making it hard for humans to detect whether the information is factual or not. A hoax refers to deliberately fabricated news designed to mislead readers, and its widespread dissemination can negatively impact organizational reputation, trust, and brand image. This study investigates the application of a Gemma 3-based large language model for Indonesian fake news classification using a custom dataset. A classification head is integrated through a custom wrapper, enabling supervised learning tailored to the task. The model is fine-tuned using Low-Rank Adaptation (LoRA), allowing efficient training under constrained computational resources. A key finding of this research is the importance of fine-tuning for domain-specific datasets. Comparative evaluation with a baseline (not-fine-tuned) Gemma 3 model demonstrates that, without adaptation, the model struggles to capture task-specific patterns. In contrast, the fine-tuned model achieves a macro-average F1-score of 99.67% on the internal test set, highlighting significant performance gains from fine-tuning. Additional comparisons with a widely used fine-tuning approach, IndoBERT combined with LSTM, further confirm that the fine-tuned Gemma 3 model consistently outperforms both the baseline and alternative approaches across evaluation metrics. External dataset evaluations reveal sensitivity to data characteristics, particularly text length, yet the model maintains stable performance under distribution shifts. These results emphasize that fine-tuning is crucial for adapting large language models to specialized classification tasks.

Sales Trend Forecasting System for Snack Distribution using Extreme Gradient Boosting Algorithm

Kezya Tanudiredja (Bina Nusantara University)*
Yanfi Yanfi (Bina Nusantara University)

Abstract

Predictive analytics is an essential tool in ensuring the sustainability of Micro Small and Medium Enterprises (MSMEs) in managing their operations. This is particularly significant in industries where there is inventory perishability as well as variability in sales trends. This study aims to develop an integrated framework in sales forecasting that transforms intuition-based management into data-informed decision-making. Using advanced ensemble learning approaches as well as non-linear regression analysis, this study addresses the challenges of managing seasonality in sales trends. Additionally, this study develops a heuristic formula in determining an inventory buffer using error metric analysis to avoid risks of inventory depletion as well as wastage. Experimental evaluation of the study using a regional distribution dataset indicates that the ensemble learning approach demonstrates improved accuracy over conventional linear regression in modeling non-linear consumption trends. The study concludes that integrating advanced predictive modeling approaches as well as error metric analysis in determining an inventory buffer is an effective solution in managing small-scale retail operations.

No Time to Waste: A Comparative Analysis of Deep Learning Models for Dynamic Garbage Classification

Richard Kurniawan (Binus)*
Evan Somangkey (Binus)
Felix Giovan (Binus)
Eko Setyo (Binus)

Abstract

Global waste management is a crisis worldwide, this research aims to address this issue by looking at and analyzing five deep learning architecture ConvNeXt, Inception V3, YOLOv8, Resnet50, and EfficientNet to determine each models' efficiency in classifying and identifying garbage with real life videos. Using trashnet and proprietary data augmentation to simulate variance found in real life scenarios. Convnext delivered the highest overall performance both in static images with an accuracy of 98.25% and a 0.9987 mAP score, convnext remained the most reliable model during video inference, effectively detecting multiple waste classes in bright and dark environments. While yolov8 offered superior localization, convnext 7 x 7 kernels were crucial in extracting key features in deformed waste. Convnext is identified as the most robust model in unpredictable real world video applications.

Long-Context Contrastive Adaptation for Automated Essay Scoring: Mitigating Information Blindness with BGE-M3

Emmanuel Widhiarto (Bina Nusantara University)*
Abba Suganda Girsang (Bina Nusantara University)

Abstract

Transformer-based Automated Essay Scoring Systems have a limited input range of 128 to 512 tokens, truncating longer essays and preventing the model from retrieving complete information from the documents. Furthermore, the models tend to be biased towards the middle of the score range and fail to detect rare high-scoring essays. Therefore, we propose an optimized BGE-M3 model capable of handling 8,192 tokens. We use score-aware contrastive learning for fine-tuning and imbalance-aware oversampling to address class imbalance. For the scoring system, we use similarity-weighted k-nearest neighbor retrieval. This experiment uses the ASAP 2.0 dataset and a 5-stage ablation study: model scaling, fine-tuning, context expansion, and comparison with baseline models. Context window expansion yields the biggest improvement from 128 to 1024 tokens, reducing the MAE from 0.7793 to 0.4658, increasing F1 Macro from 0.2077 to 0.5193, and achieving a QWK of 0.7740. This result is supported by an empirical token analysis showing that 1024 tokens account for 98.7% of the dataset. The proposed method also detects rare high-scoring essays, with Score 6 achieving an F1 score of 0.29. The generalization is also stable; the QWK gap validation vs. test is 0.006. These findings suggest that context length is a critical yet underexplored factor in AES system design.

Improving Implicit Feedback Recommendation with Matrix Factorization Model WARP Loss and SVD Prior Loss

Fakhri Aulia (Bina Nusantara University)
Emny Yossy (Bina Nusantara University)*

Abstract

Implicit feedback recommender systems typically use matrix factorization (MF) models with pairwise ranking objective such as Bayesian Personalized Ranking (BPR) . Recently, graph-based models are paid attention for its performance. However, they are computationally expensive. On the other hand, matrix factorization is still desirable due to its simplicity and efficient computation, notwithstanding that it is not explicitly designed for ranking purpose. Here we present an upgraded MF framework, which replaces the BPR loss with the Weighted Approximate-Rank Pairwise (WARP) loss and incorporates the SVD Prior loss. These methods improve ranking performance by using randomized Singular Value Decomposition (SVD) as a soft regularizer to capture global collaborative structures and focusing on hard negative samples. The proposed technique was tested on the Amazon Art dataset with a 10-core filtering mechanism. It gets the Recall@20 of 0.0665 and NDCG@20 of 0.0397 which is better than that of the baseline MF with BPR loss, i.e., 0.0421 and 0.0225 correspondingly. Furthermore, SVD Prior loss decreases the number of training epochs to converge than WAP loss alone. The results show that the combination of the ranking objectives and the SVD prior loss enhances the ranking accuracy without losing the efficiency of MF approach. This makes the proposed method fit for early stage adoption of recommender system particularly in e-commerce area.

Deep Transfer Learning for Predicting Corn Yield in Ukraine Using Remote Sensing and Climate Reanalysis Data

Angelyn Gunawan (Bina Nusantara University)*
Albert Verasius Dian Sano (Bina Nusantara University)

Abstract

War is limiting traditional ground-based crop yield surveys in Ukraine, the fourth-largest corn exporter in the world. The only viable alternative method to predict corn yield in the country may be remote-sensing-based analysis with deep learning techniques; however, historical corn yield record in Ukraine is limited, leading to a data scarcity problem. To address this challenge, this study developed and evaluated an LSTM transfer learning model pre-trained on U.S. corn belt data in 2010–2024 as the source domain and transferred to predict corn yield in Ukraine as the target domain by fine-tuning on 2015–2021 data and testing on 2022–2024 data. Multi-source data not requiring ground access comprising of remote sensing, climate reanalysis, soil properties, and historical yield were utilized. Results demonstrate that, compared to a non-transfer-learning model, transfer learning reduces RMSE by 18% from 1.06 t/ha to 0.86 t/ha and MAPE by 3.7 percentage points from 14.9% to 11.2%, while increasing R^2 by 6.8 percentage points from 79.4% to 86.2%; and proves to improve spatial consistency when predicting on various oblasts. However, transfer learning does not always improve performance, with mild negative transfer occurring when testing on 2023 data. SHAP analysis shows that both models still relied heavily on regional yield averages, rather than remote sensing and climate features, as the most important factor when predicting yield. Overall, deep transfer learning is one effective, albeit not guaranteed, solution for corn yield assessment in Ukraine under restricted ground access and limited data.

Development of Rule-Weighted Ensemble Method for Masking Personally Identifiable Information in Financial Document Using Regex and NER

Stephanie Makmur (BINUS)*
Grisella Zega (BINUS)
Richie Hartono (BINUS)
Bakti Jabar (BINUS)

Abstract

Personally Identifiable Information (PII) is information that can identify a specific individual, so its confidentiality must be maintained. Along with the development of digital transformation that adapts Generative Artificial Intelligence (GenAI) technology, as many as 51% of domestic financial institutions are focusing on using GenAI to increase the efficiency of daily tasks. However, not all cloud-based GenAI services used have policies to delete entered personal data. Therefore, without a strict data filtering mechanism before sending data to the cloud, there is a potential risk of data leakage and misuse of information contained in user financial data that is intentionally or unintentionally uploaded. Therefore, risk mitigation is carried out by anonymizing PII data. Previously, PII identification and masking were performed using the Regex method, but experienced many false positives because it was only effective for entities with consistent patterns. Entities with unstructured text patterns can only be detected using the NER model. In this paper, a combination of the Regex method and two NER models was developed using a rule-weighted ensemble approach. This development achieved evaluation results with a precision value precision of 89.50%, a recall of 81.65%, and an F1-score of 85.40%. These results indicate an increase in accuracy compared to the individual Regex results, although they are still lower than the individual NER model results.

The Impact of AI Usage For Accounting Students Learning Process And Motivation

Gilbert Rae (BINUS University)*
Alberto Rusli (BINUS University)
Rafael Kalyana (BINUS University)
Michael Angelus (BINUS University)

Abstract

The swift advancement of artificial intelligence (AI) technology has immensely influenced higher education. Specifically, it improves the effectiveness and accessibility of learning. Yet, the escalating usage of AI technology among students triggers worries about their cognitive development because they might rely too much on AI and lack critical thinking skills. This research focuses on examining the effect of the use of AI on the learning experience of accounting students, where the dependent variable will be measured through the multidimensional construct of the use of AI, encompassing AI Usage Frequency, AI Dependency, and AI Learning Support Effectiveness. The cross-sectional design was utilized in collecting data from 95 accounting students in Jakarta and South Tangerang by implementing a structured questionnaire. Subsequently, the collected data were processed with Partial Least Squares Structural Equation Modeling (PLS-SEM). The outcome of the analysis reveals that there is a significant and positive correlation between AI Usage Frequency and learning experience and a significant and negative connection between AI Dependency and learning experience. Furthermore, there is no significant association between AI Learning Support Effectiveness and learning experience. This study provides valuable information on the role of AI technology in the learning process of students.

Multimodal Drug-Target Interactions Prediction with 3-Dimensional Structure Exploitation and Computation Optimization 2

Xinyuan Chen (unikl)*

Abstract

Predicting drug–target interactions (DTI) is a key step for speeding up drug discovery, re-using existing compounds, and spotting possible side effects. A big part of any DTI model is how it encodes drugs and proteins. For instance, using the three-dimensional (3D) shape of a protein lets the model see subtle atomic contacts that are otherwise invisible, which usually translates into more reliable predictions. Most knowledge-graph-based DTI methods, however, rely on a single source of information. Existing multimodal approaches tend to ignore the fact that several objectives—such as potency, toxicity, and synthetic accessibility—pull in opposite directions, and they rarely enforce biochemical constraints in a principled way. Here we introduce MMDTI-MOEA, a framework that couples multimodal 3D structure encoding with a constrained multi-objective evolutionary optimizer. Experiments on public benchmarks show that the proposed method consistently beats strong baselines; the mean-squared error (MSE) drops by about 8.9\,\% compared with representative graph-based models. Ablation studies confirm that both the 3D features and the multi-objective component are essential. Moreover, the method yields interpretable Pareto-optimal solutions, making it easier to navigate the trade-off between activity and toxicity during lead optimisation. By cutting down on redundant screening and \textit{in silico} filtering, the approach also contributes to a more sustainable, resource-efficient drug development pipeline.

Explainable Hybrid Machine Learning Model for Daily AQI Predicting in Jakarta

Violitta Yesmaya (BINUS)*

Abstract

Air pollution is one of the serious environmental and public health problems in Jakarta, where air quality is often in the unhealthy category due to high concentrations of particulate matter and pollutant gases. This study proposes a combined approach between machine learning and explainable AI so that it can accurately explain the pollutants that influence the daily Air Quality Index (AQI) in Jakarta. The data used are historical air quality data for the period 2016-2025 with the main pollutant parameters, namely PM2.5, PM_{10}, O_3, SO_2, NO_2, and CO. The three-machine learning used are Long Short-Term Memory (LSTM) as a time series-based deep learning representation, as well as Random Forest and XGBoost as an ensemble model based on a decision tree. Performance evaluation was conducted using RMSE, MAE, R^2 Score, and MAPE metrics. The results of the experiment showed that Random Forest performed best, with an RMSE of 0.1781, MAE of 0.0179, R^2 score of 0.9669, and a MAPE of 0.57%, indicating a very high level of prediction accuracy and stability. XGBoost exhibits competitive performance, while LSTM generates greater prediction errors, confirming that deep learning models are not always superior to tabular and structured environmental data. Using Explainable AI (SHAP), this study proves that O_3 and PM2.5 are the main determinants of AQI in Jakarta. In addition to being accurate, this model provides transparency of causative factors, so it is very practical to use as a support system for air quality management decisions.

Pareto-Based Evolutionary Reward Shaping for Deep Reinforcement Learning Traffic Control

Karldritz Hanson (Bina Nusantara University)*
Eikel Barus (Bina Nusantara University)
Frederick Kuntjoro (Bina Nusantara University)
Ivan Edbert (Bina Nusantara University)

Abstract

Urban traffic congestion is a persistent problem that involves conflicting objectives, such as maximizing throughput, ensuring fairness, and maintaining flow smoothness. Traditional Deep Reinforcement Learning (DRL) approaches typically rely on single scalar reward functions that prioritize throughput at the expense of the other objectives. The authors proposes a Pareto-based Evolutionary Reward Shaping framework that integrates the Non-Dominated Sorting Genetic Algorithm II (NSGA-II) with a Long Short-Term Memory Deep Q-Network (LSTM-DQN) agent. Our approach uses NSGA-II to dynamically evolve the weights of a multi-objective reward function, which enables the LSTM-DQN agent to learn an optimal policy that are both high-volume and high-efficiency. The findings suggest that the framework successfully generates a Pareto Front of non-dominated solutions and achieved a peak throughput (1,059 vehicles), while also reducing queue variance by 35% (13.0) and total stops by 28% (14,676) compared to baselines. This research demonstrates how an evolutionary reward shaping approach outperforms single-objective DRL agents, and allows for a diverse option of optimal strategies to effectively control the unpredictable nature of traffic.

A Comparative Study on EEG Classification using FFT for Emotion Analysis with GRU, BI-GRU, LSTM, BI-LSTM

Filbert Winch (Universitas Bina Nusantara)*
Andry Chowanda (Universitas Bina Nusantara)

Abstract

Electroencephalography (EEG) is a signal that is purposed to visualized the current state of the human brain. Subbya et al. in it's previous research stated that it is very difficult to observe it in time domain just by observing it directly. It is found that brainwave activity by nature are very random, causing the data to be both non-linear and non-stationary. In this paper, advanced computational methodologies are used to process the raw numerical data resulted from the commercial headband's wave patterns. In order to navigate the complex dimensionality of EEG, it is profound that machine learning specifically in deep learning can be utilized. The purpose of the proposed experiment is to analyze the performance of how each deep learning architecture performs for predicting emotions. The experiment systematically evaluated four Recurrent Neural Network (RNN) variants, which consist of Long Short-Term Memory (LSTM), Bidirectional LSTM (Bi-LSTM), Gated Recurrent Unit (GRU), and Bidirectional GRU (Bi-GRU). These models are trained across varying batch processing configurations (128, 256, and 512). The result Shows that the BI-LSTM architecture with batch size of 512 achieved the best result or in other words, highest predictive capability, that resulted in 98.75% of accuracy with a test loss of 0.6252.

Skin Tumor Classification using Convolutional Neural Network (CNN): Benign, Pre-Cancer, or Malignant

Golda Salazar (Bina Nusantara University)*
Adila Yuhanas (Bina Nusantara University)
Karto Iskandar (Bina Nusantara University)

Abstract

Skin Tumors are lesions caused by cells growing abnormally on the skin, ranging from benign, pre-cancerous, to malignant, where the latter signifies cancer. Early diagnosis is best done to determine the type and category of tumors as well as increase the possibility of recovery in the case of cancer. However, as visual examinations to receive diagnosis require a certain amount of time and cost, a type of deep learning algorithm referred to as the Convolutional Neural Network (CNN) could be used to generate accurate image classifications regarding skin lesions, aiding the examination to receive early diagnosis. Two architectures of the CNN, ResNet and MobileNet, will be trained using the HAM10000 collection from the International Skin Imaging Collaboration (ISIC) to determine the type of skin tumor, establish the category (Benign, Pre-Cancerous, or Malignant), and provide information pertaining to each result. According to the model training, the MobileNetV2 model resulted in an overall accuracy of 73% but is heavily affected by the majority classes in the dataset. The model struggles in classifying two minority classes, particularly one class pertaining to the pre-cancerous category. The ResNet50 model outperforms the MobileNet model as it produced an overall accuracy of 69% but displays strong generalization across all 7 classes.

Explainable Ensemble Learning for Water Potability Classification Using SHAP Based Feature Interpretation

Reinardus Djafar (Binus University)
Diandra Auwlia (Binus University)
Abram Setyo (Binus University)*

Abstract

Ensuring safe drinking water is a significant problem in environmental monitoring and public health. Machine learning methods are widely used to solve water potability classification problems. However, many models lack interpretability in explaining how predictions are made. This study provides an explainable machine learning strategy for water potability classification utilizing a weighted soft voting ensemble model that incorporates numerous classifiers, including SVM, Random Forest, XGBoost, CatBoost and KNN. Hyperparameter tuning was performed using RandomSearchCV with 100 iterations and 10-fold cross validation to improve model performance. The experimental data have shown that the ensemble model has an accuracy of 75.5%, an F1-score of 75.6% and a ROC-AUC of 81.3%, suggesting a considerable improvement in overall performance compared to individual models. In order to improve the interpretability of the ensemble model, SHAP (Shapley Additive Explanations) was used to identify the most impactful features in relation to water potability classification. The results show that pH, sulfate, and solids are the most impactful features in revelation to classification results. This shows that ensemble learning improves not only performance but also interpretability using explainable AI.

Multimodal AI Framework for Personalized and Context-Aware Skin Disease Diagnosis, Monitoring, and Treatment Support

Oshidee Wijesinghe (Sri Lanka Institute of Information Technology)*
Ravindran Laksopan (Sri Lanka Institute of Information Technology)
Mihisandali Mayusha (Sri Lanka Institute of Information Technology)
Paboda Devindi (Sri Lanka Institute of Information Technology)
Lokesha Weerasinghe (Sri Lanka Institute of Information Technology)
Akshi Silva (Sri Lanka Institute of Information Technology)

Abstract

WITHDRAWN

Advancing Emotion Recognition through Cross-Modal Fusion and Gender-Aware Learning

Darren Gunawan (Binus University International)
Viona Djunaidy (Binus University International)
Rafael Angelo Christianto (Binus University International)
Nunung Nurul Qomariyah (Binus University International)
Feri Setiawan (Binus University International)*

Abstract

Emotion recognition plays a crucial role in human-computer interaction as they enable systems to respond empathetically to users. However, existing approaches rely heavily on single-modality inputs, particularly text, which limits their ability to classify emotions effectively through speech and demographic characteristics. To address these limitations, this study proposes TrueTone, which is a multimodal emotion recognition model that utilizes linguistic and paralinguistic data through a transformer-based and cross-modal architecture, with gender-aware conditioning. MELD dataset was utilized, along with strategies to handle class imbalance. The results show that TrueTone outperforms the baselines, achieving a weighted F1 score of 79.65\% and accuracy of 78.77\%. In addition, it was found that gender-aware conditioning improves macro F1 score performance by enhancing recognition of underrepresented emotions. This study highlights the effectiveness of cross-modal fusion and gender-aware learning for robust and fair emotion recognition.

A Two-Step Machine Learning Approach for Clinker Free Lime Prediction toward Energy-Efficient Cement Kiln Operation

Slamet Riyanto (Bina Nusantara University)*

Abstract

Clinker quality variations can influence energy consumption and cement strength in cement production. Free lime (f-CaO) is an indicator, which is influenced by dynamic and multivariate cement kiln operating conditions. However, accurate prediction remains challenging. In previous studies, machine learning has been widely used to predict free lime actual value at the current time. The research observes the impact of two-step prediction compared to one-step prediction on the prediction performance. Support Vector Regression (SVR), Random Forest (RF), and Extreme Gradient Boosting (XGBoost) are used in this study. Features are derived from high-frequency process variables that are obtained from the Distributed Control System (DCS). Additional features and targets are low-frequency data from laboratory measurements. Model performance was assessed using R^2, RMSE, and MAE based on time-series cross-validation. The study shows that two-step prediction performs better in all methods. The one-step XGBoost model resulted MAE of 0.3087, RMSE of 0.3968, and R^2 of 0.2011. On the other hand, the two-step XGBoost model improved to MAE of 0.0256, RMSE of 0.1088, and R^2 of 0.9442. This corresponds to a reduction of 91.7% in MAE and 72.6% in RMSE, and an increase of 369.5% in R^2. This study also analyzes feature importance and SHAP. That indicates that previous Lime Saturation Factor (LSF) and kiln motor current variations are the most influential predictors. This approach can assist operators in controlling kiln operations. The proposed approach enables high-frequency estimation of free lime, which may support more timely monitoring of kiln conditions and provide insights into process dynamics.

Evaluating the Impact of Quantization on Indonesian Question Answering Performance in Large Language Models

Vincent Ramaputra (Bina Nusantara University)*
Setiawan Joddy (Bina Nusantara University)

Abstract

Large Language Models (LLMs) have shown strong performance across many natural language processing tasks, including question answering (QA), but their high computational requirement makes them difficult to deploy in resource-constrained environments. Quantization is one of many ways to make model deployment in these environments possible, by reducing models' weights into lower precision formats to lower model size and improve inference speed. However, its impact on Indonesian QA task has not been widely studied. This study investigates how post-training quantization affects Indonesian question answering performance in LLMs. We evaluate multiple model sizes from two commonly used LLM families under different quantization settings, specifically 8-bit and 4-bit precision using AWQ and GPTQ methods. Our evaluation considers both answer quality and inference efficiency. Results show that 8-bit quantization maintains similar performance to the original model while improving inference speed, while 4-bit quantization provides greater efficiency gains but leads to a more noticeable drop in performance. We also discovered that quantization impacts extractive and generative QA differently, with generative tasks showing greater robustness. Our findings highlight important trade-offs between efficiency and accuracy, and provide practical guidance for selecting quantization strategies when deploying LLM-based QA systems for Indonesian in resource-limited environments.

Social Media Based Depression Recognition Using BERT with Hyperparameters Optimisation

Andry Chowanda (Bina Nusantara University)*
Mohamed Imran Bin Mohamed Ariff (UiTM Cawangan Perak)

Abstract

Depression is one of the top causes of suicide worldwide. It affects around 280-300 million people globally. Undiagnosed and untreated individuals cause most problems. Early mental health diagnosis allows for prevention and treatment and prevent untreated mental health from causing self-harm or suicide. Most people now share their issues and stories on social media. The exponential expansion of social media platforms to exchange stories has changed how we track public health, such as mental health. The platforms can serve as a digital diary for individuals to express feelings, struggles, and potential psychological distress. Analysing those emotionally rich textual data streams allow us to build a mental health problem early detection system. This might prevent severe consequences of undiagnosed and untreated mental health, which can leads to self harm or suicide. This paper presents experiments of transformer based models for depression recognition using text from social media. The research evaluates five architectures with different preprocessing setups, data augmentation techniques, and automated hyperparameter optimisation. A total of seventeen experiments were conducted in this research. The results shows that a model trained with bert-base-uncased pre-trained model, augmented and automated hyperparameter tuning, provides the best and balanced results with an accuracy of 68% and an F1 score of 63.63% in the test dataset.

Leveraging AI-Driven Data Analytics to Identify the Sanitation-Poverty Correlation via Comparative Clustering Analysis

Prabowo Wahyu Sudarno (Bina Nusantara University)*
Nur Sitha Afrilia (Bina Nusantara University)
Kumara Pinasthika Dharaka (Bina Nusantara University)
Nico Yonatan Wicaksana (Bina Nusantara University)
Hermawan Honggo Widagdo (Bina Nusantara University)

Abstract

The world faces critical challenges related to poverty. There is a question whether there is a correlation between basic infrastructure and people's income. It is hypothesized that when sanitation coverage continuously differs, it can lead to a structural poverty trap. Poor living conditions reduce the quality of public health and productivity, thus causing economic poverty. To identify this problem, this research conducted a proof of the correlation between poverty and infrastructure such as sanitation. This research proposed the use of unsupervised machine learning using K-Means and Agglomerative Hierarchical Clustering. Initially, the elbow method was used to find the optimal number of clusters and it was found that k=2 efficiently captured most socioeconomic differences. The K-Means algorithm achieved a superior Silhouette Score of 0.6164 compared to AHC. Cluster and centroid visualizations provide empirical evidence of the structural linkages between variables, with lower sanitation coverage correlated with higher poverty and vice versa. These results indicate that access to sanitation acts as one of the structural anchors of poverty, and the evidence provides strong guidance for policymakers to prioritize infrastructure-based interventions such as sanitation for poverty eradication.

Controllable Identity Obfuscation via Multiple SVM Boundaries for Privacy-Preserving Face Dataset Generation

Dhanushraaj Shanmuganandan (PES University)*
BG Pranav (PES University)
Suhas Chandan (PES University)
Abhijit Shanbhag (PES University)
Nitin Pujari (PES University)

Abstract

Privacy regulations such as the General Data Protection Regulation (GDPR) encourage the removal of identifiable information from face datasets, yet existing methods demand extensive retraining or lack fine-grained control to balance privacy and utility. We present a training-free approach in which trained in latent space to separate identity from general face distribution in a GAN's latent space, each trained with independent random negative sets, produce diverse transformation trajectories. A privacy parameter δ provides custom anonymization from a single input image.

Evaluating YOLOv12 for Road Pothole Detection: A Cross-Dataset Performance Comparison

Jonathan Wijaya (BINUS University)*
Suryadiputra Liawatimena (BINUS University)

Abstract

The detection and localization of road potholes are important for infrastructure maintenance, traffic safety and predictive maintenance planning. Deep learning has come a long way in the recent years with single-stage object detectors like YOLO, which have shown impressive speed-accuracy trade-offs making them suitable for real-time applications. However most current research evaluates models on individual datasets, which limits our understanding of their performance across different datasets. This study aims to fill this gap by providing a systematic cross-dataset performance evaluation of the latest YOLOv12 architecture across six distinct pothole detection benchmarks which are RDD2020, RDD2022, PothRGBD, MWPD, Nienaber's Pothole Dataset, and Mixed Pothole Dataset. Experimental findings show that YOLOv12 maintains its accuracy on large-object datasets achieving 96.2% mAP50 on PothRGBD while achieving state-of-the-art inference performance of up to 166.67 FPS. However, on small-object-dominated datasets there is a noticeable decline in performance with 64.2% mAP50 (-8.8%) on Nienaber's dataset. Varying weather and lighting conditions also contribute to performance declines as seen in the MWPD dataset, where mAP50 reduces to 75.3% (-19.7%). Our results highlight a crucial trade-off between speed and accuracy. Through error analysis we found that YOLOv12's architectural optimizations improve computational efficiency. But through feature map we found that R-ELAN backbone and Area Attention mechanism may in a way affect the preservation of the necessary fine-grained spatial features for small pothole detection. Based on our results, we suggest a hybrid two-stage detection architecture where we use YOLOv12's high-speed inference for preliminary screening. Then, we use precision-focused detectors for refined verification. This method offers deployment techniques for automated road infrastructure monitoring platforms and intelligent transportation systems. This can find the middle ground between real-time performance and detection accuracy.

Improving YOLOv11 Model for Fabric defect detection using Knowledge Distillation approach

Suwandi The (Binus University)*
Alif Tanoto (Binus University)
Joshua Rawatan (Binus University)
Feri Setiawan (Binus University)
Nunung Qomariyah (Binus University)

Abstract

AI-based Fabric defect detection is a crucial component in the textile industry to ensure consistent product quality as manual inspections are prone to human-error. However, accurate detection models are large, and deploying the models remains a challenge, especially in the environment where the computational resources are limited. Therefore, to address this issue, we proposed model pruning and fine-tuning, combined with knowledge distillation to reduce the model size and further improve the performance as well. In this study, two distillation approaches, including Masked Generative Distillation and Channel-Wise Distillation are explored to test their effectiveness in improving the performance of YOLO11n student model for fabric defect detection task by transferring the knowledge from the larger YOLO11m teacher model. Structured channel Pruning and Fine-tuning were first applied to the student model before the distillation. Our experiments were conducted on a fabric defect dataset with 6 types of defect including hole, knot, line, seam, stain, and stitch. The results have shown that both MGD and CWD are effective in increasing the overall mAP50 value of the student model by achieving 1.87% and 1.73% increase for MGD and CWD respectively. Additionally, we also achieved 7.95% model parameters reduced, which also demonstrated the effectiveness of model pruning in this study. These findings have shown that combining pruning and fine-tuning with knowledge distillation is an effective way to improve the model efficiency, without losing any performance as well. Lastly, our solution was integrated into a Gradio python app and deployed on AWS EC2 instance with Docker.

Robust Detection of AI-Generated Text via Hybrid mBERT and Stylometric Feature Modeling

Leander Farrell Suryadi (BINUS)
Kent Alber Fredson (BINUS)*
Yithro Paulus Tjendra (BINUS)
Abba Suganda Girsang (BINUS)

Abstract

The advancement of Large Language Models (LLMs) such as ChatGPT and Gemini has made it extremely difficult to distinguish between LLM-generated and human-written text, concerning integrity and authenticity for academic writing nowadays. Despite abundance detection methods and tools, they still encounter difficulties with high-level paraphrasing and adversarial techniques. To overcome that, this study proposes a strong hybrid detection framework of combining eighteen stylometric features with multilingual BERT (mBERT) contextual embeddings. The architecture consists of dual-branch which take semantic and linguistic features from the input. In addition, this study utilizes the latest DAIGT v4 dataset which consists of many human-written student papers and LLM-generated texts. The proposed method achieves an accuracy of 98.9%, an F1-score of 99.2%, a recall of 99.3%, and an AUC of 0.998. These results confirm that the integration of semantic and stylistic information offers more dependable method for differentiating between human-written text and LLM-generated content.

Real-Time Traffic Monitoring and Management Using Deep Learning-Based Vehicle Detection, Density Estimation, and Adaptive Signal Control

Suzana Basaruddin (Universiti Kuala Lumpur)*
Ritu Chauhan (Amity University)
Megat Farez Azril Zuhairi (Universiti Kuala Lumpur)
Khushi Mehta (Amity University)
Juliana Jaafar (Universiti Kuala Lumpur)

Abstract

Traffic congestion in cities is one of the most urgent issues in modern smart-city networks that creates significant economic losses, environmental damage, and increases the risk of road safety. This paper presents RT -TMMS, a Real-Time Traffic Monitoring and Management System that has combined deep-learning-based vehicle detection, multi-zone traffic density estimation, vehicle classification, speed estimation and a traffic signal control adaptive system into a single web-based operational dashboard. The system uses YOLOv9-TURBO, a state-of-the-art single-stage object detector developed with DeepSORT multi-object tracking, and provides a mean Average Precision of 94.7 per cent (mAP0.5) on the UA-DETRAC benchmark dataset, which is a real-world surveillance video collection of more than 140000 frames and 121million labelled vehicle bounding boxes at 24 intersections. RT -TMMS receives real-time camera images of 12 surveillance nodes at a rate of 30 fps, maps 16 zones of traffic density in real-time with four levels of severity, classifies cars to one of five categories and directs detection data to an adaptive traffic signal controller. Compared to UA-DETRAC and the CityFlow multi-camera benchmark, it is found that the average waiting time of vehicles decreases by 31.4% and intersection throughput increases by 18.7% compared to fixed-time signal baselines. The latency of the end-to-end system is 12ms which makes the deployment in operational smart-city traffic control centres possible.

Dynamic Pricing under Queueing Constraints in Electric Vehicle Charging Systems: A Comparative Analysis of Rule-Based and Reinforcement Learning Policies

Pandu Dwi Luhur Pambudi (BINUS University)*
Imelda Juliana BR. Sibarani (BINUS University)
Cedric Sebastian (BINUS University)

Abstract

In this work we evaluate the potential of dynamic pricing strategies to optimize the management of Electric Vehicle (EV) charging stations in the presence of fluctuating demand and inherent hardware constraints. As the use of electric vehicles is rapidly expanding, the suppliers of charging infrastructure are facing the challenge of high load variations, with large spikes during peak periods and significant underutilization during off-peak periods. Such mismatches lead to large operational inefficiencies. Hence, this study tries to propose a pricing model that can stabilize the demand distribution without sacrificing its economic feasibility. To tackle this difficulty, we built a simulation model that considers the charging process as a stochastic multi-server queuing system. In this system, vehicle arrivals are portrayed as random intervals, with user behavior mimicking real-time responsiveness to pricing signals. The pricing mechanism is also stated as a sequential decision making issue, where operators can dynamically alter the tariffs according to the real-time station occupancy to optimize the overall revenue.

Which Deep Learning Model Works Best for Emerging Market Stocks? Comparing LSTM and Transformer Models on 47 Indonesian Banks

Adi Teguh Suprapto (Bina Nusantara University)*
Panji Arisaputra (Bina Nusantara University)

Abstract

Predicting stock prices in emerging markets is notoriously difficult due to higher volatility, scarce historical data, and unique market microstructure features. This study addresses the question: which deep learning model works best for this challenging task? We compare the performance of Long Short-Term Memory (LSTM) and Transformer models using daily data from 47 Indonesian bank stocks spanning July 2022 to July 2024. Both models were evaluated using multiple metrics, including MSE, MAE, RMSE, R^2, MAPE, and MedAE. Our experimental results show that the LSTM model consistently outperforms the Transformer, achieving a substantially lower MSE of 755.93 compared to the Transformer's 5,706.89 on the combined test set. Furthermore, LSTM demonstrates superior capability in capturing both short- and long-term dependencies and exhibits greater robustness when handling volatile stock price fluctuations. These empirical findings align with recent research suggesting that simpler architectures with appropriate inductive biases often outperform more complex models in financial time-series forecasting. Visual comparisons of predicted versus actual stock prices further validate LSTM's enhanced accuracy and consistency. This study contributes to the growing literature on deep learning applications in emerging financial markets, offering practical guidance for model selection in banking sector forecasting.

Automated Interpretation of ECG to Predict Potential Cardiovascular Diseases

Amrita Naik (Don Bosco College of Engineering)*

Abstract

Cardiovascular diseases (CVDs) are among the top ten causes of mortality worldwide, calling for early diagnosis that may impact treatment and the outcome of the patient in a favorable way. The main focus of the study is to develop a machine-learning (ML)-assisted model capable of classifying and recognizing CVDs based on ECG images. The methodology proposed includes image preprocessing techniques comprising gray scale conversion, grid removal, contour detection, and normalization for effective extraction of features from the ECG signals. The dimensionality of feature is further reduced using PCA algorithm. The Mendeley ECG Image Dataset was used for model training and tested using various ML techniques such as SVM, KNN, Decision Tree, Logistic Regression, Naïve Bayes, and Random Forest. The classified 12-lead ECG signals belong to five categories, namely Normal, Myocardial Infarction (MI), Previous History of MI (PMI), COVID-19, and Abnormal ECG. Experimental results suggest that KNN, Decision Tree, and Logistic Regression classifiers showed the highest accuracy at 97%, outperforming all other implemented models. The proposed system enables remote interpretation of ECGs, allowing for the convenience of patients to consult doctors without physical presence, hence providing a scalable, accessible solution for early detection of CVDs.

A Preliminary Study on Mitigating Upcoding in Indonesia's National Health Insurance using Multi-Agent Multi-Armed Bandits

Elisabeth Nilasari (Sanata Dharma University)*
Bambang Soelistijanto (Sanata Dharma University)

Abstract

Developing countries are often faced by corruption. Corruption threatens public sector organizations' viability. For instance, recent corruption at BPJS Kesehatan (Indonesia's universal health care provider) has recently led to enormous financial losses. Adversaries causing losses intelligently game the system. Existing static solutions cannot easily keep pace with the adversary. Limited audit resources further compound the problem. With hard budget and human-resource constraints, how can auditing efforts be directed to wisely choose claims that are suspicious for human review? We introduce a decentralized auditing system built with Multi-Agent Multi-Armed Bandit (MAMAB) framework. Contextual Bandits models allow us to reason about individual claims using attributes of a patient's claim. Bandits with Knapsacks (BwK) allow us to reason about the daily limit of human reviewers. Our goal is to leverage MAMAB to efficiently utilize limited resources and choose claims with the highest likelihood of containing fraud for review by a human. Moreover, we can theoretically allow regional BPJS branch offices to act as individual agents that share their knowledge of adversary behaviors (instead of sharing patients' information) to quickly learn adversarial patterns and minimize overall regret. In this paper, we present a high-level problem formulation and system design to implement MAMAB in BPJS Kesehatan's constrained environment.

Gated Hybrid Architecture: Fusing Contextual Representations and Lexical Signals for Indonesian Hate Speech Classification

Matthew Halim (BINUS University)
Nathanael Widjaja (BINUS University)*
Derwin Suhartono (BINUS University)

Abstract

The development of social media in Indonesia has led to an increase in harmful content such as hate speech, especially on platforms like X. Although many transformer-based models such as IndoBERT are capable of extracting language information, these models often have difficulty handling informal languages such as ambiguous vocabulary, slang, and even out-of-vocabulary (OOV) phenomena. This study proposes a Gated Hybrid Architecture approach that combines contextual representation with external lexical features. This architecture uses a gradual fusion mechanism to dynamically weigh contextual and lexical features prior to classification. Through experiments with a multilabel dataset of 13,169 tweets, the hybrid model successfully achieved an accuracy of 0.8876 and a macro-f1 of 0.8596, representing a 7.96% point increase over the base model. In addition, the model successfully increased recall on the minor class "Abusive" from 0.61 to 0.78. These results indicate that the integration of external lexical features successfully boosts the capabilities of the contextual model, resulting in a more relevant hate speech detection system in informal language processing.

Improving the Accuracy of Facial Age Estimation Models for Social Media Platforms Using Hybrid CNN-Transformer Model

Daniel Amarta (Bina Nusantara University)*
Nicholas Laksono (Bina Nusantara University)
Stefanus Benhard (Bina Nusantara University)
Derwin Suhartono (Bina Nusantara University)

Abstract

Age estimation remains an ongoing issue in age verification systems in social media platforms as part of moderation towards young internet users. As regulations from UK Online Safety Act and EU Digital Services have been heightened in recent times, the need for necessary age verification technologies comes into play. However, most AI-based age estimation systems are still widely inaccurate and unsecure, usually costing internet user experience and personal data safety. This study introduces a new approach to age estimation models by incorporating a hybrid CNN (Convolutional Neural Network) and Vision Transformer model to improve the overall performance compared to the overall baseline model by improving the processing pipeline to better capture the overall facial feature map of facial images. This study has been evaluated on several
public datasets such as UTKFace, FG-NET, MORPH2, IMDB-WIKI, ONOT and a combined dataset (UTKFace + MORPH2 + IMDB-WIKI) with comparison to the baseline model. The results have shown that the hybrid model has improved accuracy results for example, reaching an within dataset MAE of 4.564 and cross dataset (evaluated using IMDB-WIKI) MAE of 8.827 when trained using UTKFace, compared to baseline model which got an MAE of 4.633 and 17.22 respectively.

A Dual-Stream Deep Learning Architecture for Deforestation Land Cover Classification in Indonesia Using Optical and Infrared Satellite Imagery

Muhammad Ferlanda (BINUS University)*

Abstract

Deforestation in Indonesia has been a problem for a long time and is hard to solve. We need a monitoring system that genuinely works—one that is accurate and can last—to solve it. Satellite photos seem like the best solution that can be used on a big basis. But normal optical data doesn't always work in the tropics. It merely can't properly tell the difference between complicated landscapes, including big palm oil plantations, smallholder plots that are spread out, and thick natural forests. Additionally, datasets on deforestation in the actual world are very unbalanced, which makes it harder to automatically classify them. To tackle these constraints and difficulties, this research introduces a deep learning architecture employing a resilient dual-stream to amalgamate optical (RGB) and infrared (IR) satellite pictures via a late-fusion approach. Label Smoothing regularization is used to fix the class imbalance and stop overfitting without punishing minority traits. Also, a 4-way Test Time Augmentation (TTA) technique is employed to make predictions more stable during inference. This system fully tests our proposed architecture against eight State-of-the-Art (SOTA) models: MobileNetV4, ConvNeXtV2, SwinV2, ResNeXt50, ResNet18, EfficientNetV2, RegNetY, and YOLOv11. The results of multiple tests with different deep learning models demonstrate that the dual-stream modality model will greatly increase the model's capacity to tell the difference between different types of land cover. The SwinV2 transformer backbone can get the best results with this design, with an overall accuracy of 80.54% and a balanced Macro F1-Score on the test set that has never been seen previously in other investigations. This study also creates a dependable and scalable system for automated deforestation monitoring that can help Indonesia make better environmental policy decisions.

Robust Deepfake Image Detection Under Image Distortion Using Fine-Tuned CNNs and Stacking Ensembles

Vincent Sebastian Gunawan (BINUS)*
Christopher Alexander Chandra (BINUS)
Said Achmad (BINUS)

Abstract

The rapid development of generative models has increased the realism of synthetic facial images, raising concerns related to misinformation, digital fraud, and identity manipulation. Consequently, effective deepfake detection methods are essential to ensure the authenticity of digital media. This study proposes a deepfake image detection framework based on a stacking ensemble of multiple Convolutional Neural Network (CNN) architectures combined with transfer learning. Several pretrained CNN backbones are used as base learners to extract discriminative visual features from facial images. Selected layers of the pretrained models are fine-tuned to adapt the networks to the deepfake detection task while preserving useful feature representations learned from large-scale datasets. The predictions from the base models are then combined using a stacking strategy to improve overall classification performance. Experiments are conducted using the Deepfake Face Images Dataset containing real and synthetically generated facial images. The results show that the proposed stacking ensemble model achieves better performance than individual CNN models in terms of accuracy, precision, recall, and F1-score. Additional experiments under image distortions demonstrate that the proposed approach maintains stable performance and improved robustness in detecting manipulated facial images.

A Genetic Algorithm-Based Optimization of Keyboard Layouts Using Indonesian Language

Christopher Purnomo (Binus University)*
Melson Perawira (Binus University)
Rani Puspita (Binus University)

Abstract

QWERTY letter placement was not designed around Indonesian letter and digraph statistics, which can lead to unnecessary finger movement and inefficient transitions during prolonged text entry. Despite the scale of Indonesian daily computer use, there is still limited corpus driven work that proposes an Indonesian oriented physical layout candidate under standardizedergonomic criteria. In this study, we optimize an Indonesian oriented keyboard layout by combining corpus preprocessing, a constrained Genetic Algorithm search, and ergonomic simulation using Oxeylyzer. We use the Leipzig Indonesian mixed corpus and evaluate candidates on a fixed ANSI 3×10 geometry. To preserve comparability with standard keyboards, three punctuation keys, comma, period, and slash, are kept fixed, while the semicolon key is not pinned in this configuration, and the letter keys are optimized through permutation. The results show that the selected GA best layout substantially outperforms QWERTY on the main penalty indicators, reducing Finger Speed by 76.09 percent and Redirects by 43.34 percent, while increasing Alternates by 41.61 percent and Rolls by 14.61 percent. Supporting indicators remain favorable, including same-finger bigrams of 0.00857 and a lateral stretch (LSBS) penalty of 0.01398, suggesting the improvement is not achieved by introducing common ergonomic regressions.

Fusion of Trajectory-Based and Behavioral Features for Session-Replay Bot Detection in Mouse Dynamics

Devin Farrell (Bina Nusantara University)*
Navizar Lubis (Bina Nusantara University)
Stefanus Benhard (Bina Nusantara University)
Derwin Suhartono (Bina Nusantara University)

Abstract

Session-replay bots have evolved to bypass traditional security by mimicking biological mouse trajectories using complex physical friction and neuromotor algorithms. This paper proposes a multi-representation fusion architecture that combines deep spatial features from a VGG16 Convolutional Neural Network with 12 handcrafted 2D Euclidean kinematic features. Evaluated on the SapiMouse dataset across five classification categories including four levels of adversarial synthesis the proposed CNN-XGBoost model achieved an accuracy of 99.48% and an AUC-ROC exceeding 0.999. The results demonstrate almost perfect detection capabilities, effectively isolating advanced Level 4 neuromotor emulations that typically bypass traditional statistical filters. Furthermore, the architecture maintains a sub-second inference latency of 105.05 ms, providing a robust and computationally efficient defense for real-time behavioral authentication.

Closing the Explainability Gap: Bridging ML and Deep Learning for Diabetes Prediction Using SHAP-Consensus Feature Selection

Jordan Wijaya (Universitas Bina Nusantara)*
Pualam Nusantara (Universitas Bina Nusantara)

Abstract

Explainability has become a non-negotiable requirement for AI systems deployed in clinical settings, yet most existing work either explains classical ML models or evaluates deep learning performance rarely both, and almost never on the same large-scale dataset. This study addresses that gap by applying a SHAP-consensus feature selection framework to compare five models across the interpretability spectrum: Logistic Regression, Random Forest, SVM, TabNet, and a custom FT-Transformer implemented in TensorFlow/Keras All models are trained and evaluated on the BRFSS 2015 dataset (253,680 records), first on the full 22-feature set, then retrained on the 10 features selected through SHAP and LIME consensus scoring. To determine clinical alignment, model explanations are cross-validated against WHO and ADA diabetes risk criteria. All five models achieve similar accuracy (0.827–0.850), according to our findings, even though SHAP and LIME consistently recognize Age, BMI, and HighChol as dominant predictors that are completely congruent with published clinical data. The most deployment-ready model is logistic regression, which offers competitive accuracy, complete SHAP transparency, and little computing cost.

Cost sensitive ordinal classification for meningitis risk stratification considering clinical aspects to minimize false negatives

Chelsea Fakhira (Bina Nusantara University)*
Latasha Ivy Khoe (Bina Nusantara University)
Clarisse Xaviera Kurniaman (Bina Nusantara University)
Dani Suandi (Bina Nusantara University)

Abstract

Meningitis is a life-threatening disease where delaying the identification of high-risk patients lead to severe clinical outcomes and increased mortality. Conventional machine learning for risk prediction usually treat outcome categories as independent classes, which ignore the ordinal nature of clinical severity. As a result, these approach does not penalize significant misclassifications well, especially false negatives in high-risk cases. This study suggest a cost-sensitive ordinal classification system for meningitis risk stratification using clinical and demographic data in order to overcome this drawback. This framework integrate modeling that considers the ordinal nature of the target. A higher penalty is applied to high-risk misclassifications using a penalty matrix. We evaluate the result using a clinical penalty, a weighted scoring metric, Matthews Correlation Coefficient (MCC), and Mean Squared Error (MSE). Cost sensitive Random Forest is the best model with a weighted score of 0.834 and an MCC of 0.821 when compared to the baseline multinominal model, ordinal logistic regression model, and other cost sensitive models, including LightGBM, and XGBoost. The best model performance reduce the clinical penalty by 53.3% when compared to the ordinal logistic regression model. These findings underscore the value of incorporating ordinal structure and cost-sensitive learning into the model. This study propose a reliable framework for early intervention and helps hospital decision making.

Real-Time Coffee Bean Defect Detection on Android Using YOLO and TensorFlow Lite: A Comparative Study

Vinsensius Sutedja (BINUS University)*
Budi Juarto (BINUS University)
Yuda Sjaerodji (Binus University)

Abstract

Coffee bean quality is the key to product value which in turn is determined by the sorting process that in many small scale coffee industries is still dependent on manual inspection which is a cause of variation and human error. We present a study which developed a real time coffee bean defect detection system for Android devices which uses multiple YOLO based models and TensorFlow Lite. We trained and evaluated five light weight object detection models which are YOLov8n, YOLov9n, YOLov10n, YOLov11n, and YOLov26n using a coffee bean image set that we divided into defect and non defect categories. We looked at detection accuracy and real time performance after we deployed the models on a mobile device using INT8 and FP16 quantization. We found that all models did very well in terms of detection performance which had an average mAP50 of over 95% which is very good for classification. Also in terms of real time performance we saw that INT8 quantization gave us similar results across CPU and GPU which achieved about 10 to 12 FPS, and that FP16 quantization did better on GPU but not as well on CPU. We also noted that quantization strategy does in fact play a large role in how the system performs at deployment and that at a system level what we put in place also plays a large role in real time inference on mobile devices. We present that this approach does indeed work in putting in place effective and accurate coffee bean defect detection systems on mobile platforms that have limited resources.

Explainability in Multi-Relational Opinion Spam Detection: The Trade-Offs of XAI-Guided Graph Pruning

Mikhael Kevin Narendra Jayadharma (BINUS)*
Jason Evansaputra Sudargo (BINUS)
Prabowo Wahyu Sudarno (BINUS)
Anang Prasetyo (BINUS)
Fredy Purnomo (BINUS)

Abstract

Online reviews have a major role in shaping consumer behavior. The growing presence of fake reviews represents a serious threat to e-commerce platforms. Fraudsters have become more sophisticated, adopting strategies that mimic real user behavior. Graph Neural Networks still have a challenge to provide a solution to this problem. This study examines how Explainable AI (XAI) can be applied to active graph pruning within a CARE-GNN-inspired multi-relational architecture by using the heavily imbalanced YelpChi dataset. Two XAI approaches were compared: GNNExplainer, a local method, and PGExplainer, a global one. The results reveal a difference in how well each method preserves graph structure. The baseline model achieved a Spam F1-Score of 0.65. GNNExplainer managed to retain important relational patterns which delivered an F1-Score of 0.63 while keeping 49.1% of edges intact. On the other hand, PGExplainer, showed a performance drop that its score was only 0.58, due to extreme sparsity at just 11.7% edge retention which dissolved 137 complete spam clusters. We theoretically attribute this failure to mathematical regularization collapse within the global explainer, compounded by the 'Faithfulness Trap' inherent in optimizing XAI purely against baseline predictions rather than ground-truth labels. Even though inductive XAI remains computationally essential for large-scale streaming networks, our findings highlight the urgent need for future methods to integrate noise-robust and sound reasoning in order to prevent destructive graph pruning.

Automated Coloring of Ulos Motif Designs via the UlosColor-GAN

Samuel Situmeang (Institut Teknologi Del)*
Arlinta Christy Barus (Institut Teknologi Del)
Humasak Simanjuntak (Institut Teknologi Del)
Rachel Riris Rotua Damanik (Institut Teknologi Del)
Reinhard Hottua Simbolon (Institut Teknologi Del)
Reinhard Hottua Simbolon (Institut Teknologi Del)

Abstract

Ulos is a traditional woven textile of the Batak people. Ulos are difficult to design as skilled laborers must ensure that colors are aligned the right way. Ulos weavers usually select colors to match the design that has already been created. While AI-driven deep learning technologies have been able to automate the production of textiles, making sure that the original pattern structure is retained while utilizing colors consistent with the local culture remains challenging. In this paper, we developed UlosColor-GAN, a Conditional Generative Adversarial Network (cGAN) that would color Ulos patterns. The model implements a U-Net generator with skip connections to keep important details in the patterns, and a PatchGAN discriminator to mimic the texture. The model is trained using 153 images, which come from three major types of Ulos: Harungguan, Puca, and Sadum. The result indicated that the model is able to keep the pattern structure very well with the high Structural Similarity Index Measure (SSIM) of 0.9929. The average colorfulness score is 152.4561, which means that the model can produce strong traditional colors. When we put together the resulting designs and compared them with the original hand-colored designs, most of them looked pretty similar. This research is a solution to standardize the design process. It contributes to traditional weaving communities and enables digital heritage conservation.

Quantization Effects on Explainability in Indonesian Transformer-Based Sentiment Analysis

Helena Sanjaya (Binus University)*
Marvel Collin (Binus University)
Bertrand Tjahyadi (Binus University)
Derwin Suhartono (Binus University)

Abstract

Deploying transformer-based NLP models in resource-constrained settings almost always requires some form of compression. However, its impact on post-hoc explanations remains poorly understood. In this work, we study this gap using fine-tuned IndoBERT with three-class Indonesian sentiment classification. We compare post-training quantization (PTQ) and quantization-aware training (QAT) across FP16, INT8, and INT4, and evaluate explanations using five attribution methods from gradient-based and perturbation-based groups, each run over three random seeds. Accuracy is mostly unaffected, with all variants within 0.4 percentage points of each other, but attribution behavior shows greater disagreement across compression pathways rather than by bit-width. PTQ-INT8 maintains high stability (Spearman $\rho = 0.967$), while QAT at full precision already drops to $\rho = 0.788$, possibly due to fake quantization reshaping parts of the gradients during training, though the additional fine-tuning stage may also contribute. Our further findings complicate the understanding of explainability in compressed models. Gradient-based and perturbation-based methods show a near-zero cross-family correlation even at full precision, indicating that they capture fundamentally different aspects of the model and should not be combined into a single metric. Additionally, gradient attributions computed through the straight-through estimator on pre-export QAT checkpoints show almost no correlation with deployed ONNX attributions ($\rho \approx 0.10$). This is not an approximation problem but a model identity problem, where the checkpoint being explained is not the same as the model being deployed. Overall, these results suggest that accuracy is a poor proxy for explanation quality under compression, and that PTQ-INT8 is the most practically defensible configuration when interpretability requirements coexist with efficiency constraints.

Benchmarking LangChain4j Retrieval-Augmented Generation for Data Warehouse Querying on AdventureWorksDW2025

Elgan Kenlie (Universitas Bina Nusantara)*
Anderies Anderies (Universitas Bina Nusantara)

Abstract

Enterprise data warehouses store a large volume of key business information, but accessing such data requires understanding of the database schema and expertise in SQL, leaving out of reach for most users. Retrieval-Augmented Generation (RAG) emerged as a promising approach to this problem, but the accuracy of RAG for analytical purposes remains understudied. This study evaluates the accuracy of a production ready RAG based Text-to-SQL pipeline developed with Spring Boot, LangChain4j, GPT-4o mini, and BGE-Small-EN-v1.5. The pipeline is then evaluated on AdventureWorksDW2025 data warehouse using 80 natural language queries spanning across Easy, Medium, and Hard difficulty. The pipeline achieves Execution Accuracy score of 0.787 with 0.934, 0.833, and 0.520 for Easy, Medium, and Hard queries respectively. Retrieval analysis reveals a tradeoff between precision and recall where the F1 score reaches its maximum value at K=3. A larger context window leads to more noise that affects the quality of generated queries. The result suggest that current RAG based pipeline performs well on analytical tasks but still require human intervention in order to fulfill business requirement, motivating future work on multi-agent self-correction and temporal metadata encoding.

A Personalized Wardrobe Management Application with Attention-Based Outfit Recommendation

Tu-Nhu Le (Can Tho University)
Van Long Nguyen Huu (Can Tho University)
An Cong Tran (Can Tho University)*

Abstract

Personal wardrobe management and outfit selection are common daily tasks. However, most of the existing studies in fashion recommendation focus more on model development than on deployable application systems. This paper presents an intelligent personal wardrobe management application that integrates outfit recommendation into a mobile client–server architecture. The proposed system provides core wardrobe functions, including item management, outfit management, suggestion handling, and recommendation. The attention-based recommendation component of the proposed system combines item embeddings, attention-based aggregation, multi-layer perceptron prediction, and ranking-based retrieval to generate outfit candidates from the user's existing wardrobe under a specified purpose. For deployment, the best-performing configuration of the recommendation module was selected, achieving a mean squared error of 0.000062 and a mean cosine similarity of 0.984107 between the predicted and target outfit embeddings. The application is implemented using React Native on the client side and NodeJS on the server side. The evaluation shows that the proposed system can effectively connect structured wardrobe data and recommendation logic.

Voiced-Segment Demarcation and Translation-Invariant Scattering Networks for Parkinson's Disease Detection

Laura Marcella Pratama (Bina Nusantara University)
Andreas Edward Putra Jatmiko (Bina Nusantara University)
Simeon Yuda Prasetyo (Bina Nusantara University)*

Abstract

Parkinson's disease (PD), a central nervous system neurodegenerative disorder, which affects millions globally and its early diagnosis is difficult as it depends on subjective clinical evaluation. Many prior studies report high performance using segment-level cross-validation, which inadvertently allows samples from the same subject to appear in both training and testing sets. This leads to optimistic bias and limits real-world generalization. In this work, we present a meta-learner stacking ensemble framework that overcomes these limitations. The raw speech recordings of MDVR-KCL dataset consisting of spontaneous dialogue and read text tasks were processed with Voice Activity Detection (VAD) to extract voiced segments. The WST, with fixed scales J=6 and Q=8, is applied on every segment and produces translation- invariant multi-scale time-frequency representation which is stable to local temporal deformations and speaker variabilities. Leave-One-Subject-Out (LOSO) Group K-Fold cross-validation (K=5), with strict subject-level separation was employed to avoid inter-subject data leakage. An ablation study was performed among seven classifiers, all trained with Optuna a Bayesian hyperparameter search to find which base learners are the best candidates. The optimal stacking ensemble with CatBoost, LightGBM and Logistic Regression as base models and Multilayer Perceptron (MLP) as meta-learner achieved 81.80% in F1-score, 79.44% in sensitivity, 90.41% in specificity and 90.31% in AUC. These findings indicate that rigorous subject-level validation coupled with a heterogeneous ensemble provides a robust and clinically feasible approach for automated PD detection from speech.

Development of a Content-Based Career Recommendation Website from CV or Resume Using Transformer Models

Budi Juarto (Bina Nusantara University)*

Abstract

Higher - education graduates in Indonesia face considerable unemployment, reaching 11.28% in August 2024, reflecting an academic curriculum that's misaligned with current industry demands. Standard human resources practices also struggle with efficiency, as they require manual sifting through numerous CVs. Herein, we've created a content - based career recommender that utilizes natural language processing (NLP) to bridge this skill gap. Through sentence - similarity techniques such as cosine similarity, user preferences parsed from CV or manual forms were aligned with job specifications gathered in a collection of 75 careers. The process was structured as a waterfall methodology, back - end features were implemented in Python/Flask, the user interface in React. js and ModernBERT, a sentence - transformer, was chosen for feature learning to match user CVs or forms against existing job profiles in the career data. We selected a BERT - family transformer and conducted extensive evaluation against a wide spectrum of models. Ten sentence transformer models (namely, ModernBERT, RoBERTa, BERT - MiniLM, E5-Large, BGE - Large, GTR - T5-Large, CareerBERT, DistilBERT, JobBERT - Base and all - mpnet - base - v2) were benchmarking on a randomly sampled subset of 50 résumés from a publicly available set of résumés. ModernBERT outperformed all other transformer - based sentence - transformers across numerous relevant evaluation metrics such as the Mean Reciprocal Rank (0.848), Hit@1 (0.72), a perfect Hit@5 (1.0), NDCG@5 (0.887) and NDCG@10 (0.887), an average Top-1 cosine similarity score (0.862) and a reasonable average rank (1.36). Usability testing through the System Usability Scale, conducted on a cohort of 31 users, yielded an overall score of 79.68 (grade B - Good). Therefore, our system not only increases recruitment effectiveness but also reduces the effort required by individuals searching for work and provides a potential tool to help overcome educated unemployment in Indonesia.

A Probabilistic ECOC-SVM Framework for Multi-Class Land Suitability Assessment

Sitarani Safitri (Badan Riset dan Inovasi Nasional)*
Akhmad Riqqi (Institut Teknologi Bandung)
Albertus Deliar (Institut Teknologi Bandung)

Abstract

In theory, each location has suitability for all types of land use at different levels. Land suitability is derived from the relationship between land use/land cover (LULC) classes and land characteristics, including topography, meteorology, and soil properties. In this study, land suitability modeling was conducted using the Error-Correcting Output Codes Support Vector Machine (ECOC-SVM). SVM aims to determine optimal decision boundaries among land cover classes, while ECOC extends SVM for multi-class classification and enables the estimation of posterior probability values representing the level of suitability. The resulting posterior probabilities range from 0.00% to 99.70%, with the highest probabilities most frequently occurring in the range of 75% to 100% (76.47%). The overall accuracy of the model is 87.728%, with a kappa value of 0.852, indicating a very high agreement between the existing LULC and the modeled results. These findings demonstrate that the ECOC-SVM approach can effectively generate posterior probability values that represent land suitability levels for each spatial unit (grid). This approach is also applicable to other regions. However, potential random errors may occur in built-up areas, as their development is not always aligned with land suitability conditions.

An Evaluation of Quantized Large Language Models and Entailment-Based Reformulation for Sentiment Classification

Renaldy Fredyan (Bina Nusantara University)*
Luis Frentzen Salim (National Taiwan University of Science and Technology)
Khairunnisa (Universitas Muhammadiyah Bima)

Abstract

Sentiment analysis is a fundamental task in natural language processing for extracting opinions and emotions from textual data. This study examines two synergistic approaches to enhance sentiment classification: (1) quantized large language models (LLMs) for resource-efficient inference, and (2) entailment task reformulation for data-efficient learning. We evaluate classical models (Naive Bayes, SVM), transformer-based models (BERT, DistilBERT, RoBERTa), and LLMs (Llama-3.2 1B and 3B) using the Stanford Large Movie Review and Sentiment Labeled Sentences datasets. We use two post-training quantization methods, bitsandbytes (4-bit, 8-bit) and exllamav2 (2-bit, 3-bit), along with QLoRA fine-tuning, to compress LLM. Our findings indicate that the 2-bit quantized Llama-3.2-1B attains 93.4% accuracy while being roughly 11 times faster and 60% smaller than the full-precision 3B variant. With only 2,000 training samples, RoBERTa fine-tuned on the entailment formulation gets 84% accuracy, while standard classification only gets 54% accuracy under the same data-scarce conditions. This is a 30 percentage-point improvement. These results show that quantization and entailment reformulation are useful, complementary solutions for deploying sentiment analysis in resource-constrained and data-limited environments.

Attention-Constrained Mask-Aware Transformer For Intelligent Reconstruction Of Missing Biomedical semg Signals

Shamna AR (Musaliar College of Engineering)*

Abstract

The signals of Surface electromyography (sEMG) are prone to missing problems, including missing data caused by faulty sensors, movement artifact, and transmission noise, thus affecting the reliability of the biomedical application of the signal. As an attempt to solve such limitations, this study introduce Deep Learning-Based Biomedical Reconstruction (DeepBioRec) Framework which makes use of Temporal Feature Embedding, Mask-Aware Attention, and Transformer Model Hierarchies to achieve precise reconstruction of missing segments in the signal. In addition, the proposed technique makes explicit distinctions between the available and the missing signal parts, thus minimizing errors that may occur during reconstruction. Based on experimental results carried out using benchmark sEMG datasets, our algorithm performs with 13.7%, 11.6%, and 9.2% improvement over existing state-of-the-art techniques under SM, BM, and EM, 50% missing scenarios. Thus, the Deep Learning Based Biomedical Signal Recovery Framework has proved reliable and robust for biomedical signal recovery applications.

Evaluating Data Augmentation and Half-Precision Efficiency in Transformer-based Hate Speech Classification

Harry Santosa (Binus University)
Garent Ecklesia (Binus University)*
Derwin Suhartono (Binus University)

Abstract

The fast spread of hate speech on social media outpaces human moderation capacity, making automated, computationally efficient solutions a necessity. This paper investigates two strategies for improving Transformer-based hate speech classifiers which are augmenting data through a 4-bit quantized Mistral-7B Large Language Model, and cutting inference costs through half-precision (FP16) casting. Four Transformer architectures were each fine-tuned on a combined corpus of 35,998 labelled samples. Because human annotators disagreed on 17.21% of individual labels, we decided to use soft-label evaluation grounded in the Law of Total Expectation rather than treating any single majority vote as ground truth. Our experiments show that LLM-generated augmentation is more effective than conventional back-translation. The best three-model ensemble that were trained on a fully augmented split, reached a macro F1 of 88.56% and 0.9026 Accuracy, clearly exceeding the estimated 82.79% ceiling implied by human consensus. More than that, a probabilistic model trained on soft voting distributions matched annotator uncertainty with a Mean Squared Error of only 0.0381, showing that the model is well-calibrated rather than just predicting the majority class. On the efficiency side, the best single model was cast with FP16 precision reduced the model size by 50% and increased throughput from 138.13 to 632.12 tweets per second with no drop in F1 macro. These results show that we can build a content filtering system that is accurate and fast enough to handle the massive traffic of live social media.

Comparative Analysis of Feature Extraction Strategies for Atrial Fibrillation Detection Using ECG Signals and Machine Learning

Vittorio DinataVittorio Dinata (Bina Nusantara University)*
Jefferson Louis (Bina Nusantara University)
Wilbert Leung (Bina Nusantara University)
Setiawan Joddy (Bina Nusantara University)
Jurike Moniaga (Bina Nusantara University)

Abstract

An automated atrial fibrillation (AF) detection from ECG signals still remains challenged by the lack of systematic evaluation of feature domain contributions. This study will comparatively evaluate three HRV feature domains, such as time-domain, frequency-domain, and nonlinear, with a total of 63 features extracted from the MIT-BIH AFDB. These features will later be evaluated across three classifiers (Random Forest, SVM, and Logistic Regression) under Leave-One-Subject-Out cross-validation. The combined multi-domain feature set consistently outperformed any single-domain subsets, with the result of F1 = 0.9618 and AUC = 0.9939. While multi-domain features outperform any single domain feature, Time-domain features performed the strongest among any single domain feature. Lastly, frequency-domain features remain underperformed. Results confirm that feature representation quality outweighs classifier complexity in ECG-based AF detection.

A Multi-Objective Reinforcement Learning Framework for Antibiotic Selection Using Drug–Pathogen Interaction

Fathima Rumaiza SM (Vels Institute of Science, Technology & Advanced Studies)*
Kamalakkannan S (Vels Institute of Science, Technology & Advanced Studies)

Abstract

Antimicrobial resistance (AMR) poses a serious threat in contemporary healthcare by compromising the efficacy of antibiotic drugs and making clinical decision-making difficult. Conventional machine learning techniques concentrate mainly on predicting resistance, without providing assistance in selecting an appropriate treatment option. Additionally, current techniques generally fail to consider the relationship between antibiotics and pathogens, which has a considerable impact on the effectiveness of treatment. This paper introduces a novel multi-objective reinforcement learning framework for antibiotic selection called MO-RL-AMR, which considers the interaction between antibiotics and pathogens in the treatment decision-making process. This work views the recommendation of antibiotics as a sequential decision-making problem, in which the agent learns to determine optimal treatment protocols, depending on the health condition of the patient and the particular pathogen involved. The reward function of the multi-objective reinforcement learning algorithm is optimized to ensure treatment efficacy, reduced resistance development, and affordability. Experimental findings reveal that the proposed method enhances treatment success, decreases resistance, and maintains reasonable costs compared to the benchmark algorithms.

Adaptive AI-Based Models for Detecting and Mitigating Evolving Threats in Cloud-Hosted E-Commerce Platforms

Balaji VISTAS (Vels Institute of Science Technology and Advanced Studies)*

Abstract

The high pace in the development of cloud-based e-commerce applications has greatly revolutionized the world through online trading through the provision of scalability, flexibility, and cost-effectiveness. Even then, this development has come with sophisticated and ever-changing cyber threats such as fraud, data breaches, account hijackings, distributed denial-of-service attacks, and other complex malware-related intrusions. The existing traditional rule-based and static security devices are not fitted to deal with these dynamic threat environments because of their low adaptability and slow response rates. The proposed study presents a threat detection and mitigation model based on adaptive artificial intelligence (AI) that is implemented in the context of cloud-hosted e-commerce infrastructure. The model suggested combines machine learning and deep learning methods to process real-time transactional, behavioral, and network-level data, helping to detect anomalous trends early and before actual threats occur. Experiment-based assessments prove that the adaptive AI model is more highly detected, has fewer false positives, and lower response times than traditional security techniques. Although the study is effective, there are practical limitations associated with data availability, model interpretability, computational overhead, and difficulty of deployment identified. The research focus in the future involves explainable AI, federated learning on privacy concessions, and autonomous association of response to increase the strength and reliability of deployments in the real world.

SAD-HSA: A Secure Framework for Detecting Malicious Client Updates in Federated Learning through Statistical Anomaly Detection and Hybrid Aggregation

Jayalakshmi Vaithiyanathan (Vels Institute of Science Technology and Advanced Studies,Chennai)*

Abstract

Federated Learning (FL) has developed as a transformative paradigm in distributed machine learning by supporting multiple clients to collaboratively train models without sharing raw data to preserve data privacy. However, FL systems are particularly prone to data poisoning attacks and byzantine failures. The malicious participants in FL deliberately influence model updates to reduce the total model performance. Detecting such adversarial behaviour is challenging because FL servers indirectly access client data. To address these challenges, this paper proposes a novel framework, SAD-HSA (Statistical Anomaly Detection with Hybrid Secure Aggregation), designed to enhance robustness and reliability in adversarial federated environments The proposed framework combines three robust aggregation techniques namely Krum, Covariance based Mean and Median aggregation to effectively detect and isolates the malicious client updates which enhances the anomaly detection accuracy and decreases the effect of poisoned gradients during global model aggregation. Experimental results demonstrate that the proposed SAD-HSA framework significantly outperforms baseline federated Learning models Krum and FedAvg in the presence of malicious clients and making it suitable for privacy sensitive distributed applications such as healthcare, smart IoT networks and finance.

Grey wolf–Firefly algorithm-based feature selection with Support Vector Machine for Pulmonary Tuberculosis detection from Chest X-ray images

Parameswari Ramachandran (Vels Institute of Science,Technology and Advanced Studies)*
Jyothilakshmi K N (Vels Institute of Science,Technology and Advanced Studies)

Abstract

In this study, a computational framework is proposed using a hybrid feature selection approach, by combining Grey Wolf Optimization (GWO) and Firefly Algorithm (FA) to optimize feature subsets for the detection of pulmonary tuberculosis (TB) from chest X-ray images. Given the computational cost of high dimensional data in medical image analysis, we extracted 794 features using a combination of deep learning (ResNet18) and multi-dimensional handcrafted descriptors (LBP, GLCM, GLRLM, and Shape). The GWFA algorithm was used to aggressively reduce the feature dimensionality, trying to retain the variance that distinguishes the classes. The optimal feature vector was classified with a Support Vector Machine (SVM). The GWFA+SVM model was evaluated on a strictly hold-out dataset of 1200 cases and achieved an accuracy of 79.17%, precision of 77.42%, recall of 81.36%, F1-score of 79.34%, and MCC of 0.5843. The presented results show that hybrid swarm intelligence can provide efficient dimensionality reduction as a computationally feasible option for resource-limited settings, but also highlight a critical trade-off between feature pruning and maximum discriminative power.

AI-Based Automated Data Quality Management System for Noise Detection in Numerical Data Pipelines

Akila A (Vels Institute of Science, Technology and Advanced Studies)*
Priya Dharshini V (Vels Institute of Science, Technology and Advanced Studies)

Abstract

Getting data right matters more than ever when numbers drive systems. When odd readings slip through - say from sensors, money flows, or daily logs - they can twist later choices and predictions. Old ways of spotting weirdness lean on fixed rules or one-size-fits-all algorithms. These methods often cry wolf, misfire across different data sets, and offer little clue why they flagged anything at all. A new approach tackles weak spots in data monitoring by combining three methods: Isolation Forest, Local Outlier Factor, and One-Class SVM work together. An agreement happens when two or more tag a reading as off track, then it gets marked strange. To show why decisions happen, SHAP tools step in – Tree Explainer handles one model while Kernel Explainer covers the rest, giving clear input reasons behind each odd result spotted. What sets this apart? It guesses how messy the data is by itself, thanks to Interquartile Range math. On top of that comes a triple-layer rating tracking clean signals, full records, duplicates, everything viewable via Streamlit, smooth enough for non-experts. Tests run on financial files about credit fitness reveal fewer false flags now - down from 123 to 106 real ones caught - with final trust level hitting 90.6%. Digging deeper showed Credit History plus Loan Amount Term kept standing out during reviews, explaining what triggered alerts.

Integrated Hybrid CNN–LSTM Framework for Multi-Omics-Based Prediction of Salinity-Responsive Genes in Plants

Nikitha V (VISTAS)*

Abstract

Abiotic stresses such as soil salinity significantly restrict agricultural productivity and growth of plants. It is necessary to identify major genes that are associated with salt tolerance to create stress-resilient crop varieties. This paper will present a deep learning (DL) based and bioinformatics combined framework to determine potential salt tolerance genes in plants. GEO2R was used to analyse publicly available datasets of halophytic gene expression to identify genes that were differentially expressed in response to salt stress. The analysis was performed on two Arabidopsis thaliana datasets GSE16765 and GSE5620 due to their relevance and quality of data. The top ten candidate genes were shortlisted out of the differentially expressed genes based on the statistical significance and expression patterns. DL model consisting of a convolutional neural network (CNN) was then used to further analyse these genes to determine their contribution towards salt tolerance. Protein-protein interaction analysis was used to perform network-based validation, and hub genes were identified with the CytoHubba plugin. DL analysis results were checked against network centrality measures to provide consistency and a biological value. The suggested methodology effectively reveals one of the important genes related to salt tolerance, proving the usefulness of the combination of GEO-based analysis, DL models, and network biology. This research offers a powerful computational framework to rank salt tolerance genes and may be applied in future functional and experimental validation research.

A Decade of Augmented Reality Integration in Malaysian TVET: Adoption, Practices, and Challenges

IRMA SYARLINA CHE ILIAS (UNIVERSITI KUALA LUMPUR)*
Muslihah Wook (National Defence University of Malaysia)
Norazimah Awang (UNIVERSITI KUALA LUMPUR)
Muhammad Ezra Ismail (UNIVERSITI KUALA LUMPUR)

Abstract

This study evaluates the use of Augmented Reality (AR) in Malaysian Technical and Vocational Education and Training (TVET). A systematic search on Google Scholar, Scopus, and Mendeley resulted in 99 articles, of which 27 peer-reviewed articles (published between 2017 and 2025) were included after screening, with thematic coding (using ATLAS.ti) on dimensions of Tool Development, Student Acceptance, Pedagogical Gap, Systematic Design, and Systemic Barrier. The study shows that AR in TVET is dominated by context-specific prototype development, most using the ADDIE model, with evaluation done primarily using expert usability/learner perception methods, with only a small number using pre-post or quasi-experimental methods to demonstrate objective learning outcomes. Although growth in output is seen after 2020, integration, scaling, and long-term evaluation are constrained by instructors, devices, connectivity, curriculum, SOP, and funding. The study recommends competency-based evaluation and adoption pathways to move AR from pilot to reproducible, systemic, and pedagogical use in Malaysian TVET.

Automated Personal Protective Equipment Detection Using Convolutional Neural Networks for Laboratory Safety Monitoring

Bustami Bustami (Universitas Syiah Kuala)*
Yuwaldi Away (Universitas Syiah Kuala)
Husni Husin (Universitas Syiah Kuala)
Asri Asri (Universitas Malikussaleh)
Muhammad Fikry (Universitas Malikussaleh)

Abstract

Personal protective equipment (PPE), especially that related to the laboratory, is a must when it comes to workplace safety. The psychological pressure in an environment where people are exposed to harmful products and equipment and where safety can potentially be at serious risk, can be a formidable burden for workers addition to the physical toll. Because there is no real-time, ongoing approach to monitoring which can be continuously validated, this monitoring system has therefore been a subjective and inconsistent method of validating safety over time. Recent developments in deep learning and computer vision have enabled the discovery of suitable locations for automated safety monitoring systems. In this study, we introduced a Convolutional Neural Network (CNN) model to automatically identify PPE in a laboratory context. Eleven individuals from a wide variety of 10 PPE categories including safety helmet, welding helmet, safety glasses, ear protection, safety gloves, low shoe, safety boot, hair cap, respirators, and mask were enrolled in a standalone, consisting of 1,254 raw photos. Data: There were 61.6% multi-PPE mix photographs and 38.4% original single-PPE images, reflecting the actual laboratory scenario, the four viewpoints (front, left, right and back). Based on the dataset, 12 data augmentation strategies, including 16,302 photos of images were applied to increase accuracy of the model and to prevent overfitting. To extract hierarchical visual information, we proposed a three-block CNN architecture with 32 filters, 64 filters, and 128 filters used per convolutional layer to achieve image resolution scaling to 64x64. Based on the experimental data, the accuracy of our model tested was 93.42%, that is more than precision (90.88-96.71%), recall (89.76-95.89%) and F1_scores (90.32-96.30%) of the ten class. Low-safety shoes show the largest detection deficit (90.32%), while welding helmet class has a higher F1-score (96.30%). Moreover, we can see that performance over unknown photos shows that 78.3% of all forecasts result in confidence values greater than 90%, meaning that the generalizability is very sufficient, and 86.67% of all predictions surpass the 75% confidence level. Data augmentation reduced the overfitting gap by 4.16 percentage points and increased performance against this benchmark by 10.49% (from 82.93% to 93.42%). These results demonstrate that CNN-based vision systems are sophisticated approaches for automated safety checks in laboratories. By conducting this research, data can be transformed into multimodal sensors to create a future intelligent safety management system. The flattening layer's 4,608-dimensional feature vector is designed specifically for multimodal fusion with these environmental sensors.

Constructing a Performance Report System as a Profiling Tool to Measure Game Object Spawning Performance on Unreal Engine 5

Daniel Wahyudi (Binus University)*
Andry Chowanda (Binus University)

Abstract

The optimization topics of video games become the most important part for the players. They always evaluate the video game based on the smoothness while it is running without waiting for the processing time. During development, sometimes the development team found a performance issue and needed a profiling tool to profile the spawn system. Unreal insight, which was included in the engine,does not explicitly give the memory (RAM) consumption and the result of exported CSV file return list of called function. The performance report system is a runtime diagnostics model embedded within the game that is an Unreal Engine Project. This system was designed as a profiling tool to measure the hardware load during the spawning object in a game session. The performance report system successfully captured the key parameters of real time workload of the system such as frame rate (FPS), frame time (ms), CPU usage as a percentage, GPU time (ms), and memory (RAM) consumption. The system was evaluated using a high intensity stress test that involved both projectile and particle spawning. Projectile spawning causes significant bottlenecks on the GPU at the 1,500 until 2,000 object threshold while particle spawn triggers simultaneous bottleneck on CPU and GPU with extreme memory consumption. This research utilizes Savitzky-Golay filtering and mean aggregation to isolate the performance trend from noise. Overall, the results demonstrate a clear inverse relationship between the number of active entities in a scene and the overall stability of the system, successfully identifying specific bottlenecks in memory allocation and rendering pipeline.

Mobile Application for Strawberry Plant Disease Classification Using CNN-Based Leaf Image Analysis

Hernawati Samosir (Institut Teknologi Del)*
Samuel Sibuea (Institut Teknologi Del)
Yoseplin Hutauruk (Institut Teknologi Del)
Novrael Marbun (Institut Teknologi Del)
Michael Joseph Christian Situ (Institut Teknologi Del)

Abstract

Strawberry plant diseases remain a major challenge to agricultural productivity, as delays in detection and improper treatment can result in significant yield losses. In this study, a mobile application was developed to classify strawberry plant diseases from leaf images using CNN. The application is designed to help farmers detect diseases quickly and accurately using widely available mobile devices. The development process began with collecting a dataset of strawberry leaf images showing various disease conditions. This dataset was then pre-processed and used to train a CNN model, which is known for its effectiveness in pattern recognition and image classification. The trained model showed strong performance in identifying and distinguishing different types of diseases. In addition, the application features a user-friendly interface that allows users to capture leaf images in real time and obtain diagnostic results within a short time. The experimental results suggest that the system can provide fast and reasonably accurate predictions for detecting strawberry plant diseases, making it a potentially useful tool for farmers in managing their crops by using 200 test images data evaluation (40 images per class) with accuracy results reaching 94%. Overall, this research helps increase the use of artificial intelligence in agriculture, especially in supporting more sustainable strawberry production through early and efficient disease detection.

Adaptive SNR-Aware Zero-DCE For Domain Generalizable Low-Light Image Enhancement (LLIE)

Rudy Hartanto (Bina Nusantara University)*
Francesco Sebastian (Bina Nusantara University)
Anderies Anderies (Bina Nusantara University)

Abstract

Low-Light Image Enhancement (LLIE) is one of the most researched topic in computer vision. In LLIE, models often struggles with color distortion, noise and poor performance when tested on real-world scenarios. In this research we focus on the Zero-DCE model, this framework offers a zero-reference approach. Though good, this model is limited by the lack of an SNR-Aware module and Domain Generalization capability for unseen dataset. To address this gap, this study is proposing a Domain Generalizable LLIE Framework based on the Zero-DCE, with an SNR-Aware module. This new SNR-Aware module will estimate local noise and prevent excessive enhancement on dark regions which produces high noise. Other than the new SNR-Aware module, we also are applying a Domain Generalization based on Mixed-up based DG which will make the model maintain a stable performance on unseen dataset. The framework's effectiveness was validated through metrics like PSNR, SSIM, MAE and BRISQUE for unseen dataset. Upon testing with the LOL Dataset we achieved a result of 18.317 dB PSNR and 0.7244 SSIM, these results outperforms the baseline Zero-DCE model. Further testing shows that, in an unseen dataset like DarkFace it is able to generate a lower BRISQUE score than the baseline Zero-DCE model. It achieved a score of 17.9020, this confirms the ability to perform in unseen dataset thus providing a strong solution for low-light image enhancement models.

Detecting Indonesian Illegal Online Gambling Site Advertisements in Social Media Images: A Comparative Study

Anastasia Putri Hartanti (BINUS University)*
Gregorius Natanael Elwirehardja (BINUS University)

Abstract

The growth of social media opens the door for gambling site operators to reach more potential gamblers. With the current number of 8.8 million illegal online gamblers in Indonesia, it is necessary to build an automated system that could identify gambling advertisements and stop their spread. This study compares 11 different models with three different machine learning approaches: OCR-extracted text-based, image-based, and multimodal approaches that integrate both textual and image features. This research also accounted for the resource usage that is consumed per model in order to generate a prediction, mainly GPU utilization. The models were trained using social media advertisement images for both gambling and non-gambling content, which is actively being used on social media. Experimental results show that the multimodal approaches, IndoBERT + EfficientNetB0 and IndoBERT + ResNet50, all achieved 99.76% in recall score. The average resource usage across all evaluated models shows that models using OCR, specifically PaddleOCR (as this study only employs PaddleOCR), utilize significantly more resources, particularly GPU memory. The model with the best recall to inference time and recall to peak GPU utilization ratio was ShuffleNetV2 0.5x, the lightest model in this study, which achieved a recall score of 94.76% while maintaining a significantly low resource utilization. This study shows how using unimodal image-based classifiers is already sufficient for identifying gambling advertisements, as adding text features increases complexity mainly in OCR text extraction, despite the near-perfect AUC.

Application of AI for Recipe Recommendations Using Images of Food Ingredients

Daniel Kurniawan (Bina Nusantara University)*
Puti Andam Suri (Bina Nusantara University)
Sani Muhammad Isa (Bina Nusantara University)

Abstract

Cooking at home is a common practice to reduce daily expenses, but selecting appropriate menus based on available ingredients remains challenging and can lead to food waste. This study proposes an integrated food recommendation system that generates recipe suggestions from food images by considering both ingredients and quantities. The system combines object detection, quantity estimation, and content-based recommendation. Three object detection models, namely YOLOv5, YOLOv8, and YOLOv11, were evaluated to determine the most suitable model for ingredient detection. YOLOv5 achieved the best performance with a precision of 0.873 and an mAP of 0.825, and was selected as the detection model used in the recommendation system. For recommendations, TF-IDF and the Quantity-Aware approach were compared using Precision@5 and Mean Reciprocal Rank (MRR). The TF-IDF approach achieved better overall performance with a Precision@5 of 72.0% and an MRR of 0.9 compared to the Quantity Aware approach. The proposed system demonstrates the feasibility of integrating ingredient detection and quantity estimation with recommendations to support menu selection and reduce food waste.

Segmentation of Liver and Tumor Regions in CT Scan Images Using U-Net

Ercent Tannius (BINUS)*
Jos Hwen Sim (BINUS)
Nicholas Lee (BINUS)
Jurike V. Moniaga (BINUS)
Setiawan Joddy (BINUS)

Abstract

Liver cancer is one of the most deadly cancers worldwide. For targeted treatment plans, tumor and liver edges need to be characterized and identified on abdominal CT scans. Manual segmentation is too slow, subjective, and inconsistent for large-scale use; therefore, developing segmentation algorithms using deep learning is not an option but a necessity.In this study, three architectures : SimpleCNN, SimpleSegNet, and U-Net were tested for developing segmentation algorithms based on the LiTS dataset. From 51 NIfTI volumes, 6,802 slices were sampled, with the training dataset containing 5,441 slices and the validation dataset containing 1,361 slices, using foreground pixel stratification. A combined Dice-BCE loss function was employed to address class imbalance, and the Adam optimizer with a learning rate of 1×10^{-4} was used for 30 epochs on an NVIDIA Tesla T4 GPU.The results obtained demonstrate the excellent performance of U-Net with an accuracy of 0.998155, a Dice score of 0.986047, and a Jaccard score of 0.972478. For the SimpleCNN and SimpleSegNet architectures, Dice scores of 0.714132 and 0.720601 were observed, both considerably lower than U-Net. These results confirm that U-Net, combined with a rigorous preprocessing pipeline and an appropriate loss function, serves as a reliable and reproducible baseline for liver and tumor segmentation in medical imaging.

Integration of Decentralized Identity (DID) and Digital Signature for Blockchain-Based Digital Copyright Storage

Andrew Colin (Bina Nusantara University)
Jason Tandiari (Bina Nusantara University)*
Yehezkiel Ferdinand (Bina Nusantara University)
Aditya Kurniawan (Bina Nusantara University)
Franz Junior (Bina Nusantara University)

Abstract

The growing of economy digital has increase production and distribution of digital works. but also make it worse the problem of copyright violation like content thief, illegal distribution, and fake ownership, the system of centralized copyright protection that still limited because of the slow registration. vulnerability of manipulation data and weak verification identity. Even though the blockchain technology give you the evidence of the real data that can't be change and transparent, this technology can not verify the real identity in behind of wallet address, so there's a gap between technical ownership and legal authorship. this research propose the integrity model that combine digital signature and decentralization identity (DID) in the work framework based in blockchain to overcome this limitations. digital signature ensure the real and integrity data, while DID give the layer identity that can be verify by cryptographic and protect the privacy dan decentralized. This system can be implemented by protoype that combine the resolution DID, creation of digital signature, and smart contract ERC-721 that enriched by specific metadata of copyright. the functional evaluation, and regulation show us that the model proposed identity authenticity, to strengthen validity ownership claim law. and aligned with requirements of Indonesian law for digital signature and digital document. integrity of DID and digital signature offer safety approach, verified and reliable by law for protection digital copyright.

AUTOMATIC ANALYSIS OF TERMS OF SERVICE: IDENTIFYING DANGEROUS CLAUSES BASED ON MACHINE LEARNING

Keitaro Alexander Herman (Bina Nusantara)*
Jeslyn Tania (Bina Nusantara)
Melvern Amadio Hidayat (Bina Nusantara)
Ade Putera Kemala (Bina Nusantara)

Abstract

In the digital era, users are required to agree to Terms and Conditions (T&C) or Terms of Service (ToS) documents before accessing digital platforms. However, these documents are often lengthy and written in complex legal language, leading to low user awareness and understanding. This study proposes an automated system based on Natural Language Processing (NLP) to identify and highlight important and potentially problematic clauses in T&C and ToS documents. The system leverages pre-trained transformer models, namely Sentence-BERT (SBERT), DistilBERT, and Legal-BERT, and applies an ensemble-based approach for sentence level classification on English legal texts. The dataset consists of T&C and ToS documents collected from various digital platforms, with problematic clauses manually annotated. Experimental results show that the proposed ensemble model achieves an accuracy of 81.82% and an F1-macro score of 0.6943, indicating balanced performance across clause categories. These findings demonstrate that transformer-based NLP models are effective in detecting important clauses within legal documents. The proposed system presents analysis results in a user-friendly format, enabling users without legal backgrounds to better understand critical information in digital legal agreements.

Mapping Stunting Prevalence: A Digital Geospatial Approach for Sustainable Health Monitoring in Nusa Tenggara Timur

Andrea Stevens Karnyoto (Bina Nusantara University)*
Meiton Boru (Universitas Nusa Cendana)
Dony Martinus Sihotang (Universitas Nusa Cendana)
Nelci Dessi Rumlaklak (Universitas Nusa Cendana)
Bertha Selviana Djahi (Universitas Nusa Cendana)
Najia Al Umri (Bina Nusantara University)
Bens Pardamean (Bina Nusantara University)

Abstract

This study aims to develop a geographical information system (GIS) to monitor the stunting distribution in Kupang City. We used the waterfall method in this research. The GIS displays the distribution of stunting in polygons at the district and village levels. The End User Computing Satisfaction (EUCS) test was used, with a user satisfaction rate of 82.31%, indicating high satisfaction. In the EUCS test, users expressed satisfaction with the content (82.1%), accuracy (81.8%), form (83.6%), convenience (81.4%), and timeliness (82.2%). This study can assist the Government of Kupang City in obtaining visual information through a digital stunting map. Furthermore, we significantly contribute to the literature on the spatial distribution of stunting in Kupang City. The system can be further enhanced by integrating real-time data collection from health posts. Such improvements will enable more timely interventions and better resource allocation to address stunting.

Data-Driven Volcanic Disaster Mitigation in Indonesia

Benedictus Lukito (Universitas Bina Nusantara)*
Jason Loderick (Universitas Bina Nusantara)
Rezki Yunanda (Universitas Bina Nusantara)

Abstract

This research aims to explore unsupervised clustering techniques that can be used to cluster Indonesian volcanoes based on their activity profiles, thus helping in disaster mitigation by monitoring and prioritizing disaster preparedness. MAGMA activity reports from 28 June 2025 to 6 March 2026 were used to aggregate report-level data into volcano-level data by calculating daily rates and peaks of structured seismic events, including eruptive events, shallow/deep volcanic earthquakes, and tectonic earthquakes, and alert level data. Three clustering algorithms, K-Means, DBSCAN, and Gaussian Mixture Model, were used and compared based on internal cluster validation, including Silhouette, Calinski-Harabasz, and Davies-Bouldin, and usability. The K-Means algorithm was identified as the best clustering algorithm for clustering Indonesian volcanoes, achieving the best separation and grouping all volcanoes into clusters. DBSCAN was able to cluster volcanoes into compact clusters but identified a large number of volcanoes as noise, thus more suitable for anomaly detection than clustering all volcanoes. GMM was identified as having poor separability for clustering. The best K-Means clustering algorithm, with $k = 3$, identified three groups of volcanoes: a baseline cluster of volcanoes with low activity, a cluster of volcanoes with high unrest activity driven by high rates of seismic/eruptive events and high daily peak counts, and a cluster of volcanoes with high priority activity driven by high alert levels.

Identification of Key Microbial Biomarkers for Colorectal Disease Prediction through Univariate Feature Selection

Linh Thuy Thi Pham (Can Tho University of technology)
Hat Nguyen Le (Can Tho University)
Thai T. Vo Nguyen (FPT University)
Bao Gia Tran (FPT University)
Ngoc Huynh Pham (Can Tho University)
Thanh-Hai Nguyen (Can Tho University)*

Abstract

Metagenomic data analysis has emerged as a promising avenue for investigating the relationship between gut microbial communities and colorectal disease; however, the high dimensionality of such data presents substantial challenges for computational modelling. In this study, we propose a univariate feature selection framework based on the SelectKBest algorithm with ANOVA F-statistics scoring to identify the most discriminative microbial biomarkers for colorectal disease prediction. Experiments were conducted on four geographically diverse cohorts — Vogtmann (American), Feng (Austrian), Yu (Chinese), and Zeller (German/French) — comprising a total of 1,988 bacterial features. To address data quality issues, standardisation-based preprocessing was applied prior to feature extraction. Feature subsets of incrementally increasing size (5, 10, 15, …, 50) were systematically evaluated using a Random Forest classifier, with performance assessed via accuracy, F1-score, and confusion matrix analysis. Results demonstrate that a compact subset of 25 features achieves classification accuracy comparable to or exceeding that of the full feature set across all four cohorts, with notable discriminative taxa including Bifidobacterium, Corynebacterium, Riemerella anatipestifer, and Chlamydia abortus. The proposed method outperforms established baseline approaches including LIME-based selection and random feature selection, underscoring the effectiveness of univariate statistical filtering for dimensionality reduction in metagenomic disease prediction tasks.

A Boosting MSME Growth with AI-Powered Social Media Analytics: A New Era for Guerrilla Marketing

Marvelous Nathanael (Bina Nusantara University)*
Murniati Murniati (Bina Nusantara University)

Abstract

Amid increasingly fierce competition, MSMEs need a boost that doesn't require significant costs but yields substantial results. With the rapid advancement of technology, the creative integration of artificial intelligence into promotional strategies offers a significant solution for the growth of MSMEs. Using a quantitative approach based on data collected from SMEs actively engaged across various social media platforms, the analysis employed Structural Equation Modeling to identify relationships between analytical capabilities, creative marketing tactics, consumer interactions, and business performance. The research findings indicate that AI-based analytics play a significant role in enhancing real-time marketing outcomes that engage audiences, delivering personalized content, and driving data-driven decision-making, all of which fundamentally foster stronger engagement. These findings align was consistent with the majority of academic research reviews on the role function of AI in enhancing brand recognition and long-term data-driven revenue measurement. Beyond mere confirmation, this study offers a digital framework that positions AI-powered analytics as a practical growth tool for businesses that has limited resources.

Data Driven Tower Placement Optimization Using Spatial-MINLP

Angga Ramadhan (University of Indonesia)*
Arian Dhini (University of Indonesia)

Abstract

Achieving reliable wireless coverage with the lowest possible deployment cost remains a major issue in modern telecommunications planning. Impressive increase of devices connected to the internet and services run on top of these, but still having a huge urban-rural digital gap to overcome while operators now have the needy as well as excessive cost-effective network extension strategies. Traditional placement algorithms, which typically rely on heuristic or rule-based solutions, do not leverage the spatial variation of demand, population distribution and propagation conditions to their full extent resulting in unequal service delivery across the network footprint and suboptimal capital allocation. The study will propose a data-driven Mixed-Integer Nonlinear Programming (MINLP) which optimizes cellular base station placement by minimizing total deployment cost while ensuring full population coverage and signal quality thresholds simultaneously. This framework embeds geospatial and demographic data using a Geohash Level 7 spatial templating, bringing alignment of population distributions, and crowdsourced RSRP measurements along with both build-to-suit (CAPEX) and collocation (OPEX) deployment costs into a single optimization problem. Forcing the constraint of signal quality using a log-distance propagation model with a minimum RSRP threshold of −95 dBm is formulated via Big-M Constraints. Three coverage radius settings (1 km, 2 km and 3 km) correspond to micro, standard and rural macro tower deployments respectively, so that capacity constraints can be applied at large radii to prevent service overload. We validate this model across a large-scale deployment scenario deployed on Java area and Kalimantan area, both of Indonesia which have different population density and backbone network maturity. The optimization achieved 5,976 towers serving 25.2M people at an approximate cost of USD 361.5 million USD and collocation representing approximately 58.7% of the deployments. The findings show that the model achieves spatial optimization of resources without loss of service quality by adaptively setting smaller coverage radii in areas with high urban density and a larger radius for sparse rural regions. The proposed methodology allows for evidence-based scalable and reproducible telecommunication network designs service quality but still maintaining effective and efficient investment.

Hormonal Imbalances: Analyzing its Role on the Menstrual Cycle in Women

Stanislaus Wirianto (Bina Nusantara University)
Sesilia Dewi (Bina Nusantara University)
Louisa Halim (Bina Nusantara University)
Nunung Qomariyah (Bina Nusantara University)
Feri Setiawan (Bina Nusantara University)*

Abstract

Hormonal fluctuations influences mood swings, menstrual irregularities, and the overall health of young women.
Adversely, individualized, data-based analysis of the pattern of hormones has not been thoroughly examined. This study investigates the connection between specific hormonal factors, menstrual cycle phases, and symptom patterns using retrospective menstrual health log records of 50 participants. Quantitative methods were adopted for statistical analysis using Pearson correlation, one-way ANOVA, chi-square tests, and K-means clustering techniques. Results reveal that there is a statistically significant phase-specific change in luteinizing hormone (LH) and estrogen levels ($p < 0.05$). Hormonal imbalance can therefore be defined as a relative state within the menstrual period of the individual. Elevated estrogen without a corresponding LH surge was associated with increased symptom severity, whereas the ovulatory LH surge correlated with reduced discomfort. Inter-symptom correlations, particularly between stress and headaches ($r = 0.91$), were substantially stronger than hormone-to-symptom correlations. The results point at the complexity of the mechanisms underlying the menstrual syndrome manifestations and underline the significance of employing a phase-aware multivariate approach in studying the issues related to menstrual health outcomes.

Indo-UnCRtainTS: Sentinel-1 and Sentinel-2 Fusion for Cloud Removal in Indonesian Urban Areas

Alexander Gunawan (Binus University)*
Roger Julianto Angryawan (Binus University)
Fabian Surya Pramudya (Binus University)

Abstract

Rapid urbanization in Indonesia has intensified the need for timely and reliable urban monitoring, yet persistent cloud cover continues to limit the utility of optical satellite imagery such as Sentinel-2. To address this challenge, this study presents Indo-UnCRtainTS, a domain-adapted cloud removal model for Indonesian urban environments, fine-tuned using a newly constructed dataset, Indonesia_Urban_MSCR, which integrates Sentinel-1 SAR data with Sentinel-2 optical imagery. Initialized from weights pre-trained on SEN12MS-CR, the model was refined to better represent the spectral and structural characteristics of urban areas in Indonesia. Performance metrics included MAE, SAM, PSNR, and SSIM. The optimized model surpassed the baseline, decreasing MAE to 0.021 and SAM to 7.095, while enhancing PSNR to 30.38 and SSIM to 0.854. These improvements demonstrate that domain-specific fine-tuning substantially enhances cloud removal performance under Indonesian urban conditions. To support practical deployment, we further developed SkyWipe, a web-based application that provides cloud-reduced satellite imagery by city and acquisition period for urban monitoring, planning, and disaster mitigation. A user evaluation revealed positive usability, especially in usability, interface uniformity, and interaction clarity.

Developer Discourse on Generative AI Programming Tools on Reddit: A Computational Grounded Theory Analysis

Alexander Gunawan (Binus University)*
KELVIN ALEXANDER BONG (Binus University)
JASON PANGESTU (Binus University)

Abstract

The rapid adoption of generative AI programming tools is reshaping software development, yet existing research has largely emphasized performance benchmarking and user perceptions rather than the socio-technical processes through which developers integrate these tools into everyday practice. In this study, we explored Reddit developers' discussions on generative AI programming tools using Computational Grounded Theory. We analyzed Reddit comments with topic modeling and sentiment analysis, followed by open, axial, and selective coding to create a conceptual framework based on the discussions. In the open coding stage, keywords, recurring expressions, and manually validated discussion elements were grouped into 11 initial concept clusters, including task, language, context, usage role, technical factors, practical factors, behavior, learning, impact, evaluation, and tool. In the axial coding stage, these initial concepts were synthesized into 10 higher-order thematic categories that captured relationships among programming tasks, usage roles, domain context, tools used, practical and technical factors, evaluation, learning influence, workflow impact, and behavioral change. In the selective coding stage, these interrelated categories were integrated into the core category of Tool-Mediated Adaptive Programming Practice, which conceptualizes AI-assisted programming as an iterative process structured by task execution, tool use, evaluation, and contextual adaptation. Quantitative pattern analysis reinforced this interpretation: the strongest co-occurrences were ProgrammingTasks-ToolsUsed (3,673 comments), ToolsUsed-Evaluation (2,251), and ProgrammingTasks-Evaluation (1,873), while the dominant triple, ProgrammingTasks-ToolsUsed-Evaluation, appeared in 1,238 comments. We found that 84% of comments on practical and technical considerations mentioned specific AI tools, with a positive overall sentiment of 48% positive, 51% neutral, and only 1% negative. These findings indicate that developers perceive AI as an active component of a dynamic working system, which they consistently evaluate and modify according to various factors.

Refining Spatial Detail in Sentinel-2 Wildfire Segmentation through a Data-Driven Post-Processing Pipeline

Alexander Gunawan (Binus University)*
Kasandika Andariefli (Binus University)
Vincentius Dewanto (Binus University)
Jennifer Patricia (Binus University)

Abstract

The inherently challenging nature of accurately segmenting wildfires from Sentinel-2 imagery is compounded by the dynamic and spatially heterogeneous behaviour of wildfire. Nonetheless, the predictions of deep learning models on wildfire mapping are invariably plagued by fragmented masks, inaccurate boundaries as well as missed burned areas in complex remote sensing scenes. In order to overcome these limitations without the need for such classifiers retraining behind existing segmentation models, this paper proposes a structured data-driven post-processing pipeline specifically designed for wildfire mask refinement. We performed experiments on the Sen2Fire dataset, a challenging benchmark for Sentinel-2 based wildfire segmentation. We applied the combined approach of CRF, thresholding and CCA to improve predictions from our U-Net & DeepLabV3+ models. Among the evaluated configurations, U-Net with EfficientNet-B4 combined with the CRF + Static Thresholding + CCA pipeline achieved the best overall performance, yielding the highest Recall of 0.8328. We emphasize the significance of accurate wildfire monitoring, prioritizing detection over missing burned areas. To facilitate practical application, we developed a block-based aggregation method and a Streamlit visualization tool for easier interpretation in operational contexts. Overall, our findings indicate that the post-processing pipeline enhances detection sensitivity and spatial consistency, resulting in more reliable wildfire mapping from Sentinel-2 imagery for practical use.

Machine Learning for Indonesian Stock Market Index Forecasting: The Role of Macroeconomic Indicators

Alexander Gunawan (Binus University)*
Wilson Gregory Pribadi (Binus University)

Abstract

The Indonesian stock market plays an important role in the national economy, yet its high volatility makes index forecasting inherently challenging. This study investigates the effectiveness of using global and domestic macroeconomic indices in forecasting three leading Indonesian stock indices, namely the IDX Composite, LQ45, and Kompas100. We use a multivariate time-series framework, applying log-return transformations to alleviate non-stationarity. We compare the three forecasting models: Random Forest, Temporal Fusion Transformer (TFT), and Time-Series Mixer (TSMixer) using seven macroeconomic covariates: Brent oil price, WTI oil price, the USD/IDR exchange rate, inflation, gross domestic product (GDP), consumer price index (CPI), and unemployment rate. Short-, medium-, and long-term forecasting horizons are used to assess model performance. The results show that Random Forest consistently outperforms the deep learning models, achieving the lowest average mean absolute percentage error (MAPE) of 1.00%, compared with 1.04% for TSMixer and 1.21% for TFT, while also demonstrating greater robustness, particularly for long-horizon forecasts. The results show that macroeconomic effects are index-specific: The IDX Composite is more sensitive to domestic inflation in the short run, while LQ45 and Kompas100 respond stronger to global energy prices, probably because of their higher exposure to commodity-related sectors. The study also shows that using all available covariates simultaneously can create unnecessary noise and degrade predictive performance, emphasizing the critical role of selective feature engineering in financial forecasting. To support practical use, the results are also implemented in a Streamlit-based dashboard for interactive analysis.

Interpretable Machine Learning for Credit Rating Classification: Evidence from Indonesian Companies

Alexander Gunawan (Binus University)*
Teuku Rusydi Khairi (Binus University)
Wawan Hermawan (Padjajaran University)

Abstract

Corporate ratings are significant to stability in the financial markets and efficient allocation of capital. Accurate credit assessments are important to attract investment and reduce default risk. Yet, market participants continue to place onerous reliance on ratings by credit rating agencies where methodologies tend to be backward looking. Their assessments are often too late to capture the rapid changes in firm fundamentals and therefore provide little real value as early warning signals of financial distress. The nonlinear and high-dimensional structure of financial data was also poorly fitted in a traditional statistical framework. While the machine learning methods present a viable solution, their usage within credit rating analysis is often hindered as many predictive algorithms are considered black boxes. This study fills one critical gap in the literature by combining machine learning classification with an interpretable framework, which is specifically developed for Indonesian firms. Based on a multi-year dataset of Indonesian companies extracted from the Bloomberg Terminal, we explore 24 financial variables that span sector, liquidity, firm size, solvency, efficiency and profitability to classify corporate credit ratings. In addition, the methodology is tested to compare performance of Random Forest and XGBoost models while integrating SHAP for global and local insights into feature contributions. XGBoost algorithm gives an accurate score of 85% on the test set which is better than Random Forest. The SHAP analysis further unveils that determinants of creditworthiness in terms of profitability, interest coverage, asset turnover and sector have the most positioning. These results show how interpretable machine learning could offer a more rapid and robust analytic method than traditional rating approaches, enhancing risk management and informing key stakeholders for improved decision-making.

Cluster-Based Labeling and Feature Reduction for Money Laundering Detection Using Hybrid Models

Jordan Jordan (Binus University)*
Rojali Rojali (Binus University)

Abstract

This study sets out to examine whether cluster-based labeling can meaningfully improve how we detect money laundering compared to the conventional approach of relying on manually defined financial rules. Identifying suspicious transactions remains a persistent challenge in financial systems, largely because fraudulent behavior is rarely straightforward — it often blends into legitimate activity in ways that rule-based methods struggle to catch. To explore this, we built a hybrid machine learning framework that brings together three clustering algorithms — K-Means, Gaussian Mixture Models (GMM), and DBSCAN — alongside two widely used dimensionality reduction techniques: Principal Component Analysis (PCA) and t-SNE. Classification was then carried out using Histogram-Based Gradient Boosting (HGB), with performance assessed through accuracy, precision, recall, and F1-score. Models that paired clustering-based labels with PCA consistently outperformed their counterparts. DBSCAN combined with PCA delivered the strongest overall results, reaching 99.79% accuracy, perfect recall, an F1-score of 96.97%, and a ROC-AUC of 100%. K-Means with PCA performed comparably well, with 99.58% accuracy and an F1-score of 98.31%. By contrast, substituting t-SNE for PCA led to a noticeable drop — K-Means with t-SNE, for instance, achieved only 94.94% accuracy and an F1-score of 78.95%, suggesting that t-SNE, despite its popularity for visualization, is less reliable as a preprocessing step for classification tasks in this context. What makes this work particularly interesting is the way it bridges unsupervised and supervised learning. Rather than depending entirely on predefined labels — which can reflect human assumptions and miss emerging fraud patterns — the framework lets the data reveal its own structure first, then uses those statistically grounded labels to train the classifier. This combination appears to give the model a sharper sense of what genuinely anomalous financial behavior looks like. Beyond the numbers, this research highlights a broader point: that thoughtful integration of machine learning techniques can push fraud detection capabilities well beyond what traditional logic-based systems offer. As financial crimes grow more sophisticated, approaches like this one may prove increasingly valuable for analysts and institutions looking to stay ahead of the curve.

Data Mining for Discovering Consumer Behavioral Patterns in E-Commerce to Inform Marketing Strategies

CORNELIUS Joe (Bina Nusantara University)
Joel Christopher (Bina Nusantara University)
Sulistyo Triutomo (Bina Nusantara University)*
Rani Puspita (Bina Nusantara University)

Abstract

The rapid growth of e-commerce has generated massive amounts of transaction data; however, high shopping cart abandonment rates remain a major challenge for the industry. This study applies data mining methods based on the CRISP-DM framework, with the aim of designing targeted marketing initiatives and identifying consumer behavior patterns. Following an evaluation of the K-Means, Gaussian Mixture Model (GMM), and Hierarchical Clustering models, the final results from Hierarchical Clustering with five clusters (K=5) emerged as the optimal model for this study, as evidenced by the most effective balance between internal consistency and clear segment separation. Additionally, the process of detecting product associations to support marketing strategies utilizes the Association Rule Mining method via the Apriori algorithm. The research results are presented in an interactive dashboard that serves operational frameworks, such as product bundling packages aimed at reducing cart abandonment and consistently increasing sales conversion, as well as platform management workflows and strategic practitioners, to support the implementation of personalized promotions.

Algorithmic Curation and Local Content Discoverability: A Survey-Based Assessment of Netflix's Recommendation Interface in Indonesia

Jasmine Chandra (Bina Nusantara University)*
Meilani Meilani (Bina Nusantara University)

Abstract

Global streaming platforms have increasingly controlled content visibility through algorithmic recommendation systems, this issue raise concerns about the discoverability of local cultural production. This study investigates how Netflix's recommendation shapes the visibility of Indonesian films among 119 Indonesian users from a descriptive-quantitative survey. The result from the investigation reveals that 46.2% of respondents discover Indonesian films from the Netflix's homepage auto recommendation, while the films row positioning significantly influences the user's viewing choices (mean = 3.80). Although 68.4% reported discovering previously unknown Indonesian films through the recommendation, however 46.2% are expressing their concerns that the local Indonesian films are getting more and more produced for global market, rather then for local audiences. This study's findings highlight an algorithmic paradox: Netflix are expanding the local films reach but doing that, risks narrowing its cultural scope.

The Impact of Social Media Analytics Capability on Entrepreneurial Performance in the Digital Era: The Role of Engagement and Trust

Luthfillah Basya (Bina Nusantara University)*
Murniati Murniati (Bina Nusantara University)

Abstract

The rapid evolution of digital technology has revolutionized business, especially with the widespread use of social media as a strategic tool for communication, market intelligence, and decision-making. However, despite its widespread use, many companies still face difficulties in transforming social media data into actionable insights, resulting in performance gaps. Social Media Analytics Capabilities (SMAC) is an emerging capability to measure performance, as well as improve social interactions and trust. Using Cochrane research methodology, SEM-PLS was used to collect and analyze data from 200 companies in Indonesia. The results show that SMAC, a digital business competency, positively and significantly influences engagement and trust in relation to corporate performance. Relationship transformation is influenced by trust, while engagement and trust play important roles as mediators. The results show that participation alone is not enough if it does not foster trust, highlighting the importance of combining behavioral and cognitive interventions. These are innovative ideas, such as integrating SMAC digitally into business models and creating a primary education engagement.

Sentiment Analysis of Toxic YouTube Comments Using Naive Bayes and Confusion Matrix

Romindo Romindo (Universitas Pelita Harapan)*
Kevin Bastian Sirait (Universitas Pelita Harapan)
Jefri Junifer Pangaribuan (Bina Nusantara University)
Valentino Riffan (Universitas Pelita Harapan)
Chailine Garcia Wijaya (Universitas Pelita Harapan)

Abstract

Toxic comments remain one of the most persistent problems in digital content moderation, particularly in open comment spaces attached to globally viewed and socially controversial media. This study examines the sentiment structure of comments posted on Hozier's “Take Me to Church” music video and evaluates the performance of the Naive Bayes algorithm for three-class sentiment classification. A total of 100,000 English-language comments were obtained through the YouTube Data API v3. After duplicate removal, language filtering, text cleaning, contraction expansion, tokenization, stopword removal, and lemmatization with the Natural Language Toolkit (NLTK), 51,348 valid comments were retained for analysis. The comments were manually annotated into positive, negative, and neutral categories, then classified using a stratified 80:20 train-test procedure and evaluated through a three-class confusion matrix. The model produced an overall accuracy of 74%. Class-level precision reached 96% for negative comments, 97% for neutral comments, and 64% for positive comments, while recall values were 57%, 49%, and 99%, respectively. The corresponding F1-scores were 71%, 65%, and 78%. The analysis identified 8,475 negative or toxic-leaning comments, equivalent to 16.50% of the processed dataset, whereas positive responses dominated the corpus at 69.75%. These findings indicate that Naive Bayes remains a practical, interpretable, and computationally efficient baseline for large-scale sentiment classification in YouTube comment environments, although class imbalance continues to constrain minority-class recall.

Optimized Hybrid Neural-Statistical Architecture for Time-Series Forecasting: An Adaptive Multicategory Evaluation

Aisyah At Thahirah (Bina Nusantara University)*
Suharjito - (Bina Nusantara University)

Abstract

In the Industry 4.0 era, accurate demand forecasting is essential for facilitating real-time decision-making in highly dynamic contexts. Even though multi-model forecasting methodologies have been extensively investigated, the majority of research continues to depend on aggregated data at monthly, weekly, or yearly intervals, whereas intraday sales forecasting is still comparatively unexamined. This study fills this gap by looking at intraday demand forecasting with a 6-hour interval, which provides a balance between temporal granularity and data stability. We suggest a dynamic model selection system that tests several configurations of several proposed forecasting models, such as ARIMA, SARIMA, LSTM, and a hybrid ARIMA-LSTM, to find the best model for each category. The framework not only proposes systematic parameter optimization but also an adaptive learning mechanism to optimize model performance across different types of demand patterns. Experimental results demonstrate that the proposed framework significantly outperforms literature baselines, achieving MAPE reductions ranging from 8% to 32% when compared to the best-performing models in each category. Analysis clearly shows that volatility in data impacts the effectiveness of algorithms in modeling time series. ARIMA models perform best when the stable data has dominant linear patterns, whereas hybrid architectures provide an optimal balance for moderate fluctuations. On the other hand, LSTM models exhibit resilience in identifying nonlinear mappings at high volatility levels. Therefore, it is evident that selecting models adaptively based on temporal attributes is imperative in ensuring consistency in predictions.

Public Emotion Analysis on Social Media and Short-horizon Stock Market and Rupiah Exchange Rate Movements A Case Study of the 17+8 Protests

Seline Loewel (Bina Nusantara University)
Gareth Aurelius Harrison (Bina Nusantara University)
Alleina Felisa Widjaya (Bina Nusantara University)
Belvareia Reva Utamandarya (Bina Nusantara University)
Benedictus Rahardjo (Bina Nusantara University)*
Evaristus Didik Madyatmadja (Bina Nusantara University)

Abstract

This study examines whether social-media emotion analysis provides useful information about short-horizon movements in Indonesian financial markets during the August–September 2025 "17+8" protest window. To this end, approximately 80,000 Indonesian tweets related to protests and market discourse were collected and processed through a noise filtering pipeline. These tweets were then categorized with GPT-5 into anger, fear, and other emotion categories. Daily shares of these emotions were aggregated into various sentiment metrics, such as a net sentiment index, and matched with daily IHSG and USD/IDR returns. Pearson correlation indicates a moderate contemporaneous association between net sentiment and IHSG returns and a weaker negative association with USD/IDR returns during the event window. Pairwise Granger causality tests provide only marginal evidence of predictive content for IHSG at a two-day lag and no significant predictive signal for the exchange rate. Overall, LLM-derived emotion features from Indonesian social media appear useful for event-aligned analysis of equity-market stress, although their forecasting value in this sample remains limited.

Quantifying Collective Intelligence as a Volatility-moderating Signal in Short-horizon Digital Asset Markets

Winston Khogres (Bina Nusantara University)
Seline Loewel (Bina Nusantara University)
Benedictus Rahardjo (Bina Nusantara University)*

Abstract

This study presents a social-signal forecasting framework for Bitcoin risk prediction using large-scale Reddit data and domain-adapted language modeling. The pipeline processes 46 million Reddit interactions collected from r/Bitcoin, r/CryptoCurrency, and r/CryptoMarkets between 2019 and 2024, applies automated noise filtering, and utilizes consensus labels from three large language model configurations to annotate 50,000 samples for DeBERTa-v3 training. The resulting sentiment signals are analyzed across community types and integrated into a PatchTST forecasting model with a 168-hour look-back window and a 22-hour prediction horizon. Empirical analysis identifies a statistically significant lead-lag relationship in which sentiment shocks from the speculative community precede Bitcoin price movements by approximately 22 hours ($p < 0.05$). In out-of-sample evaluation on the second half of 2024, the proposed framework reduces annualized portfolio volatility by 30.5%, from 52.8% to 36.7%, and lowers maximum drawdown by 24% or 6.8 percentage points. These findings indicate that filtered collective sentiment can provide a measurable predictive signal for risk-aware Bitcoin forecasting and support the use of social media analytics in financial time-series modeling.

Fuzzy TOPSIS for Digital-Readiness MSME Selection: A Multi-Criteria Decision-Making Method Employing Ministry of Manpower TKM Program Administrative Data

Sheila Hadi (BINUS)*
Fabian Surya Pramudya (BINUS)
Muhammad Fikri Hasani (BINUS)

Abstract

This study creates the Fuzzy Technique for Order Preference by Similarity to Ideal Solution (Fuzzy TOPSIS) framework in order to rank 478 micro, small, and medium enterprises (MSMEs) participating in a national workforce development program in Indonesia based on their digital readiness. A review of the literature led to the identification of seven evaluation criteria, which included financial governance, market reach, digital channel engagement, and economic performance. These criteria were then evaluated by statistical screening of administrative records. Fuzzy Direct Rating utilizing a five-level Triangular Fuzzy Number (TFN) scale was used to elicit criteria weights from four domain experts. The results were then aggregated by arithmetic mean and defuzzified using the Center of Area approach. The range of closeness coefficients (CC) is 0.135 to 0.721 where we know the mean is 0.442 and the SD is 0.112. The population was divided into three readiness categories using K-means Clustering where there is High ($n = 137$, $CC > 0.511$), Medium ($n = 196$, $0.378 < CC \leq 0.511$), and Low ($n = 145$, $CC \leq 0.378$). Spearman rank correlations of 0.999 or higher with the baseline were obtained through sensitivity analysis across 14 weight-perturbation scenarios where Spearman $\rho = 0.933$ was the result of a cross-method comparison using Simple Additive Weighting (SAW). An empirical proxy validation using five administrative indicators not included in the model confirms that tier assignments align with independently observable enterprise characteristics, including full financial statement adoption of High: 77.6% and Low: 22.4%, marketplace channel use of High: 68.8% and Low: 19.0%, and median monthly revenue of High: Rp 8,500,000 and Low: Rp 4,000,000. Program administrators can use this ranking list directly to allocate funding for digital adoption.

Forecasting Indonesia's cashless transition: a SARIMA and MIDAS approach with cross-country benchmarking

Anna Gracia Halim (Bina Nusantara University)
Valerie Angelica Dharmo (Bina Nusantara University)
Nadya Mulyarini (Bina Nusantara University)
Benedictus Rahardjo (Bina Nusantara University)*

Abstract

This study examines the development of Indonesia's cashless ratio and the factors affecting its growth towards digital payment. Data on monthly payment method usage from 2013 to 2024 are analyzed using the Seasonal Autoregressive Integrated Moving Average (SARIMA) model. The results show that Indonesia's cashless ratio remains stagnant at around 12.5% to 12.8% in 2027. Meanwhile, results using the Compound Annual Growth Rate (CAGR) predict that Indonesia will need decades for its cashless ratio targeted milestones starting from 30%. In order to further understand the factors driving the growth, the Mixed Data Sampling (MIDAS) regression is applied along with the monthly cashless ratio as targets and annual structural factors as predictors. Findings show that internet penetration, GDP per capita, urbanization, and size of the labor force have significant statistical effects on digital payment adoption, while education does not. These findings are then benchmarked against countries like Sweden, China, South Korea, and Thailand to pinpoint Indonesia’s focus of improvement.

Application of Self-Attention-based Imputation for Time Series on Sparse Morbidity Data for Deep Learning Forecasting

Rowell Hernandez (Batangas State Univeristy)*

Abstract

Maintaining public health supply chain resilience during pandemic and endemic emergencies requires highly accurate, morbidity-driven demand forecasting as the consumption and service data will not available initially. However, global epidemiological repositories (such as WHO FluNet or local HMIS & LMIS system) frequently suffer from structural sparsity, characterized by missing reporting, timeframes, and zero-inflated inter-epidemic phases. Traditional imputation methods, such as Last Observation Carried Forward (LOCF) or Mean Imputation, often resolve these gaps by applying zero-variance flatlines. Although these methods minimize absolute error mathematically during massive gaps, the non-linear biological transmission patterns and momentum of the epidemiological time series data are destroyed. To address this critical challenge, this study benchmarks the Self-Attention-based Imputation for Time Series (SAITS) architecture against traditional statistical methods across different missingness mechanisms (i.e., Point, Sequence, & Block) and missing rates (i.e., 10% to 50%) for three disease categories, COVID-19, Seasonal Influenza, and RSV datasets. SAITS demonstrated the superior topological reconstruction, specifically for random (Point) and sequential missingness. By achieving the lowest Mean Absolute Error (MAE) and Root Mean Square Error (RMSE), SAITS outperforms LOCF, K-NN, and MICE. Furthermore, the attention-based imputation ensures the integrity of the data stream by preserving the structural variance and non-linear characteristic of the outbreak curve, and proving to be a mandatory preprocessing step for deploying high-capacity recurrent neural networks in healthcare commodity forecasting during emergencies.

Out of Distribution Detection in Structured Data: A Systematic Review with Empahis on Autoencoder-Based Gatekeeper Methods

Rosni Lumbantoruan (Institut Teknologi Del)*
Katrina Tambunan (Institut Teknologi Del)

Abstract

Out-of-distribution (OOD) detection is an important part of decisions and safety applications like medical, financial and industrial monitoring. Although anomaly detection methods for images exist yet structured data consisting of multiple numeric or categorical fields faces additional challenges due to variability in field characteristics such as the variety of data types and the complexity of high dimensional interactions. This review studies 76 academic articles between 2017 and 2024 that specifically focuses on Autoencoder based gatekeeper approaches for detecting anomalies in structured data specifically on architecture, training design, synthetic anomaly techniques, and thresholds. They are compared to other statistical models, nearest neighbor models and hybrid models. Overall, our findings indicate that autoencoder gatekeepers can detect more complex relationships than previous methods, but their performance is dependent upon bottleneck size, regularized strength and the method used for calibrating them. We further note that there currently does not appear to be a standard benchmark for evaluating OOD performance across different datasets for structured data which hampers direct comparisons among studies. Finally we outline several design principles, some practical trade-offs, and provide direction for potential areas of research toward developing unified evaluation frameworks for reliably identifying anomalous instances in both numeric and string data.

A Systematic Ablation of CBAM Depth in SqueezeNet for optimizing Cross-Domain Chickpea Disease Detection

Rowell Hernandez (Batangas State Univeristy)*

Abstract

While the Attention mechanisms have been shown to enhance the performance of lightweight CNNs for plant disease detection, their optimal integration points within a network architecture remains unexplored. In this study, eight Convolutional Block Attention Module (CBAM) integration strategies within SqueezeNet v1.1 were systematically evaluated under a strict cross-geographic protocol. Under this protocol, models were trained on Turkish chickpea image dataset (n=1,894) and tested on independently collected Myanmar field image dataset (n=268, Cohen's κ=0.94). Results demonstrate that mid-level placement (V8) at Fire4 and Fire5 consistently achieves the top performing cross-domain accuracy with 82.09%. V8 outperformed deep-level placement variants and vanilla SqueezeNet. Moreover, unstable and lower cross-domain accuracy occurred in deep-only placement (V5, V6) and full-network placements (V4), indicating that mid-level texture features are more domain-invariant than high-level semantic representations. These findings provide the first empirical evidence on attention placement depth for chickpea disease detection and offer actionable guidance for designing cross-domain-robust lightweight models.

Augmented Reality Card Game as a Learning Medium for Acehnese Cultural and Geographical Education

Safitri Aprillia Putri (Bina Nusantara)*
Viona Putri Salim (Bina Nusantara)
Jevon Jeremy (Bina Nusantara)

Abstract

Cultural learning at the elementary school level has an important role in fostering students' awareness of local identity and cultural diversity. However, conventional learning media is often unable to actively involve students, resulting in low interest and understanding of the material. This research aims to design and develop an Augmented Reality (AR) based card game as an interactive learning medium for cultural and geographical introduction in elementary schools. This research uses a Design Thinking approach which includes the stages of empathize, define, ideate, prototype, and test to produce user centered learning media. The card game is designed to integrate paired card games with AR technology, where some cards are equipped with bar-codes to trigger animated displays and educational videos. Aceh Province is used as a case study to contextualize learning materials, including traditional clothing, landmark, traditional dance, and regional characteristics. The evaluation was conducted through a limited trial with fifth grade elementary school students to assess aspects of ease of use, student engagement, and understanding of the material. The results of the study show that AR-based card games are able to increase student involvement and active participation and support the understanding of cultural learning materials. This research shows the potential of AR-based card games as an interactive learning system that can be developed and adapted for the introduction of other regional cultures.

Assessing Human Adaptability Under Time and Environmental Constraints Through Time-loop Mechanic in the Game "30 Seconds"

Jeremy Dewanto (Bina Nusantara University)*
Kelvin Kelvin (Bina Nusantara University)
Galih Pratama (Bina Nusantara University)

Abstract

Adaptability supports positive psychological outcomes by helping individuals adjust their thoughts, emotions, and behaviors to change. It enables them to overcome challenges and continue functioning effectively in evolving environments. Adaptability is shaped by various factors that can influence how well a person adjusts to new situations, the example being time and environmental constraints. This study presents 30 Seconds, a short 2D platformer game. The player will navigate the character controller through a fixed map layout with obstacles such as breakables boxes, doors, and enemies. The player has to reach the end before the 30 seconds timer runs out, in which the character and most aspects of the game will be resetted from the start. The game was developed using Unity and Game Development Life Cycle (GDLC) as the development method. A total of 30 participants aged 18-25 participated by filling the Adaptability Scale questionnaire before (pre-test) and after (post-test) playing the game. The result was evaluated using the Shapiro-Wilk test followed by Paired Sample T-test if the data are normally distributed. Otherwise, Wilcoxon Signed-Rank test will be used. Evaluation results suggest that human adaptability increases after going through time and environmental constraints in the context of the game 30 Seconds.

Enhancing Cultural Awareness Through a Rhythm-Based Serious Game: A Study Using Indonesian Traditional Song

Andhika Suryanto (Bina Nusantara University)*
Reynata Praja Oktavian (Bina Nusantara University)
Mirza Ramadhani (Bina Nusantara University)

Abstract

The younger generation now has a low level of knowledge of traditional songs, especially in the 2020s, because of globalization which is threatening the preservation of intangible cultural heritage. This research proposes the creation of a rhythm-based serious game, called Project Rhythm, which aims to promote cultural awareness in students using the implicit learning mechanisms. The audio-motor synchronization paradigm in this study is not pedagogical approaches that pick up the cultural information explicitly, but it is a pedagogical approach that incorporates cultural information organically as a part of the gameplay. The study employed a quantitative experimental design using a one group pretest/posttest design with 50 university students who were assessed on three cultural awareness measures: song recognition; regional identification; and cultural understanding. A 15-item questionnaire, which had been validated using a Likert scale, was used for the evaluation. Using statistical analysis of the Wilcoxon Signed-Rank Test, a very significant increase ($p < .001$) was found in the level of cultural understanding, with participants showing a great increase in their comprehension of the stories behind the songs. Song recognition and regional identification revealed positive trends, though they were not statistically significant ($p > 0.05$). The results indicate that rhythmic rehearsal is useful for deep cultural engagement, but that the cognitive load and speed/accuracy relationship of the rhythm mechanics can inhibit the acquisition of peripheral textual information for a short period. This study shows the feasibility of serious games to connect old traditions and modern youth lifestyles in the digital world. Future developments should aim to create adaptive difficulty to achieve a balance between cognitive load and ensure that rhythm-based serious games continue to be a powerful medium for digital cultural heritage preservation and long-term cultural education.

THE EFFECT OF SOUND ON FEAR LEVELS IN THE 3D HORROR GAME FAMINE

Crescentia Prasetya (Binus University)*
Muhammad Fajar (Binus University)

Abstract

The purpose of this study is to explore the relationship between sound and fear levels in players of the horror game, Famine. This study focuses on Famine, a 3D horror game, and examines how sound design contributes to enhancing horror elements and increasing player immersion. The objective of this research is to evaluate how much sound affects players' fear reactions, identify how sound influences immersion, and contribute to the development of improved sound design techniques for horror games. This research was conducted by 40 participants filling out questionnaires after playing the game with one conditioned sound. The results indicate the significance of the use of sound to create a greater experience with the game as the anxiety response (ANX) and the physiological response (PHY) of the players of the 3D game were greater when the game was played with the use of sound.

UI/UX DESIGN OF THE "EDUPHANY" APPLICATION FOR ENHANCING THE QUALITY OF ELEMENTARY SCHOOL MATHEMATICS OLYMPIAD EDUCATION IN INDONESIA

Stephanie Vania Suwardi (Binus University)*
Safitri Aprillia Putria (Binus University)
Natalia Limantara (Binus University)

Abstract

This study investigates the mathematics olympiad from the perspective of SDG 4: Quality Education, highlighting the issue of unequal access to olympiad preparation in less-developed regions such as Java and Jakarta, which leads to uneven competition among provinces. The aim is to develop a technology-driven tool that allows children across Indonesia to access affordable preparation for mathematics olympiads. Employing a Mixed-method approach, data were gathered through surveys, interviews, observations, and literature reviews. The outcome is a mathematics olympiad learning app that can be accessed freely and at any time nationwide. In summary, there remains a disparity in access to Olympiad education between major urban centers and smaller areas, concerning both costs and availability of Olympiad coaches.

Integrating Computational Thinking into Project-Based Learning: A Collaboration Model and Minimum Design Standard for Computer Science Students

Cynthia Nababan (Institut Teknologi Del)*
Fanny Clara Sinaga (Institut Teknologi Del)
Yennisa Marianti Napitupulu (Institut Teknologi Del)
Exaudi Siregar (Institut Teknologi Del)

Abstract

Team collaboration is a central yet persistently problematic dimension of Project-Based Learning (PBL) in higher education computer science programs. This study investigates how integrating Computational Thinking (CT) into PBL can address collaboration challenges faced by undergraduate students. A mixed-methods design was employed: a validated Likert-scale questionnaire (Cronbach's $\alpha = 0.841$) was distributed to 127 students across three program cohorts at Institut Teknologi Del, and in-depth semi-structured interviews were conducted with 12 purposively selected participants. Survey results reveal that 94.5% of respondents reported unclear task distribution, 92.9% experienced ineffective team communication, 84.2% encountered difficulties reaching shared team goals, and 80.3% lacked adequate mutual support. Drawing on a systematic literature review of existing PBL-CT models and an expert interview with a CT specialist, a six-stage PBL-CT model was developed in which each PBL phase is enacted through specific CT indicators — decomposition, pattern recognition, abstraction, and algorithmic design — with observable collaboration checkpoints governing sequential progression. This model was instantiated as a Figma prototype comprising six interface modules. The prototype was qualitatively evaluated by 12 prospective users. Beyond this initial prototype, the study derives and presents a minimum design standard—a validated set of required interface elements for any CT-structured collaboration system and also as a reusable design framework for institutions developing PBL-CT collaboration tools.

Evaluating a Serious Rhythm RPG Game for Enhancing Decision Making Through Mental Workload Measurement

Raphael Ponda'ag (Bina Nusantara)
Andrew Hanjaya (Bina Nusantara)*
Dimas Ramdhan (Bina Nusantara)

Abstract

Cognitive workload is a crucial factor in cognitive performance and adaptation within digital gaming environments. This paper presents the design and evaluation of Dungeon Beats, a serious rhythm-based RPG, to study the influence of gameplay on cognitive workload and decision-making. The game combines rhythm synchronization with spatial navigation and combat mechanics in a fixed tempo, resulting in continuous cognitive demands. A within-subject quantitative method was applied to 33 undergraduate participants aged 18–24 years. Each participant played the game three consecutive times, with perceived cognitive workload measured using the NASA Task Load Index (NASA-TLX). This is based on comparing workloads in the first and third gameplay sessions to observe adaptation effects. The perceived cognitive workload significantly decreased, with mean Weighted Workload (WWL) scores dropping from 5.868 to 4.096. The paired sample t-test revealed that this difference was statistically significant, $t(32) = 5.885$, $p < .001$. These findings suggest that players became better adapted to the task demands through repeated gameplay, indicating improved cognitive efficiency in managing and allocating mental resources over time. Therefore, Dungeon Beats demonstrates potential as a serious game that supports cognitive load adaptation through structured rhythm-based gameplay. Further work could be considered on dynamic difficulty adjustment and setting different tempos to best suit cognitive involvement and learning.

User-Centered Gamification in E-Learning: A Preference-Based Personalization Approach using DPE and BWS

Ranty Deviana Siahaan (Institut Teknologi Del)*
Benhard Yudha (Institut Teknologi Del)
Archico Darius Simpar Sembiring (Institut Teknologi Del)
Rafael A. C. Manurung (Institut Teknologi Del)

Abstract

The potential of gamification has been demonstrated to be great in improving digital learning environments' user experience, with respect to the UX. Nevertheless, its utility is sometimes weak when implemented alike without customization to individual users. This study fills in the gap with three steps, being (1) Direct Preference Elicitation (DPE) with Best-Worst Scaling (BWS), which can measure the preferred gamification elements selected from 120 participants; (2) building a mobile learning application (Levelearn) by combining the five best aspects (Level, Badge, Point, Progress Bar, and Leaderboard); and (3) evaluating their impact on UX using the HEART model (Happiness, Engagement, Adoption, Retention, Task Success) with 55 participants. By using Spearman correlation and Ridge Regression and Wilcoxon Signed-Rank tests, statistical analysis confirmed that all five of them positively and significantly increased UX across all HEART dimensions ($p < 0.05$). Specifically, Progress Bar had the highest percentage association for Task Success ($r = 0.8134$) and Badge had the most with Retention ($r = 0.8076$). This is to the authors' knowledge the first to crosslink DPE to the HEART framework regarding the evaluation of gamification and presents a rigorous and user-centered methodological framework for personalized gamification design in educational technology.

Analysis of Factors Affecting The Use of Digital Libraries Through The Starry Application on The Reading Interest of Students at SMK Telkom Purwokerto

Cornelius Sarungu (Bina Nusantara University)*
Ery Lestari Cahya Ningrum (Bina Nusantara University)
Yosieka Wibawa (Bina Nusantara University)
Zalsa Kusumaningtyas (Bina Nusantara University)

Abstract

Student interest in reading in Indonesia remains very low, as evidenced by a UNESCO study showing that Indonesians' literacy rates rank second-lowest globally. Turning to the school setting, one example is a vocational school in Purwokerto. It is known that SMK Telkom Purwokerto has had a digital library application called Starry since 2023, which is actively available to students in grades 10–12. However, its utilization rate has not been optimal, and the extent of its contribution to fostering reading habits remains unclear, particularly amid the negative impacts of digital technology such as social media which have reduced interest in reading printed books. Therefore, this study aims to analyze the effect of using the Starry digital library application on the reading interest of students at SMK Telkom Purwokerto using the Technology Acceptance Model (TAM) approach. The variables analyzed include Perceived Usefulness, Perceived Ease of Use, User Experience, User Intention to Use, and Social Influence. This study employs a quantitative method using a survey technique involving 281 respondents. Data analysis was conducted using multiple linear regression. The results indicate that, collectively, all variables significantly influence reading interest. Partially, Perceived Usefulness, User Experience, User Intention to Use, and Social Influence have positive and significant effects, whereas Perceived Ease of Use does not have a significant effect. These findings confirm that technology acceptance based on the Technology Acceptance Model (TAM) plays a crucial role in enhancing students' reading interest.

Enhancing Fraction Learning in Primary Education through Curriculum-Aligned Game-Based Learning and Bloom's Digital Taxonomy

Pumudu Fernando (Informatics Institute of Technology)*
Salinda Premadasa (Sabaragamuwa University of Sri Lanka)

Abstract

Game-based learning (GBL) approaches have proven to be effective in enhancing engagement and performance among Generation Alpha learners. However, many recent GBL studies have not adequately aligned game-based learning activities with standard curriculum-based learning. Hence, the overall impact of GBL environments on learners' cognitive development is limited. To address this limitation, this study proposes the design, implementation, and evaluation of a fraction learning game aligned with the Sri Lankan Grade 6 Mathematics fractions curriculum. To support progressive learning implementation, Bloom's Digital Taxonomy was incorporated into the game design. Game-based activities are organized into three difficulty levels, namely Novice, Intermediate, and Expert. Learning tasks are designed to provide beginner to advanced-level knowledge on fractions, with increasing difficulty. The game was tested and evaluated against traditional classroom-based video simulation, using a quasi-experimental study. Grade 6 students from a Sri Lankan primary school were involved in the game testing. Student performance changes were measured using a Pretest and posttest assessments while a usability questionnaire was employed to measure motivation, engagement, and understanding in both game-based and traditional teaching approaches. As per the observations, the experimental group has shown a reasonable improvement in performance, with mean scores increasing from 2.9 to 7.6, compared to the control group's improvement from 2.6 to 6.2. Furthermore, learner perception analysis revealed improved student motivation (88%), engagement (92%), understanding (80%), and willingness to relearn (92%) in the proposed learning game compared to the traditional classroom simulation. The observations highlight that the proposed fraction learning game has enhanced students' subject knowledge. Furthermore, significant improvements have been observed in students' interaction, engagement, and motivation to learn. Hence, the findings of this study suggest that a curriculum-aligned game-based learning approach has the potential to enhance traditional classroom learning experience and improve primary mathematics education, when integrated with standard educational model.

Integrating Formative Feedback and the Technology Acceptance Model: Enhancing Student Satisfaction Through Automated Grading in Gamification-Based Digital Learning Environments

Cecilia Widayat (BINUS University)*
Edi Purnomo Putra (BINUS University)

Abstract

The current study focuses on the ever-growing requirement for instant and effective feedback in online education at higher educational levels to ensure quality access to education (SDG 4). The combination of automatic grading with gamification is becoming more common, but little empirical research exists on the impact of such a system on students' satisfaction via feedback and technology acceptance. This study looks at how automatic grading helps give students feedback during their learning and checks how it affects how happy students are with their experience in online learning games. A quantitative approach was used, collecting data from 150 Indonesian university students who had previously used gamified learning platforms that included automatic grading features. The study combines the Technology Acceptance Model (TAM) with a formative feedback system and uses Partial Least Squares Structural Equation Modeling (PLS-SEM) to examine the data. The results indicate that formative feedback significantly influences perceived ease of use ($\beta = 0.822$) and perceived usefulness ($\beta = 0.438$), while attitude toward use strongly predicts behavioral intention ($\beta = 0.837$), which turn has a substantial effect on student satisfaction ($\beta = 0.802$), with the model explaining 64.3% of the variance in student satisfaction ($R^2 = 0.643$). Theoretically, this research builds on the Technology Acceptance Model by adding formative feedback as an important factor. The results indicate that colleges and universities ought to focus on implementing automated grading tools that offer prompt and significant responses to improve student involvement, educational results, and fair access to high-quality learning in online educational.

A Virtual Reality Game for Financial Literacy among Young Learners: An Empirical Study of User Engagement

Lana Aliyah Amran Amran (Universiti Teknologi Mara)
Nurul Hidayah Mat Zain (Universiti Teknologi Mara)*
Mohd Nabil Zulhemay (Universiti Teknologi Mara)
Chiou Sheng Chew (Universiti Teknologi Mara)
Anita Mohd Yasin (Universiti Teknologi Mara)
Nor Azida Mohamed Noh (Universiti Teknologi Mara)

Abstract

People who master financial literacy understand the value of money and how to use it effectively. Financial literacy requires understanding money values while making sensible spending and financial decisions. However, lack of engaging financial education content and the limitations of traditional classroom methods in capturing young learners' attention. Hence, this project introduces an interactive learning tool that integrates VR technology to teach key financial concepts through gameplay, namely Grocery Grab. The game applied the Game Development Life Cycle (GDLC) as its methodology and implemented the User Engagement Model as a guideline. The 3D environments tailored for young learners were created using Unity and Blender tools. The User Engagement Scale-Short Form (UES-SF) was used to evaluate user engagement. The findings showed high engagement levels, with scores of 4.41 for Focused Attention, 4.37 for Aesthetic Appeal and 4.61 for Reward. The Perceived Usability score was slightly lower at 3.25, while the overall engagement score was 4.16, equivalent to an engagement rate of 83.2%. In conclusion, Grocery Grab demonstrates that VR can be an effective tool for enhancing financial literacy among young learners. By combining enjoyable gameplay with learning, this project makes educational VR games more engaging than traditional methods and opens the door for future improvements.

The Role of Interactivity on Co-Developing, Influencing, Augmenting, and Mobilizing Behaviors: A Study of Roblox Mobile Gamers

Adi Teguh Suprapto (Bina Nusantara University)*
Raizidan Rahman (Bina Nusantara University)
Revi Arnan (Bina Nusantara University)

Abstract

In the mobile gaming ecosystem, consumers serve multiple roles as both players and creators of content. One of the most prominent platforms is Roblox, which is a unique user-generated game that allows for high interaction and collaboration among users. Customer engagement theory forms the theoretical framework for the current study, which aims at determining how interactivity influences four different dimensions of customer engagement behavior (co-developing, influencing, augmenting, and mobilizing) and analyzes whether service quality acts as a moderator in this relationship. By applying PLS-SEM method to 104 Roblox users in Indonesia, this study found a significant positive influence of interactivity on influencing, augmenting, and mobilizing behaviors. However, there is no evidence that interactivity positively affects co-developing behavior. Moreover, service quality plays the role of a moderator in the relationships between interactivity and all four dimensions of customer engagement behaviors. Service quality appears to work as a substitute for interactivity since high service quality leads to a reduction in interactivity's impact on customers' engagement behaviors. This paper contributes to customer engagement literature by providing a context-specific examination of interactivity-based engagement models in user-generated gaming platforms and introducing a substituting mechanism of service quality.

Santa-Pan Nusantara: Pilot Study on Educational Serious Game About Indonesian Local Food and Geography

Kayla Lukman (Binus University)*
Jeslyn Djunaidy (Binus University)
Muhamad Fajar (Binus University)

Abstract

Indonesian culture is starting to fade from the minds of young people. This research offers a solution by providing a serious game for education about Indonesian local food and geography. Serious games are a type of video game typically built for educational purposes, able to create immersive and meaningful experiences while also providing interesting challenges. “Santa-Pan Nusantara” is a serious game with point-and-click mechanic and guessing game about Indonesian geography. Through the results of the questionnaire given out, it shows significant and positive correlation for all hypotheses except for H1. The data strongly suggests that "Santa-Pan Nusantara" successfully bridges entertainment and education. Players find the game visually engaging and easy to use, which significantly enhances their knowledge of Indonesian culture. Most critically, the perceived educational value (PU) is the strongest predictor of whether they will continue playing and recommend the game, affirming the project's core mission.

Integrating Design Thinking and Technology Acceptance Model In Developing A Sustainable E-Commerce Prototype For Carbon Emission Reduction

Evlyn Manurung (Bina Nusantara University)*
Muhammad Rachmadhani (Bina Nusantara University)
Ayu Cahyani (Bina Nusantara University)

Abstract

The growth of e-commerce activities has led to serious implications for increased carbon emissions. Despite this, consumers continue to depend on e-commerce to meet their specific needs. This scenario promotes sustainable e-commerce, incorporates inclusive features and sustainable practices into the shopping experience. The platform aims to connect sellers with consumers through four core features: eco-score, community, recycling tips, and eco-friendly delivery and packaging, to maintain a sustainable experience that spans from the decision-making stage to post-purchase. To address the setback, this study employs a five-stage design thinking method focused on user-centered platform development and utilizes the Technology Acceptance Model (TAM) to assess users' acceptance of the platform. A survey was conducted with 72 respondents in Indonesia and validated by SEM. TAM was tested through the perceived ease of use and perceived usefulness construct. The results show that perceived usefulness positively and significantly influences users' intention to use the platform. Users recognize that it offers a meaningful sustainable shopping experience, thus facilitating more frequent purchases. This study is expected to provide valuable insights for future developers, emphasizing the need to enhance user interface and experience, as well as the impact of social influences.

Development and Validation of a Three-Layer Heuristic Evaluation Framework for Faith-Aligned Human-Computer Interaction

Fadzlin Ahmadon (Universiti Teknologi MARA)*
Muna Azuddin (International Islamic University Malaysia)

Abstract

Heuristic evaluation is a widely used usability inspection method in Human-Computer Interaction (HCI), yet existing heuristic sets rarely address faith-aligned digital artefacts designed for spiritual and religious contexts. This paper presents the development and expert validation of a three-layer heuristic evaluation framework for the LUQMAN system, a prototype designed to support Muslim parents in sensemaking of online risks for their children. The framework integrates Nielsen's ten usability heuristics, Morville's seven User Experience Honeycomb values, and seven newly derived Spiritual Experience values grounded in Surah Luqman (verses 12–19). Two statement formats were operationalised: Operationalised Heuristic Statements (OHS) for UX experts and Simplified Heuristic Statements (SHS) for lay parent evaluators, yielding 62 items across 24 heuristic values. Expert content validation was conducted with six HCI/UX academics from four Malaysian public universities using the Content Validity Index (CVI) method. All 62 statements achieved a perfect Item-level CVI (I-CVI = 1.00) and Scale-level CVI (S-CVI/Ave = 1.00). Qualitative expert feedback guided targeted refinements across three revision categories: clarity, split, and present tense. The resulting instrument provides a validated, domain-specific evaluation tool that extends conventional HCI evaluation to encompass spiritual user experience.

Increasing Awareness and Knowledge of Burnout Syndrome Through Sunside Cafe: A Serious Game Study

Vincent Wijaya (Binus University)*
Evan Jonathan (Binus University)
Radovieo Daffacetta (Binus University)
Galih Pratama (Binus University)

Abstract

With the topic of mental health being more apparent since the COVID-19 pandemic, the case of burnout syndrome happening towards people has been increasing. With young adults such as university students being part of the highest age group to be affected, a need for awareness and knowledge regarding the syndrome must be affirmed to help prevent and reduce the case of burnout and enable help for people affected by it. As one of the effective mediums of cognitive behavioral therapy, serious games have a potential to give education and awareness regarding the topic. As such, this study aims to understand if game-based learning is an effective method to increase burnout knowledge to university students aged 18-25. Using the EMERGO method alongside Jesse Schell's elemental tetrad concept, the education game will then be created and then spread for people to play. The methodology used to obtain quantitative data with pre and post-test surveys regarding burnout knowledge before and after playing the game, then analyzed using the Wilcoxon test. The results show that there's a 24.14% significant increase in knowledge indicated by the p-score from both pre and post-test surveys below 0.01, before and after playing Sunside Cafe. In conclusion makes it possible to use Sunside cafe as a method of education regarding burnout knowledge.

Memetic Algorithm Approach in Code Caves for Detection Evasion

Vincent Kwee (BINUS University)
Michael Windarta (BINUS University)
Vito Putra (BINUS University)*

Abstract

Deep learning-based malware detectors such as MalConv learn directly from raw executable bytes and are often seen as more robust than traditional signature-based systems. However, they can still be fooled by adversarial modifications that preserve malware functionality while changing the model's prediction. This paper presents a black-box attack that generates in-binary adversarial examples by injecting semantics-preserving perturbations into inter-section code caves of Windows PE malware and optimizing them with a memetic algorithm. The attack operates only on bytes that are not mapped into memory at runtime, keeping the original executable content and behavior. Using MalConv as the target model, we collect PE32 samples from MalShare, MalwareBazaar, Zenodo, VirusShare, and TheZoo. We then filter them to 559 valid, detected malwares, which then we ran our memetic algorithm on. Results from our tests show that the memetic algorithm gives better results on average than the genetic algorithm in terms of evasion rate, generations, and time taken to reach an evasive malware score.

Integrating ISO 31000, ISO/IEC 27005, and NIST SP 800-30 for Risk Management in High-Classification Government Application

Dedy Setiawan (Binus University)*
Benfano Soewito (Binus University)

Abstract

Digital transformation has encouraged various sectors to adapt and innovate. One such sector is government, which is required to simplify bureaucracy through digitalization. However, the digitalization process has led to an increase in the complexity of cybersecurity risks. A comprehensive risk management strategy is needed to address vulnerabilities to cyberattacks. A stand-alone risk management framework is not entirely effective in dealing with this issue, therefore a combined framework approach may be a more relevant solution. This research is a case study of risk management implementation in Application X, which is a high-classification application in the government sector. The methods used combine ISO 31000, ISO/IEC 27005, and NIST SP 800-30. ISO 31000 is used as the main risk management framework, meanwhile ISO/IEC 27005 and NIST SP 800-30 are used as guidelines in the risk assessment process. The risk evaluation stage is carried out by comparing the results of the risk analysis with the risk acceptance criteria, with the aim of determining which risks are tolerable. The results show that 64 risks were identified, of which 13 need to be addressed because they are outside the organization's risk tolerance, while the other 51 do not need to be addressed because they are within the organization's risk tolerance. ISO/IEC 27002 is used as a technical guideline for risk treatment recommendations.

Proactive Behavior-Based Ransomware Detection Using System Monitor Log Feature Extraction (Sysmon) And Random Forest Algorithm

Priskila Tsunami (Bina Nusantara University)*
Benfano Soewito (Bina Nusantara University)
Ika Rachmawati (Bina Nusantara University)

Abstract

Ransomware is one of today's most prevalent global cybersecurity concerns. Today's ransomware uses polymorphic evasion techniques that can't be detected by signature-based techniques alone. Behavior-based ransomware detection holds promise; however, many current solutions rely on volumetric behavior such as CPU/Memory usage and network activity and fail to detect low-and-slow or offline crypto-ransom attacks. This paper presents an intentional host-based detection framework to proactively detect ransomware by applying deep behavior analytics against System Monitor (Sysmon) logs with a Random Forest classifier. Our solution leverages deterministic behavior intent such as high entropy file creation, ransom note writing, and write-disable command usage to detect ransomware at the endpoint independent of encryption speed. We used an equal Random Forest to overcome class imbalance. Our stratified 5-fold cross validation produced models with low variance (std of .0019). The proposed model attained 95.14% accuracy, and mean accuracy of 92.31% on an out-of-sample test set with low false-negative rates. We've proven that system log–based intent analysis can detect zero-day and adaptive ransomware reliably in real-time.

A New Alternative for Testing Brownian Motion in Time Series

Muhammad Fajar (Badan Pusat Statistik-Statistics Indonesia)*
Rinda Nariswari (BINUS UNIVERSITY)
Celine Ghefira Nurhaliza (Badan Pusat Statistik-Statistics Indonesia)

Abstract

This study introduces a robust framework for testing Brownian motion by combining fractal dimension estimation with the Maximum Entropy Bootstrap as the foundation of the testing procedure. The method is designed to overcome the weaknesses of conventional Brownian motion process or pure random walk tests, which typically depend on strict assumptions of independence and normality. These assumptions are often violated in real-world time series data exhibiting autocorrelation and nonlinear patterns. In this framework, Brownian motion is defined under the null hypothesis as having a fractal dimension of 1.5, and statistical inference is carried out using MEB so that the temporal dependence structure of the data is preserved in the sample bootstrap distribution. Results from Simulation I show that the proposed test maintains properly controlled Type I error rates across different sample sizes. Meanwhile, Simulation II demonstrates perfect empirical power when tested against a strongly persistent, non-Brownian alternative, indicating excellent discriminatory performance even in finite samples. These findings confirm both the statistical validity and sensitivity of the proposed testing. An empirical application to Indonesia's monthly inflation data from 2015 to 2025 produces a fractal dimension estimate of 1.657, corresponding to a Hurst exponent of 0.343. The test results strongly reject the null hypothesis of Brownian motion, suggesting that the inflation process exhibits anti-persistent, mean-reverting dynamics rather than Brownian motion process. Overall, the test proposed by this study offers a statistically rigorous and practically applicable approach for identifying deviations from Brownian motion in complex time-series data.

Development of a Freshwater Fish Cultivation IoT System Based on Correlation Analysis to Identify the Most Influential Variables

Boby Siswanto (Bina Nusantara University)*

Abstract

Freshwater fish farming can now be done on a small scale, with some using tarpaulin ponds. Water quality parameters are important to consider because cultivation is carried out in conditions where the water is not changed and relies on filter media. Some water quality parameters are pH, water temperature, turbidity, and dissolved oxygen. The purpose of this study is to develop an IoT system for freshwater fish farming using correlation analysis. This study utilized a secondary dataset from Kaggle to analyze the correlation of the parameters available in the dataset, and grouped them into morning, afternoon, evening, and night. The analysis results stated that water temperature has a positive correlation with pH, turbidity, and dissolved oxygen, where the strongest correlation is between water temperature and pH. The best time to adjust pH based on water temperature is in the morning, based on the correlation value obtained > 0.66, while the best time to adjust dissolved oxygen based on water temperature is during the day, with a correlation value > 0.23. The IoT system created in this study uses a Peltier as a water temperature regulator.

An Explainable Web-Based Application for Rule-Driven Deep Learning Intrusion Detection System

Van Long Nguyen Huu (Can Tho University)*
An Tran Cong (Can Tho University)

Abstract

In recent years, we have witnessed the emergence of Deep Learning in Intrusion Detection Systems (DL-IDS), which can analyze and adapt to complex patterns and anomalies in Internet of Things (IoT) network traffic. However, DL presents significant challenges in terms of transparency and explainability of decisions, which are critical for confidence in Security Operations Centers (SOCs). Integrating eXplainable AI (XAI) frameworks like SHAP (Shapley Additive exPlanations) provides the reasoning behind the decisions of DL models. Although SHAP aids in transparency, its results are often too complex and condensed for human interpretation. This paper's main objective is to efficiently integrate and distill SHAP values, making them interpretable to support more informed and secure decision-making in DL-IDS operations. Specifically, through deep inspections and granular insights, we leverage SHAP global and local features visualization to validate the reasonability behind the high performance of a specific DL-IDS. Based on the contribution and correlation of SHAP features, we apply seven major decision signatures for corresponding predefined analysis and Artificial Intelligence (AI)-driven reactive solutions to present a web-based decision-making support interface. Experimental results show that 1D-CNN in IDS achieves the highest performance, with a precision of 0.994, while accuracy reaches 99.42% on specific intrusion datasets. Importantly, we validate that the integration between 1D-CNN and SHAP is seamless and transparent, as it can automatically detect intrusions, summarize the signatures, provide predefined remarks, and recommend a reaction for SOC analysts.

Determinants of Cybersecurity Awareness among Indonesian Generation Z: A PLS-SEM Approach

Vivian Alicia Santoso (Universitas Bina Nusantara)*
Daniel Kurniawan Tjahjono (Universitas Bina Nusantara)
Jovanda Carissa Joewono (Universitas Bina Nusantara)

Abstract

Cyberattacks now affect not only systems but also humans. Generation Z is known to have a high level of digital literacy, but is vulnerable to various cyber threats, so it is necessary to understand the factors that affect cybersecurity awareness. This study used a quantitative approach with a survey method of 159 respondents in the Generation Z category. Data were collected through a structured questionnaire with a 6-point Likert scale and analyzed using PLS-SEM. The results of the study show that digital literacy has a positive and significant influence on cybersecurity. In contrast, social norms have no significant effect on cybersecurity awareness. These findings show that cybersecurity awareness in Generation Z is largely influenced by individual factors rather than social pressure. This implication emphasizes the importance of strengthening digital literacy to enhance cognitive and socio-emotional abilities and to promote safe behavior when using digital technology.

Implementation of Retrieval-Augmented Generation on Indonesian AI-Driven Job Interview Practice Chatbot for Software Engineering

Dexter Krisnadhi (BINUS University)*
Hosannia Hardiputri (BINUS University)
Evotianus Evotianus (BINUS University)
Muhammad Hasani (BINUS University)

Abstract

Artificial intelligence has been increasingly implemented in interview training through chatbots to provide interview simulation. However, unlike human interviewers, LLMs lack access to specific knowledge of the job, required skills, and company needs which can result in questions and assessments that are often general and less relevant. To address this limitation, this study develops an interview chatbot using the Retrieval-Augmented Generation (RAG) method, so that the system can read additional data and generate more specific questions to fill the research gap. The interview chatbot system was developed using Mistral 7B and tuned to generate responses in Indonesian. This chatbot is integrated with 471 rows of dataset embedded using BGE-large-en and indexed with FAISS. Hyperparameter tuning was performed to optimize the chatbot, while user acceptance testing through questionnaire was conducted to evaluate its performance. Our results showed that the chatbot implemented with RAG cannot exceed the non-RAG version in user acceptance testing before and after hyperparameter tuning, even though the RAG version was better when judged by Gemini 3.1 Pro while in hyperparameter tuning. This suggests a trade-off where RAG improves completeness but results in longer and less natural responses. Additionally, RAG introduces issues such as English leakage.

Software Reengineering of Village Administration Information System: Achieving Regulatory Compliance with Ministry of Home Affair Regulation No. 47 of 2016

Riyanthi Sianturi (Institut Teknologi Del)*

Abstract

Villages have the autonomy to manage their own administration as a crucial component of accountable and efficient governance. However, many villages still rely on manual systems for administrative record-keeping and reporting, leading to data delays, the loss of critical information, and a lack of transparency toward the community. The village administrative information system developed in 2024 supports only 2 of the 24 administrative registers required by Ministry of Home Affairs Regulation No. 47 of 2016, and lacks adequate technical documentation, thereby hindering sustainable development. As a solution, software reengineering was performed on the village administrative information system using the Pressman approach, which encompasses four stages: inventory analysis, document restructuring, data restructuring, and forward engineering. The system was developed with a monolithic architecture using the waterfall development method, simplifying the development scenario from two servers to a single server. The database was restructured from 9 to 28 tables to cover all general administrative, population, and development needs in accordance with applicable regulations. Functional testing using the black box testing method on six system modules resulted in a 100% pass rate. The system was successful deployed on June 20, 2025, in Hutabulu Mejan Village and has proven to be fully operational for village officials. The research findings indicate that the reengineering process successfully expanded the scope of village administration in accordance with Ministry of Home Affairs Regulation No. 47 of 2016.

Regression-Based Theoretical Calibration of Multiple MQ Gas Sensors Using Datasheet Curves for IoT Air Quality Monitoring

Adhe Lingga Dewi (Binus University)*
Anang Prasetyo (Binus University)
Jayden Samuel Kurniawan (Binus University)
Charles Nobel Wijaya Salim (Binus University)

Abstract

Air pollution is a global problem that impacts human health and environmental quality, so an accurate air quality monitoring system is needed. This study aims to apply a theoretical calibration method to MQ type gas sensors based on sensor sensitivity curves and regression equations that refer to the sensor's technical specifications. The air quality monitoring system was built using MQ – 2, MQ – 7, MQ – 8, and MQ – 135 sensors to measure the concentrations of methane, carbon dioxide, hydrogen, and carbon dioxide, and DHT11 for measuring temperature and humidity. Sensor calibration was carried out theoretically by modeling the relationship between sensor output values (ADC) and gas concentrations in PPM units using nonlinear regression in the form of a power function. The modeling results produced a regression equation with a coefficient of determination (R2) value above 0.99 for all sensors, indicating a very good level of model suitability to the sensor characteristics according to the manufacturer's data sheet. The regression equation obtained is then implemented into Arduino IDE programming as a calibration formulation, so that the ADC value can be converted directly into a theoretical concentration estimate based on the regression equation and improve the consistency and interpretability of the MQ gas sensor readings, and can be applied as a practical solution in an Internet of Things (IoT) based air quality monitoring system.

Target-Oriented Robust Optimization for Blockchain Governance Implementation Risk Management in the Philippines

Arnold Aguilar (De La Salle University)*

Abstract

Implementing blockchain governance in developing economies is characterized by complex multi-objective decision environments which include severe uncertainty regarding institutional capacity, regulatory readiness and infrastructure availability. This paper develops a new formal approach to national blockchain governance rollout planning using a robust optimization technique called target oriented robust optimization (TORO). TORO has been applied in practice for deterministic scheduling problems but not yet in the context of phased implementation planning for blockchain based public fiscal transparency. The model used in this study applies TORO to four institutional objectives that are relevant to the Philippines' CADENA act (Senate bill 1506): procurement integrity; digital inclusion; institutional compliance; cybersecurity resilience. The BG-TORO framework introduces a novel formulation that replaces classical deterministic scheduling with a robustness index $\theta \in [0,1]$. This index quantifies the maximum institutional uncertainty that any phased rollout plan can withstand while remaining feasible. An additional layer uses the analytical hierarchy process (AHP) to encode multi-stakeholder governance priorities into a weighted composite constraint. This generates a doubly parameterized Robustness-Efficiency Frontier indexed by risk posture α and priority weight vector w. five phase case studies for the Philippines demonstrates that balancing configuration $\alpha=0.5$ achieves $\theta^*=0.689$. Therefore, the plan is feasible in 68.9% of all adverse institutional scenarios.

Adaptive and Explainable Human-Robot Collaboration for Intelligent Sorting Systems Woody the Educational Robot

Uvindya Jayasundara (Sri Lanka Institute of Information Technology)*
Yeshani Kumarage (Sri Lanka Institute of Information Technology)
Hiruni Liyanagamage (Sri Lanka Institute of Information Technology)
Kithmi Muthuweera (Sri Lanka Institute of Information Technology)
Samadhi Rathnayake (Sri Lanka Institute of Information Technology)
Amila Alexander (Sri Lanka Institute of Information Technology)

Abstract

The present study proposes a novel integration of HRC, aiming at providing an inclusive platform for cognitive and sensory learning of differently abled students. The system consists of four key components, namely 1) a vision-based texture recognition component based on a self-supervised transformer (DINOv2) algorithm for object material identification, 2) a robotic learning component aimed at pattern recognition and linguistic ability improvement via object-based exercises, 3) a head and gaze-tracking object selection process that makes object interaction non-invasive due to the use of only one RGB camera and 4) an XAI-driven predictive sorting component that improves transparency in decision making when completing the task. By incorporating object detection and manipulation algorithms, attention tracking, and speech feedback, the system creates a multi-modal interactive and adaptive learning environment. Due to the presence of a human-in-the-loop component, a student can watch how the robot performs its task and gains additional knowledge about it. This way, the use of real-time feedback and explainability in robot decision making results in increased engagement, better understanding, and more trust.

Fuzzy-PID Based Motion Control for a Mecanum-Wheeled Smart Trolley

Alexander Gunawan (Binus University)*
Alicia Taslim (Binus University)
Junaedi Dede (Binus University)
Muhamad Daffa Mennawi (Binus University)

Abstract

Reliable navigation is the basic demand required for autonomous smart trolley systems, especially in achieving consistent path tracking across distinct route complexities. In this paper, a Fuzzy-PID based control strategy is proposed to enhance the line-follower mechanism of smart trolley. It incorporates infrared sensors for line detection and turn orientation control, localization based on NFC-RFID techniques, Wi-Fi communication with an Android application, and Mecanum wheels enabling omnidirectional movement. To evaluate the performance of our proposed Fuzzy-PID controller, we conducted comparative tests against proportional (P), proportional-derivative (PD), and conventional PID controllers. We utilized Average Absolute Error (AAE) as the performance metric. These findings demonstrate that adaptive gain tuning through fuzzy logic improves the accuracy and robustness of the smart trolley's line-following performance, making Fuzzy-PID the most effective control strategy for the proposed navigation system. These results translate to an overall error reduction of 5.29% against traditional PID, 7.12% over PD and 12.54% compared to proportional control The results showed that tunning the gain adaptively using fuzzy logic increased the performance of smart trolley's line-following landing both in terms of accuracy and robustness thereby proving to be the best control strategy while implementing navigation system.

Impact of Micro Services Architecture Design Patterns on Software Maintenance Cost

Gintoro Gintoro (Universitas Bina Nusantara)*

Abstract

Software maintenance typically accounts for the largest portion of total costs in the lifecycle of enterprise systems, but the mechanisms by which architectural decisions impact that cost are still poorly understood. This is especially true in microservices where design pattern choices result in costs that are direct but not tracked. This research attempts to fill this gap by creating an empirical pathway from MSA design pattern use, to changes in class-level structural metrics, and resulting COCOMO II maintenance cost estimate. A production Student Registration System with 21 classes was analyzed both before and after purposeful use of four design patterns: Service Decomposition, Anti-Corruption Layer, Strangler Fig, and Database per Service. Structural analysis identified a concentration of high coupling on a single class with 10 dependencies, a circular dependency group of three critical classes, and low cohesion in some core service classes. The proposed Architectural Maintenance Effort (AME) formula, which combines Coupling Factor, Cohesion Factor, Inheritance Factor, and Cycle Penalty, connected the structural measurement outcome to COCOMO II Annual Change Traffic estimation. Application of the design patterns resulted in a 79.1% reduction in the AME composite metrics and a reduction in predicted annual maintenance effort from 5.45 to 3.98 person-months, which is a 27.0% reduction in cost due to architectural changes only; elimination of the circular dependency accounted for 51.3% of the total reduction which highlights its impact given the scale of change.

Absolute Zero Reasoning for Autonomous Penetration Testing: A Hybrid LLM–MCTS Framework with Action-Constrained Exploration

Yulius Prabowo (Univesitas Bina Nusantara)*
Albert Wibisono (Universitas Atma Jaya Yogyakarta)
Martinus Maslim (Universitas Atma Jaya Yogyakarta)

Abstract

The increasing scale and complexity of enterprise networks make manual security assessment progressively less effective, thereby creating a pressing need for automated security validation system that can adapt to uncertain and evolving attack environtments. Large Language Model (LLM) offer promising capabilities for contextual interpretation and high-level decision making in such setting, but their standalone use remains constrain by hallucination, repetitive action pattern and weak grounding in environment specific states. This study proposed Absolute Zero Reasoning (AZR) based framework for autonomous penetration testing that combines Qwen-1.5B as an intuition engine with Monte Carlo Tree Search (MCTS) as formal reasoning engine. In the proposed framework, each candidate action is subjected to 100 internal simulations before execution, enabling the agent to verify the plausabilty od language guided decision through deterministic search. To further improve the behavioral stability, a dynamic action mechanism is introduced to prevent repetitive and unproductive action loops. Experiment on the Chain-10 scenario in MicrosoftCyberBattleSim demonstrate that the proposed hybrid agent realibly reaches local privilege escalation and achieved higher exploitation efficiency than LLM-Online Baseline. This result suggests that integrating probabilistic language-based intuition with search based verivication offers more robust and effective foundation for autonomous cyber security agents and advances development of trustworthy autonomous penetration testing systems

Network performance of multiple VPN topologies under simulated DDoS attacks using GNS3

Nur Akmal Fahmi Nuri (FSKM, UiTM)
Alya Geogiana Buja (FSKM, UiTM)
Albin Lemuel Kushan (FSKM, UiTM)*

Abstract

Network security has been evolving throughout the years and even more so with the emergence of growing technology like Artificial Intelligence (AI). One of the implementations of network security is the Virtual Private Network (VPN) and it is often used in untrusted networks to protect data confidentiality, however, the performance trade-off during a cyber-attack has not been fully explored yet. This research study offers comparative performance analysis between four network topologies which are No VPN, Site-to-Site (S2S) VPN, Dynamic Multipoint VPN (DMVPN) and a hybrid (S2S VPN and DMVPN), where all the topologies are tested with simulated Distributed Denial-of-Service (DDoS) attacks. The testbed is developed in the Graphical Network Simulator-3 (GNS3) network simulator which consists of Cisco 7200 series routers as the network device, several end host systems, and a few virtual machines that run Kali Linux that act as the attackers. The DDoS attack traffic is produced using the hping3 tool in Kali while the genuine traffic is captured using the Wireshark tool and each topology is tested five times to provide consistency. A custom Python-based Graphical User Interface (GUI) tool, named Network Performance Extractor, is developed to automate extraction and visualization of five key network performance metrics: throughput, latency, bandwidth, jitter and packet loss. The result obtained exhibit that the S2S VPN achieve the highest overall performance, it delivers the highest average on throughput and bandwidth including the lowest overall latency and jitter, all which maintain zero packet loss. Likewise, the hybrid S2S VPN and DMVPN topology also provides a balance compromise between throughput and jitter stability to provide suitability for real-time applications. However, in contrast, the DMVPN shows the worse latency and has the lowest throughput in the test, even though it is seen as a flexible and scalable architecture. Under the simulated attacks, all topologies manage to maintain 0% packet loss and the findings imply that VPN architecture selection greatly influences network performance and should be carefully decided when designing efficient and secure networks that are prone to DDoS threats.

Blockchain-based Islamic family asset management system with a secured encrypted repository

Muhammad Hariz Hakim Zakaria (FSKM, UiTM)
Alya Geogiana Buja (FSKM, UiTM)*
Raihah Aminuddin (FSKM, UiTM)
Sulaiman Mahzan (FSKM, UiTM)
Dziauddin Sharif (ACIS, UiTM)
Anwar Farhan Zolkeplay (FSKM, UiTM)

Abstract

This study addresses the issues with current Islamic asset management such as transparency in handling Islamic asset distribution, and data security problems. Centralised system often vulnerable with data tempering and breach, leading to loss of trust among users and the increase in unclaimed assets, this study offers a secure, safe and Shariah-compliant platform through the usage of blockchain technology and encryption, to improvise the current system. With blockchain, documents are ensured to be temper- proof, and all interaction are transparent, while encryptions provide additional layer of security towards sensitive data. The platform relies on Next.js, smart contract on Ethereum, and Interplanetary File System (IPFS) for decentralise storage. User Acceptance Testing (UAT) have been conducted to validate the objectives in terms of usability and functionality. This study shows that blockchain, paired with suitable encryption, would enhance Islamic assets distribution, by offering secure and trustworthy platform, with Shariah principle at its core.

Development of Secure Credential Repository System Using Multifactor Authentication and Encryption

Tengku Zafran Shah Tengku Iskandar Zulkarnain (FSKM, UiTM)
Alya Geogiana Buja (FSKM, UiTM)*
Mohd Hafifi Mohd Supir (FSKM, UiTM)
Noor Afni Deraman (FSKM, UiTM)
Nor Intan Shafini Nasaruddin (FSKM, UiTM)
Nor Aimuni Md Rashid (FSKM, UiTM)

Abstract

In today's digital economy, which is linked to the internet, passwords, personal identification numbers, and private documents are more likely to be hacked or misused illegally. Single-factor authentication methods, which typically use passwords that people remember, fail to defend against brute-force attacks or credential theft. Further, some Multifactor Authentication (MFA) solutions that lack a strong encryption can be breached and used for inappropriate reasons. This project describes about how to make and utilize a Secure Credential Repository System (SCRS) that secures a digital credentials with MFA and Laravel's encryption API. The SCRS is distinct because it can securely digitally pass on credentials to the right next of kin in case a user pass away. The system was developed according to a literature review, survey data collection, functionality testing, and User Acceptance Testing (UAT) to verify how reliable and useful it is. The findings indicate that the SCRS keeps digital credentials much safer, which enhances users' belief and trust in digital systems. This project enables to move forward bigger efforts to protect digital identities and reduce down on data breaches by providing a safe and easy-to-use alternative.

Architectural Refactoring and Container-Based Orchestration for DiTenun Platform

Arlinta Christy Barus (Institut Teknologi Del)
Humasak Simanjuntak (Institut Teknologi Del)
Lusye Triksi Pasaribu (Institut Teknologi Del)
Hedrin S. Sitorus (Institut Teknologi Del)
Greyssenly Ester Lya Simanjuntak (Institut Teknologi Del)
Samuel Situmeang (Institut Teknologi Del)*

Abstract

DiTenun is a digital weaving design platform for weavers. Initially, the platform had two main divisions. One of the pieces is Ulospedia, a knowledge base about culture. The second piece is Modul Lidi, which is a pattern generation tool using ant colony optimization. Because of how each of these components operated independently, the platform struggled with integration and performance. To address these issues, this study proposed migration to a microservices architecture to enable co-working services more easily. For example, managing different services deployed in different frameworks or using different environments requires a powerful infrastructure. Consequently, containerization is adopted to standardize the environment across services and reduce dependencies. Docker is deployed for packaging the services, and Nginx is deployed as a reverse proxy. This ability could fuse all the different technical elements into one platform that people can refer to over the cloud. The new architecture ensures that the development and production process of the system is the same, a critical process for sustaining the system and fulfilling service level contracts between startups that do digital asset management. Additionally, the new system connects all requests to the central server and joins the previously separate components of the system. In general, this approach performs well in performance tests with a response time improvement of 15.42% ($p=0.046$). Using containerized microservices will solve integration challenges in the AI systems, leading to better performance and reliability.

CRISPR-AF: A Bio-Inspired Adaptive Firewall with Self-Curating Signature Memory for Network Intrusion Detection1

Dr Muruganandham S K (Rajarajeswari college of engineering)*

Abstract

There's an alarming trend in exact-match intrusion detection systems such as Snort. They keep adding rules. Over time, firewalls become bogged down looking for thousands of expired signatures in every network packet. Worse, if you're an attacker you can bypass these hard-matching rules by altering their payload by a little bit. Bacteria do it another way. The CRISPR-Cas immune system doesn't store an endless list of previous infections. It has a compact and current set of viral snippets. When a viral invasion subsides, the bacteria forget that. We designed CRISPR-AF as a mechanic. It's a self-learning firewall that maintains a tightly bounded array of unique short sequences of By It detects altered payloads via approximate matching that traditional firewalls fail to detect. The firewall manages its own ruleset. It dynamically adds signatures during the attacks and removes no-hit signatures whenever that occurs. We tested CRISPR-AF using the CICIDS2017 dataset against Snort , Random Forest, LSTM and traditional Artificial Immune Systems. The firewall matches them in detection, but only with a small fraction of the rules. It also interfaces well with polymorphic attacks, and learns new attacks without retraining.

Hybrid CNN-BiLSTM Based Secure Predictive Maintenance Framework for Accurate Fault Detection in IIoT Systems

Revathi S (VELS UNIVERSITY)*

Abstract

Predictive maintenance plays a crucial role in enhancing reliability and reducing operational costs in Industrial Internet of Things (IIoT) environments. This paper proposes a hybrid deep learning framework that integrates Convolutional Neural Networks (CNN) and Bidirectional Long Short-Term Memory (BiLSTM) for accurate anomaly detection and fault prediction in industrial equipment. Initially, multivariate sensor data collected from industrial systems undergo preprocessing, including noise filtering, missing value handling, and normalization using Min–Max scaling to ensure uniform data distribution and improved model convergence. The CNN component is employed to extract high-level spatial features from the processed data, while the BiLSTM network captures bidirectional temporal dependencies, enabling the model to learn both past and future contextual information within time-series signals. The proposed hybrid architecture significantly enhances predictive accuracy and robustness compared to conventional machine learning and standalone deep learning models. Furthermore, a secure data transmission mechanism is implemented using the Message Queuing Telemetry Transport (MQTT) protocol, reinforced with Advanced Encryption Standard (AES) and Transport Layer Security (TLS), ensuring data confidentiality, integrity, and protection against cyber threats during communication between edge devices and cloud servers. The framework supports real-time monitoring, early fault diagnosis, and proactive maintenance decision-making, thereby minimizing unexpected downtime and operational risks. Experimental results demonstrate that the proposed approach achieves superior performance in terms of accuracy of 97.8% , reliability, and security. Overall, the developed system offers a scalable, intelligent, and secure predictive maintenance solution tailored for modern IIoT-based industrial applications.

Dual-Emotion Mining for Fake News Detection on Social Media

Van Long Nguyen Huu (Can Tho University)*
An Tran Cong (Can Tho University)

Abstract

Misinformation on social media damages institutional trust and incites social unrest during critical events. Conventional detection approaches frequently apply semantic exploitation to detect complex linguistic patterns through Machine Learning (ML) or Deep Learning (DL). However, these approaches fail when misinformation emulates the professional syntax and formal vocabulary of reputable news. To address this limitation, our paper introduces Dual Emotion Mining to combine the analytical of textual content and emotional dissonance between Publisher Emotion and Social Emotion. Dual Emotion Mining specifically targets the psychological gap between the intent behind misinformation and the community's emotional response, thus, facilitating more accurate identification of sophisticated rumors designed to evade semantic detection. We also present a dual-pathway architecture to integrate high-dimensional SBERT embeddings with a DistilRoBERTa emotion map. Experimental results validate our proposal by demonstrating that optimized XGBoost classifier of dual path way architecture achieves an accuracy of 0.73 and F1-score of 0.61.

Mobile Application for Real-Time Fresh Fruit Bunch Oil Palm Ripeness Classification

Rosniza Roslan (Universiti Teknologi MARA)*
NURUL ANIS NAJIHAH AZMAN (Universiti Teknologi MARA)
AHMAD HAFIZUDIN HAKIM DAWAM (Universiti Teknologi MARA)
Noor Hasimah Ibrahim Teo (Universiti Teknologi MARA)

Abstract

This paper presents PalmVision, a real-time mobile application that has been developed to digitalize Fresh Fruit Bunch (FFB) oil palm ripeness classification directly in plantation fields. This application embarks accurate ripeness detection across different maturity stages, reducing reliance on manual inspection and improving harvesting decisions in oil palm domain. This study employed four phases of modified Waterfall methodology i.e., requirements, design, development, and testing phases. PalmVision consists of seven key modules that are homepage, camera interface, result viewer, action recording, detection report analytics, filtering, data export, and record deletion that integrate into user-friendly platform suitable for field workers. This application offers enhanced oil palm FFB detection, and grading ripeness maturity, ripeness detection reporting functions that help track harvesting patterns, and productivity levels that have been significant to the PalmVision. This paper shows PalmVision evaluations produced the effective and efficient runs across all seven interface modules, and usability assessment. Therefore, PalmVision improves efficiency, minimizes crop losses, and enhances fruit quality for both small farms and commercial plantations.

Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 289 (2026) 851–858

Procedia Computer Science

www.elsevier.com/locate/procedia

The 11th International Conference on Computer Science and Computational Intelligence 2026

A Preliminary Study on Mitigating Upcoding in Indonesia's National Health Insurance using Multi-Agent Multi-Armed Bandits

Elisabeth Victoria Nilasari[a], Bambang Soelistijanto[a,*]

[a]*Department of Informatics, Sanata Dharma University, Yogyakarta 5528, Indonesia*

Abstract

Developing countries are often faced by corruption. Corruption threatens public sector organizations' viability. For instance, recent corruption at BPJS Kesehatan (Indonesia's universal health care provider) has recently led to enormous financial losses. Adversaries causing losses intelligently game the system. Existing static solutions cannot easily keep pace with the adversary. Limited audit resources further compound the problem. With hard budget and human-resource constraints, how can auditing efforts be directed to wisely choose claims that are suspicious for human review? We introduce a decentralized auditing system built with Multi-Agent Multi-Armed Bandit (MAMAB) framework. Contextual Bandits models allow us to reason about individual claims using attributes of a patient's claim. Bandits with Knapsacks (BwK) allow us to reason about the daily limit of human reviewers. Our goal is to leverage MAMAB to efficiently utilize limited resources and choose claims with the highest likelihood of containing fraud for review by a human. Moreover, we can theoretically allow regional BPJS branch offices to act as individual agents that share their knowledge of adversary behaviors (instead of sharing patients' information) to quickly learn adversarial patterns and minimize overall regret. In this paper, we present a high-level problem formulation and system design to implement MAMAB in BPJS Kesehatan's constrained environment.

© 2026 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0)
Peer-review under responsibility of the scientific committee of Bina Nusantara University.

Keywords: Multi-Agent Systems; Contextual Bandits; Healthcare Fraud Detection; Federated Learning

* Corresponding author. Tel.: +62-819-3172-7372.
E-mail address: b.soelistijanto@usd.ac.id

1877-0509 © 2026 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0)
Peer-review under responsibility of the scientific committee of Bina Nusantara University.

1. Introduction

Developing countries face numerous socio-economic challenges that hinder national growth, with systemic corruption remaining one of the most difficult obstacles [1]. In Indonesia, corruption can be found across various critical sectors, including banking, insurance, and healthcare [2]. Unfortunately, a major for target for this type of fraud is the Badan Penyelenggara Jaminan Sosial Kesehatan (BPJS Kesehatan) which is the state agency established to provide universal health coverage to the Indonesian population [3].

The Indonesian Ministry of Health has categorized eight main types of healthcare fraud. These include phantom billing, phantom/manipulation diagnosis, self-referrals, repeat billing, fragmentation, bribery, illegal fees, and upcoding [4]. In this study, we will focus on upcoding because it remains quite difficult to handle. This is primarily because the practice involves altering a patient's diagnosis or procedure codes to bill at a higher, medically unjustified rate. Additionally, malicious hospitals are constantly changing their tactics to avoid the detection systems. This makes it hard to detect upcoding with traditional rule-based algorithms.

For example, in July 2024, the Corruption Eradication Commission (KPK) discovered that the state had lost Rp 35 billion from upcoding at BPJS Kesehatan. This case is an obvious example of how the ongoing fraudulent claims are threatening the financial stability of BPJS Kesehatan [5]. The KPK also stated that of the 4,341 BPJS Health claims they investigated, they found 30,000 entries that were inconsistent or might have been fraudulent, and 40 percent had markups as high as 200% above the normal price. This ongoing case ultimately cost BPJS Rp 20 trillion, thus severely threatening its financial stability [6].

To overcome this problem, we adopt Multi-Armed Bandits (MAB), which is an artificial intelligence approach capable of optimally making a series of decisions under uncertain circumstances [7]. However, a standard MAB approach is not enough since the probability of fraud is deeply tied to the specific attributes of a patient and their visit. Instead, we rely on Contextual Bandits, which is an extension of the framework where the system evaluates the unique context of every claim to accurately adapt and identify fraudulent behavior over time [7].

Nevertheless, BPJS branch offices cannot audit all questionable claims since human auditors are highly limited. For this reason, we add the Bandits with Knapsacks (BwK) framework to our system. Rather than simply copying the idea of the classical knapsack problem, BwK turns the whole system against itself and forces it to be very thoughtful. The system must foresee changes and select a certain group of claims that will maximize detecting fraud. More importantly, it does all of this while making sure the required effort never exceeds what a human auditor can handle in a single day [7].

Since BPJS operates on a big scale and spread out throughout Indonesia, we need a solution that fits this reality. Therefore, we use a multi-agent system (MAS) which is defined as a system where several autonomous agents perceive their environment, behave accordingly, and interact with one another [8]. Together with these concepts, we propose a "Multi-Agent Multi-Armed Bandit" (MAMAB) architecture.

The MAMAB model proposed here involves multiple decentralized agents (the regional BPJS agencies) working in parallel and solving the same cBwK problems. This collaborative learning system allows agents to help each other by sharing fraud patterns (mathematical weights). For example, an over-coding method which an agent in Jakarta found can also be recognized by an agent in Surabaya. Hence, this will speed up detecting best audit targets and reduce collective regret very significantly.

1.1. Related Work

Traditionally, the approach to fraud prevention in the healthcare and financial sectors was predominantly focused on changing the rules of the organizations, strengthening the control system inside the organization, and training the staff extensively [9]. On the other hand, as fraud techniques evolve in complexity, fraud detection strategies are increasingly relying on supervised Machine Learning (ML) algorithms. Recently, studies comparing different ML models on imbalanced transaction data have illustrated the compromises involved with the use of standard ML models [10]. For example, Logistic Regression is a very interpretable and easily implementable model, but it is not very flexible in terms of feature interactions and nonlinear decision boundaries. Decision Trees, on the other hand, can capture nonlinear behavior; however, single trees tend to overfit the training data when there is a big imbalance in the class distribution. Ensemble methods such as Random Forest and Light Gradient Boosting Machine (LightGBM) not

only control overfitting better but also provide more stable models and less memory usage. Deep learning models such as Gated Recurrent Units (GRU) have the capability to derive sequential or contextual representations from categorical fields but are also likely to have a relatively high false-positive rate [10]. To combine the advantages of these different methods, cost-sensitive ML frameworks that emphasize high recall by adjusting decision thresholds considering the cost factor have been successfully implemented making it possible to catch 98% of the actual fraud by abandoning static, unchanging rules [11].

Traditional supervised learning models perform well when trained on historical data. Despite that, their performance often declines in real-world environments because malicious behavior changes over time. A problem known as *concept drift* [12]. As a result, researchers have recently explored the use of Multi-Armed Bandit (MAB) algorithms for adaptive fraud detection. Earlier studies noted that applying MAB methods to fraud detection was challenging due to difficulty of creating decision "arms" [12]. To address this issue, newer approaches based on Contextual Multi-Armed Bandits (CMAB) have been proposed. Some studies combine CMAB with incremental regression trees and semi-supervised learning techniques to dynamically group transactions with similar characteristics. This allows the model to continuously explore potentially suspicious patterns while still exploiting previously identified fraud indicators [13]. In addition, recent studies show that multi-agent contextual bandits offer greater efficiency in resource-constrained environments [14]. This strongly supports the use of multi-agent contextual bandits to handle upcoding within the strict audit limits of BPJS Kesehatan.

1.2. Research Gaps

Recent studies show that there are still substantial issues with the application of these technologies to a huge and a completely decentralized healthcare system such as Indonesia's BPJS Kesehatan. The paper here aims at two main research gaps.:

- Ignorance of Operational Resource Constraints: Existing high-recall ML frameworks and CMAB models optimize for fraud detection accuracy but implicitly assume an unlimited capacity to investigate suspicious claims [11], [13]. Human auditors are a strictly limited daily resource. There is a lack of literature applying Bandits with Knapsacks (BwK) framework to healthcare fraud, which is necessary to mathematically pace and optimize the allocation of limited human auditing capacity over a continuous stream of claims.
- Lack of Collaborative Multi-Agent Learning in Privacy-Restricted Environments: Current state-of-the-art fraud detection systems operate in centralized, single-agent environments. They do not address the decentralized reality of regional healthcare branch offices that must combat dynamic fraud (such as upcoding) without violating strict personal data protection laws. There is a critical gap in utilizing Federated Multi-Agent Multi-Armed Bandits (MAMAB) to allow distributed agents to collaboratively share learned adversarial fraud patterns (parameter weights) without sharing raw, sensitive patient contexts.

2. Method

To address the highly adaptive nature of upcoding fraud and the strict operational constraints of the BPJS Kesehatan verification process, we propose a decentralized, privacy-preserving Multi-Agent Multi-Armed Bandit (MAMAB) framework. The proposed model integrates Linear Contextual Bandits (LinUCB), Lagrangian-based Knapsack, and Federated Learning for multi-agent knowledge synchronization.

2.1. System Architecture

As shown in Figure 1, the end-to-end design of our proposed system runs through three main functional areas, enabling continuous real-time decision-making and periodic global learning. The initial point is in the Data Pipeline and Context Extraction section. The system gets raw claim documents from the BPJS Claim Stream at the hospital providers, and this happens before the payment is made. Context extraction is done by a dedicated data engineering pipeline which converts characterization of raw categorical and numerical features into a numeric standard form vector called Context Vector x_t. The stream is then divided and sent according to the office branch code.

Once a claim is routed, it will move into the MAMAB Local Agents area where each BPJS regional office is viewed as an autonomous agent. Upon receiving the context vector x_t, the local agent decides whether to call the claim for audit or give it an auto-approval based on the agent's local policy $\pi(\alpha|x)$ against its strictly defined local human-auditor budget B_{local}. The agent then outputs an action $\alpha_t \in 0,1$, where 0 represents the decision to approve the claim and 1 represents the decision to audit it. This central zone also contains BPJS Headquarters (Central Server) which is responsible for the federated learning process by computing and disseminating the global fraud blueprint θ_{global} without ever requiring or storing raw patient data.

The selected action instantaneously activates the Feedback and Reward Mechanism zone. If the agent decides on $\alpha_t = 0$, the claim is routed to auto-approve, requiring no human intervention. The system instantly generates a feedback loop of $(r_t = 0,\ c_t = 0)$ back to the agent. On the other hand, if the agent decides on $\alpha_t = 1$, the claim is routed to a human auditor who investigates the case and provides a ground truth label $y_t \in 0,\ 1$, where 1 denotes confirmed fraud and 0 denotes a clean claim. After that, the system mathematically calculates the financial reward of the detection $(r_t = y_t \times v_t)$ where v_t represents the monetary value of the submitted claim and consumes the local audit cost $(c_t = 1)$. In this way, a full feedback tuple is produced (r_t, c_t, y_t) and sent back to the branch office. Upon receiving this feedback, the local agents update their internal mathematical models and eventually upload these parameters to the central server, allowing the entire network to adapt to newly discovered fraud patterns.

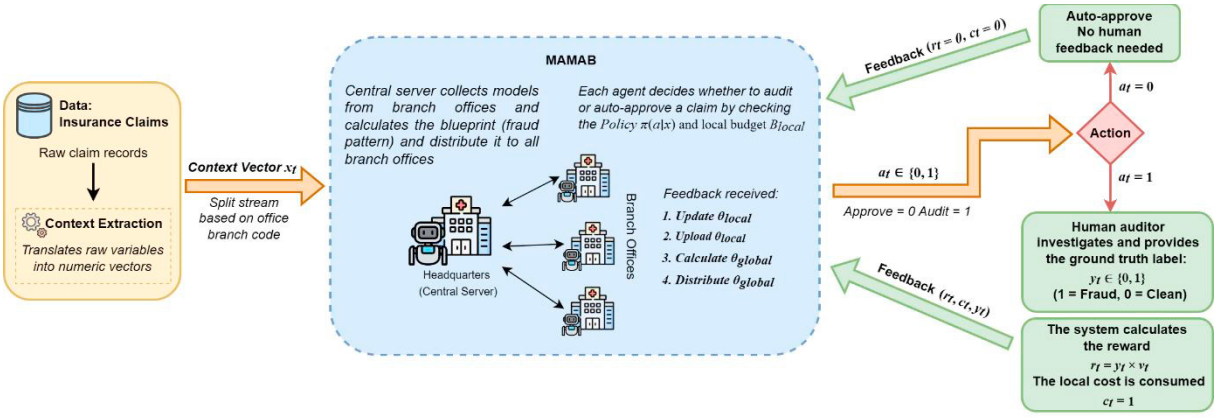


Fig. 1. Proposed system architecture of MAMAB.

2.2. Contextual Bandits with Knapsack (cBwK)

We formulate the daily claim verification process at each branch office as a finite-horizon cBwK problem defined by the tuple $T,\ A,\ X,\ D,\ B$. For a given day, the branch office receives a sequence of T claims. At each time steps $t \in |T|$, the local agent observes a context representing the claim feature vector x_t\in$X \subseteq \mathbb{R}^d$. Based on this observed context, the agent must select an action α_t from the binary action space $A = 0,1$. Selecting $\alpha_t = 0$ designates the Null Arm (Auto-Approve), which safely processes the claim without consuming any human audit resources. Conversely, selecting $\alpha_t = 1$ designates the Audit Arm, which routes the claim to a human verifier.

Each selected action incurs a resource consumption cost $c_t(\alpha_t) = \alpha_t$. This means that an automated approval costs zero and a manual audit consumes exactly one human audit slot. The agent must ensure that the cumulative cost does not exceed this limit $\sum_{t=1}^{T} c_t(\alpha_t) \leq B_{local}$. If the local audit budget is completely depleted, the agent is strictly forced to pull the Null Arm $(\alpha_t = 0)$ for the remainder of the time horizon T. Finally, the objective of the agent is driven by the reward function $r_t(\alpha_t)$, which quantifies the actual financial value of the fraud successfully prevented, is calculated as $r_t(\alpha_t) = \alpha_t \times y_t \times v_t$.

2.3. LinUCB

To avoid prematurely exhausting the budget due to an excessive number of checks, local agents must balance the probability of detecting fraud against the cost of deploying a human inspector. We model this decision-making process

using the Linear Contextual Bandits (LinUCB) algorithm [7]. As defined in the linear contextual bandits framework, the agent assumes that the expected reward in the event of fraud $\mu(\alpha|x_t)$ for an action α given a claim context x_t is a linear function of an unknown parameter vector θ_α [7].

$$\mu(\alpha|x_t) = x_t \cdot \theta_\alpha \tag{1}$$

For the Null Arm, the expected reward is known to be exactly zero $\mu(0|x_t) = 0$. For the Audit Arm $\alpha = 1$, the agent maintains a "confidence region" C_t that contains the true fraud parameter θ_1 with high probability [7]. For an incoming claim x_t, the agent calculates the Upper Confidence Bound (UCB) of the expected fraud reward by finding the parameter θ within the confidence region C_t that maximizes the expected outcome:

$$UCB_t(1|x_t) = sup_{\theta \epsilon C_t} x_t \cdot \theta \tag{2}$$

To incorporate the local auditing capacity constraint, we adapt the LagrangeBwK approach [7]. We introduce a Lagrangian multiplier $\lambda \geq 0$, which acts as a dynamic "shadow price" representing the scarcity of the remaining budget. The agent evaluates the penalized score for auditing.

$$Score_t(1) = UCB_t(1|x_t) - \lambda c_t(1) \tag{3}$$

Therefore, the agent's decision rule at step t is to pull the Audit Arm only if the UCB score exceeds the shadow price of the budget, provided capacity remains:

$$\alpha_t = \begin{cases} 1 \;\; if \; UCB_t(1|x_t) > \lambda \; and \; \sum_{t=1}^{T} c_t(\alpha_t) \leq B_{local} \\ 0 \;\; otherwise \end{cases} \tag{4}$$

2.4. Decentralized Federated Learning

To strictly adhere to Indonesia's Personal Data Protection Law (UU PDP), raw patient contexts (x_t) and ground-truth labels (y_t) are legally prohibited from being shared across regional branches. To overcome this limitation, we implement a decentralized Federated MAMAB protocol that enables fraud patterns to be shared globally while raw data learning remains entirely localized. Instead of transmitting sensitive patient records, the agents solely exchange the mathematical parameters that define their local confidence regions.

The learning cycle begins when a branch agent i audits a claim $\alpha = 1$ and receives ground-truth feedback comprising the reward, cost, and label tuple (r_t, c_t, y_t). Using this feedback, the agent independently updates the mathematical boundaries of its local confidence region $(C_{local,i})$ to reflect the newly observed fraud pattern. After the end of the synchronization period, each branch securely uploads to the BPJS Headquarters the parameters describing its locally updated confidence region $(\Delta C_{local,i})$. The Central Server then gathers the parameters from every M participating branch and merges them into one highly focused global model C_{global}. Finally, the Central Server broadcasts the updated global parameters back to all local agents. With this federated approach, when one branch discovers a new upcoding pattern, the system updates its confidence region around the related context vector. As a result, other branches can detect similar fraudulent contexts in the next operational cycle without sharing raw patient data.

3. Results

Since this paper is intended as a preliminary study, empirical simulations are still ongoing but not yet done. Nevertheless, the proposed Multi-Agent Multi-Armed Bandit (MAMAB) framework already provides several analytical insights for the BPJS Kesehatan system through its mathematical formulation and overall design.

- Optimal Utilization of Limited Audit Resources: Traditional greedy algorithms deplete auditing capacity early, leaving systems vulnerable to late-day adversarial attacks. By integrating a Lagrangian multiplier λ into the

LinUCB scoring function, our framework converts the hard constraint of human auditor capacity (B_{local}) into a dynamic shadow price. This mathematically guarantees budget pacing, ensuring a continuous defense across the entire horizon T.

- Privacy-Preserving Collaborative Adaptation: Sharing raw patient data across branches violates Indonesia's Personal Data Protection Law (UU PDP). Our Federated MAMAB architecture ensures that only non-sensitive, aggregated parameter updates are transmitted to the central server. The resulting global policy (θ_{global}) legally and mathematically protects all provincial branches against novel upcoding tactics discovered anywhere in the network.
- Value-Driven and Adaptive Fraud Detection: Traditional models optimize raw accuracy and can be misleading in highly imbalanced datasets. By defining the reward function as the actual financial value of prevented fraud, the algorithm acts as a cost-sensitive classifier. This is a theoretical proof that the system prioritizes the most expensive high-risk claims. As a result, minimizing the financial loss for BPJS Kesehatan.

4. Discussion

4.1. Contextual Bandits vs. Supervised Learning

Most automated fraud detection systems in the healthcare and financial sectors rely on supervised machine learning (ML) algorithms like Random Forest, LightGBM and deep learning architectures [10]. These models offer high precision on known patterns but their reliance on historical pre-labeled data exposes a fundamental vulnerability since they assume a static data distribution.

Healthcare fraud is highly adversarial and able to cause concept drift [12,13]. When static supervised models are deployed, they fail to recognize newly invented upcoding schemes because those schemes do not exist in the training data.

Contextual Bandits address this issue by balancing exploration and exploitation [7]. Rather than relying only on previously known fraud patterns, algorithms such as LinUCB also evaluate unfamiliar claim combinations by assigning them an uncertainty bonus. As a result, the system occasionally investigates suspicious billing patterns that have not been seen before. This allows the model to adapt more quickly to evolving fraud strategies rather than waiting for periodic software updates [12].

4.2. Single Agent vs. Multi-Agent Frameworks

Single-agent bandit architectures possess distinct advantages, primarily their computational simplicity, zero communication overhead, and absolute data privacy, as no information ever leaves the local node [15]. However, in a nationwide network like BPJS Kesehatan, their critical weakness is isolated learning. If BPJS deployed an independent single-agent Contextual Bandit at each regional branch, each node would have to "pay the cost" of exploring new fraud schemes. Recent studies on federated bandits demonstrate that in the absence of information exchange with neighbors, an isolated agent can suffer a severe linear cumulative regret of $O(T)$ [16].

Our proposed MAMAB architecture overcomes this limitation through collaborative intelligence. Literature demonstrates that even a minimal level of collaboration among different agents enables a significant reduction in per-agent regret compared to a scenario where agents do not communicate [15]. Furthermore, deploying multi-agent contextual bandits has been proven to produce faster convergence and greater adaptability than single-agent baselines in complex environments [14]. However, sharing raw patient data across branches violates Indonesia's Personal Data Protection Law (UU PDP).

By integrating Federated Learning into the MAMAB framework, we achieve the mathematical benefits of multi-agent collaboration without the privacy risks [17]. Because the agents only transmit their locally updated parameter weights θ_{local} to the central BPJS server, the "exploration cost" is distributed across the entire national network. Once a single branch identifies a new upcoding vector, the aggregated global policy θ_{global} instantly immunizes all other branches against the same adversarial strategy.

4.3. Addressing Asymmetric Costs and Class Imbalance

A significant disadvantage of the standard supervised learning for fraud detection is how it cannot deal with severe class imbalance. In national health insurance system, fraudulent claims are far less common than legitimate ones. Thus, traditional algorithms try to achieve high accuracy by classifying almost all transactions as legitimate and completely miss the rare but financially devastating fraud cases [10]. Furthermore, all prediction errors are treated as equally important in standard classification models. However, the consequences in health care finance are not evenly distributed. Failing to detect a large upcoding scheme may cost much more than temporarily sending a legitimate claim for manual review [11].

Our proposed MAMAB architecture mathematically resolves both class imbalance and asymmetric costs by abandoning standard accuracy metrics in favor of value-driven reward maximization. In our Contextual Bandit formulation, the algorithm does not merely ask, "Is this claim fraud?" Instead, the reward function evaluates the actual monetary value of the claim ($r_t = y_t \times v_t$). This inherently transforms the system into a cost-sensitive learning framework. The agent learns to prioritize the auditing of high-value, high-risk claims where the expected financial recovery outweighs the operational cost of the human auditor. By explicitly aligning the algorithm's remuneration mechanism with actual financial risk, the MAMAB environment naturally addresses statistical imbalances and directly safeguards public health budgets.

If this framework is implemented, BPJS Kesehatan will be able to quickly detect cases of fraud and optimize the use of its limited audit resources while ensuring new fraud patterns are shared with all regional branches without violating patient data privacy regulations.

4.4. Future Directions and Conclusion

BPJS Kesehatan faces threats from fraud that never stops evolving. Sadly, regular audit systems can't keep up with this change, limited by their resources and tight privacy regulations. Consequently, we suggest MAMAB to address the problem. In addition, the system can exchange info on newly identified fraud patterns without compromising patients' underlying data. By combining all the above, our creation tackles a substantial problem with extreme real-life applicability.

As this paper implies a preliminary architectural study, empirical simulation remains in progress but has not yet been finished. Future research will focus on empirical validation using real-world datasets with the aim of evaluating the accuracy of fraud detection and identifying the system's limitations. Additionally, future research aims to expand the MAMAB system to cover other fraud cases in the Indonesian healthcare system while assessing the scalability potential of the decentralized synchronization process.

Data Availability

No data was used for the research described in this article. The study focuses on the theoretical design and architectural proposition of a Multi-Agent Multi-Armed Bandit framework. Details regarding the required dataset schema and simulations are reserved for future work.

Author Contribution Statement

The study was designed and written by **E. V. Nilasari** and **B. Soelistijanto**. All authors read and approved the manuscript.

References

[1] Kumar P. Major Challenges for Developing Countries: Obstacles to Growth. *Journalism University*; 2026. https://journalism.university/journalistic-writings/major-challenges-developing-countries-growth/

[2] Media KC. Daftar 10 Kasus Korupsi Terbesar di Indonesia. *KOMPAS.com*; 2025. https://nasional.kompas.com/read/2025/02/27/14210521/daftar-10-kasus-korupsi-terbesar-di-indonesia

[3] Nicoamon. Apa Itu INA-CBGs? Panduan Sistem Klaim BPJS (2026). *DHealth*; 2024. https://www.dhealth.co.id/post/apa-itu-ina-cbgs-istilah-yang-ada-pada-bpjs

[4] CNN Indonesia. Kemenkes Ungkap Jenis Fraud Klaim BPJS, Sudah Siapkan Sanksi. *CNN Indonesia*; 2024. https://www.cnnindonesia.com/nasional/20240724181909-12-1125147/kemenkes-ungkap-jenis-fraud-klaim-bpjs-sudah-siapkan-sanksi

[5] Untari DS, Hestin P. Sederet Masalah BPJS Kesehatan: Tagihan Fiktif hingga Asuransi Swasta Mundur dari Kerja Sama. *Bisnis.com*; 2024. https://finansial.bisnis.com/read/20240725/215/1785237/sederet-masalah-bpjs-kesehatan-tagihan-fiktif-hingga-asuransi-swasta-mundur-dari-kerja-sama

[6] Milovich D. Korupsi BPJS Kesehatan: KPK Ungkap Fraud Klaim Rp 34 Miliar dari Markup dan Phantom Billing di Tiga Rumah Sakit. *Indoraya News*; 2025. https://indoraya.news/korupsi-bpjs-kesehatan-kpk-ungkap-fraud-klaim-rp-34-miliar-dari-markup-dan-phantom-billing-di-tiga-rumah-sakit

[7] Slivkins A. Introduction to Multi-Armed Bandits. *Now Publishers*; 2019. https://doi.org/10.1561/9781680836219

[8] Vlassis N. A Concise Introduction to Multiagent Systems and Distributed Artificial Intelligence. *Synthesis Lectures on Artificial Intelligence and Machine Learning*. Cham: Springer International Publishing; 2007. https://doi.org/10.1007/978-3-031-01543-4

[9] Tito J, Siregar K. Faktor Pemicu dan Penghambat Fraud dalam Program Jaminan Kesehatan Nasional dan Strategi Pencegahannya: Sebuah Scoping Review. *EKI*. 2024;**9**(2).

[10] Wang C, Nie C, Liu Y. Evaluating Supervised Learning Models for Fraud Detection: A Comparative Study of Classical and Deep Architectures on Imbalanced Transaction Data. In: Siuta-Tokarska B, Grigorescu A, Habib M, Zhu Y, editors. *Proceedings of the 2025 3rd International Academic Conference on Management Innovation and Economic Development (MIED 2025)*. Dordrecht: Atlantis Press International BV; 2025. p. 613–24.

[11] VR K, S K, S P, Sangeetha MJ. A High-Recall Cost-Sensitive Machine Learning Framework for Real-Time Online Banking Transaction Fraud Detection. *Authorea*; 2026. https://doi.org/10.22541/au.176824649.95910513/v2

[12] Gao R. Optimizing Credit Card Fraud Detection with Multi-Armed Bandit Algorithms. In: *Proceedings of the 1st International Conference on Engineering Management, Information Technology and Intelligence*. Shanghai: SCITEPRESS - Science and Technology Publications; 2024. p. 474–8. https://doi.org/10.5220/0012956000004508

[13] Soemers DJNJ, Brys T, Driessens K, Winands MHM, Nowé A. Adapting to concept drift in credit card transaction data streams using contextual bandits and decision trees. *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence*. 2018;**32**(1):7831–6. https://dl.acm.org/doi/10.5555/3504035.3504998

[14] Zhu Z. Multi-Agent Contextual Bandits for Multi-Objective Recommendation Systems. *ITM Web Conf*. 2025;**78**:01033.

[15] Sankararaman A, Ganesh A, Shakkottai S. Social Learning in Multi Agent Multi Armed Bandits. *Proc ACM Meas Anal Comput Syst*. 2019;**3**(3):53:1–35

[16] Zhang H, Xu Y, Wang X, Chen HX, Qiu H, Yang L, et al. Federated Multi-armed Bandits with Efficient Bit-Level Communications. *OpenReview*; 2025. https://openreview.net/forum?id=o3x3RV8XHd

[17] Shao J, Wang S, Fang Z. Learning with Limited Shared Information in Multi-agent Multi-armed Bandit. In: *International Joint Conference on Autonomous Agents and Multiagent Systems*. New York; 2025. p. 1914–22. https://doi.org/10.48550/arXiv.2502.15338

The 11th International Conference on Computer Science and Computational Intelligence 2026

Certificate of Appreciation

089/ICCSCI/AUTHOR/VI/2026

Presented to

Elisabeth Victoria Nilasari and Bambang Soelistijanto

Paper ID : 580

Paper Title : A Preliminary Study on Mitigating Upcoding in Indonesia's National Health Insurance using Multi-Agent Multi-Armed Bandits

at the 11th International Conference on Computer Science and Computational Intelligence 2026, held on 18 June – 19 June 2026

People
Innovation
Excellence

Dr. Mochammad Haldi Widianto
General Chair of ICCSCI 2026

Prof. Derwin Suhartono
Dean - School of Computer Science