2024 IEEE International Conference on Data Mining Workshops (ICDMW 2024)

Abu Dhabi, United Arab Emirates
9 December 2024

Pages 1-452

IEEE

IEEE Catalog Number: CFP2456B-POD
ISBN: 979-8-3315-3064-8

Copyright © 2024 by the Institute of Electrical and Electronics Engineers, Inc.
All Rights Reserved

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

****** This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number:	CFP2456B-POD
ISBN (Print-On-Demand):	979-8-3315-3064-8
ISBN (Online):	979-8-3315-3063-1
ISSN:	2375-9232

Additional Copies of This Publication Are Available From:

Curran Associates, Inc
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: (845) 758-0400
Fax: (845) 758-2633
E-mail: curran@proceedings.com
Web: www.proceedings.com

CURRAN ASSOCIATES INC.
proceedings .com

24th IEEE International Conference on Data Mining Workshops (ICDMW)

ICDMW 2024

Table of Contents

Message from Workshop Chairs

OEDM: Optimization Based Techniques for Emerging Data Mining Problems

From Connections to Conversions: Enhanced Product Knowledge Graph Embeddings for E-commerce Recommendations 1
Abhishek Kulkarni (Walmart Global Tech, India), Sumit Kumar Dutta (Walmart Global Tech, India), Sivaram Prasad Mudunuri (Walmart Global Tech, India), Kamiya Motwani (Walmart Global Tech, India), and Rahul Ghosh (Walmart Global Tech, India)

An Adaptive Method Stabilizing Activations for Enhanced Generalization 9
Hyunseok Seung (University of Georgia, USA), Jaewoo Lee (University of Georgia, USA), and Hyunsuk Ko (Hanyang University, South Korea)

Outlier Exposure in Visual Industrial Anomaly Detection - Survey 17
Shi Yong (School of Economics and Management, Chinese Academy of Sciences, China), Xu Ruijie (School of Computer Science and Technology, Chinese Academy of Sciences, China), and Qi Zhiquan (School of Economics and Management, Chinese Academy of Sciences, China)

Structure Aware Loss for Deep Unsupervised Industrial Image Registration 25
Yong Shi (Research Center on Fictitious Economy and Data Science, Chinese Academy of Sciences, China; Key Laboratory of Big Data Mining and Knowledge Management, Chinese Academy of Sciences, China; University of Nebraska at Omaha, USA), Mengjin Lyu (University of Chinese Academy of Sciences, China), and Zhiquan Qi (University of Chinese Academy of Sciences, China; Research Center on Fictitious Economy and Data Science, Chinese Academy of Sciences, China; Key Laboratory of Big Data Mining and Knowledge Management, Chinese Academy of Sciences, China)

Detecting Multi-Modal Fake AIGC from the Perspective of Emotional Manipulation Purpose 33
Linzi Zhang (Communication University of China, China)

ARRL: Second International Workshop on Adaptable, Reliable, and Responsible Learning

Urban Functional Zone Delineation Aided by Vision Foundation Models 40
Heng Lian (William & Mary) and Yi He (William & Mary)

Differential Evolution Integrated Hybrid Deep Learning Model for Object Detection in Pre-Made Dishes 44
Lujia Lv (Southwest University, China), Di Wu (Southwest University, China), Yangyi Xia (Southwest University, China), Jia Wu (Southwest University, China), Xiaojing Liu (Chongqing True Flavor Food Co LTD, China), and Yi He (William & Mary, United States of America)

Efficient and High-Quality Formal Verification for Decision Tree Ensembles 51
Saori Matsunaga (Information Technology R&D Center, Mitsubishi Electric Corporation, Japan) and Genta Yoshimura (Information Technology R&D Center, Mitsubishi Electric Corporation, Japan)

PriSM: Privacy-Preserving Social Media Text Processing and Analytics Framework 59
Sakib Shahriar (University of Guelph, Canada) and Rozita Dara (University of Guelph, Canada)

Adaptable and Reliable Text Classification using Large Language Models 67
Zhiqiang Wang (Florida Atlantic University, USA), Yiran Pang (Florida Atlantic University, USA), Yanbin Lin (Florida Atlantic University, USA), and Xingquan Zhu (Florida Atlantic University, USA)

WAIN-24: 4th International Workshop on AI for Nudging and Personalization

Care-by-Design: Enhancing Customer Experience through Empathetic AI Chatbots 75
Kadhim Hayawi (AICEO, UAE) and Sakib Shahriar (AICEO, UAE)

Customer Engagement Journey Driven Propensity Model for Home Page Composition 82
Venkataramana Kini (Prime Video, USA), Charlie Xu (Prime Video, USA), Ravi Divvela (Prime Video, USA), Unmesh Phadke (Prime Video, USA), Fei Wang (Prime Video, USA), and Zhen Wen (Prime Video, USA)

Mamba for Scalable and Efficient Personalized Recommendations 89
Andrew Starnes (Lirio, Inc., USA) and Clayton Webster (Lirio, Inc., USA)

DMF 2024: The 2nd International Workshop on Data Mining in Finance

Leveraging Large Language Models for Predicting Stock Option Valuation and Financial Risk Mitigation 97
Lester Dsouza (University of Galway, Ireland), Jamal Abdul Nasir (University of Galway, Ireland), Muhammad Mohsin Kamal (Knowledge Platform, Pakistan), and Lena Connolly (Zayed University, UAE)

Metal Price Spike Prediction via a Neurosymbolic Ensemble Approach 106
Nathaniel Lee (Arizona State University, USA), Noel Ngu (Arizona State University, USA), Harshdeep Singh Sahdev (Arizona State University, USA), Pramod Motaganahall (Arizona State University, USA), Al Mehdi Saadat Chowdhury (Arizona State University, USA), Bowen Xi (Arizona State University, USA), and Paulo Shakarian (Arizona State University, USA)

Sentiment Score of Bloomberg Market Wraps with ChatGPT 111
Baptiste Lefort (AI for Alpha and CentraleSupelec), Eric Benhamou (Ai For Alpha, France), Jean-Jacques Ohana (Ai For Alpha, France), Beatrice Guez (Ai For Alpha, France), and David Saltiel (Ai For Alpha, France)

The Chain of Narrative Framework: Advancing Controllable Generation of Financial Themes 117
Shuoling Liu (Hong Kong University of Science and Technology, China), Gaoguo Jia (E Fund Management Co., Ltd, China), Dongsheng Gu (E Fund Management Co., Ltd, China), Liyuan Chen (E Fund Management Co., Ltd, China), and Qiang Yang (Hong Kong University of Science and Technology, China)

Sequential Block Elimination for Dynamic Pricing 125
Shameem A Puthiya Parambath (University of Glasgow, UK), Saleh Abdullah M Alfahad (University of Glasgow, UK), Christos Anagnostopoulos (University of Glasgow, UK), and Kostas Kolomvatsos (University of Thessaly, Greece)

Multi-Agent Based Casual Triple Extraction for Factuality Evaluation using Large Language Models 130
Jian Zhang (University of Science and Technology of China & Commind Technology Co., Ltd., China), Zemin Xu (Commind Technology Co., Ltd., China), Lifang Liu (Commind Technology Co., Ltd., China), Zhanfeng Shen (Commind Technology Co., Ltd., China), Yue Cui (The Hong Kong University of Science and Technology, China), and Yongdong Zhang (University of Science and Technology of China, China)

MM4SG 2024: Workshop on Multimodal Content Analysis for Social Good

Context-Aware Offensive Meme Detection: A Multi-Modal Zero-Shot Approach with Caption-Enhanced Classification 137
Faseela Abdullakutty (Qatar University, Qatar), Somaya Al-Maadeed (Qatar University, Qatar), and Usman Naseem (Macquarie University, Australia)

Fine-Tuning LLMs for Reliable Medical Question-Answering Services 146
Ali Anaissi (The University of Sydney, Australia; University of Technology Sydney, Australia), Ali Braytee (University of Technology Sydney, Australia), and Junaid Akram (The University of Sydney, Australia)

Bengali Cyberbullying: Detection, Categorization, and Gender Bias Analysis 154
Md. Mithun Hossain (Bangladesh University of Business and Technology, Bangladesh), Md. Shakil Hossain (Bangladesh University of Business and Technology, Bangladesh), Sudipto Chaki (Bangladesh University of Business and Technology, Bangladesh), Md. Saifur Rahman (Bangladesh University of Business and Technology, Bangladesh), and Mohammad Ali Moni (Charles Sturt University, Australia; Washington University of Science and Technology, USA; The University of Queensland, Australia)

Optimized Biomedical Question-Answering Services with LLM and Multi-BERT Integration 164
Cheng Qian (The University of Sydney, Australia), Xianglong Shi (The University of Sydney, Australia), Shanshan Yao (The University of Sydney, Australia), Yichen Liu (The University of Sydney, Australia), Fengming Zhou (The University of Sydney, Australia), Zishu Zhang (The University of Sydney, Australia), Junaid Akram (The University of Sydney, Australia), Ali Braytee (University of Technology Sydney, Australia), and Ali Anaissi (The University of Sydney, Australia; University of Technology Sydney, Australia)

UMML: Layout-Aware Text-Image Fusion for Unified Multilingual Multimodal Learning 174
Hariram Veeramani (UCLA, USA), Surendrabikram Thapa (Virginia Tech, USA), Mohammad Salman (University of Stavanger, Norway), Mohammad Ali Moni (Charles Sturt University, Australia; Washington University of Science and Technology, USA), and Usman Naseem (Macquarie University, Australia)

Enhanced Multimodal Hate Video Detection via Channel-wise and Modality-wise Fusion 183
Yinghui Zhang (University of Exeter, UK), Tailin Chen (University of Exeter, UK), Yuchen Zhang (University of Exeter, UK; University of Essex, UK), and Zeyu Fu (University of Exeter, UK)

IncrLearn: Incremental classification and clustering, concept drift, novelty detection, active learning in big/fast data context

A New Index for Computer Science Conferences and Its Applications: The S-Index 191
Sharareh Alipour (Khatan university, Iran) and Sobhan Ebrahimi (Khatan university, Iran)

Incremental Gaussian Mixture Clustering for Data Streams 197
Aniket Bhanderi (University of Cincinnati, USA) and Raj Bhatnagar (University of Cincinnati, USA)

Non-Stationary Imbalanced Data Stream Classification with Weighted One-Class Ensemble 205
Jakub Klikowski (Wroclaw University of Science and Technology, Poland) and Michal Wozniak (Wroclaw University of Science and Technology, Poland)

Classification of Recurrent Concepts with Metafeature-Based Model Selection 213
Joanna Komorniczak (Wrocław University of Science and Technology, Poland) and Paweł Ksieniewicz (Wrocław University of Science and Technology, Poland)

Semi-Supervised Learning Trumps Active Learning in Cold-Starting Netflow-Based AI Network Intrusion Detection Systems on Scarce and Difficult Data 221
Marek Pawlicki (Bydgoszcz University of Science and Technology, Poland), Aleksandra Pawlicka (University of Warsaw, Poland), Ryszard Choraś (Bydgoszcz University of Science and Technology, Poland), Rafał Kozik (Bydgoszcz University of Science and Technology Bydgoszcz, Poland), and Michał Choraś (Bydgoszcz University of Science and Technology, Poland)

Online Adaptive Local Interpretable Tree Ensembles for Time Series Forecasting 229
Amal Saadallah (Lamarr Institute for Machine Learning and Artificial Intelligence, Germany)

Changing Lineup Classifier Ensemble for Drifting Imbalanced Data Streams 238
Weronika Węgier (Wrocław University of Science and Technology, Poland), Maciej Mączyński (Wrocław University of Science and Technology, Poland), and Michał Woźniak (Wrocław University of Science and Technology, Poland)

Cross-Modality Clustering-Based Self-Labeling for Multimodal Data Classification 246
Pawel Zyblewski (Wroclaw University of Science and Technology, Poland) and Leandro Minku (University of Birmingham, United Kingdom)

AI4FCF : Workshop on AI for Financial Crime Fight

Amatriciana: Exploiting Temporal GNNs for Robust and Efficient Money Laundering Detection ... 254
Marco Di Gennaro (Politecnico di Milano, Italy), Francesco Panebianco (Politecnico di Milano, Italy), Marco Pianta (Politecnico di Milano, Italy), Stefano Zanero (Politecnico di Milano, Italy), and Michele Carminati (Politecnico di Milano, Italy)

GAN+ABM to Simulate Suspicious Money Laundering Trajectories 262
Oscar M. Granados (Universidad Jorge Tadeo Lozano, Colombia), Olmer García-Bedoya (Universidad Jorge Tadeo Lozano, Colombia), and Edwin González (Universidad Konrad Lorenz, Colombia)

Towards Transparent AI-Powered Cybersecurity in Financial Systems: The Deployment of Federated Learning and Explainable AI in the CaixaBank Pilot 270
Aikaterini Karampasi (University of Western Macedonia, Greece), Panagiotis Radoglou-Grammatikis (University of Western Macedonia, Greece), Marek Pawlicki (ITTI sp. z o.o., Poland; PBS Bydgoszcz University of Science and Technology, Poland), Ryszard Choraś (PBS Bydgoszcz University of Science and Technology, Poland), Ramon Martin de Pozuelo (CaixaBank, Spain), Panagiotis Sarigiannidis (University of Western Macedonia, Greece), Damian Puchalski (ITTI sp. z o.o., Poland), Aleksandra Pawlicka (ITTI sp. z o.o., Poland; University of Warsaw, Poland), Rafał Kozik (ITTI sp. z o.o., Poland; PBS Bydgoszcz University of Science and Technology, Poland), and Michał Choraś (ITTI sp. z o.o., Poland; PBS Bydgoszcz University of Science and Technology, Poland)

Anomaly Detection in Financial Data using GA-Optimized MLP Models 278
Ivan Malashin (Bauman Moscow State Technical University, Russia), Igor Masich (Bauman Moscow State Technical University, Russia), Aleksei Borodulin (Bauman Moscow State Technical University, Russia), Vladimir Bukhtoyarov (Bauman Moscow State Technical University, Russia), Vladimir Nelyub (Bauman Moscow State Technical University, Russia), and Vadim Tynchenko (Bauman Moscow State Technical University, Russia)

Multi-Task CNN Behavioral Embedding Model for Transaction Fraud Detection 286
Bo Qu (eBay, China), Zhurong Wang (eBay, USA), Minghao Gu (eBay, USA), Daisuke Yagi (eBay, USA), Yang Zhao (eBay, China), Yinan Shan (eBay, China), and Frank Zahradnik (eBay, USA)

BigIS 2024: Workshop on Information Seeking with Big Models

Emotion Retrieval using Generative Models: An Exploratory Study 293
M. Luqman Jamil (University of Beira Interior, Portugal), Sebastião Pais (University of Beira Interior, Portugal), Andrés Caro (Universidad de Extremadura, Spain), and João Cordeiro (University of Beira Interior, Portugal)

Foundation Models for Course Equivalency Evaluation 300
Mark Kim (San Francisco State University, USA), Arno Puder (San Francisco State University, USA), Craig Hayward (Kern Community College District, USA), and Hui Yang (San Francisco State University, USA)

SRSA: A Cost-Efficient Strategy-Router Search Agent for Real-World Human-Machine Interactions 307
Yaqi Wang (Carnegie Mellon University, USA) and Haipei Xu (Columbia University, USA)

HDM: The 11-th ICDM Workshop on High Dimensional Data Mining

Do Good Scores Imply Good Embeddings? On the Necessity of Inspecting Manifold Learning Evaluation Results by Multiple Criteria 317
Daniyal Kazempour (Kiel University, Germany), Claudius Zelenka (Kiel University, Germany), Atakan Kara (Kiel University, Germany), Andreas Lohrer (Kiel University, Germany), and Peer Kröger (Kiel University, Germany)

Not the Same, Yet Similar? - On Commonalities and Differences between VQPCA and ORCLUS ... 325
Daniyal Kazempour (Kiel University, Germany), Claudius Zelenka (Kiel University, Germany), Artjom Mukhamedov (Kiel University, Germany), and Peer Kröger (Kiel University, Germany)

Data Complexity via Random Projection 333
Yamonporn Thummanusarn (University of Birmingham, UK) and Ata Kabán (University of Birmingham, UK)

Resource Efficient Delaunay Triangulation for Higher Dimensional Topological Studies 341
Anurag Yadav (University of Cincinnati, USA), Rohit P. Singh (Univ of Wisconsin–Milwaukee, USA), and Philip A. Wilsey (University of Cincinnati, USA)

Prediction Is NOT Classification: On Formulation and Evaluation of Hyperedge Prediction 349
Taehyung Yu (KAIST, Republic of Korea), Soo Yong Lee (KAIST, Republic of Korea), Hyunjin Hwang (KAIST, Republic of Korea), and Kijung Shin (KAIST, Republic of Korea)

AIDM-AS: Advances in AI-Driven Data Mining for Autonomous Systems

Trust-Based Positional Forgery Detection in AI-Driven Autonomous Intersection Management using Hybrid Graph-Based Reinforcement Learning 357
Shajina Anand (Seton Hall University)

Real-World Application Facilitating Trustworthy Human-Robot Collaboration with Innovative Explainable Neuro-Symbolic Reasoning 364
Rafał Kozik (Bydgoszcz University of Science and Technology, Poland; ITTI Sp. z.o.o, Poland), Szymon Buś (Sieć Badawcza Łukasiewicz Przemysłowy Instytut Automatyki i Pomiarów PIAP, Poland), Grzegorz Gawdzik (Sieć Badawcza Łukasiewicz Przemysłowy Instytut Automatyki i Pomiarów PIAP, Poland), Marek Pawlicki (Bydgoszcz University of Science and Technology, Poland; ITTI Sp. z.o.o, Poland), Aleksandra Pawlicka (University of Warsaw, Poland; ITTI Sp. z.o.o, Poland), and Michał Choraś (Bydgoszcz University of Science and Technology, Poland; ITTI Sp. z.o.o, Poland)

Energy-Efficiency Management Framework of a Ship Based on Data Classification Technique 372
Tien Anh Tran (Vietnam Maritime University, Vietnam) and Sathishkumar Veerappampalayam Easwaramoorthy (Sunway University, Malaysia)

ETDLH 2024: Workshop on Emerging Trends in Deep Learning for Healthcare

GDRNet: A Novel Graph Neural Network Architecture for Diabetic Retinopathy Detection 378
Shahed Hossain (Daffodil International University, Bangladesh), Md. Zahid Hasan (Daffodil International University, Bangladesh), Risul Islam Jim (Daffodil International University, Bangladesh), Taslima Ferdaus Shuva (Daffodil International University, Bangladesh), Md. Tanvir Rahman (The University of Queensland, Australia), Abdullah Al-Mamun Bulbul (The University of Queensland, Australia), Risala Tasin Khan (Jahangirnagar University, Bangladesh), M. Shamim Kaiser (Jahangirnagar University, Bangladesh), and Mohammad Ali Moni (Charles Sturt University, Australia; Washington University of Science and Technology, USA; The University of Queensland, Australia)

CardioNetFusion: A Deep Learning and Explainable AI Integrated System for Cardiovascular Disease Detection 388
Gunasekaran Raja (Anna University, India), Selvam Essaky (Anna University, India), Surya Raj Muthu (Anna University, India), Babith Sarish Suresh (Anna University, India), Jaehyuk Cho (Jeonbuk National University, The Republic of Korea), Seohyun Yoo (Jeonbuk National University, The Republic of Korea), and Sathiskumar Veerappampalayam Easwaramoorthy (Sunway University, Malaysia)

Smart Contract-Driven Pandemic Management using Blockchain Technology and Artificial Intelligence 396
Sathishkumar Veerappampalayam Easwaramoorthy (Sunway University, Jalan Universiti, Malaysia), Neelakandan Subramani (Department of Computer Science and Engineering, R.M.K Engineering College, India), and Angela Lee Siew Hoong (Sunway University, Jalan Universiti, Malaysia)

Accessible Health Screening using Body Fat Estimation by Image Segmentation 405
Zhipeng Yin (Florida International University, U.S.A), Sameeksha Agarwal (Carnegie Mellon University, U.S.A), Ayesha Kashif (Jose Marti MAST 6 - 12 Academy, U.S.A), Matthew Gonzalez (TERRA Environmental Research Institute, U.S.A), Zichong Wang (Florida International University, U.S.A), Suqing Liu (Guangdong Ocean University, China), Zhen Liu (Guangdong University of Foreign Studies, China), Yanzhao Wu (Florida International University, U.S.A), Ian Stockwell (University of Maryland, U.S.A), Weifeng Xu (University of Baltimore, U.S.A), Puqing Jiang (Huazhong University of Science and Technology, China), Xingyu Zhang (University of Pittsburgh, U.S.A), and Wenbin Zhang (Florida International University, U.S.A)

SENTIRE: Sentiment Elicitation from Natural Text for Information Retrieval and Extraction

Multilingual Sentiment Analysis for Investigating Perceptions of Globalization 415
Anagha Ani (Nanyang Technological University, Singapore) and Erik Cambria (Nanyang Technological University, Singapore)

Into the Unknown: Self-Learning Large Language Models 423
Teddy Ferdinan (Wroclaw Tech, Poland), Jan Kocoń (Wroclaw Tech, Poland), and Przemysław Kazienko (Wroclaw Tech, Poland)

Automating Maritime Risk Data Collection and Identification Leveraging Large Language Models 433
Donghao Huang (Singapore Management University, Singapore; Mastercard, Singapore), Xiuju Fu (Agency for Science, Technology and Research, Singapore), Xiaofeng Yin (Agency for Science, Technology and Research, Singapore), Haibo Pen (Tianjin University, China), and Zhaoxia Wang (Singapore Management University, Singapore)

A Silent Voice: A Bilingual Longitudinal Analysis of Mahsa Amini's Movement Tweet? 440
Sadegh Jafari (Iran University Science and Technology, Iran), Mahsa Farsouni Nezhad (Allameh Tabatabai University, Iran), Melika Nobakhtian (Khatam University (TeIAS), Iran), Arash Chaichi Mellatshahi (Ghent University, Belgium), Amirmohammad Kazemeini (Vector institute, Canada), and Erik Cambria (Nanyang Technological University, Singapore)

Converting Vocal Performances into Sheet Music Leveraging Large Language Models 445
Jinjing Jiang (Research and Development, Mastercard, Singapore), Nicole Teo (Singapore Management University, Singapore), Haibo Pen (Tianjin University, China), Seng-Beng Ho (AI Institute Global, Singapore), and Zhaoxia Wang (Research and Development, Mastercard, Singapore)

Comprehensive Sentiment Analysis of Polish Book Reviews using Large and Small Language Models 453
Agnieszka Karlińska (NASK National Research Institute, Poland), Piotr Miłkowski (Wroclaw Tech, Poland), Paulina Czwordon-Lis (Institute of Literary Research of the Polish Academy of Sciences, Poland), Bartłomiej Koptyra (Wroclaw Tech, Poland), and Jan Kocoń (Wroclaw Tech, Poland)

Small Language Models for Emotion Recognition in Polish Stock Market Investor Opinions 463
Bartłomiej Koptyra (Wroclaw Tech, Poland), Marcin Oleksy (Wroclaw Tech, Poland), Ewa Dzięcioł (Wroclaw Tech, Poland), and Jan Kocoń (Wroclaw Tech, Poland)

Investigating the Role of Metaphors in Human-AI Conversation and Engagement 471
Pragati Maheshwary (Carnegie Mellon University, United States), Qiawen Liu (University of Wisconsin-Madison, United States), and Rui Mao (Nanyang Technological University, Singapore)

Sentiment Analysis on Climate Change for Sustainable Investment 479
Keane Ong (National University of Singapore), Rui Mao (Nanyang Technological University), Deeksha Varshney (National University of Singapore), Ranjan Satapathy (Institute of High Performance Computing, Agency for Science, Technology and Research), Erik Cambria (Nanyang Technological University), and Gianmarco Mengaldo (National University of Singapore)

Information Extraction: Unstructured to Structured for ESG Reports 487
*Zounachuan Sun (National University of Singapore, Singapore), Ranjan Satapathy (Institute of High-Performance Computing (IHPC), Agency for Science, Technology and Research (A*STAR), Singapore), Daixue Guo (National University of Singapore, Singapore), Bo Li (National University of Singapore, Singapore), Xinyuan Liu (National University of Singapore, Singapore), Yangchen Zhang (National University of Singapore, Singapore), Cheng-ann Tan (DHI Group, Singapore), Ricardo Shirota Filho (Institute of High-Performance Computing (IHPC), Agency for Science, Technology and Research (A*STAR), Singapore), and Rick Siow Mong Goh (Institute of High-Performance Computing (IHPC), Agency for Science, Technology and Research (A*STAR), Singapore)*

Evaluating Vision Language Models in Detecting Learning Engagement 496
Jayant Teotia (Nanyang Technological University, Singapore), Xulang Zhang (Nanyang Technological University, Singapore), Rui Mao (Nanyang Technological University, Singapore), and Erik Cambria (Nanyang Technological University, Singapore)

Neurosymbolic AI for Mining Key Aspects of Socially Responsible Investing 503
*Wihan van der Heever (Nanyang Technological University, Singapore), Ranjan Satapathy (Agency for Science, Technology and Research (A*STAR), Singapore), Erik Cambria (Nanyang Technological University, Singapore), and Rick Siow Mong Goh (Agency for Science, Technology and Research (A*STAR), Singapore)*

Personalized Large Language Models 511
Stanisław Woźniak (Department of Artificial Intelligence, Wroclaw Tech, Poland), Bartłomiej Koptyra (Department of Artificial Intelligence, Wroclaw Tech, Poland), Arkadiusz Janz (Department of Artificial Intelligence, Wroclaw Tech, Poland), Przemysław Kazienko (Department of Artificial Intelligence, Wroclaw Tech, Poland), and Jan Kocoń (Department of Artificial Intelligence, Wroclaw Tech, Poland)

Understanding the Hidden State of Language Models from the Perspective of Sentiment Analysis 521
Kaicheng Xue (University of Wisconsin–Madison, United States), Xulang Zhang (Nanyang Technological University, Singapore), Rui Mao (Nanyang Technological University, Singapore), and Erik Cambria (Nanyang Technological University, Singapore)

EDMML: 4th Workshop on Evolutionary Data Mining and Machine Learning

Optimizing MLP Network Structure for Classification Problems using PSO and Dominating Vertex Set 529
Ines Hadjadji (University of Badji Mokhtar-Annaba, Algeria), Nour Elislem Karabadji (National Higher School of Technology and Engineering, Algeria), Hassina Seridi (University of Badji Mokhtar-Annaba, Algeria), Naouel Manaa (University of Badji Mokhtar-Annaba, Algeria), Mohamed Elati (University of Lille, France), and Wajdi Dhifli (University of Lille, France)

TRUST: Responsible AI to Increase Clinical Decision Trust: Explainability & Reliability of Machine Learning Models

Trust in Machine Learning Models: Global and Individual Assessment 535
Simão Paredes (CISUC), Teresa Rocha (CISUC), Sónia Sousa (Center for Informatics and Systems of University of Coimbra), Jorge Henriques (Center for Informatics and Systems of University of Coimbra), José Sousa (Instituto Português de Oncologia do Porto, Porto), and Lino Gonçalves (Centro Hospitalar e Universitário de Coimbra)

Enhancing Clinical Trust: The Role of AI Explainability in Transforming Healthcare 543
Aleksandra Pawlicka (University of Warsaw, Poland), Marek Pawlicki (Bydgoszcz University of Science and Technology, Poland), Rafał Kozik (Bydgoszcz University of Science and Technology, Poland), Dagmara Jaroszewska-Choraś (Kazimierz Wielki University, Poland), and Michał Choraś (Bydgoszcz University of Science and Technology, Poland)

Pointwise Reliability Metrics for Machine Learning Model Predictions: Three Different Approaches 550
José Sousa (Coimbra Engineering Institute, Portugal), Simão Paredes (CISUC, Portugal), Jorge Henriques (Center for Informatics and Systems of University of Coimbra, Portugal), Lino Gonçalves (Coimbra Hospital and University Centre, Portugal), and Teresa Rocha (CISUC, Portugal)

Pointwise Reliability Assessment: Density vs. Local Fit Methods 556
João Trovão (CISUC, Center for Informatics and Systems of the University of Coimbra, Portugal), Jorge Henriques (CISUC, Center for Informatics and Systems of the University of Coimbra, Portugal), Lorena Petrella (CISUC, Center for Informatics and Systems of the University of Coimbra, Portugal), Madalena Trindade (CISUC, Center for Informatics and Systems of the University of Coimbra, Portugal), and João Loureiro (CISUC, Center for Informatics and Systems of the University of Coimbra, Portugal)

Reliability Assessment of Models for Risk Stratification: A Knowledge-Based Approach 564
Z. Zhao (University of Coimbra, Portugal), P. Gil (University of Coimbra, Portugal; NOVA School of Sci. and Technology, Portugal), J. Loureiro (University of Coimbra, Portugal), L. Petrella (University of Coimbra, Portugal), and J. Henriques (University of Coimbra, Portugal)

DMS2024: Data Mining for Service

Generative-Graph Method for Project Duration Estimation 572
Sandeep Chougule (TCS Research, India) and Vaishali Sadaphal (TCS Research, India)

A Comparative Analysis of User-Generated Content Across Gaming Platforms using Deep Learning-Based Topic Modeling 580
Robert B. Gommerman (Chuo University, Japan) and Eiji Motohashi (Yokohama National University, Japan)

Prediction and Interpretation of Vacancy Duration in Rental Properties using Geographic Data 588
Genki Katsumata (Yokohama National University, Japan) and Eiji Motohashi (Yokohama National University, Japan)

Data Augmentation using Generative AI In Tomato Leaf Lesion Classification Problems 595
Masaya Oirase (Nagoya University; Chubu Electric Power Co. Inc, Japan) and Eisuke Kita (Nagoya University, Japan)

Mining Patterns for Proactive Management of Procure-to-Pay Exceptions 600
Satya Narayana Samudrala (Digitate, Tata Consultancy Services), Uday Chandra Bhookya (Digitate, Tata Consultancy Services), and Maitreya Natu (Digitate, Tata Consultancy Services)

A Data-Driven Approach for Ancillary Bundle Recommendation to Segmented Users 606
Aakash Swami (TCS Research, India), Narendhar Gugulothu (TCS Research, India), Tirumala V (TCS Research, India), and Sanjay Bhat (TCS Research, India)

Temporal Data Mining in AI-Based Patient Navigation Service 613
Shusaku Tsumoto (Shimane University, Japan), Tomohiro Kimura (Shimane University, Japan), Shoji Hirano (Shimane University, Japan), and Katsutoshi Yada (Kansai University, Japan)

Covid-19 and Multichannel Customers 622
Katsutoshi Yada (Kansai University, Japan), Tomoya Ikeda (Kansai University, Japan), Genki Yoshida (Kansai University, Japan), Takahiro Hakozaki (Kansai University, Japan), Ken Matsumoto (Kansai University, Japan), and Shusaku Tsumoto (Shimane University, Japan)

Discrimination Of Brown Spot Leaf Disease From RGB Data Of Image 626
Nao Yuasa (Nagoya University, Japan), Daisuke Minamisawa (Nagoya University, Japan), Ayana Takai (Nagoya University, Japan), Masaya Oirase (Nagoya University, Japan), and Eisuke Kita (Nagoya University, Japan)

MLC: The 4th Machine Learning for Cybersecurity

Visually Analyze SHAP Plots to Diagnose Misclassifications in ML-Based Intrusion Detection..... 632
Maraz Mia (Department of Computer Science, Tennessee Tech, USA), Mir Mehedi Pritom (Department of Computer Science, Tennessee Tech), Tariqul Islam (Syracuse University, USA), and Kamrul Hasan (Tennessee State University, USA)

Enhancing IoT Intrusion Detection through Vertical Federated Learning and Latent Space Mapping 642
Krishna Yadav (SAI Lab, India), Harshit Chaubey (IIIT Naya Raipur, India), and Rajesh Sah (SAI Lab, USA)

Uncovering Attacks and Defenses in Secure Aggregation for Federated Deep Learning 650
Yiwei Zhang (Purdue University, USA), Rouzbeh Behnia (University of South Florida, USA), Attila Yavuz (University of South Florida, USA), Reza Ebrahimi (University of South Florida, USA), and Elisa Bertino (Purdue University, USA)

DCAI: The 3rd InternationalWorkshop on Data-Centric AI

Anomaly Detection and Interpretation from Tabular Data using Transformer Architecture 657
Hajar Homayouni (San Diego State University), Hamed Aghayarzadeh (Colorado State University), Indrakshi Ray (Colorado State University), and Hossein Shirazi (San Diego State University)

scReader: Prompting Large Language Models to Interpret scRNA-seq Data 665
Cong Li (Chinese Academy of Sciences; University of Chinese Academy of Sciences), Qingqing Long (Chinese Academy of Sciences), Yuanchun Zhou (Chinese Academy of Sciences; University of Chinese Academy of Sciences), and Meng Xiao (Chinese Academy of Sciences)

Heisenberg Algebra Implementation to AAG Algorithm for Secure Data Transfer in IoT Networks 673
Jolanta Mizera-Pietraszko (Military University of Land Forces, Poland) and Jolanta Tańcula (Opole University, Poland)

Formal Classifier of a Rhythmic Structure Model for Inference about Linguistic Text Properties 682
Jolanta Mizera-Pietraszko (Military University of Land Forces, Poland) and Jolanta Tańcula (Opole University, Poland)

MRDiff: Time Series Anomaly Detection using Multi-Level Reconstruction Diffusion 688
Dagyeong Na (Chung-Ang University, Korea) and Junseok Kwon (Chung-Ang University, Korea)

On the Classification of Weather Based on the Production of Photovoltaic Installations 696
Paweł Parczyk (Wrocław University of Science and Technology, Poland), Tobiasz Puślecki (Wrocław University of Science and Technology, Poland), and Robert Burduk (Wrocław University of Science and Technology, Poland)

DMU2: Workshop on Data Mining for User Understanding

CUPID: A Real-Time Session-Based Reciprocal Recommendation System for a One-on-One Social Discovery Platform 704
Beomsu Kim (Hyperconnect), Sangbum Kim (Hyperconnect), Minchan Kim (Hyperconnect), Joonyoung Yi (Hyperconnect), Sungjoo Ha (Hyperconnect), Suhyun Lee (Hyperconnect), Youngsoo Lee (Hyperconnect), Gihoon Yeom (Hyperconnect), Buru Chang (Sogang University), and Gihun Lee (Hyperconnect)

Measuring the User Experience on Network Assurance 714
Wutao Wei (Cisco Meraki, USA), Sara Isaac (Cisco Meraki, USA), Alex Ting (Cisco Meraki, USA), Yuanqi Wang (Cisco Meraki, USA), Ke Zhu (Cisco Meraki, USA), and Sophine Zhang (Cisco Meraki, USA)

DLC'24: Deep Learning and Clustering

Cross-Tissue, Cross-Platform Prediction and Validation of Toxicogenomics Profiles with Deep Learning and Clustering 718
Trang Do (Purdue University), Richard Kim (Pomona College), Scott Auerbach (NIEHS), Warren Casey (NIEHS), Andrew Rooney (NIEHS), and Guojing Cong (Oak Ridge National Laboratory)

Dying Clusters Is All You Need - Deep Clustering with an Unknown Number of Clusters 726
Collin Leiber (LMU Munich, Germany), Niklas Strauß (LMU Munich, Germany), Matthias Schubert (LMU Munich, Germany), and Thomas Seidl (LMU Munich, Germany)

Unsupervised Graph-Based Anomaly Detection using Variational Autoencoder 734
Zeinab Sadat Sadrolhefazi (University of Tehran, Iran), Hadi Zare (University of Tehran, Iran), Seyyed Amir Asghari (Kharazmi University, Iran), Parham Moradi (RMIT University, Australia), and Mahdi Jalili (RMIT University, Australia)

Finite Mixture Models for an Underlying Zero-One Inflated Beta Distribution 742
Cédric Noel (University of Lorraine, France) and Jang Schiltz (University of Luxembourg, Luxembourg)

SSTDM-24: International Workshop on Spatial and Spatio-Temporal Data Mining

A New Approach Towards Identification of Individuals Belonging to Different Subpopulations using Machine Learning Models and CLC Data 748
Małgorzata Charytanowicz (Lublin University of Technology, Poland; Polish Academy of Sciences, Poland), Kajetan Perzanowski (John Paul II Catholic University of Lublin, Poland), Maciej Januszczak (Polish Academy of Sciences, Poland), Aleksandra Wołoszyn-Gałęza (Polish Academy of Sciences, Poland), Maria Sobczuk (Warsaw University of Life Sciences, Poland), and Piotr Kulczycki (AGH University of Krakow, Poland; Academy of Sciences, Poland)

Landslide-as-a-Service: Standards-Driven Landslide Mapping Framework using Machine Learning and GeoDataCube 756
Chetan Mahajan (Indian Institute of Technology Bombay, India), Rajat Shinde (The University of Alabama in Huntsville, U.S.A), Gérald Fenoy (GeoLabs SARL, France), and Surya S. Durbha (Indian Institute of Technology Bombay, India)

Fed-LDR: Federated Local Data-Infused Graph Creation with Node-Centric Model Refinement 762
Jiechao Gao (University of Virginia, USA), Yuangang Li (University of Southern California, US), and Syeda Faiza Ahmed (University of Dhaka, Bangladesh)

Neighborhood Spatio-Temporal Self-Attention for Crowd Flow Prediction 770
Souhei Kodama (Tokyo University of Science, Japan) and Shin Ando (Tokyo University of Science, Japan)

Sirius: A Density-Based Colocation Mining Framework for Sets of Continuous Spatial Variables 778
Md Mahin (University of Houston, TX), Aparna Budhavarapu (University of Houston, TX), Karima Elgarroussi (University of Houston, TX), Ernst L Leiss (University of Houston, TX), and Christoph F Eick (University of Houston, TX)

A Deep Learning Based Cloud Imputation Framework 786
Harshitaa Yarramsetti (North Carolina State University, USA), Kritika Javali (North Carolina State University, USA), Mukund Logasundar (North Carolina State University, USA), Richard Li (North Carolina State University, USA), and Ranga Raju Vatsavai (North Carolina State University, USA)

CRLW: Causal Representation Learning Workshop

Deep Entire Space Cross Networks for Estimating Treatment Effect Under Unmeasured Confounding 794
Han Qiao (South China Normal University, China)

M3-InvRL: Multimedia Model Merging Invariant Representation Learning for Recommendation .. 800
Xinghang Hu (Sichuan University, China)

Learning Feasible Causal Algorithmic Recourse without Prior Structure 805
Haotian Wang (National University of Defense Technology), Hao Zou (ZGC Lab), Xueguang Zhou (Naval University of Engineering), Shangwen Wang (National University of Defense Technology), Long Lan (National University of Defense Technology), Wenjing Yang (National University of Defense Technology), and Peng Cui (Tsinghua Unversity)

Projection Pursuit Based Density Ratio Estimation 814
Meilin Wang (Renmin University of China, China), Wei Huang (University of Melbourne, Australia), and Zheng Zhang (Renmin University of China, China)

Learning Invariant Features on Graphs by Disentangling Spurious Features 820
Tianjun Yao (Mohamed bin Zayed University of Artificial Intelligence), Yongqiang Chen (Mohamed bin Zayed University of Artificial Intelligence), Zhiqiang Shen (Mohamed bin Zayed University of Artificial Intelligence), and Kun Zhang (Mohamed bin Zayed University of Artificial Intelligence)

UMVUE-DR: Uniformly Minimum Variance Unbiased Doubly Robust Learning with Reduced Bias and Variance for Debiased Recommendation 828
Chunyuan Zheng (Peking University, China), Taojun Hu (Peking University, China), Chuan Zhou (Peking University, China), and Xiao-Hua Zhou (Peking University, China)

NML: Neverending Machine Learning

Beyond Overall Accuracy: Exploring Per-Class Performance Reveals Robustness and Safety Gaps in Deep Models 833
Kamil Szyc (Wroclaw University of Science and Technology, Poland), Tomasz Walkowiak (Wroclaw University of Science and Technology, Poland), and Henryk Maciejewski (Wroclaw University of Science and Technology, Poland)

Combining Active Learning and Self-Labeling for Deep Learning from Data Streams 842
Jan Wasilewski (Rochester Institute of Technology, USA; OKE Poland), Michał Woźniak (Wrocław University of Science and Technology, Poland), and Bartosz Krawczyk (Rochester Institute of Technology, USA)

DMBIH'24: Data Mining in Biomedical Informatics and Healthcare

Detecting Gene Regulatory Network Rewiring as Changes in Marker Regulatory Associations using Only Single-Cell Gene Expression Data 850
Lisa Chabrier (Centre de Lyon, France; Universite Claude Bernard Lyon 1, France), Anton Crombach (Centre de Lyon, France; Universite Claude Bernard Lyon 1, France), Sergio Peignier (INSA Lyon, France), and Christophe Rigotti (Universite Claude Bernard Lyon 1, France; Centre de Lyon, France)

Summarizing Clinical Notes using LLMs for ICU Bounceback and Length-of-Stay Prediction 859
Akash Choudhuri (University of Iowa), Philip Polgreen (University of Iowa), Alberto Segre (University of Iowa), and Bijaya Adhikari (University of Iowa)

Structural Breakpoint Detection in Noisy Time Series using an Augmented Matrix Profile Index 867
Marco Fuchs (Catholic University of Eichstätt-Ingolstadt, Germany)

Automated Detection and Quantitative Assessment of Extranodal Extension in Lymph Node Breast Tumors using Dual-Tier Deep Learning and Density-Based Contour Algorithms 877
Zhichao Yu (University of Macau, Macau SAR), Simon James Fong (University of Macau, Macau SAR), Liming Zhang (University of Macau, Macau SAR), Tengyue Li (North China University of Technology, China), and Xin Zhang (First People's Hospital of Foshan, China)

Demo Track

VIA AI: Reliable Deep Reinforcement Learning for Traffic Signal Control 887
Matvey Gerasyov (HSE University, Russia), Dmitrii Kiselev (HSE University, Russia), Maxim Beketov (HSE University, Russia), and Ilya Makarov (AIRI, Russia; ISP RAS, Russia)

BiTSA: Leveraging Time Series Foundation Model for Building Energy Analytics 891
Xiachong Lin (University of New South Wales, Australia), Arian Prabowo (University of New South Wales, Australia), Imran Razzak (University of New South Wales, Australia), Hao Xue (University of New South Wales, Australia), Matthew Amos (CSIRO Energy Centre, Australia), Sam Behrens (CSIRO Energy Centre, Australia), and Flora D. Salim (University of New South Wales, Australia)

Go-Kart Racing Simulator for Reinforcement Learning with Augmented Sim2Real Adaptation 895
Ildar Nurgaliev (Innopolis University), Andre Kuzminykh (Dataism Laboratory), Andrey V. Savchenko (Sber AI Lab), and Ilya Makarov (AIRI)

Video-Based Learning of Sign Languages: One Pre-Train to Fit them All 899
Maxim Novopoltsev (Sber AI), Aleksandr Tulenkov (Sber AI), Ruslan Murtazin (Sber AI), Roman Akhidov (Sber AI), Iuliia Zemtsova (Sber AI), Emilia Bojarskaja (ITMO University), Daria Bondarenko (ITMO University), Andrey Savchenko (Sber AI Lab), and Ilya Makarov (ITMO University)

LLM-KT: A Versatile Framework for Knowledge Transfer from Large Language Models to Collaborative Filtering 903
Nikita Severin (HSE University), Aleksei Ziablitsev (MIPT), Yulia Savelyeva (MIPT), Valeriy Tashchilin (Ural Federal University), Ivan Bulychev (HSE University), Mikhail Yushkov (HSE University), Artem Kushneruk (HSE University), Amaliya Zaryvnykh (HSE University), Dmitrii Kiselev (HSE University), Andrey Savchenko (HSE University), and Ilya Makarov (AIRI, ISP RAS)

ByteScience: Bridging Unstructured Scientific Literature and Structured Data with Auto Fine-tuned Large Language Model in Token Granularity 907
Tong Xie (GreenDynamics, Australia; University of New South Wales, Australia), Hanzhi Zhang (University of North Texas, United States), Shaozhou Wang (GreenDynamics, Australia), Yuwei Wan (GreenDynamics, Australia), Imran Razzak (University of New South Wales, Australia), Chunyu Kit (City University of Hong Kong, China), Wenjie Zhang (University of New South Wales, Australia), and Bram Hoex (University of New South Wales, Australia)

Author Index