Rozaida Ghazali · Nazri Mohd Nawi · Mustafa Mat Deris · Jemal H. Abawajy · Nureize Arbaiy *Editors*

Recent Advances on Soft Computing and Data Mining

Proceedings of the Sixth International Conference on Soft Computing and Data Mining (SCDM 2024), August 21–22, 2024



Lecture Notes in Networks and Systems

Series Editor

Janusz Kacprzyk , Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland

Advisory Editors

Fernando Gomide, Department of Computer Engineering and Automation—DCA, School of Electrical and Computer Engineering—FEEC, University of Campinas—UNICAMP, São Paulo, Brazil

Okyay Kaynak, Department of Electrical and Electronic Engineering, Bogazici University, Istanbul, Türkiye

Derong Liu, Department of Electrical and Computer Engineering, University of Illinois at Chicago, Chicago, USA

Institute of Automation, Chinese Academy of Sciences, Beijing, China Witold Pedrycz, Department of Electrical and Computer Engineering, University of Alberta, Canada

Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland Marios M. Polycarpou, Department of Electrical and Computer Engineering, KIOS Research Center for Intelligent Systems and Networks, University of Cyprus, Nicosia, Cyprus

Imre J. Rudas, Óbuda University, Budapest, Hungary Jun Wang, Department of Computer Science, City University of Hong Kong, Kowloon, Hong Kong The series "Lecture Notes in Networks and Systems" publishes the latest developments in Networks and Systems—quickly, informally and with high quality. Original research reported in proceedings and post-proceedings represents the core of LNNS.

Volumes published in LNNS embrace all aspects and subfields of, as well as new challenges in, Networks and Systems.

The series contains proceedings and edited volumes in systems and networks, spanning the areas of Cyber-Physical Systems, Autonomous Systems, Sensor Networks, Control Systems, Energy Systems, Automotive Systems, Biological Systems, Vehicular Networking and Connected Vehicles, Aerospace Systems, Automation, Manufacturing, Smart Grids, Nonlinear Systems, Power Systems, Robotics, Social Systems, Economic Systems and other. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution and exposure which enable both a wide and rapid dissemination of research output.

The series covers the theory, applications, and perspectives on the state of the art and future developments relevant to systems and networks, decision making, control, complex processes and related areas, as embedded in the fields of interdisciplinary and applied sciences, engineering, computer science, physics, economics, social, and life sciences, as well as the paradigms and methodologies behind them.

Indexed by SCOPUS, EI Compendex, INSPEC, WTI Frankfurt eG, zbMATH, SCImago.

All books published in the series are submitted for consideration in Web of Science.

For proposals from Asia please contact Aninda Bose (aninda.bose@springer.com).

Rozaida Ghazali · Nazri Mohd Nawi · Mustafa Mat Deris · Jemal H. Abawajy · Nureize Arbaiy
Editors

Recent Advances on Soft Computing and Data Mining

Proceedings of the Sixth International Conference on Soft Computing and Data Mining (SCDM 2024), August 21–22, 2024





Editors
Rozaida Ghazali
Faculty of Computer Science and Information
Technology
Universiti Tun Hussein Onn Malaysia
Parit Raja, Malaysia

Mustafa Mat Deris Faculty of Business and Information Technology Universiti Muhammadiah Malaysia Padang Besar, Perlis, Malaysia

Nureize Arbaiy
Faculty of Computer Science and Information
Technology
Universiti Tun Hussein Onn Malaysia
Parit Raia, Johor, Malaysia

Nazri Mohd Nawi Faculty of Computer Science and Information Technology Universiti Tun Hussein Onn Malaysia Parit Raja, Malaysia

Jemal H. Abawajy School of Information Technology Deakin University Wandana Heights, VIC, Australia

ISSN 2367-3370 ISSN 2367-3389 (electronic) Lecture Notes in Networks and Systems ISBN 978-3-031-66964-4 ISBN 978-3-031-66965-1 (eBook) https://doi.org/10.1007/978-3-031-66965-1

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2024

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

If disposing of this product, please recycle the paper.

Preface

Advancements in data storage and accessibility have fueled the growth of data science, which employs various methods to extract insights and patterns from data. Despite the demand for skilled data scientists, extracting actionable insights remains challenging, particularly with complex data systems. To address this, data mining has emerged as a crucial approach, offering potential for discovering patterns across diverse data types. By leveraging data and soft computing techniques, researchers can explore extensive databases to uncover hidden patterns. Ongoing research focuses on developing advanced statistical interpretations and innovative technologies. Soft computing techniques address imprecision and uncertainty, enhancing tractability and robustness. These techniques, whether used individually or combined, are emerging as robust options for various tasks in fields such as business and marketing, healthcare, finance, e-commerce, manufacturing, telecommunications, transportation, environmental science, agriculture, education, and more. They aim to transform data into innovative solutions that offer new value propositions for customers.

Following the successful organization of five previous SCDM conferences from 2014 to 2022, we are pleased to continue this journey of achievements with our sixth international conference, SCDM 2024. This year's conference, held in a virtual format on August 21–22, 2024, facilitated global participation, providing live interactive networking opportunities and content access to attendees worldwide. We received 75 paper submissions from 15 countries, each of which underwent rigorous screening and peerreview processes. Ultimately, 42 papers of the highest quality and merit were selected for oral presentation and publication in this volume proceeding, representing an acceptance rate of 56%.

We would like to express our sincere appreciation to the conference organizer, Faculty of Computer Science & Information Technology, UTHM, and the Soft Computing & Data Mining research group, as well as to the Steering Committee, Conference Chair, Program Committee Chair, Organizing Chairs, and all Program and Reviewer Committee members of SCDM 2024. Their invaluable contributions to the review process have ensured the highest quality of selected papers for the conference.

We also extend our appreciation to our esteemed keynote speakers, Associate Professor Dr. Harish Garg from the Thapar Institute of Engineering & Technology, Punjab, India, and Mr. Azhar Kassim Mustapha from Nervesis, Malaysia. Special thanks are due to Dr. Thomas Ditzinger for facilitating the publication of the proceeding in Lecture Notes in Networks and Systems, Springer. We acknowledge the Organizing Committee members for their significant contributions, particularly those in pivotal roles.

Finally, we want to extend our heartfelt appreciation to all authors for their valuable contributions and to all participants for their enthusiastic engagement. We are truly

vi Preface

grateful for your dedication and commitment, which have greatly contributed to the success of this conference.

Rozaida Ghazali Nazri Mohd Nawi Mustafa Mat Deris Jemal H. Abawajy Nureize Arbaiy

Conference Organization

Patron

Ruzairi bin Abdul Rahim Vice Chancellor, Universiti Tun Hussein Onn

Malaysia, Malaysia

Advisory Committee

Ajith Abraham Machine Intelligence Research Labs, USA

Hamido Fujita Iwate Prefectural University, Japan

Junzo Watada Waseda University, Japan

Nikola Kasabov KEDRI, Auckland University of Technology,

New Zealand

Rajkumar Buyya University of Melbourne, Australia Witold Pedrycz University of Alberta, Canada

Steering Committee

Mustafa Mat Deris Universiti Tun Hussein Onn Malaysia

Jemal H Abawajy Deakin University, Australia

Nazri Mohd Nawi Universiti Tun Hussein Onn Malaysia Rozaida Ghazali Universiti Tun Hussein Onn Malaysia Hairulnizam Mahdin Universiti Tun Hussein Onn Malaysia

Conference Chair

Nazri Mohd Nawi Universiti Tun Hussein Onn Malaysia

Conference Co-chair

Norhalina Senan Universiti Tun Hussein Onn Malaysia

Proceeding Chair

Rozaida Ghazali Universiti Tun Hussein Onn Malaysia

Program Committee Chair

Nureize Arbaiy Universiti Tun Hussein Onn Malaysia

Website, Promotion, and Publicity Chair

Mohd Norasri Ismail Universiti Tun Hussein Onn Malaysia

Organizing Committee

Nurezayana Zainal Universiti Tun Hussein Onn Malaysia Zuraida Bosri Universiti Tun Hussein Onn Malaysia Suzivanti Marjudi Universiti Tun Hussein Onn Malaysia Rabatul Aduni Sulaiman Universiti Tun Hussein Onn Malaysia Zam Zarina Zainal Abidin Universiti Tun Hussein Onn Malaysia Mohd Zanes Sahid Universiti Tun Hussein Onn Malaysia Shahreen Kasim Universiti Tun Hussein Onn Malaysia Norashid Hassan Universiti Tun Hussein Onn Malaysia Mohd Helmy Abd Wahab Universiti Tun Hussein Onn Malaysia

Program Committee

Adila Firdaus Arbain Universiti Teknologi Malaysia

Adnan Abid University of Management and Technology,

Pakistan

Afnizanfaizal Abdullah Universiti Teknologi Malaysia
Agouti Tarik Cadi Ayyad University, Morocco
Ahmed A. Elngar Beni-Suef University, Egypt

Aida Mustapha Universiti Tun Hussein Onn Malaysia

Alessandro D'Amelio University of Milan, Italy Alessandro Giuliani University of Cagliari, Italy Ali Ahmadian Universiti Putra Malaysia

Ali Mohammadi Isfahan University of Technology, Iran

Amelia Zafra Gomez University of Cordoba, Spain

Ammar Awad Mutlag Ayodele Lasisi

Bazeer Ahamed Bagrudeen

Carlos Pereora Chuah Chai Wen

Cik Feresa Mohd Foozy

Daoudi Najima El Habib Benlahmar Elena Benderskaya Ender Özcan Ezak Ahmad Fairouz Zendaoui

Fatima Zahra Fagroud

Hanaa Hachimi Hardeo Kumar Thakur

Hazlina Hamdan

Isredza Rahmi Ab Hamid

Jawad Ali

José Ramón Villar

Jyotir Moy Chatterjee

Kashif Hussain

Katsuhiro Honda Khalil Ghathwan

Mario José Diván

Maslina Zolkepli Md. Raihan Uddin

Mohamad Aizi Salamat Mohammad Zubair Rehman Mohd Amin Mohd Yunus

Mohd. Farhan Md. Fudzee Mohd. Najib Mohd. Salleh Mohit Jain

Mouad Banane Muhammad Adnan Khan

Muhammad Faheem Mushtaq Nadjet Kamel Universiti Teknikal Malaysia Melaka

Augustine University, Nigeria

University of Technology and Applied Sciences

Al Musannah, Oman

ISEC, Portugal

Universiti Tun Hussein Onn Malaysia Universiti Tun Hussein Onn Malaysia

ESI, Morocco

University of Hassan II, Casablanca, Morocco St.Petersburg State Polytechnic University, Russia

University of Nottingham, UK

Universiti Tun Hussein Onn Malaysia Institut National de la Poste et des TIC,

Eucalyptus, Alger, Algérie

Hassan II University, Casablanca, Morocco

Ibn Tofail University, Morocco

Manav Rachna University, Faridabad, India

Universiti Putra Malaysia

Universiti Tun Hussein Onn Malaysia COMSATS University ISB Lahore Campus,

Pakistan

University of Oviedo, Spain

 $Lord\ Buddha\ Education\ Foundation,\ Kathmandu,$

Nepal

University of Electronic Science and Technology

of China

Osaka Prefecture University, Japan University of Technology, Iraq

National University of La Pampa (UNLPam),

Argentina

Universiti Putra Malaysia

Daffodil International University, Bangladesh

Universiti Tun Hussein Onn Malaysia Agriculture University Peshawar, Pakistan Universiti Tun Hussein Onn Malaysia Universiti Tun Hussein Onn Malaysia Universiti Tun Hussein Onn Malaysia NSIT (University of Delhi), India

University Hassan 2, Morocco

Riphah International University, Pakistan

The Islamia University of Bahawalpur, Pakistan

University Ferhat Abbas Setif 1, Algeria

Noor Azah Samsudin Noor Zuraidin Mohd Safar

Nor Azura Husin Nordiana Rahim Norfaradilla Wahid Noryusliza Abdullah Nureize Arbaiy Nurezayana Zainal

Okan Duru

Oumaima Hourrane Palaniappan Shamala

Pei-Chun Lin

Pramit Brata Chanda

Rabei Al-Jawary

Rachid Saadane Radiah Mohamad Radzi Ambar Rahayu Hamid Rahmat Hidayat Rajdeep Chowdhury

Riyaz Ahamed

Sadiq Abdelalim Salama A. Mostafa

Sanaa El Filali Saoud Sahar

Sasalak Tongkaw Sathya Bursic Shahreen Kasim

Shamsollah Ghanbari

Sofia Najwa Ramli Suhaimi Abd Ishak Szymon Lukasik Tadashi Nomoto Uma N. Dulhare

Venkatesh Gauri Shankar

Vitalii Nitsenko

Vittorio Cuculo Waddah Waheeb Universiti Tun Hussein Onn Malaysia Universiti Tun Hussein Onn Malaysia

Universiti Putra Malaysia

Universiti Tun Hussein Onn Malaysia Universiti Tun Hussien Onn Malaysia Universiti Tun Hussein Onn Malaysia Universiti Tun Hussein Onn Malaysia Universiti Tun Hussein Onn Malaysia

Nanyang Technological University, Singapore University of Hassan II, Casablanca, Morocco

Universiti Teknologi MARA Feng Chia University, Taiwan

Kalyani Government Engineering College, Kalyani

American University of Ras Al Khaimah, UAE

LETI, EHTP, Morocco Universiti Tun Hussein Onn Malaysia Universiti Tun Hussein Onn Malaysia Universiti Tun Hussein Onn Malaysia Politeknik Negeri Padang, Indonesia

West Bengal, India

International University of Malaya-Wales

Ibn Tofail University, Morocco

Universiti Tun Hussein Onn Malaysia Hassan II University, Casablanca, Morocco National Business School, Ibn Zohr University, Agadir, Morocco

Songkhla Rajabhat University, Thailand

University of Milan, Italy

Universiti Tun Hussein Onn Malaysia

Islamic Azad University, Ashtian Branch, Iran

Universiti Tun Hussein Onn Malaysia Universiti Tun Hussein Onn Malaysia Cracow University of Technology, Poland National Institute of Japanese Literature, Japan

Muffakham Jah College of Engineering & Technology (MJCET), Hyderabad, India

Manipal University Jaipur, India

Odessa I.I.Mechnikov National University,

Odessa, Ukraine

University of Milan, Italy University of Agder, Norway Wamiq Raza University of Trento, Italy
Waseem Mohssen Alhasan Al-Sham Private University, Syria
Yana Mazwin Mohmad Hassim Universiti Tun Hussein Onn Malaysia
Youness Tabii Labo ADMIR–ENSIAS UM5 Rabat, Morocco
Zubaile Abdullah Universiti Tun Hussein Onn Malaysia

Organizer

Faculty of Computer Science & Information Technology, Universiti Tun Hussein Onn Malaysia



Contents

Prediction of OPEC Carbon Dioxide Emissions Using K-Means Clustering and Ensemble Algorithm Ayodele Lasisi, Nur Ariffin Mohd Zin, Rozaida Ghazali, and Modupe Agagu	1
Detection of Phishing Websites from URLs Using Hybrid Ensemble-Based Machine Learning Technique	11
Minimal Data for Maximum Impact: An Indonesian Part-of-Speech Tagging Case Study Chi Log Chua, Tong Ming Lim, and Kwee Teck See	23
Alleviating Sparsity to Enhance Group Recommendation with Cross-Linked Domain Model Yui Chee Xuan, Rosmamalmi Mat Nawi, Nurul Aida Osman, and Nur Ziadah Harun	33
Evaluating Deep Transfer Learning Models for Detecting Various Face Mask Wearings Pei-Jin Goh, Meei-Hao Hoo, and Kok-Chin Khor	43
Classification of Stunting Events: Case Study in West Java, Indonesia	53
The Effects of Data Reduction Using Rough Set Theory on Logistic Regression Model Izzati Rahmi, Riswan Efendi, Nor Azah Samat, Hazmira Yozza, and Muhammad Wahyudi	64
Robust Heart Disease Prognosis: Integrating Extended Isolation Forest Outlier Detection with Advanced Prediction Models Irfan Javid, Norlida Hassan, Rozaida Ghazali, Yana Mazwin Mohmad Hassim, Tuba Batool, Noor Aida Husaini, and Syed Irteza Hussain Jafri	74

Classification	84
Improved Rough-Multiple Regression for Unemployment Rate Model in Indonesia	94
Utilizing Machine Learning for Gene Expression Data: Incorporating Gene Sequencing, K-Mer Counting and Asymmetric N-Grams Features Chai-Wen Chuah, WanXian He, De-Shuang Huang, and Janaka Alawatugoda	105
Text Sentiment Analysis on VIX's Impact on Market Sentiment Dynamics Zhuqin Liang, Mohd Tahir Ismail, and Huimin Qu	115
Multilevel Monte Carlo Simulation Model for Air Pollution Index Prediction of a Smart Network Mustafa Hamid Hassan, Salama A. Mostafa, Rozaida Ghazali, Mohd Zainuri Saringat, Noor Aida Husaini, Aida Mustapha, Mohammed Ahmed Jubair, and Hussein Muhi Hariz	125
An In-Depth Strategy using Deep Generative Adversarial Networks for Addressing the Cold Start in Movie Recommendation Systems	136
Predicting Undergraduate Academic Success with Machine Learning Approaches	144
Comparative Assessment of Facial Expression Recognition Models for Unraveling Emotional Signals with Convolutional Neural Networks Afia Zafar, Nazri Mohd Nawi, Noushin Saba, Kainat Zafar, Mohsin Suleman, and Shahneer Zafar	154
Evaluating Path-Finding Algorithms for Real-Time Route Recommendation System Built using FreeRTOS	165
Machine Learning-Based Phishing Website Detection: A Comparative Analysis and Web Application Development Jia Xin Yau and Kai Lin Chia	175

Comparative Performance of Multi-level Pre-trained Embeddings on CNN, LSTM and CNN-LSTM for Hate Speech and Offensive Language Detection Noor Azeera Abdul Aziz, Anazida Zainal, Bander Ali Saleh Al-Rimy, and Fuad Abdulgaleel Abdoh Ghaleb	186
Improved Classifier Chain Method Based on Particle Swarm Optimization and Genetic Algorithm for Multilabel Classification Problem	196
Sentiment Analysis on Umrah Packages Review in Malaysia	207
Opinion Mining System for Influence Detection Using Machine Learning to Secure Business Reputation	219
A Presentation Mining Framework: From Text Mining to to Mind Mapping Vinothini Kasinathan and Aida Mustapha	233
Enhancing Network Intrusion Detection Systems Through Dimensionality Reduction Mosleh M. Abualhaj, Sumaya N. Al-Khatib, Ali Al-Allawee, Alhamza Munther, and Mohammed Anbar	244
Performance Evaluation of Whale and Harris Hawks Optimization Algorithms with Intrusion Prevention Systems Mosleh M. Abualhaj, Ahmad Adel Abu-Shareha, Ali Al-Allawee, Alhamza Munther, and Mohammed Anbar	254
Domestic Solid Waste Prediction with an Enhanced LSTM with SigmoReLU and RAdam Optimizer Abdulrahman Sharaf Mohammed Fadhel, Rozaida Ghazali, Mohd Razali Md Tomari, Yana Mazwin Mohmad Hassim, Abdullahi Abdi Abubakar Hassan, and Lokman Hakim Ismail	266
Sounds Prediction Instruments Based Using K-Means and Bat Algorithm Rozlini Mohamed, Noor Azah Samsuddin, and Munirah Mohd Yusof	276
A Comparative Study on Ant-Colony Algorithm and Genetic Algorithm for Mobile Robot Planning	286

Enhanced Air Quality Index Prediction Using a Hybrid Convolutional Network Pei-Chun Lin, Nureize Arbaiy, Chen-Yu Yu, and Mohd Zaki Mohd Salikon	296
Filter Method Feature Selection Techniques for Solid Waste Prediction Based on GRU Deep Learning Model Tuba Batool, Siti Hajar Arbain, Rozaida Ghazali, Lokman Hakim Ismail, and Irfan Javid	307
Spiking Neural Network for Microseismic Events Detection Using Distributed Acoustic Sensing Data Mohd Safuwan Bin Shahabudin, Nor Farisha Binti Muhamad Krishnan, and Farahida Hanim Binti Mausor	317
Battery Electric Vehicle Charging Load Forecasting Using LSTM on STL Trend, Seasonality, and Residual Decomposition	327
Convolutional Neural Network Using Regularized Conditional Entropy Loss (CNNRCoE) for MNIST Handwritten Digits Classification Ashikin Ali, Norhalina Senan, and Norhanifah Murli	337
Optimizing Team Formation for Welfare Activities: A Study Using Four Metaheuristic Optimization Algorithms	349
Detection of Paddy Plant Diseases Using Google Teachable Machine	360
Comparative Analysis of ResNet Models for Skin Cancer Diagnosis: Performance Evaluation and Insights Razan Alharith, Ashraf Osman Ibrahim, Noorhaniza Wahid, Rozaida Ghazali, and Abubakar Elsafi	370
The Predictive Modelling of Student Academic Performance Using Machine Learning Approaches Nurul Habibah Abdul Rahman, Sahimel Azwal Sulaiman, and Nor Azuana Ramli	379

Contents	xvii
Predictive Modeling of Gold Prices: Integrating Technical Indicators for Enhanced Accuracy Noor Aida Husaini, Yee Jing Gan, Rozaida Ghazali, Yana Mazwin Mohmad Hassim, Jie Shen Yeap, and Jerome Subash Joseph	390
Portfolio Optimization with Percentage Error-Based Fuzzy Random Data for Industrial Production	400
The Football Matches Outcome Prediction for English Premier League (EPL): A Comparative Analysis of Multi-class Models	411
An Automated Quasi-Identification (QID) for Re-identification	421
Author Index	433