

Proceedings

**21st IEEE International Conference
on Data Mining Workshops**

ICDMW 2021

Proceedings

21st IEEE International Conference on Data Mining Workshops

7–10 December 2021
Virtual Conference

Editors

Bing Xue, Mykola Pechenizkiy, and Yun Sing Koh



Los Alamitos, California
Washington • Tokyo



All rights reserved.

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

Other copying, reprint, or republication requests should be addressed to: IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, P.O. Box 133, Piscataway, NJ 08855-1331.

The papers in this book comprise the proceedings of the meeting mentioned on the cover and title page. They reflect the authors' opinions and, in the interests of timely dissemination, are published as presented and without change. Their inclusion in this publication does not necessarily constitute endorsement by the editors, the IEEE Computer Society, or the Institute of Electrical and Electronics Engineers, Inc.

BMS Part Number CFP2156B-ART
ISBN 978-1-6654-2427-1

Additional copies may be ordered from:

IEEE Computer Society
Customer Service Center
10662 Los Vaqueros Circle
P.O. Box 3014
Los Alamitos, CA 90720-1314
Tel: + 1 800 272 6657
Fax: + 1 714 821 4641
<http://computer.org/cspress>
csbooks@computer.org

IEEE Service Center
445 Hoes Lane
P.O. Box 1331
Piscataway, NJ 08855-1331
Tel: + 1 732 981 0060
Fax: + 1 732 981 9667
[http://shop.ieee.org/store/
customer-service@ieee.org](http://shop.ieee.org/store/customer-service@ieee.org)

IEEE Computer Society
Asia/Pacific Office
Watanabe Bldg., 1-4-2
Minami-Aoyama
Minato-ku, Tokyo 107-0062
JAPAN
Tel: + 81 3 3408 3118
Fax: + 81 3 3408 3553
tokyo.ofc@computer.org

Individual paper REPRINTS may be ordered at: <reprints@computer.org>

Editorial production by Lisa O'Conner
Cover art production by Hector Torres



*IEEE Computer Society
Conference Publishing Services (CPS)*

<http://www.computer.org/cps>

2021 International Conference on Data Mining Workshops (ICDMW) **ICDMW 2021**

Table of Contents

| | |
|---|-------|
| Message from the ICDM 2021 General Chairs | xxiii |
| Message from the ICDM 2021 Program Chairs | xxv |
| Message from the Workshops Chairs | xxvii |
| Organizing Committee | xxix |

NeuRec: Advanced Neural Algorithms and Theories for Recommender Systems

| | |
|--|----|
| Incorporating Adjacent User Modeling into Session-Based Recommendation with Graph Neural Networks | 1 |
| <i>Minghao Chen (Zhejiang University, China) and Jiale Zheng (Zhejiang University, China)</i> | |
| Sequential Item Recommendation in the MOBA Game Dota 2 | 10 |
| <i>Alexander Dallmann (University of Würzburg, Germany), Johannes Kohlmann (University of Würzburg, Germany), Daniel Zoller (University of Würzburg, Germany), and Andreas Hotho (University of Würzburg, Germany)</i> | |
| Dynamic Sequential Recommendation: Decoupling User Intent from Temporal Context | 18 |
| <i>Wei Jiang (Alibaba Group, China), Fangquan Lin (Alibaba Group, China), Jihai Zhang (Alibaba Group, China), Cheng Yang (Alibaba Group, China), Hanwei Zhang (Alibaba Group, China), and Ziqiang Cui (Alibaba Group, China)</i> | |
| A Probabilistic Perspective on Nearest Neighbor for Implicit Recommendation | 27 |
| <i>Domokos Miklós Kelen (Institute for Computer Science and Control (SZTAKI), Hungary) and Andras A. Benczúr (Institute for Computer Science and Control (SZTAKI), Hungary)</i> | |
| DynaPosGNN: Dynamic-Positional GNN for Next POI Recommendation | 36 |
| <i>Junbeom Kim (Samsung Research, Samsung Electronics. Co., Ltd, Korea), Sihyun Jeong (Samsung Research, Samsung Electronics. Co., Ltd, Korea), Goeon Park (Samsung Research, Samsung Electronics. Co., Ltd, Korea), Kihoon Cha (Samsung Research, Samsung Electronics. Co., Ltd, Korea), Ilhyun Suh (Samsung Research, Samsung Electronics. Co., Ltd, Korea), and Byungkook Oh (Samsung Research, Samsung Electronics. Co., Ltd, Korea)</i> | |

| | |
|---|-----|
| CoBERT: Scientific Collaboration Prediction via Sequential Recommendation | 45 |
| <i>Tobias Koopmann (University of Würzburg, Germany), Konstantin Kobs (University of Würzburg, Germany), Konstantin Herud (University of Würzburg, Germany), and Andreas Hotho (University of Würzburg, Germany)</i> | |
| SynEvaRec: A Framework for Evaluating Recommender Systems on Synthetic Data Classes | 55 |
| <i>Vladimir Provalov (National Center for Cognitive Research, ITMO University, Russia), Elizaveta Stavinova (National Center for Cognitive Research, ITMO University, Russia), and Petr Chunaev (National Center for Cognitive Research, ITMO University, Russia)</i> | |
| Balanced News Neural Network for a News Recommender System | 65 |
| <i>Shaina Raza (Ryerson University, Canada), Syed Raza Bashir (Ryerson University, Canada), Dora D. Liu (DeepBlue Academy of Sciences, China), and Usman Naseem (The University of Sydney, Australia)</i> | |
| Embedding Normalization: Significance Preserving Feature Normalization for Click-Through Rate Prediction | 75 |
| <i>Joonyoung Yi (Hyperconnect, South Korea), Beomsu Kim (Hyperconnect, South Korea), and Buru Chang (Hyperconnect, South Korea)</i> | |
| IBFM: An Instance-Weight Balanced Factorization Machine for Sparse Prediction | 85 |
| <i>Wenxiong Wu (FuZhou University, China), Yongli Cheng (FuZhou University, China; HuaZhong University of Science and Technology, China), Hong Jiang (Department of Computer Science & Engineering, University of Texas at Arlington, USA), Fang Wang (Huazhong University of Science and Technology, China), Xianghao Xu (Nanjing University of Science and Technology, China), Dan He (Information Engineering School of Nanchang Hangkong University, China), and Jing Yu (FuZhou University, China)</i> | |
| Challenging the Long Tail Recommendation on Heterogeneous Information Network | 94 |
| <i>Chuanyan Zhang (Shandong University, China) and Xiaoguang Hong (Shandong University, China)</i> | |
| ContentHE: Content-Enhanced Network Embedding for Hashtag Representation | 102 |
| <i>Shaoliang Zhang (Beihang Hangzhou Innovation Institute Yuhang, China), Xuefeng Liu (Beihang university, China), Jianwei Niu (Beihang university, China), and Huiyong Li (Beihang university, China)</i> | |

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

| | |
|--|-----|
| Sentiment Analysis using Part-of-Speech-Based Feature Extraction and Game-Theoretic Rough Sets | 110 |
| <i>Yixing Chen (University of Regina, Canada) and JingTao Yao (University of Regina, Canada)</i> | |

| | |
|--|-----|
| A Real-Time Platform for Contextualized Conspiracy Theory Analysis | 118 |
| <i>David Chong (University of California, Los Angeles), Erl Lee (University of California, Los Angeles), Matthew Fan (University of California, Los Angeles), Pavan Holur (University of California, Los Angeles), Shadi Shahsavari (University of California, Los Angeles), Timothy Tangherlini (University of California, Berkeley), and Vwani Roychowdhury (University of California, Los Angeles)</i> | |
| Automated Pipeline for Sentiment Analysis of Political Tweets | 128 |
| <i>Atrik Das (Nanyang Technological University, Singapore), Kushal Sai Gunturi (Nanyang Technological University, Singapore), Aditya Chandrasekhar (Nanyang Technological University, Singapore), Abhinandan Padhi (Nanyang Technological University, Singapore), and Qian Liu (Nanyang Technological University, Singapore)</i> | |
| Deep Neural Language-Agnostic Multi-task Text Classifier | 136 |
| <i>Karol Gawron (Wroclaw University of Science and Technology, Poland), Michal Pogoda (Wroclaw University of Science and Technology, Poland), Norbert Ropiak (Wroclaw University of Science and Technology, Poland), Michal Swedrowski (Wroclaw University of Science and Technology, Poland), and Jan Kocon (Wroclaw University of Science and Technology, Poland)</i> | |
| Sentiment Analysis Framework using Data Driven Approach | 143 |
| <i>Md Jahedul Islam (Brac University, Bangladesh), Md Shubiour Shuvo (Brac University, Bangladesh), Tonmoy Sarker (Brac University, Bangladesh), Mohammad Zavid Parvez (Brac University, Bangladesh), and Md Anisur Rahman (Charles Sturt University, Australia)</i> | |
| A Multitask Learning Framework for Multimodal Sentiment Analysis | 151 |
| <i>Dazhi Jiang (Shantou University, China), Runguo Wei (Shantou University, China), Hao Liu (Shantou University, China), Jintao Wen (Shantou University, China), Geng Tu (Shantou University, China), Lin Zheng (Shantou University, China), and Erik Cambria (Nanyang Technological University, Singapore)</i> | |
| Interpretable Representation Learning for Personality Detection | 158 |
| <i>Amirmohammad Kazemeini (University of Western Ontario, Canada), Sudipta Singha Roy (University of Western Ontario, Canada), Robert E. Mercer (University of Western Ontario, Canada), and Erik Cambria (Nanyang Technological University, Singapore)</i> | |
| AspectEmo: Multi-Domain Corpus of Consumer Reviews for Aspect-Based Sentiment Analysis | 166 |
| <i>Jan Kocon (Wroclaw University of Science and Technology, Poland), Jarema Radom (Wroclaw University of Science and Technology, Poland), Ewa Kaczmarz-Wawryk (Institute of Polish Studies, University of Wroclaw, Poland), Kamil Wabnic (Institute of Polish Studies, University of Wroclaw, Poland), Ada Zajackowska (Wroclaw University of Science and Technology, Poland), and Monika Zasko-Zielinska (Institute of Polish Studies, University of Wroclaw, Poland)</i> | |
| DUSE: A New Benchmark Dataset for Drug User Sentiment Extraction | 174 |
| <i>Ashok J. Kumar (Information Science and Technology, Anna University, India), Erik Cambria (Nanyang Technological University, Singapore), and Tina Esther Trueman (Information Science and Technology, Anna University, India)</i> | |

| | |
|--|-----|
| Enhancing Negation Scope Detection using Multitask Learning | 179 |
| <i>Harsh Patel (Indian Institute of Technology Gandhinagar, India), Xulang Zhang (Nanyang Technological University Singapore, India), and Qian Liu (Nanyang Technological University Singapore, India)</i> | |
| Contextualized Embedding Based Approaches for Social Media-Specific Sentiment Analysis | 186 |
| <i>Harsh Sakhrani (Pune Institute of Computer Technology, India), Saloni Parekh (Pune Institute of Computer Technology, India), and Pratik Ratadiya (vCreaTek Consulting Services Pvt. Ltd., India)</i> | |
| Modelling Context with Graph Convolutional Networks for Aspect-Based Sentiment Analysis | 194 |
| <i>Maoyuan Zhang (Central China Normal University, PR China), Jieqiong Zhang (Central China Normal University, PR China), and Lisha Liu (Central China Normal University, PR China)</i> | |

DMS: Data Mining for Service

| | |
|---|-----|
| Prediction Diagnostics – Addressing Data Veracity in Predicting Batch Processes | 201 |
| <i>Parag Agrawal (Digitate, Tata Consultancy Services, India), Priyadarshi Priyadarshi (Digitate, Tata Consultancy Services, India), Vikrant Shimpi (Digitate, Tata Consultancy Services, India), Neha Behl (Digitate, Tata Consultancy Services, India), Deepa Vaidyanathan (Digitate, Tata Consultancy Services, India), and Maitreya Natu (Digitate, Tata Consultancy Services, India)</i> | |
| Application of Machine Learning for Growth Environment Prediction in Agriculture | 208 |
| <i>Momoko Fujita (Nagoya University, Japan), Akari Yamada (Nagoya University, Japan), Mao Susuki (Nagoya University, Japan), Hiroya Makino (Nagoya University, Japan), and Eisuke Kita (Nagoya University, Japan)</i> | |
| A Data-Driven Approach to Predict Hourly Bill Rates for US Contingent Workers | 214 |
| <i>Giuliano Giuliani (Randstad Sourceright, Netherlands), Fabrizio Giuliani (Randstad Sourceright, Netherlands), and Summer Husband (Randstad Sourceright, Netherlands)</i> | |
| A Time-Series Analysis of How Google Trends Searches Affect Cryptocurrency Prices for Decentralized Finance and Non-Fungible Tokens | 222 |
| <i>Yuta Kaneko (Doshisha Women’s College of Liberal Arts, Japan)</i> | |
| Intent-Based Product Collections for E-Commerce using Pretrained Language Models | 228 |
| <i>Hiun Kim (Naver Clova), Jisu Jeong (Naver Clova, Naver AI Lab), Kyung-Min Kim (Naver Clova; Naver AI Lab), Dongjun Lee (LBox Co., Ltd.), Hyun Dong Lee (Naver Clova), Dongpil Seo (Naver Clova), Jeeseung Han (Naver Clova), Dong Wook Park (Naver Clova), Ji Ae Heo (Naver Clova), and Rak Yeong Kim (Naver Clova)</i> | |
| Stochastic Schemata Exploiter-Based AutoML | 238 |
| <i>Hiroya Makino (Nagoya University, Japan) and Eisuke Kita (Nagoya University, Japan)</i> | |
| Confident Collaborative Metric Learning | 246 |
| <i>Ryo Matsui (Tokyo Institute of Technology, Japan), Taketo Naito (SMN, Corp., Japan), Suguru Yaginuma (MC Digital, inc., Japan), and Kazuhide Nakata (Tokyo Institute of Technology, Japan)</i> | |

| | |
|---|-----|
| Legitimacy: An Ensemble Learning Model for Credibility Based Fake News Detection | 254 |
| <i>Amit Neil Ramkissoon (The University of the West Indies at St Augustine, Trinidad & Tobago) and Wayne Goodridge (The University of the West Indies at St Augustine, Trinidad & Tobago)</i> | |
| Application of Fractal Analysis for Customer Classification Based on Path Data | 262 |
| <i>Fengmei Sun (Collaborative Innovation Center for Transport Studies, Dalian Maritime University, China), Yi Zuo (Collaborative Innovation Center of Maritime Big Data & Shipping Artificial General Intelligence, Dalian Maritime University, China), Licheng Zhao (Collaborative Innovation Center for Transport Studies, Dalian Maritime University, China), and Yuta Kaneko (Department of Social System Studies, Faculty of Contemporary Social Studies, Doshisha Women's College of Liberal Arts, Japan)</i> | |
| Data-Driven Divide-and-Conquer for Estimating Build Times of 3D Objects | 268 |
| <i>Mahdi Tabassian (EluciDATA Lab of Sirris, Belgium), Robbert Verbeke (EluciDATA Lab of Sirris, Belgium), Tom Tourwé (EluciDATA Lab of Sirris, Belgium), and Elena Tsiporkova (EluciDATA Lab of Sirris, Belgium)</i> | |
| The Vehicle Routing Problem with Time Windows and Time Costs | 278 |
| <i>Zhao Wang (NTT Network Technologies Laboratories, NTT Corporation, Japan), Yuusuke Nakano (NTT Network Technologies Laboratories, NTT Corporation, Japan), and Ken Nishimatsu (NTT Network Technologies Laboratories, NTT Corporation, Japan)</i> | |
| Application of LSTM Models to Predict In-Store Trajectory of Customers | 288 |
| <i>Weizheng Zhao (Collabrative Innovation Center for Transport Studies, Dalian Maritime University, China), Yi Zuo (Navigation College, Collaborative Innovation Center of Maritime Big Data & Shipping Artificial General Intelligence, Dalian Maritime University, China), Licheng Zhao (Collabrative Innovation Center for Transport Studies, Dalian Maritime University, China), and Junhao Jiang (Navigation College, Dalian Maritime University, China)</i> | |

CLEATED: Continual Learning and Adaptation for Time Evolving Data

| | |
|--|-----|
| Evaluating and Explaining Generative Adversarial Networks for Continual Learning Under Concept Drift | 295 |
| <i>Filip Guzy (Wroclaw University of Science and Technology, Poland), Michal Wozniak (Wroclaw University of Science and Technology, Poland), and Bartosz Krawczyk (Virginia Commonwealth University, USA)</i> | |
| Lightweight Alternatives for Hyper-Parameter Tuning in Drifting Data Streams | 304 |
| <i>Jesus L. Lobo (TECNALIA, Basque Research and Technology Alliance (BRTA), Spain), Javier Del Ser (TECNALIA, Basque Research and Technology Alliance (BRTA), Spain; University of the Basque Country (UPV/EHU), Spain), and Eneko Osaba (TECNALIA, Basque Research and Technology Alliance (BRTA), Spain)</i> | |

| | |
|--|-----|
| A Fully Unsupervised and Efficient Anomaly Detection Approach with Drift Detection Capability | 312 |
| <i>Chang How Tan (Monash University, Australia), Vincent CS Lee (Monash University, Australia), Mahsa Salehi (Monash University, Australia), Slaven Marusic (Aurecon, Australia), Srimal Jayawardena (Aurecon, Australia), and Dion Lucke (Aurecon, Australia)</i> | |

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

| | |
|---|-----|
| SGDOL: Self-Evolving Generative and Discriminative Online Learning for Data Stream Classification | 322 |
| <i>Deeksha Aggarwal (Spatial Computing Laboratory, Center for Data Sciences, IIT Bangalore, India), J. Senthilnath (Institute for Infocomm Research, Agency for Science, Technology and Research (A*STAR), Singapore), Uttam Kumar (Spatial Computing Laboratory, Center for Data Sciences, IIT Bangalore, India), Vivek Yadav (Spatial Computing Laboratory, Center for Data Sciences, IIT Bangalore, India), Sushant Kulkarni (Indian Institute of Technology, India), Md Meftahul Ferdaus (Institute for Infocomm Research, Agency for Science, Technology and Research (A*STAR), Singapore), and Li Xiaoli (Institute for Infocomm Research, Agency for Science, Technology and Research (A*STAR), Singapore)</i> | |
| IEBench: Benchmarking Streaming Learners on Imbalanced Evolving Data Streams | 331 |
| <i>Alessio Bernardo (DEIB - Politecnico di Milano, Italy), Giacomo Ziffer (DEIB - Politecnico di Milano, Italy), and Emanuele Della Valle (DEIB - Politecnico di Milano, Italy)</i> | |
| DRIFT LENS: Real-Time Unsupervised Concept Drift Detection by Evaluating per-Label Embedding Distributions | 341 |
| <i>Salvatore Greco (Politecnico di Torino, Italy) and Tania Cerquitelli (Politecnico di Torino, Italy)</i> | |
| NimbleLearn: A Scalable and Fast Batch-Mode Active Learning Approach | 350 |
| <i>Ruoyan Kong (University of Minnesota, USA), Zhanlong Qiu (Amazon, USA), Yang Liu (Amazon, USA), and Qi Zhao (Amazon, USA)</i> | |
| Incremental Clustering Algorithms for Massive Dynamic Graphs | 360 |
| <i>Johannes Langguth (Simula Research Laboratory, BI Norwegian Business School, Norway), Aigars Tumanis (University of Oslo, Norway), and Ariful Azad (Indiana University Bloomington, USA)</i> | |
| Customs Fraud Detection in the Presence of Concept Drift | 370 |
| <i>Tung-Duong Mai (Korea Advanced Institute of Science and Technology), Kien Hoang (Korea Advanced Institute of Science and Technology), Aitolkyn Baigutanova (Korea Advanced Institute of Science and Technology), Gaukhartas Alina (Korea Advanced Institute of Science and Technology), and Sundong Kim (Institute for Basic Science)</i> | |

| | |
|--|-----|
| Fast and Lightweight Binary and Multi-branch Hoeffding Tree Regressors | 380 |
| <i>Saulo Martiello Mastelini (Institute of Mathematics and Computer Science, University of São Paulo, Brazil), Jacob Montiel (AI Institute, University of Waikato, New Zealand), Heitor Murilo Gomes (AI Institute, University of Waikato, New Zealand), Albert Bifet (AI Institute, University of Waikato, New Zealand), Bernhard Pfahringer (Department of Computer Science, University of Waikato, New Zealand), and Andre C. P. L. F. de Carvalho (Institute of Mathematics and Computer Science, University of São Paulo, Brazil)</i> | |
| Crowd Behavior Detection in Videos using Statistical Physics | 389 |
| <i>Huiyu Mu (China agriculture university, China), Ruizhi Sun (China agriculture university, China), Gang Yuan (China agriculture university, China), Jiayao Li (China agriculture university, China), and Miao Wang (China agriculture university, China)</i> | |
| Accelerating Active Learning Image Labeling Through Bulk Shift Recommendations | 398 |
| <i>Philipp Scharpf (BMT AG and University of Konstanz, Switzerland and Germany), Chi Lap Hong (BMT AG, Switzerland), and Oliver Duerr (University of Applied Science (HTWG) Konstanz, Germany)</i> | |
| Multi-Label kNN Classifier with Online Dual Memory on Data Stream | 405 |
| <i>Xihui Wang (Laboratory of Digital Sciences of Nantes, France), Pascale Kuntz (Laboratory of Digital Sciences of Nantes, France), Frank Meyer (Orange Labs, France), and Vincent Lemaire (Orange Labs, France)</i> | |
| Online Changepoint Detection on a Budget | 414 |
| <i>Zhaohui Wang (Applied Research, Splunk, USA), Xiao Lin (Applied Research, Splunk, USA), Abhinav Mishra (Applied Research, Splunk, USA), and Ram Sriharsha (Applied Research, Splunk, USA)</i> | |
| Few-Shot Class-Incremental Learning with Meta-Learned Class Structures | 421 |
| <i>Guangtao Zheng (University of Virginia, USA) and Aidong Zhang (University of Virginia, USA)</i> | |

HDM: High Dimensional Data Mining

| | |
|---|-----|
| LUCKe- Connecting Clustering and Correlation Clustering | 431 |
| <i>Anna Beer (LMU Munich), Lisa Stephan (LMU Munich), and Thomas Seidl (LMU Munich)</i> | |
| Implicit Hough Transform Neural Networks for Subspace Clustering | 441 |
| <i>Julian Busch (LMU Munich, Germany), Maximilian Hünemörder (CAU Kiel, Germany), Janis Held (LMU Munich, Germany), Peer Kröger (CAU Kiel, Germany), and Thomas Seidl (LMU Munich, Germany)</i> | |
| TensorMode Algorithm for Network Embedding in Dynamic Environments | 449 |
| <i>Chris Connell (Indiana University) and Yang Wang (Indiana University)</i> | |
| ReTriM: Reconstructive Triplet Loss for Learning Reduced Embeddings for Multi-variate Time Series | 460 |
| <i>Yash Garg (Nokia Bell Labs - AI Research, USA)</i> | |
| Anomaly Detection with Dual Adversarial Training | 466 |
| <i>Shuo Liu (North China University of Technology, China) and Liwen Xu (North China University of Technology, China)</i> | |

| | |
|--|-----|
| Accelerating Density-Based Subspace Clustering in High-Dimensional Data | 474 |
| <i>Jürgen Prinzbach (Offenburg University of Applied Sciences, Germany), Tobias Lauer (Offenburg University of Applied Sciences, Germany), and Nicolas Kiefer (VEGA Grieshaber KG, Germany)</i> | |
| Random Projection Through the Lens of Data Complexity Indicators | 482 |
| <i>Yamonporn Thummanusarn (University of Birmingham) and Ata Kaban (University of Birmingham)</i> | |
| Causal Structure Learning of Nonlinear Additive Noise Model Based on Streaming Feature | 490 |
| <i>Jing Yang (Key Laboratory of Knowledge Engineering with Big Data(Hefei University of Technology), Ministry of Education; Hefei University of Technology, P. R. China), Gaojin Fan (Key Laboratory of Knowledge Engineering with Big Data(Hefei University of Technology), Ministry of Education; Hefei University of Technology, P. R. China), Anbo Shen (Key Laboratory of Knowledge Engineering with Big Data(Hefei University of Technology), Ministry of Education; Hefei University of Technology, P. R. China), and Aiguo Wang (Foshan University, China)</i> | |

DMBIH: Data Mining in Biomedical Informatics and Healthcare

| | |
|---|-----|
| Metagenome2Vec: Building Contextualized Representations for Scalable Metagenome Analysis | 500 |
| <i>Sathyanarayanan Aakur (Oklahoma State University), Vineela Indla (Oklahoma State University), Vennela Indla (Oklahoma State University), Sai Narayanan (Oklahoma State University), Arunkumar Bagavathi (Oklahoma State University), Vishalini Laguduva Ramnath (Oklahoma State University), and Akhilesh Ramachandran (Oklahoma State University)</i> | |
| Multimodal Machine Learning for 30-Days Post-Operative Mortality Prediction of Elderly Hip Fracture Patients | 508 |
| <i>Berk Yenidogan (University of Twente, Netherlands), Shreyasi Pathak (University of Twente, Netherlands; Hospital Group Twente (ZGT), Netherlands), Jeroen Geerdink (Hospital Group Twente(ZGT)), Johannes H. Hegeman (University of Twente, Netherlands; Hospital Group Twente (ZGT), Netherlands), and Maurice van Keulen (University of Twente, Netherlands)</i> | |
| Empirical Quantitative Analysis of COVID-19 Forecasting Models | 517 |
| <i>Yun Zhao (University of California, Santa Barbara, USA), Yuqing Wang (University of California, Santa Barbara, USA), Junfeng Liu (University of California, Santa Barbara, USA), Haotian Xia (University of California, Santa Barbara, USA), Zhenni Xu (University of California, Santa Barbara, USA), Qinghang Hong (University of California, Santa Barbara, USA), Zhiyang Zhou (Northwestern University, USA), and Linda Petzold (University of California, Santa Barbara, USA)</i> | |

LITSA: Large-scale Industrial Time Series Analysis

| | |
|--|-----|
| EnsembleNTLDetect: An Intelligent Framework for Electricity Theft Detection in Smart Grid | 527 |
| <i>Yogesh Kulkarni (Pune Institute of Computer Technology, India), Sayf Hussain Z (Anna University, India), Krithi Ramamritham (Robert Bosch Centre for Data Science & Artificial Intelligence, IIT Madras, India), and Nivethitha Somu (IIT Bombay, India)</i> | |
| Attention Augmented Convolutional Transformer for Tabular Time-Series | 537 |
| <i>Sharath M Shankaranarayana (Airt Research, Croatia) and Davor Runje (Airt Research, Croatia; Algebra University College, Croatia)</i> | |
| Anomaly Detection for Multivariate Time Series on Large-Scale Fluid Handling Plant using Two-Stage Autoencoder | 542 |
| <i>Susumu Naito (Corporate R&D Center, Toshiba Corporation, Japan), Yasunori Taguchi (Corporate R&D Center, Toshiba Corporation, Japan), Kouta Nakata (Corporate R&D Center, Toshiba Corporation, Japan), and Yuichi Kato (Corporate R&D Center, Toshiba Corporation, Japan)</i> | |
| Towards Dynamic Structure Changes Detection in Financial Series via Causal Analysis | 552 |
| <i>Patrick Owusu (LORIA, University of Lorraine, University of Sherbrooke, Canada), Armelle Brun (LORIA, University of Lorraine, France), and Shengrui Wang (Department of Computer Science, University of Sherbrooke, Canada)</i> | |
| Forecasting of Reservoir Inflow by the Combination of Deep Learning and Conventional Machine Learning | 558 |
| <i>Topon Paul (Corporate R&D Center, Toshiba Corporation, Japan), Sreeharsha Raghavendra (R&D Division, Toshiba Software (India) Pvt. Ltd., India), Ken Ueno (Corporate R&D Center, Toshiba Corporation, Japan), Fang Ni (Corporate R&D Center, Toshiba Corporation, Japan), Hiromasa Shin (Corporate R&D Center, Toshiba Corporation, Japan), Kaneharu Nishino (Corporate R&D Center, Toshiba Corporation, Japan), and Ryusei Shingaki (Corporate R&D Center, Toshiba Corporation, Japan)</i> | |

UDML: Utility Driven Mining and Learning

| | |
|---|-----|
| Mining High Utility Subgraphs | 566 |
| <i>Md. Tanvir Alam (University of Dhaka, Bangladesh), Amit Roy (University of Dhaka, Bangladesh), Chowdhury Farhan Ahmed (University of Dhaka, Bangladesh), Md. Ashraful Islam (University of Dhaka, Bangladesh), and Carson Leung (University of Manitoba, Canada)</i> | |
| Sequence Prediction using Partially-Ordered Episode Rules | 574 |
| <i>Yangming Chen (Harbin Inst. of Technology (Shenzhen), China), Philippe Fournier-Viger (Harbin Inst. of Technology (Shenzhen), China), Farid Nouioua (University of Bordj Bou Arreridj, Algeria), and Youxi Wu (Hebei University of Technology, China)</i> | |
| Personalized Neural Architecture Search | 581 |
| <i>Cedric Kulbach (FZI Research Center for Information Technology, Germany) and Steffen Thoma (FZI Research Center for Information Technology, Germany)</i> | |

| | |
|--|-----|
| Large-Scale Closed High-Utility Itemset Mining | 591 |
| <i>Jerry Chun-Wei Lin (Western Norway University of Applied Sciences, Norway), Youcef Djenouri (SINTEF Digital, Norway), Gautam Srivastava (Brandon University, Canada; China Medical University, Taiwan), and Jimmy Ming-Tai Wu (Shandong University of Science and Technology, China)</i> | |
| CHUQI-Miner: Mining Correlated Quantitative High Utility Itemsets | 599 |
| <i>Mourad Nouioua (Harbin Inst. of Technology (Shenzhen), China; University of Bordj Bou Arreridj, Algeria), Philippe Fournier-Viger (Harbin Inst. of Technology (Shenzhen), China), Jun-Feng Qu (Hubei University of Arts and Science, China), Jerry Chun-Wei Lin (Western Norway University of Applied Sciences (HVL), Norway), Wensheng Gan (Jinan University, China), and Wei Song (North China University of Technology, China)</i> | |
| A Unified Framework to Discover Partial Periodic-Frequent Patterns in Row and Columnar Temporal Databases | 607 |
| <i>Veena Pamalla (Sri Balaji PG college, India), So Nakamura (The University of Aizu, Japan), Likhitha Palla (The University of Aizu, Japan), Uday Kiran Rage (The University of Aizu, Japan), Yutaka Watanobe (The University of Aizu, Japan), and Koji Zettsu (NICT, Japan)</i> | |
| TopUMS: Top-k Utility Mining in Stream Data | 615 |
| <i>Wei Song (North China University of Technology, Beijing, China), Caiyu Fang (North China University of Technology, Beijing, China), and Wensheng Gan (Jinan University, Guangzhou, China)</i> | |
| Optimal Segmented Linear Regression for Financial Time Series Segmentation | 623 |
| <i>Chi-Jen Wu (Chang Gung University, Taiwan), Wei-Sheng Zeng (Institute of Information Science, Academia Sinica, Taiwan), and Jan-Ming Ho (Institute of Information Science, Academia Sinica, Taiwan)</i> | |

MLLD: Mining and Learning in the Legal Domain

| | |
|--|-----|
| Simplify Your Law: using Information Theory to Deduplicate Legal Documents | 631 |
| <i>Corinna Coupette (Max Planck Institute for Informatics, Germany), Jyotsna Singh (CISPA Helmholtz Center for Information Security, Germany), and Holger Spamann (Harvard Law School, USA)</i> | |
| Detection of Similar Legal Cases on Personal Injury | 639 |
| <i>Jason Lam (Queen's University, Canada), Yuhao Chen (Queen's University, Canada), Farhana Zulkernine (Queen's University, Canada), and Samuel Dahan (Queen's University, Canada)</i> | |
| Determining Standard Occupational Classification Codes from Job Descriptions in Immigration Petitions | 647 |
| <i>Sourav Mukherjee (Fairleigh Dickinson University, Canada), David Widmark (Berry Appleman & Leiden LLP, USA), Vince DiMascio (Berry Appleman & Leiden LLP, USA), and Tim Oates (Synaptiq, USA)</i> | |

| | |
|--|-----|
| Legal Entity Extraction using a Pointer Generator Network | 653 |
| <i>Stavroula Skylaki (Thomson Reuters Labs, Switzerland), Ali Oskooei (Thomson Reuters Labs, Switzerland), Omar Bari (Thomson Reuters Labs, Canada), Nadja Herger (Thomson Reuters Labs, Switzerland), and Zac Kriegman (Thomson Reuters Labs, US)</i> | |

DLC: Deep Learning and Clustering

| | |
|---|-----|
| SSPF: A Simple and Scalable Parameter Free Clustering Method | 659 |
| <i>Xingyu Cai (Baidu Research USA), Yuchen Bian (Baidu Research USA), Jiayi Huang (Baidu Research USA), Boxiang Liu (Baidu Research USA), Jiahong Yuan (Baidu Research USA), and Kenneth Church (Baidu Research USA)</i> | |
| Versatile Feature Learning with Graph Convolutions and Graph Structures | 669 |
| <i>Guojing Cong (Oakridge National Laboratory) and Seung-Hwan Lim (Oakridge National Laboratory)</i> | |
| Sparse Subspace K-Means | 678 |
| <i>Abdoul Wahab Diallo (African Institute for Mathematical Sciences (AIMS), Senegal), Ndèye Niang (CEDRIC LAB, CNAM, France), and Mory Ouattara (UFR SFA, Université NANGUI ABROGOUA)</i> | |
| Deep Embedded K-Means Clustering | 686 |
| <i>Wengang Guo (Tongji University, China), Kaiyan Lin (Tongji University, China), and Wei Ye (Tongji University, China)</i> | |
| Convolutional Variational Autoencoders for Image Clustering | 695 |
| <i>Ioannis Nellas (University of Thessaly, Greece), Sotiris Tasoulis (University of Thessaly, Greece), and Vassilis Plagianakos (University of Thessaly, Greece)</i> | |
| Model-Based Poisson co-Clustering for Attributed Networks | 703 |
| <i>Paul Riverain (Thales Research & Technology, Université de Paris, France), Simon Fossier (Thales Research & Technology, France), and Mohamed Nadif (Université de Paris, France)</i> | |
| Graph Representation Learning with Adaptive Mixtures | 711 |
| <i>Da Sun Handason Tam (The Chinese University of Hong Kong, Hong Kong), Siyue Xie (The Chinese University of Hong Kong, Hong Kong), and Wing Cheong Lau (The Chinese University of Hong Kong, Hong Kong)</i> | |
| Unsupervised Graph-Clustering Learning Framework for Financial news Summarization | 719 |
| <i>Jun Wang (Southwestern University Of Finance And Economics, China), Jinghua Tan (Southwestern University Of Finance And Economics, China), Hanlei Jin (Southwestern University Of Finance And Economics, China), and Shuo Qi (Southwestern University Of Finance And Economics, China)</i> | |

SSTDM: Spatial and Spatio-Temporal Data Mining

| | |
|--|-----|
| Detecting Wandering Behavior of People with Dementia | 727 |
| <i>Nicklas Sindlev Andersen (Institute of Mathematics and Computer Science, University of Southern Denmark, Denmark), Marco Chiarandini (Institute of Mathematics and Computer Science, University of Southern Denmark, Denmark), Stefan Jänicke (Institute of Mathematics and Computer Science, University of Southern Denmark, Denmark), Panagiotis Tampakis (Institute of Mathematics and Computer Science, University of Southern Denmark, Denmark), and Arthur Zimek (Institute of Mathematics and Computer Science, University of Southern Denmark, Denmark)</i> | |
| Passenger Flow Forecasting on Transportation Network: Sensitivity Analysis of the Spatiotemporal Features | 734 |
| <i>Johanna Baro (IRT SystemX, France) and Mostepha Khouadjia (IRT SystemX, France)</i> | |
| On the Unreasonable Efficiency of State Space Clustering in Personalization Tasks | 742 |
| <i>Anton Dereventsov (Lirio, LLC, USA), Raju Vatsavai (Lirio, LLC, USA), and Clayton Webster (Lirio, LLC, USA)</i> | |
| Deriving Spatio-Temporal Trajectory Fingerprints from Mobility Data using Non-Negative Matrix Factorisation | 750 |
| <i>Michiel Dhont (EluciDATA Lab of Sirris, Electronics and Information Processing (ETRO) - VUB, Belgium), Elena Tshiporkova (EluciDATA Lab of Sirris, Belgium), and Nicolás González-Deleito (EluciDATA Lab of Sirris, Belgium)</i> | |
| Disjoint-CNN for Multivariate Time Series Classification | 760 |
| <i>Seyed Navid Mohammadi Foumani (Monash University, Australia), Chang Wei Tan (Monash University, Australia), and Mahsa Salehi (Monash University, Australia)</i> | |
| Functional Foot Segmentation Based on Plantar Pressure Measurements for Profiling Subjects Performing a Running Exercise | 770 |
| <i>Peter Van Hese (EluciDATA Lab, Sirris, Belgium), Henrique Cabral (EluciDATA Lab, Sirris, Belgium), Wouter Aerts (R&D Department Materialise Motion, Materialise Motion, Belgium), and Elena Tshiporkova (EluciDATA Lab, Sirris, Belgium)</i> | |
| STONE: Signal Temporal Logic Neural Network for Time Series Classification | 778 |
| <i>Ruixuan Yan (Rensselaer Polytechnic Institute, USA), Agung Julius (Rensselaer Polytechnic Institute, USA), Maria Chang (IBM T.J. Watson Research Center, IBM Research, USA), Achille Fokoue (IBM T.J. Watson Research Center, IBM Research, USA), Tengfei Ma (IBM T.J. Watson Research Center, IBM Research, USA), and Rosario Uceda-Sosa (IBM T.J. Watson Research Center, IBM Research, USA)</i> | |

DMC: Data Mining and Machine Learning in Cybersecurity

| | |
|---|-----|
| Identifying Darknet Vendor Wallets by Matching Feedback Reviews with Bitcoin Transactions | 788 |
| <i>Xucan Chen (Georgia State University, USA), Wei Cheng (NEC Laboratories America, Inc, USA), Marie Ouellet (Georgia State University, USA), Yuan Li (North China University of Technology, China), David Maimon (Georgia State University, USA), and Yubao Wu (Georgia State University, USA)</i> | |

| | |
|---|-----|
| Shedding Light in the Tunnel: Counting Flows in Encrypted Network Traffic | 798 |
| <i>Fares Meghdouri (Technische Universitat Wien), Félix Iglesias Vázquez (Technische Universitat Wien), and Tanja Zseby (Technische Universitat Wien)</i> | |
| Static Analysis for Android Malware Detection with Document Vectors | 805 |
| <i>Utkarsh Raghav (University of Canberra, Australia), Elisa Martinez-Marroquin (University of Canberra, Australia), and Wanli Ma (University of Canberra, Australia)</i> | |
| Faster Classification using Compression Analytics | 813 |
| <i>Christina Ting (Sandia National Laboratories), Nicholas Johnson (Sandia National Laboratories), Uzoma Onunkwo (Sandia National Laboratories), and J. Derek Tucker (Sandia National Laboratories)</i> | |

EDMML: Evolutionary Data Mining and Machine Learning

| | |
|---|-----|
| Multi-objective Feature Selection with a Sparsity-Based Objective Function and Gradient Local Search for Multi-label Classification | 823 |
| <i>Kaan Demir (Victoria University of Wellington, New Zealand), Bach Nguyen (Victoria University of Wellington, New Zealand), Bing Xue (Victoria University of Wellington, New Zealand), and Mengjie Zhang (Victoria University of Wellington, New Zealand)</i> | |
| Automated and Efficient Sparsity-Based Feature Selection via a Dual-Component Vector | 833 |
| <i>Bach Nguyen (Victoria University of Wellington, New Zealand), Bing Xue (Victoria University of Wellington, New Zealand), and Mengjie Zhang (Victoria University of Wellington, New Zealand)</i> | |
| Instance Selection for Multi-Label Learning Based on a Scalable Evolutionary Algorithm | 843 |
| <i>Juan Antonio Romero del Castillo (University of Córdoba), Domingo Ortiz-Boyer (University of Córdoba), and Nicolás García-Pedrajas (University of Córdoba)</i> | |
| Transformer-Based Hierarchical Encoder for Document Classification | 852 |
| <i>Harsh Sakhrani (Pune Institute of Computer Technology, India), Saloni Parekh (Pune Institute of Computer Technology, India), and Pratik Ratadiya (vCreaTek Consulting Services Pvt. Ltd., India)</i> | |

SDM: Social Data Mining in the Post-pandemic Era

| | |
|---|-----|
| Cross-Lingual COVID-19 Fake News Detection | 859 |
| <i>Jiangshu Du (University of Illinois at Chicago, USA), Yingtong Dou (University of Illinois at Chicago, USA), Congying Xia (University of Illinois at Chicago, USA), Limeng Cui (Pennsylvania State University, USA), Jing Ma (Hong Kong Baptist University, Hong Kong), and Philip Yu (University of Illinois at Chicago, USA)</i> | |
| Patient Preferences: An Unexplored Area in the Post-Pandemic Era | 863 |
| <i>Haridimos Kondylakis (Institute of Computer Science, FORTH, Greece), Angelina Kouroubali (Institute of Computer Science, FORTH, Greece), and Dimitrios Katehakis (Institute of Computer Science, FORTH, Greece)</i> | |

| | |
|--|-----|
| A Sentiment-Aware Delightful Walking Route Recommendation System Considering the Scenery and Season | 867 |
| <i>Da Li (Kyoto Sangyo University, Japan), Shiho Ishitsubo (Kyoto Sangyo University, Japan), Katsuyuki Yamauchi (Kyoto Sangyo University, Japan), Panote Siriaraya (Kyoto Institute of Technology, Japan), Shinsuke Nakajima (Kyoto Sangyo University, Japan), and Yukiko Kawai (Kyoto Sangyo University, Japan)</i> | |
| Effects of Stimulus Checks on Spending Patterns of Different Economic Groups | 873 |
| <i>Nidhi Mulay (AI Garage, Mastercard), Vikas Bishnoi (AI Garage, Mastercard), Yatin Katyaj (AI Garage, Mastercard), Mohsen Bahrami (Massachusetts Institute of Technology), Esteban Moro (Massachusetts Institute of Technology), Ankur Saraswat (AI Garage, Mastercard), and Alex Pentland (Massachusetts Institute of Technology)</i> | |
| Analysis of User Behavior in a C2C Platform During COVID-19 Pandemic | 878 |
| <i>Taisei Ninomiya (The University of Tokyo, Japan), Fujio Toriumi (The University of Tokyo, Japan), Hiroka Zaitso (GMO Pepabo, Inc., Japan), and Yusuke Miyake (GMO Pepabo, Inc., Japan)</i> | |
| Analyzing the Bad-Words in Tweets of Twitter Users to Discover the Mental Health Happiness Index and Feel-Good-Factors | 882 |
| <i>Sudha Tushara Sadasivuni (Georgia State University, USA) and Yanqing Zhang (Georgia State University, USA)</i> | |
| HappyRec: Evaluation of a "Happy Spot" Recommendation System Aimed at Improving Mental Well-Being | 889 |
| <i>Kenta Suzuki (Kyoto Sangyo University, Japan), Panote Siriaraya (Kyoto Institute of Technology, Japan), Wan Jou She (Weill Cornell Medicine, United States), Ryotaro Tanaka (Kyoto Sangyo University, Japan), Da Li (Kyoto Sangyo University, Japan), and Shinsuke Nakajima (Kyoto Sangyo University, Japan)</i> | |
| Online Partisan Polarization of COVID-19 | 893 |
| <i>Zachary Yang (McGill University, Mila - Quebec AI Institute, Canada), Anne Imouza (Université de Montréal, Canada), Kellin Pelrine (McGill University, Mila - Quebec AI Institute, Canada), Sacha Lévy (McGill University, Mila - Quebec AI Institute, Canada), Jiewen Liu (McGill University, Mila - Quebec AI Institute, Canada), Gabrielle Desrosiers-Brisebois (Université de Montréal, Canada), Jean-François Godbout (Université de Montréal, Canada), André Blais (Université de Montréal, Canada), and Reihaneh Rabbany (McGill University, Mila - Quebec AI Institute, Canada)</i> | |

OEDM: Optimization Based Techniques for Emerging Data Mining Problems

| | |
|--|-----|
| Mixture Gaussian Prototypes for Few-Shot Learning | 902 |
| <i>Ruijin Jiang (Beijing University of Posts and Telecommunications, China) and Zhaohui Cheng (China University of Geosciences, China)</i> | |

| | |
|---|-----|
| Overview of Optimization Algorithms for Large-Scale Support Vector Machines | 909 |
| <i>Xuchan Ju (Shenzhen University, China), Zhenghao Yan (Shenzhen University, China), and Tianhe Wang (Shenzhen Branch, China Merchants Bank, China)</i> | |
| Machine Learning and Deep Learning Methods used in Safety Management of Nuclear Power Plants: A Survey | 917 |
| <i>Yong Shi (Chinese Academy of Sciences, China; University of Nebraska at Omaha, USA), Xiaodong Xue (University of Chinese Academy of Sciences, China; Chinese Academy of Sciences, China), Yi Qu (Chinese Academy of Sciences, China), Jiayu Xue (Chinese Academy of Sciences, China), and Linzi Zhang (Chinese Academy of Sciences, China)</i> | |
| SCORER-Gap: Sequentially Correlated Rules for Event Recommendation Considering Gap Size | 925 |
| <i>Ludwig Zellner (LMU Munich, Germany), Janina Sontheim (LMU Munich, Germany), Florian Richter (LMU Munich, Germany), Gabriel Lindner (LMU Munich, Germany), and Thomas Seidl (LMU Munich, Germany)</i> | |

IAAA: Intelligence-Augmented Anomaly Analytics

| | |
|--|-----|
| Fake Reviewer Group Detection in Online Review Systems | 935 |
| <i>Chen Cao (Dalian University of Technology, China), Shihao Li (Dalian University of Technology, China), Shuo Yu (Dalian University of Technology, China), and Zhikui Chen (Dalian University of Technology, China)</i> | |
| A Human-in-the-Loop Approach Based on Explainability to Improve NTL Detection | 943 |
| <i>Bernat Coma-Puig (Universitat Politècnica de Catalunya, Spain) and Josep Carmona (Universitat Politècnica de Catalunya, Spain)</i> | |
| Cross Network Representation Matching with Outliers | 951 |
| <i>Mingliang Hou (Dalian University of Technology, China), Jing Ren (Federation University Australia, Australia), Falih Febrinanto (Federation University Australia, Australia), Ahsan Shehzad (Dalian University of Technology, China), and Feng Xia (Federation University Australia, Australia)</i> | |
| Deep Video Anomaly Detection: Opportunities and Challenges | 959 |
| <i>Jing Ren (Federation University Australia), Feng Xia (Federation University Australia), Yemeng Liu (Dalian University of Technology), and Ivan Lee (STEM, University of South Australia)</i> | |
| Early Prediction of Hate Speech Propagation | 967 |
| <i>Ken-Yu Lin (National Chiao Tung University), Roy Ka-Wei Lee (Singapore University of Technology and Design), Wei Gao (Singapore Management University), and Wen-Chih Peng (National Chiao Tung University)</i> | |
| Surrogate Supervision-Based Deep Weakly-Supervised Anomaly Detection | 975 |
| <i>Zhiyue Wu (National University of Defense Technology, China), Hongzuo Xu (National University of Defense Technology, China), Yijie Wang (National University of Defense Technology, China), and Yongjun Wang (National University of Defense Technology, China)</i> | |

| | |
|---|------|
| Temporal Graph Representation Learning for Detecting Anomalies in E-Payment Systems | 983 |
| <i>Yiming Li (The Chinese University of Hong Kong), Da Sun Handason Tam (The Chinese University of Hong Kong), Siyue Xie (The Chinese University of Hong Kong), Xiaxin Liu (Tencent Technology Co.Ltd, China), Qiu Fang Ying (Tencent Technology Co.Ltd, China), Wing Cheong Lau (The Chinese University of Hong Kong), Dah Ming Chiu (The Chinese University of Hong Kong), and Shou Zhi Chen (Tencent Technology Co.Ltd, China)</i> | |
| OAB — An Open Anomaly Benchmark Framework for Unsupervised and Semisupervised Anomaly Detection on Image and Tabular Data Sets | 991 |
| <i>Andreas Lohrer (Kiel University, Germany), Jan Deller (Kiel University, Germany), Maximilian Hünemörder (Kiel University, Germany), and Peer Kröger (Kiel University, Germany)</i> | |
| Neural Architecture Search and Multi-Objective Evolutionary Algorithms for Anomaly Detection | 1001 |
| <i>Chakkrit Termritthikun (Naresuan University, Thailand), Lin Xu (University of South Australia, Australia), Yemeng Liu (Dalian University of Technology, China), and Ivan Lee (University of South Australia, Australia)</i> | |
| SFE-TSDM: Systematic Feature Engineering for Time-Series Data Mining | |
| A Fast Sorting-Based Aggregation Method for Symbolic Time Series Representation | 1009 |
| <i>Xinye Chen (University of Manchester, United Kingdom) and Stefan Güttel (University of Manchester, United Kingdom)</i> | |
| Feature Selection for Multivariate Time Series via Network Pruning | 1017 |
| <i>Kang Gu (Dartmouth College), Soroush Vosoughi (Dartmouth College), and Temiloluwa Prioleau (Dartmouth College)</i> | |
| Time Series Ordinal Regression for Supporting the Storage of Temperature Sensitive Medication in Domestic Refrigerators | 1025 |
| <i>Ali Hammadeh (University of Auckland, New Zealand), Amin Zayani (MedAngel, Netherlands), and Andreas W. Kempa-Liehr (The University of Auckland, New Zealand)</i> | |
| An Empirical Evaluation of Time-Series Feature Sets | 1032 |
| <i>Trent Henderson (The University of Sydney, Australia) and Ben Fulcher (The University of Sydney, Australia)</i> | |
| Evaluating Time Series Predictability via Transition Graph Analysis | 1039 |
| <i>Anton Kovantsev (National Center for Cognitive Research, ITMO University, Russia), Petr Chunaev (National Center for Cognitive Research, ITMO University, Russia), and Klavdiya Bochenina (National Center for Cognitive Research, ITMO University, Russia)</i> | |
| Comparison of Variant Principal Component Analysis using New RNN-Based Framework for Stock Prediction | 1047 |
| <i>Shuyan Liu (Southwest Petroleum University, China) and Yue Chen (dept. of Mathematics, National University of Singapore, Singapore)</i> | |

| | |
|---|------|
| Data Mining on Extremely Long Time-Series | 1057 |
| <i>Scott Simmons (University of Auckland, New Zealand), Louis Jarvis (University of Auckland, New Zealand), David Dempsey (University of Canterbury, New Zealand), and Andreas W. Kempa-Liehr (University of Auckland, New Zealand; Freiburg Materials Research Center, University of Freiburg, Germany)</i> | |
| Feature Selection on a Flare Forecasting Testbed: A Comparative Study of 24 Methods | 1067 |
| <i>Atharv Yeolekar (Georgia State University, USA), Sagar Patel (Georgia State University, USA), Shreejaa Talla (Georgia State University, USA), Krishna Rukmini Puthucode (Georgia State University, USA), Azim Ahmadzadeh (Georgia State University, USA), Viacheslav M. Sadykov (Georgia State University, USA), and Rafal A. Angryk (Georgia State University, USA)</i> | |

PhD Forum

| | |
|---|------|
| Deep Reinforcement Learning Task for Portfolio Construction | 1077 |
| <i>Boris Belyakov (National Research University Higher School of Economics, Russia) and Dmitry Sizykh (National Research University Higher School of Economics, Russia)</i> | |
| Self-Supervised Source Code Annotation from Related Research Papers | 1083 |
| <i>Pierre Haritz (Technical University Dortmund, Germany), Lukas Pfahler (Technical University Dortmund, Germany), Thomas Liebig (Technical University Dortmund, Germany), and Helena Kotthaus (ML2R, Germany)</i> | |
| A Gamified Approach to Automatically Detect Biased Wording and Train Critical Reading | 1085 |
| <i>Smilla Hinterreiter (Institute of Industrial Design, University of Applied Sciences Magdeburg-Stendal, Germany)</i> | |
| Early Detection of Atmospheric Turbulence for Civil Aircraft: A Data Driven Approach | 1087 |
| <i>Tianyi Li (université de Toulouse & Airbus, France), Philippe Goupil (Aircraft Control, Airbus, France), Josiane Mothe (Université de Toulouse, France), and Olivier Teste (Université de Toulouse, France)</i> | |
| Multi-channel Convolution Neural Network for Gas Mixture Classification | 1094 |
| <i>YongKyung Oh (Ulsan National Institute of Science and Technology (UNIST), Republic of Korea) and Sungil Kim (Ulsan National Institute of Science and Technology (UNIST), Republic of Korea)</i> | |
| An Interdisciplinary Approach for the Automated Detection and Visualization of Media Bias in News Articles | 1096 |
| <i>Timo Spinde (University of Wuppertal)</i> | |
| A Knowledge-Aware and Time-Sensitive Financial News Recommendation System Based on Firm Relation Derivation | 1104 |
| <i>Jun Wang (Southwestern University of Finance and Economics, China), Jinghua Tan (Southwestern University of Finance and Economics, China), Jiujiu Chen (Southwestern University of Finance and Economics, China), and Hanlei Jin (Southwestern University of Finance and Economics, China)</i> | |

| | |
|--|-------------|
| Optimal Option Hedging with Policy Gradient | 1112 |
| <i>Bo Xiao (City University of Hong Kong, China), Wuguannan Yao (City University of Hong Kong, China), and Xiang Zhou (City University of Hong Kong, Hong Kong Institute of Data Science, China)</i> | |
| Author Index | 1121 |

Message from the ICDM 2021 General Chairs

On behalf of the organizing committee of the IEEE ICDM 2021 conference and our virtual host Auckland, New Zealand, it is our great pleasure to welcome you to the 2021 IEEE International Conference on Data Mining. Due to the COVID-19 global pandemic, IEEE ICDM 2021 will be hosted virtually for the second time.

Our goal is to run the conference in a way that replicates an in-person conference as best as we can, while leveraging events that can only happen virtually. We have live keynotes, tutorials, and special sessions. We also have speed-networking sessions, online activities, and are trialing a job matching program for the first time. Additionally, this is the first time there have been Reproducibility Chairs, and a Diversity and Inclusion Chair at ICDM. We hope to ultimately encapsulate this IEEE ICDM 2021 conference with our unique New Zealand hospitality.

The organization of a successful conference would not be possible without the dedicated efforts from many individuals. In particular, we would like to express our gratitude to the Program Chairs *Pauli Miettinen*, University of Eastern Finland (Finland), *James Bailey*, The University of Melbourne (Australia); Steering Committee Chair *Xindong Wu*, Mininglamp Academy of Sciences (China); Workshop Chairs *Bing Xue*, Victoria University of Wellington (New Zealand), *Mykola Pechenizkiy*, Eindhoven University of Technology (The Netherlands); Tutorial Chairs *Wei Liu*, University of Technology Sydney (Australia), *Katharina Morik*, TU Dortmund (Germany); Online Experience/Virtual Chair *Heitor Gomes*, University of Waikato (New Zealand); Publicity Chair *Diana Benavides Prado*, The University of Auckland (New Zealand); Finance Chair *Gillian Dobbie*, The University of Auckland (New Zealand); Sponsorship Chairs *Kaiqi Zhao*, The University of Auckland (New Zealand), *Philippe Fournier-Viger*, Harbin Institute of Technology (China), *Eva García-Martín*, Ekkono Solutions (Sweden); PhD Forum Chairs *Sibylle Hess*, Eindhoven University of Technology (The Netherlands), *Meng-Fen Chiang*, The University of Auckland (New Zealand), *Lisi Chen*, Inception Institute of Artificial Intelligence (IIAI) (UAE); Diversity and Inclusion Chair *Richi Nayak*, Queensland University of Technology (Australia); Job Matching Chairs *Albert Bifet*, University of Waikato (New Zealand), *Lin Liu*, University of South Australia (Australia); Panel Chairs *Chenliang Li*, Wuhan University (China), *Michael Witbrock*, The University of Auckland (New Zealand); Newcomers Chairs *Huan Liu*, Arizona State University (USA), *Katerina Taskova*, The University of Auckland (New Zealand); Local Arrangement Chair *Joerg Wicker*, The University of Auckland (New Zealand); Reproducibility Chairs *Dragi Kocev*, Jožef Stefan Institute (Slovenia), *Jacob Montiel*, University of Waikato (New Zealand); Award Committee Chair *Xia Ning*, The Ohio State University (USA).

We owe special thanks to our sponsors of the conference, including the US National Science Foundation (NSF), School of Computer Science - The University of Auckland, IEEE Technical Committee on Intelligent Informatics, Google and Two Sigma. We thank the authors, keynote speakers, special session speakers, panelists, and tutorial speakers for agreeing to present their sessions virtually. We also thank the workshop organizers for running their workshops virtually.

Finally, we thank all researchers, practitioners and students who are working in the field of data mining for their support and promotion of ICDM over the years. We wish you a productive conference with new discoveries, new collaborations and a very enjoyable virtual experience.

Yun Sing Koh and Dacheng Tao

IEEE ICDM 2021 General Co-Chairs

Message from the ICDM 2021 Program Chairs

Since its inception in 2001, the IEEE International Conference on Data Mining (ICDM) has become a premier forum for researchers, users, practitioners, and developers to exchange and disseminate not only original research results but also new research directions in data mining. The 21st IEEE ICDM is being hosted this year in Auckland, New Zealand and run as an entirely virtual conference, due to the COVID-19 pandemic

It is our great pleasure to welcome you to ICDM 2021 and to present its proceedings to you. The ICDM conference is truly an international forum. During its nineteen-year history, the conference has been held in ten countries around the world. This year's conference continues this global trend: Our organizing and program committee members represent around 36 countries/regions. This year's conference was extremely competitive. We are pleased to announce 990 paper submissions from 46 different countries/regions for review. Best efforts were made to ensure each paper was reviewed by at least three program committee members and the selection was made on the basis of discussion among the reviewers, an area chair, and the program co-chairs. Like previous years, we implemented a triple-blind review process, ensuring that the reviewers do not know the identity of the authors or of the other individuals reviewing the same submission. This process is intended to remove bias during the paper discussions. This year, 98 regular papers (9.9% acceptance rate) and 100 short papers were selected for inclusion in the proceedings and program (giving a total acceptance rate of 20%). Of the papers that were submitted, 573 (57.9%) had student first authors. These authors represent the future of our field.

In keeping with the goal of advancing the state-of-the-art in data mining, paper topics span a range of areas including: novel data mining algorithms in traditional areas, models and algorithms for new, structured, data types; deep learning and its applications; mining sequences and sequential data; mining spatial and temporal datasets; mining textual and unstructured datasets; high performance implementations of data mining algorithms; stream data mining; mining and link analysis in networked settings; data mining in electronic commerce (e.g., recommendation); web search, advertising, and marketing tasks; methodological aspects and the KDD process; and healthcare, epidemic modeling, and clinical research.

In addition to the technical presentations, our program also highlights three outstanding keynotes given by internationally renowned, distinguished scientists Jian Pei (Simon Fraser University), Masashi Sugiyama (RIKEN Center for Advanced Intelligence Project / The University of Tokyo) and Svetha Venkatesh (Deakin University). Four tutorials will be offered and 21 workshops will be run in conjunction with the main conference. A job matching program will also be run as part of the conference. We would like to thank all those who invested their substantial efforts into making this conference what it is, starting with all the authors of the 990 manuscripts for submitting content for the conference. Reviewing and selecting papers from a large set of submissions required the coordinated effort of many individuals. We want to thank the 55 Area Chairs and 398 Program Committee members who provided insightful feedback to the authors and helped with this selection process.

Organizing the ICDM 2021 program required the time and expertise of numerous contributors. We are very thankful for the outstanding work of Bing Xue and Mykola Pechenizkiy who served as Workshop Co-Chairs, Wei Liu and Katharina Morik who were the Tutorial Co-Chairs, Lisi Chen, Sibylle Hess and Meng-Fen Chiang who organized the PhD Forum, Chenliang Li and Michael Witbrock who were the Panel Co-Chairs, Diana Benvides Prado who served as Publicity Chair, sponsorship chairs Kaiqi Zhao, Philippe Fournier-Viger and Eva Garcia Martin, newcomers chairs Huan Liu and Katerina Taskova, job matching chairs Albert Bifet and Lin Liu, reproducibility chairs Dragi Kocev and Jacon Montiel, Xia Ning who chaired the Best Paper Award Committee, and Joerg Wicker, Gillian Dobbie, Heitor Gomes who served as Local Arrangement Chair, Finance Chair and Online experience/virtual Chair respectively.

The guidance of the Steering Committee Chair, Xindong Wu, and the General Co-Chairs, Yun Sing Koh and Dacheng Tao, were invaluable throughout each step of the conference organization and we wish to express our appreciation to them for their tireless efforts. We also would like to extend our special thanks to Juzhen Dong for the many hours she put in to maintain and enhance the Cyberchair web system to support the conference.

Finally, we thank the ICDM community for their support of this premier conference. We hope you enjoy the ICDM conference and that you are inspired by the ideas found in these papers.

James Bailey and Pauli Miettinen
ICDM 2021 Program Co-Chairs

Message from the Workshop Chairs

The 21th IEEE International Conference on Data Mining (IEEE ICDM 2021) is a premier and truly international conference for researchers and practitioners in the broad area of data mining. The ICDM Workshops program (IEEE ICDMW) aims to provide a platform for multiple workshops with a range of more focused topics to be discussed and explored, where attendees can present their original results, exchange research ideas, identify limitations, and explore new opportunities on the theoretical development and real-world applications of data mining techniques.

Due to the global COVID-19 pandemic, IEEE ICDMW was held on virtually on 7th December 2021, followed by the IEEE ICDM 2021 conference. This year, we receive 24 proposals. After the workshop proposal review, paper review and merging of workshops, the final IEEE ICDMW 2021 program consisted of 18 workshops, including 6 full-day and 12 half-day workshops. Overall, IEEE ICDMW received 266 papers while 131 papers (i.e. 50%) being presented in the final program. All the papers have been through a peer-review process to ensure high-quality papers being presented in the ICDMW 2021 proceedings.

A wide range of research topics in data mining have been included in IEEE ICDMW 2021, covering both theoretical research and real-world applications. Among the more traditional areas of data mining, ICDMW 2021 includes recommendation systems, information retrieval, information extraction, natural text, high-dimensional data mining, multi-source data, incremental learning, continual learning, sentiment analysis, deep learning, clustering, concept drift, novelty detection, feature engineering, time-series data mining, and data optimization. Furthermore, emerging research areas include data mining for bioinformatics, healthcare, engineering service, anomaly analytics, utility-driven data mining and learning, data in legal domain, spatial and spatio-temporal data mining, evolutionary computation based data mining, and cybersecurity.

It is no doubt that the success of ICDMW 2021 is a collective effort from many colleagues. Our particular thanks to all the workshop organizers for their delegate time and effort in preparing, and submitting their workshop proposals, managing their technical program committees, paper review, presentations, and invited talks from experts of their domains. Last but not least, we would like to thank all the reviewers for their thorough evaluations and constructive feedback to the papers, which ensure the high quality of the papers appeared in ICDMW 2021 and the support authors receive in further improving their work.

The PhD forum is featured by IEEE ICDM and continued to this year for the 10th year. As a tradition, the PhD forum aims to provide more opportunities for research students to present their works and communicate with their peers and senior researchers in the areas related to data mining. This forum particularly beneficial for PhD students in the early stage of their doctoral study, and MSc students who are planning to PhD study. The 2021 edition of the PhD forum includes 8 high-quality papers, spanning various topics of data mining and work in related fields including deep learning, crowdsourcing, statistical learning, anomaly detection, recommendation, and

reinforcement learning. Many thanks to the program committee members of the PhD forum this year, who helped with the selection of these high quality papers and who provided high quality review comments, under the guidance of the PhD forum Chairs Lisi Chen, Meng-Fen Chiang and Sibylle Hess.

The Workshop Chairs of IEEE ICDM 2021 would like to thank the Steering Committee Chair, Xindong Wu, and the General Chairs, Yun Sing Koh and Dacheng Tao, for their precious advices and leadership. We would also like to thank the Online Experience/Virtual Chair, Heitor Gomes, for the arrangement of the Workshops day. Furthermore, we also would like to thank the Program Chairs, Pauli Miettinen and James Bailey, for their substantial contribution to the organization of the conference, and the Finance Chair, Gillian Dobbie, for managing the finance and registration of the conference. Furthermore, special thanks to the Publicity Chair, Diana Benavides Prado, for her great effort on managing the conference web and social medias, and her valuable help with the call for workshop proposals and call for papers. Last but not least we would like to thank Giuseppe Di Fatta for passing their experience as Workshop Chairs from the past year edition. Thank you all for your great effort in making IEEE ICDMW 2021 a success!

ICDM 2021 Workshop Chairs:

Bing Xue, *Victoria University of Wellington, New Zealand*

Mykola Pechenizkiy, *Eindhoven University of Technology, Netherlands*

Organizing Committee

General Chairs

Yun Sing Koh, *The University of Auckland, New Zealand*

Dacheng Tao, *The University of Sydney, Australia*

Steering Committee Chair

Xindong Wu, *Hefei University of Technology and Mininglamp Academy of Sciences, China*

Program Chairs

Pauli Miettinen, *University of Eastern Finland, Finland*

James Bailey, *The University of Melbourne, Australia*

Local Arrangement Chair

Joerg Wicker, *The University of Auckland, New Zealand*

Finance Chair

Gillian Dobbie, *The University of Auckland, New Zealand*

Sponsorship Chairs

Kaiqi Zhao, *The University of Auckland, New Zealand*

Philippe Fournier-Viger, *Harbin Institute of Technology, China*

Eva García-Martín, *Ekkono Solutions, Sweden*

Tutorial Chairs

Wei Liu, *University of Technology Sydney, Australia*

Katharina Morik, *TU Dortmund, Germany*

Workshop Chairs

Bing Xue, *Victoria University of Wellington, New Zealand*

Mykola Pechenizkiy, *Eindhoven University of Technology, Netherlands*

Panel Chairs

Chenliang Li, *Wuhan University, China*

Michael Witbrock, *The University of Auckland, New Zealand*

PhD Forum Chairs

Lisi Chen, *Inception Institute of Artificial Intelligence (IIAI), UAE*

Meng-Fen Chiang, *The University of Auckland, New Zealand*

Sibylle Hess, *Eindhoven University of Technology, Netherlands*

Diversity and Inclusion Chair

Richi Nayak, *Queensland University of Technology, Australia*

Online Experience/Virtual Chair

Heitor Gomes, *University of Waikato, New Zealand*

Newcomers Chairs

Huan Liu, *Arizona State University, USA*

Katerina Taskova, *The University of Auckland, New Zealand*

Job Matching Chairs

Albert Bifet, *University of Waikato, New Zealand*

Lin Liu, *University of South Australia, Australia*

Reproducibility Chairs

Dragi Kocev, *Jožef Stefan Institute, Slovenia*

Jacob Montiel, *University of Waikato, New Zealand*

Award Committee Chair

Xia Ning, *The Ohio State University, USA*

Publicity Chair

Diana Benavides Prado, *The University of Auckland, New Zealand*

Area Chairs and Program Committee

Area Chairs

Charu Aggarwal, *IBM T. J. Watson Research Center*
Leman Akoglu, *Carnegie Mellon University*
Elena Baralis, *Politecnico di Torino*
Bettina Berendt, *K.U. Leuven*
Elisa Bertino, *Purdue University*
Francesco Bonchi, *ISI Foundation*
Wray Buntine, *Monash University*
Ricardo Campello, *University of Newcastle*
Jeffrey Chan, *RMIT University*
Ming-Syan Chen, *National Taiwan University*
Chris Clifton, *Purdue University at West Lafayette*
Diane Cook, *Washington State University*
Gianmarco De Francisci Morales, *ISI Foundation*
Carlotta Domeniconi, *George Mason University*
Guozhu Dong, *Wright State University*
Eibe Frank, *University of Waikato*
Esther Galbrun, *University of Eastern Finland*
Aristides Gionis, *KTH*
Quanquan Gu, *University of California, Los Angeles*
Michael E. Houle, *National Institute of Informatics*
George Karypis, *University of Minnesota*
Hisashi Kashima, *Kyoto University*
Eamonn Keogh, *University of California, Riverside*
Latifur Khan, *University of Texas at Dallas*
Jiuyong Li, *University of South Australia*
Jessica Lin, *George Mason University*
Huan Liu, *ASU*
Wei Liu, *University of Technology Sydney*
Wagner Meira Jr., *Universidade Federal de Minas Gerais*
Salvatore Orlando, *Ca' Foscari University of Venice*
Evangelos Papalexakis, *UC Riverside*
Panagiotis Papapetrou, *Stockholm University*
Srinivasan Parthasarathy, *The Ohio State University*
Jian Pei Simon, *Fraser University*
Dinh Phung, *Monash University*
Nico Piatkowski, *TU Dortmund*
Naren Ramakrishnan, *Virginia Tech*
Huzefa Rangwala, *George Mason University*
Matteo Riondato, *Amherst College*
Joerg Sander, *University of Alberta*
Kyuseok Shim, *Seoul National University*
Myra Spiliopoulou, *University of Magdeburg*
Karthik Subbian, *Amazon*
Nikolaj Tatti, *University of Helsinki*

Kai Ming Ting, *Nanjing University*
Panayiotis Tsaparas, *University of Ioannina*
Vincent Tseng, *National Yang Ming Chiao Tung University*
Matthijs van Leeuwen, *Leiden Institute of Advanced Computer*
Jilles Vreeken, *CISPA Helmholtz Center for Information Security*
Jia Wu, *Macquarie University*
Hui Xiong, *Rutgers University*
Min-Ling Zhang, *Southeast University*
Weinan Zhang, *Shanghai Jiao Tong University*
Xingquan Zhu, *Florida Atlantic University*
Fuzhen Zhuang, *Institute of Artificial Intelligence, Beihang University*

Program Committee

Pedro Henriques Abreu, *University of Coimbra*
Florian Adriaens, *KTH*
Ankit Agrawal, *Northwestern University*
Aijun An, *York University*
Shin Ando, *Tokyo University of Science*
Xiang Ao, *Institute of Computing Technology, CAS*
Annalisa Appice, *Universita' degli Studi di Bari Aldo Moro*
Renato Assunção, *Universidade Federal de Minas Gerais*
Martin Atzmueller, *Osnabrück University*
Kun Bai, *Tencent*
Roberto Bayardo, *Google*
Sadok Ben Yahia, *Universita de tunis ElManar, LIPAH*
Andras A. Benczur, *Hungarian Academy of Sciences*
Indrajit Bhattacharya, *TCS Research Lab Kolkata*
Siddhartha Bhattacharyya, *University of Illinois, Chicago*
Christian Boehm, *University of Munich*
Klemens Böhm, *Karlsruhe Institute of Technology (KIT)*
Ludovico Boratto, *EURECAT, Centre Tecnològic de Catalunya*
Robin Burke, *University of Colorado*
Rajmonda Caceres, *MIT Lincoln Laboratory*
Luca Cagliero, *Politecnico di Torino*
Jian Cao, *Shanghai Jiao Tong University*
Michelangelo Ceci, *University of Bari*
Loïc Cerf, *Universidade Federal de Minas Gerais*
Tania Cerquitelli, *Politecnico di Torino*
Philip Chan, *Florida Institute of Technology*
Kevin C. Chang, *University of Illinois at Urbana Champaign*
Mandar Chaudhary, *eBay*
Abon Chaudhuri, *Walmart Labs*
Arbee L.P. Chen, *Asia University*
Huiyuan Chen, *Visa Research*
Jinjun Chen, *Swinburne University of Technology*
Lei Chen, *Hong Kong University of Science and Technology*

Lingwei Chen, *Pennsylvania State University*
Ping Chen, *University of Massachusetts Boston*
Rui Chen, *Samsung Research America, Mountain View*
Tao Chen, *China University of Geosciences*
Xi Chen, *Ghent University*
Zhenzhong Chen, *Wuhan University*
Zhiyuan Chen, *University of Maryland Baltimore County*
Xueqi Cheng, *Institute of Computing Technology, Chinese Academy of Science*
Sriram P. Chockalingam, *Georgia Institute of Technology*
Byron Choi, *Hong Kong Baptist University*
Fu-Lai Chung, *Hong Kong Polytechnic University*
Alfredo Cuzzocrea, *ICAR-CNR and University of Calabria*
Ernest Damiani, *Khalifa University of Science and Technology*
Xuan-Hong Dang, *IBM Research*
Anirban Dasgupta, *Indian Institute of Technology Gandhinagar*
Pieter Delobelle, *KU Leuven*
Shuiguang Deng, *College of Computer Science and Technology, Zhejiang University*
Anne Denton, *North Dakota State University*
Djellel Difallah, *New York University*
Wei Ding, *University of Massachusetts Boston*
Ying Ding, *University of Texas at Austin*
Nemanja Djuric, *Aurora Innovation*
Josep Domingo-Ferrer, *Universitat Rovira i Virgili*
Yanni Dong, *China University of Geosciences*
Dejing Dou, *University of Oregon*
Bo Du, *Wuhan University*
Lei Duan, *Sichuan University*
Saso Dzeroski, *Jozef Stefan Institute*
Christoph F. Eick, *University of Houston*
Mohammad El-Hajj, *Grant Macewan University*
Tapio Elomaa, *Tampere University*
Yujie Fan, *Case Western Reserve University*
Hua (Julia) Fang, *Computational Statistics and Data Science (CSDS) lab*
wei Feng, *Hulu*
Yunlong Feng, *State University of New York at Albany*
Stefano Ferilli, *University of Bari Aldo Moro*
Edouard Fouché, *Karlsruhe Institute of Technology (KIT)*
Yanjie Fu, *University of Central Florida*
Yun Fu, *Northeastern University*
Benjamin C. M. Fung, *McGill University*
Johannes Fürnkranz, *JKU Linz*
Matjaz Gams, *Jozef Stefan Institute*
Aryya Gangopadhyay, *University of Maryland, Baltimore County*
Byron Gao, *Texas State University*
Yang Gao, *Beijing Institute of Technology*
Yue Gao, *Tsinghua University*
Rainer Gemulla, *University of Mannheim*

Aris Gkoulalas-Divanis, *IBM Watson Health*
Chen Gong, *Nanjing University of Science and Technology*
Mingming Gong, *University of Melbourne*
Ananth Grama, *Purdue University*
Le Gruenwald, *University of Oklahoma*
Ziyu Guan, *Xidian University*
Francesco Gullo, *UniCredit, R&D Dept.*
Dimitrios Gunopulos, *University of Athens*
Ting Guo, *University of Technology, Sydney*
Maria Halkidi, *University of Piraeus*
Lawrence Hall, *University of South Florida*
Sibylle Hess, *TU Eindhoven*
Shen-Shyang Ho, *Rowan University*
Mahmud Shahriar Hossain, *University of Texas at El Paso*
Andreas Hotho, *Data Science Chair, University of Würzburg*
Chun-Nan Hsu, *UC San Diego*
Haibo Hu, *Hong Kong Polytechnic University*
Minqing Hu, *MySpace*
Chao Huang, *University of Notre Dame*
Heng Huang, *University of Pittsburgh; JD Finance America Corporation*
Jianbin Huang, *Xidian University*
Jin Huang, *Tsinghua University*
Nina Hubig, *Clemson University*
Tsuyoshi Ide, *IBM T. J. Watson Research Center*
Dino Ienco, *UMR Tetis, IRSTEA*
Nitin Indurkha, *AI Data-Miner LLC*
Stratis Ioannidis, *Northeastern University*
Vandana P. Janeja, *University of Maryland, Baltimore County*
Ling Jian, *China University of Petroleum*
Lili Jiang, *Umeå University*
Zhe Jiang, *University of Alabama*
Bo Jin, *Dalian University of Technology*
Huidong Jin, *CSIRO*
Ruoming Jin, *Kent State University*
Panos Kalnis, *King Abdullah University of Science and Technology*
Vana Kalogeraki, *Athens University of Economics and Business*
Konstantinos Kalpakis, *UMBC*
Toshihiro Kamishima, *National Institute of Advanced Industrial Science and Technology (AIST)*
Andrey Kan, *University of Adelaide*
Bhargav Kanagal, *Google*
Zhao Kang, *University of Electronic Science and Technology of China*
Hung-Yu Kao, *National Cheng Kung University*
Kamal Karlapalem, *International Institute of Information Technology, Hyderabad*
Saurav Karmakar, *GreyKarma Technologies*
Panagiotis Karras, *Aarhus University*
Sang Wook Kim, *Hanyang University, Seoul, Korea*
Deguang Kong, *Yahoo Research*

Youyong Kong, *Southeast University*
Olivera Kotevska, *Oak Ridge National Laboratory*
Stefan Kramer, *Johannes Gutenberg University*
Nils M. Kriege, *TU Dortmund University*
Peer Kröger, *Kiel University*
Hye-Chung Kum, *Texas A&M*
Jean-Charles Lamirel, *TALARIS- LORIA*
Long Lan, *National University of Defense Technology*
Anne Laurent, *Univ. Montpellier 2, LIRMM*
Hady Lauw, *Singapore Management University*
Aleksandar Lazarevic, *Stanley Black & Decker*
Jae-Gil Lee, *KAIST*
Roy Ka-Wei Lee, *University of Saskatchewan*
Mark Levene, *Birkbeck, University of London*
Cheng-Te Li, *Institute of Data Science, National Cheng Kung University*
Gang Li, *Deakin University*
Hui Li, *Xidian University*
Jianxin Li, *Deakin University*
Jianxin Li, *Beihang University*
Ming Li, *Nanjing University*
Peipei Li, *Hefei University of Technology*
Shuai Li, *Cambridge University*
Xue Li, *University of Queensland*
Xuelong Li, *Northwestern Polytechnical University*
Yanhua Li, *Worcester Polytechnic Institute*
Yuefeng Li, *Queensland University of Technology*
Yun Li, *Nanjing University of Posts and Telecommunications*
Qing Liao, *Harbin Institute of Technology (Shenzhen)*
Chien-Liang Liu, *National Chiao Tung University*
Chuanren Liu, *Drexel University*
Guannan Liu, *Beihang University*
Hongfu Liu, *Brandeis University*
Hongyan Liu, *Tsinghua University*
Jialu Liu, *Google Research*
Jinfei Liu, *Emory University*
Junqiang Liu, *Zhejiang Gongshang University*
Li Liu, *Chongqing University*
Shenghua Liu, *Institute of Computing Technology, Chinese Academy of Science*
Tongliang Liu, *The University of Sydney*
Weifeng Liu, *China University of Petroleum (East China)*
Zhonghua Liu, *Henan University of Science and Technology*
Guodong Long, *University of Technology Sydney*
Yin Lou, *Ant Group*
Grigorios Loukides, *King's College London*
Chang-Tien Lu, *Virginia Tech*
Haibing Lu, *Santa Clara University*
Simone Ludwig, *North Dakota State University*

Fulin Luo, *Wuhan University*
Jiebo Luo, *University of Rochester*
Ping Luo, *Institute of Computing Technology, Chinese Academy of Sciences*
Yong Luo, *Nanyang Technological University*
Shuai Ma, *Beihang University*
Xingjun Ma, *Deakin University*
Richard Maclin, *University of Minnesota-Duluth*
Son T. Mai, *Queens University Belfast*
Arun Maiya, *Institute for Defense Analyses*
Bradley Malin, *Vanderbilt University*
Hiroshi Mamitsuka, *Kyoto University / Aalto University*
Giuseppe Manco, *ICAR-CNR*
Panagiotis Mandros, *CISPA Helmholtz Center for Information Security*
Alexandru Mara, *Ghent University*
Francesco Marcelloni, *University of Pisa*
Florent Masegla, *INRIA*
Gabriele Mencagli, *University of Pisa*
Weiyi Meng, *Binghamton University*
Matej Mihelcic, *University of Zagreb*
Shin-ichi Minato, *Kyoto University*
Pasquale Minervini, *University College London*
Anna Monreale, *Pisa*
Bongki Moon, *Seoul National University*
Sebastian Moreno, *Universidad Adolfo Ibanez*
Shinichi Morishita, *University of Tokyo*
Abdullah Mueen, *University of New Mexico*
Tsuyoshi Murata, *Tokyo Institute of Technology*
Mirco Nanni, *KDDLab, ISTI-CNR Pisa*
Stefan Neumann, *KTH Royal Institute of Technology*
Wilfred NG, *HKUST*
Cam-Tu Nguyen, *Nanjing University*
Feiping Nie, *Northwestern Polytechnical University*
Xia Ning, *The Ohio State University*
Tim Oates, *University of Maryland, Baltimore County*
Riccardo Ortale, *ICAR-CNR*
Amichai Painsky, *MIT*
Shirui Pan, *Monash University*
Shalini Pandey, *University of Minnesota*
Guansong Pang, *University of Adelaide*
Laurence Park, *Western Sydney University*
Noseong Park, *George Mason University*
Hao Peng, *Beihang University*
Wen-Chih Peng, *National Chiao Tung University*
Ruggero G. Pensa, *University of Torino*
Jean-Marc Petit, *Universite de Lyon, CNRS, INSA-Lyon*
NhatHai Phan, *NJIT*
Gianvito Pio, *University of Bari Aldo Moro*

Claudia Plant, *University of Vienna*
Marc Plantevit, *Universite de Lyon*
Ronaldo Prati, *Universidade Federal do ABC – UFABC*
Jianzong Qi, *University of Melbourne*
Liu Qi, *University of Science and Technology of China*
Buyue Qian, *Xi'an Jiaotong University*
Qi Qian, *Alibaba Group*
Tieyun Qian, *Wuhan University*
Maoying Qiao, *the Commonwealth Scientific and Industrial Research Organisation*
Mu Qiao, *IBM Almaden Research Center*
Lu Qin, *University of Technology, Sydney*
Milos Radavanovic, *University of Novi Sad*
Vijay Raghavan, *University of Louisiana at Lafayette*
Sutharshan Rajasegarar, *Deakin University*
Ganesh Ramesh, *SAP Ariba*
Jan Ramon, *INRIA*
Rajeev Rastogi, *Amazon*
Chotirat Ann Ratanamahatana, *Chulalongkorn University*
Jan Rauch, *Prague University of Economics and Business*
Matthias Renz, *Universitaet Kiel*
Justine Reynaud, *Uni Caen*
Senjuti Basu Roy, *New Jersey Institute of Technology*
Carolina Ruiz, *Worcester Polytechnic Institute*
Lorenza Saitta, *Universita del Piemonte Orientale*
David Sánchez, *Universitat Rovira i Virgili*
Maria Luisa Sapino, *U. Torino*
Lars Schmidt-Thieme, *University of Hildesheim*
Matthias Schubert, *Ludwig-Maximilians-Universität München*
Friedhelm Schwenker, *University of Ulm*
Kristen Scott, *KU Leuven*
Pavel Senin, *Los Alamos National Laboratory*
Neil Shah, *Snap Inc.*
Huasong Shan, *JD.com*
Fanhua Shang, *Xidian University*
Junming Shao, *University of Electronic Science and Technology of China*
Ming Shao, *University of Massachusetts Dartmouth*
Huawei Shen, *Institute of Computing Technology, Chinese Academy of Science*
Jialie Shen, *Queen's University Belfast*
Wei Shen, *Nankai University*
Xiaobo Shen, *Nanyang Technological University*
Kijung Shin, *Korea Advanced Institute of Science and Technology*
Mei-Ling Shyu, *University of Miami*
Dan Simovici, *University of Massachusetts at Boston*
Lisa Singh, *Georgetown University*
Krishnamoorthy Sivakumar, *Washington State University*
Andrzej Skowron, *University of Warsaw and Polish Academy of Sciences*
Dongjin Song, *University of Connecticut*

Guojie Song, *Key Laboratory of Machine Perception, Peking University*
Sucheta Soundarajan, *Syracuse University*
Anna Squicciarini, *The Pennsylvania State University*
Jaideep Srivastava, *University of Minnesota*
Giovanni Stilo, *University of L'Aquila*
Ja-Hwung Su, *Cheng Shiu University*
Guangzhong Sun, *University of Science and Technology of China*
Liang Sun, *OPERA Solutions*
Xin Sun, *Oceans University of China*
Andrea Tagarelli, *University of Calabria*
Chih-Hua Tai, *National Taipei University*
Atsuhiko Takasu, *National Institute of Informatics*
Pang-Ning Tan, *Michigan State University*
David Taniar, *Monash University*
Jianrong Tao, *Netease Research Institute*
Maguelonne Teisseire, *TETIS – Irstea*
Masahiro Terabe, *Chugai Pharmaceutical Co., Ltd.*
Lini Thomas, *IITH*
Bin Tong, *Central Research Laboratory, Hitachi, Ltd.*
Yongxin Tong, *Beihang University*
Vicenc Torra, *University of Skovde*
Shusaku Tsumoto, *Shimane University*
Antti Ukkonen, *Speechly*
Fabio Vandin, *University of Padova*
Ranga Raju Vatsavai, *NC State University*
Chang-Dong Wang, *Sun Yat-sen University*
Dianhui Wang, *La Trobe University*
Fei Wang, *Cornell University*
Honggang Wang, *University of Massachusetts*
Jason T. L. Wang, *New Jersey Institute of Technology*
Jin Wang, *Megagon Labs*
Jingyuan Wang, *Beihang University*
Ke Wang, *Simon Fraser University*
Sen Wang, *University of Queensland*
Senzhang Wang, *Nanjing University of Aeronautics and Astronautics*
Suhang Wang, *The Pennsylvania State University*
Wei Wang, *Fudan University*
Xiao Wang, *Beijing University of Posts and Telecommunications*
Yisen Wang, *Peking University*
Zhi Wang, *Tsinghua University, China*
Takashi Washio, *The Institute of Scientific and Industrial Research, Osaka University*
Ingmar Weber, *Qatar Computing Research Institute*
Zhi Wei, *New Jersey Institute of Technology*
Tim Weninger, *University of Notre Dame*
Raymond Chi Wing Wong, *Hong Kong University of Science and Technology*
Chuan Wu, *The University of Hong Kong*
Gongqing Wu, *Hefei University of Technology*

Runze Wu, *Netease Research Institute*
Wensheng Wu, *University of Southern California*
Yinghui Wu, *Case Western Reserve University*
Lingyun Xiang, *Changsha University of Science and Technology*
Keli Xiao, *Stony Brook University*
Xiaokui Xiao, *National University of Singapore*
Yanghua Xiao, *Fudan University*
Sihong Xie, *Lehigh University*
Yun Xiong, *Fudan University*
Chang Xu, *The University of Sydney*
Guandong Xu, *University of Technology Sydney*
Yan Xu, *SAS Institute Inc.*
Zenglin Xu, *University of Electronic Science and Technology of China*
Zhao Xu, *NEC Laboratories Europe*
Bo Yang, *Jilin University*
Hui Yang, *The Ohio State University*
Lianghuai Yang, *Zhejiang University of Technology*
Lina Yao, *University of New South Wales*
Yuan Yao, *State Key Laboratory for Novel Software Technology, Nanjing University*
Yanfang (Fanny) Ye, *Case Western Reserve University*
Hongzhi Yin, *The University of Queensland*
Jianhua Yin, *Shandong University*
Jie Yin, *The University of Sydney*
Jin Soung Yoo, *Purdue University Fort Wayne*
Dantong Yu, *Brookhaven National Laboratory, Stony Brook University*
Guoxian Yu, *Shandong University*
Hwanjoyu Yu, *Pohang University of Science and Technology*
Jeffrey Xu Yu, *The Chinese University of Hong Kong*
Jun Yu, *Hangzhou Dianzi University*
Kui Yu, *Hefei University of Technology*
Philip S. Yu, *University of Illinois at Chicago*
Qiangqiang Yuan, *Wuhan University*
Reza Zafarani, *Syracuse University*
Nayyar Zaidi, *Monash University*
Chengxi Zang, *Cornell University*
Petros Zerfos, *IBM T. J. Watson Research Centre*
Aidong Zhang, *University of Virginia*
Chao Zhang, *Dalian Univ of Technology*
Chengqi Zhang, *University of Technology Sydney*
Chuxu Zhang, *Brandeis University*
Daokun Zhang, *Monash University*
Daoqiang Zhang, *Nanjing University of Aeronautics and Astronautics*
Jianfei Zhang, *University of Alberta*
Jiawei Zhang, *University of Illinois at Chicago*
Jing Zhang, *Nanjing University of Science and Technology*
Kun Zhang, *Xavier Univ. of Louisiana*
Kun Zhang, *CMU*

Lefei Zhang, *Wuhan University*
Ping Zhang, *The Ohio State University*
Shichao Zhang, *Guangxi Normal University*
Wei Zhang, *University of Central Florida*
Wei Zhang, *East China Normal University*
Wenlu Zhang, *Old Dominion University*
Xiang Zhang, *The Pennsylvania State University*
Xiang Zhang, *National University of Defense Technology*
Xiangliang Zhang, *King Abdullah University of Science and Technology*
Xuyun Zhang, *Macquarie University*
Ya Zhang, *Shanghai Jiao Tong University*
Yan Zhang, *Peking University*
Yanchun Zhang, *Victoria University*
Zhao Zhang, *Hefei University of Technology*
Zheng Zhang, *Harbin Institute of Technology (Shenzhen)*
Chen Zhao, *University of Texas, Dallas*
Liang Zhao, *George Mason University*
Peixiang Zhao, *Florida State University*
Ying Zhao, *Tsinghua University*
Yun Zhao, *University of California, Santa Barbara*
Zhou Zhao, *Zhejiang University*
Guanjie Zheng, *Shanghai Jiaotong University*
Aoying Zhou, *East China Normal University, Shanghai, China*
Chuan Zhou, *Academy of Mathematics and Systems Science*
Dawei Zhou, *University of Illinois at Urbana-Champaign*
Guanqun Zhou, *Stanford University*
Jiayu Zhou, *Michigan State University*
Xiaofang Zhou, *The University of Queensland, Brisbane*
Xun Zhou, *University of Iowa*
Yang Zhou, *Auburn University*
Feida Zhu, *Singapore Management University*
Hengshu Zhu, *Baidu Research-Big Data Lab, Baidu Inc.*
Blaz Zupan, *University of Ljubljana*
Andreas Zufle, *George Mason University*