2022 IEEE International Conference on Data Mining Workshops (ICDMW 2022)

Orlando, Florida, USA 28 November - 1 December 2022

Pages 1-625



IEEE Catalog Number: CFP2256B-POD **ISBN:**

979-8-3503-4610-7

Copyright © 2022 by the Institute of Electrical and Electronics Engineers, Inc. All Rights Reserved

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

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

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

CFP2256B-POD
979-8-3503-4610-7
979-8-3503-4609-1
2375-9232

Additional Copies of This Publication Are Available From:

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



2022 IEEE International Conference on Data Mining Workshops (ICDMW) **ICDMW 2022**

Table of Contents

Message from the ICDM 2022 General Chairs	xxv
Message from the ICDM 2022 Program Chairs	. xxvii
Message from the Workshops Chairs	xxix
Organizing Committee	xxxi

UNIT: Urban Internet-of-Things Intelligence

A Multi-source Information Learning Framework for Airbnb Price Prediction Lu Jiang (Northeast Normal University), Yuanhan Li (Northeast Normal University), Na Luo (Northeast Normal University), Jianan Wang (Northeast Normal University), and Qiao Ning (Dalian Maritime University)	L
A Comparison of Ambulance Redeployment Systems on Real-World Data	3
Streaming Traffic Flow Prediction Based on Continuous Reinforcement Learning	5

DLC: Deep Learning and Clustering

Augmenting Graph Convolution with Distance Preserving Embedding for Improved Learning 23 Guojing Cong (Oakridge National Laboratory), Seung-Hwan Lim (Oakridge National Laboratory), and Steven Young (Oakridge National Laboratory)
Emerging Properties from Bayesian Non-Parametric for Multiple Clustering: Application for
Multi-view Image Dataset
Reda Khoufache (Université Sorbonne Paris Nord, France), Mohamed
Djallel Dilmi (Université Sorbonne Paris Nord, France), Hanene Azzag
(Université Sorbonne Paris Nord, France), Etienne Goffinet (Technology
Innovation Institute, United Arab Emirates, France), and Mustapha
Lebbah (Université de Versailles, Université Paris-Saclay, France)

Unsupervised Deep One-Class Classification with Adaptive Threshold Based on Training Dynamics
cSmartML-Glassbox: Increasing Transparency and Controllability in Automated Clustering
Joint Debiased Representation Learning and Imbalanced Data Clustering
Joint Low-Rank and Orthogonal Deep Multi-view Subspace Clustering Based on Local Fusion 63 Guixiang Wang (Huzhou University, China), Hongwei Yin (Huzhou University, China), Wenjun Hu (Huzhou University, China), Yong Liu (Zhejiang University, China), and Ruiqin Wang (Huzhou University, China)

DMS: Data Mining for Service

User Value in Modern Payment Platforms: A Graph Approach
An Application of Customer Embedding for Clustering
Simplifying Process Navigations – Divide and Rule way
 Scene and Texture Based Feature Set for DeepFake Video Detection

Using Spatial Data and Cluster Analysis to Automatically Detect non-Trivial Relationships Between Environmental Transgressors
José Alberto Sousa Torres (University of Brasilia, Brazil), Paulo
Henrique dos Santos (Advocacy-general of the Union of Brazil, Brazil),
Daniel Alves da Silva (University of Brasília, Brazil), Carlos Eduardo
Lacerda Veiga (Advocacy-general of the Union of Brazil, Brazil),
Márcio Bastos Medeiros (Advocacy-general of the Union of Brazil,
Brazil), Guilherme Fay Vergara (University of Brasília, Brazil), Fábio
Lúcio Lopes Mendonça (University of Brasilia, Brazil), and Rafael
Timóteo de Sousa Júnior (University of Brasilia, Brazil)
Identify Malfunctions and Their Possible Causes Using Rules, Application to Process Mining 105 Benoit Vuillemin (University of Technology of Troyes, France) and Frédéric Bertrand (University of Technology of Troyes, France)
Weight-Training Ensemble Model for Stock Price Forecast

ISAW: International Sustainable AI Workshop

Domain Adaptation Through Cluster Integration and Correlation Vishnu Manasa Devagiri (Blekinge Institute of Technology, Sweden), Veselka Boeva (Blekinge Institute of Technology, Sweden), and Shahrooz Abghari (Blekinge Institute of Technology, Sweden)	119
 Data-Driven Usage Profiling and Anomaly Detection in Support of Sustainable Machining Processes Fabian Fingerhut (EluciDATA Lab of Sirris, Belgium), Chaitra Harsha (EluciDATA Lab of Sirris, Belgium), Amirmohammad Eghbalian (EluciDATA Lab of Sirris, Belgium), Tom Jacobs (Precision Manufacturing, Sirris, Belgium), Mahdi Tabassian (EluciDATA Lab of Sirris, Belgium), Robbert Verbeke (EluciDATA Lab of Sirris, Belgium), and Elena Tsiporkova (EluciDATA Lab of Sirris, Belgium) 	127
Equal Confusion Fairness: Measuring Group-Based Disparities in Automated Decision Systems Furkan Gursoy (University of Houston, USA) and Ioannis A. Kakadiaris (University of Houston, USA)	137
 FastFlow: AI for Fast Urban Wind Velocity Prediction	147
Empirical Analysis of Fairness-Aware Data Segmentation Seiji Okura (Fujitsu Limited, Japan) and Takao Mohri (Fujitsu Limited, Japan)	155

OEDM: Optimization Based Techniques for Emerging Data Mining Problems

DeepDive: Deep Latent Factor Model for Enhancing Diversity in Recommender Systems
Understanding Concept Identification as Consistent Data Clustering Across Multiple Feature Spaces
Felix Lanfermann (Honda Research Institute Europe, Germany), Sebastian Schmitt (Honda Research Institute Europe, Germany), and Patricia Wollstadt (Honda Research Institute Europe, Germany)
 FLAG: Stock Movement Prediction via Fusing Logic and Semantic Graphs of Financial News 190 Angela Li (Peking University), Jiduan Liu (Wangxuan Institute of Computer Technology, Peking University; Center for Data Science, AAIS, Peking University), Yuyong Li (Peking University), and Fan Meng (Peking University)
Solving Non-Linear Optimization Problem in Engineering by Model-Informed Generative Adversarial Network (MI-GAN) Yuxuan Li (Oklahoma State University, USA), Chaoyue Zhao (University of Washington, USA), and Chenang Liu (Oklahoma State University, USA)
Towards Practical Explainability with Cluster Descriptors206Xiaoyuan Liu (Fujitsu Research of America, Inc., USA), Ilya Tyagin206(University of Delaware, USA), Hayato Ushijima-Mwesigwa (FujitsuResearch of America, Inc., USA), Indradeep Ghosh (Fujitsu Research of America, Inc., USA), and Ilya Safro (University of Delaware, USA)
MetaSieve: Performance vs. Complexity Sieve for Time Series Forecasting

A Study of Automatic Speech Recognition in Portuguese by the Brazilian General Attorney of the Union	226
Rodrigo Fay Vergara (Cyber Security INCT Unit 6, Laboratory for	220
Decision-Making Technologies (LATITUDE), University of Brasília	
(UnB)), Paulo Henrique dos Santos (Advocacy-general of the Union of	
Brazil), Guilherme Fay Vergara (University of Brasília (UnB)), Fábio	
Lúcio Lopes Mendonça (University of Brasília (UnB)), Carlos Eduardo	
Lacerda Veiga (Advocacy-general of the Union of Brazil), Bruno Justino	
Garcia Praciano (University of Brasília (UnB)), Daniel Alves da Silva	
(University of Brasília (UnB)), and Rafael Timóteo de Sousa Júnior	
(University of Brasília (UnB))	
SV-Learn: Learning Matrix Singular Values with Neural Networks	232
Derek Xu (Ayala High School, USA), William Shiao (UC Riverside, USA),	
Jia Chen (UC Riverside, USA), and Evangelos Papalexakis (UC Riverside,	
USA)	

ARIAL: AI for Aging, Rehabilitation, and Intelligent Assisted Living

 MAISON — Multimodal AI-Based Sensor Platform for Older Individuals	8
Domain-Specific Deep Learning Feature Extractor for Diabetic Foot Ulcer Detection	3
Reza Basiri (Institute of Biomedical Engineering, University of	
Toronto, Canada; KITE, Toronto Rehabilitation Institute, University	
Health Network, Canada), Milos R. Popovic (Institute of Biomedical	
Engineering, University of Toronto, Canada; KITE, Toronto	
Rehabilitation Institute, University Health Network, Canada), and	
Shehroz S. Khan (Institute of Biomedical Engineering, University of	
Toronto, Canada; KITE, Toronto Rehabilitation Institute, University	
Health Network, Canada)	
Detection of Mild Cognitive Impairment from Quantitative Analysis of Timed Up and Go (TUG)24 Mahmoud Seifallahi (Florida Atlantic University Boca Raton, USA),	8
James E. Galvin (University of Miami, USA), and Behnaz Ghoraani	
(Florida Atlantic University Boca Raton, USA)	

DMLS: Data Mining in Learning Science

Diagonally Colorized iVAT Images for Labeled Data	254
Elizabeth Hathaway (University of Tennessee, USA) and Richard Hathaway	
(Hartwell Research Institute, ŬŠĂ)	

Students Temporal Profiling and e-Learning Resources Recommendation Based on NLP's Terms Extraction
Sentence-BERT Distinguishes Good and Bad Essays in Cross-Prompt Automated Essay Scoring 274 Toru Sasaki (Rikkyo University, Japan) and Tomonari Masada (Rikkyo University, Japan)
Deep Knowledge Tracing with Learning Curves
Online Deep Knowledge Tracing292Wenxin Zhang (Northwestern Polytechnical University; Key Laboratory of Data Storage and Management of MIIT, China), Yupei Zhang (Northwestern Polytechnical University; Key Laboratory of Data Storage and Management of MIIT, China), Shuhui Liu (Northwestern Polytechnical University; Key Laboratory of Data Storage and Management of MIIT, China), and Xuequn Shang (Northwestern Polytechnical University; Key Laboratory of Data Storage and Management of MIIT, China)
 Programming Knowledge Tracing: A Comprehensive Dataset and A New Model

DMBIH: Data Mining in Biomedical Informatics and Healthcare

AWS-EP: A Multi-Task Prediction Approach for MBTI-Big5 Personality Tests Fahed Elourajini (université de montréal, Canada) and Esma Aïmeur (université de montréal, Canada)	. 308
HMM-Boost: Improved Time Series State Prediction via Supervised Hidden Markov Models: Case	
Studies in Epileptic Seizure and Complex Care Management	316
Georgios Mavroudeas (Rensselaer Polytechnic Institute, USA), Malik	
Magdon-Ismail (Rensselaer Polytechnic Institute, USA), Xiao Shou	
(Rensselaer Polytechnic Institute, USA), and Kristin Bennett	
(Rensselaer Polytechnic Institute, USA)	

 Model Personalization with Static and Dynamic Patients' Data	324
Adversarial Removal of Population Bias in Genomics Phenotype Prediction	334
U-Net Transfer Learning for Image Restoration on Sparse CT Reconstruction in Pre-Clinical Research	341
 VGG-FusionNet: A Feature Fusion Framework from CT Scan and Chest X-ray Images Based Deep Learning for COVID-19 Detection	349

EDMML: Evolutionary Data Mining and Machine Learning

A Paraphrase Identification Approach in Paragraph Length Texts Arwa Al Saqaabi (University of Durham, United Kingdom), Eleni Akrida (University of Durham, United Kingdom), Alexandra Cristea (University of Durham, United Kingdom), and Craig Stewart (University of Durham, United Kingdom)	358
Using Genetic Programming to Identify Probability Distribution Behind Data: A Preliminary Trial	368
A Genetic Programming Approach to Automatically Construct Informative Attributes for Mammographic Density Classification <i>Qurrat Ul Ain (Victoria University of Wellington, New Zealand), Bing</i> <i>Xue (Victoria University of Wellington, New Zealand), Harith Al-Sahaf</i> <i>(Victoria University of Wellington, New Zealand), and Mengjie Zhang</i> <i>(Victoria University of Wellington, New Zealand)</i>	378

Deep Neural Networks and Data Accentuation for Standoff Detection of Dangerous Chemicals 388 Eric Yao (Winston Churchill High School, USA), Christopher A. Kendziora (U.S. Naval Research Laboratory, USA), Robert Furstenberg (U.S. Naval Research Laboratory, USA), Christopher J. Breshike (U.S. Naval Research Laboratory, USA), Drew M. Finton (U.S. Naval Research Laboratory, USA), and Drew C. Kendziora (U.S. Naval Research Laboratory, USA)

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

Linguistic Knowledge Application to Neuro-Symbolic Transformers in Sentiment Analysis) 5
 Sentic Parser: A Graph-Based Approach to Concept Extraction for Sentiment Analysis)3
Making Sense of Sentiments for Aesthetic Plastic Surgery	1
Compression Methods for Transformers in Multidomain Sentiment Analysis	19
Stress Identification in Online Social Networks	27
Feature Extraction and Prediction of Combined Text and Survey Data Using Two-Staged 43 Modeling 43 Asif Ahmed Neloy (University of Manitoba, Canada) and Maxime Turgeon 43 (University of Manitoba, Canada) 10	35
MultiAspectEmo: Multilingual and Language-Agnostic Aspect-Based Sentiment Analysis	43

MRL: Multi-view Representation Learning

Multi-view Representation Learning from Malware to Defend Against Adversarial Variants 451 James Hu (University of Arizona, USA), Mohammadreza Ebrahimi (University of South Florida, USA), Weifeng Li (University of Georgia, USA), Xin Li (University of Arizona, USA), and Hsinchun Chen (University of Arizona, USA)

Deep Transfer Tensor Factorization for Multi-View Learning	159
MORI-RAN: Multi-view Robust Representation Learning via Hybrid Contrastive Fusion	L67
Personalized User Recommendation Based on Various User Behavior in Local Domain	175
Graph Convolutional Networks with Dependency Parser Towards Multiview Representation Learning for Sentiment Analysis	182

UDML: 5th International Workshop on Utility-Driven Mining and Learning

Mining High Utility Itemset with Multiple Minimum Utility Thresholds Based on Utility Deviation	490
Cut the Peaches: Image Segmentation for Utility Pattern Mining in Food Processing Diletta Chiaro (University of Naples Federico II, Italy), Edoardo Prezioso (University of Naples Federico II, Italy), Stefano Izzo (University of Naples Federico II, Italy), Fabio Giampaolo (University of Naples Federico II, Italy), Salvatore Cuomo (University of Naples Federico II, Italy), and Francesco Piccialli (University of Naples Federico II, Italy)	497
Active Keyword Selection to Track Evolving Topics on Twitter Sacha Lévy (McGill University; Mila - Quebec AI Institute, Canada), Farimah Poursafaei (McGill University; Mila - Quebec AI Institute, Canada), Kellin Pelrine (McGill University; Mila - Quebec AI Institute, Canada), and Reihaneh Rabbany (McGill University; Mila - Quebec AI Institute, Canada)	507
Towards Efficient Discovery of Target High Utility Itemsets Vincent Mwintieru Nofong (University of Mines and Technology, Ghana), Priscilla Okai Owiredu (Deloitte Ghana, Ghana), Hamidu Abdel-Fatao (University of Mines and Technology, Ghana), Selasi Kwashie (CSIRO Data61, Australia), Michael Bewong (Charles Sturt University, Australia), and John Wondoh (University of South Australia, Australia)	517

A Machine Learning-Based Approach for Mercury Detection in Marine Waters	
Mining Valuable Fuzzy Patterns via the RFM Model	
 Skyline Pattern Mining by Qantity-Utility Constraints in Large-Scale Databases	
Large-Scale Sequential Utility Pattern Mining in Uncertain Environments	

MLC: Machine Learning for Cybersecurity

 EW-Tune: A Framework for Privately Fine-Tuning Large Language Models with Differential Privacy Rouzbeh Behnia (University of South Florida, USA), Mohammadreza (Reza) Ebrahimi (University of South Florida, USA), Jason Pacheco (University of Arizona, USA), and Balaji Padmanabhan (University of South Florida, USA) 	0
An Adversarial Reinforcement Learning Framework for Robust Machine Learning-Based Malware Detection	7
Mohammadreza Ebrahimi (University of South Florida, USA), Weifeng Li (University of Georgia, USA), Yidong Chai (Hefei University of Technology, China), Jason Pacheco (University of Arizona, USA), and Hsinchun Chen (University of Arizona)	
Backdoor Poisoning of Encrypted Traffic Classifiers	7

Cascaded Multi-class Network Intrusion Detection with Decision Tree and Self-Attentive Model	586
 ENiD: An Encrypted Web Pages Traffic Identification Based on Web Visiting Behavior	93
 Extracting Entities and Events from Cyber-Physical Security Incident Reports	02
Identifying Patterns of Vulnerability Incidence in Foundational Machine Learning Repositories on GitHub: An Unsupervised Graph Embedding Approach	10

HDM: High Dimensional Data Mining

AARS: A Novel Adaptive Archive-Based Efficient Counting Method for Machine Learning Applications Sajib K. Biswas (South Asian University, India), Pranab K. Muhuri (South Asian University, India), and Uttam K. Roy (South Asian University, India)	618
Unsupervised DeepView: Global Uncertainty Visualization for High Dimensional Data Carina Newen (TU Dortmund University) and Emmanuel Müller (TU Dortmund University)	626
Polytopal Complex Construction and Use in Persistent Homology Rohit P. Singh (University of Cincinnati, USA) and Philip A. Wilsey (University of Cincinnati, USA)	634
An Efficient and Reliable Tolerance-Based Algorithm for Principal Component Analysis Michael Yeh (University of California, Berkeley) and Ming Gu (University of California, Berkeley)	642
Unknown Type Streaming Feature Selection via Maximal Information Coefficient Peng Zhou (Anhui University, China), Yunyun Zhang (Anhui University, China), Yuanting Yan (Anhui University, China), and Shu Zhao (Anhui University, China)	650

MLoG: Machine Learning on Graphs

 Extensive Attention Mechanisms in Graph Neural Networks for Materials Discovery	3
SGKD: A Scalable and Effective Knowledge Distillation Framework for Graph Representation Learning	5
Institute of Technology)	
EnD: Enhanced Dedensification for Graph Compressing and Embedding	ł
Improving Graph Neural Network with Learnable Permutation Pooling	2
LSP: Acceleration of Graph Neural Networks via Locality Sensitive Pruning of Graphs)
 HeteroGuard: Defending Heterogeneous Graph Neural Networks Against Adversarial Attacks 698 Udesh Kumarasinghe (University of Colombo, Sri Lanka; SCoRe Lab, Sri Lanka), Mohamed Nabeel (Palo Alto Networks Inc., USA), Kasun De Zoysa (University of Colombo, Sri Lanka), Kasun Gunawardana (University of Colombo, Sri Lanka), and Charitha Elvitigala (SCoRe Lab, Sri Lanka) 	3
Self-Organizing Map-Based Graph Clustering and Visualization on Streaming Graphs	5
Representing Social Networks as Dynamic Heterogeneous Graphs	1
 Graph Convolutional Neural Network Based on the Combination of Multiple Heterogeneous Graphs	1
 Physics-Informed Graph Learning	2
Forecasting Unobserved Node States with Spatio-Temporal Graph Neural Networks)

Improving Your Graph Neural Networks: A High-Frequency Booster Jiaqi Sun (Shenzhen International Graduate School, Tsinghua University, China), Lin Zhang (Tencent, China), Shenglin Zhao (Tencent, China), and Yujiu Yang (Shenzhen International Graduate School, Tsinghua University, China)	748
Degree-Related Bias in Link Prediction Yu Wang (Vanderbilt University, USA) and Tyler Derr (Vanderbilt University, USA)	757
Deep Heterogeneous Graph Neural Networks via Similarity Regularization Loss and Hierarchical Fusion Zhilong Xiong (Guangdong University of Finance & Economics, China) and Jia Cai (Guangdong University of Finance & Economics, China)	759
Behavioral Graph Fraud Detection in E-Commerce Hang Yin (eBay China, China), Zitao Zhang (eBay China, China), Zhurong Wang (eBay China, China), Yilmazcan Özyurt (ETH Zürich, Switzerland), Weiming Liang (eBay China, China), Wenyu Dong (eBay China, China), Yang Zhao (eBay China, China), and Yinan Shan (eBay China, China)	769

IncrLeam: Incremental Classification and Clustering, Concept Drift, Novelty Detection, Active Learning in big/fastdatacontext

Efficient Distributed Algorithms for Minimum Spanning Tree in Dense Graphs	77
InQMAD: Incremental Quantum Measurement Anomaly Detection	37
Discovering Unknown Labels for Multi-label Image Classification	<i>)</i> 7
Reconstruction-Based Unsupervised Drift Detection over Multivariate Streaming Data)7
Incremental Constrained Clustering with Application to Remote Sensing Images Time Series 81 Baptiste Lafabregue (Université de Strasbourg, France), Pierre Gançarski (Université de Strasbourg, France), Jonathan Weber (Université de Haute-Alsace, France), and Germain Forestier (Université de Haute-Alsace, France)	4

A Study of the Dream Net Model Robustness Across Continual Learning Scenarios	324
Nearest Neighbors with Incremental Learning for Real-Time Forecasting of Electricity Demand	834
Laura Melgar-García (Pablo de Olavide University, Spain), David Gutiérrez-Avilés (University of Seville, Spain), Cristina Rubio-Escudero (University of Seville, Spain), and Alicia Troncoso (Pablo de Olavide University, Spain)	
DragStream: An Anomaly and Concept Drift Detector in Univariate Data Streams	342
On Robust Incremental Learning over Many Multilingual Steps	352
Knowledge Distillation-Enabled Multi-stage Incremental Learning for Online Process Monitoring in Advanced Manufacturing	360
Incremental Learning in Time-Series Data Using Reinforcement Learning	868
Data-Driven Kernel Subspace Clustering with Local Manifold Preservation	376

KG: Knowledge Graph

ZeroKBC: A Comprehensive Benchmark for Zero-Shot Knowledge Base Completion Pei Chen (Texas A&M University), Wenlin Yao (Tencent AI Lab), Hongming Zhang (Tencent AI Lab), Xiaoman Pan (Tencent AI Lab), Dian Yu (Tencent AI Lab), Dong Yu (Tencent AI Lab), and Jianshu Chen (Tencent AI Lab)	885
Abnormal Entity-Aware Knowledge Graph Completion Ke Sun (Dalian University of Technology, China), Shuo Yu (Dalian University of Technology, China), Ciyuan Peng (Federation University Australia, Australia), Xiang Li (Dalian University of Technology, China), Mehdi Naseriparsa (Federation University Australia, Australia), and Feng Xia (Federation University Australia, Australia)	891
Towards Fair Representation Learning in Knowledge Graph with Stable Adversarial Debiasing Yihe Wang (University of North Carolina), Mohammad Mahdi Khalili	901

(Yahoo! Research, USA), and Xiang Zhang (University of North Carolina)

FOMO-VL: Foundation Models in Vision and Language

 What Do Audio Transformers Hear? Probing Their Representations for Language Delivery & Structure <i>Yaman Kumar (Adobe Media and Data Science Research (MDSR); SUNY at Buffalo; IIIT-Delhi), Jui Shah (IIIT-Delhi), Changyou Chen (SUNY at Buffalo), and Rajiv Ratn Shah (IIIT-Delhi)</i> 	0
Zero-Shot Object Detection Through Vision-Language Embedding Alignment	6
 STT: Soft Template Tuning for Few-Shot Adaptation	1
Join-Chain Network: A Logical Reasoning View of the Multi-head Attention in Transformer94 Jianyi Zhang (Duke University), Yiran Chen (Duke University), and Jianshu Chen (Tencent AI Lab)	7
A Multi-level Alignment Training Scheme for Video-and-Language Grounding	8

WAIN: AI for Nudging and Personalization

 Deep-SHEEP: Sense of Humor Extraction from Embeddings in the Personalized Context
Multi-purpose Recommender Platform Using Perceiver IO
Simulated Contextual Bandits for Personalization Tasks from Recommendation Datasets
 Learning About People's Attitude Towards Food Available in India and Its Implications for Fair AI-Based Systems

Persona-Based Conversational AI: State of the Art and Challenges Junfeng Liu (Lirio LLC, USA), Christopher Symons (Lirio LLC, USA), and Ranga Raju Vatsavai (Lirio LLC, USA)	.993
ULTRA: A Data-Driven Approach for Recommending Team Formation in Response to Proposal Calls	1002
Biplav Srivastava (University of South Carolina, USA), Tarmo Koppel	
(Tallinn University of Technology, Estonia), Sai Paladi (University of	
South Carolina, USA), Likitha Valluru (University of South Carolina,	
USA), Rohit Sharma (University of South Carolina, USA), and Owen Bond	
(University of South Carolina, USA)	
Modeling Non-Deterministic Human Behaviors in Discrete Food Choices	1010
(Lirio AI Research, Lirio LLC, USA), E. Susanne Blazek (Behavioral	
Reinforcement Learning Lab, Lirio LLC, USA), and Folasade Phillips	
(Behavioral Reinforcement Learning Lab, Lirio LLC, USA)	

SSTDM: Spatial and Spatiotemporal Data Mining

Scalable Joins over Big Data Streams: Actual and Future Trends
Mining of Spatiotemporal Trajectory Profiles Derived from Mobility Data
 On the Deployment of Post-Disaster Building Damage Assessment Tools Using Satellite Imagery: A Deep Learning Approach
A Case Study on Periodic Spatio-Temporal Hotspot Detection in Azure Traffic Data
Edit Distance with Quasi Real Penalties: A Hybrid Distance for Network-Constrained Trajectories

Corn Grain Yield Prediction Using UAV-Based High Spatiotemporal Resolution Multispectral Imagery	4
Deep-Learning-Based Precipitation Nowcasting with Ground Weather Station Data and Radar Data	3
Jihoon Ko (Korea Advanced Institute of Science and Technology, South Korea), Kyuhan Lee (Korea Advanced Institute of Science and Technology, South Korea), Hyunjin Hwang (Korea Advanced Institute of Science and Technology, South Korea), and Kijung Shin (Korea Advanced Institute of Science and Technology, South Korea)	
Distributed LSTM-Learning from Differentially Private Label Proportions	1
 Efficient Spatio-Temporal Randomly Wired Neural Networks for Traffic Forecasting	9
 End-to-End Modeling of Hierarchical Time Series Using Autoregressive Transformer and Conditional Normalizing Flow-Based Reconciliation	7

DMESS: Data Mining in Earth System Science

Macaw: The Machine Learning Magnetometer Calibration Workflow	1095
Identifying Hydrometeorological Factors Influencing Reservoir Releases Using Machine	
Learning Methods 1	1102
Ming Fan (Oak Ridge National Laboratory, USA), Lujun Zhang (The	
University of Oklahoma, USA), Siyan Liu (Oak Ridge National	
Laboratory, USA), Tiantian Yang (The University of Oklahoma, USA), and	
Dan Lu (Oak Ridge National Laboratory, USA)	
Improving net Ecosystem CO_2 Flux Prediction Using Memory-Based Interpretable Machine	
Learning	1111
Siyan Liu (Computational Sciences and Engineering Division, Oak Ridge	
National Laboratory, USA), Dan Lu (Computational Sciences and	
Engineering Division, Oak Ridge National Laboratory, USA), Daniel	
Ricciuto (Environmental Sciences Division, Oak Ridge National	
Laboratory, USA), and Anthony Walker (Environmental Sciences Division,	
Oak Ridge National Laboratory, USA)	

A Hybrid ConvLSTM Deep Neural Network for Noise Reduction and Data Augmentation for
Prediction of Non-Linear Dynamics of Streamflow
Juan F. Ramirez Rochac (University of the District of Columbia, USA),
Nian Zhang (University of the District of Columbia, USA), Tolessa
Deksissa (University of the District of Columbia, USA), Jiajun Xu
(University of the District of Columbia, USA), and Lara Thompson
(University of the District of Columbia, USA)
Above Ground Biomass Estimation of a Cocoa Plantation using Machine Learning
Using Image Processing Techniques to Identify and Quantify Spatiotemporal Carbon Cycle
Extremes
Bharat Sharma (Oak Ridge National Laboratory, USA), Jitendra Kumar
(Oak Ridge National Laboratory, USA), Auroop R. Ganguly (Northeastern
University, USA), and Forrest M. Hoffman (Oak Ridge National
Laboratory, USA)

NeuRec: The Third International Workshop on Advanced Neural Algorithms and Theories for Recommender Systems

TPE-AutoClust: A Tree-Based Pipline Ensemble Framework for Automated Clustering 1144 Radwa ElShawi (University of Tartu, Estonia) and Sherif Sakr (University of Tartu, Estonia)
Next POI Recommender System: Multi-view Representation Learning for Outstanding
Performance in Various Context
Yeonghwan Jeon (HyperLocal Personalization Team, NAVER Corporation,
Republic of Korea) and Junhyung Kim (HyperLocal Personalization Team,
NAVER Corporation, Republic of Korea)
A Recommendation System Framework to Generalize AutoRec and Neural Collaborative Filtering
1162
Ramin Raziperchikolaei (Rakuten, Inc., USA) and Young-joo Chung

(Rakuten, Inc., USA)

SocDM: Social Data Mining in the Post-Pandemic Era

Post-Pandemic Economic Transformations in the United States of America
Visual Analytics of Mobility Network Changes Observed using Mobile Phone Data during COVID-19 Pandemic
Karimipour (Institute of Science and Technology Austria (ISTA) Austria), and Georg Heiler (Vienna University of Technology, Austria)
 Human Mobility Driven Modeling of an Infectious Disease

ML-HOS: Machine Learning on Higher-Order Structured Data

Influence Maximization on Hypergraphs via Similarity-based Diffusion Mehmet Aktas (Georgia State University), Sidra Jawaid (University of Central Oklahoma), Ihsan Gokalp (North American University), and Esra Akbas (Georgia State University)	1197
Dynamic Combination of Heterogeneous Models for Hierarchical Time Series Xing Han (UT-Austin, USA), Jing Hu (Intuit, USA), and Joydeep Ghosh (UT-Austin, USA)	1207
Hybrid Oversampling Technique Based on Star Topology and Rejection Methodology for Classifying Imbalanced Data	1217
Exploiting Cross-Order Patterns and Link Prediction in Higher-Order Networks	1227

Forums

GPUCSL: GPU-Based Library for Causal Structure Learning	າg
Tom Braun (Hasso Plattner Institute, Germany), Ben Hurdelhey (Hasso	
Plattner Institute, Germany), Dominik Meier (Hasso Plattner Institute,	
Germany), Petr Tsayun (Hasso Plattner Institute, Germany), Christopher	
Hagedorn (Hasso Plattner Institute, Germany), Johannes Huegle (Hasso	
Plattner Institute, Germany), and Rainer Schlosser (Hasso Plattner	
Institute, Germany)	

Author Index