

SIAM International Conference on Data Mining (SDM24)

Houston, Texas, USA
18-20 April 2024

Editors:

**Shashi Shekhar
Vagelis Papalexakis
Jing Gao**

**Zhe Jiang
Matteo Riondato**

ISBN: 978-1-7138-9595-4

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2024) by SIAM: Society for Industrial and Applied Mathematics
All rights reserved.

Printed with permission by Curran Associates, Inc. (2024)

For permission requests, please contact SIAM: Society for Industrial and Applied Mathematics
at the address below.

SIAM
3600 Market Street, 6th Floor
Philadelphia, PA 19104-2688 USA

Phone: (215) 382-9800

siambooks@siam.org

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
Email: curran@proceedings.com
Web: www.proceedings.com

TABLE OF CONTENTS

CP01

Local Differential Privacy in Graph Neural Networks: A Reconstruction Approach	1
<i>Karuna Bhaila, Wen Huang, Yongkai Wu and Xintao Wu</i>	
GENIUS: Subteam Replacement with Clustering-Based Graph Neural Networks	10
<i>Chuxuan Hu, Qinghai Zhou and Hanghang Tong</i>	
Laplacian Score Benefit Adaptive Filter Selection for Graph Neural Networks	19
<i>Yewen Wang, Shichang Zhang, John (Junghoo) Cho and Yizhou Sun</i>	
Self-Similar Graph Neural Network for Hierarchical Graph Learning	28
<i>Zheng Zhang and Liang Zhao</i>	

CP02

PUPAE: Intuitive and Actionable Explanations for Time Series Anomalies	37
<i>Audrey Der, Chin-Chia Michael Yeh, Yan Zheng, Junpeng Wang, Zhongfang Zhuang, Liang Wang, Wei Zhang and Eamonn Keogh</i>	
An Exemplars-Based Approach for Explainable Clustering: Complexity and Efficient Approximation Algorithms	46
<i>Ian Davidson, Michael Livanos, Antoine Gourru, Peter Walker, Julien Velcin and S.S. Ravi</i>	
MISS: Multiclass Interpretable Scoring Systems	55
<i>Michal K. Grzeszczyk, Tomasz Trzcíński and Arkadiusz Sitek</i>	
XGExplainer: Robust Evaluation-Based Explanation for Graph Neural Networks	64
<i>Ryoji Kubo and Djellel Difallah</i>	
STES: A Spatiotemporal Explanation Supervision Framework	73
<i>Dazhou Yu, Binbin Chen, Yun Li, Suman Dhakal, Yifei Zhang, Zhenke Liu, Minxing Zhang, Jie Zhang and Liang Zhao</i>	

CP03

CoLafier: Collaborative Noisy Label Purifier with Local Intrinsic Dimensionality Guidance	82
--	-----------

Dongyu Zhang, Ruofan Hu and Elke Rundensteiner

Differences between Hard and Noisy-Labeled Samples: An Empirical Study	91
---	-----------

Mahsa Forouzesh and Patrick Thiran

Feature-Engineered Random Forests	100
--	------------

Saket Sathe, Charu Aggarwal, Horst Samulowitz and Deepak Turaga

Multi-polytope Machine for Classification	109
--	------------

Dzung T. Phan, Lam Nguyen, Jayant Kalagnanam and Chandra Reddy

YOLO-OCR: End-to-End Compound Figure Separation and Label Recognition of Images in Scientific Publications	118
---	------------

Shou Meng, Xinshuo Liang, Shuai Zhang, Leqi Lei, Hanbai Wu, Saira Iqbal and Jinlian Hu

CP04

Tensorized Hypergraph Neural Networks	127
--	------------

Maolin Wang, Yaoming Zhen, Yu Pan, Yao Zhao, Chenyi Zhuang, Zenglin Xu, Ruocheng Guo and Xiangyu Zhao

Training Sparse Graph Neural Networks via Pruning and Sprouting	136
--	------------

Xueqi Ma, Xingjun Ma, Sarah Erfani and James Bailey

Prompt Based Tri-channel Graph Convolution Neural Network for Aspect Sentiment Triplet Extraction	145
--	------------

Kun Peng, Lei Jiang, Hao Peng, Rui Liu, Zhengtao Yu, Jiaqian Ren, Zhifeng Hao and Philip S. Yu

Non-euclidean Spatial Graph Neural Network	154
---	------------

Zheng Zhang, Sirui Li, Jingcheng Zhou, Junxiang Wang, Abhinav Angirekula, Allen Zhang and Liang Zhao

CP05

EsaCL: An Efficient Continual Learning Algorithm	163
---	------------

Weijieying Ren and Vasant G. Honavar

On the Number of Iterations of the DBA Algorithm	172
---	------------

Frederik Brünig, Anne Driemel, Alperen Ergür and Heiko Röglin

Robust Sparse Online Learning for Data Streams with Streaming Features	181
---	------------

Zhong Chen, Yi He, Di Wu, Huixin Zhan, Victor Sheng and Kun Zhang

Utility-Oriented String Mining	190
---	------------

Giulia Bernardini, Huiping Chen, Alessio Conte, Roberto Grossi, Veronica Guerrini, Grigorios Loukides, Nadia Pisanti and Solon P. Pissis

Early Multiple Temporal Patterns Based Event Prediction in Heterogeneous Multivariate Temporal Data	199
--	------------

Nevo Itzhak, Szymon Jaroszewicz and Robert Moskovitch

CP06

Semi-supervised Clustering via Structural Entropy with Different Constraints	208
---	------------

Guangjie Zeng, Hao Peng, Angsheng Li, Zhiwei Liu, Runze Yang, Chunyang Liu and Lifang He

Towards Tuning-Free Minimum-Volume Nonnegative Matrix Factorization	217
--	------------

Duc Toan Nguyen and Eric C. Chi

Decentralized Stochastic Compositional Gradient Descent for AUPRC Maximization	226
---	------------

Hongchang Gao, Yubin Duan, Yihan Zhang and Jie Wu

On Robust Wasserstein Barycenter: The Model and Algorithm	235
--	------------

Xu Wang, Jiawei Huang, Qingyuan Yang and Jinpeng Zhang

Towards Optimization and Model Selection for Domain Generalization: A Mixup-Guided Solution	244
--	------------

Wang Lu, Jindong Wang, Yidong Wang and Xing Xie

CP07

Helper Recommendation with Seniority Control in Online Health Community	253
--	------------

Junruo Gao, Chen Ling, Carl Yang and Liang Zhao

Active Learning for Graphs with Noisy Structures	262
---	------------

Hongliang Chi, Cong Qi, Suhang Wang and Yao Ma

Graph Summarization for Preserving Spectral Characteristics	271
--	------------

Houquan Zhou, Shenghua Liu, Huawei Shen and Xueqi Cheng

H²ABM: Heterogeneous Agent-Based Model on Hypergraphs to Capture Group Interactions	280
---	------------

Vivek Anand, Jiaming Cui, Jack Heavey, Anil Vullikanti and B. Aditya Prakash

Treatment-Aware Hyperbolic Representation Learning for Causal Effect Estimation with Social Networks	289
<i>Ziqiang Cui, Xing Tang, Yang Qiao, Bowei He, Liang Chen, Xiuqiang He and Chen Ma</i>	

CP08

EBV: Electronic Bee-Veterinarian for Principled Mining and Forecasting of Honeybee Time Series	298
<i>Mst. Shamima Hossain, Christos Faloutsos, Boris Baer, Hyoseung Kim and Vassilis J. Tsotras</i>	

Message Propagation through Time: An Algorithm for Sequence Dependency Retention in Time Series Modeling	307
<i>Shaoming Xu, Ankush Khandelwal, Arvind Renganathan and Vipin Kumar</i>	

Pattern-Based Time Series Semantic Segmentation with Gradual State Transitions	316
<i>Louis Carpentier, Len Feremans, Wannes Meert and Mathias Verbeke</i>	

Time-Transformer: Integrating Local and Global Features for Better Time Series Generation	325
<i>Yuansan Liu, Sudanthi Wijewickrema, Ang Li, Christofer Bester, Stephen O'Leary and James Bailey</i>	

Towards Entity-Aware Conditional Variational Inference for Heterogeneous Time-Series Prediction: An Application in Hydrology	334
<i>Rahul Ghosh, Arvind Renganathan, Wallace McAliley, Michael Steinbach, Christopher Duffy and Vipin Kumar</i>	

CP09

3D Molecular Geometry Analysis with 2D Graphs	343
<i>Zhao Xu, Yaochen Xie, Youzhi Luo, Xuan Zhang, Xinyi Xu, Meng Liu, Kaleb Dickerson, Cheng Deng, Maho Nakata and Shuiwang Ji</i>	

A Novel Hybrid Graph Learning Method for Inbound Parcel Volume Forecasting in Logistics System	352
<i>Lisha Ye, Jianfeng Zhou, Zhe Yin, Kunpeng Han, Haoyuan Hu and Dongjin Song</i>	

Automated Fusion of Multimodal Electronic Health Records for Better Medical Predictions	361
<i>Suhan Cui, Jiaqi Wang, Yuan Zhong, Han Liu, Ting Wang and Fenglong Ma</i>	

Dual-Stage Flows-Based Generative Modeling for Traceable Urban Planning	370
<i>Xuanming Hu, Wei Fan, Dongjie Wang, Pengyang Wang, Yong Li and Yanjie Fu</i>	

Graph-Based Student Knowledge Profile for Online Intelligent Education	379
<i>Jinze Wu, Haotian Zhang, Zhenya Huang, Liang Ding, Qi Liu, Jing Sha, Enhong Chen and Shijin Wang</i>	

CP10

Data Silences: How to Unsilence the Uncertainties in Data Science	388
--	------------

Michael Muller

Foundation Models for Spatiotemporal Tasks in the Physical World	392
---	------------

Zhe Jiang, Yu Wang and Zelin Xu

Blue Sky: Multilingual, Multimodal Domain Independent Deception Detection	396
--	------------

Dainis Boumber, Rakesh M. Verma and Fatima Zhara Qachfar

CP11

Label Distribution Learning-Enhanced Dual-KNN for Text Classification	400
--	------------

Bo Yuan, Yulin Chen, Zhen Tan, Wang Jinyan, Huan Liu and Yin Zhang

Refining Pre-trained Language Models for Domain Adaptation with Entity-Aware Discriminative and Contrastive Learning	409
---	------------

Jian Yang, Xinyu Hu, Yulong Shen and Gang Xiao

An Image Dataset for Benchmarking Recommender Systems with Raw Pixels	418
--	------------

Yu Cheng, Yunzhu Pan, Jiaqi Zhang, Yongxin Ni, Aixin Sun and Fajie Yuan

Disinformation Detection: An Evolving Challenge in the Age of LLMs	427
---	------------

Bohan Jiang, Zhen Tan, Ayushi Nirmal and Huan Liu

User Migration across Multiple Social Media Platforms	436
--	------------

Ujun Jeong, Ayushi Nirmal, Kritshekhar Jha, Susan Xu Tang, H. Russell Bernard and Huan Liu

CP12

Combining Satellite and Weather Data for Crop Type Mapping: An Inverse Modelling Approach	445
--	------------

Praveen Ravirathinam, Rahul Ghosh, Ankush Khandelwal, Xiaowei Jia, David Mulla and Vipin Kumar

CAMLO: Cross-Attentive Multi-view Network for Long-Term Origin-Destination Flow Prediction	454
---	------------

Liang Wang, Hao Fu, Shu Wu, Qiang Liu, Xuelei Tan, Fangsheng Huang, Mengdi Zhang and Wei Wu

Test-Time Training for Spatial-Temporal Forecasting	463
--	------------

Changlu Chen, Yanbin Liu, Ling Chen and Chengqi Zhang

Geospatial Topological Relation Extraction from Text with Knowledge Augmentation	472
---	------------

Wei Hu, Bowen Jin, Minhao Jiang, Sizhe Zhou, Zhaonan Wang, Jiawei Han and Shaowen Wang

Only Attending What Matter within Trajectories – Memory-Efficient Trajectory Attention	481
---	------------

Mingzhi Hu, Xin Zhang, Yanhua Li, Yiqun Xie, Xiaowei Jia, Xun Zhou and Jun Luo

CP13

Lattice Convolutional Networks for Learning Ground States of Quantum Many-Body Systems	490
---	------------

Cong Fu, Xuan Zhang, Huixin Zhang, Hongyi Ling, Shenglong Xu and Shuiwang Ji

MedDiffusion: Boosting Health Risk Prediction via Diffusion-Based Data Augmentation	499
--	------------

Yuan Zhong, Suhan Cui, Jiaqi Wang, Xiaochen Wang, Ziyi Yin, Yaqing Wang, Houping Xiao, Mengdi Huai, Ting Wang and Fenglong Ma

Meta-Adaptive Stock Movement Prediction with Two-Stage Representation Learning	508
---	------------

Donglin Zhan, Yusheng Dai, Yiwei Dong, Jinghai He, Zhenyi Wang and James Anderson

Pretraining Molecules with Explicit Substructure Information	517
---	------------

Yuting Ma, Shuo Yu and Yanming Shen

RHINE: A Regime-Switching Model with Nonlinear Representation for Discovering and Forecasting Regimes in Financial Markets	526
---	------------

Kunpeng Xu, Lifei Chen, Jean-Marc Patenaude and Shengrui Wang

CP14

FedDCSR: Federated Cross-Domain Sequential Recommendation via Disentangled Representation Learning	535
---	------------

Hongyu Zhang, Dongyi Zheng, Xu Yang, Jiyuan Feng and Qing Liao

Denoising Long- and Short-Term Interests for Sequential Recommendation	544
---	------------

Xinyu Zhang, Beibei Li and Beihong Jin

CausalCDR: Causal Embedding Learning for Cross-Domain Recommendation	553
---	------------

Fengxin Li, Hongyan Liu, Jun He and Xiaoyong Du

DimReg: Embedding Dimension Search via Regularization for Recommender Systems	562
--	------------

Mingjun Zhao, Liyao Jiang, Yakun Yu, Xinmin Wang, Yi Yuan, Zheng Wei and Di Niu

DualVAE: Dual Disentangled Variational AutoEncoder for Recommendation	571
--	------------

Zhiqiang Guo, Guohui Li, Jianjun Li, Chaoyang Wang and Si Shi

Exploiting Multifaceted Nature of Items and Users for Session-Based Recommendation	580
---	------------

Tianqi Sun, Hongrui Guo, Zihan Zhang, Hongzhi Liu and Zhonghai Wu

CP15

Prescribed Fire Modeling Using Knowledge-Guided Machine Learning for Land Management	589
---	------------

Somya Sharma Chatterjee, Kelly Lindsay, Neel Chatterjee, Rohan Patil, Ilkay Altintas De Callafon, Michael Steinbach, Daniel Giron, Mai H. Nguyen and Vipin Kumar

Spatial-Temporal Augmented Adaptation via Cycle-Consistent Adversarial Network: An Application in Streamflow Prediction	598
--	------------

Nasrin Kalanat, Yiqun Xie, Yanhua Li and Xiaowei Jia

Bridging Semantics: Mobility Analytics Framework for Knowledge Transfer	607
--	------------

Shreya Ghosh and Prasenjit Mitra

Towards Spatially-Lucid AI Classification in Non-euclidean Space: An Application for MxIF Oncology Data	616
--	------------

Majid Farhadloo, Arun Sharma, Jayant Gupta, Alexey Leontovich, Svetomir N. Markovic and Shashi Shekhar

Unified Modeling and Clustering of Mobility Trajectories with Spatiotemporal Point Processes	625
---	------------

Haowen Lin, Yao-Yi Chiang, Li Xiong and Cyrus Shahabi

CP16

Ada-VAD: Domain Adaptable Video Anomaly Detection	634
--	------------

Dongliang Guo, Yun Fu and Sheng Li

COMBOOD: A Semiparametric Approach for Detecting Out-of-Distribution Data for Image Classification	643
---	------------

Magesh Rajasekaran, Md Saiful Islam Sajol, Frej Berglind, Supratik Mukhopadhyay and Kamalika Das

Dimensionality-Aware Outlier Detection	652
---	------------

Alastair Anderberg, James Bailey, Ricardo J.G.B. Campello, Michael E. Houle, Henrique O. Marques, Miloš Radovanović and Arthur Zimek

MultiNetAD: Multiplex Network-Based Anomaly Access Detection Featuring Semantic Hierarchies	661
--	------------

Ziqi Yuan, Qingyun Sun, Haoyi Zhou, Zukun Zhu and Jianxin Li

Semi-supervised Isolation Forest for Anomaly Detection	670
---	------------

Luca Stradiotti, Lorenzo Perini and Jesse Davis

CP17

Dual-Disentangled Deep Multiple Clustering	679
---	------------

Jiawei Yao and Juhua Hu

DualToken-ViT: Position-Aware Efficient Vision Transformer with Dual Token Fusion	688
--	------------

Zhenzhen Chu, Jiayu Chen, Cen Chen, Chengyu Wang, Ziheng Wu, Jun Huang and Weining Qian

Identifying and Uses of Deep Learning Backbones via Pattern Mining	697
---	------------

Michael Livanos and Ian Davidson

GE-AdvGAN: Improving the Transferability of Adversarial Samples by Gradient Editing-Based Adversarial Generative Model	706
---	------------

Zhiyu Zhu, Huaming Chen, Xinyi Wang, Jiayu Zhang, Zhibo Jin, Kim-Kwang Raymond Choo, Jun Shen and Dong Yuan

Knowledge Guided Machine Learning for Extracting, Preserving, and Adapting Physics-Aware Features	715
--	------------

Erhu He, Yiqun Xie, Licheng Liu, Zhenong Jin, Dajun Zhang and Xiaowei Jia

CP18

Light POI-Guided Conversational Recommender System Based on Adaptive Space	724
---	------------

Meng Yuan, Fuwei Zhang, Yiqi Tong, Yuxin Ying, Fuzhen Zhuang, Deqing Wang, Baoxing Huai, Yi Zhang and Jia Su

Multi-interest Network with Simple Diffusion for Multi-behavior Sequential Recommendation	734
--	------------

Qingfeng Li, Huifang Ma, Wangyu Jin, Yugang Ji and Zhixin Li

Towards More Robust and Accurate Sequential Recommendation with Cascade-Guided Adversarial Training	743
--	------------

Juntao Tan, Shelby Heinecke, Zhiwei Liu, Yongjun Chen, Yongfeng Zhang and Huan Wang

Variational Invariant Representation Learning for Multimodal Recommendation	752
--	------------

Wei Yang, Haoran Zhang and Li Zhang

Vietoris-Rips Complex: A New Direction for Cross-Domain Cold-Start Recommendation	761
--	------------

Ajay Krishna Vajjala, Dipak Falgun Meher, Shrinal Pothagoni, Ziwei Zhu and David S. Rosenblum

CP19

Stable Synthetic Control with Anomaly Detection for Causal Inference	770
---	------------

Qiang Li, Yiqiao Sun, Linsey Pang, Liang Sun and Qingsong Wen

CaMU: Disentangling Causal Effects in Deep Model Unlearning	779
--	------------

Shaofei Shen, Chenhao Zhang, Alina Bialkowski, Weitong Chen and Miao Xu

Robust Estimation of Causal Heteroscedastic Noise Models	788
---	------------

Quang-Duy Tran, Bao Duong, Phuoc Nguyen and Thin Nguyen

Analysis of Causal and Non-causal Convolution Networks for Time Series Classification	797
--	------------

Uday Singh Saini, Zhongfang Zhuang, Chin-Chia Michael Yeh, Wei Zhang and Evangelos E. Papalexakis

CP20

Deep Efficient Private Neighbor Generation for Subgraph Federated Learning	806
---	------------

Ke Zhang, Lichao Sun, Bolin Ding, Siu Ming Yiu and Carl Yang

Distributed Collapsed Gibbs Sampler for Dirichlet Process Mixture Models in Federated Learning	815
---	------------

Reda Khoufache, Mustapha Lebbah, Hanene Azzag, Etienne Goffinet and Djamel Bouchaffra

HyperFLORA: Federated Learning with Instantaneous Personalization	824
--	------------

Qikai Lu, Di Niu, Mohammadamin Samadi Khoshkho and Baochun Li

AEDFL: Efficient Asynchronous Decentralized Federated Learning with Heterogeneous Devices	833
--	------------

Ji Liu, Tianshi Che, Yang Zhou, Ruoming Jin, Huaiyu Dai, Dejing Dou and Patrick Valduriez

Personalized Federated Learning with Contextual Modulation and Meta-Learning	842
---	------------

Anna Vettoruzzo, Mohamed-Rafik Bouguelia and Thorsteinn Rögnvaldsson

UPFL: Unsupervised Personalized Federated Learning towards New Clients	851
---	------------

Tiandi Ye, Cen Chen, Yinggui Wang, Xiang Li and Ming Gao

CP21

HADT: Human-AI Diagnostic Team via Hierarchical Reinforcement Learning	860
---	------------

Xuehan Zhao, Jiaqi Liu, Zhiwen Yu and Bin Guo

Spatial-Aware Deep Reinforcement Learning for the Traveling Officer Problem	869
--	------------

Niklas Strauß and Matthias Schubert

Feature Interaction Aware Automated Data 878
Ehtesamul Azim, Dongjie Wang, Kunpeng Liu, Wei Zhang and Yanjie Fu

Neural Locality Sensitive Hashing for Entity Blocking 887
Runhui Wang, Luyang Kong, Yefan Tao, Andrew Borthwick, Davor Golac, Henrik Johnson, Shadie Hijazi, Dong Deng and Yongfeng Zhang

Word Embedding with Neural Probabilistic Prior 896
Shaogang Ren, Dingcheng Li and Ping Li