

SIAM International Conference on Data Mining (SDM'23)

Minneapolis-St. Paul Twin Cities, Minnesota, USA
27 - 29 April 2023

ISBN: 978-1-7138-8294-7

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© (2023) 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

CADENCE: Community-Aware Detection of Dynamic Network States	1
---	----------

Maxwell McNeil, Carolina Mattsson, Frank W. Takes, Petko Bogdanov

Influence without Authority: Maximizing Information Coverage in Hypergraphs	10
--	-----------

Peiyang Li, Honglian Wang, Kai Li, Christian Böhm

A Temporal Graphlet Kernel for Classifying Dissemination in Evolving Networks	19
--	-----------

Lutz Oettershagen, Nils M. Kriege, Claude Jordan, Petra Mutzel

Causal Discovery by Graph Attention Reinforcement Learning	28
---	-----------

Dezhi Yang, Guoxian Yu, Jun Wang, Zhongmin Yan, Maozu Guo

Heterogeneous Graph Contrastive Learning with Meta-Path Contexts and Weighted Negative Samples	37
---	-----------

Jianxiang Yu, Xiang Li

CP02

UFNRec: Utilizing False Negative Samples for Sequential Recommendation	46
---	-----------

Xiaoyang Liu, Chong Liu, Pinzheng Wang, Rongqin Zheng, Lixin Zhang, Leyu Lin, Zhijun Chen, Liangliang Fu

Adaptive Label Smoothing to Regularize Large-Scale Graph Training	55
--	-----------

Kaixiong Zhou, Soo-Hyun Choi, Zirui Liu, Ninghao Liu, Fan Yang, Rui Chen, Li Li, Xia Hu

Pluggable Deep Thompson Sampling with Applications to Recommendation	64
---	-----------

Lu Wang, Yuhai Song, Zhe Wang, Haoxiang Wang, Yu Li, Weiwei Zhou, Haoming Dang, Mona Shao, Xiwei Zhao, Zhangang Lin, Jinghe Hu, Jingping Shao

Embedding Transfer with Enhanced Correlation Modeling for Cross-Domain Recommendation	73
--	-----------

Shilei Cao, Yujie Lin, Xianli Zhang, Yufu Chen, Zhen Zhu, Yuxin Chen, Buyue Qian, Feng Wang, Zang Li

HARVESTER: Principled Factorization-Based Temporal Tensor Granularity Estimation	82
---	-----------

Ravdeep S. Pasricha, Uday Singh Saini, Nicholas D. Sidiropoulos, Fei Fang, Kevin Chan, Evangelos E. Papalexakis

CP03

Max-Min Diversification with Fairness Constraints: Exact and Approximation Algorithms	91
<i>Yanhao Wang, Michael Mathioudakis, Jia Li, Francesco Fabbri</i>	
Beyond the Evidence Lower Bound: Dual Variational Graph Auto-Encoders for Node Clustering	100
<i>Nairouz Mrabah, Mohamed Bouguessa, Riadh Ksantini</i>	
Extension of the Dip-Test Repertoire – Efficient and Differentiable p-Value Calculation for Clustering	109
<i>Lena G.M. Bauer, Collin Leiber, Christian Böhm, Claudia Plant</i>	
Reinforced EM Algorithm for Clustering with Gaussian Mixture Models	118
<i>Joshua Tobin, Chin Pang Ho, Mimi Zhang</i>	
A Lagrangian-Based Approach to Learn Distance Metrics for Clustering with Minimal Data Transformation	127
<i>Rodrigo Randel, Daniel Aloise, Alain Hertz</i>	

CP04

Heterogeneous Graph Contrastive Multi-View Learning	136
<i>Zehong Wang, Qi Li, Donghua Yu, Xiaolong Han, Xiao-Zhi Gao, Shigen Shen</i>	
Multimodal Graph Learning for Cross-Modal Retrieval	145
<i>Jingyou Xie, Zishuo Zhao, Zhenzhou Lin, Ying Shen</i>	
RELIANT: Fair Knowledge Distillation for Graph Neural Networks	154
<i>Yushun Dong, Binchi Zhang, Yiling Yuan, Na Zou, Qi Wang, Jundong Li</i>	
Adversarial Hard Negative Generation for Complementary Graph Contrastive Learning	163
<i>Senzhang Wang, Hao Yan, Jinlong Du, Jun Yin, Junxing Zhu, Chaozhuo Li, Jianxin Wang</i>	
It's PageRank All the Way Down: Simplifying Deep Graph Networks	172
<i>Dominic Jack, Sarah Erfani, Jeffrey Chan, Sutharshan Rajasegarar, Christopher Leckie</i>	

CP05

Estimating Latent Population Flows from Aggregated Data via Inversing Multi-Marginal Optimal Transport	181
<i>Sikun Yang, Hongyuan Zha</i>	

Towards Learning in Grey Spatiotemporal Systems: A Prophet to Non-consecutive Spatiotemporal Dynamics 190
Zhengyang Zhou, Kuo Yang, Wei Sun, Binwu Wang, Min Zhou, Yunan Zong, Yang Wang

StAGN: Spatial-Temporal Adaptive Graph Network via Constranstive Learning for Sleep Stage Classification 199
Junyang Chen, Yidan Dai, Xianhui Chen, Yingshan Shen, Yan Luximon, Hailiang Wang, Yuxin He, Wenjun Ma, Xiaomao Fan

STM-GAIL: Spatial-Temporal Meta-GAIL for Learning Diverse Human Driving Strategies 208
Yingxue Zhang, Yanhua Li, Xun Zhou, Ziming Zhang, Jun Luo

Physics-Guided Meta-Learning Method in Baseflow Prediction over Large Regions 217
Shengyu Chen, Yiqun Xie, Xiang Li, Xu Liang, Xiaowei Jia

CP06

ML4C: Seeing Causality through Latent Vicinity 226
Haoyue Dai, Rui Ding, Yuanyuan Jiang, Shi Han, Dongmei Zhang

Uncertainty in Selective Bagging: A Dynamic Bi-objective Optimization Model 235
Mansoureh Maadi, Hadi Akbarzadeh Khorshidi, Uwe Aickelin

Saliency-Augmented Memory Completion for Continual Learning 244
Guangji Bai, Chen Ling, Yuyang Gao, Liang Zhao

An Interpretable Measure of Dataset Complexity for Imbalanced Classification Problems 253
Jonatan Møller Nuutinen Gøttcke, Colin Bellinger, Paula Branco, Arthur Zimek

CESED: Exploiting Hyperspherical Predefined Evenly-Distributed Class Centroids for OOD Detection 262
Shuai Feng, Wenyu Jiang, Mingcai Chen, Yuntao Du, Hao Cheng, Yuxin Ge, Chongjun Wang

CP07

AlignGraph: A Group of Generative Models for Graphs 271
Kimia Shayestehfard, Dana Brooks, Stratis Ioannidis

An Index for Temporal Closeness Computation in Evolving Graphs 280
Lutz Oettershagen, Petra Mutzel

An Efficient Algorithm for Assessing the Number of St-Paths in Large Graphs 289
Giulia Punzi, Alessio Conte, Roberto Grossi, Andrea Marino

Representation Learning on Dynamic Network of Networks 298
Si Zhang, Yinglong Xia, Yan Zhu, Hanghang Tong

Heavy Nodes in a Small Neighborhood: Algorithms and Applications 307
Huiping Chen, Grigorios Loukides, Robert Gwadera, Solon P. Pissis

CP08

A Hidden Markov Forest Model for Terrain-Aware Flood Inundation Mapping from Earth Imagery 316
Zhe Jiang, Yupu Zhang, Saugat Adhikari, Da Yan, Arpan Man Sainju, Xiaowei Jia, Yiqun Xie

Time-Delayed Multivariate Time Series Predictions 325
Hao Niu, Guillaume Habault, Roberto Legaspi, Chuizheng Meng, Defu Cao, Shinya Wada, Chihiro Ono, Yan Liu

A Two-View EEG Representation for Brain Cognition by Composite Temporal-Spatial Contrastive Learning 334
Zheng Chen, Lingwei Zhu, Haohui Jia, Takashi Matsubara

Hierarchical Reinforced Urban Planning: Jointly Steering Region and Block Configurations 343
Pengfei Wang, Daniel Wang, Kunpeng Liu, Dongjie Wang, Yuanchun Zhou, Leilei Sun, Yanjie Fu

Why Are We Waiting? Discovering Interpretable Models for Predicting Sojourn and Waiting Times 352
Boris Wiegand, Dietrich Klakow, Jilles Vreeken

CP09

Physics-Guided Graph Diffusion Network for Combining Heterogeneous Simulated Data: An Application in Predicting Stream Water Temperature 361
Xiaowei Jia, Shengyu Chen, Can Zheng, Yiqun Xie, Zhe Jiang, Nasrin Kalanat

CISum: Learning Cross-Modality Interaction to Enhance Multimodal Semantic Coverage for Multimodal Summarization 370
Litian Zhang, Xiaoming Zhang, Ziming Guo, Zhipeng Liu

ProGrEST: Prototypical Graph Regression Soft Trees for Molecular Property Prediction 379
Dawid Rymarczyk, Daniel Dobrowolski, Tomasz Danel

A Physics-Guided NN-Based Approach for Tropical Cyclone Intensity Estimation 388
Ziheng Zhou, Ying Zhao, Yiyu Qing, Wenming Jiang, Yihan Wu, Wenguang Chen

Anomaly Detection Networks and Fuzzy Control Modules for Energy Grid Management with Q-Learning-Based Decision Making	397
--	------------

Jia-Hao Syu, Jerry Chun-Wei Lin, Philip S. Yu

CP10

STABLE: Identifying and Mitigating Instability in Embeddings of the Degenerate Core	406
--	------------

David Liu, Tina Eliassi-Rad

Node Ranking in Labeled Networks	415
---	------------

Chamalee Wickrama Arachchi, Nikolaj Tatti

Optimal Intervention on Weighted Networks via Edge Centrality	424
--	------------

Dongyue Li, Tina Eliassi-Rad, Hongyang R. Zhang

GIST: Graph Inference for Structured Time Series	433
---	------------

Boya Ma, Maxwell McNeil, Petko Bogdanov

DyFormer: A Scalable Dynamic Graph Transformer with Provable Benefits on Generalization Ability	442
--	------------

Weilin Cong, Yanhong Wu, Yuandong Tian, Mengting Gu, Yinglong Xia, Chun-Cheng Jason Chen, Mehrdad Mahdavi

CP11

Exact and Heuristic Approaches to Speeding Up the MSM Time Series Distance Computation	451
---	------------

Jana Holznigenkemper, Christian Komusiewicz, Bernhard Seeger

Debiased Imitation Learning for Modulated Temporal Point Processes	460
---	------------

Zhuoqun Li, Zihan Zhou, Mingxuan Sun, Hongteng Xu

Attention-Based Multi-Modal Missing Value Imputation for Time Series Data with High Missing Rate	469
---	------------

Khandakar Tanvir Ahmed, Sudipto Baul, Yanjie Fu, Wei Zhang

Probabilistic Decomposition Transformer for Time Series Forecasting	478
--	------------

Junlong Tong, Liping Xie, Kanjian Zhang

Spatiotemporal Classification with Limited Labels Using Constrained Clustering for Large Datasets	487
--	------------

Praveen Ravirathinam, Rahul Ghosh, Ke Wang, Keyang Xuan, Ankush Khandelwal, Hilary Dugan, Paul Hanson, Vipin Kumar

CP12

Knowledge-Enhanced Semi-Supervised Federated Learning for Aggregating Heterogeneous Lightweight Clients in IoT	496
<i>Jiaqi Wang, Shenglai Zeng, Zewei Long, Yaqing Wang, Houping Xiao, Fenglong Ma</i>	
A Multi-Scale Interaction Motion Network for Action Recognition Based on Capsule Network	505
<i>Xiangping Zheng, Xun Liang, Bo Wu, Jun Wang, Yuhui Guo, Xuan Zhang, Yuefeng Ma</i>	
MoVAE: A Variational AutoEncoder for Molecular Graph Generation	514
<i>Zerun Lin, Yuhan Zhang, Lixin Duan, Le Ou-Yang, Peilin Zhao</i>	
Extrinsic-Intrinsic Representation Learning Framework for Drug Discovery	523
<i>Tian Xia, Sarp Aykent, Wei-Shinn Ku</i>	
Penalized Non-linear Canonical Correlation Analysis for Ordinal Data with Application to the International Classification of Functioning, Disability and Health	532
<i>Jan Gertheiss, Russell Shinohara</i>	

CP13

Domain Disentangled Meta-Learning	541
<i>Xin Zhang, Yanhua Li, Ziming Zhang, Zhi-Li Zhang</i>	
BDA: Bandit-Based Transferable AutoAugment	550
<i>Shan Lu, Mingjun Zhao, Songling Yuan, Xiaoli Wang, Lei Yang, Di Niu</i>	
Adaptive Precision Training (AdaPT): A Dynamic Quantized Training Approach for DNNs	559
<i>Lorenz Kummer, Kevin Sidak, Tabea Reichmann, Wilfried Gansterer</i>	
Robust Learning via Golden Symmetric Loss of (un)Trusted Labels	568
<i>Amirmasoud Ghiassi, Robert Birke, Lydia Y. Chen</i>	
Concept Discovery for Fast Adaptation	577
<i>Shengyu Feng, Hanghang Tong</i>	

CP14

Multi-Task Learning with Prior Information	586
<i>Mengyuan Zhang, Kai Liu</i>	

Optimal Smooth Approximation for Quantile Matrix Factorization 595
Peng Liu, Yi Liu, Rui Zhu, Linglong Kong, Bei Jiang, Di Niu

Ranking with Submodular Functions on the Fly 604
Guangyi Zhang, Nikolaj Tatti, Aristides Gionis

Learning Causal Structure on Mixed Data with Tree-Structured Functional Models 613
Tian Qin, Tian-Zuo Wang, Zhi-Hua Zhou

MC-SQ: A Highly Accurate Ensemble for Multi-Class Quantification 622
Zahra Donyavi, Adriane Serapio, Gustavo Batista

CP15

Taxonomy-Guided Fine-Grained Entity Set Expansion 631
Jinfeng Xiao, Mohab Elkaref, Nathan Herr, Geeth De Mel, Jiawei Han

Balancing Task Coverage and Expert Workload in Team Formation 640
Karan Vombatkere, Evimaria Terzi

Mini-Batch Learning Strategies for Modeling Long Term Temporal Dependencies: A Study in Environmental Applications 649
Shaoming Xu, Ankush Khandelwal, Xiang Li, Xiaowei Jia, Licheng Liu, Jared Willard, Rahul Ghosh, Kelly Cutler, Michael Steinbach, Christopher Duffy, John Nieber, Vipin Kumar

A Linkage-Based Double Imbalanced Graph Learning Framework for Face Clustering 658
Huafeng Yang, Qijie Shen, Xingjian Chen, Fangyi Zhang, Rong Du

NEEDED: Introducing Hierarchical Transformer to Eye Diseases Diagnosis 667
Xu Ye, Meng Xiao, Zhiyuan Ning, Weiwei Dai, Wenjuan Cui, Yi Du, Yuanchun Zhou

CP16

Context-Aware Domain Adaptation for Time Series Anomaly Detection 676
Kwei-Herng Lai, Lan Wang, Huiyuan Chen, Kaixiong Zhou, Fei Wang, Hao Yang, Xia Hu

Matrix Profile XXVIII: Discovering Multi-Dimensional Time Series Anomalies with K of N Anomaly Detection 685
Sadaf Tafazoli, Eamonn Keogh

Deep Contrastive One-Class Time Series Anomaly Detection 694
Rui Wang, Chongwei Liu, Xudong Mou, Kai Gao, Xiaohui Guo, Pin Liu, Tianyu Wo, Xudong Liu

Subgraph Centralization: A Necessary Step in Graph Anomaly Detection 703
Zhong Zhuang, Kai Ming Ting, Guansong Pang, Shuaibin Song

Abnormal Event Detection via Hypergraph Contrastive Learning 712
Bo Yan, Cheng Yang, Chuan Shi, Jiawei Liu, Xiaochen Wang

CP17

A General-Purpose Transferable Predictor for Neural Architecture Search 721
Fred X. Han, Keith G. Mills, Fabian Chudak, Parsa Riahi, Mohammad Salameh, Jialin Zhang, Wei Lu, Shangling Jui, Di Niu

Estimating Propensity Scores with Deep Adaptive Variable Selection 730
Zhixuan Chu, Mechelle Claridy, Jose Cordero, Sheng Li, Stephen L. Rathbun

Scalable Batch Acquisition for Deep Bayesian Active Learning 739
Aleksandr Rubashevskii, Daria Kotova, Maxim Panov

Fisher Scoring Method for Neural Networks Optimization 748
Jackson de Faria, Renato Assunção, Fabricio Murai

Multi-State Survival Analysis Using Pseudo Value-Based Deep Neural Networks 757
Md Mahmudur Rahman, Sanjay Purushotham

CP18

Understand Influence Maximization via Higher-Order Decomposition 766
Zonghan Zhang, Zhiqian Chen

Traceable Automatic Feature Transformation via Cascading Actor-Critic Agents 775
Meng Xiao, Dongjie Wang, Min Wu, Ziyue Qiao, Pengfei Wang, Kunpeng Liu, Yuanchun Zhou, Yanjie Fu

Learning to Learn Task Transformations for Improved Few-Shot Classification 784
Guangtao Zheng, Qiuling Suo, Mengdi Huai, Aidong Zhang

Sign-Regularized Multi-Task Learning 793
Guangji Bai, Johnny Torres, Junxiang Wang, Liang Zhao, Cristina Abad, Carmen Vaca

Toward Theoretical Guidance for Two Common Questions in Practical Cross-Validation based Hyperparameter Selection	802
--	------------

Parikshit Ram, Alexander G. Gray, Horst C. Samulowitz, Gregory Bramble

CP19

Meta-Learning with Motif-based Task Augmentation for Few-Shot Molecular Property Prediction	811
--	------------

Ziqiao Meng, Yaoman Li, Peilin Zhao, Yang Yu, Irwin King

A Novel Reject Option Applied to Sleep Stage Scoring	820
---	------------

Dries Van Der Plas, Wannes Meert, Johan Verbraecken, Jesse Davis

Reinforcement Learning Guided Multi-Objective Exam Paper Generation	829
--	------------

Yuhu Shang, Xuexiong Luo, Lihong Wang, Hao Peng, Xiankun Zhang, Yimeng Ren, Kun Liang

Eco-PiNN: A Physics-informed Neural Network for Eco-toll Estimation	838
--	------------

Yan Li, Mingzhou Yang, Matthew Eagon, Majid Farhadloo, Yiqun Xie, William F. Northrop, Shashi Shekhar

Probabilistic Inverse Modeling: An Application in Hydrology	847
--	------------

Somya Sharma, Rahul Ghosh, Arvind Renganathan, Xiang Li, Snigdhasu Chatterjee, John Nieber, Christopher Duffy, Vipin Kumar

CP20

Fairness-aware Multi-view Clustering	856
---	------------

Lecheng Zheng, Yada Zhu, Jingrui He

Group AdaBoost with Fairness Constraint	865
--	------------

Zhiyu Xue

On Improving Fairness of AI Models with Synthetic Minority Oversampling Techniques	874
---	------------

Yan Zhou, Murat Kantarcioglu, Chris Clifton

TEM: High Utility Metric Differential Privacy on Text	883
--	------------

Ricardo Silva Carvalho, Theodore Vasiloudis, Oluwaseyi Feyisetan, Ke Wang

PMP: Privacy-Aware Matrix Profile against Sensitive Pattern Inference for Time Series	891
--	------------

Li Zhang, Jiahao Ding, Yifeng Gao, Jessica Lin

CP21

Feature Enhanced Zero-Shot Stance Detection via Contrastive Learning	900
<i>Xuechen Zhao, Jiaying Zou, Zhong Zhang, Feng Xie, Bin Zhou, Lei Tian</i>	
GOTTA: Generative Few-Shot Question Answering by Prompt-Based Cloze Data Augmentation	909
<i>Xiusi Chen, Yu Zhang, Jinliang Deng, Jyun-Yu Jiang, Wei Wang</i>	
Coarse-to-Fine Open Information Extraction via Relation Oriented Reading Comprehension	918
<i>Tingxin Li, Rui Meng, Feng Chen, Jianming Wu</i>	
XDC: Adaptive Cross Domain Short Text Clustering	927
<i>Saed Rezayi, Handong Zhao, Ronghang Zhu, Sheng Li</i>	
Hierarchical Neural Topic Model with Embedding Cluster and Neural Variational Inference	936
<i>Ningjing Wang, Deqing Wang, Ting Jiang, Chenguang Du, Chuyu Fang, Fuzhen Zhuang</i>	

CP22

Data-Centric AI: Perspectives and Challenge	945
<i>Daochen Zha, Zaid Pervaiz Bhat, Kwei-Herng Lai, Fan Yang, Xia Hu</i>	
Statistically-Sound Knowledge Discovery from Data	949
<i>Matteo Riondato</i>	
Data Mining Challenges and Opportunities to Achieve Net Zero Carbon Emissions: Focus on Electrified Vehicles	953
<i>Mingzhou Yang, Bharat Jayaprakash, Matthew Eagon, Hyeonjung Tari Jung, William F. Northrop, Shashi Shekhar</i>	
Towards Trustworthy Representation Learning	957
<i>Sheng Li</i>	
Missed Opportunities in Fair AI	961
<i>Nripsuta Ani Saxena, Wenbin Zhang, Cyrus Shahabi</i>	
Making a Computational Attorney	965
<i>Dell Zhang, Frank Schilder, Jack G. Conrad, Masoud Makrehchi, David von Rickenbach, Isabelle Moulinier</i>	