

SIAM International Conference on Data Mining (SDM'22)

Alexandria, Virginia, USA and Online
28 – 30 April 2022

ISBN: 978-1-7138-5209-4

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© (2022) by SIAM: Society for Industrial and Applied Mathematics
All rights reserved.

Printed with permission by Curran Associates, Inc. (2022)

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
Fax: (215) 386-7999

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

Session 1: Foundations I

Self-supervision based Semantic Alignment for Unsupervised Domain Adaptation	1
<i>Ronghang Zhu, Sheng Li</i>	
Meta Adversarial Weight for Unsupervised Domain Adaptation	10
<i>Chang Liu, Lichen Wang, Yun Fu</i>	
Positive Unlabeled Learning with a Sequential Selection Bias	19
<i>Walter Gerych, Tom Hartvigsen, Luke Buquicchio, Abdulaziz Alajaji, Kavin Chandrasekaran, Hamid Mansoor, Elke Rundensteiner, Emmanuel Agu</i>	
Continuously Generalized Ordinal Regression for Linear and Deep Models	28
<i>Fred Lu, Francis Ferraro, Edward Raff</i>	
Entropy Weight Allocation: Positive-unlabeled Learning via Optimal Transport	37
<i>Wen Gu, Teng Zhang, Hai Jin</i>	

Session 2: Graph Mining and Analysis I

DualCast: Friendship-Preference Co-evolution Forecasting for Attributed Networks	46
<i>Hiroyoshi Ito, Christos Faloutsos</i>	
Neighbor2Seq: Deep Learning on Massive Graphs by Transforming Neighbors to Sequences	55
<i>Meng Liu, Shuiwang Ji</i>	
Leveraging Contextual Graphs for Stochastic Weight Completion in Sparse Road Networks	64
<i>Xiaolin Han, Reynold Cheng, Tobias Grubenmann, Silviu Maniu, Chenhao Ma, Xiaodong Li</i>	
Interpretable Molecular Graph Generation via Monotonic Constraints	73
<i>Yuanqi Du, Xiaojie Guo, Amarda Shehu, Liang Zhao</i>	
Structure-Enhanced Heterogeneous Graph Contrastive Learning	82
<i>Yanqiao Zhu, Yichen Xu, Hejie Cui, Carl Yang, Qiang Liu, Shu Wu</i>	

Session 3: Applications I

Modeling Reservoir Release Using Pseudo-Pro prospective Learning and Physical Simulations to Predict Water Temperature	91
---	----

Xiaowei Jia, Shengyu Chen, Yiqun Xie, Haoyu Yang, Alison Appling, Samantha Oliver, Zhe Jiang

Early Forecasting of the Impact of Traffic Accidents Using a Single Shot Observation	100
---	-----

Guangyu Meng, Qisheng Jiang, Kaiqun Fu, Beiyu Lin, Chang-Tien Lu, Zhqian Chen

On Predicting and Generating a Good Break Shot in Billiards Sports	109
---	-----

Qianru Zhang, Zheng Wang, Cheng Long, Siu-Ming Yiu

ConLearn: Contextual-knowledge-aware Concept Prerequisite Relation Learning with Graph Neural Network	118
--	-----

Hao Sun, Yuntao Li, Yan Zhang

Hypergraph Contrastive Learning for Electronic Health Records	127
--	-----

Derun Cai, Chenxi Sun, Moxian Song, Baofeng Zhang, Shenda Hong, Hongyan Li

Session 4: Graph Mining and Analysis II

SAGA: Signal-Aware Graph Aggregation	136
---	-----

Maxwell McNeil, Boya Ma, Petko Bogdanov

Diverse and Experienced Group Discovery via Hypergraph Clustering	145
--	-----

Ilya Amburg, Nate Veldt, Austin R. Benson

Transition Matrix Representation of Trees with Transposed Convolutions	154
---	-----

Jaemin Yoo, Lee Sael

On the Persistence of Higher-Order Interactions in Real-World Hypergraphs	163
--	-----

Hyunjin Choo, Kijung Shin

Neural Graph Matching for Pre-training Graph Neural Networks	172
---	-----

Yupeng Hou, Binbin Hu, Wayne Xin Zhao, Zhiqiang Zhang, Jun Zhou, Ji-Rong Wen

Session 5: Time Series I

Error-bounded Approximate Time Series Joins using Compact Dictionary Representations of Time Series	181
--	-----

Chin-Chia Michael Yeh, Yan Zheng, Junpeng Wang, Huiyuan Chen, Zhongfang Zhuang, Wei Zhang, Eamonn Keogh

Learning Time-series Shapelets Enhancing Discriminability	190
<i>Akihiro Yamaguchi, Ken Ueno, Hisashi Kashima</i>	
Towards Similarity-Aware Time-Series Classification	199
<i>Daochen Zha, Kwei-Herng Lai, Kaixiong Zhou, Xia Hu</i>	
Joint Time Series Chain: Detecting Unusual Evolving Trend across Time Series	208
<i>Li Zhang, Nital Patel, Xiuqi Li, Jessica Lin</i>	
IB-GAN: A Unified Approach for Multivariate Time Series Classification under Class Imbalance	217
<i>Grace Deng, Cuize Han, Tommaso Dreossi, Clarence Lee, David S. Matteson</i>	
 Session 6: Clustering and Anomaly Detection	
Automatic Parameter Selection for Non-Redundant Clustering	226
<i>Collin Leiber, Dominik Mautz, Claudia Plant , Christian Böhm</i>	
Constrained Mean Shift Clustering	235
<i>Maximilian Schier, Christoph Reinders, Bodo Rosenhahn</i>	
SSSNET: Semi-Supervised Signed Network Clustering	244
<i>Yixuan He, Gesine Reinert, Songchao Wang, Mihai Cucuringu</i>	
Multi-class Label Noise Learning via Loss Decomposition and Centroid Estimation	253
<i>Yongliang Ding, Tao Zhou, Chuang Zhang, Yijing Luo, Juan Tang, Chen Gong</i>	
Unseen Anomaly Detection on Networks via Multi-Hypersphere Learning	262
<i>Shuang Zhou, Xiao Huang, Ninghao Liu, Qiaoyu Tan, Fu-Lai Chung</i>	
 Session 7: Deep Learning I	
Capturing Model Uncertainty with Data Augmentation in Deep Learning	271
<i>Wenming Jiang, Ying Zhao, Yihan Wu, Haojia Zuo</i>	
Input Selection for Bandwidth-Limited Neural Network Inference	280
<i>Stefan Oehmcke, Fabian Gieseke</i>	
Hybrid Tokenization and Datasets for Solving Mathematics and Science Problems Using Transformers	289
<i>Pratik Mandlecha, Snehith Kumar Chatakonda, Neeraj Kollepara, Pawan Kumar</i>	

GANBLR++: Incorporating Capacity to Generate Numeric Attributes and Leveraging Unrestricted Bayesian Networks	298
<i>Yishuo Zhang, Nayyar Zaidi, Jiahui Zhou, Gang Li</i>	
Flexible Image Captioning via Internal Understanding and External Reasoning	307
<i>Jiahui Wei, Zhixin Li, Jianwei Zhu, Huifang Ma</i>	
Session 8: Applications II	
Multiperiod Corporate Default Prediction Through Neural Parametric Family Learning	316
<i>Wei-Lun Luo, Yu-Ming Lu, Jheng-Hong Yang, Jin-Chuan Duan, Chuan-Ju Wang</i>	
Efficient Wind Speed Nowcasting with GPU-Accelerated Nearest Neighbors Algorithm	325
<i>Arnaud Pannatier, Ricardo Picatoste, François Fleuret</i>	
HintNet: Hierarchical Knowledge Transfer Networks for Traffic Accident Forecasting on Heterogeneous Spatio-Temporal Data	334
<i>Bang An, Amin Vahedian, Xun Zhou, W. Nick Street, Yanhua Li</i>	
BRNet: Branched Residual Network for Fast and Accurate Predictive Modeling of Materials Properties	343
<i>Vishu Gupta, Wei-keng Liao, Alok Choudhary, Ankit Agrawal</i>	
Textbook Enhanced Student Learning Outcome Prediction	352
<i>Tianqi Wang, Fenglong Ma, Tang Tang, Longfei Zhang, Jing Gao</i>	
Session 9: Graph Mining and Analysis III	
Dense and Well-Connected Subgraph Detection in Dual Networks	361
<i>Tianyi Chen, Francesco Bonchi, David Garcia-Soriano, Atsushi Miyauchi, Charalampos E. Tsourakakis</i>	
Fine-grained Attributed Graph Clustering	370
<i>Zhao Kang, Zhanyu Liu, Shirui Pan, Ling Tian</i>	
Computation of the Sample Fréchet Mean for Sets of Large Graphs with Applications to Regression	379
<i>Daniel Ferguson, François G. Meyer</i>	
Classic Graph Structural Features Outperform Factorization-Based Graph Embedding Methods on Community Labeling	388
<i>Andrew Stolman, Caleb Levyy, C. Seshadhriz, Aneesh Sharmax</i>	

HG-Meta: Graph Meta-learning over Heterogeneous Graphs	397
<i>Qiannan Zhang, Xiaodong Wu, Qiang Yang, Chuxu Zhang, Xiangliang Zhang</i>	
Session 10: Causal Learning	
Multi-interest Sequence Modeling for Recommendation with Causal Embedding	406
<i>Caiqi Sun, Penghao Lu, Lei Cheng, Zhenfu Cao, Xiaolei Dong, Yili Tang, Jun Zhou, Linjian Mo</i>	
Estimating Mutual Information via Geodesic kNN	415
<i>Alexander Marx, Jonas Fischer</i>	
Masked Gradient-Based Causal Structure Learning	424
<i>Ignavier Ng, Shengyu Zhu, Zhuangyan Fang, Haoyang Li, Zhitang Chen, Jun Wang</i>	
Learning Infomax and Domain-Independent Representations for Causal Effect Inference with Real-World Data	433
<i>Zhixuan Chu, Stephen L. Rathbun, Sheng Li</i>	
Cycle-Balanced Representation Learning For Counterfactual Inference	442
<i>Guanglin Zhou, Lina Yao, Xiwei Xu, Chen Wang, Liming Zhu</i>	
Session 11: Optimization	
GRP-FED: Addressing Client Imbalance in Federated Learning via Global-Regularized Personalization	451
<i>Yen-Hsiu Chou, Shenda Hong, Chenxi Sun, Derun Cai, Moxian Song, Hongyan Li</i>	
FedMe: Federated Learning via Model Exchange	459
<i>Koji Matsuda, Yuya Sasaki, Chuan Xiao, Makoto Onizuka</i>	
Noisy Truncated SGD: Optimization and Generalization	468
<i>Yingxue Zhou, Xinyan Li, Arindam Banerjee</i>	
StepDIRECT - A Derivative-Free Optimization Method for Stepwise Functions	477
<i>Dzung T. Phan¹, Hongsheng Liu¹, Lam M. Nguyen</i>	
Adaptive Principal Component Analysis	486
<i>Xiangyu Li, Hua Wang</i>	

Session 12: Time Series II

Collaborative Attention Mechanism for Multi-Modal Time Series Classification 495

Yue Bai, Zhiqiang Tao, Lichen Wang, Sheng Li, Yu Yin, Yun Fu

Leveraging Dependencies among Learned Temporal Subsequences 504

Shoumik Roychoudhury, Fang Zhou, Zoran Obradovic

Contextual Bandits for Advertising Campaigns: A Diffusion-Model Independent Approach 513

Alexandra Iacob, Bogdan Cautis, Silviu Maniu

Measuring Disentangled Generative Spatio-Temporal Representation 522

Sichen Zhao, Wei Shao, Jeffrey Chan, Flora D. Salim

Session 13: Recommender Systems

A Non-sequential Approach to Deep User Interest Model for CTR Prediction 531

Keke Zhao, Xing Zhaoy, Qi Cao, Linjian Mo

Localized Graph Collaborative Filtering 540

Yiqi Wang, Chaozhuo Li, Mingzheng Li, Wei Jin, Yuming Liu, Hao Sun, Xing Xie, Jiliang Tang

Private Recommender Systems: How Can Users Build Their Own Fair Recommender Systems without Log Data? 549

Ryoma Sato

Time-aware Multi-interest Capsule Network for Sequential Recommendation 558

Zhikai Wang, Yanyan Shen

Cyclic Transfer Learning for Recommender Systems with Heterogeneous Feedbacks 567

Xuelian Ni, Fei Xiong, Yutian Hu , Shirui Pan , Hongshu Chen, Liang Wang

Session 14: Foundations II

Sensitivity-aware Distance Measurement for Boosting Metric Learning 576

Haoliang Liu, Tan Yu, Ping Li

Exceptional Model Mining for Repeated Cross-Sectional Data (EMM-RCS) 585

Rianne Margaretha Schouten, Wouter Duivesteijn, Mykola Pechenizkiy

Learning from Imbalanced Crowdsourced Labeled Data	594
<i>Wentao Wang, Joseph Thekinen, Xiaorui Liu, Zitao Liu, Jiliang Tang</i>	
Deperturbation of Online Social Networks via Bayesian Label Transition	603
<i>Jun Zhuang, Mohammad Al Hasan</i>	
Improved Knowledge Distillation via Full Kernel Matrix Transfer	612
<i>Qi Qian, Hao Li, Juhua Hu</i>	
Session 15: Text/NLP	
Mining Logical Arithmetic Expressions From Proper Representations	621
<i>Eitan Kosman, Ilya Kolchinsky, Assaf Schuster</i>	
CPTAM: Constituency Parse Tree Aggregation Method	630
<i>Adithya Kulkarni, Nasim Sabetpour, Alexey Markin, Oliver Eulenstein, Qi Li</i>	
Knowledge-enhanced Spherical Representation Learning for Text Classification	639
<i>Hafsa Ennajari, Nizar Bouguila, Jamal Bentahar</i>	
Session 16: Graph Mining and Analysis IV	
A Strong Node Classification Baseline for Temporal Graphs	648
<i>Farimah Poursafaei, Zeljko Zilic, Reihaneh Rabbany</i>	
Heterogeneous Temporal Graph Neural Network	657
<i>Yujie Fan, Mingxuan Ju, Chuxu Zhang, Yanfang Ye</i>	
Adapting Distilled Knowledge for Few-shot Relation Reasoning over Knowledge Graphs	666
<i>Yiming Zhang, Yiyue Qian, Yanfang Yey, Chuxu Zhang</i>	
Explainable Concept Graph Completion by Bridging Open-Domain Relations and Concepts	675
<i>Yue Zhang, Mingming Sun, Jingyuan Zhang, Ping Li</i>	
YACC: A Framework Generalizing TuránShadow for Counting Large Cliques	684
<i>Shweta Jain, Hanghang Tong</i>	

Session 17: Applications III

Adversarially Reprogramming Pretrained Neural Networks for Data-limited and Cost-efficient Malware Detection 693

Lingwei Chen, Xiaoting Li, Dinghao Wu

Invertibility aware Integration of Static and Time-series data: An application to Lake Temperature Modeling 702

Kshitij Tayal, Xiaowei Jia, Rahul Ghosh, Jared Willard, Jordan Read, Vipin Kumar

Unifying Human Mobility Forecasting and Trajectory Semantics Augmentation via Hawkes Process Based LSTM 711

Pengfei Wang, Kunpeng Liu, Yuanchun Zhou, Yanjie Fu

FAME: Fragment-based Conditional Molecular Generation for Phenotypic Drug Discovery 720

Thai-Hoang Pham, Lei Xie, Ping Zhang

DeepGLSTM: Deep Graph Convolutional Network and LSTM Based Approach for Predicting Drug-Target Binding Affinity 729

Shrimon Mukherjee, Madhusudan Ghosh, Partha Basuchowdhuri