

SIAM International Conference on Data Mining (SDM19)

Calgary, Canada
2-4 May 2019

Editors:

Tanya Berger-Wolf

Nitesh Chawla

ISBN: 978-1-5108-8593-6

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© (2019) by SIAM: Society for Industrial and Applied Mathematics
All rights reserved.

Printed by Curran Associates, Inc. (2019)

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

Incremental Multi-view Support Vector Machine	1
<i>Peng Zhou, Yi-Dong Shen, Liang Du, Fan Ye</i>	
Deep Multimodality Model for Multi-task Multi-view Learning	10
<i>Lecheng Zheng, Yu Cheng, Jingrui He</i>	
Augmented Multi-Task Learning by Optimal Transport	19
<i>Boyang Liu, Pang-Ning Tan, Jiayu Zhou</i>	
Calibrating Probability Estimation Trees using Venn-Abers Predictors.....	28
<i>Ulf Johansson, Tuwe Löfström, Henrik Boström</i>	
Deep Multi-view Information Bottleneck	37
<i>Qi Wang, Claire Boudreau, Qixing Luo, Pang-Ning Tan, Jiayu Zhou</i>	
Forest Packing: Fast Parallel, Decision Forests	46
<i>James Browne, Disa Mhembere, Tyler M. Tomita, Joshua T. Vogelstein, Randal Burns</i>	
Fast Unsupervised Location Category Inference from Highly Inaccurate Mobility Data.....	55
<i>Jinfeng Yi, Qi Lei, Wesley M. Gifford, Ji Liu, Junchi Yan, Bowen Zhou</i>	
GeoAttn: Localization of Social Media Messages via Attentional Memory Network	64
<i>Sha Li, Chao Zhang, Dongming Lei, Ji Li, Jiawei Han</i>	
Bus Travel Speed Prediction using Attention Network of Heterogeneous Correlation Features	73
<i>Yidan Sun, Guiyuan Jiang, Siew-Kei Lam, Shicheng Chen, Peilan He</i>	
A Nondeterministic Normalization Based Scan Statistic (NN-scan) towards Robust Hotspot Detection: A Summary of Results.....	82
<i>Yiqun Xie, Shashi Shekhar</i>	
A Fast Algorithm for Combinatorial Hotspot Mining Based on Spatial Scan Statistic	91
<i>Shin-Ichi Minato, Jun Kawahara, Fumio Ishioka, Masahiro Mizuta, Koji Kurihara</i>	
A Deep Spatio-Temporal Fuzzy Neural Network for Passenger Demand Prediction.....	100
<i>Xiaoyuan Liang, Guiling Wang, Martin Renqiang Min, Yi Qi, Zhu Han</i>	
AMAS: Attention Model for Attributed Sequence Classification.....	109
<i>Zhongfang Zhuang, Xiangnan Kong, Elke Rundensteiner</i>	
MM-Pred: A Deep Predictive Model for Multi-attribute Event Sequence	118
<i>Li Lin, Lijie Wen, Jianmin Wang</i>	
Constraint Programming for Association Rules	127
<i>Mohamed-Bachir Belaid, Christian Bessiere, Nadjib Lazaar</i>	
Linear Time Motif Discovery in Time Series.....	136
<i>Xiaosheng Li, Jessica Lin</i>	
Deep Skip-Gram Networks for Text Classification.....	145
<i>Chaochun Liu, Yaliang Li, Hongliang Fei, Ping Li</i>	
Mining Rules Incrementally Over Large Knowledge Bases.....	154
<i>Xiaofeng Zhou, Ali Sadeghian, Daisy Zhe Wang</i>	
Optimally Rotated Coordinate Systems for Adaptive Least-squares Regression on Sparse Grids	163
<i>Bastian Bohn, Michael Griebel, Jens Oettershagen</i>	
ℓ_0-Regularized Sparsity for Probabilistic Mixture Models.....	172
<i>Dzung T. Phan, Tsuyoshi Idé</i>	
Intrinsic Dimensionality Estimation within Tight Localities	181
<i>Laurent Amsaleg, Oussama Chelly, Michael E. Houle, Ken-Ichi Kawarabayashi, Miloš Radovanovic, Weeris Treeratanajaru</i>	
Identifying When Effect Restoration Will Improve Estimates of Causal Effect	190
<i>Hüseyin Oktay, Akanksha Atrey, David Jensen</i>	
We Are Not Your Real Parents: Telling Causal from Confounded using MDL	199
<i>David Kaltenpoth, Jilles Vreeken</i>	
Query-Aware Bayesian Committee Machine for Scalable Gaussian Process Regression	208
<i>Jiayuan He, Jianzhong Qi, Kotagiri Ramamohanarao</i>	
A Novel Regularizer for Temporally Stable Learning with an Application to Twitter Topic Classification	217
<i>Yakun Wang, Ga Wu, Mohamed Reda Bouadjenek, Scott Sanner, Sen Su, Zhongbao Zhang</i>	
Efficient Summarizing of Evolving Events from Twitter Streams	226
<i>Thi-Huyen Nguyen, Tuan-Anh Hoang, Wolfgang Nejdl</i>	
Hierarchical Attention Networks for Cyberbullying Detection on the Instagram Social Network	235
<i>Lu Cheng, Ruocheng Guo, Yasin Silva, Deborah Hall, Huan Liu</i>	

SPMM: A Soft Piecewise Mapping Model for Bilingual Lexicon Induction	244
<i>Yan Fan, Chengyu Wang, Boxing Chen, Zhongkai Hu, Xiaofeng He</i>	
Aspect-Based Sentiment Analysis with Minimal Guidance	253
<i>Honglei Zhuang, Timothy Hanratty, Jiawei Har</i>	
Joint Post and Link-level Influence Modeling on Social Media	262
<i>Liangzhe Chen, B. Aditya Prakash</i>	
Positive-Unlabeled Classification under Class Prior Shift and Asymmetric Error	271
<i>Nontawat Charoenphakdee, Masashi Sugiyama</i>	
Fast Training for Large-Scale One-versus-All Linear Classifiers using Tree-Structured Initialization	280
<i>Huang Fang, Minhao Cheng, Cho-Jui Hsieh, Michael Friedlander</i>	
Towards Identifying Causal Relation Between Instances and Labels	289
<i>Tian-Zuo Wang, Sheng-Jun Huang, Zhi-Hua Zhou</i>	
Discovering and Controlling for Latent Confounds in Text Classification Using Adversarial Domain Adaptation	298
<i>Virgile Landeiro, Tuan Tran, Aron Culotta</i>	
A Bayesian Approach for Accurate Classification-Based Aggregates	306
<i>Q. A. Meertens, C. G. H. Diks, H. J. Van Den Herik, F. W. Takes</i>	
Decision List Optimization Based on Continuous Relaxation	315
<i>Yuzuru Okajima, Kunihiko Sadamasa</i>	
Mining Maximal Induced Bicliques using Odd Cycle Transversals	324
<i>Kyle Kloster, Andrew Van Der Poel, Blair D. Sullivan</i>	
DSL: Discriminative Subgraph Learning via Sparse Self-Representation	333
<i>Lin Zhang, Petko Bogdanov</i>	
Relationship Profiling Over Social Networks: Reverse Smoothness from Similarity to Closeness	342
<i>Carl Yang, Kevin Chang</i>	
Edge Replacement Grammars: A Formal Language Approach for Generating Graphs	351
<i>Revanth Reddy, Sarath Chandar, Balaraman Ravindran</i>	
Temporal Graph Regression via Structure-Aware Intrinsic Representation Learning	360
<i>Chao Han, Xi Hang Cao, Marija Stanojevic, Mohamed Ghalwash, Zoran Obradovic</i>	
Multi-Dimensional, Multilayer, Nonlinear and Dynamic HITS	369
<i>Francesca Arrigo, Francesco Tudisco</i>	
Flow-Based Local Graph Clustering with Better Seed Set Inclusion	378
<i>Nate Veldt, Christine Klymko, David F. Gleich</i>	
Spherical Principal Component Analysis	387
<i>Kai Liu, Qiuwei Li, Hua Wang, Gongguo Tang</i>	
Auto-Summarization: A Step Towards Unsupervised Learning of a Submodular Mixture	396
<i>Chandrashekhar Lavania, Jeff Bilmes</i>	
Learning Diverse Gaussian Graphical Models and Interpreting Edges	405
<i>Roosbeh Yousefzadeh, Diane Oyen, Nina Lanza</i>	
Deep Co-Clustering	414
<i>Dongkuan Xu, Wei Cheng, Bo Zong, Jingchao Ni, Dongjin Song, Wenchao Yu, Yuncong Chen, Haifeng Chen, Xiang Zhang</i>	
Discovering Multiple Co-Clusterings in Subspaces	423
<i>Shixin Yao, Guoxian Yu, Xing Wang, Jun Wang, Carlotta Domeniconi, Maozu Guo</i>	
Accelerated Experimental Design for Pairwise Comparisons	432
<i>Yuan Guo, Jennifer Dy, Deniz Erdogmus, Jayashree Kalpathy-Cramer, Susan Ostmo, J. Peter Campbell, Michael F. Chiang, Stratis Ioannidis</i>	
Region-Based Active Learning with Hierarchical and Adaptive Region Construction	441
<i>Zhipeng Luo, Milos Hauskrecht</i>	
Feature Selection as Monte-Carlo Search in Growing Single Rooted Directed Acyclic Graph by Best Leaf Identification	450
<i>Aurélien Pélassier, Atsuyoshi Nakamura, Koji Tabata</i>	
Relative Distance Comparisons with Confidence Judgements	459
<i>Stefan Mojsilovic, Antti Ukkonen</i>	
Doubly Robust Prediction and Evaluation Methods Improve Uplift Modeling for Observational Data	468
<i>Yuta Saito, Hayato Sakata, Kazuhide Nakata</i>	
EstImAgg: A Learning Framework for Groupwise Aggregated Data	477
<i>Avradeep Bhowmik, Minmin Chen, Zhengming Xing, Suju Rajan</i>	
DTEC: Distance Transformation Based Early Time Series Classification	486
<i>Liuyi Yao, Yaliang Li, Yezheng Li, Hengtong Zhang, Mengdi Huai, Jing Gao, Aidong Zhang</i>	
Leveraging Subsequence-orders for Univariate and Multivariate Time-series Classification	495
<i>Shoumik Roychoudhury, Fang Zhou, Zoran Obradovic</i>	

Branch and Border: Partition-Based Change Detection in Multivariate Time Series	504
<i>Bryan Hooi, Christos Faloutsos</i>	
Spatial Context-Aware Networks for Mining Temporal Discriminative Period in Land Cover Detection	513
<i>Xiaowei Jia, Sheng Li, Ankush Khandelwal, Guruprasad Nayak, Anuj Karpatne, Vipin Kumar</i>	
Elastic Bands Across the Path: A New Framework and Method to Lower Bound DTW	522
<i>Chang Wei Tan, François Petitjean, Geoffrey I. Webb</i>	
Coupled Clustering of Time-Series and Networks	531
<i>Yike Liu, Linhong Zhu, Pedro Szekely, Aram Galstyan, Danai Koutra</i>	
Classifying Heterogeneous Sequential Data by Cyclic Domain Adaptation: An Application in Land Cover Detection	540
<i>Xiaowei Jia, Guruprasad Nayak, Ankush Khandelwal, Anuj Karpatne, Vipin Kumar</i>	
Incomplete Conditional Density Estimation for Fast Materials Discovery	549
<i>Phuoc Nguyen, Truyen Tran, Sunil Gupta, Santu Rana, Matthew Barnett, Svetha Venkatesh</i>	
Physics Guided RNNs for Modeling Dynamical Systems: A Case Study in Simulating Lake Temperature Profiles	558
<i>Xiaowei Jia, Jared Willard, Anuj Karpatne, Jordan Read, Jacob Zwart, Michael Steinbach, Vipin Kumar</i>	
Reading Customer Reviews to Answer Product-related Questions	567
<i>Miao Fan, Chao Feng, Mingming Sun, Ping Li, Haifeng Wang</i>	
An Interpretable Fast Model for Predicting The Risk of Heart Failure	576
<i>Xianli Zhang, Buyue Qian, Xiaoyu Li, Jishang Wei, Yingqian Zheng, Lingyun Song, Qinghua Zheng</i>	
LSCP: Locally Selective Combination in Parallel Outlier Ensembles	585
<i>Yue Zhao, Zain Nasrullah, Maciej K. Hryniewicki, Zheng Li</i>	
Deep Anomaly Detection on Attributed Networks	594
<i>Kaize Ding, Jundong Li, Rohit Bhanushali, Huan Liu</i>	
ABACUS: Unsupervised Multivariate Change Detection via Bayesian Source Separation	603
<i>Wenyu Zhang, Daniel Gilbert, David S. Matteson</i>	
Learning On-the-Job to Re-rank Anomalies from Top-1 Feedback	612
<i>Hemank Lamba, Leman Akoglu</i>	
SMF: Drift-Aware Matrix Factorization with Seasonal Patterns	621
<i>Bryan Hooi, Kijung Shin, Shenghua Liu, Christos Faloutsos</i>	
Multi-Stage Variational Auto-Encoders for Coarse-to-Fine Image Generation	630
<i>Lei Cai, Hongyang Gao, Shuiwang Ji</i>	
Path-based Attribute-aware Representation Learning for Relation Prediction	639
<i>Ying Shen, Desi Wen, Yaliang Li, Nan Du, Hai-Tao Zheng, Min Yang</i>	
Spatial Variational Auto-Encoding via Matrix-Variate Normal Distributions	648
<i>Zhengyang Wang, Hao Yuan, Shuiwang Ji</i>	
Multi-dimensional Graph Convolutional Networks	657
<i>Yao Ma, Suhang Wang, Chara C. Aggarwal, Dawei Yin, Jiliang Tang</i>	
Autonomous Deep Learning: Continual Learning Approach for Dynamic Environments	666
<i>Andri Ashfahani, Mahardhika Pratama</i>	
Deep Transfer Reinforcement Learning for Text Summarization	675
<i>Yaser Keneshloo, Naren Ramakrishnan, Chandan K. Reddy</i>	
TMSA: A Mutual Learning Model for Topic Discovery and Word Embedding	684
<i>Dingcheng Li, Jingyuan Zhang, Ping Li</i>	
Attentional Heterogeneous Graph Neural Network: Application to Program Reidentification	693
<i>Shen Wang, Zhengzhang Chen, Ding Li, Zhichun Li, Lu-An Tang, Jingchao Ni, Junghwan Rhee, Haifeng Chen, Philip S. Yu</i>	
GPU Accelerated Sub-Sampled Newton's Method for Convex Classification Problems	702
<i>Sudhir Kylasa, Fred (Farbod) Roosta, Michael W. Mahoney, Ananth Grama</i>	
CP Tensor Decomposition with Cannot-Link Intermod Constraints	711
<i>Jette Henderson, Bradley A Malin, Joshua C Denny, Abel N Kho, Jimeng Sun, Joydeep Ghosh, Joyce C Ho</i>	
Find the Dimension That Counts: Fast Dimension Estimation and Krylov PCA	720
<i>Shashanka Ubaru, Abd-Krim Seghouane, Yousef Saad</i>	
Boolean Matrix Factorization Meets Consecutive Ones Property	729
<i>Nikolaj Tatti, Pauli Miettinen</i>	
Robust Factorization Machine: A Doubly Capped Norms Minimization	738
<i>Chenghao Liu, Teng Zhang, Jundong Li, Jianwen Yin, Peilin Zhao, Jianling Sun, Steven C. H. Hoi</i>	
Collaborative Filtering with Noisy Ratings	747
<i>Dongsheng Li, Chao Chen, Zhilin Gong, Tun Lu, Stephen M. Chu, Ning Gu</i>	
A Full Probabilistic Model for Yes/No Type Crowdsourcing in Multi-Class Classification	756
<i>Belen Saldias-Fuentes, Pavlos Protopapas, Karim Pichara</i>	

Predicting Multiple Demographic Attributes with Task Specific Embedding Transformation and Attention Network	765
<i>Raehyun Kim, Hyunjae Kim, Janghyuk Lee, Jaewoo Kang</i>	
Semantics-Aware Hidden Markov Model for Human Mobility	774
<i>Hongzhi Shi, Hancheng Cao, Xiangxin Zhou, Yong Li, Chao Zhang, Vassilis Kostakos, Funing Sun, Fanchao Meng</i>	
Dissecting the Learning Curve of Taxi Drivers: A Data-Driven Approach.....	783
<i>Menghai Pan, Yanhua Li, Xun Zhou, Zhenming Liu, Rui Song, Hui Lu, Jun Luo</i>	
Recommendations for Optimizing the Collective User Experience	792
<i>Behzad Golshan, Evimaria Terzi, Panayiotis Tsaparas</i>	
Fairness in Representation: Quantifying Stereotyping As a Representational Harm	801
<i>Mohsen Abbasi, Sorelle A. Friedler, Carlos Scheidegger, Suresh Venkatasubramanian</i>	
Author Index	