

Contents – Part I

Classification

Joint Classification with Heterogeneous Labels Using Random Walk with Dynamic Label Propagation	3
<i>Yongxin Liao, Shenxi Yuan, Jian Chen, Qingyao Wu, and Bin Li</i>	
Hybrid Sampling with Bagging for Class Imbalance Learning	14
<i>Yang Lu, Yiu-ming Cheung, and Yuan Yan Tang</i>	
Sparse Adaptive Multi-hyperplane Machine	27
<i>Khanh Nguyen, Trung Le, Vu Nguyen, and Dinh Phung</i>	
Exploring Heterogeneous Product Networks for Discovering Collective Marketing Hying Behavior	40
<i>Qinzhe Zhang, Qin Zhang, Guodong Long, Peng Zhang, and Chengqi Zhang</i>	
Optimal Training and Efficient Model Selection for Parameterized Large Margin Learning	52
<i>Yuxun Zhou, Jae Yeon Baek, Dan Li, and Costas J. Spanos</i>	
Locally Weighted Ensemble Learning for Regression	65
<i>Man Yu, Zongxia Xie, Hong Shi, and Qinghua Hu</i>	
Reliable Confidence Predictions Using Conformal Prediction	77
<i>Henrik Linusson, Ulf Johansson, Henrik Boström, and Tuve Löfström</i>	
Grade Prediction with Course and Student Specific Models	89
<i>Agoritsa Polyzou and George Karypis</i>	
Flexible Transfer Learning Framework for Bayesian Optimisation	102
<i>Tinu Theckel Joy, Santu Rana, Sunil Kumar Gupta, and Svetha Venkatesh</i>	
A Simple Unlearning Framework for Online Learning Under Concept Drifts	115
<i>Sheng-Chi You and Hsuan-Tien Lin</i>	
User-Guided Large Attributed Graph Clustering with Multiple Sparse Annotations	127
<i>Jianping Cao, Senzhang Wang, Fengcai Qiao, Hui Wang, Feiyue Wang, and Philip S. Yu</i>	

Early-Stage Event Prediction for Longitudinal Data	139
<i>Mahtab J. Fard, Sanjay Chawla, and Chandan K. Reddy</i>	
Toxicity Prediction in Cancer Using Multiple Instance Learning in a Multi-task Framework	152
<i>Cheng Li, Sunil Gupta, Santu Rana, Wei Luo, Svetha Venkatesh, David Ashely, and Dinh Phung</i>	
Shot Boundary Detection Using Multi-instance Incremental and Decremental One-Class Support Vector Machine	165
<i>Hanhe Lin, Jeremiah D. Deng, and Brendon J. Woodford</i>	
Will I Win Your Favor? Predicting the Success of Altruistic Requests	177
<i>Hsun-Ping Hsieh, Rui Yan, and Cheng-Te Li</i>	
Feature Extraction and Pattern Mining	
Unsupervised and Semi-supervised Dimensionality Reduction with Self-Organizing Incremental Neural Network and Graph Similarity Constraints.	191
<i>Zhiyang Xiang, Zhu Xiao, Yourong Huang, Dong Wang, Bin Fu, and Wenjie Chen</i>	
Cross-View Feature Hashing for Image Retrieval	203
<i>Wei Wu, Bin Li, Ling Chen, and Chengqi Zhang</i>	
Towards Automatic Generation of Metafeatures	215
<i>Fábio Pinto, Carlos Soares, and João Mendes-Moreira</i>	
Hash Learning with Convolutional Neural Networks for Semantic Based Image Retrieval	227
<i>Jinma Guo, Shifeng Zhang, and Jianmin Li</i>	
Bayesian Group Feature Selection for Support Vector Learning Machines . . .	239
<i>Changde Du, Changying Du, Shandian Zhe, Ali Luo, Qing He, and Guoping Long</i>	
Active Distance-Based Clustering Using K-Medoids	253
<i>Amin Aghaee, Mehrdad Ghadiri, and Mahdieh Soleymani Baghshah</i>	
Analyzing Similarities of Datasets Using a Pattern Set Kernel.	265
<i>A. Ibrahim, P.S. Sastry, and Shivakumar Sastry</i>	
Significant Pattern Mining with Confounding Variables	277
<i>Aika Terada, David duVerle, and Koji Tsuda</i>	

Building Compact Lexicons for Cross-Domain SMT by Mining Near-Optimal Pattern Sets	290
<i>Pankaj Singh, Ashish Kulkarni, Himanshu Ojha, Vishwajeet Kumar, and Ganesh Ramakrishnan</i>	
Forest CERN: A New Decision Forest Building Technique	304
<i>Md. Nasim Adnan and Md. Zahidul Islam</i>	
Sparse Logistic Regression with Logical Features	316
<i>Yuan Zou and Teemu Roos</i>	
A Nonlinear Label Compression and Transformation Method for Multi-label Classification Using Autoencoders	328
<i>Jörg Wicker, Andrey Tyukin, and Stefan Kramer</i>	
Preconditioning an Artificial Neural Network Using Naive Bayes	341
<i>Nayyar A. Zaidi, François Petitjean, and Geoffrey I. Webb</i>	
OCEAN: Fast Discovery of High Utility Occupancy Itemsets	354
<i>Bilong Shen, Zhaoduo Wen, Ying Zhao, Dongliang Zhou, and Weimin Zheng</i>	

Graph and Network Data

Leveraging Emotional Consistency for Semi-supervised Sentiment Classification	369
<i>Minh Luan Nguyen</i>	
The Effect on Accuracy of Tweet Sample Size for Hashtag Segmentation Dictionary Construction	382
<i>Laurence A.F. Park and Glenn Stone</i>	
Social Identity Link Across Incomplete Social Information Sources Using Anchor Link Expansion	395
<i>Yuxiang Zhang, Lulu Wang, Xiaoli Li, and Chunjing Xiao</i>	
Discovering the Network Backbone from Traffic Activity Data.	409
<i>Sanjay Chawla, Kiran Garimella, Aristides Gionis, and Dominic Tsang</i>	
A Fast and Complete Enumeration of Pseudo-Cliques for Large Graphs.	423
<i>Hongjie Zhai, Makoto Haraguchi, Yoshiaki Okubo, and Etsuji Tomita</i>	
Incorporating Heterogeneous Information for Mashup Discovery with Consistent Regularization	436
<i>Yao Wan, Liang Chen, Qi Yu, Tingting Liang, and Jian Wu</i>	
Link Prediction in Schema-Rich Heterogeneous Information Network	449
<i>Xiaohuan Cao, Yuyan Zheng, Chuan Shi, Jingzhi Li, and Bin Wu</i>	

FastStep: Scalable Boolean Matrix Decomposition	461
<i>Miguel Araujo, Pedro Ribeiro, and Christos Faloutsos</i>	

Applications

An Expert-in-the-loop Paradigm for Learning Medical Image Grouping	477
<i>Xuan Guo, Qi Yu, Rui Li, Cecilia Ovesdotter Alm, Cara Calvelli, Pengcheng Shi, and Anne Haake</i>	

Predicting Post-operative Visual Acuity for LASIK Surgeries	489
<i>Manish Gupta, Prashant Gupta, Pravin K. Vaddavalli, and Asra Fatima</i>	

LBMF: Log-Bilinear Matrix Factorization for Recommender Systems	502
<i>Yunhui Guo, Xin Wang, and Congfu Xu</i>	

An Empirical Study on Hybrid Recommender System with Implicit Feedback	514
<i>Sunhwan Lee, Anca Chandra, and Divyesh Jadav</i>	

Who Will Be Affected by Supermarket Health Programs? Tracking Customer Behavior Changes via Preference Modeling	527
<i>Ling Luo, Bin Li, Shlomo Berkovsky, Irena Koprinksa, and Fang Chen</i>	

TrafficWatch: Real-Time Traffic Incident Detection and Monitoring Using Social Media.	540
<i>Hoang Nguyen, Wei Liu, Paul Rivera, and Fang Chen</i>	

Automated Setting of Bus Schedule Coverage Using Unsupervised Machine Learning	552
<i>Jihed Khiari, Luis Moreira-Matias, Vitor Cerqueira, and Oded Cats</i>	

Effective Local Metric Learning for Water Pipe Assessment.	565
<i>Mojgan Ghanavati, Raymond K. Wong, Fang Chen, Yang Wang, and Simon Fong</i>	

Classification with Quantification for Air Quality Monitoring	578
<i>Sanad Al-Maskari, Eve Bélisle, Xue Li, Sébastien Le Digabel, Amin Nawahda, and Jiang Zhong</i>	

Predicting Unknown Interactions Between Known Drugs and Targets via Matrix Completion.	591
<i>Qing Liao, Naiyang Guan, Chengkun Wu, and Qian Zhang</i>	

Author Index	605
-------------------------------	------------

Contents – Part II

Spatiotemporal and Image Data

Denoising Time Series by Way of a Flexible Model for Phase Space Reconstruction	3
<i>Minhazul Islam Sk and Arunava Banerjee</i>	
Distributed Sequential Pattern Mining in Large Scale Uncertain Databases . . .	17
<i>Jiaqi Ge and Yuni Xia</i>	
DeepCare: A Deep Dynamic Memory Model for Predictive Medicine	30
<i>Trang Pham, Truyen Tran, Dinh Phung, and Svetha Venkatesh</i>	
Indoor Positioning System for Smart Homes Based on Decision Trees and Passive RFID	42
<i>Frédéric Bergeron, Kevin Bouchard, Sébastien Gaboury, Sylvain Giroux, and Bruno Bouchard</i>	
Deep Feature Extraction from Trajectories for Transportation Mode Estimation.	54
<i>Yuki Endo, Hiroyuki Toda, Kyosuke Nishida, and Akihisa Kawanobe</i>	
Online Learning for Accurate Real-Time Map Matching	67
<i>Biwei Liang, Tengjiao Wang, Shun Li, Wei Chen, Hongyan Li, and Kai Lei</i>	
Multi-hypergraph Incidence Consistent Sparse Coding for Image Data Clustering.	79
<i>Xiaodong Feng, Sen Wu, Wenjun Zhou, and Zhiwei Tang</i>	
Robust Multi-view Manifold Ranking for Image Retrieval	92
<i>Jun Wu, Jianbo Yuan, and Jiebo Luo</i>	
Image Representation Optimization Based on Locally Aggregated Descriptors.	104
<i>Shijiang Chen, Guiguang Ding, Chenxiao Li, and Yuchen Guo</i>	
Reusing Extracted Knowledge in Genetic Programming to Solve Complex Texture Image Classification Problems.	117
<i>Muhammad Iqbal, Bing Xue, and Mengjie Zhang</i>	

Personal Credit Profiling via Latent User Behavior Dimensions on Social Media	130
<i>Guangming Guo, Feida Zhu, Enhong Chen, Le Wu, Qi Liu, Yingling Liu, and Minghui Qiu</i>	
Linear Upper Confidence Bound Algorithm for Contextual Bandit Problem with Piled Rewards	143
<i>Kuan-Hao Huang and Hsuan-Tien Lin</i>	
Incremental Hierarchical Clustering of Stochastic Pattern-Based Symbolic Data	156
<i>Xin Xu, Jiaheng Lu, and Wei Wang</i>	
Computing Hierarchical Summary of the Data Streams	168
<i>Zubair Shah, Abdun Naser Mahmood, and Michael Barlow</i>	
Anomaly Detection and Clustering	
Unsupervised Parameter Estimation for One-Class Support Vector Machines	183
<i>Zahra Ghafoori, Sutharshan Rajasegarar, Sarah M. Erfani, Shanika Karunasekera, and Christopher A. Leckie</i>	
Frequent Pattern Outlier Detection Without Exhaustive Mining	196
<i>Arnaud Giacometti and Arnaud Soulet</i>	
Ensembles of Interesting Subgroups for Discovering High Potential Employees	208
<i>Girish Keshav Palshikar, Kuleshwar Sahu, and Rajiv Srivastava</i>	
Dynamic Grouped Mixture Models for Intermittent Multivariate Sensor Data	221
<i>Naoya Takeishi, Takehisa Yairi, Naoki Nishimura, Yuta Nakajima, and Noboru Takata</i>	
Parallel Discord Discovery	233
<i>Tian Huang, Yongxin Zhu, Yishu Mao, Xinyang Li, Mengyun Liu, Yafei Wu, Yajun Ha, and Gillian Dobbie</i>	
Dboost: A Fast Algorithm for DBSCAN-Based Clustering on High Dimensional Data	245
<i>Yuxiao Zhang, Xiaorong Wang, Bingyang Li, Wei Chen, Tengjiao Wang, and Kai Lei</i>	
A Precise and Robust Clustering Approach Using Homophilic Degrees of Graph Kernel	257
<i>Haolin Yang, Deli Zhao, Lele Cao, and Fuchun Sun</i>	

Constraint Based Subspace Clustering for High Dimensional Uncertain Data	271
<i>Xianchao Zhang, Lu Gao, and Hong Yu</i>	
A Clustering-Based Framework for Incrementally Repairing Entity Resolution	283
<i>Qing Wang, Jingyi Gao, and Peter Christen</i>	
Adaptive Seeding for Gaussian Mixture Models	296
<i>Johannes Blömer and Kathrin Bujna</i>	
A Greedy Algorithm to Construct L1 Graph with Ranked Dictionary	309
<i>Shuchu Han and Hong Qin</i>	

Novel Models and Algorithms

A Rule Based Open Information Extraction Method Using Cascaded Finite-State Transducer	325
<i>Hailun Lin, Yuanzhuo Wang, Peng Zhang, Weiping Wang, Yinliang Yue, and Zheng Lin</i>	
Active Learning Based Entity Resolution Using Markov Logic.	338
<i>Jeffrey Fisher, Peter Christen, and Qing Wang</i>	
Modeling Adversarial Learning as Nested Stackelberg Games.	350
<i>Yan Zhou and Murat Kantarcioglu</i>	
Fast and Semantic Measurements on Collaborative Tagging Quality	363
<i>Yuqing Sun, Haiqi Sun, and Reynold Cheng</i>	
Matrices, Compression, Learning Curves: Formulation, and the GROUPNTEACH Algorithms.	376
<i>Bryan Hooi, Hyun Ah Song, Evangelos Papalexakis, Rakesh Agrawal, and Christos Faloutsos</i>	
Privacy Aware K-Means Clustering with High Utility	388
<i>Thanh Dai Nguyen, Sunil Gupta, Santu Rana, and Svetha Venkatesh</i>	
Secure k -NN Query on Encrypted Cloud Data with Limited Key-Disclosure and Offline Data Owner.	401
<i>Youwen Zhu, Zhikuan Wang, and Yue Zhang</i>	
Hashing-Based Distributed Multi-party Blocking for Privacy-Preserving Record Linkage	415
<i>Thilina Ranbaduge, Dinusha Vatsalan, Peter Christen, and Vassilios Verykios</i>	

Text Mining and Recommender Systems

Enabling Hierarchical Dirichlet Processes to Work Better for Short Texts at Large Scale.	431
<i>Khai Mai, Sang Mai, Anh Nguyen, Ngo Van Linh, and Khoat Than</i>	
Query-Focused Multi-document Summarization Based on Concept Importance.	443
<i>Hai-Tao Zheng, Ji-Min Guo, Yong Jiang, and Shu-Tao Xia</i>	
Mirror on the Wall: Finding Similar Questions with Deep Structured Topic Modeling	454
<i>Arpita Das, Manish Shrivastava, and Manoj Chinnakotla</i>	
An Efficient Dynamic Programming Algorithm for STR-IC-STR-IC-LCS Problem.	466
<i>Daxin Zhu, Yingjie Wu, and Xiaodong Wang</i>	
Efficient Page-Level Data Extraction via Schema Induction and Verification	478
<i>Chia-Hui Chang, Tian-Sheng Chen, Ming-Chuan Chen, and Hung-Li Ding</i>	
Transfer-Learning Based Model for Reciprocal Recommendation	491
<i>Chia-Hsin Ting, Hung-Yi Lo, and Shou-De Lin</i>	
Enhanced SVD for Collaborative Filtering	503
<i>Xin Guan, Chang-Tsun Li, and Yu Guan</i>	
Social Group Based Video Recommendation Addressing the Cold-Start Problem	515
<i>Chunfeng Yang, Yipeng Zhou, Liang Chen, Xiaopeng Zhang, and Dah Ming Chiu</i>	
FeRoSA: A Faceted Recommendation System for Scientific Articles.	528
<i>Tanmoy Chakraborty, Amrith Krishna, Mayank Singh, Niloy Ganguly, Pawan Goyal, and Animesh Mukherjee</i>	
Dual Similarity Regularization for Recommendation	542
<i>Jing Zheng, Jian Liu, Chuan Shi, Fuzhen Zhuang, Jingzhi Li, and Bin Wu</i>	
Collaborative Deep Ranking: A Hybrid Pair-Wise Recommendation Algorithm with Implicit Feedback.	555
<i>Haochao Ying, Liang Chen, Yuwen Xiong, and Jian Wu</i>	
Author Index	569

Advances in Knowledge Discovery and Data Mining
20th Pacific-Asia Conference, PAKDD 2016, Auckland,
New Zealand, April 19-22, 2016, Proceedings, Part I
Bailey, J.; Khan, L.; Washio, T.; Dobbie, G.; Huang, J.Z.;
Wang, R. (Eds.)
2016, XXIV, 608 p. 184 illus., Softcover
ISBN: 978-3-319-31752-6