

Contents – Part II

Clustering and Anomaly Detection

A Targeted Retraining Scheme of Unsupervised Word Embeddings for Specific Supervised Tasks	3
<i>Pengda Qin, Weiran Xu, and Jun Guo</i>	
A Neural Joint Model for Extracting Bacteria and Their Locations	15
<i>Fei Li, Meishan Zhang, Guohong Fu, and Donghong Ji</i>	
Advanced Computation of Sparse Precision Matrices for Big Data	27
<i>Abdelkader Baggag, Halima Bensmail, and Jaideep Srivastava</i>	
Accurate Recognition of the Current Activity in the Presence of Multiple Activities.	39
<i>Weihao Cheng, Sarah Erfani, Rui Zhang, and Ramamohanarao Kotagiri</i>	
Modeling Information Sharing Behavior on Q&A Forums	51
<i>Biru Cui, Shanchieh Jay Yang, and Christophan M. Homan</i>	
Effective Multiclass Transfer for Hypothesis Transfer Learning.	64
<i>Shuang Ao, Xiang Li, and Charles X. Ling</i>	
Clustering Based on Dominant Set and Cluster Expansion	76
<i>Jian Hou and Weixue Liu</i>	

Recommender Systems

Friend Recommendation Considering Preference Coverage in Location-Based Social Networks	91
<i>Fei Yu, Nan Che, Zhijun Li, Kai Li, and Shouxu Jiang</i>	
Contrast Pattern Based Collaborative Behavior Recommendation for Life Improvement	106
<i>Yan Chen, Margot Lisa-Jing Yann, Heidar Davoudi, Joy Choi, Aijun An, and Zhen Mei</i>	
Exploiting Location Significance and User Authority for Point-of-Interest Recommendation	119
<i>Yonghong Yu, Hao Wang, Shuanzhu Sun, and Yang Gao</i>	
Personalized Ranking Recommendation via Integrating Multiple Feedbacks	131
<i>Jian Liu, Chuan Shi, Binbin Hu, Shenghua Liu, and Philip S. Yu</i>	

Fairness Aware Recommendations on Behance	144
<i>Natwar Modani, Deepali Jain, Ujjawal Soni, Gaurav Kumar Gupta, and Palak Agarwal</i>	

A Performance Evaluation Model for Taxi Cruising Path Recommendation System	156
<i>Huimin Lv, Fang Fang, Yishi Zhao, Yuanyuan Liu, and Zhongwen Luo</i>	

MaP2R: A Personalized Maximum Probability Route Recommendation Method Using GPS Trajectories	168
<i>Ge Cui and Xin Wang</i>	

Feature Selection

SNE: Signed Network Embedding.	183
<i>Shuhan Yuan, Xintao Wu, and Yang Xiang</i>	

mHUIMiner: A Fast High Utility Itemset Mining Algorithm for Sparse Datasets	196
<i>Alex Yuxuan Peng, Yun Sing Koh, and Patricia Riddle</i>	

Partial Tree-Edit Distance: A Solution to the Default Class Problem in Pattern-Based Tree Classification.	208
<i>Maciej Piernik and Tadeusz Morzy</i>	

A Domain-Agnostic Approach to Spam-URL Detection via Redirects	220
<i>Heeyoung Kwon, Mirza Basim Baig, and Leman Akoglu</i>	

Automatic and Effective Mining of Coevolving Online Activities	233
<i>Thin Minh Do, Yasuko Matsubara, and Yasushi Sakurai</i>	

Keeping Priors in Streaming Bayesian Learning	247
<i>Anh Nguyen Duc, Ngo Van Linh, Anh Nguyen Kim, and Khoat Than</i>	

Text and Opinion Mining

Efficient Training of Adaptive Regularization of Weight Vectors for Semi-structured Text	261
<i>Tomoya Iwakura</i>	

Behavior-Based Location Recommendation on Location-Based Social Networks	273
<i>Seyyed Mohammadreza Rahimi, Xin Wang, and Behrouz Far</i>	

Integer Linear Programming for Pattern Set Mining; with an Application to Tiling	286
<i>Abdelkader Ouali, Albrecht Zimmermann, Samir Loudni, Yahia Lebbah, Bruno Cremilleux, Patrice Boizumault, and Lakhdar Loukil</i>	
Secured Privacy Preserving Data Aggregation with Semi-honest Servers	300
<i>Zhigang Lu and Hong Shen</i>	
Efficient Pedestrian Detection in the Low Resolution via Sparse Representation with Sparse Support Regression.	313
<i>Wenhua Fang, Jun Chen, and Ruimin Hu</i>	
Multi-task Representation Learning for Enhanced Emotion Categorization in Short Text	324
<i>Anirban Sen, Manjira Sinha, Sandya Mannarswamy, and Shourya Roy</i>	
Fine-Grained Emotion Detection in Contact Center Chat Utterances	337
<i>Shreshtha Mundra, Anirban Sen, Manjira Sinha, Sandya Mannarswamy, Sandipan Dandapat, and Shourya Roy</i>	
Dependency-Tree Based Convolutional Neural Networks for Aspect Term Extraction	350
<i>Hai Ye, Zichao Yan, Zhunchen Luo, and Wenhan Chao</i>	
Topic Modeling over Short Texts by Incorporating Word Embeddings	363
<i>Jipeng Qiang, Ping Chen, Tong Wang, and Xindong Wu</i>	
Mining Drug Properties for Decision Support in Dental Clinics	375
<i>Wee Pheng Goh, Xiaohui Tao, Ji Zhang, and Jianming Yong</i>	
PURE: A Novel Tripartite Model for Review Sentiment Analysis and Recommendation	388
<i>Yue Xue, Liutong Xu, Hai Huang, and Yao Cheng</i>	

Clustering and Matrix Factorization

Multi-View Matrix Completion for Clustering with Side Information.	403
<i>Peng Zhao, Yuan Jiang, and Zhi-Hua Zhou</i>	
Weighted NMF-Based Multiple Sparse Views Clustering for Web Items	416
<i>Xiaolong Gong, Fuwei Wang, and Linpeng Huang</i>	
Parallel Visual Assessment of Cluster Tendency on GPU.	429
<i>Tao Meng and Bo Yuan</i>	
Clustering Complex Data Represented as Propositional Formulas	441
<i>Abdelhamid Boudane, Said Jabbour, Lakhdar Sais, and Yakoub Salhi</i>	

Deep Bayesian Matrix Factorization	453
<i>Sotirios P. Chatzis</i>	

Dynamic, Stream Data Mining

Mining Competitive Pairs Hidden in Co-location Patterns from Dynamic Spatial Databases	467
<i>Junli Lu, Lizhen Wang, Yuan Fang, and Momo Li</i>	
Utility Aware Clustering for Publishing Transactional Data	481
<i>Michael Bewong, Jixue Liu, Lin Liu, and Jiuyong Li</i>	
Self-tuning Filers — Overload Prediction and Preventive Tuning Using Pruned Random Forest	495
<i>Kumar Dheenadayalan, Gopalakrishnan Srinivasaraghavan, and V.N. Muralidhara</i>	
A Centrality-Based Local-First Approach for Analyzing Overlapping Communities in Dynamic Networks	508
<i>Ximan Chen, Heli Sun, Hongxia Du, Jianbin Huang, and Ke Liu</i>	
Web-Scale Personalized Real-Time Recommender System on Suumo	521
<i>Shiyingxue Li, Shimpei Nomura, Yohei Kikuta, and Kazuma Arino</i>	
Modeling Contextual Changes in User Behaviour in Fashion e-Commerce . . .	539
<i>Ashay Tamhane, Sagar Arora, and Deepak Warriar</i>	
Weighted Ensemble Classification of Multi-label Data Streams	551
<i>Lulu Wang, Hong Shen, and Hui Tian</i>	

Novel Models and Algorithms

Improving One-Class Collaborative Filtering with Manifold Regularization by Data-driven Feature Representation	565
<i>Yen-Chieh Lien and Pu-Jen Cheng</i>	
Stable Bayesian Optimization	578
<i>Thanh Dai Nguyen, Sunil Gupta, Santu Rana, and Svetha Venkatesh</i>	
An Exponential Time-Aware Recommendation Model for Mobile Notification Services	592
<i>Chenglin Zeng, Laizhong Cui, and Zhi Wang</i>	
Discovering Periodic Patterns in Non-uniform Temporal Databases	604
<i>R. Uday Kiran, J.N. Venkatesh, Philippe Fournier-Viger, Masashi Toyoda, P. Krishna Reddy, and Masaru Kitsuregawa</i>	

Discovering Both Explicit and Implicit Similarities for Cross-Domain Recommendation	618
<i>Quan Do, Wei Liu, and Fang Chen</i>	
Mining Recurrent Patterns in a Dynamic Attributed Graph	631
<i>Zhi Cheng, Frédéric Flouvat, and Nazha Selmaoui-Folcher</i>	
SS-FIM: Single Scan for Frequent Itemsets Mining in Transactional Databases	644
<i>Youssef Djenouri, Marco Comuzzi, and Djamel Djenouri</i>	
Multi-view Regularized Gaussian Processes	655
<i>Qiuyang Liu and Shiliang Sun</i>	
A Neural Network Model for Semi-supervised Review Aspect Identification	668
<i>Ying Ding, Changlong Yu, and Jing Jiang</i>	

Behavioral Data Mining

Unsupervised Embedding for Latent Similarity by Modeling Heterogeneous MOOC Data.	683
<i>Zhuoxuan Jiang, Shanshan Feng, Weizheng Chen, Guangtao Wang, and Xiaoming Li</i>	
Matrix-Based Method for Inferring Variable Labels Using Outlines of Data in Data Jackets	696
<i>Teruaki Hayashi and Yukio Ohsawa</i>	
Integrating Reviews into Personalized Ranking for Cold Start Recommendation	708
<i>Guang-Neng Hu and Xin-Yu Dai</i>	
Taste or Addiction?: Using Play Logs to Infer Song Selection Motivation . . .	721
<i>Kosetsu Tsukuda and Masataka Goto</i>	
Understanding Drivers' Safety by Fusing Large Scale Vehicle Recorder Dataset and Heterogeneous Circumstantial Data	734
<i>Daisaku Yokoyama, Masashi Toyoda, and Masaru Kitsuregawa</i>	

Graph Clustering and Community Detection

Query-oriented Graph Clustering.	749
<i>Li-Yen Kuo, Chung-Kuang Chou, and Ming-Syan Chen</i>	
CCCG: Clique Conversion Ratio Driven Clustering of Graphs	762
<i>Prathyush Sambaturu and Kamalakar Karlapalem</i>	

Mining Cohesive Clusters with Interpretations in Labeled Graphs	774
<i>Hongxia Du, Heli Sun, Jianbin Huang, Zhongbin Sun, Liang He, and Hong Cheng</i>	
A SAT-Based Framework for Overlapping Community Detection in Networks	786
<i>Said Jabbour, Nizar Mhadhbi, Badran Raddaoui, and Lakhdar Sais</i>	
Dimensionality Reduction	
Denoising Autoencoder as an Effective Dimensionality Reduction and Clustering of Text Data	801
<i>Milad Leyli-Abadi, Lazhar Labiod, and Mohamed Nadif</i>	
Gradable Adjective Embedding for Commonsense Knowledge	814
<i>Kyungjae Lee, Hyunsouk Cho, and Seung-won Hwang</i>	
Combining Dimensionality Reduction with Random Forests for Multi-label Classification Under Interactivity Constraints	828
<i>Noureddine-Yassine Nair-Benrekia, Pascale Kuntz, and Frank Meyer</i>	
A Generalized Model for Multidimensional Intransitivity	840
<i>Jiuding Duan, Jiye Li, Yukino Baba, and Hisashi Kashima</i>	
Author Index	853

Contents – Part I

Classification and Deep Learning

Convolutional Bi-directional LSTM for Detecting Inappropriate Query Suggestions in Web Search.	3
<i>Harish Yenala, Manoj Chinnakotla, and Jay Goyal</i>	
A Fast and Easy Regression Technique for k -NN Classification Without Using Negative Pairs.	17
<i>Yutaro Shigeto, Masashi Shimbo, and Yuji Matsumoto</i>	
Deep Network Regularization via Bayesian Inference of Synaptic Connectivity.	30
<i>Harris Partaourides and Sotirios P. Chatzis</i>	
Adaptive One-Class Support Vector Machine for Damage Detection in Structural Health Monitoring.	42
<i>Ali Anaissi, Nguyen Lu Dang Khoa, Samir Mustapha, Mehrisadat Makki Alamdari, Ali Braytee, Yang Wang, and Fang Chen</i>	
A Classification Model for Diverse and Noisy Labelers	58
<i>Hao-En Sung, Cheng-Kuan Chen, Han Xiao, and Shou-De Lin</i>	
Automatic Discovery of Common and Idiosyncratic Latent Effects in Multilevel Regression	70
<i>Sk Minhazul Islam and Arunvava Banerjee</i>	
A Deep Neural Network for Pairwise Classification: Enabling Feature Conjunctions and Ensuring Symmetry	83
<i>Kyohei Atarashi, Satoshi Oyama, Masahito Kurihara, and Kazune Furudo</i>	
Feature Ranking of Large, Robust, and Weighted Clustering Result	96
<i>Mirka Saarela, Joonas Hämäläinen, and Tommi Kärkkäinen</i>	
Purchase Signatures of Retail Customers	110
<i>Clement Gautrais, René Quiniou, Peggy Cellier, Thomas Guyet, and Alexandre Termier</i>	
Volatility Adaptive Classifier System.	122
<i>Ruolin Jia, Yun Sing Koh, and Gillian Dobbie</i>	

Distributed Representations for Words on Tables	135
<i>Minoru Yoshida, Kazuyuki Matsumoto, and Kenji Kita</i>	
Link Prediction for Isolated Nodes in Heterogeneous Network by Topic-Based Co-clustering.	147
<i>Katsufumi Tomobe, Masafumi Oyamada, and Shinji Nakadai</i>	
Predicting Destinations from Partial Trajectories Using Recurrent Neural Network	160
<i>Yuki Endo, Kyosuke Nishida, Hiroyuki Toda, and Hiroshi Sawada</i>	
Preventing Inadvertent Information Disclosures via Automatic Security Policies	173
<i>Tanya Goyal, Sanket Mehta, and Balaji Vasan Srinivasan</i>	
Personalized Deep Learning for Tag Recommendation.	186
<i>Hanh T.H. Nguyen, Martin Wistuba, Josif Grabocka, Lucas Rego Drumond, and Lars Schmidt-Thieme</i>	
Information-Theoretic Non-redundant Subspace Clustering	198
<i>Nina Hubig and Claudia Plant</i>	
Cost Matters: A New Example-Dependent Cost-Sensitive Logistic Regression Model	210
<i>Nikou Günnemann and Jürgen Pfeffer</i>	
Social Network and Graph Mining	
Beyond Assortativity: Proclivity Index for Attributed Networks (ProNE)	225
<i>Reihaneh Rabbany, Dhivya Eswaran, Artur W. Dubrawski, and Christos Faloutsos</i>	
Hierarchical Mixed Neural Network for Joint Representation Learning of Social-Attribute Network	238
<i>Weizheng Chen, Jinpeng Wang, Zhuoxuan Jiang, Yan Zhang, and Xiaoming Li</i>	
Cost-Effective Viral Marketing in the Latency Aware Independent Cascade Model.	251
<i>Robert Gwadera and Grigorios Loukides</i>	
DSBPR: Dual Similarity Bayesian Personalized Ranking	266
<i>Longfei Shi, Bin Wu, Jing Zheng, Chuan Shi, and Mengxin Li</i>	

Usage Based Tag Enhancement of Images	278
<i>Balaji Vasan Srinivasan, Noman Ahmed Sheikh, Roshan Kumar, Saurabh Verma, and Niloy Ganguly</i>	
Edge Role Discovery via Higher-Order Structures	291
<i>Nesreen K. Ahmed, Ryan A. Rossi, Theodore L. Willke, and Rong Zhou</i>	
Efficient Bi-level Variable Selection and Application to Estimation of Multiple Covariance Matrices	304
<i>Duy Nhat Phan, Hoai An Le Thi, and Dinh Tao Pham</i>	
Entity Set Expansion with Meta Path in Knowledge Graph	317
<i>Yuyan Zheng, Chuan Shi, Xiaohuan Cao, Xiaoli Li, and Bin Wu</i>	
Using Network Flows to Identify Users Sharing Extremist Content on Social Media	330
<i>Yifang Wei and Lisa Singh</i>	
MC3: A Multi-class Consensus Classification Framework	343
<i>Tanmoy Chakraborty, Des Chandhok, and V.S. Subrahmanian</i>	
Monte Carlo Based Incremental PageRank on Evolving Graphs	356
<i>Qun Liao, ShuangShuang Jiang, Min Yu, Yulu Yang, and Tao Li</i>	
Joint Weighted Nonnegative Matrix Factorization for Mining Attributed Graphs	368
<i>Zhichao Huang, Yunming Ye, Xutao Li, Feng Liu, and Huajie Chen</i>	
Predicting Happiness State Based on Emotion Representative Mining in Online Social Networks	381
<i>Xiao Zhang, Wenzhong Li, Hong Huang, Cam-Tu Nguyen, Xu Chen, Xiaoliang Wang, and Sanglu Lu</i>	
Exploring Celebrities on Inferring User Geolocation in Twitter	395
<i>Mohammad Ebrahimi, Elaheh ShafieiBavani, Raymond Wong, and Fang Chen</i>	
Do Rumors Diffuse Differently from Non-rumors? A Systematically Empirical Analysis in Sina Weibo for Rumor Identification	407
<i>Yahui Liu, Xiaolong Jin, Huawei Shen, and Xueqi Cheng</i>	
Investigating the Dynamics of Religious Conflicts by Mining Public Opinions on Social Media	421
<i>Swati Agarwal and Ashish Sureka</i>	

Mining High-Utility Itemsets with Both Positive and Negative Unit Profits from Uncertain Databases	434
<i>Wensheng Gan, Jerry Chun-Wei Lin, Philippe Fournier-Viger, Han-Chieh Chao, and Vincent S. Tseng</i>	
Sparse Stochastic Inference with Regularization	447
<i>Tung Doan and Khoat Than</i>	
Exploring Check-in Data to Infer Social Ties in Location Based Social Networks	460
<i>Gunarto Sindoro Njoo, Min-Chia Kao, Kuo-Wei Hsu, and Wen-Chih Peng</i>	
Scalable Twitter User Clustering Approach Boosted by Personalized PageRank	472
<i>Anup Naik, Hideyuki Maeda, Vibhor Kanojia, and Sumio Fujita</i>	
An Enhanced Markov Clustering Algorithm Based on <i>Physarum</i>	486
<i>Mingxin Liang, Chao Gao, Xianghua Li, and Zili Zhang</i>	
Exploiting Geographical Location for Team Formation in Social Coding Sites.	499
<i>Yuqiang Han, Yao Wan, Liang Chen, Guandong Xu, and Jian Wu</i>	
Weighted Simplicial Complex: A Novel Approach for Predicting Small Group Evolution	511
<i>Ankit Sharma, Terrence J. Moore, Ananthram Swami, and Jaideep Srivastava</i>	
A P-LSTM Neural Network for Sentiment Classification	524
<i>Chi Lu, Heyan Huang, Ping Jian, Dan Wang, and Yi-Di Guo</i>	
Learning What Matters – Sampling Interesting Patterns	534
<i>Vladimir Dzyuba and Matthijs van Leeuwen</i>	
PNE: Label Embedding Enhanced Network Embedding	547
<i>Weizheng Chen, Xianling Mao, Xiangyu Li, Yan Zhang, and Xiaoming Li</i>	
Improving Temporal Record Linkage Using Regression Classification	561
<i>Yichen Hu, Qing Wang, Dinusha Vatsalan, and Peter Christen</i>	
Community Detection in Graph Streams by Pruning Zombie Nodes	574
<i>Yue Ding, Ling Huang, Chang-Dong Wang, and Dong Huang</i>	
Bilingual Lexicon Extraction from Comparable Corpora Based on Closed Concepts Mining	586
<i>Mohamed Chebel, Chiraz Latiri, and Eric Gaussier</i>	

Collective Geographical Embedding for Geolocating Social Network Users	599
<i>Fengjiao Wang, Chun-Ta Lu, Yongzhi Qu, and Philip S. Yu</i>	

Privacy-Preserving Mining and Security/Risk Applications

Partitioning-Based Mechanisms Under Personalized Differential Privacy	615
<i>Haoran Li, Li Xiong, Zhanglong Ji, and Xiaoqian Jiang</i>	
Efficient Cryptanalysis of Bloom Filters for Privacy-Preserving Record Linkage	628
<i>Peter Christen, Rainer Schnell, Dinusha Vatsalan, and Thilina Ranbaduge</i>	
Energy-Based Localized Anomaly Detection in Video Surveillance	641
<i>Hung Vu, Tu Dinh Nguyen, Anthony Travers, Svetha Venkatesh, and Dinh Phung</i>	
A Fast Fourier Transform-Coupled Machine Learning-Based Ensemble Model for Disease Risk Prediction Using a Real-Life Dataset.	654
<i>Raid Lafta, Ji Zhang, Xiaohui Tao, Yan Li, Wessam Abbas, Yonglong Luo, Fulong Chen, and Vincent S. Tseng</i>	
Assessing Death Risk of Patients with Cardiovascular Disease from Long-Term Electrocardiogram Streams Summarization.	671
<i>Shenda Hong, Meng Wu, Jinbo Zhang, and Hongyan Li</i>	

Spatio-Temporal and Sequential Data Mining

On the Robustness of Decision Tree Learning Under Label Noise.	685
<i>Aritra Ghosh, Naresh Manwani, and P.S. Sastry</i>	
Relevance-Based Evaluation Metrics for Multi-class Imbalanced Domains . . .	698
<i>Paula Branco, Luís Torgo, and Rita P. Ribeiro</i>	
Location Prediction Through Activity Purpose: Integrating Temporal and Sequential Models.	711
<i>Dongliang Liao, Yuan Zhong, and Jing Li</i>	
Modeling Temporal Behavior of Awards Effect on Viewership of Movies . . .	724
<i>Basmah Altaf, Faisal Kamiran, and Xiangliang Zhang</i>	
A <i>Physarum</i> -Inspired Ant Colony Optimization for Community Mining. . . .	737
<i>Mingxin Liang, Chao Gao, Xianghua Li, and Zili Zhang</i>	

A Novel Diversity Measure for Understanding Movie Ranks in Movie Collaboration Networks	750
<i>Manqing Ma, Wei Pang, Lan Huang, and Zhe Wang</i>	
Local-to-Global Unsupervised Anomaly Detection from Temporal Data	762
<i>Seif-Eddine Benkabou, Khalid Benabdeslem, and Bruno Canitia</i>	
Mining Temporal Fluctuating Patterns	773
<i>Shan-Yun Teng, Cheng-Kuan Ou, and Kun-Ta Chuang</i>	
Marked Temporal Dynamics Modeling Based on Recurrent Neural Network	786
<i>Yongqing Wang, Shenghua Liu, Huawei Shen, Jinhua Gao, and Xueqi Cheng</i>	
Mining of Location-Based Social Networks for Spatio-Temporal Social Influence	799
<i>Yu-Ting Wen, Yi Yuan Fan, and Wen-Chih Peng</i>	
Multi-perspective Hierarchical Dirichlet Process for Geographical Topic Modeling	811
<i>Yuan He, Cheng Wang, and Changjun Jiang</i>	
Enumerating Non-redundant Association Rules Using Satisfiability	824
<i>Abdelhamid Boudane, Said Jabbour, Lakhdar Sais, and Yakoub Salhi</i>	
Author Index	837

Advances in Knowledge Discovery and Data Mining
21st Pacific-Asia Conference, PAKDD 2017, Jeju, South
Korea, May 23-26, 2017, Proceedings, Part II

Kim, J.; Shim, K.; Cao, L.; Lee, J.-G.; Lin, X.; Moon, Y.-S.
(Eds.)

2017, XXXII, 857 p. 252 illus., Softcover

ISBN: 978-3-319-57528-5