Contents – Part II

Clustering and Anomaly Detection

A Targeted Retraining Scheme of Unsupervised Word Embeddings for Specific Supervised Tasks .............................. 3
Pengda Qin, Weiran Xu, and Jun Guo

A Neural Joint Model for Extracting Bacteria and Their Locations .......... 15
Fei Li, Meishan Zhang, Guohong Fu, and Donghong Ji

Advanced Computation of Sparse Precision Matrices for Big Data ............. 27
Abdelkader Baggag, Halima Bensmail, and Jaideep Srivastava

Accurate Recognition of the Current Activity in the Presence of Multiple Activities .............................................. 39
Weihao Cheng, Sarah Erfani, Rui Zhang, and Ramamohanarao Kotagiri

Modeling Information Sharing Behavior on Q&A Forums ...................... 51
Biru Cui, Shanchieh Jay Yang, and Christophan M. Homan

Effective Multiclass Transfer for Hypothesis Transfer Learning ................ 64
Shuang Ao, Xiang Li, and Charles X. Ling

Clustering Based on Dominant Set and Cluster Expansion ...................... 76
Jian Hou and Weixue Liu

Recommender Systems

Friend Recommendation Considering Preference Coverage in Location-Based Social Networks ........................................... 91
Fei Yu, Nan Che, Zhijun Li, Kai Li, and Shouxu Jiang

Contrast Pattern Based Collaborative Behavior Recommendation for Life Improvement .................................................. 106
Yan Chen, Margot Lisa-Jing Yann, Heidar Davoudi, Joy Choi, Aijun An, and Zhen Mei

Exploiting Location Significance and User Authority for Point-of-Interest Recommendation ............................................. 119
Yonghong Yu, Hao Wang, Shuanzhu Sun, and Yang Gao

Personalized Ranking Recommendation via Integrating Multiple Feedbacks ................................................................. 131
Jian Liu, Chuan Shi, Binbin Hu, Shenghua Liu, and Philip S. Yu
Fairness Aware Recommendations on Behance .......................... 144
  Natwar Modani, Deepali Jain, Ujjawal Soni, Gaurav Kumar Gupta,
  and Palak Agarwal

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

MaP2R: A Personalized Maximum Probability Route Recommendation
Method Using GPS Trajectories ................................. 168
  Ge Cui and Xin Wang

Feature Selection

SNE: Signed Network Embedding ................................. 183
  Shuhan Yuan, Xintao Wu, and Yang Xiang

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

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

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

Automatic and Effective Mining of Coevolving Online Activities ....... 233
  Thinh Minh Do, Yasuko Matsubara, and Yasushi Sakurai

Keeping Priors in Streaming Bayesian Learning ...................... 247
  Anh Nguyen Duc, Ngo Van Linh, Anh Nguyen Kim, and Khoat Than

Text and Opinion Mining

Efficient Training of Adaptive Regularization of Weight Vectors
for Semi-structured Text ........................................ 261
  Tomoya Iwakura

Behavior-Based Location Recommendation on Location-Based
Social Networks .................................................. 273
  Seyyed Mohammadreza Rahimi, Xin Wang, and Behrouz Far
Integer Linear Programming for Pattern Set Mining; with an Application to Tiling ................................................. 286
Abdelkader Ouali, Albrecht Zimmermann, Samir Loudni, Yahia Lebbah, Bruno Cremilleux, Patrice Boizumault, and Lakhdar Loukil

Secured Privacy Preserving Data Aggregation with Semi-honest Servers ...... 300
Zhigang Lu and Hong Shen

Efficient Pedestrian Detection in the Low Resolution via Sparse Representation with Sparse Support Regression ................................. 313
Wenhua Fang, Jun Chen, and Ruimin Hu

Multi-task Representation Learning for Enhanced Emotion Categorization in Short Text .................................................. 324
Anirban Sen, Manjira Sinha, Sandya Mannarswamy, and Shourya Roy

Fine-Grained Emotion Detection in Contact Center Chat Utterances ............ 337
Shreshtha Mundra, Anirban Sen, Manjira Sinha, Sandya Mannarswamy, Sandipan Dandapat, and Shourya Roy

Dependency-Tree Based Convolutional Neural Networks for Aspect Term Extraction .......................................................... 350
Hai Ye, Zichao Yan, Zhunchen Luo, and Wenhan Chao

Topic Modeling over Short Texts by Incorporating Word Embeddings .......... 363
Jipeng Qiang, Ping Chen, Tong Wang, and Xindong Wu

Mining Drug Properties for Decision Support in Dental Clinics ................. 375
Wee Pheng Goh, Xiaohui Tao, Ji Zhang, and Jianming Yong

PURE: A Novel Tripartite Model for Review Sentiment Analysis and Recommendation ......................................................... 388
Yue Xue, Liutong Xu, Hai Huang, and Yao Cheng

Clustering and Matrix Factorization

Multi-View Matrix Completion for Clustering with Side Information ............ 403
Peng Zhao, Yuan Jiang, and Zhi-Hua Zhou

Weighted NMF-Based Multiple Sparse Views Clustering for Web Items ......... 416
Xiaolong Gong, Fuwei Wang, and Linpeng Huang

Parallel Visual Assessment of Cluster Tendency on GPU .......................... 429
Tao Meng and Bo Yuan

Clustering Complex Data Represented as Propositional Formulas ............... 441
Abdelhamid Boudane, Said Jabbour, Lakhdar Sais, and Yakoub Salhi
Deep Bayesian Matrix Factorization ........................................... 453
   Sotirios P. Chatzis

**Dynamic, Stream Data Mining**

Mining Competitive Pairs Hidden in Co-location Patterns from Dynamic
Spatial Databases ................................................................. 467
   Junli Lu, Lizhen Wang, Yuan Fang, and Momo Li

Utility Aware Clustering for Publishing Transactional Data .......... 481
   Michael Bewong, Jixue Liu, Lin Liu, and Jiuyong Li

Self-tuning Filers — Overload Prediction and Preventive Tuning
Using Pruned Random Forest ............................................... 495
   Kumar Dheenadayalan, Gopalakrishnan Srinivasaraghavan,
   and V.N. Muralidhara

A Centrality-Based Local-First Approach for Analyzing Overlapping
Communities in Dynamic Networks ....................................... 508
   Ximan Chen, Heli Sun, Hongxia Du, Jianbin Huang, and Ke Liu

Web-Scale Personalized Real-Time Recommender System on Suumo .... 521
   Shiyinxue Li, Shimpei Nomura, Yohei Kikuta, and Kazuma Arino

Modeling Contextual Changes in User Behaviour in Fashion e-Commerce . . 539
   Ashay Tamhane, Sagar Arora, and Deepak Warrier

Weighted Ensemble Classification of Multi-label Data Streams ........... 551
   Lulu Wang, Hong Shen, and Hui Tian

**Novel Models and Algorithms**

Improving One-Class Collaborative Filtering with Manifold Regularization
by Data-driven Feature Representation ..................................... 565
   Yen-Chieh Lien and Pu-Jen Cheng

Stable Bayesian Optimization ................................................. 578
   Thanh Dai Nguyen, Sunil Gupta, Santu Rana, and Svetha Venkatesh

An Exponential Time-Aware Recommendation Model for Mobile
Notification Services ......................................................... 592
   Chenglin Zeng, Laizhong Cui, and Zhi Wang

Discovering Periodic Patterns in Non-uniform Temporal Databases .... 604
   R. Uday Kiran, J.N. Venkatesh, Philippe Fournier-Viger,
   Masashi Toyoda, P. Krishna Reddy, and Masaru Kitsuregawa
Discovering Both Explicit and Implicit Similarities for Cross-Domain Recommendation .................................................. 618
Quan Do, Wei Liu, and Fang Chen

Mining Recurrent Patterns in a Dynamic Attributed Graph ............. 631
Zhi Cheng, Frédéric Flouvat, and Nazha Selmaoui-Folcher

SS-FIM: Single Scan for Frequent Itemsets Mining in Transactional Databases .......................................................... 644
Youcef Djenouri, Marco Comuzzi, and Djamel Djenouri

Multi-view Regularized Gaussian Processes .................................. 655
Qiuyang Liu and Shiliang Sun

A Neural Network Model for Semi-supervised Review Aspect Identification ............................................................ 668
Ying Ding, Changlong Yu, and Jing Jiang

Behavioral Data Mining

Unsupervised Embedding for Latent Similarity by Modeling Heterogeneous MOOC Data .................................................. 683
Zhuoxuan Jiang, Shanshan Feng, Weizheng Chen, Guangtao Wang, and Xiaoming Li

Matrix-Based Method for Inferring Variable Labels Using Outlines of Data in Data Jackets ........................................... 696
Teruaki Hayashi and Yukio Ohsawa

Integrating Reviews into Personalized Ranking for Cold Start Recommendation .......................................................... 708
Guang-Neng Hu and Xin-Yu Dai

Taste or Addiction?: Using Play Logs to Infer Song Selection Motivation ................................................................. 721
Kosetsu Tsukuda and Masataka Goto

Understanding Drivers’ Safety by Fusing Large Scale Vehicle Recorder Dataset and Heterogeneous Circumstantial Data .................................................. 734
Daisaku Yokoyama, Masashi Toyoda, and Masaru Kitsuregawa

Graph Clustering and Community Detection

Query-oriented Graph Clustering .................................................. 749
Li-Yen Kuo, Chung-Kuang Chou, and Ming-Syan Chen

CCCG: Clique Conversion Ratio Driven Clustering of Graphs ............. 762
Prathyush Sambaturu and Kamalakar Karlapalem
Mining Cohesive Clusters with Interpretations in Labeled Graphs 774
   Hongxia Du, Heli Sun, Jianbin Huang, Zhongbin Sun, Liang He, and Hong Cheng

A SAT-Based Framework for Overlapping Community Detection in Networks 786
   Said Jabbour, Nizar Mhadhbi, Badran Raddaoui, and Lakhdar Sais

Dimensionality Reduction

Denoising Autoencoder as an Effective Dimensionality Reduction and Clustering of Text Data 801
   Milad Leyli-Abadi, Lazhar Labiod, and Mohamed Nadif

Gradable Adjective Embedding for Commonsense Knowledge 814
   Kyungjae Lee, Hyunsouk Cho, and Seung-won Hwang

Combining Dimensionality Reduction with Random Forests for Multi-label Classification Under Interactivity Constraints 828
   Noureddine-Yassine Nair-Benrekia, Pascale Kuntz, and Frank Meyer

A Generalized Model for Multidimensional Intransitivity 840
   Jiuding Duan, Jiyi Li, Yukino Baba, and Hisashi Kashima

Author Index 853
Contents – Part I

Classification and Deep Learning

Convolutional Bi-directional LSTM for Detecting Inappropriate Query Suggestions in Web Search. ................................................................. 3
  Harish Yenala, Manoj Chinnakotla, and Jay Goyal

A Fast and Easy Regression Technique for $k$-NN Classification Without Using Negative Pairs. ................................................................. 17
  Yutaro Shigeto, Masashi Shimbo, and Yuji Matsumoto

Deep Network Regularization via Bayesian Inference of Synaptic Connectivity .......................................................................................... 30
  Harris Partaourides and Sotirios P. Chatzis

Adaptive One-Class Support Vector Machine for Damage Detection in Structural Health Monitoring. ............................................................ 42
  Ali Anaissi, Nguyen Lu Dang Khoa, Samir Mustapha, Mehrisadat Makki Alamdari, Ali Braytee, Yang Wang, and Fang Chen

A Classification Model for Diverse and Noisy Labelers ............................ 58
  Hao-En Sung, Cheng-Kuan Chen, Han Xiao, and Shou-De Lin

Automatic Discovery of Common and Idiosyncratic Latent Effects in Multilevel Regression ................................................................. 70
  Sk Minhazul Islam and Arunvava Banerjee

A Deep Neural Network for Pairwise Classification: Enabling Feature Conjunctions and Ensuring Symmetry ........................................... 83
  Kyohei Atarashi, Satoshi Oyama, Masahito Kurihara, and Kazune Furudo

Feature Ranking of Large, Robust, and Weighted Clustering Result .......... 96
  Mirka Saarela, Joonas Hämäläinen, and Tommi Kärkkäinen

Purchase Signatures of Retail Customers ................................................. 110
  Clement Gautrais, René Quiniou, Peggy Cellier, Thomas Guyet, and Alexandre Termier

Volatility Adaptive Classifier System .................................................... 122
  Ruolin Jia, Yun Sing Koh, and Gillian Dobbie
Distributed Representations for Words on Tables ........................................... 135
Minoru Yoshida, Kazuyuki Matsumoto, and Kenji Kita

Link Prediction for Isolated Nodes in Heterogeneous Network by Topic-Based Co-clustering ................................................................. 147
Katsufumi Tomobe, Masafumi Oyamada, and Shinji Nakadai

Predicting Destinations from Partial Trajectories Using Recurrent Neural Network ................................................................. 160
Yuki Endo, Kyosuke Nishida, Hiroyuki Toda, and Hiroshi Sawada

Preventing Inadvertent Information Disclosures via Automatic Security Policies ................................................................. 173
Tanya Goyal, Sanket Mehta, and Balaji Vasan Srinivasan

Personalized Deep Learning for Tag Recommendation ................................................................. 186
Hanh T.H. Nguyen, Martin Wistuba, Josif Grabocka, Lucas Rego Drumond, and Lars Schmidt-Thieme

Information-Theoretic Non-redundant Subspace Clustering ................................................................. 198
Nina Hubig and Claudia Plant

Cost Matters: A New Example-Dependent Cost-Sensitive Logistic Regression Model ................................................................. 210
Nikou Günnewann and Jürgen Pfeffer

Social Network and Graph Mining

Beyond Assortativity: Proclivity Index for Attributed Networks (PRoNE) ................................................................. 225
Reihan Rabbany, Dhivya Eswaran, Artur W. Dubrawski, and Christos Faloutsos

Hierarchical Mixed Neural Network for Joint Representation Learning of Social-Attribute Network ................................................................. 238
Weizheng Chen, Jinpeng Wang, Zhuoxuan Jiang, Yan Zhang, and Xiaoming Li

Cost-Effective Viral Marketing in the Latency Aware Independent Cascade Model ................................................................. 251
Robert Gwadera and Grigorios Loukides

DSBPR: Dual Similarity Bayesian Personalized Ranking ................................................................. 266
Longfei Shi, Bin Wu, Jing Zheng, Chuan Shi, and Mengxin Li
Usage Based Tag Enhancement of Images ................................. 278

balaji vasan srinivasan, noman ahmed sheikh, roshan kumar, saurabh verma, and niloy ganguly

Edge Role Discovery via Higher-Order Structures .......................... 291

nesreen k. ahmed, ryan a. rossi, theodore l. willke, and rong zhou

Efficient Bi-level Variable Selection and Application to Estimation of Multiple Covariance Matrices .......................... 304

duy nhat phan, hoai an le thi, and dinh tao pham

Entity Set Expansion with Meta Path in Knowledge Graph .............. 317

yuyan zheng, chuan shi, xiaohuan cao, xiaoli li, and bin wu

Using Network Flows to Identify Users Sharing Extremist Content on Social Media .......................... 330

yifang wei and lisa singh

MC3: A Multi-class Consensus Classification Framework .................. 343

tanmoy chakraborty, des chandhok, and v.s. subrahmanian

Monte Carlo Based Incremental PageRank on Evolving Graphs ........... 356

qun liao, shuang shuang jiang, min yu, yulu yang, and tao li

Joint Weighted Nonnegative Matrix Factorization for Mining Attributed Graphs .................................................. 368

zhichao huang, yunming ye, xutao li, feng liu, and huajie chen

Predicting Happiness State Based on Emotion Representative Mining in Online Social Networks .......................... 381

xiao zhang, wenzhong li, hong huang, cam-tu nguyen, xu chen, xiaoliang wang, and sanglu lu

Exploring Celebrities on Inferring User Geolocation in Twitter .......... 395

mohammad ebrahimi, elaheh shafieibavani, raymond wong, and fang chen

Do Rumors Diffuse Differently from Non-rumors? A Systematically Empirical Analysis in Sina Weibo for Rumor Identification ........ 407

yahui liu, xiaolong jin, huawei shen, and xueqi cheng

Investigating the Dynamics of Religious Conflicts by Mining Public Opinions on Social Media .......................... 421

swati agarwal and ashish sureka
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mining High-Utility Itemsets with Both Positive and Negative Unit Profits from Uncertain Databases</td>
<td>434</td>
</tr>
<tr>
<td>Wensheng Gan, Jerry Chun-Wei Lin, Philippe Fournier-Viger, Han-Chieh Chao, and Vincent S. Tseng</td>
<td></td>
</tr>
<tr>
<td>Sparse Stochastic Inference with Regularization</td>
<td>447</td>
</tr>
<tr>
<td>Tung Doan and Khoat Than</td>
<td></td>
</tr>
<tr>
<td>Exploring Check-in Data to Infer Social Ties in Location Based Social Networks</td>
<td>460</td>
</tr>
<tr>
<td>Gunarto Sindoro Njoo, Min-Chia Kao, Kuo-Wei Hsu, and Wen-Chih Peng</td>
<td></td>
</tr>
<tr>
<td>Scalable Twitter User Clustering Approach Boosted by Personalized PageRank</td>
<td>472</td>
</tr>
<tr>
<td>Anup Naik, Hideyuki Maeda, Vibhor Kanojia, and Sumio Fujita</td>
<td></td>
</tr>
<tr>
<td>An Enhanced Markov Clustering Algorithm Based on Physarum</td>
<td>486</td>
</tr>
<tr>
<td>Mingxin Liang, Chao Gao, Xianghua Li, and Zili Zhang</td>
<td></td>
</tr>
<tr>
<td>Exploiting Geographical Location for Team Formation in Social Coding Sites</td>
<td>499</td>
</tr>
<tr>
<td>Yuqiang Han, Yao Wan, Liang Chen, Guandong Xu, and Jian Wu</td>
<td></td>
</tr>
<tr>
<td>Weighted Simplicial Complex: A Novel Approach for Predicting Small Group Evolution</td>
<td>511</td>
</tr>
<tr>
<td>Ankit Sharma, Terrence J. Moore, Ananthram Swami, and Jaideep Srivastava</td>
<td></td>
</tr>
<tr>
<td>A P-LSTM Neural Network for Sentiment Classification</td>
<td>524</td>
</tr>
<tr>
<td>Chi Lu, Heyan Huang, Ping Jian, Dan Wang, and Yi-Di Guo</td>
<td></td>
</tr>
<tr>
<td>Learning What Matters – Sampling Interesting Patterns</td>
<td>534</td>
</tr>
<tr>
<td>Vladimir Dzyuba and Matthijs van Leeuwen</td>
<td></td>
</tr>
<tr>
<td>PNE: Label Embedding Enhanced Network Embedding</td>
<td>547</td>
</tr>
<tr>
<td>Weizheng Chen, Xianling Mao, Xiangyu Li, Yan Zhang, and Xiaoming Li</td>
<td></td>
</tr>
<tr>
<td>Improving Temporal Record Linkage Using Regression Classification</td>
<td>561</td>
</tr>
<tr>
<td>Yichen Hu, Qing Wang, Dinusha Vatsalan, and Peter Christen</td>
<td></td>
</tr>
<tr>
<td>Community Detection in Graph Streams by Pruning Zombie Nodes</td>
<td>574</td>
</tr>
<tr>
<td>Yue Ding, Ling Huang, Chang-Dong Wang, and Dong Huang</td>
<td></td>
</tr>
<tr>
<td>Bilingual Lexicon Extraction from Comparable Corpora Based on Closed Concepts Mining</td>
<td>586</td>
</tr>
<tr>
<td>Mohamed Chebel, Chiraz Latiri, and Eric Gaussier</td>
<td></td>
</tr>
</tbody>
</table>
Collective Geographical Embedding for Geolocating
Social Network Users ................................................. 599
   Fengjiao Wang, Chun-Ta Lu, Yongzhi Qu, and Philip S. Yu

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

Spatio-Temporal and Sequential Data Mining
On the Robustness of Decision Tree Learning Under Label Noise ............ 685
   Aritra Ghosh, Naresh Manwani, and P.S. Sastry
Relevance-Based Evaluation Metrics for Multi-class Imbalanced Domains . . . 698
   Paula Branco, Luís Torgo, and Rita P. Ribeiro
Location Prediction Through Activity Purpose: Integrating Temporal and Sequential Models ........................................ 711
   Dongliang Liao, Yuan Zhong, and Jing Li
Modeling Temporal Behavior of Awards Effect on Viewership of Movies . . . 724
   Basmah Altaf, Faisal Kamiran, and Xiangliang Zhang
A Physarum-Inspired Ant Colony Optimization for Community Mining ........ 737
   Mingxin Liang, Chao Gao, Xianghua Li, and Zili Zhang
A Novel Diversity Measure for Understanding Movie Ranks in Movie Collaboration Networks .......................................................... 750
   Manqing Ma, Wei Pang, Lan Huang, and Zhe Wang

Local-to-Global Unsupervised Anomaly Detection from Temporal Data .... 762
   Seif-Eddine Benkabou, Khalid Benabdeslem, and Bruno Canitia

Mining Temporal Fluctuating Patterns ........................................... 773
   Shan-Yun Teng, Cheng-Kuan Ou, and Kun-Ta Chuang

Marked Temporal Dynamics Modeling Based on Recurrent Neural Network .......................................................... 786
   Yongqing Wang, Shenghua Liu, Huawei Shen, Jinhua Gao,
   and Xueqi Cheng

Mining of Location-Based Social Networks for Spatio-Temporal Social Influence .......................................................... 799
   Yu-Ting Wen, Yi Yuan Fan, and Wen-Chih Peng

Multi-perspective Hierarchical Dirichlet Process for Geographical Topic Modeling .......................................................... 811
   Yuan He, Cheng Wang, and Changjun Jiang

Enumerating Non-redundant Association Rules Using Satisfiability ...... 824
   Abdelhamid Boudane, Said Jabbour, Lakhdar Sais, and Yakoub Salhi

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