

# Contents

## Spotlight Research Papers

Effective Monotone Knowledge Integration in Kernel Support Vector Machines . . . . .	3
<i>Christopher Bartley, Wei Liu, and Mark Reynolds</i>	
Textual Cues for Online Depression in Community and Personal Settings . . .	19
<i>Thin Nguyen, Svetha Venkatesh, and Dinh Phung</i>	
Confidence-Weighted Bipartite Ranking. . . . .	35
<i>Majdi Khalid, Indrakshi Ray, and Hamidreza Chitsaz</i>	
Mining Distinguishing Customer Focus Sets for Online Shopping Decision Support . . . . .	50
<i>Lu Liu, Lei Duan, Hao Yang, Jyrki Nummenmaa, Guozhu Dong, and Pan Qin</i>	
Community Detection in Networks with Less Significant Community Structure . . . . .	65
<i>Ba-Dung Le, Hung Nguyen, and Hong Shen</i>	
Prediction-Based, Prioritized Market-Share Insight Extraction . . . . .	81
<i>Renato Keshet, Alina Maor, and George Kour</i>	
Interrelationships of Service Orchestrations. . . . .	95
<i>Victor W. Chu, Raymond K. Wong, Fang Chen, and Chi-Hung Chi</i>	
Outlier Detection on Mixed-Type Data: An Energy-Based Approach. . . . .	111
<i>Kien Do, Truyen Tran, Dinh Phung, and Svetha Venkatesh</i>	
Low-Rank Feature Reduction and Sample Selection for Multi-output Regression . . . . .	126
<i>Shichao Zhang, Lifeng Yang, Yonggang Li, Yan Luo, and Xiaofeng Zhu</i>	
Biologically Inspired Pattern Recognition for E-nose Sensors . . . . .	142
<i>Sanad Al-Maskari, Wenping Guo, and Xiaoming Zhao</i>	
Addressing Class Imbalance and Cost Sensitivity in Software Defect Prediction by Combining Domain Costs and Balancing Costs. . . . .	156
<i>Michael J. Siers and Md Zahidul Islam</i>	

Unsupervised Hypergraph Feature Selection with Low-Rank and Self-Representation Constraints . . . . . 172  
*Wei He, Xiaofeng Zhu, Yonggang Li, Rongyao Hu, Yonghua Zhu, and Shichao Zhang*

Improving Cytogenetic Search with GPUs Using Different String Matching Schemes . . . . . 188  
*Chantana Chantrapornchai and Chidchanok Choksuchat*

CEIoT: A Framework for Interlinking Smart Things in the Internet of Things . . . . . 203  
*Ali Shemshadi, Quan Z. Sheng, Yongrui Qin, and Ali Alzubaidi*

Adopting Hybrid Descriptors to Recognise Leaf Images for Automatic Plant Specie Identification . . . . . 219  
*Ali A. Al-kharaz, Xiaohui Tao, Ji Zhang, and Raid Lafta*

Efficient Mining of Pan-Correlation Patterns from Time Course Data . . . . . 234  
*Qian Liu, Jinyan Li, Limsoon Wong, and Kotagiri Ramamohanarao*

Recognizing Daily Living Activity Using Embedded Sensors in Smartphones: A Data-Driven Approach . . . . . 250  
*Wenjie Ruan, Leon Chea, Quan Z. Sheng, and Lina Yao*

Dynamic Reverse Furthest Neighbor Querying Algorithm of Moving Objects . . . . . 266  
*Bohan Li, Chao Zhang, Weitong Chen, Yingbao Yang, Shaohong Feng, Qiqian Zhang, Weiwei Yuan, and Dongjing Li*

**Research Papers**

Relative Neighborhood Graphs Uncover the Dynamics of Social Media Engagement . . . . . 283  
*Natalie Jane de Vries, Ahmed Shamsul Arefin, Luke Mathieson, Benjamin Lucas, and Pablo Moscato*

An Ensemble Approach for Better Truth Discovery . . . . . 298  
*Xiu Susie Fang, Quan Z. Sheng, and Xianzhi Wang*

Single Classifier Selection for Ensemble Learning . . . . . 312  
*Guangtao Wang, Xiaomei Yang, and Xiaoyan Zhu*

Community Detection in Dynamic Attributed Graphs . . . . . 329  
*Gonzalo A. Bello, Steve Harenberg, Abhishek Agrawal, and Nagiza F. Samatova*

Secure Computation of Skyline Query in *MapReduce* . . . . . 345  
*Asif Zaman, Md. Anisuzzaman Siddique, Annisa, and Yasuhiko Morimoto*

Recommending Features of Mobile Applications for Developer . . . . . 361  
*Hong Yu, Yahong Lian, Shuotao Yang, Linlin Tian, and Xiaowei Zhao*

Adaptive Multi-objective Swarm Crossover Optimization for Imbalanced Data Classification . . . . . 374  
*Jinyan Li, Simon Fong, Meng Yuan, and Raymond K. Wong*

Causality-Guided Feature Selection . . . . . 391  
*Mandar S. Chaudhary, Doel L. Gonzalez II, Gonzalo A. Bello, Michael P. Angus, Dhara Desai, Steve Harenberg, P. Murali Doraiswamy, Fredrick H. M. Semazzi, Vipin Kumar, and Nagiza F. Samatova, for the Alzheimer’s Disease Neuroimaging Initiative*

Temporal Interaction Biased Community Detection in Social Networks . . . . . 406  
*Noha Alduaiji, Jianxin Li, Amitava Datta, Xiaolu Lu, and Wei Liu*

Extracting Key Challenges in Achieving Sobriety Through Shared Subspace Learning . . . . . 420  
*Haripriya Harikumar, Thin Nguyen, Santu Rana, Sunil Gupta, Ramachandra Kaimal, and Svetha Venkatesh*

Unified Weighted Label Propagation Algorithm Using Connection Factor . . . . . 434  
*Xin Wang, Songlei Jian, Kai Lu, and Xiaoping Wang*

MetricRec: Metric Learning for Cold-Start Recommendations. . . . . 445  
*Furong Peng, Xuan Lu, Jianfeng Lu, Richard Yi-Da Xu, Cheng Luo, Chao Ma, and Jingyu Yang*

Time Series Forecasting on Engineering Systems Using Recurrent Neural Networks . . . . . 459  
*Dongxu Shao, Tianyou Zhang, Kamal Mannar, and Yue Han*

EDAHT: An Expertise Degree Analysis Model for Mass Comments in the E-Commerce System . . . . . 472  
*Jiang Zhong, You Xiong, Weili Guo, and Jingyi Xie*

A Scalable Document-Based Architecture for Text Analysis. . . . . 481  
*Ciprian-Octavian Truicǎ, Jérôme Darmont, and Julien Velcin*

DAPFFC: Density-Based Affinity Propagation for Parameter Free Clustering . . . . . 495  
*Hanning Yuan, Shuliang Wang, Yang Yu, and Ming Zhong*

Effective Traffic Flow Forecasting Using Taxi and Weather Data . . . . .	507
<i>Xiujuan Xu, Benzhe Su, Xiaowei Zhao, Zhenzhen Xu, and Quan Z. Sheng</i>	
Understanding Behavioral Differences Between Short and Long-Term Drinking Abstainers from Social Media . . . . .	520
<i>Haripriya Harikumar, Thin Nguyen, Sunil Gupta, Santu Rana, Ramachandra Kaimal, and Svetha Venkatesh</i>	
Discovering Trip Hot Routes Using Large Scale Taxi Trajectory Data . . . . .	534
<i>Linjiang Zheng, Qisen Feng, Weining Liu, and Xin Zhao</i>	
Discovering Spatially Contiguous Clusters in Multivariate Geostatistical Data Through Spectral Clustering . . . . .	547
<i>Francky Fouedjio</i>	
On Improving Random Forest for Hard-to-Classify Records . . . . .	558
<i>Md Nasim Adnan and Md Zahidul Islam</i>	
Scholarly Output Graph: A Graphical Article-Level Metric Indicating the Impact of a Scholar’s Publications . . . . .	567
<i>Yu Liu, Dan Lin, Jing Li, and Shimin Shan</i>	
Distributed Lazy Association Classification Algorithm Based on Spark . . . . .	580
<i>Xueming Li, Chaoyang Zhang, Guangwei Chen, Xiaoteng Sun, Qi Zhang, and Haomin Yang</i>	
Event Evolution Model Based on Random Walk Model with Hot Topic Extraction . . . . .	591
<i>Chunzi Wu, Bin Wu, and Bai Wang</i>	
Knowledge-Guided Maximal Clique Enumeration . . . . .	604
<i>Steve Harenberg, Ramona G. Seay, Gonzalo A. Bello, Rada Y. Chirkova, P. Murali Doraiswamy, and Nagiza F. Samatova</i>	
Got a Complaint?- Keep Calm and Tweet It! . . . . .	619
<i>Nitish Mittal, Swati Agarwal, and Ashish Sureka</i>	
Query Classification by Leveraging Explicit Concept Information . . . . .	636
<i>Fang Wang, Ze Yang, Zhoujun Li, and Jianshe Zhou</i>	
Stabilizing Linear Prediction Models Using Autoencoder . . . . .	651
<i>Shivapratap Gopakumar, Truyen Tran, Dinh Phung, and Svetha Venkatesh</i>	
Mining Source Code Topics Through Topic Model and Words Embedding. . .	664
<i>Wei Emma Zhang, Quan Z. Sheng, Ermyas Abebe, M. Ali Babar, and Andi Zhou</i>	

IPC Multi-label Classification Based on the Field Functionality of Patent Documents . . . . . 677  
*Sora Lim and YongJin Kwon*

Unsupervised Component-Wise EM Learning for Finite Mixtures of Skew  $t$ -distributions . . . . . 692  
*Sharon X. Lee and Geoffrey J. McLachlan*

Supervised Feature Selection by Robust Sparse Reduced-Rank Regression . . . 700  
*Rongyao Hu, Xiaofeng Zhu, Wei He, Jilian Zhang, and Shichao Zhang*

PUEPro: A Computational Pipeline for Prediction of Urine Excretory Proteins . . . . . 714  
*Yan Wang, Wei Du, Yanchun Liang, Xin Chen, Chi Zhang, Wei Pang, and Ying Xu*

Partitioning Clustering Based on Support Vector Ranking . . . . . 726  
*Qing Peng, Yan Wang, Ge Ou, Yuan Tian, Lan Huang, and Wei Pang*

Global Recursive Based Node Importance Evaluation . . . . . 738  
*Lu Zhao, Li Xiong, and Shan Xue*

Extreme User and Political Rumor Detection on Twitter . . . . . 751  
*Cheng Chang, Yihong Zhang, Claudia Szabo, and Quan Z. Sheng*

Deriving Public Sector Workforce Insights: A Case Study Using Australian Public Sector Employment Profiles . . . . . 764  
*Shameek Ghosh, Yi Zheng, Thorsten Lammers, Ying Ying Chen, Carolyn Fitzmaurice, Scott Johnston, and Jinyan Li*

Real-Time Stream Mining Electric Power Consumption Data Using Hoeffding Tree with Shadow Features . . . . . 775  
*Simon Fong, Meng Yuen, Raymond K. Wong, Wei Song, and Kyungeun Cho*

Real-Time Investigation of Flight Delays Based on the Internet of Things Data . . . . . 788  
*Abdulwahab Aljubairy, Ali Shemshadi, and Quan Z. Sheng*

**Demo Papers**

IRS-HD: An Intelligent Personalized Recommender System for Heart Disease Patients in a Tele-Health Environment . . . . . 803  
*Raid Lafta, Ji Zhang, Xiaohui Tao, Yan Li, and Vincent S. Tseng*

Sentiment Analysis for Depression Detection on Social Networks . . . . . 807  
*Xiaohui Tao, Xujuan Zhou, Ji Zhang, and Jianming Yong*

Traffic Flow Visualization Using Taxi GPS Data . . . . . 811  
*Xiujuan Xu, Zhenzhen Xu, and Xiaowei Zhao*

Erratum to: Advanced Data Mining and Applications . . . . . E1  
*Jinyan Li, Xue Li, Shuliang Wang, Jianxin Li, and Quan Z. Sheng*

**Author Index** . . . . . 815



<http://www.springer.com/978-3-319-49585-9>

Advanced Data Mining and Applications  
12th International Conference, ADMA 2016, Gold Coast,  
QLD, Australia, December 12-15, 2016, Proceedings  
Li, J.; Li, X.; Wang, S.; Li, J.; Sheng, Q.Z. (Eds.)  
2016, XVI, 817 p. 280 illus., Softcover  
ISBN: 978-3-319-49585-9