Contents

Spotlight Research Papers

Effective Monotone Knowledge Integration in Kernel Support Vector Machines ................................................................. 3
Christopher Bartley, Wei Liu, and Mark Reynolds

Textual Cues for Online Depression in Community and Personal Settings ................................................................. 19
Thin Nguyen, Svetha Venkatesh, and Dinh Phung

Confidence-Weighted Bipartite Ranking ................................................. 35
Majdi Khalid, Indrakshi Ray, and Hamidreza Chitsaz

Mining Distinguishing Customer Focus Sets for Online Shopping Decision Support .......................................................... 50
Lu Liu, Lei Duan, Hao Yang, Jyrki Nummenmaa, Guozhu Dong, and Pan Qin

Community Detection in Networks with Less Significant Community Structure .......................................................... 65
Ba-Dung Le, Hung Nguyen, and Hong Shen

Prediction-Based, Prioritized Market-Share Insight Extraction ................................................................. 81
Renato Keshet, Alina Maor, and George Kour

Interrelationships of Service Orchestrations ................................................................. 95
Victor W. Chu, Raymond K. Wong, Fang Chen, and Chi-Hung Chi

Outlier Detection on Mixed-Type Data: An Energy-Based Approach ................................................................. 111
Kien Do, Truyen Tran, Dinh Phung, and Svetha Venkatesh

Low-Rank Feature Reduction and Sample Selection for Multi-output Regression ................................................................. 126
Shichao Zhang, Lifeng Yang, Yonggang Li, Yan Luo, and Xiaofeng Zhu

Biologically Inspired Pattern Recognition for E-nose Sensors ................................................................. 142
Sanad Al-Maskari, Wenping Guo, and Xiaoming Zhao

Addressing Class Imbalance and Cost Sensitivity in Software Defect Prediction by Combining Domain Costs and Balancing Costs ................................................................. 156
Michael J. Siers and Md Zahidul Islam
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 Truică, Jérôme Darmont, and Julien Velcin

DAPPFC: 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
Xiujuan Xu, Benzhe Su, Xiaowei Zhao, Zhenzhen Xu, and Quan Z. Sheng

Understanding Behavioral Differences Between Short and Long-Term Drinking Abstainers from Social Media ........................................ 520
Haripriya Harikumar, Thin Nguyen, Sunil Gupta, Santu Rana, Ramachandra Kaimal, and Svetha Venkatesh

Discovering Trip Hot Routes Using Large Scale Taxi Trajectory Data ............... 534
Linjiang Zheng, Qisen Feng, Weining Liu, and Xin Zhao

Discovering Spatially Contiguous Clusters in Multivariate Geostatistical Data Through Spectral Clustering ............................................. 547
Francky Fouedjio

On Improving Random Forest for Hard-to-Classify Records .......................... 558
Md Nasim Adnan and Md Zahidul Islam

Scholarly Output Graph: A Graphical Article-Level Metric Indicating the Impact of a Scholar’s Publications ............................................. 567
Yu Liu, Dan Lin, Jing Li, and Shimin Shan

Distributed Lazy Association Classification Algorithm Based on Spark ............ 580
Xueming Li, Chaoyang Zhang, Guangwei Chen, Xiaoteng Sun, Qi Zhang, and Haomin Yang

Event Evolution Model Based on Random Walk Model with Hot Topic Extraction ................................................................. 591
Chunzi Wu, Bin Wu, and Bai Wang

Knowledge-Guided Maximal Clique Enumeration ....................................... 604
Steve Harenberg, Ramona G. Seay, Gonzalo A. Bello, Rada Y. Chirkova, P. Murali Doraiswamy, and Nagiza F. Samatova

Got a Complaint? - Keep Calm and Tweet It! ......................................... 619
Nitish Mittal, Swati Agarwal, and Ashish Sureka

Query Classification by Leveraging Explicit Concept Information .................. 636
Fang Wang, Ze Yang, Zhoujun Li, and Jianshe Zhou

Stabilizing Linear Prediction Models Using Autoencoder ............................ 651
Shivapratap Gopakumar, Truyen Tran, Dinh Phung, and Svetha Venkatesh

Mining Source Code Topics Through Topic Model and Words Embedding .......... 664
Wei Emma Zhang, Quan Z. Sheng, Ermyas Abebe, M. Ali Babar, and Andi Zhou
IPC Multi-label Classification Based on the Field Functionality of Patent Documents

Sora Lim and YongJin Kwon

Unsupervised Component-Wise EM Learning for Finite Mixtures of Skew t-distributions

Sharon X. Lee and Geoffrey J. McLachlan

Supervised Feature Selection by Robust Sparse Reduced-Rank Regression

Rongyao Hu, Xiaofeng Zhu, Wei He, Jilian Zhang, and Shichao Zhang

PUEPro: A Computational Pipeline for Prediction of Urine Excretory Proteins

Yan Wang, Wei Du, Yanchun Liang, Xin Chen, Chi Zhang, Wei Pang, and Ying Xu

Partitioning Clustering Based on Support Vector Ranking

Qing Peng, Yan Wang, Ge Ou, Yuan Tian, Lan Huang, and Wei Pang

Global Recursive Based Node Importance Evaluation

Lu Zhao, Li Xiong, and Shan Xue

Extreme User and Political Rumor Detection on Twitter

Cheng Chang, Yihong Zhang, Claudia Szabo, and Quan Z. Sheng

Deriving Public Sector Workforce Insights: A Case Study Using Australian Public Sector Employment Profiles

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

Simon Fong, Meng Yuen, Raymond K. Wong, Wei Song, and Kyungeoon Cho

Real-Time Investigation of Flight Delays Based on the Internet of Things Data

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

Raid Lafta, Ji Zhang, Xiaohui Tao, Yan Li, and Vincent S. Tseng
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