

# Contents

## Database and Distributed Machine Learning

Querying and Mining Strings Made Easy . . . . .	3
<i>Majed Sahli, Essam Mansour, and Panos Kalnis</i>	
Distributed Training Large-Scale Deep Architectures . . . . .	18
<i>Shang-Xuan Zou, Chun-Yen Chen, Jui-Lin Wu, Chun-Nan Chou, Chia-Chin Tsao, Kuan-Chieh Tung, Ting-Wei Lin, Cheng-Lung Sung, and Edward Y. Chang</i>	
Fault Detection and Localization in Distributed Systems Using Recurrent Convolutional Neural Networks . . . . .	33
<i>Guangyang Qi, Lina Yao, and Anton V. Uzunov</i>	
Discovering Group Skylines with Constraints by Early Candidate Pruning . . .	49
<i>Ming-Yen Lin, Yueh-Lin Lin, and Sue-Chen Hsueh</i>	
Comparing MapReduce-Based k-NN Similarity Joins on Hadoop for High-Dimensional Data. . . . .	63
<i>Přemysl Čech, Jakub Maroušek, Jakub Lokoč, Yasin N. Silva, and Jeremy Starks</i>	
A Higher-Fidelity Frugal Quantile Estimator. . . . .	76
<i>Anis Yazidi, Hugo Lewi Hammer, and B. John Oommen</i>	

## Recommender System

Fair Recommendations Through Diversity Promotion . . . . .	89
<i>Pierre-René Lhérisson, Fabrice Muhlenbach, and Pierre Maret</i>	
A Hierarchical Bayesian Factorization Model for Implicit and Explicit Feedback Data . . . . .	104
<i>ThaiBinh Nguyen and Atsuhiko Takasu</i>	
Empirical Analysis of Factors Influencing Twitter Hashtag Recommendation on Detected Communities . . . . .	119
<i>Areej Alsini, Amitava Datta, Jianxin Li, and Du Huynh</i>	
Group Recommender Model Based on Preference Interaction . . . . .	132
<i>Wei Zheng, Bohan Li, Yanan Wang, Hongzhi Yin, Xue Li, Donghai Guan, and Xiaolin Qin</i>	

Identification of Grey Sheep Users by Histogram Intersection in Recommender Systems . . . . .	148
<i>Yong Zheng, Mayur Agnani, and Mili Singh</i>	

**Social Network and Social Media**

A Feature-Based Approach for the Redefined Link Prediction Problem in Signed Networks . . . . .	165
<i>Xiaoming Li, Hui Fang, and Jie Zhang</i>	
From Mutual Friends to Overlapping Community Detection: A Non-negative Matrix Factorization Approach . . . . .	180
<i>Xingyu Niu, Hongyi Zhang, Micheal R. Lyu, and Irwin King</i>	
Calling for Response: Automatically Distinguishing Situation-Aware Tweets During Crises . . . . .	195
<i>Xiaodong Ning, Lina Yao, Xianzhi Wang, and Boualem Benatallah</i>	
Efficient Revenue Maximization for Viral Marketing in Social Networks . . . .	209
<i>Yuan Su, Xi Zhang, Sihong Xie, Philip S. Yu, and Binxing Fang</i>	
Generating Life Course Trajectory Sequences with Recurrent Neural Networks and Application to Early Detection of Social Disadvantage . . . . .	225
<i>Lin Wu, Michele Haynes, Andrew Smith, Tong Chen, and Xue Li</i>	
FRISK: A Multilingual Approach to Find twitter InterestS via wiKipedia . . .	243
<i>Coriane Nana Jipmo, Gianluca Quercini, and Nacéra Bennacer</i>	
A Solution to Tweet-Based User Identification Across Online Social Networks . . . . .	257
<i>Yongjun Li, Zhen Zhang, and You Peng</i>	

**Machine Learning**

Supervised Feature Selection Algorithm Based on Low-Rank and Manifold Learning . . . . .	273
<i>Yue Fang, Jilian Zhang, Shichao Zhang, Cong Lei, and Xiaoyi Hu</i>	
Mixed Membership Sparse Gaussian Conditional Random Fields . . . . .	287
<i>Jie Yang, Henry C.M. Leung, S.M. Yiu, and Francis Y.L. Chin</i>	
Effects of Dynamic Subspacing in Random Forest . . . . .	303
<i>Md Nasim Adnan and Md Zahidul Islam</i>	
Diversity and Locality in Multi-Component, Multi-Layer Predictive Systems: A Mutual Information Based Approach . . . . .	313
<i>Bassma Al-Jubouri and Bogdan Gabrys</i>	

Hybrid Subspace Mixture Models for Prediction and Anomaly Detection in High Dimensions . . . . .	326
<i>Jenn-Bing Ong and Wee-Keong Ng</i>	

## Classification and Clustering Methods

StruClus: Scalable Structural Graph Set Clustering with Representative Sampling. . . . .	343
<i>Till Schäfer and Petra Mutzel</i>	

Employing Hierarchical Clustering and Reinforcement Learning for Attribute-Based Zero-Shot Classification. . . . .	360
<i>Bin Liu, Li Yao, Junfeng Wu, and Xiaosheng Feng</i>	

Environmental Sound Recognition Using Masked Conditional Neural Networks. . . . .	373
<i>Fady Medhat, David Chesmore, and John Robinson</i>	

Analyzing Performance of Classification Techniques in Detecting Epileptic Seizure. . . . .	386
<i>Mohammad Khubeb Siddiqui, Md Zahidul Islam, and Muhammad Ashad Kabir</i>	

A Framework for Clustering and Dynamic Maintenance of XML Documents . . . . .	399
<i>Ahmed Al-Shammari, Chengfei Liu, Mehdi Naseriparsa, Bao Quoc Vo, Tarique Anwar, and Rui Zhou</i>	

Language-Independent Twitter Classification Using Character-Based Convolutional Networks. . . . .	413
<i>Shiwei Zhang, Xiuzhen Zhang, and Jeffrey Chan</i>	

## Behavior Modeling and User Profiling

Modeling Check-In Behavior with Geographical Neighborhood Influence of Venues . . . . .	429
<i>Thanh-Nam Doan and Ee-Peng Lim</i>	

An Empirical Study on Collective Online Behaviors of Extremist Supporters . . . . .	445
<i>Jung-jae Kim, Yong Liu, Wee Yong Lim, and Vrizlynn L.L. Thing</i>	

Your Moves, Your Device: Establishing Behavior Profiles Using Tensors . . .	460
<i>Eric Falk, J��r��my Charlier, and Radu State</i>	

An Approach for Identifying Author Profiles of Blogs . . . . .	475
<i>Chunxia Zhang, Yu Guo, Jiayu Wu, Shuliang Wang, Zhendong Niu, and Wen Cheng</i>	
Generating Topics of Interests for Research Communities . . . . .	488
<i>Nagendra Kumar, Rahul Utkoor, Bharath K.R. Appareddy, and Manish Singh</i>	
An Evolutionary Approach for Learning Conditional Preference Networks from Inconsistent Examples . . . . .	502
<i>Mohammad Haqqani and Xiaodong Li</i>	

## **Bioinformatic and Medical Data Analysis**

Predicting Clinical Outcomes of Alzheimer’s Disease from Complex Brain Networks. . . . .	519
<i>Xingjuan Li, Yu Li, and Xue Li</i>	
Doctoral Advisor or Medical Condition: Towards Entity-Specific Rankings of Knowledge Base Properties . . . . .	526
<i>Simon Razniewski, Vevake Balaraman, and Werner Nutt</i>	
Multiclass Lung Cancer Diagnosis by Gene Expression Programming and Microarray Datasets. . . . .	541
<i>Hasseeb Azzawi, Jingyu Hou, Russul Alanni, Yong Xiang, Rana Abdu-Aljabar, and Ali Azzawi</i>	
Drug-Drug Interaction Extraction via Recurrent Neural Network with Multiple Attention Layers . . . . .	554
<i>Zibo Yi, Shasha Li, Jie Yu, Yusong Tan, Qingbo Wu, Hong Yuan, and Ting Wang</i>	

## **Spatio-Temporal Data**

People-Centric Mobile Crowdsensing Platform for Urban Design . . . . .	569
<i>Shili Xiang, Lu Li, Si Min Lo, and Xiaoli Li</i>	
Long-Term User Location Prediction Using Deep Learning and Periodic Pattern Mining . . . . .	582
<i>Mun Hou Wong, Vincent S. Tseng, Jerry C.C. Tseng, Sun-Wei Liu, and Cheng-Hung Tsai</i>	
An Intelligent Weighted Fuzzy Time Series Model Based on a Sine-Cosine Adaptive Human Learning Optimization Algorithm and Its Application to Financial Markets Forecasting. . . . .	595
<i>Ruixin Yang, Mingyang Xu, Junyi He, Stephen Ranshous, and Nagiza F. Samatova</i>	

Mobile Robot Scheduling with Multiple Trips and Time Windows . . . . .	608
<i>Shudong Liu, Huayu Wu, Shili Xiang, and Xiaoli Li</i>	

## Natural Language Processing and Text Mining

Feature Analysis for Duplicate Detection in Programming QA Communities . . . . .	623
<i>Wei Emma Zhang, Quan Z. Sheng, Yanjun Shu, and Vanh Khuyen Nguyen</i>	
A Joint Human/Machine Process for Coding Events and Conflict Drivers . . .	639
<i>Bradford Heap, Alfred Krzywicki, Susanne Schmeidl, Wayne Wobcke, and Michael Bain</i>	
Quality Prediction of Newly Proposed Questions in CQA by Leveraging Weakly Supervised Learning . . . . .	655
<i>Yuanhao Zheng, Bifan Wei, Jun Liu, Meng Wang, Weitong Chen, Bei Wu, and Yihe Chen</i>	
Improving Chinese Sentiment Analysis via Segmentation-Based Representation Using Parallel CNN . . . . .	668
<i>Yazhou Hao, Qinghua Zheng, Yangyang Lan, Yufei Li, Meng Wang, Sen Wang, and Chen Li</i>	
Entity Recognition by Distant Supervision with Soft List Constraint . . . . .	681
<i>Hongkui Tu, Zongyang Ma, Aixin Sun, Zhiqiang Xu, and Xiaodong Wang</i>	
Structured Sentiment Analysis . . . . .	695
<i>Abdulqader Almars, Xue Li, Xin Zhao, Ibrahim A. Ibrahim, Weiwei Yuan, and Bohan Li</i>	

## Data Mining Applications

Improving Real-Time Bidding Using a Constrained Markov Decision Process . . . . .	711
<i>Manxing Du, Redouane Sassioui, Georgios Varisteas, Radu State, Mats Brorsson, and Omar Cherkaoui</i>	
PowerLSTM: Power Demand Forecasting Using Long Short-Term Memory Neural Network . . . . .	727
<i>Yao Cheng, Chang Xu, Daisuke Mashima, Vrizlynn L.L. Thing, and Yongdong Wu</i>	
Identifying Unreliable Sensors Without a Knowledge of the Ground Truth in Deceptive Environments . . . . .	741
<i>Anis Yazidi, B. John Oommen, and Morten Goodwin</i>	

Color-Sketch Simulator: A Guide for Color-Based Visual Known-Item Search. . . . .	754
<i>Jakub Lokoč, Anh Nguyen Phuong, Marta Vomlelová, and Chong-Wah Ngo</i>	

## Applications

Making Use of External Company Data to Improve the Classification of Bank Transactions. . . . .	767
<i>Erlend Vollset, Eirik Folkestad, Marius Rise Gallala, and Jon Atle Gulla</i>	
Mining Load Profile Patterns for Australian Electricity Consumers . . . . .	781
<i>Vanh Khuyen Nguyen, Wei Emma Zhang, Quan Z. Sheng, and Jason Merefild</i>	
STA: A Spatio-Temporal Thematic Analytics Framework for Urban Ground Sensing . . . . .	794
<i>Guizi Chen, Liang Yu, Wee Siong Ng, Huayu Wu, and Usha Nanthani Kunasegaran</i>	
Privacy and Utility Preservation for Location Data Using Stay Region Analysis . . . . .	808
<i>Manoranjan Dash and Sin G. Teo</i>	
Location-Aware Human Activity Recognition. . . . .	821
<i>Tam T. Nguyen, Daniel Fernandez, Quy T.K. Nguyen, and Ebrahim Bagheri</i>	

## Demos

SWYSWYK: A New Sharing Paradigm for the Personal Cloud . . . . .	839
<i>Paul Tran-Van, Nicolas Anciaux, and Philippe Pucheral</i>	
Tools and Infrastructure for Supporting Enterprise Knowledge Graphs. . . . .	846
<i>Sumit Bhatia, Nidhi Rajshree, Anshu Jain, and Nitish Aggarwal</i>	
An Interactive Web-Based Toolset for Knowledge Discovery from Short Text Log Data . . . . .	853
<i>Michael Stewart, Wei Liu, Rachell Cardell-Oliver, and Mark Griffin</i>	
Carbon: Forecasting Civil Unrest Events by Monitoring News and Social Media . . . . .	859
<i>Wei Kang, Jie Chen, Jiuyong Li, Jixue Liu, Lin Liu, Grant Osborne, Nick Lothian, Brenton Cooper, Terry Moschou, and Grant Neale</i>	

A System for Querying and Analyzing Urban Regions . . . . .	866
<i>Wee Boon Koh and Xiaolei Li</i>	
Detect Tracking Behavior Among Trajectory Data . . . . .	872
<i>Jianqiu Xu and Jiangang Zhou</i>	
<b>Author Index</b> . . . . .	879

Advanced Data Mining and Applications

13th International Conference, ADMA 2017, Singapore,

November 5–6, 2017, Proceedings

Cong, G.; Peng, W.-C.; Zhang, W.E.; Li, C.; Sun, A. (Eds.)

2017, XVII, 881 p. 264 illus., Softcover

ISBN: 978-3-319-69178-7