

LNAI Series Editors

Randy Goebel

University of Alberta, Edmonton, Canada

Yuzuru Tanaka

Hokkaido University, Sapporo, Japan

Wolfgang Wahlster

DFKI and Saarland University, Saarbrücken, Germany

LNAI Founding Series Editor

Joerg Siekmann

DFKI and Saarland University, Saarbrücken, Germany

More information about this series at <http://www.springer.com/series/1244>

James Bailey · Latifur Khan
Takashi Washio · Gillian Dobbie
Joshua Zhexue Huang · Ruili Wang (Eds.)

Advances in Knowledge Discovery and Data Mining

20th Pacific-Asia Conference, PAKDD 2016
Auckland, New Zealand, April 19–22, 2016
Proceedings, Part I

Editors

James Bailey
The University of Melbourne
Melbourne, VIC
Australia

Latifur Khan
The University of Texas at Dallas
Richardson, TX
USA

Takashi Washio
Osaka University
Osaka
Japan

Gillian Dobbie
University of Auckland
Auckland
New Zealand

Joshua Zhexue Huang
Shenzhen University
Shenzhen
China

Ruili Wang
Massey University
Auckland
New Zealand

ISSN 0302-9743

ISSN 1611-3349 (electronic)

Lecture Notes in Artificial Intelligence

ISBN 978-3-319-31752-6

ISBN 978-3-319-31753-3 (eBook)

DOI 10.1007/978-3-319-31753-3

Library of Congress Control Number: 2016934425

LNCS Sublibrary: SL7 – Artificial Intelligence

© Springer International Publishing Switzerland 2016

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

This Springer imprint is published by Springer Nature

The registered company is Springer International Publishing AG Switzerland

PC Chairs' Preface

PAKDD 2016 is the 20th conference of the Pacific Asia Conference series on Knowledge Discovery and Data Mining. For the first time, the conference is being held in New Zealand. The conference provides a forum for researchers and practitioners to present and discuss new research results and practical applications.

There were 307 papers submitted to PAKDD 2016 and they underwent a rigorous double blind review process. Each paper was reviewed by three Program Committee (PC) members and meta-reviewed by one Senior Program Committee (SPC) member who also conducted discussions with the reviewers. The Program Chairs then considered the recommendations from SPC members, looked into each paper and its reviews, to make final paper selections. At the end, 91 papers were selected for the conference program and proceedings, resulting in an acceptance rate below 30 %, among which 39 papers were assigned as long presentation and 52 papers were assigned as regular presentation. The review process was supported by the Microsoft CMT system.

The conference started with a day of five high-quality workshops and five tutorials. During the next three days, the Technical Program included 19 paper presentation sessions covering various subjects of knowledge discovery and data mining, a data mining contest, and three keynote talks by world-renowned experts.

We would like to thank all the Program Committee members and external reviewers for their hard work to provide timely and comprehensive reviews and recommendations, which were crucial to the final paper selection and production of a high-quality Technical Program. We would also like to express our sincere thanks to Huiping Cao and Jinyan Li together with the individual Workshop Chairs for organizing the workshop program; Hisashi Kashima and Leman Akoglu together with the individual tutorial speakers for arranging the tutorial program; Ruili Wang for compiling all the accepted papers and for working with the Springer team to produce these proceedings.

We hope that participants in the conference in Auckland, as well as subsequent readers of the proceedings, will find the technical program of PAKDD 2016 to be both inspiring and rewarding.

February 2016

James Bailey
Latifur Khan
Takashi Washio

General Chairs' Preface

It is our great pleasure to welcome you to the 20th Conference of the Pacific Asia Conference series on Knowledge Discovery and Data Mining. PAKDD has successfully brought together researchers and developers since 1997, with the purpose of identifying challenging problems facing the development of advanced knowledge discovery. The 20th edition of PAKDD continues this tradition.

We are delighted to present three outstanding keynote speakers: Naren Ramakrishnan from Virginia Tech, Mark Sagar from The University of Auckland, and Svetha Venkatesh from Deakin University.

We are grateful to the many authors who submitted their work to the PAKDD technical program. The Program Committee was led by James Bailey, Latifur Khan and Takashi Washio. A report on the paper selection process appears in the PC Chairs' Preface.

We also thank the other Chairs in the organization team: Muhammad Asif Naeem for running the Contest; David Tse Jung Huang for publicizing to attract submissions and managing the website; Ranjini Swaminathan for handling the registration process and Yun Sing Koh and Ranjini Swaminathan for the local arrangements ensuring the conference runs smoothly.

We are grateful to the sponsors of the conference, Auckland Tourism Events and Economic Development, and BECA, for their generous sponsorship and support, and the PAKDD Steering Committee for its guidance and Best Paper Award, Student Travel Award and Early Career Research Award sponsorship. We would also like to express our gratitude to The University of Auckland for hosting and organizing this conference. Last but not least, our sincere thanks go to all the local team members and volunteer helpers for their hard work to make the event possible. We hope you enjoy PAKDD 2016 and your time in Auckland, New Zealand.

Gillian Dobbie
Joshua Zhexue Huang

Organization

Organizing Committee

General Co-chairs

Gillian Dobbie	University of Auckland, New Zealand
Joshua Zhexue Huang	Shenzhen University, China

Program Committee Co-chairs

James Bailey	The University of Melbourne, Australia
Latifur Khan	University of Texas at Dallas, USA
Takashi Washio	Institute of Scientific and Industrial Research, Osaka University, Japan

Workshop Co-chairs

Huiping Cao	New Mexico State University, USA
Jinyan Li	University of Technology Sydney, Australia

Tutorial Co-chairs

Leman Akoglu	Stony Brook University, USA
Hisashi Kashima	Kyoto University, Japan

Local Arrangements Co-chairs

Yun Sing Koh	University of Auckland, New Zealand
Ranjini Swaminathan	University of Auckland, New Zealand

Proceedings Chair

Ruili Wang	Massey University, New Zealand
------------	--------------------------------

Contest Chair

Muhammad Asif Naeem	AUT University, New Zealand
---------------------	-----------------------------

Publicity and Website Chair

David Tse Jung Huang	University of Auckland, New Zealand
----------------------	-------------------------------------

Registration Chair

Ranjini Swaminathan	University of Auckland, New Zealand
---------------------	-------------------------------------

Steering Committee

Chairs

Tu Bao Ho (Chair)	Japan Advanced Institute of Science and Technology, Japan
Ee-Peng Lim (Co-Chair)	Singapore Management University, Singapore

Treasurer

Graham Williams	Togaware, Australia (see also under Life Members)
-----------------	---

Members

Tu Bao Ho	Japan Advanced Institute of Science and Technology, Japan (Member since 2005, Co-Chair 2012–2014, Chair 2015–2017, Life Member since 2013)
Ee-Peng Lim	Singapore Management University, Singapore (Member since 2006, Co-Chair 2015–2017)
Jaideep Srivastava	University of Minnesota, USA (Member since 2006)
Zhi-Hua Zhou	Nanjing University, China (Member since 2007)
Takashi Washio	Institute of Scientific and Industrial Research, Osaka University (Member since 2008)
Thanaruk Theeramunkong	Thammasat University, Thailand (Member since 2009)
P. Krishna Reddy	International Institute of Information Technology, Hyderabad (IIIT-H), India (Member since 2010)
Joshua Z. Huang	Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China (Member since 2011)
Longbing Cao	Advanced Analytics Institute, University of Technology, Sydney (Member since 2013)
Jian Pei	School of Computing Science, Simon Fraser University (Member since 2013)
Myra Spiliopoulou	Information Systems, Otto-von-Guericke-University Magdeburg (Member since 2013)
Vincent S. Tseng	National Cheng Kung University, Taiwan (Member since 2014)

Life Members

Hiroshi Motoda	AFOSR/AOARD and Osaka University, Japan (Member since 1997, Co-Chair 2001–2003, Chair 2004–2006, Life Member since 2006)
Rao Kotagiri	University of Melbourne, Australia (Member since 1997, Co-Chair 2006–2008, Chair 2009–2011, Life Member since 2007, Treasury Co-Sign since 2006)
Huan Liu	Arizona State University, U.S. (Member since 1998, Treasurer 1998–2000, Life Member since 2012)

Ning Zhong	Maebashi Institute of Technology, Japan (Member since 1999, Life member since 2008)
Masaru Kitsuregawa	Tokyo University, Japan (Member since 2000, Life Member since 2008)
David Cheung	University of Hong Kong, China (Member since 2001, Treasurer 2005–2006, Chair 2006–2008, Life Member since 2009)
Graham Williams	Australian National University, Australia (Member since 2001, Treasurer since 2006, Co-Chair 2009–2011, Chair 2012–2014, Life Member since 2009)
Ming-Syan Chen	National Taiwan University, Taiwan, ROC (Member since 2002, Life Member since 2010)
Kyu-Young Whang	Korea Advanced Institute of Science & Technology, Korea (Member since 2003, Life Member since 2011)
Chengqi Zhang	University of Technology Sydney, Australia (Member since 2004, Life Member since 2012)

Past Members

Hongjun Lu	Hong Kong University of Science and Technology (Member 1997–2005)
Arbee L.P. Chen	National Chengchi University, Taiwan, ROC (Member 2002–2009)
Takao Terano	Tokyo Institute of Technology, Japan (Member 2000–2009)

Program Committee

Senior Program Committee Members

Michael Berthold	University of Konstanz, Germany
Tru Cao	Ho Chi Minh City University of Technology, Vietnam
Ming-Syan Chen	National Taiwan University, Taiwan
Peter Christen	The Australian National University, Australia
Ian Davidson	UC Davis, USA
Guozhu Dong	Wright State University
Bart Goethals	University of Antwerp, Belgium
Xiaohua Hu	Drexel University, USA
Joshua Huang	Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China
George Karypis	University of Minnesota, USA
Ming Li	Nanjing University, China
Jiuyong Li	University of South Australia, Australia
Jinyan Li	University of Technology, Sydney
Chih-Jen Lin	National Taiwan University, Taiwan
Nikos Mamoulis	University of Hong Kong, Hong Kong
Wee Keong Ng	Nanyang Technological University, Singapore

Jian Pei	Simon Fraser University, Canada
Wen-Chih Peng	National Chiao Tung University, Taiwan
Rajeev Raman	University of Leicester, United Kingdom
P. Reddy	International Institute of Information Technology, Hyderabad (IIIT-H), India
Dou Shen	Baidu, China
Kyuseok Shim	Seoul National University, Korea
Myra Spiliopoulou	Otto-von-Guericke-University, Germany
Masashi Sugiyama	The University of Tokyo, Japan
Kai Ming Ting	Federation University, Australia
Hanghang Tong	City University of New York, USA
Vincent S. Tseng	National Cheng Kung University, Taiwan
Koji Tsuda	University of Tokyo, Japan
Wei Wang	University of California at Los Angeles, USA
Haixun Wang	Google, USA
Jianyong Wang	Tsinghua University, China
Xindong Wu	University of Vermont, USA
Xing Xie	Microsoft Research Asia, China
Hui Xiong	Rutgers University, USA
Xifeng Yan	UC Santa Barbara, USA
Jeffrey Yu	The Chinese University of Hong Kong, Hong Kong
Osmar Zaiane	University of Alberta, Canada
Yanchun Zhang	Victoria University, Australia
Min-Ling Zhang	Southeast University, China
Yu Zheng	Microsoft Research Asia, China
Ning Zhong	Maebashi Institute of Technology, Japan
Xiaofang Zhou	The University of Queensland, Australia
Zhi-Hua Zhou	Nanjing University, China

Program Committee Members

Mohammad Al Hasan	Purdue University, USA
Shafiq Alam	University of Auckland, New Zealand
Aijun An	York University, Canada
Gustavo Batista	University of Sao Paulo, Brazil
Chiranjib Bhattachar	Indian Institute of Science, India
Albert Bifet	Universite Paris-Saclay, France
Marut Buranarach	National Electronics and Computer Technology Center, Thailand
Krisztian Buza	Budapest University of Technology and Economics, Hungary
Rui Camacho	Universidade do Porto, Portugal
K. Selcuk Candan	Arizona State University, USA
Jeffrey Chan	RMIT University, Australia
Chia-Hui Chang	National Central University, Taiwan
Muhammad Cheema	Monash University, Australia

Meng Chang Chen	Academia Sinica, Taiwan
Shu-Ching Chen	Florida International University, USA
Songcan Chen	Nanjing University of Aeronautics and Astronautics, China
Zhiyuan Chen	University of Maryland Baltimore County, USA
Yi-Ping Phoebe Chen	La Trobe University, Australia
Zheng Chen	Microsoft Research Asia, China
Chun-Hao Chen	Tamkang University, Taiwan
Enhong Chen	University of Science and Technology of China, China
Ling Chen	University of Technology Sydney, Australia
Jake Chen	Indiana University-Purdue University Indianapolis, USA
Yiu-ming Cheung	Hong Kong Baptist University, Hong Kong
Silvia Chiusano	Politecnico di Torino, Italy
Kun-Ta Chuang	National Cheng Kung University, Taiwan
Bruno Cremilleux	Universite de Caen, France
Alfredo Cuzzocrea	ICAR-CNR and University of Calabria, Italy
Bing Tian Dai	Singapore Management University, Singapore
Dao-Qing Dai	Sun Yat-Sen University, China
Xuan-Hong Dang	UC Santa Barbara, USA
Anne Denton	North Dakota State University, USA
Bolin Ding	Microsoft Research, USA
Wei Ding	University of Massachusetts Boston, USA
Dejing Dou	University of Oregon, USA
Liang Du	Chinese Academy of Sciences, China
Lei Duan	Sichuan University
Christoph Eick	University of Houston
Vladimir Estivill-Castro	Griffith University
Philippe Fournier-Viger	University of Moncton, Canada
Dragan Gamberger	Rudjer Boskovic Institute, Croatia
Junbin Gao	Charles Sturt University, Australia
Jun Gao	Peking University, China
Yong Guan	Iowa State University, USA
Stephan Gunnemann	TU Munich, Germany
Sunil Gupta	Deakin University
Michael Hahsler	Southern Methodist University, USA
Saotshi Hara	IBM Research Tokyo, Japan
Choochart Haruechaiy	National Electronics and Computer Technology Center, Thailand
Jingrui He	IBM Research, USA
Shoji Hirano	Shimane University, Japan
Jaakko Hollmen	Aalto University, Finland
Tzung-Pei Hong	National University of Kaohsiung, Taiwan
Michael Houle	NII, Japan
Wynne Hsu	National University of Singapore, Singapore
Jun Huan	University of Kansas, USA
Jen-Wei Huang	National Cheng Kung University, Taiwan
Sheng-Jun Huang	NUAA, China

Nam Huynh	Japan Advanced Institute of Science and Technology, Japan
Akihiro Inokuchi	Kwansei Gakuin University
Motoharu Iwata	NTT Communication Science Laboratories, Japan
Sanjay Jain	National University of Singapore, Singapore
Toshihiro Kamishima	National Institute of Advanced Industrial Science and Technology, Japan
Murat Kantarcioglu	University of Texas at Dallas, USA
Hung-Yu Kao	National Cheng Kung University, Taiwan
Yoshinobu Kawahara	Osaka University, Japan
Irena Koprinka	University of Sydney, Australia
Walter Kosters	Universiteit Leiden, Netherlands
Marzena Kryszkiewicz	Warsaw University of Technology, Poland
Satoshi Kurihara	Osaka University, Japan
Hady Lauw	Singapore Management University, Singapore
Wang-Chien Lee	Pennsylvania State University, USA
Yue-Shi Lee	Ming Chuan University, Taiwan
Philippe Lenca	Telecom Bretagne, France
Carson K. Leung	University of Manitoba, Canada
Geng Li	Oracle Corporation, USA
Chun-hung Li	Hong Kong Baptist University, Hong Kong
Zhenhui Li	Pennsylvania State University, USA
Yidong Li	Beijing Jiaotong University, China
Xiaoli Li	Institute for Infocomm Research, Singapore
Wu-Jun Li	Nanjing University, China
Xuelong Li	University of London, UK
Hsuan-Tien Lin	National Taiwan University, Taiwan
Jerry Chun-Wei Lin	Harbin Institute of Technology Shenzhen, China
Xu-Ying Liu	Southeast University, China
Wei Liu	University of Technology Sydney, Australia
Qingshan Liu	NLPR Institute of Automation Chinese Academy of Science, China
Hua Lu	Aalborg University, Denmark
Jun Luo	Hua Wei Noahs Ark Lab, Hong Kong
Shuai Ma	Beihang University, China
Marco Maggini	Universita degli Studi di Siena, Italy
Hiroshi Mamitsuka	Kyoto University, Japan
Giuseppe Manco	Universita' della Calabria, Italy
Florent Masegla	INRIA, France
Mohammad Mehedy Masud	United Arab Emirates University
Tomoko Matsui	Institute of Statistical Mathematics, Japan
Xiaofeng Meng	Renmin University of China, China
Nguyen Le Minh	JAIST, Japan
Pabitra Mitra	Indian Institute of Technology Kharagpur, India
Yang-Sae Moon	Kangwon National University, Korea

Yasuhiko Morimoto	Hiroshima University, Japan
Emmanuel Mueller	Hasso-Plattner-Institut
J. Nath	Indian Institute of Technology, India
Richi Nayak	Queensland University of Technologies, Australia
Wilfred Ng	Hong Kong University of Science and Technology, Hong Kong
Xuan Vinh Nguyen	University of Melbourne, Australia
Ngoc-Thanh Nguyen	Wroclaw University of Technology, Poland
Hung-Son Nguyen	University of Warsaw, Poland
Duc Dung Nguyen	Institute of Information Technology, Vietnam
Tadashi Nomoto	National Institute of Japanese Literature, Japan
Manabu Okumura	Japan Advanced Institute of Science and Technology, Japan
Salvatore Orlando	University of Venice, Italy
Jia-Yu Pan	Google, USA
Shirui Pan	University of Technology Sydney, Australia
Dhaval Patel	Indian Institute of Technology, Roorkee, India
Dinh Phung	Deakin University, Australia
Vincenzo Piuri	Universita degli Studi di Milano, Italy
Chedy Raissi	INRIA, France
Santu Rana	Deakin University, Australia
Chandan Reddy	Wayne State University, USA
Patricia Riddle	University of Auckland, New Zealand
Hiroshi Sakamoto	Kyushu Institute of Technology, Japan
Yi-Dong Shen	Chinese Academy of Sciences, China
Jialie Shen	Singapore Management University, Singapore
Hong Shen	Adelaide University, Australia
Masashi Shimbo	Nara Institute of Science and Technology, Japan
Andrzej Skowron	University of Warsaw, Poland
Mahito Sugiyama	Osaka University, Japan
Aixin Sun	Nanyang Technological University, Singapore
Yasuo Tabei	Presto, Japan Science and Technology Agency
David Taniar	Monash University, Australia
Xiaohui (Daniel) Tao	The University of Southern Queensland, Australia
Khoat Than	Hanoi University of Science and Technology, Vietnam
Jeffrey Ullman	Stanford University, USA
Ranga Vatsavai	North Carolina University
Kitsana Waiyamai	Kasetsart University, Thailand
Lipo Wang	Nanyang Technological University, Singapore
Jason Wang	New Jersey Science and Technology University, USA
Xin Wang	University of Calgary, Canada
Raymond Chi-Wing Wong	Hong Kong University of Science and Technology, Hong Kong
Xintao Wu	University of Arkansas, USA
Jia Wu	University of Technology Sydney, Australia
Junjie Wu	Beihang University, China

Guandong Xu	University of Technology Sydney, Australia
Takehisa Yairi	University of Tokyo, Japan
De-Nian Yang	Academia Sinica, Taiwan
Min Yao	Zhejiang University, China
Mi-Yen Yeh	Academia Sinica, Taiwan
Tetsuya Yoshida	Nara Womens University
Yang Yu	Nanjing University, China
De-Chuan Zhan	Nanjing University, China
Daoqiang Zhang	Nanjing University of Aeronautics and Astronautics, China
Du Zhang	California State University, USA
Bo Zhang	Tsinghua University, China
Junping Zhang	Fudan University, China
Wenjie Zhang	University of New South Wales, Australia
Ying Zhang	University of New South Wales, Australia
Zhongfei Zhang	Binghamton University, USA
Zili Zhang	Deakin University, Australia
Mengjie Zhang	Victoria University of Wellington, New Zealand
Zhao Zhang	Soochow University, China
Xiuzhen Zhang	RMIT University, Australia
Peixiang Zhao	Florida State University, USA
Shuigeng Zhou	Fudan University, China
Bin Zhou	University of Maryland Baltimore County, USA
Feida Zhu	Singapore Management University, Singapore
Xingquan Zhu	Florida Atlantic University, USA
Arthur Zimek	Ludwig-Maximilians-University Munchen, Germany

Sponsors

**Auckland
Tourism, Events and
Economic Development**
An Auckland Council Organisation



Contents – Part I

Classification

Joint Classification with Heterogeneous Labels Using Random Walk with Dynamic Label Propagation	3
<i>Yongxin Liao, Shenxi Yuan, Jian Chen, Qingyao Wu, and Bin Li</i>	
Hybrid Sampling with Bagging for Class Imbalance Learning	14
<i>Yang Lu, Yiu-ming Cheung, and Yuan Yan Tang</i>	
Sparse Adaptive Multi-hyperplane Machine	27
<i>Khanh Nguyen, Trung Le, Vu Nguyen, and Dinh Phung</i>	
Exploring Heterogeneous Product Networks for Discovering Collective Marketing Hyping Behavior	40
<i>Qinzhe Zhang, Qin Zhang, Guodong Long, Peng Zhang, and Chengqi Zhang</i>	
Optimal Training and Efficient Model Selection for Parameterized Large Margin Learning	52
<i>Yuxun Zhou, Jae Yeon Baek, Dan Li, and Costas J. Spanos</i>	
Locally Weighted Ensemble Learning for Regression	65
<i>Man Yu, Zongxia Xie, Hong Shi, and Qinghua Hu</i>	
Reliable Confidence Predictions Using Conformal Prediction	77
<i>Henrik Linusson, Ulf Johansson, Henrik Boström, and Tuve Löfström</i>	
Grade Prediction with Course and Student Specific Models	89
<i>Agoritsa Polyzou and George Karypis</i>	
Flexible Transfer Learning Framework for Bayesian Optimisation	102
<i>Tinu Theckel Joy, Santu Rana, Sunil Kumar Gupta, and Svetha Venkatesh</i>	
A Simple Unlearning Framework for Online Learning Under Concept Drifts	115
<i>Sheng-Chi You and Hsuan-Tien Lin</i>	
User-Guided Large Attributed Graph Clustering with Multiple Sparse Annotations	127
<i>Jianping Cao, Senzhang Wang, Fengcai Qiao, Hui Wang, Feiyue Wang, and Philip S. Yu</i>	

Early-Stage Event Prediction for Longitudinal Data	139
<i>Mahtab J. Fard, Sanjay Chawla, and Chandan K. Reddy</i>	
Toxicity Prediction in Cancer Using Multiple Instance Learning in a Multi-task Framework	152
<i>Cheng Li, Sunil Gupta, Santu Rana, Wei Luo, Svetha Venkatesh, David Ashely, and Dinh Phung</i>	
Shot Boundary Detection Using Multi-instance Incremental and Decremental One-Class Support Vector Machine	165
<i>Hanhe Lin, Jeremiah D. Deng, and Brendon J. Woodford</i>	
Will I Win Your Favor? Predicting the Success of Altruistic Requests	177
<i>Hsun-Ping Hsieh, Rui Yan, and Cheng-Te Li</i>	
Feature Extraction and Pattern Mining	
Unsupervised and Semi-supervised Dimensionality Reduction with Self-Organizing Incremental Neural Network and Graph Similarity Constraints.	191
<i>Zhiyang Xiang, Zhu Xiao, Yourong Huang, Dong Wang, Bin Fu, and Wenjie Chen</i>	
Cross-View Feature Hashing for Image Retrieval	203
<i>Wei Wu, Bin Li, Ling Chen, and Chengqi Zhang</i>	
Towards Automatic Generation of Metafeatures	215
<i>Fábio Pinto, Carlos Soares, and João Mendes-Moreira</i>	
Hash Learning with Convolutional Neural Networks for Semantic Based Image Retrieval	227
<i>Jinma Guo, Shifeng Zhang, and Jianmin Li</i>	
Bayesian Group Feature Selection for Support Vector Learning Machines . . .	239
<i>Changde Du, Changying Du, Shandian Zhe, Ali Luo, Qing He, and Guoping Long</i>	
Active Distance-Based Clustering Using K-Medoids	253
<i>Amin Aghaee, Mehrdad Ghadiri, and Mahdieh Soleymani Baghshah</i>	
Analyzing Similarities of Datasets Using a Pattern Set Kernel.	265
<i>A. Ibrahim, P.S. Sastry, and Shivakumar Sastry</i>	
Significant Pattern Mining with Confounding Variables	277
<i>Aika Terada, David duVerle, and Koji Tsuda</i>	

Building Compact Lexicons for Cross-Domain SMT by Mining Near-Optimal Pattern Sets	290
<i>Pankaj Singh, Ashish Kulkarni, Himanshu Ojha, Vishwajeet Kumar, and Ganesh Ramakrishnan</i>	
Forest CERN: A New Decision Forest Building Technique	304
<i>Md. Nasim Adnan and Md. Zahidul Islam</i>	
Sparse Logistic Regression with Logical Features	316
<i>Yuan Zou and Teemu Roos</i>	
A Nonlinear Label Compression and Transformation Method for Multi-label Classification Using Autoencoders	328
<i>Jörg Wicker, Andrey Tyukin, and Stefan Kramer</i>	
Preconditioning an Artificial Neural Network Using Naive Bayes	341
<i>Nayyar A. Zaidi, François Petitjean, and Geoffrey I. Webb</i>	
OCEAN: Fast Discovery of High Utility Occupancy Itemsets	354
<i>Bilong Shen, Zhaoduo Wen, Ying Zhao, Dongliang Zhou, and Weimin Zheng</i>	

Graph and Network Data

Leveraging Emotional Consistency for Semi-supervised Sentiment Classification	369
<i>Minh Luan Nguyen</i>	
The Effect on Accuracy of Tweet Sample Size for Hashtag Segmentation Dictionary Construction	382
<i>Laurence A.F. Park and Glenn Stone</i>	
Social Identity Link Across Incomplete Social Information Sources Using Anchor Link Expansion	395
<i>Yuxiang Zhang, Lulu Wang, Xiaoli Li, and Chunjing Xiao</i>	
Discovering the Network Backbone from Traffic Activity Data.	409
<i>Sanjay Chawla, Kiran Garimella, Aristides Gionis, and Dominic Tsang</i>	
A Fast and Complete Enumeration of Pseudo-Cliques for Large Graphs.	423
<i>Hongjie Zhai, Makoto Haraguchi, Yoshiaki Okubo, and Etsuji Tomita</i>	
Incorporating Heterogeneous Information for Mashup Discovery with Consistent Regularization	436
<i>Yao Wan, Liang Chen, Qi Yu, Tingting Liang, and Jian Wu</i>	
Link Prediction in Schema-Rich Heterogeneous Information Network	449
<i>Xiaohuan Cao, Yuyan Zheng, Chuan Shi, Jingzhi Li, and Bin Wu</i>	

FastStep: Scalable Boolean Matrix Decomposition 461
Miguel Araujo, Pedro Ribeiro, and Christos Faloutsos

Applications

An Expert-in-the-loop Paradigm for Learning Medical Image Grouping 477
*Xuan Guo, Qi Yu, Rui Li, Cecilia Ovesdotter Alm, Cara Calvelli,
Pengcheng Shi, and Anne Haake*

Predicting Post-operative Visual Acuity for LASIK Surgeries 489
Manish Gupta, Prashant Gupta, Pravin K. Vaddavalli, and Asra Fatima

LBMF: Log-Bilinear Matrix Factorization for Recommender Systems 502
Yunhui Guo, Xin Wang, and Congfu Xu

An Empirical Study on Hybrid Recommender System
with Implicit Feedback 514
Sunhwan Lee, Anca Chandra, and Divyesh Jadav

Who Will Be Affected by Supermarket Health Programs? Tracking
Customer Behavior Changes via Preference Modeling 527
Ling Luo, Bin Li, Shlomo Berkovsky, Irena Koprinksa, and Fang Chen

TrafficWatch: Real-Time Traffic Incident Detection and Monitoring
Using Social Media. 540
Hoang Nguyen, Wei Liu, Paul Rivera, and Fang Chen

Automated Setting of Bus Schedule Coverage Using Unsupervised
Machine Learning 552
Jihed Khiari, Luis Moreira-Matias, Vitor Cerqueira, and Oded Cats

Effective Local Metric Learning for Water Pipe Assessment. 565
*Mojgan Ghanavati, Raymond K. Wong, Fang Chen, Yang Wang,
and Simon Fong*

Classification with Quantification for Air Quality Monitoring 578
*Sanad Al-Maskari, Eve Bélisle, Xue Li, Sébastien Le Digabel,
Amin Nawahda, and Jiang Zhong*

Predicting Unknown Interactions Between Known Drugs and Targets
via Matrix Completion. 591
Qing Liao, Naiyang Guan, Chengkun Wu, and Qian Zhang

Author Index 605

Contents – Part II

Spatiotemporal and Image Data

Denoising Time Series by Way of a Flexible Model for Phase Space Reconstruction	3
<i>Minhazul Islam Sk and Arunava Banerjee</i>	
Distributed Sequential Pattern Mining in Large Scale Uncertain Databases . . .	17
<i>Jiaqi Ge and Yuni Xia</i>	
DeepCare: A Deep Dynamic Memory Model for Predictive Medicine	30
<i>Trang Pham, Truyen Tran, Dinh Phung, and Svetha Venkatesh</i>	
Indoor Positioning System for Smart Homes Based on Decision Trees and Passive RFID	42
<i>Frédéric Bergeron, Kevin Bouchard, Sébastien Gaboury, Sylvain Giroux, and Bruno Bouchard</i>	
Deep Feature Extraction from Trajectories for Transportation Mode Estimation.	54
<i>Yuki Endo, Hiroyuki Toda, Kyosuke Nishida, and Akihisa Kawanobe</i>	
Online Learning for Accurate Real-Time Map Matching	67
<i>Biwei Liang, Tengjiao Wang, Shun Li, Wei Chen, Hongyan Li, and Kai Lei</i>	
Multi-hypergraph Incidence Consistent Sparse Coding for Image Data Clustering.	79
<i>Xiaodong Feng, Sen Wu, Wenjun Zhou, and Zhiwei Tang</i>	
Robust Multi-view Manifold Ranking for Image Retrieval	92
<i>Jun Wu, Jianbo Yuan, and Jiebo Luo</i>	
Image Representation Optimization Based on Locally Aggregated Descriptors.	104
<i>Shijiang Chen, Guiguang Ding, Chenxiao Li, and Yuchen Guo</i>	
Reusing Extracted Knowledge in Genetic Programming to Solve Complex Texture Image Classification Problems.	117
<i>Muhammad Iqbal, Bing Xue, and Mengjie Zhang</i>	

Personal Credit Profiling via Latent User Behavior Dimensions on Social Media	130
<i>Guangming Guo, Feida Zhu, Enhong Chen, Le Wu, Qi Liu, Yingling Liu, and Minghui Qiu</i>	
Linear Upper Confidence Bound Algorithm for Contextual Bandit Problem with Piled Rewards	143
<i>Kuan-Hao Huang and Hsuan-Tien Lin</i>	
Incremental Hierarchical Clustering of Stochastic Pattern-Based Symbolic Data	156
<i>Xin Xu, Jiaheng Lu, and Wei Wang</i>	
Computing Hierarchical Summary of the Data Streams	168
<i>Zubair Shah, Abdun Naser Mahmood, and Michael Barlow</i>	
Anomaly Detection and Clustering	
Unsupervised Parameter Estimation for One-Class Support Vector Machines	183
<i>Zahra Ghafoori, Sutharshan Rajasegarar, Sarah M. Erfani, Shanika Karunasekera, and Christopher A. Leckie</i>	
Frequent Pattern Outlier Detection Without Exhaustive Mining	196
<i>Arnaud Giacometti and Arnaud Soulet</i>	
Ensembles of Interesting Subgroups for Discovering High Potential Employees	208
<i>Girish Keshav Palshikar, Kuleshwar Sahu, and Rajiv Srivastava</i>	
Dynamic Grouped Mixture Models for Intermittent Multivariate Sensor Data	221
<i>Naoya Takeishi, Takehisa Yairi, Naoki Nishimura, Yuta Nakajima, and Noboru Takata</i>	
Parallel Discord Discovery	233
<i>Tian Huang, Yongxin Zhu, Yishu Mao, Xinyang Li, Mengyun Liu, Yafei Wu, Yajun Ha, and Gillian Dobbie</i>	
Dboost: A Fast Algorithm for DBSCAN-Based Clustering on High Dimensional Data	245
<i>Yuxiao Zhang, Xiaorong Wang, Bingyang Li, Wei Chen, Tengjiao Wang, and Kai Lei</i>	
A Precise and Robust Clustering Approach Using Homophilic Degrees of Graph Kernel	257
<i>Haolin Yang, Deli Zhao, Lele Cao, and Fuchun Sun</i>	

Constraint Based Subspace Clustering for High Dimensional Uncertain Data	271
<i>Xianchao Zhang, Lu Gao, and Hong Yu</i>	
A Clustering-Based Framework for Incrementally Repairing Entity Resolution	283
<i>Qing Wang, Jingyi Gao, and Peter Christen</i>	
Adaptive Seeding for Gaussian Mixture Models	296
<i>Johannes Blömer and Kathrin Bujna</i>	
A Greedy Algorithm to Construct L1 Graph with Ranked Dictionary	309
<i>Shuchu Han and Hong Qin</i>	

Novel Models and Algorithms

A Rule Based Open Information Extraction Method Using Cascaded Finite-State Transducer	325
<i>Hailun Lin, Yuanzhuo Wang, Peng Zhang, Weiping Wang, Yinliang Yue, and Zheng Lin</i>	
Active Learning Based Entity Resolution Using Markov Logic.	338
<i>Jeffrey Fisher, Peter Christen, and Qing Wang</i>	
Modeling Adversarial Learning as Nested Stackelberg Games.	350
<i>Yan Zhou and Murat Kantarcioglu</i>	
Fast and Semantic Measurements on Collaborative Tagging Quality	363
<i>Yuqing Sun, Haiqi Sun, and Reynold Cheng</i>	
Matrices, Compression, Learning Curves: Formulation, and the GROUPNTEACH Algorithms.	376
<i>Bryan Hooi, Hyun Ah Song, Evangelos Papalexakis, Rakesh Agrawal, and Christos Faloutsos</i>	
Privacy Aware K-Means Clustering with High Utility	388
<i>Thanh Dai Nguyen, Sunil Gupta, Santu Rana, and Svetha Venkatesh</i>	
Secure k -NN Query on Encrypted Cloud Data with Limited Key-Disclosure and Offline Data Owner.	401
<i>Youwen Zhu, Zhikuan Wang, and Yue Zhang</i>	
Hashing-Based Distributed Multi-party Blocking for Privacy-Preserving Record Linkage	415
<i>Thilina Ranbaduge, Dinusha Vatsalan, Peter Christen, and Vassilios Verykios</i>	

Text Mining and Recommender Systems

Enabling Hierarchical Dirichlet Processes to Work Better for Short Texts at Large Scale.	431
<i>Khai Mai, Sang Mai, Anh Nguyen, Ngo Van Linh, and Khoat Than</i>	
Query-Focused Multi-document Summarization Based on Concept Importance.	443
<i>Hai-Tao Zheng, Ji-Min Guo, Yong Jiang, and Shu-Tao Xia</i>	
Mirror on the Wall: Finding Similar Questions with Deep Structured Topic Modeling	454
<i>Arpita Das, Manish Shrivastava, and Manoj Chinnakotla</i>	
An Efficient Dynamic Programming Algorithm for STR-IC-STR-IC-LCS Problem.	466
<i>Daxin Zhu, Yingjie Wu, and Xiaodong Wang</i>	
Efficient Page-Level Data Extraction via Schema Induction and Verification	478
<i>Chia-Hui Chang, Tian-Sheng Chen, Ming-Chuan Chen, and Hung-Li Ding</i>	
Transfer-Learning Based Model for Reciprocal Recommendation	491
<i>Chia-Hsin Ting, Hung-Yi Lo, and Shou-De Lin</i>	
Enhanced SVD for Collaborative Filtering	503
<i>Xin Guan, Chang-Tsun Li, and Yu Guan</i>	
Social Group Based Video Recommendation Addressing the Cold-Start Problem	515
<i>Chunfeng Yang, Yipeng Zhou, Liang Chen, Xiaopeng Zhang, and Dah Ming Chiu</i>	
FeRoSA: A Faceted Recommendation System for Scientific Articles.	528
<i>Tanmoy Chakraborty, Amrith Krishna, Mayank Singh, Niloy Ganguly, Pawan Goyal, and Animesh Mukherjee</i>	
Dual Similarity Regularization for Recommendation	542
<i>Jing Zheng, Jian Liu, Chuan Shi, Fuzhen Zhuang, Jingzhi Li, and Bin Wu</i>	
Collaborative Deep Ranking: A Hybrid Pair-Wise Recommendation Algorithm with Implicit Feedback.	555
<i>Haochao Ying, Liang Chen, Yuwen Xiong, and Jian Wu</i>	
Author Index	569