

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>

Petra Perner (Ed.)

Machine Learning and Data Mining in Pattern Recognition

12th International Conference, MLDM 2016
New York, NY, USA, July 16–21, 2016
Proceedings

Editor
Petra Perner
Institute of Computer Vision
and applied Computer Sciences, IBaI
Leipzig, Saxony
Germany

ISSN 0302-9743 ISSN 1611-3349 (electronic)
Lecture Notes in Artificial Intelligence
ISBN 978-3-319-41919-0 ISBN 978-3-319-41920-6 (eBook)
DOI 10.1007/978-3-319-41920-6

Library of Congress Control Number: 2016943411

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

Preface

The 12th event of the International Conference on Machine Learning and Data Mining MLDM 2016 was held in New York (www.mldm.de) running under the umbrella of the World Congress “The Frontiers in Intelligent Data and Signal Analysis, DSA2016” (www.worldcongressdsa.com).

For this edition the Program Committee received 169 submissions. After the peer-review process, we accepted 56 high-quality papers for oral presentation. The topics range from theoretical topics for classification, clustering, association rule, and pattern mining to specific data-mining methods for the different multimedia data types such as image mining, text mining, video mining, and Web mining. Extended versions of selected papers will appear in the international journal *Transactions on Machine Learning and Data Mining* (www.ibai-publishing.org/journal/mldm).

A tutorial on “Data Mining,” a tutorial on “Case-Based Reasoning,” a tutorial on “Intelligent Image Interpretation and Computer Vision in Medicine, Biotechnology, Chemistry, and Food Industry,” and a tutorial on “Standardization in Immunofluorescence” were held before the conference.

We were pleased to give out the best paper award for MLDM for the fifth time this year. Three announcements are mentioned at www.mldm.de. The final decision was made by the Best Paper Award Committee based on the presentation by the authors and the discussion with the auditorium. The ceremony took place during the banquet. This prize is sponsored by ibai solutions (www.ibai-solutions.de), one of the leading companies in data mining for marketing, Web mining, and e-commerce.

We would like to thank all reviewers for their highly professional work and their effort in reviewing the papers. We would also like to thank the members of Institute of Applied Computer Sciences, Leipzig, Germany (www.ibai-institut.de), who handled the conference as secretariat. We appreciate the help and understanding of the editorial staff at Springer, and in particular Alfred Hofmann, who supported the publication of these proceedings in the LNAI series.

Last, but not least, we wish to thank all the speakers and participants who contributed to the success of the conference. We hope to see you in 2017 in New York at the next World Congress (www.worldcongressdsa.com) on “The Frontiers in Intelligent Data and Signal Analysis, DSA2017,” which combines under its roof the following three events: International Conferences on Machine Learning and Data Mining (MLDM), the Industrial Conference on Data Mining (ICDM), and the International Conference on Mass Data Analysis of Signals and Images in Medicine, Biotechnology, Chemistry and Food Industry (MDA).

Organization

Program Chair

Petra Perner IBAI Leipzig, Germany

Program Committee

Sergey V. Ablameyko	Belarus State University, Belarus
Patrick Bouthemy	Inria-VISTA, France
Michelangelo Ceci	University of Bari, Italy
Xiaoqing Ding	Tsinghua University, China
Christoph F. Eick	University of Houston, USA
Ana Fred	Technical University of Lisbon, Portugal
Giorgio Giacinto	University of Cagliari, Italy
Makato Haraguchi	Hokkaido University of Sapporo, Japan
Dimitrios A. Karras	Chalkis Institute of Technology, Greece
Adam Krzyzak	Concordia University, Montreal, Canada
Thang V. Pham	Intelligent Systems Lab Amsterdam (ISLA), The Netherlands
Linda Shapiro	University of Washington, USA
Tamas Sziranyi	MTA-SZTAKI, Hungary
Alexander Ulanov	HP Labs, Russian Federation
Patrick Wang	Northeastern University, USA

Additional Reviewers

Zejin Jason Ding	Hewlett Packard Enterprise
Long Ma	Georgia State University
Lumin Zhang	Facebook

Contents

Evolving a Low Price Recovery Strategy for Distressed Securities	1
<i>Robert E. Marmelstein, Alexander L. Hunt, and Christopher Eroh</i>	
A Spectral Clustering Based Outlier Detection Technique	15
<i>Yuan Wang, Xiaochun Wang, and Xia Li Wang</i>	
A Learning Framework to Improve Unsupervised Gene Network Inference. . .	28
<i>Turki Turki, William Bassett, and Jason T.L. Wang</i>	
A Closed Frequent Subgraph Mining Algorithm in Unique Edge Label Graphs.	43
<i>Nour El Islem Karabadji, Sabeur Aridhi, and Hassina Seridi</i>	
Using Glocal Event Alignment for Comparing Sequences of Significantly Different Lengths	58
<i>Vinh-Trung Luu, Mathis Ripken, Germain Forestier, Frédéric Fondement, and Pierre-Alain Muller</i>	
Using Support Vector Machines for Intelligent Service Agents Decision Making	73
<i>Babak Khosravifar and Mohamed Bouguessa</i>	
List Price Optimization Using Customized Decision Trees	88
<i>Kiran Rama, Shashank Shekhar, John Kiran, Raghava Rau, Sam Pritchett, Anit Bhandari, and Parag Chitalia</i>	
AdaMS: Adaptive Mountain Silhouette Extraction from Images	98
<i>Daniel Braun, Michael Singhof, and Stefan Conrad</i>	
K-Means over Incomplete Datasets Using Mean Euclidean Distance	113
<i>Loai AbdAllah and Ilan Shimshoni</i>	
A Time Series Model of the Writing Process	128
<i>Zeev Volkovich</i>	
Semantic Aware Bayesian Network Model for Actionable Knowledge Discovery in Linked Data	143
<i>Hasanein Alharbi and Mohamad Saraee</i>	
Driving Style Identification with Unsupervised Learning	155
<i>Vladimir Nikulin</i>	

A Hybrid Framework for News Clustering Based on the DBSCAN-Martingale and LDA.	170
<i>Ilias Gialampoukidis, Stefanos Vrochidis, and Ioannis Kompatsiaris</i>	
Diagnosis of Metabolic Syndrome: A Diversity Based Hybrid Model	185
<i>Nahla Barakat</i>	
EFIM-Closed: Fast and Memory Efficient Discovery of Closed High-Utility Itemsets	199
<i>Philippe Fournier-Viger, Souleymane Zida, Jerry Chun-Wei Lin, Cheng-Wei Wu, and Vincent S. Tseng</i>	
Fast Detection of Block Boundaries in Block-Wise Constant Matrices	214
<i>Vincent Brault, Julien Chiquet, and Céline Lévy-Leduc</i>	
Inferring Censored Geo-Information with Non-Representative Data	229
<i>Yu Zhang, Tse-Chuan Yang, and Stephen A. Matthews</i>	
Efficient Mining of Weighted Frequent Itemsets in Uncertain Databases	236
<i>Jerry Chun-Wei Lin, Wensheng Gan, Philippe Fournier-Viger, and Tzung-Pei Hong</i>	
Automatic Detection of Latent Common Clusters of Groups in MultiGroup Regression	251
<i>Minhazul Islam Sk and Arunava Banerjee</i>	
Content Based Identification of Talk Show Videos Using Audio Visual Features	267
<i>Atta Muhammad and Sher Muhammad Daudpota</i>	
IncMSTS-PP: An Algorithm to the Incremental Extraction of Significant Sequences from Environmental Sensor Data	284
<i>Carlos Roberto Silveira Junior, Marcela Xavier Ribeiro, and Marilde Terezinha Prado Santos</i>	
DSCo: A Language Modeling Approach for Time Series Classification	294
<i>Daoyuan Li, Li Li, Tegawendé F. Bissyandé, Jacques Klein, and Yves Le Traon</i>	
Statistical Learning on Manifold-Valued Data	311
<i>Alexander Kuleshov and Alexander Bernstein</i>	
A New Strategy for Case-Based Reasoning Retrieval Using Classification Based on Association	326
<i>Ahmed Aljuboori, Farid Meziane, and David Parsons</i>	
Tracking Communities over Time in Dynamic Social Network	341
<i>Etienne Gael Tajeuna, Mohamed Bouguessa, and Shengrui Wang</i>	

Generalized Hand Gesture Recognition for Wearable Devices in IoT: Application and Implementation Challenges	346
<i>Darshan Iyer, Fahim Mohammad, Yuan Guo, Ebrahim Al Safadi, Benjamin J. Smiley, Zhiqiang Liang, and Nilesh K. Jain</i>	
A Review on Artificial Intelligence Based Parameter Forecasting for Soil-Water Content.	356
<i>Ferhat Özçep, Eray Yıldırım, Okan Tezel, Metin Aşçı, Savaş Karabulut, and Tazegül Özçep</i>	
Dictionary Comparison for Anomaly Detection on Aircraft Engine Spectrograms	362
<i>Mina Abdel-Sayed, Daniel Duclos, Gilles Fay, Jérôme Lacaille, and Mathilde Mougeot</i>	
Location-Aware Ad Recommendation to Bid for Impressions	377
<i>Satish Kumar Verma</i>	
Building FP-Tree on the Fly: Single-Pass Frequent Itemset Mining	387
<i>Nima Shahbazi, Rohollah Soltani, Jarek Gryz, and Aijun An</i>	
Energy Disaggregation Based on Semi-Binary NMF	401
<i>Masako Matsumoto, Yu Fujimoto, and Yasuhiro Hayashi</i>	
A Probabilistic Matrix Factorization Method for Link Sign Prediction in Social Networks	415
<i>Qiang You, Ou Wu, Guan Luo, and Weiming Hu</i>	
Metadata-Based Clustered Multi-task Learning for Thread Mining in Web Communities	421
<i>Qiang You, Ou Wu, Guan Luo, and Weiming Hu</i>	
On Genetic Algorithms for Detecting Frequent Item Sets And Large Bite Sets.	435
<i>Roman A. Sizov and Dan A. Simovici</i>	
Multi-view Learning for Classification of X-Ray Crystallography Images. . . .	446
<i>B.M. Thamali Lekamge, Arcot Sowmya, and Janet Newman</i>	
C-KPCA: Custom Kernel PCA for Cancer Classification	459
<i>Van-Sang Ha and Ha-Nam Nguyen</i>	
LiveDoc: Showing Contextual Information Using Topic Modeling Techniques.	468
<i>Jayati Deshmukh, K.M. Annervaz, Shubhashis Sengupta, and Neetu Pathak</i>	

Genetic-Based Thresholds for Multi Histogram Equalization Image Enhancement	483
<i>Saeed Sedighi, Mehdi Roopaei, and Sos Agaian</i>	
A Clustering Approach for Discovering Intrinsic Clusters in Multivariate Geostatistical Data.	491
<i>Francky Fouedjio</i>	
On Learning and Exploiting Time Domain Traffic Patterns in Cellular Radio Access Networks	501
<i>Jordi Pérez-Romero, Juan Sánchez-González, Oriol Sallent, and Ramon Agusti</i>	
Pruning a Random Forest by Learning a Learning Algorithm	516
<i>Kumar Dheenadayalan, G. Srinivasaraghavan, and V.N. Muralidhara</i>	
Interestingness Hotspot Discovery in Spatial Datasets Using a Graph-Based Approach	530
<i>Fatih Akdag and Christoph F. Eick</i>	
MapReduce-Based Growing Neural Gas for Scalable Cluster Environments . . .	545
<i>Johannes Fliege and Wolfgang Benn</i>	
Comprehensible Enzyme Function Classification Using Reactive Motifs with Negative Patterns	560
<i>Thanapat Kangkachit and Kitsana Waiyamai</i>	
Enumerating Maximal Isolated Cliques Based on Vertex-Dependent Connection Lower Bound	569
<i>Yoshiaki Okubo, Makoto Haraguchi, and Etsuji Tomita</i>	
Simulator Prototyping Through Graphical Dependency Modeling	584
<i>Kumar Dheenadayalan, V.N. Muralidhara, and G. Srinivasaraghavan</i>	
Applying Data Mining to Healthcare: A Study of Social Network of Physicians and Patient Journeys	599
<i>Shruti Kaushik, Abhinav Choudhury, Kaustubh Mallik, Anzer Moid, and Varun Dutt</i>	
Feature Selection for Handling Concept Drift in the Data Stream Classification	614
<i>Pavel Turkov, Olga Krasotkina, Vadim Mottl, and Alexey Sychugov</i>	
Feature Reduction for Multi Label Classification of Discrete Data.	630
<i>V. Susheela Devi and Bhupesh Akhand</i>	

Accurate On-Road Vehicle Detection with Deep Fully Convolutional Networks	643
<i>Zequn Jie, Wen Feng Lu, and Eng Hock Francis Tay</i>	
On the Abstraction of a Categorical Clustering Algorithm	659
<i>Mina Sheikhalishahi, Mohamed Mejri, and Nadia Tawbi</i>	
Gossip-Based Behavioral Group Identification in Decentralized OSNs	676
<i>Naeimeh Laleh, Barbara Carminati, Elena Ferrari, and Sarunas Girdzijauskas</i>	
A Generalized Framework for Quantifying Trust of Social Media Text Documents	692
<i>Tuhin Sharma and Durga Toshniwal</i>	
A Comprehensive Analysis: Automated Ovarian Tissue Detection Using Type P63 Pathology Color Images	714
<i>T.M. Shahriar Sazzad, L.J. Armstrong, and A.K. Tripathy</i>	
Representation of 1-D Signals by a 0_1 Sequence and Similarity-Based Interpretation: A Case-Based Reasoning Approach	728
<i>Petra Perner</i>	
Rank Aggregation Algorithm Selection Meets Feature Selection	740
<i>Alexey Zabashta, Ivan Smetannikov, and Andrey Filchenkov</i>	
Attention Region Based Approach for Tracking Individuals in a Small School of Fish for Water Quality Monitoring	756
<i>Gang Xiao, Tengfei Shao, Tianqi Zhu, Yi Li, Jiafa Mao, and Zhenbo Cheng</i>	
FSMS: A Frequent Subgraph Mining Algorithm Using Mapping Sets	761
<i>Armita Abedijaberi and Jennifer Leopold</i>	
Can Software Project Maturity Be Accurately Predicted Using Internal Source Code Metrics?	774
<i>Mark Grechanik, Nitin Prabhu, Daniel Graham, Denys Poshyvanyk, and Mohak Shah</i>	
Multiple Consensuses Clustering by Iterative Merging/Splitting of Clustering Patterns	790
<i>Atheer Al-najdi, Nicolas Pasquier, and Frédéric Precioso</i>	
Author Index	805