

Petra Perner (Ed.)

Advances in Data Mining

Applications and Theoretical Aspects

19th Industrial Conference on Data Mining,
ICDM 2019

New York, USA, July 17 – July 21, 2019

Poster Proceedings

Volume Editor

Petra Perner
Institute of Computer Vision and Applied Computer Sciences, IBaI
PF 30 11 14
04251 Leipzig
E-mail: ppermer@ibai-institut.de

ISSN 1864-9734
ISBN 978-3-942952-61-3



The German National Library listed this publication in the German National Bibliography.
Detailed bibliographical data can be downloaded from <http://dnb.ddb.de>.

ibai-publishing
Prof. Dr. Petra Perner
PF 30 11 38
04251 Leipzig, Germany
E-mail: info@ibai-publishing.org
<http://www.ibai-publishing.org>

Copyright © 2019 ibai-publishing
ISSN 1864-9734
ISBN 978-3-942952-61-3

All rights reserved.
Printed in Germany, 2019

Preface

The nineteenth event of the Industrial Conference on Data Mining ICDM was held in New York (www.data-mining-forum.de) running under the umbrella of the World Congress on “The Frontiers in Intelligent Data and Signal Analysis, DSA 2019” (www.worldcongressdsa.com).

After the peer-review process, we accepted 39 high-quality papers for oral presentation, which are published in the ICDM Proceeding by ibai-publishing (www.ibai-publishing.org). The topics range from theoretical aspects of data mining to applications of data mining, such as in multimedia data, in marketing, in medicine, and in process control, industry, and society. Extended versions of selected papers will appear in the international journal *Transactions on Machine Learning and Data Mining* (www.ibai-publishing.org/journal/mldm).

In all, twenty one papers were selected for poster presentations, which are published in the ICDM Poster Proceeding by ibai-publishing (www.ibai-publishing.org).

A tutorial on Data Mining and a tutorial on Case-Based Reasoning were held after the conference.

The conference was running in an inspiring atmosphere. The presenters of the oral presentations gave excellent talks and the audience acknowledged each talk with excellent questions. The poster presenters presented well prepared posters and gave in front of their posters a summary of their work in five minutes for the audiences. Afterwards the audience could step to the poster that they were interested in. The audience had a lot of questions and they gave valuable comments and new directions for the work in progress. In all was the poster session a very successful session.

We would like to thank all reviewers for their highly professional work and their effort in reviewing the papers.

We also thank the members of the 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 ibai-publishing publishing house (www.ibai-publishing.org) that handled the papers and published the proceedings.

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 2020 in New York at the next World Congress on “The Frontiers in Intelligent Data and Signal Analysis, DSA 2020” (www.worldcongressdsa.com), which combines under its roof the following three events: International Conferences Machine Learning and Data Mining MLDM (www.mldm.de), the Industrial Conference on Data Mining ICDM (www.data-mining-forum.de), and the International Conference on Mass Data Analysis of Signals and Images in Artificial Intelligence and Pattern Recognition with Applications in Medicine, Biotechnology, Chemistry and Food Industry, MDA-AI&PR (www.mda-signals.de).

July 2019

Petra Pernert

19th Industrial Conference on Data Mining ICDM 2019

www.data-mining-forum.de
July 17 – 21, 2019, New York, USA

Chair

Prof. Dr. Petra Perner

Institute of Computer Vision and applied Computer Sciences, IBAI

Program Committee

Ajith Abraham	Machine Intelligence Research Labs (MIR Labs), USA
Mohamed, Bourguessa	Universite du Quebec a Montreal - UQAM, Canada
Bernard Chen	University of Central Arkansas, USA
Jeroen de Bruin	University of Applied Sciences JOANNEUM, Austria
Antonio Dourado	University of Coimbra, Portugal
Stefano Ferilli	University of Bari, Italy
Geert Gins	Glaxo Smith Kline, Belgium
Warwick Graco	Australian Tax Office ATO, Australia
Aleksandra Gruca	Silesian University of Technology, Poland
Pedro Isaias	The University of Queensland, Australia
Piotr Jedrzejowicz	Gdynia Maritime University, Poland
Martti Juhola	University of Tampere, Finland
Janusz Kacprzyk	Polish Academy of Sciences, Poland
Mehmed Kantardzic	University of Louisville, USA
Lui Xiaobing	Google Inc., USA
Eduardo F. Morales	National Institute of Astrophysics, Optics, and Electronics, Mexico
Samuel Noriega	Universitat de Barcelona, Spain
Wieslaw Paja	University of Rzeszow, Poland
Juliane Perner	Novartis Institutes for BioMedical Research (NIBR), Switzerland
Rainer Schmidt	University of Rostock, Germany
Moti Schneider	PCCW Global, Greece
Victor Sheng	University of Central Arkansas, USA
Kaoru Shimada	Fukuoka Dental College, Japan
Gero Szepannek	University of Applied Sciences Stralsund, Germany
Joao Miguel Costa Sousa	Technical University of Lisbon, Portugal
Markus Vattulainen	Tampere University, Finland
Zhu Bing	Sichuan University, China

Table of Content

The LuNa Open Toolbox for the Luxembourgish Language <i>Joshgun Sirajzade and Christoph Schommer</i>	1
Text Classification Using Combinations of Different Versions of Input Files <i>Yaakov Hacoheh-Kerner, Adir Prager and Adiel Cohen</i>	16
Use of Big Data for Disaster Management <i>Gijoo Yang, Kangwoo Lee and Dongho Kim</i>	31
Does Ambient Intelligence for Mobile User Assistance dissolve in Privacy Oriented Policies? <i>Denys Proux and Frederic Roulland</i>	37
Machine Learning-Assisted Prediction of Surgical Mortality of Lung Cancer Patients <i>Sidra Xu</i>	46
The Impact of Preprocessing on the Classification of Mental Disorders <i>Yaakov Hacoheh-Kerner, Yair Yigal and Daniel Miller</i>	52
Suggestion of Density-based Geometric One-Class Classifier using Genetic Algorithm <i>Do Gyun Kim and Jin Young Choi</i>	67
Comprehensibility for Decision Intelligence Engineering <i>Inna Kolyshkina and Simeon Simoff</i>	80
Recognition of the Experimental Design of Laboratory Scale Wine Fermentations Using ANN <i>Gonzalo Hernández, Alejandra Urtubia and Roberto León</i>	95
Data Mining for the Internet of Things <i>Tausifa Jan Saleem and Mohammad Ahsan Chishti</i>	103
Large-Scale Vertex-Centric Network Embedding via Apache Spark <i>Sara Riazi and Boyana Norris</i>	111
Leveraging Multiple Online Sources for Accurate Income Verification <i>Chirag Mahapatra and Kedar Bellare</i>	123

Methods for Solving the Challenges Observed in the Multi-platform Setup of the Self driving Cars <i>Mirza Mujtaba Baig</i>	137
Text Classification and Sentiment Analysis in Social Media for the Marketing Domain <i>Yaakov Hacohen-Kerner</i>	150
Feature Impact for Prediction Explanation <i>Mohammad Bataineh, David Steenhard and Harpreet Singh</i>	160
Data-driven American Option Pricing using Artificial Neural Networks <i>Hanen Borchani, Wojciech Michal Pawlak, Steffen Holmslykke and Allan Peter Engsig-Karup</i>	168
CLUES: Detection of User Intent through Interactive Comparison <i>Chitra Phadke, Paishun Ting, Jin Cao and Huseyin Uzunalioglu</i>	173
A Common Gene Expression Signature Analysis Method for Multiple Types of Cancer <i>Yingcheng Sun, Xiangru Liang and Kenneth Loparo</i>	185
A Deep Learning Framework for Classification of in vitro Multi-Electrode Array Recordings <i>Yun Zhao, Elmer Guzman, Morgane Audouard, Zhuowei Cheng, Paul K. Hansma, Kenneth S. Kosik and Linda Petzold</i>	197
Knowledge-guided Text Structuring in Clinical Trials <i>Yingcheng Sun and Kenneth Loparo</i>	211
Authors Index	220