

Petra Perner (Ed.)

Advances in Data Mining

Applications and Theoretical Aspects

20th Industrial Conference, ICDM 2020
Amsterdam, The Netherlands
July 20 – July 21 2020
Proceedings

Volume Editor

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

P-ISSN 1864-9734
E-ISSN 2699-5220
ISBN 978-3-942952-76-7

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 © 2020 ibai-publishing
P-ISSN 1864-9734
E-ISSN 2699-5220
ISBN 978-3-942952-76-7

All rights reserved. Printed in Germany, 2020

20th Industrial Conference on Data Mining ICDM 2020

www.data-mining-forum.de

July 20 – 21, 2020

Amsterdam, The Netherlands

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

Preface

The pandemic "Corona" has put us this year before a difficult time. With care we have kept to the hygiene rules not to get an infection with the virus Covid-19. With mask we have got into coaches and trains, have made our purchases or on work worked. Home office was the catchword of these days. The universities and research facilities have maintained only a small emergency company and lectures were held as online lectures. From home we have tried to do our scientific works. In 1-to-1 telephone calls or phone conferences we have organized with our colleagues the work and have discussed important results of the research. Under it the efficiency suffers what is easy to understand.

In the beginning of the pandemic fell the deadline of our conference. Insecurity spread. The figures of the infected persons increased rapidly. The virus spreads out in more and more countries and was further carried by continent to continent. Soon stood the whole world in the spell of Corona. A conference was the last to this in this situation most thought.

In this situation appeared once again which high demands for a scientist are made. It belongs to the job of a scientist that he presents his scientific results in conferences and makes thus his results of a wide public immediately available. A scientist should have well organized his research, should be able to do his scientific tasks and duties in a flexible way, and should have financed his research with suitable financial means. Only those who were meeting these rules could successfully continue in their professional research work.

The best of the best of us are represented with their papers in this volume. They presented themselves personal or in online presentations in the conference. The acceptance rate for the submitted paper of our conference was 33% percent for long paper as well as short papers. Because of many refusals because of missing financial means or other reasons the acceptance rate decreased to few percent. This shows once more the excellent quality of these scientists. Their papers are of most excellent quality and expand the state-of-the-art in an excellent way. The topics of the long papers range from event log file analysis, predictive maintenance, medical application, telecom application, fraud detection to a paper on how we should present the results to stakeholders so that they accept the findings of the data mining methods. The new arising topic we see here is predictive maintenance. All other topics follow the main topics of ICDM but present new excellent results and go over the recent questions to be solved with data mining for the specific applications. The short paper is a fine theoretical paper on optimal kernel density estimation.

The proceedings will be freely accessible as an OPEN-ACCESS Proceedings of a wide public so that, the new acquired knowledge on the different subjects is able to spread around quickly worldwide. You can find the proceedings for long papers and the poster proceedings for short papers at <http://www.ibai-publishing.org/html/proceeding2020.php>.

In this time, flexibility was a must Because the situation in the USA was still difficult, we have moved the conference to Amsterdam in the Netherlands. Here a variety

of the participants was able to do outward journeys. The ones who could not travel, were online present.

Extended versions of selected papers will appear in the international journal Transactions on Machine Learning and Data Mining (www.ibai-publishing.org/journal/mldm).

We hope to see you in 2021 in New York at the 21th Industrial Conference on Data Mining ICDM (www.data-mining-forum.de) again.

The conference runs under the umbrella of the World Congress on “The Frontiers in Intelligent Data and Signal Analysis, DSA 2021” (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).

We will give then the tutorials on Data Mining, Case-Based Reasoning, and Intelligent Image Analysis again (<http://www.data-mining-forum.de/tutorials.php>) again. The workshops running in connection with ICDM will also be given (<http://www.data-mining-forum.de/workshops.php>).

We would warmly invite you with pleasure to contribute to this conference. Please come and join us. We are awaiting you.

July, 2020

Petra Pernert

Table of Content

Unified Expression Ripple Down Rules based Fraud Detection Technique for Scalable Data <i>Ikram Ul Haq, Iqbal Gondal and Peter Vamplew</i>	1
Root causes labelling of industrial assets via relevancy estimation of event logs <i>Pierre Dagnely, Tom Tourwe and Elena Tsiporkova</i>	17
Application of Machine Learning to Predictive Maintenance <i>Feranmi Akanni and Martti Juhola</i>	33
What Stakeholders Expect from Analytics? <i>Warwick Graco</i>	49
Interpreting influence of feature ranking in derivation of prediction models for screening questionnaires optimization <i>Leona Cilar, Majda Pajnkihar and Gregor Stiglic</i>	67
Predicting inactivity users in telecom <i>Cong Dan Pham, Phi Hung Nguyen, Xuan Vinh Chu, Van Hung Trinh, Duc Hai Nguyen</i>	79
Inference via Conditional Kolmogorov Complexity <i>Daniel Goldfarb and Scott Evans Causal</i>	91
Survival Analysis of Breast Cancer Utilizing Integrated Features with Ordinal Cox Model and Auxiliary Loss <i>Isabelle Bichindaritz, Gunaghui Liu and Christopher Bartlett</i>	105