

Petra Perner (Ed.)

Machine Learning and Data Mining in Pattern Recognition

16th International Conference on Machine
Learning and Data Mining, MLDM 2020,
Amsterdam, The Netherlands, July 20-21, 2020
Proceedings

ibai Publishing

www.ibai-publishing.org

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

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

P-ISSN 1864-9734
E-ISSN 2699-5220
ISBN 978-3-942952-75-0

All rights reserved.
Printed in Germany, 2019

16th International Conference on Machine Learning and Data Mining, MLDM 2020

July 20-21, 2020, Amsterdam, The Netherlands
www.mldm.de

Chair

Petra Perner
Institute of Computer Vision and Applied Computer Sciences, IBAI
Germany

Program Committee

Reneta Barneva	The State University of New York at Fredonia, USA
Michelangelo Ceci	Universtiy of Bari, Italy
Ireneusz Czarnowski	Gdynia Maritime University, Poland
Roberto Corrizo	Universtiy of Bari, Italy
Christoph F. Eick	Universtiy of Houston, USA
Mark J. Embrechts	Rensselaer Polytechnic Institute and CardioMag Imaging, Inc, USA
Ana Fred	Technical University of Lisboa, Portugal
Giorgio Giacinto	University of Cagliari, Italy
Aminata Kane	Concordia University, Canada
Piet Kommers	University of Twente, The Netherlands
Olga Krasotkina	Russian State University, Russia
Dimitris Karras	Chalkis Institute of Technology, Greece
Adam Krzyzak	Concordia University, Canada
Valerio Pascucci	University of Utah, USA
Gianvito Pio	University of Bari, Italy
Francis E.H. Tay	National University of Singapore, Singapore
Turki Turki	King Abdulaziz University, Saudi Arabia
Zeev Volkovich	ORT Braude College of Engineering, Israel
Patrick Wang	Northeastern University, USA

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 tele-phone 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. This proceeding volume present mostly theoretical work on known topics such as clustering, classification and prediction, graph mining but also application-oriented theoretical work for different purposes based deep learning and others. One invited talk was given by Prof. Dan A. Simovici on a theoretical subject on "Information theoretic approaches in data mining". His paper is also included in the proceedings.

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 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 invite you to join us in 2021 in New York again to the next International Conference on Machine Learning and Data Mining MLDM. The conference will run again under the umbrella of the Worldcongress (www.worldcongressdsa.com) “The Frontiers in Intelligent Data and Signal Analysis, DSA2020” that combines under his roof the following three events: International Conferences 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 Artificial Intelligence and Pattern Recognition with Application in with Applications in Medicine, r/g/b Biotechnology, Food Industries and Dietetics, Biometry and Security, Agriculture, Drug Discover, and System Biology MDA-AI&PR.

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

Invited Talk

Information-theoretical Approaches in Data Mining <i>Dan A. Simovici</i>	1
---	---

Regular Papers

Identification of Urban Functional Areas Based on POI Cluster Analysis <i>Huifang Feng, Chong Jia, Zhenjuan Yang, Youji Xu</i>	25
Measuring the effectiveness of code review comments in GitHub repositories: A machine learning approach <i>Shadikur Rahman, Umme Ayman Koana, Hasibul Karim Shanto, Mahmuda Akter, Chitra Roy, Aras M. Ismael</i>	35
The complex role of location and time in prediction models <i>Mahdi Hashemi</i>	49
Detecting Periodic Correlated Hashtags in Twitter <i>Sultan Alzahrani, Saud Alashri, Manal Alhassoun, Sara Alghunaim, Sarah Alasraj, Muneera Alhoshan</i>	63
A Real Time Accident Detection Framework for Traffic Video Analysis <i>Hadi Ghahremannezhad, Hang Shi, Chengjun Liu</i>	77
Similarity Learning by Random Projection Forests with Application to Clustering- <i>Donghui Yan, Songxiang Gu, Ying Xu, Zhiwei Qin</i>	93
Automated Topic Modelling for Unlabelled Network Data <i>Sadaf Azad, Ci Lei, David Farrar, Thomas Mangin</i>	109
Laplacian Centrality of Sets of Vertices of Graphs <i>Cristina Maier, Dan Simovici</i>	123
New Machine Learning Approaches to Improve Software Bug Prediction <i>Abdullah Algarni, Emad Kaen, Turki Turki</i>	139
How is Your Team Spirit? Cluster Over-Time Stability Evaluation <i>Martha Tatusch, Gerhard Klassen, Marcus Bravidor, Stefan Conrad</i>	155
Coarse- and fine-scale geometric information content of Multiclass Classification and implied Data-driven Intelligence <i>Fushing Hsieh, Xiaodong Wang</i>	171

A New Ensemble Method for Convolutional Neural Networks and its Application to Plant Disease Detection <i>Vaibhav Katturu, Parampuneet Kaur Thind, Sung-Hyuk Cha, Teryn Cha</i>	187
Multiple Imputation with Denoising Autoencoder using Metamorphic Truth and Imputation Feedback <i>Haw-Minn Lu, Giancarlo Perrone, Jose Unpingco</i>	197
Deep 3D Face Recognition using 3D Data Augmentation and Transfer Learning- <i>Lyndon Smith, Ning Huang, Mark Hansen, Melvyn Smith</i>	209
Hybrid Gravitational Clustering with Centroids <i>Nathaniel Zeiger, Raed Seetan</i>	219
Deep Learning Based Android Malware Detection Framework <i>Devashish Khulbe, Soumya Sourar, Naman Kapoor</i>	231
Predicting resource utilization in Cloud using Multivariate CNN LSTM model <i>Soukaina Ouham, Youssef Hadi</i>	239
Author's Index	253