

Petra Perner (Ed.)

# **Machine Learning and Data Mining in Pattern Recognition**

15th International Conference on Machine  
Learning and Data Mining, MLDM 2019,  
New York, NY, USA, July 20-25, 2019  
Proceedings  
Volume I

**ibai** Publishing

---

[www.ibai-publishing.org](http://www.ibai-publishing.org)

Volume Editor

Petra Perner  
Institute of Computer Vision and Applied Computer Sciences, IBaI  
PF 30 11 14  
04251 Leipzig  
E-mail: [pperner@ibai-institut.de](mailto:pperner@ibai-institut.de)

ISSN 1864-9734  
ISBN 978-3-942952-62-0



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](mailto:info@ibai-publishing.org)  
<http://www.ibai-publishing.org>

**Copyright** © 2019 ibai-publishing  
ISSN 1864-9734  
ISBN 978-3-942952-62-0

All rights reserved.  
Printed in Germany, 2019

# 15th International Conference on Machine Learning and Data Mining, MLDM 2019

July 20-24, 2019, New York, USA  
[www.mldm.de](http://www.mldm.de)

## Chair

Petra Perner  
Institute of Computer Vision and Applied Computer Sciences, IBA  
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 fifteenth event of the International Conference on Machine Learning and Data Mining MLDM was held in New York ([www.mldm.de](http://www.mldm.de)) running under the umbrella of the Worldcongress “The Frontiers in Intelligent Data and Signal Analysis, DSA2019” ([www.worldcongressdsa.com](http://www.worldcongressdsa.com)).

For this edition the Program Committee received 245 submissions. After the peer-review process, we accepted 65 high-quality papers for oral presentation. The topics range from theoretical topics for classification, clustering, 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](http://www.ibai-publishing.org/journal/mldm)).

A tutorial on Data Mining and a tutorial on Case-Based Reasoning were held before the conference that took pleasure to high participation of researchers and practitioners from industry, social and public services.

We like to thank all presenter for your high-quality presentations and the audience for your high-professional questions and inspiring comments. All that has made the conference to a living and dreadful event. The presenters and the audience went home with a full bag of new insights into different topics the research and inspiring ideas for new work and research. Besides that, gave the banquet an excellent opportunity to network among the participants and set up new co-operations.

We like to thank all reviewers for their highly professional work and their effort in reviewing the papers. We also thank members of Institute of Applied Computer Sciences, Leipzig, Germany ([www.ibai-institut.de](http://www.ibai-institut.de)) who handed the conference as secretariat. We appreciate the help and understanding of the editorial staff of ibai-publishing house ([www.ibai-publishing.org](http://www.ibai-publishing.org)) that prepared and published the proceeding books in two volumes.

We invite you to join us in 2020 in New York to the next Worldcongress ([www.worldcongressdsa.com](http://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 Diagnostics, Biometry and Security, Agriculture, Drug Discover, and System Biology MDA-AI&PR.

## Table of Content

Arabic Rule-based Named Entity Recognition System Using GATE <i>Hatem M. ElSherif, Khaled Mohammad Alomari, Ahmad Qasim AlHamad and Khaled Shaalan</i> .....	1
An Unsupervised Statistical Moving Shadow Detection Method for Video Analysis <i>Hang Shi and Chengjun Liu</i> .....	16
Social Activism Analysis: An Application of Machine Learning in the World Values Survey <i>Francielle M. Nascimento, Dante A. C. Barone, and Henrique Carlos de Castro</i> .....	28
Deep Dilated Convolutional Nets for the Automatic Segmentation of Retinal Vessels <i>Ali Hatamizadeh, Hamid Hosseini, Zhengyuan Liu, Steven D. Schwartz, and Demetri Terzopoulos</i> .....	39
Clonal Selection Algorithm Applied to Object Recognition in Mobile Robots <i>Jose Guillermo Guarnizo and Luis Fernando Nino</i> .....	49
Customized use of Expectation Maximization for Glioma Identification from Brain MRIs <i>Nidhi Gupta, Pushpraj Bhatele, and Pritee Khanna</i> .....	63
A Novel Phonetic Algorithm for Predicting Chinese Names using Chinese PinYin <i>Hua Zhao and Fairouz Kamareddine</i> .....	78
Detecting Customer Churn Signals for Telecommunication Industry Through Analyzing Phone Call Transcripts with Recurrent Neural Networks <i>Junmei Zhong, William Li</i> .....	93
A comparison of linear and machine learning models for the simulation of soil moisture <i>Milan Cisty, Frantisek Cyprich and Veronika Soldanova</i> .....	104
Vehicle Classification in Video Using Deep Learning <i>Mohammad O. Faruque, Hadi Ghahremanzhad, and Chengjun Liu</i> .....	117
A Model for Diagnosis of Alzheimer’s Disease Using Haralick Texture Features and Meta Feature Selection <i>Iago Richard Rodrigues Silva, Marilia Nayara Clemente de Almeida Lima, Wellington Pinheiro dos Santos, and Roberta Andrade de Araújo Fagundes</i> .....	132

Mining Data Stream to Detect Behavior Change in a Real-Time Strategy Game <i>Eldane Vieira, Rita Maria Silva Julia, and Elaine Ribeiro de Faria</i> .....	146
Packet2Vec: Utilizing Word2Vec for Feature Extraction in Packet Data <i>Eric L. Goodman, Chase Zimmerman, and Corey Hudson</i> .....	161
Handbags Classification Model via Deep Learning <i>Dieinison J. F. Braga, Leodecio Braz S.S., Criston Pereira de Souza1, and Ticiano L. Coelho da Silva</i> .....	176
Long Short-Term Memory-based Multi-Period Price Prediction for Portfolio Man- agement <i>Zhengyong Jiang and Frans Coenen</i> .....	187
Fraud Detection Using Explainable Machine Learning Algorithms <i>Luciano C M Andrade, Andr'e C P L F Carvalho</i> .....	201
A Fast AEMST Algorithm for High-Dimensional Datasets by Removing Redundant Distance Computations <i>Xiaochun Wang, Xia Li Wang, Xuan Xiong Lin</i> .....	216
Server Failure Prediction Framework for IT Infrastructure Supporting Critical Appli- cations <i>Mythili Krishnan and Madhan Kumar Srinivasan</i> .....	229
A New Minimum Spanning Tree Algorithm Constructed in A Backward Fashion <i>Aozhong Wang, Xiaochun Wang, Xia Li Wang</i> .....	243
Analysis of Linear and Non-Linear Classiers in Imbalanced Data to Predict Diabetes Induced Complications <i>Tahsinur Rahman, Aniq Zaida Khanom, Sheikh Mastura Farzana, Sharowar Md. Shahriar Khan and Dr. Md. Ashrafal Alam</i> .....	255
FF-SVM: New FireFly-based Gene Selection Algorithm for Microarray Cancer Clas- sification <i>Nada Almugren and Hala Alshamlan</i> .....	270
Fast K-medoids Clustering Algorithm Using Triangle Inequality <i>Xiaochun Wang, Xia Li Wang, Xuan Xiong Lin</i> .....	282
Design and Implementation of an Intrusion Detection System using Deep Neural Network <i>Hyun-chul Chang and Sungbum Park</i> .....	296

VII

Community detection in networks with fuzzy boundaries <i>Pradumn Kumar Pandey, T. Ramalingeswara Rao and Satyajeet Maharana ...</i>	304
Precise Feature Selection and Case Study of Intrusion Detection in an Industrial Control System (ICS) Environment? <i>Terry Guo, Animesh Dahal, and Ambareen Siraj .....</i>	319
Entropy-based Approach for Parameter-free Attribute Clustering <i>Adison Khomprasert, Thanawin Rakthamanon, Kitsana Waiyamai .....</i>	333
Seq2SQL - Evaluating Different Deep Learning Architectures Using Word Embeddings <i>Kevin Stower and Dirk Krechel .....</i>	343
Sparse Multimodal Classification of EEG Signals from Rapid Serial Visual Presentation of Diagnostic Images <i>Valentina Sulimova, Sergey Bukhonov, Olga Krasotkina, Vadim Mottl, Annette Sterr, Kevin Wells, David Windridge .....</i>	355
Neural-Attention Multi-Instance Learning for Predicting User Demographics from Highly Noisy Tweets <i>Qing Chen, Mingxuan Sun, and Jian Zhang .....</i>	367
Explainable Artificial Intelligence for Match Analysis in Association Football <i>Bruno Marques and Dante Barone .....</i>	382
Using Machine Learning Algorithms to Predict the Likelihood of Recurrent Falls in Older Adults <i>Leeanne Lindsay, Sonya Coleman, Brian Taylor, Dermot Kerr, Anne Moorhead .....</i>	392
Tree Based Clustering On Large, High Dimensional Datasets? <i>Lee A. Carraher, Sayantan Dey, and Philip A. Wilsey .....</i>	397
Linear complexity algorithms for high dimensional SVM and regression problems with smart sparse regularization <i>Vadim Mottl, Olga Krasotkina, Valentina Sulimova, Alexey Morozov, Ilya Pugach, Alexander Tatarchuk .....</i>	412
Authors Index .....	431