Communications in Computer and Information Science 1167

Commenced Publication in 2007
Founding and Former Series Editors:
Phoebe Chen, Alfredo Cuzzocrea, Xiaoyong Du, Orhun Kara, Ting Liu, Krishna M. Sivalingam, Dominik Ślęzak, Takashi Washio, Xiaokang Yang, and Junsong Yuan

Editorial Board Members

Simone Diniz Junqueira Barbosa
Pontifical Catholic University of Rio de Janeiro (PUC-Rio), Rio de Janeiro, Brazil

Joaquim Filipe
Polytechnic Institute of Setúbal, Setúbal, Portugal

Ashish Ghosh
Indian Statistical Institute, Kolkata, India

Igor Kotenko
St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences, St. Petersburg, Russia

Lizhu Zhou
Tsinghua University, Beijing, China

More information about this series at http://www.springer.com/series/7899

Peggy Cellier · Kurt Driessens (Eds.)

Machine Learning and Knowledge Discovery in Databases

International Workshops of ECML PKDD 2019
Würzburg, Germany, September 16–20, 2019
Proceedings, Part I

Springer

Editors
Peggy Cellier
Institut National des Sciences Appliquées
Rennes, France

Kurt Driessens
Maastricht University
Maastricht, The Netherlands

ISSN 1865-0929 ISSN 1865-0937 (electronic)
Communications in Computer and Information Science
ISBN 978-3-030-43822-7 ISBN 978-3-030-43823-4 (eBook)
https://doi.org/10.1007/978-3-030-43823-4

© Springer Nature Switzerland AG 2020
The chapter "Supervised Human-Guided Data Exploration" is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/). For further details see license information in the chapter.
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, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD) is the premier European machine learning and data mining conference. In 2019, ECML PKDD was held in Würzburg, Germany, during September 16–20.

During the first and last day of the conference, the workshop program allowed a number of specialized and/or new topics to take the fore-front.

A record 46 workshop and tutorial topics were submitted to the 2019 conference. The selection and merging process resulted in 25 workshops taking place over the two days, of which 3 were combined with a tutorial.

The workshop program included the following workshops:

1. The 12th International Workshop on Machine Learning and Music (MML 2019)
2. Workshop on Multiple-aspect analysis of semantic trajectories (MASTER 2019)
3. The 4th Workshop on MIning DAta for financial applicationS (MIDAS 2019)
4. The Second International Workshop on Knowledge Discovery and User Modelling for Smart Cities (UMCit 2019)
5. New Frontiers in Mining Complex Patterns (NFMCP 2019)
6. New Trends in Representation Learning with Knowledge Graphs
7. The Second International Workshop on Energy Efficient Scalable Data Mining and Machine Learning (Green Data Mining)
8. Workshop on Deep Continuous-Discrete Machine Learning (DeCoDeML 2019)
9. Decentralised Machine Learning at the Edge (DMLE 2019)
10. Applications of Topological Data Analysis (ATDA 2019)
11. GEM: Graph Embedding and Mining
12. Interactive Adaptive Learning (AIL 2019)
13. IoT Stream for Data Driven Predictive Maintenance (IoT Steam 2019)
14. Machine Learning for Cybersecurity (MLCS 2019)
15. BioASQ: Large-scale biomedical semantic indexing and question answering
16. The 6th Workshop on Sports Analytics: Machine Learning and Data Mining for Sports Analytics (MLSA 2019)
17. The 4th Workshop on Advanced Analytics and Learning on Temporal Data (AALTD 2019)
18. MACLEAN: MAChine Learning for EArth ObservatioN
19. Automating Data Science
20. The 4th Workshop on Data Science for Social Good (DSSG 2019)
21. The Third Workshop on Advances in managing and mining Large Evolving Graphs (LEG 2019)
22. Data and Machine Learning Advances with Multiple Views (DAMVL 2019)
23. Workshop on Data Integration and Applications (DINA 2019)
24. XKDD Tutorial and XKDD-AIMLAI Workshop
25. The First Workshop SocIaL Media And Harassment (SIMAH 2019)

Of these 25 workshops, 17 workshops decided to select and publish their best papers with Springer. Two workshops were large enough to publish their own proceedings: (i) MIDAS – the 4th Workshop on MIning DAta for financial applicationS and (ii) AALTD – the 4th workshop on Advanced Analytics and Learning on Temporal Data. The 15 other workshops received a total of 200 submitted papers, out of which 70 long and 46 short papers were selected for publication after the conference. These papers are spread over two proceedings volumes.

This two-volume set contains the papers from the following workshops:

1. Automating Data Science
2. XKDD Tutorial and XKDD-AIMLAI Workshop
3. Decentralised Machine Learning at the Edge (DMLE 2019)
4. The Third Workshop on Advances in managing and mining Large Evolving Graphs (LEG 2019)
5. Data and Machine Learning Advances with Multiple Views (DAMVL 2019)
6. New Trends in Representation Learning with Knowledge Graphs
7. The 4th Workshop on Data Science for Social Good (DSSG 2019)
8. The Second International Workshop on Knowledge Discovery and User Modelling for Smart Cities (UMCit 2019)
9. Workshop on Data Integration and Applications (DINA 2019)
10. Machine Learning for Cybersecurity (MLCS 2019)
11. The 6th Workshop on Sports Analytics: Machine Learning and Data Mining for Sports Analytics (MLSA 2019)
12. The First Workshop on SocIaL Media And Harassment (SIMAH 2019)
13. IoT Stream for Data Driven Predictive Maintenance (IoT Stream 2019)
14. The 12th International Workshop on Machine Learning and Music (MML 2019)
15. BioASQ: Large-scale biomedical semantic indexing and question answering

We would like to thank all participants and invited speakers, the workshop organizers and the reviewers, as well as the local organizers for making the workshop program of ECML PKDD 2019 the success that it was. Sincere thanks also goes to Springer for their help in publishing the proceedings.

January 2020

Peggy Cellier
Kurt Driessens

Organization

ECML Workshop Chairs/Editors

Peggy Celier	INSA Rennes, France
Kurt Driessens	Maastricht University, The Netherlands

Individual Workshop Chairs/Editors

Tijl De Bie	UGent, Belgium
Luc De Raedt	KU Leuven, Belgium
Jose Hernandez-Orallo	Universitat Politecnica de Valencia, Spain
Adrien Bibal	University of Namur, Belgium
Tassadit Bouadi	University of Rennes/IRISA, France
Benoît Frénay	University of Namur, Belgium
Luis Galárraga	Inria/IRISA, France
Stefan Kramer	Universität Mainz, Germany
Ruggero G. Pensa	University of Turin, Italy
Michael Kamp	University of Bonn, Germany
Yamuna Krishnamurthy	Rolay Holloway University of London, England
Daniel Paurat	Fraunhofer IAIS, Germany
Sabeur Aridhi	University of Lorraine, France
José Antonio de Macedo	Universidade Federal do Ceará, Brazil
Engelbert Mephu Nguifo	University Clermont Auvergne, France
Karine Zeitouni	Université de Versailles Saint-Quentin, France
Stéphane Ayache	Aix-Marseille University, France
Cécile Capponi	Aix-Marseille University, France
Rémi Emonet	Jean-Monnet University, France
Usabelle Guyon	Orsay University, France
Volker Tresp	Ludwig-Maximilians University and Siemens, Germany
Jens Lehmann	Bonn University and Fraunhofer IAIS, Germany
Aditya Mogadala	Saarland University, Germany
Achim Rettinger	Trier University, Germany
Afshin Sadeghi	Fraunhofer IAIS, Germany
Mehdi Ali	Bonn University and Fraunhofer IAIS, Germany
Ricard Gavalda	UPC BarcelonaTech, Spain
Irena Koprinska	University of Sydney, Australia
Joao Gama	University of Porto, Portugal
Rabeah Alzaidy	King Abdullah University of Science and Technology, Saudi Arabia
Marcelo G. Armentano	ISISTAN, CONICET-UNICEN, Argentina
Antonela Tommasel	ISISTAN, CONICET-UNICEN, Argentina

Ludovico Boratto	Eurecat, Barcelona, Spain
Clyde L. Giles	College of Information Sciences and Technology, Pennsylvania State University
Luiza Antonie	University of Guelph, Canada
Peter Christen	The Australian National University, Australia
Erhard Rahm	University of Leipzig, Germany
Osmar Zaïane	University of Alberta, Canada
Annalisa Appice	Università degli Studi di Bari, Italy
Battista Biggio	Università degli Studi di Cagliari, Italy
Donato Malerba	Università degli Studi di Bari, Italy
Fabio Roli	Università degli Studi di Cagliari, Italy
Ibéria Medeiros	Universidade de Lisboa, Portugal
Pedro Ferreira	Universidade de Lisboa, Portugal
Michael Kamp	University of Bonn, Germany
Jesse Davis	KU Leuven, Belgium
Ulf Brefeld	Leuphana University, Germany
Jan Van Haaren	SciSports, The Netherlands
Albrecht Zimmermann	University of Caen, France
Rita Ribeiro	University of Porto, Porto
Albert Bifet	Telecom-ParisTech, France
João Gama	University of Porto, Porto, Portugal
Anders Holst	RISE SICS, Sweden
Sepideh Pashami	Halmstad University, Sweden
Sima Sharifirad	Dalhousie University, Canada
Stan Matwin	Dalhousie University, Canada
Rafael Ramirez	Universitat Pompeu Fabra, Spain
Darrell Conklin	Universidad del País Vasco, Spain
José Manuel Iñesta	Alicante University, Spain
George Paliouras	University of Houston, USA
Anastasia Krithara	NCSR “Demokritos”, Greece
Anastasios Nentidis	Aristotle University of Thessaloniki, Greece

Contents – Part I

Automating Data Science

The ABC of Data: A Classifying Framework for Data Readiness 3
Laurens A. Castelijns, Yuri Maas, and Joaquin Vanschoren

Automating Common Data Science Matrix Transformations 17
Lidia Contreras-Ochando, Cèsar Ferri, and José Hernández-Orallo

DeepNotebooks: Deep Probabilistic Models Construct Python Notebooks for Reporting Datasets 28
Claas Völcker, Alejandro Molina, Johannes Neumann, Dirk Westermann, and Kristian Kersting

HyperUCB: Hyperparameter Optimization Using Contextual Bandits 44
Maryam Tavakol, Sebastian Mair, and Katharina Morik

Learning Parsers for Technical Drawings 51
Dries Van Daele, Nicholas Decleyre, Herman Dubois, and Wannes Meert

Meta-learning of Textual Representations 57
Jorge G. Madrid, Hugo Jair Escalante, and Eduardo Morales

ReinBo: Machine Learning Pipeline Conditional Hierarchy Search and Configuration with Bayesian Optimization Embedded Reinforcement Learning 68
Xudong Sun, Jiali Lin, and Bernd Bischl

Supervised Human-Guided Data Exploration 85
Emilia Oikarinen, Kai Puolamäki, Samaneh Khoshrou, and Mykola Pechenizkiy

SynthLog: A Language for Synthesising Inductive Data Models (Extended Abstract) 102
Yann Dauxais, Clément Gautrais, Anton Dries, Arcchit Jain, Samuel Kolb, Mohit Kumar, Stefano Teso, Elia Van Wolputte, Gust Verbruggen, and Luc De Raedt

The `autofeat` Python Library for Automated Feature Engineering and Selection 111
Franziska Horn, Robert Pack, and Michael Rieger

The Extended Dawid-Skene Model: Fusing Information from Multiple Data Schemas 121
Michael P. J. Camilleri and Christopher K. I. Williams

Towards Automated Configuration of Stream Clustering Algorithms 137
Matthias Carnein, Heike Trautmann, Albert Bifet, and Bernhard Pfahringer

Advances in Interpretable Machine Learning and Artificial Intelligence & eXplainable Knowledge Discovery in Data Mining (AIMLAI-XKDD)

Effect of Superpixel Aggregation on Explanations in LIME – A Case Study with Biological Data 147
Ludwig Schallner, Johannes Rabold, Oliver Scholz, and Ute Schmid

Global Explanations with Local Scoring. 159
Mattia Setzu, Riccardo Guidotti, Anna Monreale, and Franco Turini

Adversarial Robustness Curves 172
Christina Göpfert, Jan Philip Göpfert, and Barbara Hammer

Enriching Visual with Verbal Explanations for Relational Concepts – Combining LIME with Aleph 180
Johannes Rabold, Hannah Deininger, Michael Siebers, and Ute Schmid

Quantifying Model Complexity via Functional Decomposition for Better Post-hoc Interpretability. 193
Christoph Molnar, Giuseppe Casalicchio, and Bernd Bischl

Sampling, Intervention, Prediction, Aggregation: A Generalized Framework for Model-Agnostic Interpretations 205
Christian A. Scholbeck, Christoph Molnar, Christian Heumann, Bernd Bischl, and Giuseppe Casalicchio

Learning and Interpreting Potentials for Classical Hamiltonian Systems 217
Harish S. Bhat

Finding Interpretable Concept Spaces in Node Embeddings Using Knowledge Bases 229
Maximilian Idahl, Megha Khosla, and Avishek Anand

Local Interpretation Methods to Machine Learning Using the Domain of the Feature Space 241
Tiago Botari, Rafael Izbicki, and Andre C. P. L. F. de Carvalho

Measuring Unfairness Through Game-Theoretic Interpretability 253
Juliana Cesaro and Fabio Gagliardi Cozman

LioNets: Local Interpretation of Neural Networks Through Penultimate Layer Decoding 265
Ioannis Mollas, Nikolaos Bassiliades, and Grigorios Tsoumakas

Decentralized Machine Learning at the Edge

Distributed Generative Modelling with Sub-linear Communication Overhead 281
Nico Piatkowski

Distributed Learning of Neural Networks with One Round of Communication. 293
Mike Izbicki and Christian R. Shelton

Decentralized Learning with Budgeted Network Load Using Gaussian Copulas and Classifier Ensembles 301
John Klein, Mahmoud Albardan, Benjamin Guedj, and Olivier Colot

Decentralized Recommendation Based on Matrix Factorization: A Comparison of Gossip and Federated Learning 317
István Hegedűs, Gábor Danner, and Márk Jelasity

Ring-Star: A Sparse Topology for Faster Model Averaging in Decentralized Parallel SGD 333
Mohsan Jameel, Josif Grabocka, Mofassir ul Islam Arif, and Lars Schmidt-Thieme

Hardware Acceleration of Machine Learning Beyond Linear Algebra 342
Sascha Mücke, Nico Piatkowski, and Katharina Morik

Advances in Managing and Mining Large Evolving Graphs - 3rd Edition (LEG)

Detecting Stable Communities in Link Streams at Multiple Temporal Scales 353
Souâad Boudebza, Rémy Cazabet, Omar Nouali, and Faiçal Azouaou

A Comparative Study of Community Detection Techniques for Large Evolving Graphs. 368
Lauranne Coppens, Jonathan De Venter, Sandra Mitrović, and Jochen De Weerdt

Dynamic Joint Variational Graph Autoencoders 385
Sedigheh Mahdavi, Shima Khoshraftar, and Aijun An

Evolution Analysis of Large Graphs with Gradoop 402
Christopher Rost, Andreas Thor, Philip Fritzsche, Kevin Gomez, and Erhard Rahm

MHDNE: Network Embedding Based on Multivariate Hawkes Process 409
Ying Yin, Jianpeng Zhang, Yulong Pei, Xiaotao Cheng, and Lixin Ji

Data and Machine Learning Advances with Multiple Views

Multimodal Deep Networks for Text and Image-Based Document Classification. 427
Nicolas Audebert, Catherine Herold, Kuider Slimani, and Cédric Vidal

Manifold Mixing for Stacked Regularization. 444
João Pereira, Erik S. G. Stroes, Albert K. Groen, Aeilko H. Zwinderman, and Evgeni Levin

A Wide and Deep Neural Network for Survival Analysis from Anatomical Shape and Tabular Clinical Data 453
Sebastian Pölsterl, Ignacio Sarasua, Benjamín Gutiérrez-Becker, and Christian Wachinger

Deep Generative Multi-view Learning 465
Mahdi Karami

New Trends in Representation Learning with Knowledge Graphs (KGRL)

SDE-KG: A Stochastic Dynamic Environment for Knowledge Graphs. 483
Varun Ranganathan and Natarajan Subramanyam

Iterative Representation Learning for Entity Alignment Leveraging Textual Information. 489
Weixin Zeng, Jiuyang Tang, and Xiang Zhao

Fourth Workshop on Data Science for Social Good (SoGood 2019)

#MeTooMaastricht: Building a Chatbot to Assist Survivors of Sexual Harassment 503
Tobias Bauer, Emre Devrim, Misha Glazunov, William Lopez Jaramillo, Balaganesh Mohan, and Gerasimos Spanakis

Analysis of Vocational Education and Training and the Labour Market in Catalonia. A Data-Driven Approach 522
José Mena, Marc Torrent-Moreno, Daniel González, Laura Portell, Oriol Pujol, and Jordi Vitrià

SOS-EW: System for Overdose Spike Early Warning Using Drug Mover's Distance-Based Hawkes Processes 538
Wen-Hao Chiang, Baichuan Yuan, Hao Li, Bao Wang, Andrea Bertozzi, Jeremy Carter, Brad Ray, and George Mohler

Improving GP-UCB Algorithm by Harnessing Decomposed Feedback 555
Kai Wang, Bryan Wilder, Sze-chuan Suen, Bistra Dilkina, and Milind Tambe

Optimizing Waste Collection: A Data Mining Approach 570
Guilherme Londres, Nuno Filipe, and João Gama

Mobile Game Theory with Street Gangs 579
Sarah Cooney, Wendy Gomez, Kai Wang, Jorja Leap, P. Jeffrey Brantingham, and Milind Tambe

Paired-Consistency: An Example-Based Model-Agnostic Approach to Fairness Regularization in Machine Learning 590
Yair Horesh, Noa Haas, Elhanan Mishraky, Yehezkel S. Resheff, and Shir Meir Lador

Transferring Clinical Prediction Models Across Hospitals and Electronic Health Record Systems 605
Alicia Curth, Patrick Thoral, Wilco van den Wildenberg, Peter Bijlstra, Daan de Bruin, Paul Elbers, and Mattia Fornasa

Linking Physicians to Medical Research Results via Knowledge Graph Embeddings and Twitter 622
Afshin Sadeghi and Jens Lehmann

Prediction of Frequent Out-Of-Hours' Medical Use 631
Duncan Wallace and Tahar Kechadi

Forecast of Study Success in the STEM Disciplines Based Solely on Academic Records 647
Lukas Pensel and Stefan Kramer

Improving Access to Science for Social Good 658
Mehdi Ali, Sahar Vahdati, Shruti Singh, Sourish Dasgupta, and Jens Lehmann

Author Index 675

Contents – Part II

Second International Workshop on Knowledge Discovery and User Modeling for Smart Cities (UMCit)

District Heating Substation Behaviour Modelling for Annotating the Performance 3
Shahrooz Abghari, Veselka Boeva, Jens Brage, and Christian Johansson

Modeling Evolving User Behavior via Sequential Clustering 12
Veselka Boeva and Christian Nordahl

Recognizing User's Activity and Transport Mode Detection: Maintaining Low-Power Consumption 21
Fitore Muharemi, Egzon Syka, and Doina Logofatu

Can Twitter Help to Predict Outcome of 2019 Indian General Election: A Deep Learning Based Study 38
Amit Agarwal, Durga Toshniwal, and Jatin Bedi

Towards Sensing and Sharing Auditory Context Information Using Wearable Device 54
Akio Sashima and Mitsuru Kawamoto

Workshop on Data Integration and Applications (DINA)

Noise Reduction in Distant Supervision for Relation Extraction Using Probabilistic Soft Logic 63
Birgit Kirsch, Zamira Niyazova, Michael Mock, and Stefan Rüping

Privacy-Preserving Record Linkage to Identify Fragmented Electronic Medical Records in the All of Us Research Program 79
Abel N. Kho, Jingzhi Yu, Molly Scannell Bryan, Charon Gladfelter, Howard S. Gordon, Shaun Grannis, Margaret Madden, Eneida Mendonca, Vesna Mitrovic, Raj Shah, Umberto Tachinardi, and Bradley Taylor

Data Integration for the Development of a Seismic Loss Prediction Model for Residential Buildings in New Zealand 88
Samuel Roeslin, Quincy Ma, Joerg Wicker, and Liam Wotherspoon

Linking IT Product Records 101
Katsiaryna Mirylenka, Paolo Scotton, Christoph Miksovic, and Salah-Eddine Bariol Alaoui

Pharos: Query-Driven Schema Inference for the Semantic Web 112
David Haller and Richard Lenz

Informativeness-Based Active Learning for Entity Resolution 125
Victor Christen, Peter Christen, and Erhard Rahm

Encoding Hierarchical Classification Codes for Privacy-Preserving Record Linkage Using Bloom Filters 142
Rainer Schnell and Christian Borgs

Machine Learning for Cybersecurity (MLCS)

Are Network Attacks Outliers? A Study of Space Representations and Unsupervised Algorithms 159
Félix Iglesias, Alexander Hartl, Tanja Zseby, and Arthur Zimek

Auto Semi-supervised Outlier Detection for Malicious Authentication Events 176
Georgios Kaiafas, Christian Hammerschmidt, Sofiane Lagraa, and Radu State

Defense-VAE: A Fast and Accurate Defense Against Adversarial Attacks . . . 191
Xiang Li and Shihao Ji

Analyzing and Storing Network Intrusion Detection Data Using Bayesian Coresets: A Preliminary Study in Offline and Streaming Settings 208
Fabio Massimo Zennaro

6th Workshop on Sports Analytics: Machine Learning and Data Mining for Sports Analytics (MLSA)

Analyzing Soccer Players' Skill Ratings Over Time Using Tensor-Based Methods 225
Kenneth Verstraete, Tom Decroos, Bruno Coussement, Nick Vannieuwenhoven, and Jesse Davis

Exploring Successful Team Tactics in Soccer Tracking Data 235
L. A. Meerhoff, F. R. Goes, A-.W. De Leeuw, and A. Knobbe

Soccer Team Vectors 247
Robert Müller, Stefan Langer, Fabian Ritz, Christoph Roch, Steffen Illium, and Claudia Linnhoff-Popien

Tactical Analyses in Professional Tennis 258
Arie-Willem de Leeuw, Aldo Hoekstra, Laurentius Meerhoff, and Arno Knobbe

Difficulty Classification of Mountainbike Downhill Trails Utilizing Deep Neural Networks 270
Stefan Langer, Robert Müller, Kyrill Schmid, and Claudia Linnhoff-Popien

First Workshop on Categorizing Different Types of Online Harassment Languages in Social Media

Categorizing Online Harassment on Twitter 283
Mozhgan Saeidi, Samuel Bruno da S. Sousa, Evangelos Milios, Norbert Zeh, and Lilian Berton

Learning to Detect Online Harassment on Twitter with the Transformer 298
Margarita Bugueño and Marcelo Mendoza

Detection of Harassment on Twitter with Deep Learning Techniques 307
Ignacio Espinoza and Fernanda Weiss

Gradient Boosting Machine and LSTM Network for Online Harassment Detection and Categorization in Social Media 314
Fabíola S. F. Pereira, Thiago Andrade, and André C. P. L. F. de Carvalho

Attention-Based Method for Categorizing Different Types of Online Harassment Language 321
Christos Karatsalos and Yannis Panagiotakis

IoT Stream for Data Driven Predictive Maintenance

SPICE: Streaming PCA Fault Identification and Classification Engine in Predictive Maintenance 333
Cristian Axenie, Radu Tudoran, Stefano Bortoli, Mohamad Al Hajj Hassan, Alexander Wieder, and Goetz Brasche

Event-Based Predictive Maintenance on Top of Sensor Data in a Real Industry 4.0 Case Study 345
Athanasios Naskos, Georgia Kougka, Theodoros Toliopoulos, Anastasios Gounaris, Cosmas Vamvalis, and Daniel Caljouw

Forecasting of Product Quality Through Anomaly Detection 357
Mehmet Dinç, Şeyda Ertekin, Hadi Özkan, Can Meydanlı, and Volkan Atalay

Data Preprocessing and Dynamic Ensemble Selection for Imbalanced Data Stream Classification 367
Paweł Zyblewski, Robert Sabourin, and Michał Woźniak

A Study on Imbalanced Data Streams 380
Ehsan Aminian, Rita P. Ribeiro, and João Gama

Mining Human Mobility Data to Discover Locations and Habits 390
Thiago Andrade, Brais Cancela, and João Gama

Imbalanced Data Stream Classification Using Hybrid Data Preprocessing 402
Barbara Bobowska, Jakub Klikowski, and Michał Woźniak

A Machine Learning-Based Approach for Predicting Tool Wear in Industrial Milling Processes 414
Mathias Van Herreweghe, Mathias Verbeke, Wannes Meert, and Tom Jacobs

12th International Workshop on Machine Learning and Music (MML 2019)

Cross-version Singing Voice Detection in Opera Recordings: Challenges for Supervised Learning 429
Stylianos I. Mimilakis, Christof Weiss, Vlora Arifi-Müller, Jakob Abeßer, and Meinard Müller

Neural Symbolic Music Genre Transfer Insights 437
Gino Brunner, Mazda Moayeri, Oliver Richter, Roger Wattenhofer, and Chi Zhang

Familiar Feelings: Listener-Rated Familiarity in Music Emotion Recognition 446
Lloyd May and Michael Casey

Rhythm, Chord and Melody Generation for Lead Sheets Using Recurrent Neural Networks 454
Cedric De Boom, Stephanie Van Laere, Tim Verbelen, and Bart Dhoedt

Bacher than Bach? On Musicologically Informed AI-Based Bach Chorale Harmonization 462
Alexander Leemhuis, Simon Waloschek, and Aristotelis Hadjakos

Adaptively Learning to Recognize Symbols in Handwritten Early Music 470
Luisa Micó, Jose Oncina, and José M. Iñesta

Feature-Based Classification of Electric Guitar Types 478
Renato de Castro Rabelo Profeta and Gerald Schuller

RecurSIA-RRT: Recursive Translatable Point-Set Pattern Discovery with Removal of Redundant Translators 485
David Meredith

Bow Gesture Classification to Identify Three Different Expertise Levels: A Machine Learning Approach 494
David Dalmazzo and Rafael Ramírez

Symbolic Music Classification Based on Multiple Sequential Patterns 502
Kerstin Neubarth and Darrell Conklin

OPTISIA: An Evolutionary Approach to Parameter Optimisation in a Family of Point-Set Pattern-Discovery Algorithms 509
Viktor Schmuck and David Meredith

Predicting Dynamics in Violin Pieces with Features from Melodic Motifs . . . 517
Fábio Jose Muneratti Ortega, Alfonso Perez-Carrillo, and Rafael Ramírez

Sequence Generation Using Unwords 524
Darrell Conklin

A Machine Learning Approach to Study Expressive Performance Deviations in Classical Guitar 531
Sergio Giraldo, Alberto Nasarre, Isabelle Heroux, and Rafael Ramirez

Enhanced De-Essing via Neural Networks 537
Simon Hestermann and Niklas Deffner

Representation, Exploration and Recommendation of Playlists 543
Piyush Papreja, Hemanth Venkateswara, and Sethuraman Panchanathan

Large-Scale Biomedical Semantic Indexing and Question Answering (BioASQ)

Results of the Seventh Edition of the BioASQ Challenge 553
Anastasios Nentidis, Konstantinos Bougiatiotis, Anastasia Krithara, and Georgios Paliouras

Selected Approaches Ranking Contextual Term for the BioASQ Multi-label Classification (Task6a and 7a) 569
Bernd Müller and Dietrich Rebholz-Schuhmann

Convolutional Neural Network for Automatic MeSH Indexing 581
Alastair R. Rae, James G. Mork, and Dina Demner-Fushman

A Mixed Information Source Approach for Biomedical Question Answering: MindLab at BioASQ 7B 595
Mónica Pineda-Vargas, Andrés Rosso-Mateus, Fabio A. González, and Manuel Montes-y-Gómez

AUEB at BioASQ 7: Document and Snippet Retrieval 607
Dimitris Pappas, Ryan McDonald, Georgios-Ioannis Brokos, and Ion Androutsopoulos

Classification Betters Regression in Query-Based Multi-document Summarisation Techniques for Question Answering: Macquarie University at BioASQ7b 624
Diego Mollá and Christopher Jones

Structured Summarization of Academic Publications 636
Alexios Gidiotis and Grigorios Tsoumakas

How to Pre-train Your Model? Comparison of Different Pre-training Models for Biomedical Question Answering 646
Sanjay Kamath, Brigitte Grau, and Yue Ma

Yes/No Question Answering in BioASQ 2019 661
Dimitris Dimitriadis and Grigorios Tsoumakas

Semantically Corroborating Neural Attention for Biomedical Question Answering 670
Marilena Oita, K. Vani, and Fatma Oezdemir-Zaech

Measuring Domain Portability and Error Propagation in Biomedical QA 686
Stefan Hosein, Daniel Andor, and Ryan McDonald

UNCC Biomedical Semantic Question Answering Systems. BioASQ: Task-7B, Phase-B 695
Sai Krishna Telukuntla, Aditya Kapri, and Wlodek Zadrozny

Transformer Models for Question Answering at BioASQ 2019 711
Michele Resta, Daniele Arioli, Alessandro Fagnani, and Giuseppe Attardi

Pre-trained Language Model for Biomedical Question Answering 727
Wonjin Yoon, Jinhyuk Lee, Donghyeon Kim, Minbyul Jeong, and Jaewoo Kang

Author Index 741