Andreas Holzinger
Peter Kieseberg
A Min Tjoa
Edgar Weippl (Eds.)

LNCS 13480

# Machine Learning and Knowledge Extraction

6th IFIP TC 5, TC 12, WG 8.4, WG 8.9, WG 12.9
International Cross-Domain Conference, CD-MAKE 2022
Vienna, Austria, August 23–26, 2022
Proceedings

Springer

# Lecture Notes in Computer Science 13480


Founding Editors

Gerhard Goos
*Karlsruhe Institute of Technology, Karlsruhe, Germany*

Juris Hartmanis
*Cornell University, Ithaca, NY, USA*

Editorial Board Members

Elisa Bertino
*Purdue University, West Lafayette, IN, USA*

Wen Gao
*Peking University, Beijing, China*

Bernhard Steffen
*TU Dortmund University, Dortmund, Germany*

Moti Yung
*Columbia University, New York, NY, USA*

More information about this series at https://link.springer.com/bookseries/558

Andreas Holzinger · Peter Kieseberg ·
A Min Tjoa · Edgar Weippl (Eds.)

# Machine Learning and Knowledge Extraction

6th IFIP TC 5, TC 12, WG 8.4, WG 8.9, WG 12.9
International Cross-Domain Conference, CD-MAKE 2022
Vienna, Austria, August 23–26, 2022
Proceedings

Springer

*Editors*
Andreas Holzinger
University of Natural Resources and Life Sciences Vienna
Vienna, Austria

Medical University Graz and Graz University of Technology
Graz, Austria

University of Alberta
Edmonton, Canada

A Min Tjoa
TU Wien
Vienna, Austria

Peter Kieseberg
St. Pölten University of Applied Sciences
St. Pölten, Austria

Edgar Weippl
SBA Research
Vienna, Austria

University of Vienna
Vienna, Austria

ISSN 0302-9743 ISSN 1611-3349 (electronic)
Lecture Notes in Computer Science
ISBN 978-3-031-14462-2 ISBN 978-3-031-14463-9 (eBook)
https://doi.org/10.1007/978-3-031-14463-9

© IFIP International Federation for Information Processing 2022
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 International Cross Domain Conference for Machine Learning & Knowledge Extraction CD-MAKE is a joint effort of IFIP TC 5, TC 12, IFIP WG 8.4, IFIP WG 8.9 and IFIP WG 12.9, and is held in conjunction with the International Conference on Availability, Reliability and Security (ARES) – this time in beautiful Vienna, Austria, during August 23–26, 2022. Thanks to the current improved situation of the COVID-19 pandemic, which affected us all heavily, we were happy that we could meet all our international colleagues and friends in person again.

For those who are new to our conference, the letters CD in CD-MAKE stand for "Cross-Domain" and describe the integration and appraisal of different fields and application domains to provide an atmosphere to foster different perspectives and opinions. We are strongly convinced that this cross-domain approach is very fruitful for new developments and novel discoveries. The conference fosters an integrative machine learning approach, considering the importance of data science and visualization for the algorithmic pipeline with a strong emphasis on privacy, data protection, safety, and security. It is dedicated to offering an international platform for novel ideas and a fresh look on methodologies to put crazy ideas into business for the benefit of humans. Serendipity is a desired effect for the cross-fertilization of methodologies and transfer of algorithmic developments.

The acronym MAKE stands for "MAchine Learning and Knowledge Extraction", a field of artificial intelligence (AI) that, while quite old in its fundamentals, has just recently begun to thrive based on both novel developments in the algorithmic area and the availability of vast computing resources at a comparatively low cost.

Machine learning (ML) studies algorithms that can learn from data to gain knowledge from experience and to generate decisions and predictions. A grand goal is in understanding intelligence for the design and development of algorithms that work autonomously (ideally without a human-in-the-loop) and can improve their learning behavior over time. The challenge is to discover relevant structural and/or temporal patterns ("knowledge") in data, which is often hidden in arbitrarily high dimensional spaces, and thus simply not accessible to humans. Knowledge extraction is one of the oldest fields in AI and is seeing a renaissance, particularly in the combination of statistical methods with classical ontological approaches.

AI is currently undergoing a kind of Cambrian explosion and is the fastest growing field in computer science today thanks to the successes in machine learning to help to solve real-world problems. There are many application domains, e.g., in agriculture, climate research, forestry, etc. with many use cases from our daily lives, which can be useful to help to solve various problems Examples include recommender systems, speech recognition, autonomous driving, cyber-physical systems, robotics, etc.

However, in our opinion the grand challenges are in sensemaking, in context understanding, and in decision making under uncertainty, as well as solving the problems of human interpretability, explainability, and verification.

Our real world is full of uncertainties and probabilistic inference enormously influences AI generally and ML specifically. The inverse probability allows us to infer unknowns, to learn from data, and to make predictions to support decision-making. Whether in social networks, recommender systems, health applications, or industrial applications, the increasingly complex data sets require a joint interdisciplinary effort bringing the human-in-control and to manage ethical and social issues, accountability, retractability, explainability, causability and privacy, safety, and security!

The International Federation for Information Processing (IFIP) is the leading multinational, non-governmental, apolitical organization in information and communications technologies and computer sciences. IFIP is recognized by the United Nations (UN) and was established in 1960 under the auspices of UNESCO as an outcome of the first World Computer Congress held in Paris in 1959.

IFIP is incorporated in Austria by decree of the Austrian Foreign Ministry (September 20, 1996, GZ 1055.170/120-I.2/96) granting IFIP the legal status of a non-governmental international organization under the Austrian Law on the Granting of Privileges to Non-Governmental International Organizations (Federal Law Gazette 1992/174). IFIP brings together more than 3500 scientists without boundaries from both academia and industry, organized in more than 100 Working Groups (WGs) and 13 Technical Committees (TCs).

To acknowledge all those who contributed to the efforts and stimulating discussions at CD-MAKE 2022 would be impossible in a preface like this. Many people contributed to the development of this volume, either directly or indirectly, so it is impossible to list all of them. We herewith thank all local, national, and international colleagues and friends for their positive and supportive encouragement. Finally, yet importantly, we thank the Springer management team and the Springer production team for their professional support. This year CD-MAKE received 45 submissions, which all have been carefully reviewed by our program committee in a double-blind review. Finally 23 papers have been accepted and were presented at the conference in Vienna.

Thank you to all! Let's MAKE it cross-domain!

August 2022

Andreas Holzinger
Peter Kieseberg
Edgar Weippl
A Min Tjoa

# Organization

## Conference Organizers

| | |
|---|---|
| Andreas Holzinger | University of Natural Resources and Life Sciences Vienna, Medical University of Graz, and Graz University of Technology, Austria/University of Alberta, Canada |
| Peter Kieseberg | FH St. Pölten, Austria |
| Edgar Weippl (IFIP WG 8.4 Chair) | SBA Research and University of Vienna, Austria |
| A Min Tjoa (IFIP WG 8.9. Chair, Honorary Secretary IFIP) | TU Vienna, Austria |

## Conference Office Team

| | |
|---|---|
| Bettina Jaber | SBA Research, Austria |
| Daniela Freitag David | SBA Research, Austria |
| Barbara Friedl | St. Pölten University of Applied Sciences, Austria |

## Proceedings Manager

| | |
|---|---|
| Bettina Jaber | SBA Research, Austria |

## Program Committee

| | |
|---|---|
| Amin Anjomshoaa | National University of Ireland Galway, Ireland |
| John A. Atkinson-Abutridy | Universidad de Concepcion, Chile |
| Frantisek Babic | Technical University of Košice, Slovakia |
| Christian Bauckhage | University of Bonn, Germany |
| Smaranda Belciug | University of Craiova, Romania |
| Mounir Ben Ayed | Ecole Nationale d'Ingenieurs de Sfax, Tunesia |
| Chris Biemann | Universität Hamburg, Germany |
| Guido Bologna | Université de Genève, Switzerland |
| Ivan Bratko | University of Ljubljana, Slovenia |
| Federico Cabitza | Università degli Studi di Milano-Bicocca, Italy |
| Andre Calero-Valdez | RWTH Aachen University, Germany |
| Andrea Campagner | University of Milano-Bicocca, Italy |
| Angelo Cangelosi | University of Manchester, UK |
| Mirko Cesarini | Università degli Studi di Milano-Bicocca, Italy |
| Nitesh V. Chawla | University of Notre Dame, USA |

| | |
|---|---|
| Roberto Confalonieri | Free University of Bozen-Bolzano, Italy |
| Tim Conrad | Zuse Institute Berlin, Germany |
| Gloria Cerasela Crisan | Vasile Alecsandri University of Bacau, Romania |
| Beatriz De La Iglesia | University of East Anglia, UK |
| Javier Del Ser | Universidad del País Vasco/Euskal Herriko Unibertsitatea, Spain |
| Natalia Dias Rodriguez | University of Granada, Spain |
| Josep Domingo-Ferrer | Universitat Rovira i Virgili, Catalonia |
| Max J. Egenhofer | University of Maine, USA |
| Massimo Ferri | University of Bologna, Italy |
| Hugo Gamboa | PLUX Wireless Biosignals and Universidade Nova de Lisboa, Portugal |
| Panagiotis Germanakos | SAP, Germany |
| Siegfried Handschuh | University of Passau, Germany |
| Pitoyo Hartono | Chukyo University, Japan |
| Barna Laszlo Iantovics | George Emil Palade University of Medicine, Pharmacy, Science and Technology of Targu Mures, Romania |
| Igor Jurisica | IBM Life Sciences Discovery Centre and Princess Margaret Cancer Centre, Canada |
| Epaminodas Kapetanios | University of Westminster, UK |
| Andreas Kerren | Linköping University and Linnaeus University, Sweden |
| Freddy Lecue | Inria, France |
| Lenka Lhotska | Czech Technical University in Prague, Czech Republic |
| Shujun Li | University of Kent, UK |
| Brian Lim | National University of Singapore, Singapore |
| Ljiljana Majnaric-Trtica | University of Osijek, Croatia |
| Bradley Malin | Vanderbilt University, USA |
| Ernestina Menasalvas | Polytechnic University of Madrid, Spain |
| Fabio Mercorio | University of Milano-Bicocca, Italy |
| Yoan Miche | Nokia Bell Labs, Finland |
| Vasile Palade | Coventry University, UK |
| Jan Paralic | Technical University of Kosice, Slovakia |
| Francesco Piccialli | University of Naples Federico II, Italy |
| Camelia-M. Pintea | Technical University of Cluj-Napoca, Romania |
| Fabrizio Riguzzi | Università di Ferrara, Italy |
| Luca Romeo | Università Politecnica delle Marche and Istituto Italiano di Tecnologia, Italy |
| Irena Spasic | Cardiff University, UK |
| Ivan Štajduhar | University of Rijeka, Croatia |
| Bharath Sudharsan | National University of Ireland Galway, Ireland |

| | |
|---|---|
| Dimitar Trajanov | Ss. Cyril and Methodius University of Skopje, Macedonia |
| Isaac Triguero | University of Nottingham, UK |
| Cagataj Turkay | University of Warwick, UK |
| Dmitry Vetrov | Yandex Research, Russia |
| Markus Wagner | University of Applied Sciences St. Pölten, Austria |
| Xuezhong Zhou | Beijing Jiaotong University, China |

# Contents

Explain to Not Forget: Defending Against Catastrophic Forgetting with XAI ... 1
*Sami Ede, Serop Baghdadlian, Leander Weber, An Nguyen, Dario Zanca, Wojciech Samek, and Sebastian Lapuschkin*

Approximation of SHAP Values for Randomized Tree Ensembles ..... 19
*Markus Loecher, Dingyi Lai, and Wu Qi*

Color Shadows (Part I): Exploratory Usability Evaluation of Activation Maps in Radiological Machine Learning ..... 31
*Federico Cabitza, Andrea Campagner, Lorenzo Famiglini, Enrico Gallazzi, and Giovanni Andrea La Maida*

Effects of Fairness and Explanation on Trust in Ethical AI ..... 51
*Alessa Angerschmid, Kevin Theuermann, Andreas Holzinger, Fang Chen, and Jianlong Zhou*

Towards Refined Classifications Driven by SHAP Explanations ..... 68
*Yusuf Arslan, Bertrand Lebichot, Kevin Allix, Lisa Veiber, Clément Lefebvre, Andrey Boytsov, Anne Goujon, Tegawendé F. Bissyandé, and Jacques Klein*

Global Interpretable Calibration Index, a New Metric to Estimate Machine Learning Models' Calibration ..... 82
*Federico Cabitza, Andrea Campagner, and Lorenzo Famiglini*

The ROC Diagonal is Not Layperson's Chance: A New Baseline Shows the Useful Area ..... 100
*André M. Carrington, Paul W. Fieguth, Franz Mayr, Nick D. James, Andreas Holzinger, John W. Pickering, and Richard I. Aviv*

Debiasing MDI Feature Importance and SHAP Values in Tree Ensembles ..... 114
*Markus Loecher*

The Influence of User Diversity on Motives and Barriers when Using Health Apps - A Conjoint Investigation of the Intention-Behavior Gap ..... 130
*Eva Rössler, Patrick Halbach, Laura Burbach, Martina Ziefle, and André Calero Valdez*

Identifying Fraud Rings Using Domain Aware Weighted Community Detection . . . . . 150
*Shaik Masihullah, Meghana Negi, Jose Matthew, and Jairaj Sathyanarayana*

Capabilities, Limitations and Challenges of Style Transfer with CycleGANs: A Study on Automatic Ring Design Generation . . . . . 168
*Tomas Cabezon Pedroso, Javier Del Ser, and Natalia Díaz-Rodríguez*

Semantic Causal Abstraction for Event Prediction . . . . . 188
*Sasha Strelnikoff, Aruna Jammalamadaka, and Tsai-Ching Lu*

An Evaluation Study of Intrinsic Motivation Techniques Applied to Reinforcement Learning over Hard Exploration Environments . . . . . 201
*Alain Andres, Esther Villar-Rodriguez, and Javier Del Ser*

Towards Generating Financial Reports from Tabular Data Using Transformers . . . . . 221
*Clayton Leroy Chapman, Lars Hillebrand, Marc Robin Stenzel, Tobias Deußer, David Biesner, Christian Bauckhage, and Rafet Sifa*

Evaluating the Performance of SOBEK Text Mining Keyword Extraction Algorithm . . . . . 233
*Eliseo Reategui, Marcio Bigolin, Michel Carniato, and Rafael Antunes dos Santos*

Classification of Screenshot Image Captured in Online Meeting System . . . . . 244
*Minoru Kuribayashi, Kodai Kamakari, and Nobuo Funabiki*

A Survey on the Application of Virtual Reality in Event-Related Potential Research . . . . . 256
*Vladimir Marochko, Richard Reilly, Rachel McDonnell, and Luca Longo*

Visualizing Large Collections of URLs Using the Hilbert Curve . . . . . 270
*Poornima Belavadi, Johannes Nakayama, and André Calero Valdez*

How to Reduce the Time Necessary for Evaluation of Tree-Based Models . . . . . 290
*Viera Anderková and František Babič*

An Empirical Analysis of Synthetic-Data-Based Anomaly Detection . . . . . 306
*Majlinda Llugiqi and Rudolf Mayer*

SECI Model in Data-Based Procedure for the Assessment of the Frailty State in Diabetic Patients . . . . . 328
*František Babič, Viera Anderková, Zvonimir Bosnić, Mile Volarić, and Ljiljana Trtica Majnarić*

Comparing Machine Learning Correlations to Domain Experts' Causal Knowledge: Employee Turnover Use Case . . . . . 343
*Eya Meddeb, Christopher Bowers, and Lynn Nichol*

Machine Learning and Knowledge Extraction to Support Work Safety for Smart Forest Operations . . . . . 362
*Ferdinand Hoenigsberger, Anna Saranti, Alessa Angerschmid, Carl Orge Retzlaff, Christoph Gollob, Sarah Witzmann, Arne Nothdurft, Peter Kieseberg, Andreas Holzinger, and Karl Stampfer*

**Author Index** . . . . . 377