

Irena Koprinska · Michael Kamp ·
Annalisa Appice · Corrado Loglisci ·
Luiza Antonie · Albrecht Zimmermann ·
Riccardo Guidotti · Özlem Özgöbek et al. (Eds.)

ECML PKDD 2020 Workshops

Workshops of the European Conference on Machine Learning and Knowledge Discovery in Databases (ECML PKDD 2020): SoGood 2020, PDFL 2020, MLCS 2020, NFMCP 2020, DINA 2020, EDML 2020, XKDD 2020 and INRA 2020
Ghent, Belgium, September 14–18, 2020
Proceedings

Contents

Fifth Workshop on Data Science for Social Good (SoGood 2020)

On Modeling Labor Markets for Fine-Grained Insights	9
<i>Hendrik Santoso Sugiarto and Ee-Peng Lim</i>	
Reasoning About Neural Network Activations: An Application in Spatial Animal Behaviour from Camera Trap Classifications	26
<i>Benjamin C. Evans, Allan Tucker, Oliver R. Wearn, and Chris Carbone</i>	
Practical Lessons from Generating Synthetic Healthcare Data with Bayesian Networks	38
<i>Juan de Benedetti, Namir Oues, Zhenchen Wang, Puja Myles, and Allan Tucker</i>	
Building Trajectories Over Topology with TDA-PTS: An Application in Modelling Temporal Phenotypes of Disease	48
<i>Seyed Erfan Sajjadi, Barbara Draghi, Lucia Sacchi, Arianna Dagliani, John Holmes, and Allan Tucker</i>	
Data Decomposition Based Learning for Load Time-Series Forecasting	62
<i>Jatin Bedi and Durga Toshniwal</i>	
Mitigating Bias in Online Microfinance Platforms: A Case Study on Kiva.org	75
<i>Soumajyoti Sarkar and Hamidreza Alvani</i>	
A Left Realist Critique of the Political Value of Adopting Machine Learning Systems in Criminal Justice	92
<i>Fabio Massimo Zennaro</i>	

Workshop on Parallel, Distributed and Federated Learning (PDFL 2020)

Knowledge Discovery on Blockchains: Challenges and Opportunities for Distributed Event Detection Under Constraints	113
<i>Cedric Sanders and Thomas Liebig</i>	
Resource-Constrained On-Device Learning by Dynamic Averaging	129
<i>Lukas Heppel, Michael Kamp, Linara Adilova, Danny Heinrich, Nico Piatkowski, and Katharina Morik</i>	

**Second Workshop on Machine Learning for Cybersecurity
(MLCS 2020)**

MitM Attack Detection in BLE Networks Using Reconstruction
and Classification Machine Learning Techniques. 149
Abdelkader Lahmadi, Alexis Duque, Nathan Heraief, and Julien Franco

Advocating for Multiple Defense Strategies Against Adversarial Examples. . . 165
*Alexandre Araujo, Laurent Meunier, Rafael Pinot, and Benjamin
Negrevergne*

Hybrid Connection and Host Clustering for Community Detection
in Spatial-Temporal Network Data 178
Mark Patrick Roeling, Azqa Nadeem, and Sicco Verwer

Collaborative Learning Based Effective Malware Detection System. 205
Narendra Singh, Harsh Kasyap, and Somanath Tripathy

**Ninth International Workshop on New Frontiers in Mining
Complex Patterns (NFMCP 2020)**

A Hybrid Recommendation System Based on Bidirectional
Encoder Representations 225
Irem Islek and Sule Gunduz Oguducu

Leveraging Multi-target Regression for Predicting the Next Parallel
Activities in Event Logs. 237
Michelangelo Ceci, Angelo Impedovo, and Antonio Pellicani

A Multi-view Ensemble of Deep Models for the Detection of Deviant
Process Instances 249
*Francesco Folino, Gianluigi Folino, Massimo Guarascio,
and Luigi Pontieri*

Exploiting Temporal Convolution for Activity Prediction
in Process Analytics 263
*Francesco Folino, Massimo Guarascio, Angelica Liguori,
Giuseppe Manco, Luigi Pontieri, and Ettore Ritacco*

Workshop on Data Integration and Applications (DINA 2020)

Hyper-Parameter Optimization for Privacy-Preserving Record Linkage. 281
*Joyce Yu, Jakub Nabaglo, Dinusha Vatsalan, Wilko Henecka,
and Brian Thorne*

Group-Specific Training Data. 297
Ben Busath, Jalen Morgan, and Joseph Price

Scalable Blocking for Very Large Databases	303
<i>Andrew Borthwick, Stephen Ash, Bin Pang, Shehzad Qureshi, and Timothy Jones</i>	
Address Validation in Transportation and Logistics: A Machine Learning Based Entity Matching Approach	320
<i>Yassine Guerhazi, Sana Sellami, and Omar Boucelma</i>	
Linking Heterogeneous Data for Food Security Prediction	335
<i>Hugo Deléglise, Agnès Bégué, Roberto Interdonato, Elodie Maître d’Hôtel, Mathieu Roche, and Maguelonne Teisseire</i>	
Second Workshop on Evaluation and Experimental Design in Data Mining and Machine Learning (EDML 2020)	
Towards Better Evaluation of Multi-target Regression Models	353
<i>Evgeniya Korneva and Hendrik Blockeel</i>	
Assessing the Difficulty of Labelling an Instance in Crowdsourcing.	363
<i>Neetha Jambigi, Tirtha Chanda, Vishnu Unnikrishnan, and Myra Spiliopoulou</i>	
Experimental Evaluation of Scale, and Patterns of Systematic Inconsistencies in Google Trends Data.	374
<i>Philipp Behnen, Rene Kessler, Felix Kruse, Jorge Marx Gómez, Jan Schoenmakers, and Sergej Zerr</i>	
Assessing the Uncertainty of the Text Generating Process Using Topic Models	385
<i>Jonas Rieger, Carsten Jentsch, and Jörg Rahnenführer</i>	
A Ranking Stability Measure for Quantifying the Robustness of Anomaly Detection Methods	397
<i>Lorenzo Perini, Connor Galvin, and Vincent Vercauysen</i>	
Second International Workshop on eXplainable Knowledge Discovery in Data Mining (XKDD 2020)	
Interpretable Machine Learning – A Brief History, State-of-the-Art and Challenges	417
<i>Christoph Molnar, Giuseppe Casalicchio, and Bernd Bischl</i>	
Efficient Estimation of General Additive Neural Networks: A Case Study for CTG Data.	432
<i>P. J. G. Lisboa, S. Ortega-Martorell, M. Jayabalan, and I. Olier</i>	

What Would You Ask the Machine Learning Model? Identification of User Needs for Model Explanations Based on Human-Model Conversations 447
Michał Kuźba and Przemysław Biecek

Analyzing Forward Robustness of Feedforward Deep Neural Networks with LeakyReLU Activation Function Through Symbolic Propagation 460
Giulio Masetti and Felicita Di Giandomenico

LimeOut: An Ensemble Approach to Improve Process Fairness 475
Vaishnavi Bhargava, Miguel Couceiro, and Amedeo Napoli

Interpretable Privacy with Optimizable Utility 492
Jan Ramon and Moitree Basu

Prediction and Explanation of Privacy Risk on Mobility Data with Neural Networks 501
Francesca Naretto, Roberto Pellungrini, Franco Maria Nardini, and Fosca Giannotti

Approximate Explanations for Classification of Histopathology Patches 517
Iam Palatnik de Sousa, Marley M. B. R. Vellasco, and Eduardo Costa da Silva

Eighth International Workshop on News Recommendation and Analytics (INRA 2020)

Multi-stakeholder News Recommendation Using Hypergraph Learning 531
Alireza Gharahighehi, Celine Vens, and Konstantinos Pliakos

Monitoring Technoscientific Issues in the News 536
Alberto Cammozzo, Emanuele Di Buccio, and Federico Neresini

Pitch Proposal: Recommenders with a Mission - Assessing Diversity in News Recommendations 554
Sanne Vrijenhoek and Natali Helberger

An Educational News Dataset for Recommender Systems 562
Yujie Xing, Itishree Mohallick, Jon Atle Gulla, Özlem Özgöbek, and Lemei Zhang

Exploring Thematic Coherence in Fake News 571
Martins Samuel Dogo, Deepak P., and Anna Jurek-Loughrey

Media Bias in German News Articles: A Combined Approach 581
Timo Spinde, Felix Hamborg, and Bela Gipp

On the Coherence of Fake News Articles. 591
Iknoor Singh, Deepak P., and Anoop K.

Short Paper – DINA 2020

Integrating Spreadsheets by Identifying and Solving Layout Constraints. 611
Dirko Coetsee, McElory Hoffmann, and Luc De Raedt

Author Index 613