

Contents

DSCo-NG: A Practical Language Modeling Approach for Time Series Classification	1
<i>Daoyuan Li, Tegawendé F. Bissyandé, Jacques Klein, and Yves Le Traon</i>	
Ranking Accuracy for Logistic-GEE Models	14
<i>Nasser Davarzani, Ralf Peeters, Evgueni Smirnov, Joël Karel, and Hans-Peter Brunner-La Rocca</i>	
The Morality Machine: Tracking Moral Values in Tweets	26
<i>Livia Teernstra, Peter van der Putten, Liesbeth Noordegraaf-Eelens, and Fons Verbeek</i>	
A Hybrid Approach for Probabilistic Relational Models Structure Learning . . .	38
<i>Mouna Ben Ishak, Philippe Leray, and Nahla Ben Amor</i>	
On the Impact of Data Set Size in Transfer Learning Using Deep Neural Networks.	50
<i>Deepak Soekhoe, Peter van der Putten, and Aske Plaats</i>	
Obtaining Shape Descriptors from a Concave Hull-Based Clustering Algorithm.	61
<i>Christian Braune, Marco Dankel, and Rudolf Kruse</i>	
Visual Perception of Discriminative Landmarks in Classified Time Series . . .	73
<i>Tobias Sobek and Frank Höppner</i>	
Spotting the Diffusion of New Psychoactive Substances over the Internet . . .	86
<i>Fabio Del Vigna, Marco Avenuti, Clara Bacciu, Paolo Deluca, Marinella Petrocchi, Andrea Marchetti, and Maurizio Tesconi</i>	
Feature Selection Issues in Long-Term Travel Time Prediction	98
<i>Syed Murtaza Hassan, Luis Moreira-Matias, Jihed Khiari, and Oded Cats</i>	
A Mean-Field Variational Bayesian Approach to Detecting Overlapping Communities with Inner Roles Using Poisson Link Generation.	110
<i>Gianni Costa and Riccardo Ortale</i>	
Online Semi-supervised Learning for Multi-target Regression in Data Streams Using AMRules	123
<i>Ricardo Sousa and João Gama</i>	

A Toolkit for Analysis of Deep Learning Experiments.	134
<i>Jim O'Donoghue and Mark Roantree</i>	
The Optimistic Method for Model Estimation.	146
<i>James Brofos, Rui Shu, and Frank Zhang</i>	
Does Feature Selection Improve Classification? A Large Scale Experiment in OpenML.	158
<i>Martijn J. Post, Peter van der Putten, and Jan N. van Rijn</i>	
Learning from the News: Predicting Entity Popularity on Twitter	171
<i>Pedro Saleiro and Carlos Soares</i>	
Multi-scale Kernel PCA and Its Application to Curvelet-Based Feature Extraction for Mammographic Mass Characterization.	183
<i>Sami Dhahbi, Walid Barhoumi, and Ezzeddine Zagrouba</i>	
Weakly-Supervised Symptom Recognition for Rare Diseases in Biomedical Text	192
<i>Pierre Holat, Nadi Tomeh, Thierry Charnois, Delphine Battistelli, Marie-Christine Jaulent, and Jean-Philippe Métivier</i>	
Estimating Sequence Similarity from Read Sets for Clustering Sequencing Data.	204
<i>Petr Ryšavý and Filip Železný</i>	
Widened Learning of Bayesian Network Classifiers.	215
<i>Oliver R. Sampson and Michael R. Berthold</i>	
Vote Buying Detection via Independent Component Analysis.	226
<i>Antonio Neme and Omar Neme</i>	
Unsupervised Relation Extraction in Specialized Corpora Using Sequence Mining.	237
<i>Kata Gábor, Haïfa Zargayouna, Isabelle Tellier, Davide Buscaldi, and Thierry Charnois</i>	
A Framework for Interpolating Scattered Data Using Space-Filling Curves. . .	249
<i>David J. Weston</i>	
Privacy-Awareness of Distributed Data Clustering Algorithms Revisited	261
<i>Josenildo C. da Silva, Matthias Klusch, and Stefano Lodi</i>	
Bi-stochastic Matrix Approximation Framework for Data Co-clustering	273
<i>Lazhar Labiod and Mohamed Nadif</i>	
Sequential Cost-Sensitive Feature Acquisition.	284
<i>Gabriella Contardo, Ludovic Denoyer, and Thierry Artières</i>	

Explainable and Efficient Link Prediction in Real-World Network Data	295
<i>Jesper E. van Engelen, Hanjo D. Boekhout, and Frank W. Takes</i>	
DGRMiner: Anomaly Detection and Explanation in Dynamic Graphs	308
<i>Karel Vaculík and Luboš Popelínský</i>	
Similarity Based Hierarchical Clustering with an Application to Text Collections	320
<i>Julien Ah-Pine and Xinyu Wang</i>	
Determining Data Relevance Using Semantic Types and Graphical Interpretation Cues	332
<i>Eduardo Haruo Kamioka, André Freitas, Frederico Caroli, and Siegfried Handschuh</i>	
A First Step Toward Quantifying the Climate's Information Production over the Last 68,000 Years.	343
<i>Joshua Garland, Tyler R. Jones, Elizabeth Bradley, Ryan G. James, and James W.C. White</i>	
HAUCA Curves for the Evaluation of Biomarker Pilot Studies with Small Sample Sizes and Large Numbers of Features.	356
<i>Frank Klawonn, Junxi Wang, Ina Koch, Jörg Eberhard, and Mohamed Omar</i>	
Stability Evaluation of Event Detection Techniques for Twitter.	368
<i>Andreas Weiler, Joeran Beel, Bela Gipp, and Michael Grossniklaus</i>	
IDA 2016 Industrial Challenge: Using Machine Learning for Predicting Failures	381
<i>Camila Ferreira Costa and Mario A. Nascimento</i>	
An Optimized k-NN Approach for Classification on Imbalanced Datasets with Missing Data.	387
<i>Ezgi Can Ozan, Ekaterina Riabchenko, Serkan Kiranyaz, and Moncef Gabbouj</i>	
Combining Boosted Trees with Metafeature Engineering for Predictive Maintenance	393
<i>Vitor Cerqueira, Fábio Pinto, Claudio Sá, and Carlos Soares</i>	
Prediction of Failures in the Air Pressure System of Scania Trucks Using a Random Forest and Feature Engineering	398
<i>Christopher Gondek, Daniel Hafner, and Oliver R. Sampson</i>	
Author Index	403

Advances in Intelligent Data Analysis XV

15th International Symposium, IDA 2016, Stockholm,
Sweden, October 13-15, 2016, Proceedings

Boström, H.; Knobbe, A.; Soares, C.; Papapetrou, P.
(Eds.)

2016, XIII, 404 p. 146 illus., Softcover

ISBN: 978-3-319-46348-3