Contents

DSCo-NG: A Practical Language Modeling Approach for Time Series Classification .................................................. 1
   Daoyuan Li, Tegawendé F. Bissyandé, Jacques Klein, and Yves Le Traon

Ranking Accuracy for Logistic-GEE Models ........................................... 14
   Nasser Davarzani, Ralf Peeters, Evgueni Smirnov, Joël Karel, and Hans-Peter Brunner-La Rocca

The Morality Machine: Tracking Moral Values in Tweets .......................... 26
   Livia Teernstra, Peter van der Putten, Liesbeth Noordegraaf-Eelens, and Fons Verbeek

A Hybrid Approach for Probabilistic Relational Models Structure Learning ... 38
   Mouna Ben Ishak, Philippe Leray, and Nahla Ben Amor

On the Impact of Data Set Size in Transfer Learning Using Deep Neural Networks .................................................. 50
   Deepak Soekhoe, Peter van der Putten, and Aske Plaat

Obtaining Shape Descriptors from a Concave Hull-Based Clustering Algorithm .................................................. 61
   Christian Braune, Marco Dankel, and Rudolf Kruse

Visual Perception of Discriminative Landmarks in Classified Time Series .... 73
   Tobias Sobek and Frank Höppner

Spotting the Diffusion of New Psychoactive Substances over the Internet .... 86
   Fabio Del Vigna, Marco Avvenuti, Clara Bacci, Paolo Deluca, Marinella Petrocchi, Andrea Marchetti, and Maurizio Tesconi

Feature Selection Issues in Long-Term Travel Time Prediction .................. 98
   Syed Murtaza Hassan, Luis Moreira-Matias, Jihed Khiari, and Oded Cats

A Mean-Field Variational Bayesian Approach to Detecting Overlapping Communities with Inner Roles Using Poisson Link Generation ............................. 110
   Gianni Costa and Riccardo Ortale

Online Semi-supervised Learning for Multi-target Regression in Data Streams Using AMRules ...................................... 123
   Ricardo Sousa and João Gama
A Toolkit for Analysis of Deep Learning Experiments. .......................... 134
   Jim O’Donoghue and Mark Roantree

The Optimistic Method for Model Estimation ................................. 146
   James Brofos, Rui Shu, and Frank Zhang

Does Feature Selection Improve Classification? A Large Scale
Experiment in OpenML .......................................................... 158
   Martijn J. Post, Peter van der Putten, and Jan N. van Rijn

Learning from the News: Predicting Entity Popularity on Twitter ....... 171
   Pedro Saleiro and Carlos Soares

Multi-scale Kernel PCA and Its Application to Curvelet-Based Feature
Extraction for Mammographic Mass Characterization .................... 183
   Sami Dhahbi, Walid Barhoumi, and Ezzeddine Zagrouba

Weakly-Supervised Symptom Recognition for Rare Diseases
in Biomedical Text ............................................................ 192
   Pierre Holat, Nadi Tomeh, Thierry Charnois, Delphine Battistelli,
   Marie-Christine Jaulent, and Jean-Philippe Méttivier

Estimating Sequence Similarity from Read Sets for Clustering
Sequencing Data .................................................................... 204
   Petr Ryšavý and Filip Železný

Widened Learning of Bayesian Network Classifiers ....................... 215
   Oliver R. Sampson and Michael R. Berthold

Vote Buying Detection via Independent Component Analysis ........... 226
   Antonio Neme and Omar Neme

Unsupervised Relation Extraction in Specialized Corpora
Using Sequence Mining ......................................................... 237
   Kata Gábor, Haïfa Zargayouna, Isabelle Tellier, Davide Buscaldi,
   and Thierry Charnois

A Framework for Interpolating Scattered Data Using Space-Filling Curves. . 249
   David J. Weston

Privacy-Awareness of Distributed Data Clustering Algorithms Revisited .... 261
   Josenildo C. da Silva, Matthias Klusch, and Stefano Lodi

Bi-stochastic Matrix Approximation Framework for Data Co-clustering .... 273
   Lazhar Labiod and Mohamed Nadif

Sequential Cost-Sensitive Feature Acquisition ............................. 284
   Gabriella Contardo, Ludovic Denoyer, and Thierry Artières
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explainable and Efficient Link Prediction in Real-World Network Data</td>
<td>295</td>
</tr>
<tr>
<td>Jesper E. van Engelen, Hanjo D. Boekhout, and Frank W. Takes</td>
<td></td>
</tr>
<tr>
<td>DGRMiner: Anomaly Detection and Explanation in Dynamic Graphs</td>
<td>308</td>
</tr>
<tr>
<td>Karel Vaculík and Luboš Popelínský</td>
<td></td>
</tr>
<tr>
<td>Similarity Based Hierarchical Clustering with an Application to Text Collections</td>
<td>320</td>
</tr>
<tr>
<td>Julien Ah-Pine and Xinyu Wang</td>
<td></td>
</tr>
<tr>
<td>Determining Data Relevance Using Semantic Types and Graphical Interpretation Cues</td>
<td>332</td>
</tr>
<tr>
<td>Eduardo Haruo Kamioka, André Freitas, Frederico Caroli, and Siegfried Handschuh</td>
<td></td>
</tr>
<tr>
<td>A First Step Toward Quantifying the Climate’s Information Production over the Last 68,000 Years</td>
<td>343</td>
</tr>
<tr>
<td>Joshua Garland, Tyler R. Jones, Elizabeth Bradley, Ryan G. James, and James W.C. White</td>
<td></td>
</tr>
<tr>
<td>HAUCA Curves for the Evaluation of Biomarker Pilot Studies with Small Sample Sizes and Large Numbers of Features</td>
<td>356</td>
</tr>
<tr>
<td>Frank Klawonn, Junxi Wang, Ina Koch, Jörg Eberhard, and Mohamed Omar</td>
<td></td>
</tr>
<tr>
<td>Stability Evaluation of Event Detection Techniques for Twitter</td>
<td>368</td>
</tr>
<tr>
<td>Andreas Weiler, Joeran Beel, Bela Gipp, and Michael Grossniklaus</td>
<td></td>
</tr>
<tr>
<td>IDA 2016 Industrial Challenge: Using Machine Learning for Predicting Failures</td>
<td>381</td>
</tr>
<tr>
<td>Camila Ferreira Costa and Mario A. Nascimento</td>
<td></td>
</tr>
<tr>
<td>An Optimized k-NN Approach for Classification on Imbalanced Datasets with Missing Data</td>
<td>387</td>
</tr>
<tr>
<td>Ezgi Can Ozan, Ekaterina Riabchenko, Serkan Kiranyaz, and Moncef Gabbouj</td>
<td></td>
</tr>
<tr>
<td>Combining Boosted Trees with Metafeature Engineering for Predictive Maintenance</td>
<td>393</td>
</tr>
<tr>
<td>Vítor Cerqueira, Fábio Pinto, Claudio Sá, and Carlos Soares</td>
<td></td>
</tr>
<tr>
<td>Prediction of Failures in the Air Pressure System of Scania Trucks Using a Random Forest and Feature Engineering</td>
<td>398</td>
</tr>
<tr>
<td>Christopher Gondek, Daniel Hafner, and Oliver R. Sampson</td>
<td></td>
</tr>
<tr>
<td>Author Index</td>
<td>403</td>
</tr>
</tbody>
</table>
Advances in Intelligent Data Analysis XV
Boström, H.; Knobbe, A.; Soares, C.; Papapetrou, P. (Eds.)
2016, XIII, 404 p. 146 illus., Softcover
ISBN: 978-3-319-46348-3