DATA ANALYTICS 2018

The Seventh International Conference on Data Analytics

November 18 - 22, 2018

Athens, Greece

DATA ANALYTICS 2018 Editors

Sandjai Bhulai, Vrije Universiteit Amsterdam, the Netherlands

Dimitris Kardaras, Athens University of Economics and Business, Greece

Ivana Semanjski, University of Zagreb, Croatia/ Ghent University, Belgium

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571

Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2018) by International Academy, Research, and Industry Association (IARIA)
Please refer to the Copyright Information page.

Printed by Curran Associates, Inc. (2019)

International Academy, Research, and Industry Association (IARIA)
412 Derby Way
Wilmington, DE 19810

Phone: (408) 893-6407
Fax: (408) 527-6351

petre@iaria.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

Utilizing Data Analytics to Support Process Implementation in Knowledge-intensive Domains 1
Gregor Grambow

Developing Transparent Credit Risk Scorecards More Effectively: An Explainable Artificial Intelligence Approach 7
Gerald Fahner

Towards a Scalable Data-Intensive Text Processing Architecture with Python and Cassandra 15
Gregor-Patrick Heine, Thomas Woltron, and Alexander Wohrer

Big Data Analytics for the Small Social Enterprise: How to Create a Data-driven Approach to Address Social Challenges 19
Soraya Sedkaoui and Salim Moualdi

Dynamic Scenario-based Selection of Data Aggregation Techniques 27
Oladotun Omosebi, Nik Bessis, Yannis Korkontzelos, Evangelos Pournaras, Quanbin Sun, and Stelios Sotiriadis

Analysis of Twitter Communication During the 2017 German Federal Election 33
Marek Opuszko, Laura Bode, and Stephan Ulbricht

Social Media and Google Trends in Support of Audience Analytics: Methodology and Architecture 39
Nikos Kalatzis, Ioanna Roussaki, Christos Matsoukas, Marios Paraskevopoulos, Symeon Papavassiliou, and Simona Tonoli

Stance Classification Using Political Parties in Tokyo Metropolitan Assembly Minutes 46
Yasutomo Kimura and Minoru Sasaki

The Impact of Social Media on User's Travel Purchase Intention 50
Stavros Kaperonis

Profiling Using Fuzzy Set QCA for Medical Diagnosis-The Case of Anemia 55
Stavroula Barbounaki, Nikos Dimitrioglou, Dimitris Kardaras, and Ilias Petrounias

A Modeling Tool for Equipment Health Estimation Using a System Identification Approach 61
Alexander Dementjev, Ilkay Wunderlich, Klaus Kabitzschh, and Germar Schneider

Shine on Transport Model Simulation Data: Web-based Visualization in R using Shiny 67
Antje von Schmidt, Rita Cyganski, and Matthias Heinrichs

Machine Learning for Cyber Defense and Attack 73
Manjeet Rege and Raymond Blanch K. Mbah

Algorithms for Electrical Power Time Series Classification and Clustering 79
Gaia Ceresa, Andrea Pitto, Diego Cirio, Emanuele Ciapessoni, and Nicolas Omont

An Integration Module for Mining Textual and Visual Social Media Data 85
Mohammed Ali Eltaher

A Predictive Data Analytic for the Hardness of Hamiltonian Cycle Problem Instances 91
Gijs van Horn, Richard Olij, Joeri Sleegers, and Daan van den Berg

Predictive Analytics in Utility Vehicle Maintenance 97
Juergen Prinzbach and Stephan Trahasch

Evaluation of Ship Energy Efficiency Predictive and Optimization Models Based on Noon Reports and Condition Monitoring Datasets 103
Christos Spandonidis, Nikos Themelis, George Christopoulos, and Christos Giordamlis

Forecasting Burglary Risk in Small Areas Via Network Analysis of City Streets 109
Maria Mahfoud, Sandjai Bhulai, Rob van der Mei, Dimitry Erkin, and Elenna Dugundji

Optimal Taxi Fleet Management: a Linear Programming Approach to the Taxi Capacity Problem 115
Jacky Li and Sandjai Bhulai

Dynamic Models for Knowledge Tracing & Prediction of Future Performance 121
Androniki Sapountzi, Sandjai Bhulai, Ilja Cornelisz, and Chris van Klaveren

Efficient Use of Geographical Information Systems for Improving Transport Mode Classification 130
Jorge Rodriguez-Echeverria, Sidharta Gautama, Nico Van de Weghe, Daniel Ochoa, and Benhur Ortiz-Jaramillo

Cyber-threats Analytics for Detection of GNSS Spoofing 136
Silvio Semanjski, Ivana Semanjski, Wim De Wilde, and Alain Muls

Comparing Route Deviation Bus Operation With Respect to Dial-a-Ride Service for a Low-Demand Residential Area 141
Antonio Pratelli, Marino Lupi, Alessandro Farina, and Chiara Pratelli

Big Data-driven Multimodal Traffic Management: Trends and Challenges 149
Ivana Semanjski and Sidharta Gautama