



DATA ANALYTICS 2014

The Third International Conference on Data Analytics

August 24 - 28, 2014

Rome, Italy

DATA ANALYTICS 2014 Editors

Fritz Laux, Reutlingen University, Germany

Panos M. Pardalos, University of Florida, USA

Alain Crolotte, Teradata Corporation - El Segundo, USA

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© (2014) by International Academy, Research, and Industry Association (IARIA)
Please refer to the Copyright Information page.

Printed by Curran Associates, Inc. (2014)

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: 845-758-2634
Email: curran@proceedings.com
Web: www.proceedings.com

TABLE OF CONTENTS

DATA ANALYTICS 1: FUNDAMENTALS FOR DATA ANALYTICS

Comparison of Linear Discriminant Functions by K-fold Cross Validation	1
<i>Shuichi Shinmura</i>	
Use of Social Microblogging to Motivate Young People (NEETs) to Participate in Distance Education Through www.eBig3.eu	7
<i>Dace Ratniece</i>	
Circadian Patterns in Twitter	12
<i>Marijn Ten Thij, Sandjai Bhulai, Peter Kampstra</i>	
Enhancement of Trajectory Ontology Inference Over Domain and Temporal Rules	18
<i>Rouaa Wannous, Jamal Malki, Alain Bouju, Cecile Vincent</i>	
An Annotation Process for Data Visualization Techniques	24
<i>Geraldo Franciscani Jr., Rodrygo L. T. Santos, Raphael Ottoni, João Paulo Pesce, Wagner Meira Jr., Raquel Melo-Minardi</i>	

DATA ANALYTICS 2: APPLICATION-ORIENTED ANALYTICS I

Prediction Model Framework for Imbalanced Datasets	33
<i>Maria Rossana De Leon, Eugene Rex Jalao</i>	
A Study of VEPSO Approaches for Multiobjective Real World Applications	42
<i>Omar Andres Carmona Cortes, Andrew Rau-Chaplin, Duane Wilson, Jürgen Gaiser-Porter</i>	
Application of Change-Point Detection in Image Retrieval	49
<i>Dongwei Wei, Yuehua Wu, Xiaoping Shi</i>	

DATA ANALYTICS 3: APPLICATION-ORIENTED ANALYTICS II

Co-movement of European Stock Markets based on Association Rule Mining	54
<i>Youqin Pan, Yong Hu, Elizabeth Haran, Saverio Manago</i>	
A Method for Measuring Similarity of Simulation Time-Series Data Based on Dynamic Time Warping	59
<i>Ping Ma, Zhong Zhang, Kaibin Zhao, Yuning Li</i>	
Validation of Simulation Model based on Combined Consistency Analysis of Data and Feature	65
<i>Ming Yang, Wei Li, Lingyun Lu, Song Jiao</i>	
Early Detection of Critical Faults Using Time-Series Analysis on Heterogeneous Information Systems in the Automotive Industry	70
<i>Thomas Leitner, Christina Feilmayr, Wolfram Woess</i>	

DATA ANALYTICS 4: APPLICATION-ORIENTED ANALYTICS III

GeoTagView: Visualizing Geographic Tags Easily	76
<i>Gianpaolo Pigliasco, Gaetano Zazzaro</i>	
Document Identification with MapReduce Framework	81
<i>Yenumula Reddy</i>	
The Economic Benefits of Allocating Spectrum for Mobile Broadband in Korea	87
<i>Jae Hyouk Jahng</i>	
Statistical Uncertainty of Market Network Structures	91
<i>Petr Koldanov, Panos M. Pardalos, Victor Zamaraev</i>	

DATA ANALYTICS 5: APPLICATION-ORIENTED ANALYTICS IV

Scalable System for Textual Analysis of Stock Market Prediction	95
<i>Roy Guanyu Lin, Tzu-Chieh Tsai</i>	

Evolutionary Clustering Analysis of Multiple Edge Set Networks used for Modeling Ivory Coast Mobile Phone Data and Sensemaking	100
<i>Thomas J. Klemas, Daniel B. Rajchwald</i>	
Property Preservation in Reduction of Data Volume for Mining: A Neighborhood System Approach	105
<i>Ray Hashemi, Azita Bahrami, Nicholas Tyler, Matthew Antonelli, Bryan Dahlqvist</i>	

DATA ANALYTICS 6: BIG DATA

A Decision Support Approach for Quality Management based on Artificial Intelligence Applications	112
<i>Nafissa Yussupova, Maxim Boyko, Diana Bogdanova, Andreas Hilbert</i>	
WLBench: A Benchmark for WebLog Data	122
<i>Ahmad Ghazal, Alain Crolotte, Mohammed Al-Kateb</i>	
Comparative Analysis of Data Structures for Approximate Nearest Neighbor Search	125
<i>Alexander Ponomarenko, Nikita Avrelín, Bilegsaikhan Naidan, Leonid Boytsov</i>	
A Novel Privacy Preserving Association Rule Mining using Hadoop	131
<i>Kangsoo Jung, Sehwa Park, Sungyong Cho, Seog Park</i>	
Author Index	