PROCEEDINGS OF
THE 2011 INTERNATIONAL CONFERENCE ON
DATA MINING

DMIN 2011

Editor

Robert Stahlbock

Associate Editors

**Mahmoud Abou-Nasr, Hamid R. Arabnia
Nikolaos Kourentzes, Philippe Lenca
Wolfram-M. Lippe, Gary M. Weiss**

WORLDCOMP'11
July 18-21, 2011
Las Vegas Nevada, USA
www.world-academy-of-science.org

This volume contains papers presented at The 2011 International Conference on Data Mining (DMIN'11). Their inclusion in this publication does not necessarily constitute endorsements by editors or by the publisher.

Copyright and Reprint Permission

ISBN: 1-60132-168-6
Printed in the United States of America

CSREA Press
U. S. A.

Foreword

We are pleased to present this collection of papers submitted to the 7th International Conference on Data Mining 2011, DMIN’11 (www.dmin--2011.com), July 18-21, 2011, Monte Carlo Resort, Las Vegas, Nevada, USA.

Data mining attracts innovative and influential contributions to both research and practice, across a wide range of academic disciplines and application domains. DMIN conferences seek to acknowledge and facilitate excellence in research and applications in the area of data mining. DMIN conferences are held annually within WORLDCOMP, the largest annual gathering of researchers in computer science, computer engineering and applied computing. WORLDCOMP'11 assembles a spectrum of 22 affiliated research conferences, workshops, and symposiums into a coordinated research meeting. Each conference has its own program committee as well as referees and own indexed proceedings. Attendees have full access to all 22 conferences' sessions, tracks, and tutorials. DMIN seeks to reflect the multi- and interdisciplinary nature of data mining and to facilitate the exchange and development of novel ideas, open communication and networking amongst researchers and practitioners in different research domains. We hope that the 2011 International Conference on Data Mining will provide a forum for you to present your research in a professional environment, exchange ideas, and network and interact across research areas. DMIN conferences actively support students and beginning researchers from lesser developed countries by funding registration and accommodation, in order to allow for a truly international networking and understanding. The 2011 conference has provided an international and multicultural experience with contributions from 28 different countries. We consider the resulting diversity in attendees and the mixture of established and starting researchers as a particular advantage of an engaging conference format.

DMIN’11 attracted a large number of submissions of theoretical research papers as well as industrial reports and case studies on applications. The program committee would like to thank all those who submitted papers for review. To reflect upon feedback from previous years we will strive to further extend the quality and rigor of the review process and the constructive feedback given within the reviews. To ensure a fair, objective and transparent review process all review criteria were published on the website. Papers were evaluated regarding their relevance to DMIN, originality, significance, information content, clarity, and soundness on an international level. Each aspect was objectively evaluated, with alternative aspects finding consideration for application papers. Each paper was refereed by at least two researchers in the topical area, taken the reviewers’ expertise and confidence into consideration, with most of the papers receiving three and up to five reviews. The review process was highly competitive. We are very grateful to the many colleagues who helped in organizing the conference. In particular, we would like to thank the members of the DMIN'11 program committee. Their continuing support has been essential to further improve the quality of accepted submissions and the resulting success of the conference. The DMIN'11 program committee members are (in alphabetical order):

Mahmoud Abou-Nasr, Paulo Adeodato, Leonidas Anastasakis, Lamine M. Aouad, Jérôme Azé, Souhaib Ben Taieb, Khalid Benabdeslem, Daniel Berrar, Jane Margot Binner, Alina Campan, Pedro A. Castillo, Peng Chen, Christian Dawson, Paulo Cortez, Kevin Daimi, Qin Ding, William Eberle, Georgios Evangelidis, Mohammed Farquad, Mengling Feng, Philippe Fournier-Viger, Shunkai Fu, Peter Geczy, Giorgio Corani, Arun Gupta, Zahid Halim, Hongmei He, Thomas J. Heiman, Kenneth E. Hild II, Tzung-Pei Hong, Wei-Chiang Hong, Yo-Ping Huang, Catherine Huang, Danish Irfan, Mehrdad Jalali, Ulf Johansson, Rikard König, Amir Hossein Keyhanipour, Madjid Khalilian, Sebastian Klenk, Nikolaos Kourentzes, Terje Kristensen, Yue-Shi Lee, Jaewook Lee, Philippe Lenca, Chuan Li, Wen-Yang Lin, Tran Hoai Linh, Wolfram-M. Lippe, Jing Liu, Qi Liu, Weifeng Liu, Tanja Magoc, Jun Meng, José M. Merigo Lindahl, Gaolin Zheng Milledge, Mohamed Nadif, Alberto Ochoa-Zezzatti, David L. Olson, Parag C. Pendharkar, Antonio Luigi Perrone, Hossein Peyvandi, Vassilis Plagianakos, Guangzhi Qu, Mekki Rachida, Rabie A. Ramadan, Robert G. Reynolds, Lotfi Ben Romdhane, Motaz Saad, Ram Prasad Reddy Sadi, Gerald Schaefer, Zhang Sen, Sabrina Senatore, Xuequn Shang, Victor Sheng, Yong Shi, Tamanna Siddiqui, Vijendra Singh, Lay-Ki Soon, Robert Stahlbock, Shiliang Sun, Sundaram Suresh, Ryszard Tadeusiewicz, Jaakko Talonen, Ying Tan, Traian Marius Truta, Shun-Hung Tsai, NB Venkateswarlu, Nicole Vincent, Baoying Wang, Simon Wang, Fan Wang, Chamont Wang, Xuewei Wang, Gary Weiss, Tianyi Wu, Zijiang Yang, Bingru Yang, Faisal Zaman, Yun Zhai, Yan Zhang, Defu Zhang, Songfeng Zheng, and Shang-Ming Zhou.

We would also like to thank our publicity co-chair Ashu M. G. Solo (Fellow of British Computer Society, Principal/R&D Engineer, Maverick Technologies America Inc., Intelligent Systems Instructor, Trailblazer Intelligent Systems, Inc.), for circulating information on the conference.

Considering the increasing efforts of all towards the quality of the review process, the conference sessions and the social program of DMIN'11 we are confident that you can look forward to participating and attending a leading and reputable international conference. It is a particular pleasure to provide data mining oriented invited talks and tutorials presented by the following esteemed members of the data mining community: Nitesh V. Chawla, Peter Geczy, Michael Mahoney and Gary M. Weiss. Futhermore, it is planned to publish a Special Issue on Real World Data Mining Applications in Springer’s Annals of Information Systems. In 2010 we have published a Special Issue on Data Mining in that series.

The DMIN'11 conference organizers are also thankful to a number of co-sponsors, without whom the conference would not have been possible. The Academic and Technical Co-Sponsors of this year’s conference include: The Berkeley Initiative in Soft Computing (BISC), University of California, Berkeley, USA (http://www-bisc.cs.berkeley.edu/); Biomedical Cybernetics Laboratory, HST of Harvard University and Massachusetts Institute of Technology (MIT), USA (http://bcl.med.harvard.edu/); Intelligent Data Exploration and Analysis Laboratory, University of Texas at Austin, Austin, Texas, USA (http://www.ideal.ece.utexas.edu/); Collaboratory for Advanced Computing and Simulations (CACS), University of Southern California, USA (http://cacs.usc.edu/); Minnesota Supercomputing Institute, University of Minnesota, USA (http://www.msi.umn.edu/); Knowledge Management & Intelligent System Center (KMIS) of University of Siegen, Germany (http:// www.kmis.uni-siegen.de); UMIT, Institute of Bioinformatics and Translational Research, Austria (http://www.umit.at/page.cfm?vpath=departments/technik/bioinf&expanddiv=subDeptItem343); BioMedical Informatics & Bio-Imaging Laboratory, Georgia Institute of Technology and Emory University, Atlanta, Georgia, USA (http://www.bio-miblab.org/); Hawkeye Radiology Informatics, Department of Radiology, College of Medicine, University of Iowa, Iowa, USA (http://www.uiowa.edu/~hri/); NDSU-CIIT Green Computing and Communications Laboratory, USA (http://gcclab.org/); High Performance Computing for Nanotechnology (HPCNano) (http://www.hpcnano.org); Supercomputer Software Department (SSD), Institute of Computational Mathematics & Mathematical Geophysics, Russian Academy of Sciences (http://ssd.sscc.ru); SECLAB - An inter-university research group (University of Naples Federico II, the University of Naples Parthenope, and the Second University of Naples, Italy, http://www.seclab.unina.it); Medical Image HPC & Informatics Lab (MiHi Lab), University of Iowa, Iowa, USA (http://www.uiowa.edu/mihpclab/); International Society of Intelligent Biological Medicine (http://www.isibm.org/); World Academy of Biomedical Sciences and Technologies (http://www.worldwabt.org/wabt); Intelligent Cyberspace Engineering Lab., ICEL, Texas A&M University (Com./Texas), USA; Model-Based Engineering Laboratory, University of North Dakota, North Dakota, USA; The International Council on Medical and Care Compunetics (http://www.icmcc.org); The UK Department for Business, Enterprise & Regulatory Reform (http://www.berr.gov.uk); Scientific Technologies Corporation (http://www.stchome.com); HoIP - Health without Boundaries (http://www.hoip.eu/).

We are also grateful for the general co-sponsors and organizers including university faculty members from the Institute of Information Systems at Hamburg University, Germany (www.uni-hamburg.de/IWI), CSREA Computer Science Research, Education, and Applications Press, and the Business Intelligence Laboratory, B I^{3}S lab, Hamburg, Germany (www.bis-lab.com).

Furthermore, we want to thank all members of the steering committee of Worldcomp 2011: Selim Aissi (Intel Corp., USA), Ruzena Bajcsy (Univ. of California, Berkeley, USA), Hyunseung Choo (Sungkyunkwan Univ., South Korea), (Winston) Wai-Chi Fang, (National Chiao Tung University, Taiwan), Kun Chang Lee (Sungkyunkwan Univ., South Korea), Andy Marsh (HoIP, ICET), Layne T. Watson (Virginia Polytechnic Institute & State Univ., Virginia, USA), and Lotfi A. Zadeh (Univ. of California, Berkeley, USA).

Most importantly, we wish to express again our sincere gratitude and respect towards Professor Hamid R. Arabnia (Univ. of Georgia, USA), General Chair of all WORLDCOMP conferences, for his excellent and tireless support, organization and coordination of all affiliated events. His exemplary and professional effort in 2011 and all the years before in the WORLDCOMP steering committee makes these events possible!

Thank you all for your contribution to DMIN'11! We hope that you will experience a stimulating conference with many opportunities for future contacts, research and applications.

Robert Stahlbock
DMIN'11 General Conference Chair

Contents

SESSION: REAL-WORLD DATA MINING APPLICATIONS, CHALLENGES, AND PERSPECTIVES

SESSION: SEGMENTATION, CLUSTERING, ASSOCIATION

SESSION: REGRESSION, CLASSIFICATION

SESSION: EXPLORATIVE DATA MINING, DATA PREPROCESSING, FEATURE SELECTION

SESSION: WEB AND TEXT MINING

SESSION: FEATURE SELECTION + CLUSTERING METHODS + TESTING APPLICATIONS