

# **2011 IEEE Symposium on Computational Intelligence and Data Mining**

**(CIDM 2011)**

**Paris, France  
11 – 15 April 2011**



**IEEE Catalog Number: CFP11IDM-PRT  
ISBN: 978-1-4244-9926-7**

## TABLE OF CONTENTS

<b>CIDM-CIVI 2011 COMMITTEE</b> .....	<i>xi</i>
<b>CIDM-CIVI KEYNOTE</b> .....	<i>xiii</i>
<b>CIDM-CIVI TUTORIALS</b> .....	<i>xx</i>
<b>CIDM-CIVI 2011 TECHNICAL SESSIONS</b> .....	<i>1</i>

**Wednesday, April 13**

**8:30 – 10:30**

### **S50: Information Retrieval**

Chair: Patrick Gallinari (University Pierre and Marie Curie, France)

<b>Multiple Query-dependent RankSVM Aggregation for Document Retrieval</b> .....	<b>1</b>
Yang Wang (Nankai University, China)	
Min Lu (Nankai University, China)	
Xiaodong Pang (Nankai University, China)	
Maoqiang Xie (Nankai University, China)	
Yalou Huang (Nankai University, China)	
<b>Active Learning Using the Data Distribution for Interactive Image Classification and Retrieval</b> .....	<b>7</b>
Pierre Blanchart (Télécom ParisTech, France)	
Marin Ferecatu (CNAM, France)	
Mihai Datcu (DLR Germany)	
<b>Geodesic Distances for Web Document Clustering</b> .....	<b>15</b>
Selma Tekir (Izmir Institute of Technology, Turkey)	
Florian Mansmann (University of Konstanz, Germany)	
Daniel Keim (University of Konstanz, Germany)	
<b>Distinguishing Defined Concepts from Prerequisite Concepts in Learning Resources</b> .....	<b>22</b>
Sahar Changuel (Université Pierre et Marie Curie - Paris 6, France)	
Nicolas Labroche (Université Pierre et Marie Curie - Paris 6, France)	
<b>Feature Extraction for Multi-label Learning in the Domain of Email Classification</b> .....	<b>30</b>
José M. Carmona-Cejudo (University of Malaga, Spain)	
Manuel Baena-García (Universidad de Málaga, Spain)	
Jose del Campo-Avila (University of Malaga, Spain)	
Rafael Morales-Bueno (Universidad de Málaga, Spain)	

11:00 - 12:40

**S51: Analysis of Data for Health**

Chair: Alfredo Vellido (Universitat Politècnica de Catalunya, Spain)

**On the Use of Decision Trees for ICU Outcome Prediction in Sepsis Patients Treated with Statins.....37**

Vicent J. Ribas (Universitat Politècnica de Catalunya, Spain)  
Jesús Caballero López (Vall d'Hebron University Hospital, Spain)  
Juan Carlos Ruiz-Rodriguez (Vall d'Hebron University Hospital, Spain)  
Adolf Ruiz-Sanmartín (Vall d'Hebron University Hospital, Spain)  
Jordi Rello (Vall d'Hebron University Hospital, Spain)  
Alfredo Vellido (Universitat Politècnica de Catalunya, Spain)

**Adaptive Treatment of Anemia on Hemodialysis Patients: a Reinforcement Learning Approach.....44**

Pablo Escandell-Montero (University of Valencia, Spain)  
José M. Martínez-Martínez (University of Valencia, Spain)  
José D. Martín-Guerrero (University of Valencia, Spain)  
Emilio Soria-Olivas (University of Valencia, Spain)  
Joan Vila-Francés (University of Valencia, Spain)  
Rafael Magdalena-Benedito (University of Valencia, Spain)

**Mining Nutrition Survey Data ..... B#5**

Lydia Manikonda (International Institute of Information Technology, India)  
Raghvendra Mall (International Institute of Information Technology, India)  
Vikram Pudi (International Institute of Information Technology, India)  
Raghunatha Rao D (National Institute of Nutrition, India)

**Clustering Categorical Data: A Stability Analysis Framework ..... 58**

I.H. Jarman (LJMU, United Kingdom)  
T.A. Etchells (Liverpool John Moores University, United Kingdom)  
P.J.G. Lisboa (Liverpool John Moores University, United Kingdom)  
C.M. Beynon (Liverpool John Moores University, United Kingdom)  
J.D. Martín-Guerrero (University of Valencia, Spain)

**A Multi-biclustering Combinatorial Based Algorithm..... 66**

Ekaterina Nosova (University of Salerno, Italy)  
Giancarlo Raiconi (University of Salerno, Italy)  
Roberto Tagliaferri (University of Salerno, Italy)

14:00 - 14:40

**S52: Recommender Systems**

Chair: Irwin King (The Chinese University of Hong Kong, Hong Kong)

**Recommendation Algorithm Using Positive and Negative Latent Models..... 72**

Atsuhiko Takasu (National Institute of Informatics, Japan)  
Saranya Maneeroj (Chulalongkorn University, Thailand)

**Using Gaming Strategies For Attacker and Defender in Recommender Systems ..... 80**

Justin Zhan (North Carolina A & T State University, USA)  
Lijo Thomas (Dakota State University, USA)  
Venkata Pasumarthi (Dakota State University, USA)

15:00 - 16:00

**S53: CIDM-CIVI - Keynote**

Chair: Alessandro Sperduti (Università degli Studi di Padova, Italy)

**Process Mining: Discovering and Improving Spaghetti and Lasagna Processes**

Wil van der Aalst (Eindhoven University of Technology, The Netherlands)

16:30 - 18:30

**S54: CIDM-CIVI - Tutorial**

**Mining Networked Data**

Nicola Di Mauro (University of Bari, Italy); Donato Malerba (University of Bari, Italy)

**Thursday, April 14**

08:50 - 10:30

**S55: Classification and Clustering**

Chair: Alessandro Sperduti (Università degli Studi di Padova, Italy)

**Active Classifier Training With the 3DS Strategy ..... 88**

Tobias Reitmaier (University of Passau, Germany)

Bernhard Sick (University of Applied Sciences Deggendorf, Germany)

**Partially Supervised k-Harmonic Means Clustering ..... 96**

Thomas A. Runkler (Siemens AG, Germany)

**Local Neighbourhood Extension of SMOTE for Mining Imbalanced Data ..... 104**

Tomasz Maciejewski (Poznan University of Technology, Poland)

Jerzy Stefanowski (Poznan University of Technology, Poland)

**FGMAC: Frequent subGraph Mining with Arc Consistency ..... 112**

Brahim Douar (LIRMM-University of Montpellier 2, France)

Michel Liquiere (LIRMM-University of Montpellier 2, France)

Chiraz Latiri (URPAH-EI Manar University, Tunisia)

Yahya Slimani (URPAH-EI Manar University, Tunisia)

**Mining toxicity structural alerts from SMILES: a New Way to Derive Structure Activity Relationships ..... 120**

Thomas Ferrari (Politecnico di Milano, Italy)

Giuseppina Gini (Politecnico di Milano, Italy)

Nazanin Golbamaki Bakhtyari (Mario Negri Institute, Italy)

Emilio Benfenati (Mario Negri Institute, Italy)

10:30 - 16:30

PS6: CIDM - 2011

Poster Session

<b>Logistic Sub-models for Small Size Populations in Credit Scoring.....</b>	<b>128</b>
Bouaguel Waad (University of Tunis, Tunisia)	
Beninel Farid (CREST-ENSAI, France)	
Bel Mufti Ghazi (University of Tunis, Tunisia)	
<b>Periodic Quick Test for Classifying Long-term Activities .....</b>	<b>135</b>
Pekka Siirtola (University of Oulu, Finland)	
Heli Koskimäki (University of Oulu, Finland)	
Juha Röning (University of Oulu, Finland)	
<b>Data Mining Driven agents for Predicting Online Auction's End Price.....</b>	<b>141</b>
Preetinder Kaur (University of Technology, Australia)	
Madhu Goyal (University of Technology, Australia)	
Jie Lu (University of Technology, Australia)	
<b>A Robust F-measure for Evaluating Discovered Process Models .....</b>	<b>148</b>
Jochen De Weerd (Katholieke Universiteit Leuven, Belgium)	
Manu De Backer (Universiteit Gent, Belgium)	
Jan Vanthienen (Katholieke Universiteit Leuven, Belgium)	
Bart Baesens (Katholieke Universiteit Leuven, Belgium)	
<b>Efficient Accelerometer-Based Swimming Exercise Tracking .....</b>	<b>156</b>
Pekka Siirtola (University of Oulu, Finland)	
Perttu Laurinen (University of Oulu, Finland)	
Juha Röning (University of Oulu, Finland)	
Hannu Kinnunen (Polar Electro Oy, Finland)	
<b>Active Learning for Aspect Model in Recommender Systems .....</b>	<b>162</b>
Rasoul Karimi (University of Hildesheim, Germany)	
Christoph Freudenthaler (University of Hildesheim, Germany)	
Alexandros Nanopoulos (University of Hildesheim, Germany)	
Lars Schmidt-Thieme (University of Hildesheim, Germany)	
<b>Trend Cluster Based Compression of Geographically Distributed Data Streams .....</b>	<b>168</b>
Anna Ciampi (University of Bari, Italy)	
Annalisa Appice (University of Bari, Italy)	
Donato Malerba (University of Bari, Italy)	
Pietro Guccione (Politecnico di Bari, Italy)	

11:00 – 12:00

**S56: Special Session on "Process Mining"**

Chair: Wil van der Aalst (Eindhoven University of Technology, The Netherlands)

<b>A Framework for Semi-Automated Process Instance Discovery From Decorative Attributes.....</b>	<b>176</b>
Andrea Burattin (University of Padua, Italy)	
Roberto Vigo (Siav S.p.A, Italy)	

<b>Enhancing Precision in Process Conformance: Stability, Confidence and Severity .....</b>	<b>184</b>
Jorge Munoz-Gama (Universitat Politecnica de Catalunya, Spain)	
Josep Carmona (Universitat Politecnica de Catalunya, Spain)	
<b>User-guided discovery of declarative process models .....</b>	<b>192</b>
Fabrizio M. Maggi (Eindhoven University of Technology, The Netherlands)	
Arjan J. Mooij (Eindhoven University of Technology, The Netherlands)	
Wil M.P. van der Aalst (Eindhoven University of Technology, The Netherlands)	
<b>Discovering Process Models through Relational Disjunctive Patterns Mining .....</b>	<b>200</b>
Corrado Loglisci (University of Bari, Italy)	
Michelangelo Ceci (University of Bari, Italy)	
Annalisa Appice (University of Bari, Italy)	
Donato Malerba (University of Bari, Italy)	
 <b>14:00 – 16:00</b>	
<b>S57: Mining Spatial and Temporal Data</b>	
Chair: Nitesh Chawla (University of Notre Dame, USA)	
<b>Mining Unlabeled GPS Logs for Modes of Transport by Using Kolmogorov-Smirnov Test .....</b>	<b>B#5</b>
Miao Lin (Nanyang Technological University, Singapore)	
Wen Jing Hsu (Nanyang Technological University, Singapore)	
<b>IQ Estimation for Accurate Time-Series Classification .....</b>	<b>216</b>
Krisztian Buza (University of Hildesheim, Germany)	
Alexandros Nanopoulos (University of Hildesheim, Germany)	
Lars Schmidt-Thieme (University of Hildesheim, Germany)	
<b>Computational Intelligence Methods for Processing Misaligned, Unevenly Sampled Time Series Containing Missing Data.....</b>	<b>224</b>
Federico Cismondi (MIT, USA)	
André S. Fialho (MIT, USA)	
Susana M. Vieira (Technical University of Lisbon, Portugal)	
João M.C. Sousa (Technical University of Lisbon, Portugal)	
Shane R. Reti (Harvard Medical School, USA)	
Michael D. Howell (Harvard Medical School, USA)	
Stan N. Finkelstein (MIT, USA)	
<b>GSOM Sequence:An Unsupervised Dynamic Approach for Knowledge Discovery in Temporal Data.....</b>	<b>232</b>
Asanka Fonseka (Monash University, Australia)	
Daminda Alahakoon (Monash University, Australia)	
Susan Bedingfield (Monash University, Australia)	
<b>Empirical Comparison of Correlation Measures and Pruning Levels in Complex Networks Representing the Global Climate System .....</b>	<b>239</b>
Alex Pelan (University of Notre Dame, USA)	
Karsten Steinhaeuser (University of Notre Dame, USA)	
Nitesh V. Chawla (University of Notre Dame, USA)	
Dilkushi A. de Alwis Pitts (University of Notre Dame, USA)	
Auroop R. Ganguly (Oak Ridge National Laboratory, USA)	

**Friday, April 15**

**08:50 - 10:30**

**S58: Evolutionary and Fuzzy Computation**

Chair: Paulo Cortez (University of Minho, Portugal)

**An Intelligent Load Forecasting Expert System by Integration of Ant Colony Optimization, Genetic Algorithms and Fuzzy Logic..... B#5**

Arash Ghanbari (University of Tehran, Iran)

Salman Abbasian-Naghneh (Islamic Azad University Najafabad Branch, Iran)

Esmaeil Hadavandi (Sharif University of Technology, Iran)

**Gaussian Kernel Based Automatic Immune Clustering Algorithm..... B#5**

Ruo Chen Liu (Xidian University, China)

Jian Jia (Northwest University, China)

Zhengchun Sheng (Xidian University, China)

Jingjing Ma (Xidian University, China)

Yangyang Li (Xidian University, China)

**OLAP Navigation in the Granular Linguistic Model of a Phenomenon..... 260**

Carlos Menendez-Gonzalez (HC Energia, Spain)

Gracian Trivino (European Centre for Soft Computing, Spain)

**A GPU-based Interactive Bio-inspired Visual Clustering ..... 268**

Ugo Erra (University of Basilicata, Italy)

Bernardino Frola (University of Salerno, Italy)

Vittorio Scarano (University of Salerno, Italy)

**Generating Materialized Views Using Ant Based Approaches And Information Retrieval Technologies ..... 276**

Habiba Drias (USTHB-University of Algiers, Algeria)

**10:30 - 16:30**

**PS7: CIDM - 2011**

**Poster Session**

**KB-CB-N Classification: Towards Unsupervised Approach for Supervised Learning ..... 283**

Zahraa Said Abdallah (Monash University, Australia)

Mohamed Medhat Gaber (University of Portsmouth, United Kingdom)

**A Banner Recommendation System Based on Web Navigation History..... 291**

Giovanni Giuffrida (University of Catania, Italy)

Diego Reforgiato (University of Catania, Italy)

Giuseppe Tribulato (University of Catania, Italy)

Calogero Zarba (Neodata Group s.r.l., Italy)

**Online Autoregressive Prediction in Time Series with Delayed Disclosure ..... 297**

Jean-Marc Andreoli (Xerox Research Centre Europe, France)

Marie-Luise Schneider (Xerox Research Centre Europe, France)

<b>Sectors on Sectors (SonS): A New Hierarchical Clustering Visualization Tool .....</b>	<b>304</b>
José María Martínez-Martínez (University of Valencia, Spain)	
Pablo Escandell-Montero (University of Valencia, Spain)	
Emilio Soria-Olivas (University of Valencia, Spain)	
José D. Martín-Guerrero (University of Valencia, Spain)	
Marcelino Martínez-Sober (University of Valencia, Spain)	
Juan Gómez-Sanchis (Universidad de Valencia, Spain)	
<b>Flexible Heuristics Miner (FHM) .....</b>	<b>310</b>
A.J.M.M. Weijters (Eindhoven, University of Technology, The Netherlands)	
J.T.S. Ribeiro (Eindhoven University of Technology, The Netherlands)	
<b>Increased Classification Accuracy and Speedup Through Pair-wise Feature Selection for Support Vector Machines .....</b>	<b>318</b>
Kurt Kramer (University of South Florida, USA)	
Dmitry B. Goldgof (University of South Florida, USA)	
Lawrence O. Hall (University of South Florida, USA)	
Andrew Remsen (University of South Florida, USA)	
<b>About the Analysis of Time Series with Temporal Association Rule Mining.....</b>	<b>325</b>
Tim Schlüter (Heinrich Heine University, Germany)	
Stefan Conrad (Heinrich Heine University, Germany)	
<b>Link Pattern Prediction with Tensor Decomposition in Multi-relational Networks .....</b>	<b>333</b>
Sheng Gao (LIP6-UPMC, France)	
Ludovic Denoyer (LIP6-UPMC, France)	
Patrick Gallinari (LIP6-UPMC, France)	
 <b>11:00 - 12:20</b>	
<b>S59: Perspectives of Dimensionality Reduction and Visual Analytics</b>	
Chair: Barbara Hammer (Technical University Clausthal, Germany)	
<b>Opening Black Box Data Mining Models Using Sensitivity Analysis.....</b>	<b>341</b>
Paulo Cortez (University of Minho, Portugal)	
Mark J. Embrechts (Rensselaer Polytechnic Institute, USA)	
<b>Dimensionality Reduction Mappings.....</b>	<b>349</b>
Kerstin Bunte (University of Groningen, The Netherlands)	
Michael Biehl (University Groningen, The Netherlands)	
Barbara Hammer (University of Bielefeld, Germany)	
<b>Visual Tracking of the Millennium Development Goals with a Self-Organizing Neural Network.....</b>	<b>357</b>
Peter Sarlin (Åbo Akademi University, Finland)	
<b>Partial Generalized Correlation for Hyperspectral Data .....</b>	<b>365</b>
Marc Strickert (University of Siegen, Germany)	
Björn Labitzke (University of Siegen, Germany)	
Volker Blanz (University of Siegen, Germany)	



14:00 - 16:00

**S60: CIDM-CIVI - Tutorial**

**Mining in distribution sensitive environments: Tackling the problem of class imbalance and predictive uncertainties**

Nitesh Chawla (University of Notre Dame, USA)

**AUTHOR INDEX .....373**