

Table of Contents

The Fourth International Symposium on Intelligent Data Analysis

Feature Characterization in Scientific Datasets	1
<i>Elizabeth Bradley (University of Colorado), Nancy Collins (University of Colorado), W. Philip Kegelmeyer (Sandia National Laboratories)</i>	
Relevance Feedback in the Bayesian Network Retrieval Model: An Approach Based on Term Instantiation	13
<i>Luis M. de Campos (University of Granada), Juan M. Fernández-Luna (University of Jaén), Juan F. Huete (University of Granada)</i>	
Generating Fuzzy Summaries from Fuzzy Multidimensional Databases	24
<i>Anne Laurent (Université Pierre et Marie Curie)</i>	
A Mixture-of-Experts Framework for Learning from Imbalanced Data Sets	34
<i>Andrew Estabrooks (IBM), Nathalie Japkowicz (University of Ottawa)</i>	
Predicting Time-Varying Functions with Local Models	44
<i>Achim Lewandowski (Chemnitz University), Peter Protzel (Chemnitz University)</i>	
Building Models of Ecological Dynamics Using HMM Based Temporal Data Clustering – A Preliminary Study	53
<i>Cen Li (Tennessee State University), Gautam Biswas (Vanderbilt University), Mike Dale (Griffith University), Pat Dale (Griffith University)</i>	
Tagging with Small Training Corpora	63
<i>Nuno C. Marques (Universidade Aberta), Gabriel Pereira Lopes (Centria)</i>	
A Search Engine for Morphologically Complex Languages	73
<i>Udo Hahn (Universität Freiburg), Martin Honeck (Universitätsklinikum Freiburg), Stefan Schulz (Universitätsklinikum Freiburg)</i>	
Errors Detection and Correction in Large Scale Data Collecting	84
<i>Renato Bruni (Università di Roma), Antonio Sassano (Università di Roma)</i>	

A New Framework to Assess Association Rules	95
<i>Fernando Berzal (University of Granada), Ignacio Blanco (University of Granada), Daniel Sánchez (University of Granada), María-Amparo Vila (University of Granada)</i>	
Communities of Interest	105
<i>Corinna Cortes (AT&T Shannon Research Labs), Daryl Pregibon (AT&T Shannon Research Labs), Chris Volinsky (AT&T Shannon Research Labs)</i>	
An Evaluation of Grading Classifiers	115
<i>Alexander K. Seewald (Austrian Research Institute for Artificial Intelligence), Johannes Fürnkranz (Austrian Research Institute for Artificial Intelligence)</i>	
Finding Informative Rules in Interval Sequences	125
<i>Frank Höppner (University of Applied Sciences Emden), Frank Klawonn (University of Applied Sciences Emden)</i>	
Correlation-Based and Contextual Merit-Based Ensemble Feature Selection	135
<i>Seppo Puuronen (University of Jyväskylä), Alexey Tsymbal (University of Jyväskylä), Iryna Skrypnyk (University of Jyväskylä)</i>	
Nonmetric Multidimensional Scaling with Neural Networks	145
<i>Michiel C. van Wezel (Universiteit Leiden), Walter A. Kosters (Universiteit Leiden), Peter van der Putten (Universiteit Leiden), Joost N. Kok (Universiteit Leiden)</i>	
Functional Trees for Regression	156
<i>João Gama (University of Porto)</i>	
Data Mining with Products of Trees	167
<i>José Tomé A.S. Ferreira (Imperial College), David G.T. Denison (Imperial College), David J. Hand (Imperial College)</i>	
S^3 Bagging: Fast Classifier Induction Method with Subsampling and Bagging	177
<i>Masahiro Terabe (Mitsubishi Research Institute, Inc.), Takashi Washio (I.S.I.R., Osaka University), Hiroshi Motoda (I.S.I.R., Osaka University)</i>	
RNA-Sequence-Structure Properties and Selenocysteine Insertion	187
<i>Rolf Backofen (University of Munich)</i>	
An Algorithm for Segmenting Categorical Time Series into Meaningful Episodes	198
<i>Paul Cohen (University of Massachusetts), Niall Adams (Imperial College)</i>	

An Empirical Comparison of Pruning Methods for Ensemble Classifiers . . .	208
<i>Terry Windeatt (School of Electronics Engineering Guildford), Gholamreza Ardeshir (School of Electronics Engineering Guildford)</i>	
A Framework for Modelling Short, High-Dimensional Multivariate Time Series: Preliminary Results in Virus Gene Expression Data Analysis	218
<i>Paul Kellam (University College London), Xiaohui Liu (Brunel University), Nigel Martin (Birkbeck College), Christine Orengo (University College London), Stephen Swift (Brunel University), Allan Tucker (Brunel University)</i>	
Using Multiattribute Prediction Suffix Graphs for Spanish Part-of-Speech Tagging	228
<i>José L. Triviño-Rodríguez (University of Málaga), Rafael Morales-Bueno (University of Málaga)</i>	
Self-Supervised Chinese Word Segmentation	238
<i>Fuchun Peng (University of Waterloo), Dale Schuurmans (University of Waterloo)</i>	
Analyzing Data Clusters: A Rough Sets Approach to Extract Cluster-Defining Symbolic Rules	248
<i>Syed Sibte Raza Abidi (Universiti Sains Malaysia), Kok Meng Hoe (Universiti Sains Malaysia), Alwyn Goh (Universiti Sains Malaysia)</i>	
Finding Polynomials to Fit Multivariate Data Having Numeric and Nominal Variables	258
<i>Ryohei Nakano (Nagoya Institute of Technology), Kazumi Saito (NTT Communication Science Laboratories)</i>	
Fluent Learning: Elucidating the Structure of Episodes	268
<i>Paul R. Cohen (University of Massachusetts)</i>	
An Intelligent Decision Support Model for Aviation Weather Forecasting . . .	278
<i>Sérgio Viademonte (Monash University), Frada Burstein (Monash University)</i>	
MAMBO: Discovering Association Rules Based on Conditional Independencies	289
<i>Robert Castelo (Utrecht University), Ad Feelders (Utrecht University), Arno Siebes (Utrecht University)</i>	
Model Building for Random Fields	299
<i>R.H. Glendinning (Defence Evaluation and Research Agency)</i>	
Active Hidden Markov Models for Information Extraction	309
<i>Tobias Scheffer (University of Magdeburg), Christian Decomain (SemanticEdge), Stefan Wrobel (University of Magdeburg)</i>	

Adaptive Lightweight Text Filtering 319
Gabriel L. Somlo (Colorado State University), Adele E. Howe (Colorado State University)

A General Algorithm for Approximate Inference in Multiply Sectioned Bayesian Networks 330
Zhang Hongwei (Tsinghua University), Tian Fengzhan (Tsinghua University), Lu Yuchang (Tsinghua University)

Investigating Temporal Patterns of Fault Behaviour within Large Telephony Networks 340
Dave Yearling (BTexact Technologies), David J. Hand (Imperial College)

Closed Set Based Discovery of Representative Association Rules 350
Marzena Kryszkiewicz (Warsaw University of Technology)

Intelligent Sensor Analysis and Actuator Control 360
Matthew Easley (Rockwell Scientific), Elizabeth Bradley (University of Colorado)

Sampling of Highly Correlated Data for Polynomial Regression and Model Discovery 370
Grace W. Rumantir (Monash University), Chris S. Wallace (Monash University)

The IDA'01 Robot Data Challenge

The IDA'01 Robot Data Challenge 378
Paul Cohen (University of Massachusetts), Niall Adams (Imperial College), David J. Hand (Imperial College)

Author Index 383

Advances in Intelligent Data Analysis
4th International Conference, IDA 2001, Cascais,
Portugal, September 13-15, 2001. Proceedings
Hoffmann, F.; Hand, D.; Adams, N.M.; Fisher, D.;
Guimaraes, G. (Eds.)
2001, XII, 388 p., Softcover
ISBN: 978-3-540-42581-6