

Contents

Part I. Clustering and Discrimination

Robust Classification Through the Forward Search	3
<i>Anthony Atkinson, Andrea Cerioli, Marco Riani</i>	
Fitting and Smoothing Properties of Length Constrained Smoothers Applied to Time Series	13
<i>Estela Bee Dagum, Alessandra Luati</i>	
Interaction Terms in Non-linear PLS via Additive Spline Transformation	22
<i>Jean-François Durand, Rosaria Lombardo</i>	
Discriminant Analysis With Categorical Variables: A Biplot Based Approach	30
<i>Sugnet Gardner, Niël le Roux</i>	
Efficient Density Clustering Using Basin Spanning Trees	39
<i>Sören Hader, Fred A. Hamprecht</i>	
Some Issues on Clustering of Functional Data	49
<i>Salvatore Ingrassia, Andrea Cerioli, Aldo Corbellini</i>	
Scalable Clustering of High Dimensional Data	57
<i>David Littau, Daniel Boley</i>	
Methods to Combine Classification Trees	65
<i>Rossella Miglio, Gabriele Soffritti</i>	
Core-Based Clustering Techniques	74
<i>Hans-Joachim Mucha, Hans-Georg Bartel, Jens Dolata</i>	
Combining Regression Trees and Radial Basis Function Networks in Longitudinal Data Modelling	83
<i>Marilena Pillati, Daniela G. Calò, Giuliano Galimberti</i>	
Implementing a new Method for Discriminant Analysis When Group Covariance Matrices are Nearly Singular	92
<i>Britta Pouwels, Winfried Theis, Christian Röver</i>	
Two Approaches for Discriminant Partial Least Squares	100
<i>Robert Sabatier, Myrtille Vivien, Pietro Amenta</i>	

Improving the Classification Performance of a Discriminant Rule by Dealing With Data Cases Having a Substantial Influence on Variable Selection	109
<i>Sarel Steel, Nelmarie Louw</i>	
Classification and Clustering of Vocal Performances	118
<i>Claus Weihs, Uwe Ligges, Jörg Güttner, Petra Hasse-Becker, Sonja Berghoff</i>	
Cluster- and Discriminant Analysis With Both Metric and Categorical Data	127
<i>Michael Wodny, Bernd Jäger, Karl-Ernst Biebler</i>	
<hr/>	
Part II. Data Analysis and Statistics	
<hr/>	
Multivariate Mixture Models Estimation: A Genetic Algorithm Approach.....	133
<i>Roberto Baragona, Francesco Battaglia</i>	
Two-Way Clustering for Contingency Tables: Maximizing a Dependence Measure	143
<i>Hans-Hermann Bock</i>	
Localised Mixtures of Experts for Mixture of Regressions	155
<i>Guillaume Bouchard</i>	
A Bayesian Approach for the Estimation of the Covariance Structure of Separable Spatio-Temporal Stochastic Processes .	165
<i>Silvia Bozza, Anthony O'Hagan</i>	
An Algorithm for the HICLAS-R Model.....	173
<i>Eva Ceulemans, Iven Van Mechelen</i>	
Some Applications of Time-Varying Coefficient Models to Count Data	182
<i>Monica Chiogna, Carlo Gaetan</i>	
On Ziggurats and Dendrograms.....	191
<i>Frank Critchley</i>	
A Dimensionality Reduction Method Based on Simple Linear Regressions	201
<i>Luigi D'Ambra, Pietro Amenta, Rosaria Lombardo</i>	
Evolutionary Strategies to Avoid Local Mimima in Multidimensional Scaling	209
<i>Stefan Etschberger, Andreas Hilbert</i>	

Principal Component Analysis of Boolean Symbolic Objects ..	218
<i>Paolo Giordani</i>	
Multivariate Analysis for Variables of Arbitrary Information Level	226
<i>Jürgen Hansohm</i>	
How Wrong Models Become Useful - and Correct Models Become Dangerous	235
<i>Christian Hennig</i>	
Visualizing Symbolic Data by Closed Shapes	244
<i>Antonio Irpino, Carlo Lauro, Rosanna Verde</i>	
LBGU-EM Algorithm for Mixture Density Estimation.....	252
<i>Bambang Heru Iswanto, Bernd Fritzsche</i>	
Change-Points in Bernoulli Trials With Dependence	261
<i>Joachim Krauth</i>	
Two-Mode Clustering Methods: Compare and Contrast	270
<i>Sabine Krolak-Schwerdt</i>	
Sensitivity of Graphical Modeling Against Contamination	279
<i>Sonja Kuhnt, Claudia Becker</i>	
New Graphical Symbolic Objects Representations in Parallel Coordinates	288
<i>Carlo N. Lauro, Francesco Palumbo, Alfonso Iodice D'Enza</i>	
A Hierarchical Classes Approach to Discriminant Analysis....	296
<i>Luigi Lombardi, Eva Ceulemans, Iven Van Mechelen</i>	
Prediction Optimal Data Analysis by Means of Stochastic Search	305
<i>Karsten Luebke, Claus Weihs</i>	
System Error Variance Tuning in State-Space Models	313
<i>Pietro Mantovan, Andrea Pastore</i>	
Corrado Gini and Multivariate Statistical Analysis: The (so far) Missing Link	321
<i>Paola Monari, Angela Montanari</i>	
On Correspondence Analysis of Incomplete Orderings	329
<i>Kalev Pärna</i>	
Some Statistical Issues in Microarray Data Analysis	337
<i>Stéphane Robin</i>	

Background of the Variability in Past Human Populations: Selected Methodological Issues	348
<i>Arkadiusz Soltysiak</i>	

Spatial Prediction With Space-Time Models	358
<i>Stefano F. Tonellato</i>	

**Part III. Data Mining, Information Processing,
and Automation**

Neural Budget Networks of Sensorial Data	369
<i>Massimo Aria, Ab Mooijaart, Roberta Siciliano</i>	

Evolutionary Model Selection in Bayesian Neural Networks ..	378
<i>Silvia Bozza, Pietro Mantovan, Rosa A. Schiavo</i>	

Extreme Datamining	387
<i>Valérie C. Chavez-Demoulin, Stephen A. Jarvis, Rafael Perera, Armin S.A. Roehrl, Stefan W. Schmiedl, M.P. Sondergaard</i>	

Information Filtering Based on Modal Symbolic Objects	395
<i>Francisco de A. T. de Carvalho, Byron L. D. Bezerra</i>	

Analyzing Learner Behavior and Performance	405
<i>Karsten Friesen, Hans Schmitz</i>	

An Integration Strategy for Distributed Recommender Services in Legacy Library Systems	412
<i>Andreas Geyer-Schulz, Michael Hahsler, Andreas Neumann, Anke Thede</i>	

Comparing Simple Association-Rules and Repeat-Buying Based Recommender Systems in a B2B Environment	421
<i>Andreas Geyer-Schulz, Michael Hahsler, Anke Thede</i>	

Students' Preferences Related to Web Based E-Learning: Results of a Survey	430
<i>Marc Göcks, Daniel Baier</i>	

MURBANDY: The (so far) Missing Link: User-Friendly Retrieval and Visualization of Geographic Information	438
<i>Bernd Hermes, Maximilian Stempfhuber, Luca Demicheli, Carlo Lavalle</i>	

ASCAID: Using an Asymmetric Correlation Measure for Automatic Interaction Detection	447
<i>Andreas Hilbert</i>	

Discriminative Clustering: Vector Quantization in Learning Metrics	456
<i>Samuel Kaski, Janne Sinkkonen</i>	
Neural Network Hybrid Learning: Genetic Algorithms & Levenberg-Marquardt	464
<i>Ricardo B. C. Prudêncio, Teresa B. Ludermir</i>	
Syntagmatic and Paradigmatic Associations in Information Retrieval	473
<i>Reinhard Rapp</i>	
Finding the Most Useful Clusters: Clustering and the Usefulness Metric	483
<i>Caroline St. Clair</i>	
Internet Thesaurus - Extracting Relevant Terms from WWW Pages using Weighted Threshold Function over a Cross-Reference Matrix	492
<i>Raz Tamir</i>	
Reflections on a Supervised Approach to Independent Component Analysis	501
<i>Cinzia Viroli</i>	
Simulated Annealing and Tabu Search for Optimization of Neural Networks	510
<i>Akio Yamazaki, Teresa B. Ludermir, Marcilio C. P. de Souto</i>	
<hr/>	
Part IV. Finance, Marketing, and Management Science	
<hr/>	
Decentralizing Risk Management in the Case of Quadratic Hedging	521
<i>Julia Bondarenko, Nicole Branger, Angelika Esser, Christian Schlag</i>	
Multimedia Stimulus Presentation Methods for Conjoint Studies in Marketing Research	530
<i>Michael Brusch, Daniel Baier</i>	
Forecasting the Customer Development of a Publishing Company with Decision Trees	538
<i>Andreas Hilbert, Alexander Spatz</i>	
Estimation of Default Probabilities in a Single-Factor Model ..	546
<i>Steffi Höse, Stefan Huschens</i>	

Simultaneous Confidence Intervals for Default Probabilities...	555
<i>Steffi Höse, Stefan Huschens</i>	
Model-Based Clustering With Hidden Markov Models and its Application to Financial Time-Series Data	561
<i>Bernhard Knab, Alexander Schliep, Barthel Steckemetz, Bernd Wichern</i>	
Classification of Multivariate Data With Missing Values Using Expected Discriminant Scores	570
<i>Wolfgang Kossa</i>	
Assessment of the Polish Manufacturing Sector Attractiveness: An End-User Approach	578
<i>Dorota Kwiatkowska-Ciotucha, Józef Dziechciarz</i>	
Developing a Layout of a Supermarket Through Asymmetric Multidimensional Scaling and Cluster Analysis of Purchase Data	587
<i>Akinori Okada, Tadashi Imaizumi</i>	
Market Segmentation Method From the Bayesian Viewpoint .	595
<i>Kazuo Shigemasu, Takuya Ohmori, Takahiro Hoshino</i>	
Support Vector Machines for Credit Scoring: Comparing to and Combining With Some Traditional Classification Methods	604
<i>Ralf Stecking, Klaus B. Schebesch</i>	
<hr/>	
Part V. Biology, Archaeology, and Medicine	
<hr/>	
Classification of Amino-Acid Sequences Using State-Space Models	615
<i>Marcus Brunnert, Tillmann Krahnke, Wolfgang Urfer</i>	
Quantitative Study of Images in Archaeology: I. Textual Coding	624
<i>Sergio Camiz, Elena Rova</i>	
Quantitative Study of Images in Archaeology: II. Symbolic Coding	633
<i>Sergio Camiz, Elena Rova, Vanda Tulli</i>	
Power Functions in Multiple Sampling Plans Using DNA Computation	643
<i>Carles Capdevila, M. Angels Colomer, Josep Conde, Josep Miret, Alba Zaragoza</i>	

Methodological Issues in the Baffling Relationship Between Hepatitis C Virus and Non Hodgkin's Lymphomas ..	652
<i>Stefano De Cantis, Daria Mendola, Emilio Iannitto</i>	
Class Discovery in Gene Expression Data: Characterizing Splits by Support Vector Machines	662
<i>Florian Markowetz, Anja von Heydebreck</i>	
Efficient Clustering Methods for Tumor Classification With Microarrays	670
<i>Jörg Rahnenführer</i>	
A Method to Classify Microarray Data	680
<i>Marco Scarnó, Donatella Sforzini, Alessandra Ulivieri, Sergio Nasi</i>	
Index	689

Between Data Science and Applied Data Analysis
Proceedings of the 26th Annual Conference of the
Gesellschaft für Klassifikation e.V., University of
Mannheim, July 22-24, 2002
Schader, M.; Gaul, W.; Vichi, M. (Eds.)
2003, XIV, 694 p. 70 illus., Softcover
ISBN: 978-3-540-40354-8