

Studies in Classification, Data Analysis, and Knowledge Organization

Managing Editors

H.-H. Bock, Aachen
W. Gaul, Karlsruhe
M. Vichi, Rome

Editorial Board

Ph. Arabie, Newark
D. Baier, Cottbus
F. Critchley, Milton Keynes
R. Decker, Bielefeld
E. Diday, Paris
M. Greenacre, Barcelona
C. Lauro, Naples
J. Meulman, Leiden
P. Monari, Bologna
S. Nishisato, Toronto
N. Ohsumi, Tokyo
O. Opitz, Augsburg
G. Ritter, Passau
M. Schader, Mannheim
C. Weihs, Dortmund

Titles in the Series

E. Diday, Y. Lechevallier, M. Schader, P. Bertrand, and B. Burtschy (Eds.)
New Approaches in Classification and Data Analysis. 1994 (out of print)

W. Gaul and D. Pfeifer (Eds.)
From Data to Knowledge. 1995

H.-H. Bock and W. Polasek (Eds.)
Data Analysis and Information Systems. 1996

E. Diday, Y. Lechevallier, and O. Opitz (Eds.)
Ordinal and Symbolic Data Analysis. 1996

R. Klar and O. Opitz (Eds.)
Classification and Knowledge Organization. 1997

C. Hayashi, N. Ohsumi, K. Yajima, Y. Tanaka, H.-H. Bock, and Y. Baba (Eds.)
Data Science, Classification, and Related Methods. 1998

I. Balderjahn, R. Mathar, and M. Schader (Eds.)
Classification, Data Analysis, and Data Highways. 1998

A. Rizzi, M. Vichi, and H.-H. Bock (Eds.)
Advances in Data Science and Classification. 1998

M. Vichi and O. Opitz (Eds.)
Classification and Data Analysis. 1999

W. Gaul and H. Locarek-Junge (Eds.)
Classification in the Information Age. 1999

H.-H. Bock and E. Diday (Eds.)
Analysis of Symbolic Data. 2000

H. A. L. Kiers, J.-P. Rasson, P. J. F. Groenen, and M. Schader (Eds.)
Data Analysis, Classification, and Related Methods. 2000

W. Gaul, O. Opitz, and M. Schader (Eds.)
Data Analysis. 2000

R. Decker and W. Gaul (Eds.)
Classification and Information Processing at the Turn of the Millenium. 2000

S. Borra, R. Rocci, M. Vichi, and M. Schader (Eds.)
Advances in Classification and Data Analysis. 2001

W. Gaul and G. Ritter (Eds.)
Classification, Automation, and New Media. 2002

K. Jajuga, A. Sokołowski, and H.-H. Bock (Eds.)
Classification, Clustering and Data Analysis. 2002

M. Schwaiger and O. Opitz (Eds.)
Exploratory Data Analysis in Empirical Research. 2003

M. Schader, W. Gaul, and M. Vichi (Eds.)
Between Data Science and Applied Data Analysis. 2003

H.-H. Bock, M. Chiodi, and A. Mineo (Eds.)
Advances in Multivariate Data Analysis. 2004

D. Banks, L. House, F.R. McMorris, P. Arabie, and W. Gaul (Eds.)
Classification, Clustering, and Data Mining Applications. 2004

D. Baier and K.-D. Wernecke (Eds.)
Innovations in Classification, Data Science, and Information Systems. 2005

M. Vichi, P. Monari, S. Mignani and A. Montanari (Eds.)
New Developments in Classification and Data Analysis. 2005

D. Baier, R. Decker, and L. Schmidt-Thieme (Eds.)
Data Analysis and Decision Support. 2005

C. Weihs and W. Gaul (Eds.)
Classification – the Ubiquitous Challenge. 2005

Myra Spiliopoulou · Rudolf Kruse
Christian Borgelt · Andreas Nürnberger
Wolfgang Gaul
Editors

From Data and Information Analysis to Knowledge Engineering

Proceedings of the 29th Annual Conference
of the Gesellschaft für Klassifikation e.V.
University of Magdeburg, March 9–11, 2005

With 239 Figures and 120 Tables

Professor Dr. Myra Spiliopoulou
Otto-von-Guericke-Universität
Magdeburg
Institut für Technische und
Betriebliche Informationssysteme
Universitätsplatz 2
39106 Magdeburg
Germany
myra@iti.cs.uni-magdeburg.de

Professor Dr. Wolfgang Gaul
Universität Karlsruhe (TH)
Institut für Entscheidungstheorie
und Unternehmensforschung
76128 Karlsruhe
wolfgang.gaul@wiwi.uni-karlsruhe.de

Professor Dr. Rudolf Kruse
Dr. Christian Borgelt
Jun.-Professor Dr. Andreas
Nürnberger
Otto-von-Guericke-Universität
Magdeburg
Institut für Wissens-
und Sprachverarbeitung
Universitätsplatz 2
39106 Magdeburg
Germany
kruse@iws.cs.uni-magdeburg.de
borgelt@iws.cs.uni-magdeburg.de
nuernb@iws.cs.uni-magdeburg.de

ISSN 1431-8814

ISBN-10 3-540-31313-3 Springer-Verlag Berlin Heidelberg New York

ISBN-13 978-3-540-31313-7 Springer-Verlag Berlin Heidelberg New York

Cataloging-in-Publication Data

Library of Congress Control Number: 2005938846

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag. Violations are liable for prosecution under the German Copyright Law.

Springer · Part of Springer Science+Business Media

springeronline.com

© Springer-Verlag Berlin · Heidelberg 2006

Printed in Germany

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Softcover-Design: Erich Kirchner, Heidelberg

SPIN 11584247 43/3153 – 5 4 3 2 1 0 – Printed on acid-free paper

Preface

This volume contains revised versions of selected papers presented during the 29th Annual Conference of the German Classification Society (Gesellschaft für Klassifikation, GfKI'2005). The conference was held at the Otto-von-Guericke-University Magdeburg in March 2005. The theme of the GfKI'2005 was "From Data and Information Analysis to Knowledge Engineering" and encompassed 230 presentations in 74 sessions, including 11 plenary and semi-plenary talks. With 324 attendants from 23 countries, the 29th GfKI conference established a new participation record for the conference series. The conference again provided an attractive interdisciplinary forum for discussions and mutual exchange of knowledge. It was organized in cooperation with the Slovenian Artificial Intelligence Society (SLAIS).

The conference was accompanied by several colocated events. In addition to the Librarians Workshop and the traditional meetings of the working groups, a new important event took place for the first time — the Doctoral Workshop for PhD students. Starting at the GfKI'2004, a Data Mining Competition took place for the second time; for the particularly challenging data analysis problem posed this year 40 solutions were submitted.

The papers in this volume were selected in a second reviewing process after the conference. Each of the 131 submitted long versions of conference contributions was reviewed by two reviewers, and 92 were accepted for this volume. In addition to papers in the fundamental areas Classification, Clustering, and Data Analysis, this volume contains many papers on a wide range of topics with a strong relation to Computer Science. Examples are Text Mining (largest track of the conference as well as in this post-conference volume), Web Mining, Fuzzy Data Analysis, IT Security, Adaptivity and Personalization, and Visualization. Application-oriented topics were addressed in several conference talks. In this volume, the corresponding papers are grouped into the clusters: (1) Economics, Marketing, Banking and Finance, (2) Medicine, Bioinformatics, Biostatistics, (3) Music Analysis. The last paper in this volume reports on the solutions of the winning data mining contestants.

The editors of these proceedings would like to thank the members of the program committee, all reviewers for their vigorous and timely reviewing process, and the authors for their contributions. Special thanks go to the area chairs, who have undertaken the coordination of the reviewing process for their individual tracks and worked under a rigorous time schedule.

The success of the GfKI'2005 conference is due to the effort and involvement of many people. We would like to thank foremostly the local organization team of Silke Reifgerste, Marko Brunzel, Dirk Dreschel, Tanja Falkowski, Folker Folkens, Henner Graubitz, Roland Müller and Rene Schult and their student support team for their hard work in the preparation of this conference and for their support during the event itself. Most cordial thanks go to the organizers of the colocated events: Werner Esswein (TU Dresden) for the organization of the Doctoral Workshop, Hans-J. Hermes (TU Chemnitz) and

Bernd Lorenz (FH München) who organized the Librarians Workshop, Christian Klein (SPSS GmbH Software) and Michael Thess (prudsys AG) for their involvement in the organization of the industrial track and to Jens Strackeljan (Otto-von-Guericke-University Magdeburg) as well as Roland Jonscher and Sigurd Prieur (Sparkassen Rating und Risikosysteme GmbH, Berlin) for the coordination of the Data Mining Competition. The awards for the competition were sponsored by the Deutscher Sparkassen- und Giroverband.

Institutional support has been of paramount importance for the success of the GfKI'2005. Our first thanks go to the Faculty of Computer Science and the Otto-von-Guericke-Universität Magdeburg for providing rooms, facilities, support and assistance to the organization of this conference. We are particularly indebted to the University Rector Klaus Erich Pollmann for his support and involvement. We gratefully acknowledge the support of the city of Magdeburg in organizing the city reception event. In addition, we would like to thank DaimlerChrysler AG and our sponsors Deutscher Sparkassen- und Giroverband, Heins+Partner GmbH, prudsys AG, Springer Verlag GmbH and SPSS GmbH Software for their support. Finally, we would like to thank Christiane Beisel and Martina Bihn of Springer-Verlag, Heidelberg, for their support and dedication to the production of this volume.

The German Classification Society entrusted us with the organization of the GfKI'2005. We are grateful for this honor and for all institutional and personal support provided to us in all phases of the GfKI'2005, from the first planning phase until the print of this volume.

Myra Spiliopoulou,
Rudolf Kruse,
Christian Borgelt,
Andreas Nürnberger,
Wolfgang Gaul

Magdeburg and Karlsruhe,
January 2006

Organization

Chairs

Local Chair

Myra Spiliopoulou (Otto-von-Guericke-University Magdeburg, Germany)

Publication Chair

Rudolf Kruse (Otto-von-Guericke-University Magdeburg, Germany)

Publicity Chair

Andreas Nürnberger (Otto-von-Guericke-University Magdeburg, Germany)

Submission and Book Preparation

Christian Borgelt (Otto-von-Guericke-University Magdeburg, Germany)

Program Chair

Wolfgang Gaul (University of Karlsruhe, Germany)

Program Committee

Hans-Hermann Bock (RWTH Aachen, Germany)

Reinhold Decker (University of Bielefeld, Germany)

Bernard Fichet (University of Aix-Marseille II, France)

Wolfgang Gaul (University of Karlsruhe, Germany)

Rudolf Kruse (Otto-von-Guericke-University Magdeburg, Germany)

Hans-Joachim Lenz (Free University of Berlin, Germany)

Dunja Mladenić (J. Stefan Institute, Slovenia)

Otto Opitz (University of Augsburg, Germany)

Myra Spiliopoulou (Otto-von-Guericke-University Magdeburg, Germany)

Maurizio Vichi (University of Roma — “La Sapienza”, Italy)

Claus Weihs (University of Dortmund, Germany)

Klaus-Dieter Wernecke (Charité Berlin, Germany)

Program Sections and Area Chairs

Clustering

Hans-Hermann Bock (RWTH Aachen, Germany)

Discrimination

Gunter Ritter (University Passau, Germany)

Multiway Classification and Data Analysis

Sabine Krolak-Schwerdt (Saarland University, Germany)

Henk A.L. Kiers (University of Groningen, Netherlands)

Multimode Clustering and Dimensionality Reduction

Maurizio Vichi (University Roma — “La Sapienza”, Italy)

Robust Methods in Multivariate Statistics

Andrea Cerioli (University of Parma, Italy)

Dissimilarities and Clustering Structures

Bernard Fichet (University of Aix-Marseille II, France)

PLS Path Modeling, PLS Regression and Classification

Natale C. Lauro (University “Federico II” of Napoli, Italy)

V. Esposito Vinzi (University “Federico II” of Napoli, Italy)

Ranking, Multi-label Classification, Preferences

Johannes Fürnkranz (Technical University Darmstadt, Germany)

Eyke Hüllermeier (Philipps-University Marburg, Germany)

Computational Advances in Data Analysis

Hans-Joachim Lenz (Free University Berlin, Germany)

Fuzzy Data Analysis

Rudolf Kruse (Otto-von-Guericke-University Magdeburg, Germany)

Visualization

Patrick J.F. Groenen (Erasmus University Rotterdam, Netherlands)

Classification and Analysis in Data Intensive Scenarios

Gunter Saake (Otto-von-Guericke-University Magdeburg, Germany)

Data Mining and Explorative Multivariate Data Analysis

Luigi D’Ambrà (University “Federico II” of Napoli, Italy)

Paulo Giudici (University of Pavia, Italy)

Text Mining

Andreas Nürnberger (Otto-von-Guericke-University Magdeburg, Germany)

Dunja Mladenič (Jozef Stefan Institute Ljubljana, Slovenia)

Web Mining

Myra Spiliopoulou (Otto-von-Guericke-University Magdeburg, Germany)

Adaptivity and Personalization

Andreas Geyer-Schulz (University Karlsruhe, Germany)

Lars Schmidt-Thieme (Albert-Ludwigs-University Freiburg, Germany)

User and Data Authentication in IT Security

Jana Dittmann (Otto-von-Guericke-University Magdeburg, Germany)

Banking and Finance

Hermann Locarek-Junge (Technical University Dresden, Germany)

Marketing

Daniel Baier (Brandenburg University of Technology Cottbus, Germany)
 Matthias Meyer (Ludwig-Maximilians-University Mnchen, Germany)

Economics

Otto Opitz (University Augsburg, Germany)

Mining in Business Processes

Claus Rautenstrauch (Otto-von-Guericke-University Magdeburg, Germany)

Bioinformatics and Biostatistics

Berthold Lausen

(Friedrich-Alexander University Erlangen-Nuremberg, Germany)

Classification of High-dimensional Biological and Medical Data

Siegfried Kropf (Otto-von-Guericke-University Magdeburg, Germany)

Johannes Bernarding (Otto-von-Guericke-University Magdeburg, Germany)

Classification with Latent Variable Models

Angela Montanari (University Bologna, Italy)

Medical and Health Sciences

Klaus-Dieter Wernecke (Charité Berlin, Germany)

Music Analysis

Claus Weihs (University Dortmund, Germany)

Industrial Applications and Solutions

Myra Spiliopoulou (Otto-von-Guericke-University Magdeburg, Germany)

Additional Reviewers (in alphabetical order)

Mark Ackermans	Enrico Hauer	Paola Monari
Sven Apel	Hartmut Hecker	Fabian Mörchen
Michael Berthold	Christian Hennig	Hans-Joachim Mucha
Eva Ceulemans	Andreas Hilbert	Daniel Müllensiefen
Steffen Bickel	Andreas Hotho	Gerhard Paaß
Ulf Brefeld	Frank Klawonn	Marco Riani
Christian Döring	Juergen Kleffe	Gunter Ritter
Daniel Enache	Meike Klettke	Fabrice Rossi
Tanja Falkowski	Peter Kuhbier	Kai-Uwe Sattler
María Teresa Gallegos	Andreas Lang	Eike Schallehn
Michael Gertz	Berthold Lausen	Ingo Schmitt
Hans Goebel	Wolfgang Lehner	Benno Stein
Gerard Govaert	Wolfgang May	Gerd Stumme
Peter Grzybek	Iven Van Mechelen	Michiel van Wezel
Larry Hall	Alexander Mehler	Adalbert Wilhelm
Fred A. Hamprecht		

Contents

Plenaries and Semi-plenaries

Boosting and ℓ^1 -Penalty Methods for High-dimensional Data with Some Applications in Genomics	1
<i>P. Bühlmann</i>	
Striving for an Adequate Vocabulary: Next Generation 'Metadata'	13
<i>D. Fellner and S. Havemann</i>	
Scalable Swarm Based Fuzzy Clustering	21
<i>L.O. Hall and P.M. Kanade</i>	
SolEuNet: Selected Data Mining Techniques and Applications	32
<i>N. Lavrač</i>	
Inferred Causation Theory: Time for a Paradigm Shift in Marketing Science?	40
<i>J.A. Mazanec</i>	
Text Mining in Action	52
<i>D. Mladenič</i>	
Identification of Real-world Objects in Multiple Databases	63
<i>M. Neiling</i>	
Kernels for Predictive Graph Mining	75
<i>S. Wrobel, T. Gärtner, and T. Horváth</i>	

Clustering

PRISMA: Improving Risk Estimation with Parallel Logistic Regression Trees	87
<i>B. Arnrich, A. Albert, and J. Walter</i>	
Latent Class Analysis and Model Selection	95
<i>J.G. Dias</i>	

An Indicator for the Number of Clusters: Using a Linear Map to Simplex Structure	103
<i>M. Weber, W. Rungsarityotin, and A. Schliep</i>	

Discriminant Analysis

On the Use of Some Classification Quality Measure to Construct Mean Value Estimates Under Nonresponse	111
<i>W. Gamrot</i>	

A Wrapper Feature Selection Method for Combined Tree-based Classifiers	119
<i>E. Gatnar</i>	

Input Variable Selection in Kernel Fisher Discriminant Analysis	126
<i>N. Louw and S.J. Steel</i>	

The Wavelet Packet Based Cepstral Features for Open Set Speaker Classification in Marathi	134
<i>H.A. Patil, P.K. Dutta, and T.K. Basu</i>	

A New Effective Algorithm for Stepwise Principle Components Selection in Discriminant Analysis	142
<i>E. Serikova and E. Zhuk</i>	

A Comparison of Validation Methods for Learning Vector Quantization and for Support Vector Machines on Two Biomedical Data Sets	150
<i>D. Sommer and M. Golz</i>	

Discriminant Analysis of Polythetically Described Older Palaeolithic Stone Flakes: Possibilities and Questions	158
<i>T. Weber</i>	

Classification with Latent Variable Models

Model-based Density Estimation by Independent Factor Analysis	166
<i>D.G. Calò, A. Montanari, and C. Viroli</i>	

Identifying Multiple Cluster Structures Through Latent Class Models	174
<i>G. Galimberti and G. Soffritti</i>	

Gene Selection in Classification Problems via Projections onto a Latent Space	182
<i>M. Pillati and C. Viroli</i>	

Multiway Classification and Data Analysis

The Recovery Performance of Two-mode Clustering Methods: Monte Carlo Experiment	190
<i>S. Krolak-Schwerdt and M. Wiedenbeck</i>	

On the Comparability of Reliability Measures: Bifurcation Analysis of Two Measures in the Case of Dichotomous Ratings	198
<i>T. Ostermann and R. Schuster</i>	

Ranking, Multi-label Classification, Preferences

On Active Learning in Multi-label Classification	206
<i>K. Brinker</i>	

From Ranking to Classification: A Statistical View	214
<i>S. Cl��men��on, G. Lugosi, and N. Vayatis</i>	

PLS Path Modeling, PLS Regression and Classification

Assessing Unidimensionality within PLS Path Modeling Framework . . .	222
<i>K. Sahmer, M. Hanafi, and E.M. Qannari</i>	

The Partial Robust M-approach	230
<i>S. Serneels, C. Croux, P. Filzmoser, and P.J. Van Espen</i>	

Classification in PLS Path Models and Local Model Optimisation	238
<i>S. Squillacciotti</i>	

Robust Methods in Multivariate Statistics

Hierarchical Clustering by Means of Model Grouping	246
<i>C. Agostinelli and P. Pellizzari</i>	

Deepest Points and Least Deep Points: Robustness and Outliers with MZE	254
<i>C. Becker and S.P. Scholz</i>	

Robust Transformations and Outlier Detection with Autocorrelated Data	262
<i>A. Cerioli and M. Riani</i>	

Robust Multivariate Methods: The Projection Pursuit Approach	270
<i>P. Filzmoser, S. Serneels, C. Croux, and P.J. Van Espen</i>	

Finding Persisting States for Knowledge Discovery in Time Series	278
<i>F. Mörchen and A. Ultsch</i>	

Data Mining and Explorative Multivariate Data Analysis

Restricted Co-inertia Analysis	286
<i>P. Amenta and E. Ciavolino</i>	

Hausman Principal Component Analysis	294
<i>V. Choulakian, L. Dambra, and B. Simonetti</i>	

Nonlinear Time Series Modelling: Monitoring a Drilling Process	302
<i>A. Messaoud, C. Weihs, and F. Hering</i>	

Text Mining

Word Length and Frequency Distributions in Different Text Genres . . .	310
<i>G. Antić, E. Stadlober, P. Grzybek, and E. Kelih</i>	

Bootstrapping an Unsupervised Morphemic Analysis	318
<i>C. Benden</i>	

Automatic Extension of Feature-based Semantic Lexicons via Contextual Attributes	326
<i>C. Biemann and R. Osswald</i>	

Learning Ontologies to Improve Text Clustering and Classification . . .	334
<i>S. Bloehdorn, P. Cimiano, and A. Hotho</i>	

Discovering Communities in Linked Data by Multi-view Clustering . . .	342
<i>I. Drost, S. Bickel, and T. Scheffer</i>	

Crosslinguistic Computation and a Rhythm-based Classification of Languages	350
<i>A. Fenk and G. Fenk-Oczlon</i>	

Using String Kernels for Classification of Slovenian Web Documents . . .	358
<i>B. Fortuna and D. Mladenič</i>	

Semantic Decomposition of Character Encodings for Linguistic Knowledge Discovery	366
<i>D. Gibbon, B. Hughes, and T. Trippel</i>	
Applying Collaborative Filtering to Real-life Corporate Data	374
<i>M. Grcar, D. Mladenič, and M. Grobelnik</i>	
Quantitative Text Typology: The Impact of Sentence Length	382
<i>E. Kelih, P. Grzybek, G. Antić, and E. Stadlober</i>	
A Hybrid Machine Learning Approach for Information Extraction from Free Text	390
<i>G. Neumann</i>	
Text Classification with Active Learning	398
<i>B. Novak, D. Mladenič, and M. Grobelnik</i>	
Towards Structure-sensitive Hypertext Categorization	406
<i>A. Mehler, R. Gleim, and M. Dehmer</i>	
Evaluating the Performance of Text Mining Systems on Real-world Press Archives	414
<i>G. Paaß and H. de Vries</i>	
Part-of-Speech Induction by Singular Value Decomposition and Hierarchical Clustering	422
<i>R. Rapp</i>	
Near Similarity Search and Plagiarism Analysis	430
<i>B. Stein and S.M. zu Eissen</i>	

Fuzzy Data Analysis

Objective Function-based Discretization	438
<i>F. Höppner</i>	
Understanding and Controlling the Membership Degrees in Fuzzy Clustering	446
<i>F. Klawonn</i>	
Autonomous Sensor-based Landing Systems: Fusion of Vague and Incomplete Information by Application of Fuzzy Clustering Techniques	454
<i>B. Korn</i>	
Outlier Preserving Clustering for Structured Data Through Kernels	462
<i>M.-J. Lesot</i>	

Economics and Mining in Business Processes

Classification-relevant Importance Measures for the West German Business Cycle	470
<i>D. Enache, C. Weihs, and U. Garczarek</i>	
The Classification of Local and Branch Labour Markets in the Upper Silesia	478
<i>W. Hantke</i>	
An Overview of Artificial Life Approaches for Clustering	486
<i>D. Kämpf and A. Ultsch</i>	
Design Problems of Complex Economic Experiments	494
<i>J. Kunze</i>	
Traffic Sensitivity of Long-term Regional Growth Forecasts	502
<i>W. Polasek and H. Berrer</i>	
Spiralling in BTA Deep-hole Drilling: Models of Varying Frequencies	510
<i>N. Raabe, O. Webber, W. Theis, and C. Weihs</i>	
Analysis of the Economic Development of Districts in Poland as a Basis for the Framing of Regional Policies	518
<i>M. Rozkrut and D. Rozkrut</i>	

Banking and Finance

The Classification of Candlestick Charts: Laying the Foundation for Further Empirical Research	526
<i>S. Etschberger, H. Fock, C. Klein, and B. Zwergel</i>	
Modeling and Estimating the Credit Cycle by a Probit-AR(1)-Process	534
<i>S. Höse and K. Vogl</i>	
Comparing and Selecting SVM-Kernels for Credit Scoring	542
<i>R. Stecking and K.B. Schebesch</i>	
Value at Risk Using the Principal Components Analysis on the Polish Power Exchange	550
<i>G. Trzpiot and A. Ganczarek</i>	

Marketing

A Market Basket Analysis Conducted with a Multivariate Logit Model	558
<i>Y. Boztuğ and L. Hildebrandt</i>	
Solving and Interpreting Binary Classification Problems in Marketing with SVMs	566
<i>G. Nalbantov, J.C. Bioch, and P.J.F. Groenen</i>	
Modeling the Nonlinear Relationship Between Satisfaction and Loyalty with Structural Equation Models	574
<i>M. Paulssen and A. Sommerfeld</i>	
Job Choice Model to Measure Behavior in a Multi-stage Decision Process	582
<i>T. Spengler and J. Malmendier</i>	
Semiparametric Stepwise Regression to Estimate Sales Promotion Effects	590
<i>W.J. Steiner, C. Belitz, and S. Lang</i>	

Adaptivity and Personalization

Implications of Probabilistic Data Modeling for Mining Association Rules	598
<i>M. Hahsler, K. Hornik, and T. Reutterer</i>	
Copula Functions in Model Based Clustering	606
<i>K. Jajuga and D. Papla</i>	
Attribute-aware Collaborative Filtering	614
<i>K. Tso and L. Schmidt-Thieme</i>	

User and Data Authentication in IT Security

Towards a Flexible Framework for Open Source Software for Handwritten Signature Analysis	622
<i>R. Guest, M. Fairhurst, and C. Vielhauer</i>	
Multimodal Biometric Authentication System Based on Hand Features	630
<i>N. Pavešić, T. Savič, and S. Ribarić</i>	

Labelling and Authentication for Medical Imaging Through Data Hiding	638
<i>A. De Rosa, R. Caldelli, and A. Piva</i>	

Hand-geometry Recognition Based on Contour Landmarks	646
<i>R. Veldhuis, A. Bazen, W. Booi, and A. Hendrikse</i>	

A Cross-cultural Evaluation Framework for Behavioral Biometric User Authentication.....	654
<i>F. Wolf, T.K. Basu, P.K. Dutta, C. Vielhauer, A. Oermann, and B. Yegnanarayana</i>	

Bioinformatics and Biostatistics

On External Indices for Mixtures: Validating Mixtures of Genes.....	662
<i>I.G. Costa and A. Schliep</i>	

Tests for Multiple Change Points in Binary Markov Sequences	670
<i>J. Krauth</i>	

UnitExpressions: A Rational Normalization Scheme for DNA Microarray Data.....	678
<i>A. Ultsch</i>	

Classification of High-dimensional Biological and Medical Data

A Ridge Classification Method for High-dimensional Observations	684
<i>M. Grüning and S. Kropf</i>	

Assessing the Trustworthiness of Clustering Solutions Obtained by a Function Optimization Scheme	692
<i>U. Möller and D. Radke</i>	

Variable Selection for Discrimination of More Than Two Classes Where Data are Sparse	700
<i>G. Szepannek and C. Weihs</i>	

Medical and Health Sciences

The Assessment of Second Primary Cancers (SPCs) in a Series of Splenic Marginal Zone Lymphoma (SMZL) Patients	708
<i>S. De Cantis and A.M. Taormina</i>	

Heart Rate Classification Using Support Vector Machines	716
<i>M. Vogt, U. Moissl, and J. Schaab</i>	

Music Analysis

Visual Mining in Music Collections	724
<i>F. Mörchén, A. Ultsch, M. Nöcker, and C. Stamm</i>	

Modeling Memory for Melodies	732
<i>D. Müllensiefen and C. Hennig</i>	

Parameter Optimization in Automatic Transcription of Music	740
<i>C. Weihs and U. Ligges</i>	

Data Mining Competition

GfKl Data Mining Competition 2005:	
Predicting Liquidity Crises of Companies	748
<i>J. Strackeljan, R. Jonscher, S. Prieur, D. Vogel, T. Deselaers,</i>	
<i>D. Keysers, A. Mauser, I. Bezrukov, and A. Hegerath</i>	
Author Index	759