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. Sokolowski, and H.-H. Bock (Eds.)
Classification, Clustering and Data Analysis. 2002

M. Schweiger 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
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
Preface

This volume contains revised versions of selected papers presented during the 29th Annual Conference of the German Classification Society (Gesellschaft für Klassifikation, GfKl’2005). The conference was held at the Otto-von-Guericke-University Magdeburg in March 2005. The theme of the GfKl’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 GfKl 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 collocated 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 GfKl’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 GfKl’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 collocated 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 GfKl’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 GfKl’2005. We are grateful for this honor and for all institutional and personal support provided to us in all phases of the GfKl’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 N"urnberger (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 Mladenčić (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’Ambra (University “Federico II” of Napoli, Italy)
Paulo Giudici (University of Pavia, Italy)

Text Mining
Andreas Nünberger (Otto-von-Guericke-University Magdeburg, Germany)
Dunja Mladenič (Jožef 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 Muenchen, 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 (Charite 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
Sven Apel
Michael Berthold
Eva Ceulemans
Steffen Bickel
Ulf Brefeld
Christian Doring
Daniel Enache
Tanja Falkowski
Maria Teresa Gallegos
Michael Gertz
Hans Goebl
Gerard Govaert
Peter Grzybek
Larry Hall
Fred A. Hamprecht
Enrico Hauer
Hartmut Hecker
Christian Hennig
Andreas Hilbert
Andreas Hotho
Frank Klawonn
Juergen Kleffe
Meike Klettke
Peter Kuhbier
Andreas Lang
Berthold Lausen
Wolfgang Lehner
Wolfgang May
Iven Van Mechelen
Alexander Mehler
Paola Monari
Fabian Morchen
Hans-Joachim Mucha
Daniel Mullensiefen
Gerhard Paas
Marco Riani
Gunter Ritter
Fabrice Rossi
Kai-Uwe Sattler
Eike Schallehn
Ingo Schmitt
Benno Stein
Gerd Stumme
Michiel van Wezel
Adalbert Wilhelm
## Contents

### Plenaries and Semi-plenaries

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

### Clustering

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRISMA: Improving Risk Estimation</td>
<td>87</td>
</tr>
<tr>
<td>with Parallel Logistic Regression Trees</td>
<td></td>
</tr>
<tr>
<td><em>B. Arnrich, A. Albert, and J. Walter</em></td>
<td></td>
</tr>
<tr>
<td>Latent Class Analysis and Model Selection</td>
<td>95</td>
</tr>
<tr>
<td><em>J.G. Dias</em></td>
<td></td>
</tr>
</tbody>
</table>
An Indicator for the Number of Clusters:
Using a Linear Map to Simplex Structure ......................... 103
_M. Weber, W. Rungsarityotin, and A. Schliep_

**Discriminant Analysis**

On the Use of Some Classification Quality Measure
to Construct Mean Value Estimates Under Nonresponse ............ 111
_W. Gamrot_

A Wrapper Feature Selection Method
for Combined Tree-based Classifiers ............................. 119
_E. Gatnar_

Input Variable Selection
in Kernel Fisher Discriminant Analysis ........................... 126
_N. Louw and S.J. Steel_

The Wavelet Packet Based Cepstral Features
for Open Set Speaker Classification in Marathi .................. 134

A New Effective Algorithm for Stepwise
Principle Components Selection in Discriminant Analysis .......... 142
_E. Serikova and E. Zhuk_

A Comparison of Validation Methods
for Learning Vector Quantization and for Support Vector Machines
on Two Biomedical Data Sets ........................................ 150
_D. Sommer and M. Golz_

Discriminant Analysis of Polythetically Described
Older Palaeolithic Stone Flakes: Possibilities and Questions .... 158
_T. Weber_

**Classification with Latent Variable Models**

Model-based Density Estimation by Independent Factor Analysis ..... 166
_D.G. Calò, A. Montanari, and C. Viroli_

Identifying Multiple Cluster Structures
Through Latent Class Models ........................................ 174
_G. Galimberti and G. Soffritti_
Gene Selection in Classification Problems
via Projections onto a Latent Space .......................... 182
M. Pillati and C. Viroli

Multiway Classification and Data Analysis

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

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

Ranking, Multi-label Classification, Preferences

On Active Learning in Multi-label Classification ............... 206
K. Brinker

From Ranking to Classification: A Statistical View .......... 214
S. Clémençon, G. Lugosi, and N. Vayatis

PLS Path Modeling, PLS Regression and Classification

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

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

Classification in PLS Path Models and Local Model Optimisation . 238
S. Squillacciotti

Robust Methods in Multivariate Statistics

Hierarchical Clustering by Means of Model Grouping .......... 246
C. Agostinelli and P. Pellizzari

Deepest Points and Least Deep Points:
Robustness and Outliers with MZE ............................ 254
C. Becker and S.P. Scholz
Robust Transformations and Outlier Detection
with Autocorrelated Data ........................................ 262
A. Cerioli and M. Riani

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

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

Data Mining and Explorative Multivariate Data Analysis

Restricted Co-inertia Analysis ........................................ 286
P. Amenta and E. Ciavolino

Hausman Principal Component Analysis .......................... 294
V. Choulakian, L. Dambra, and B. Simonetti

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

Text Mining

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

Bootstrapping an Unsupervised Morphemic Analysis .................. 318
C. Benden

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

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

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

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

Using String Kernels for Classification of Slovenian Web Documents ... 358
B. Fortuna and D. Mladenič
Semantic Decomposition of Character Encodings for Linguistic Knowledge Discovery .................................................. 366
D. Gibbon, B. Hughes, and T. Trippel

Applying Collaborative Filtering to Real-life Corporate Data .... 374
M. Grcar, D. Mladenčič, and M. Grobelnik

Quantitative Text Typology: The Impact of Sentence Length ...... 382
E. Kelih, P. Grzybek, G. Antić, and E. Stadlober

A Hybrid Machine Learning Approach for Information Extraction from Free Text ................................................. 390
G. Neumann

Text Classification with Active Learning ........................................... 398
B. Novak, D. Mladenčič, and M. Grobelnik

Towards Structure-sensitive Hypertext Categorization .............. 406
A. Mehler, R. Gleim, and M. Dehmer

Evaluating the Performance of Text Mining Systems on Real-world Press Archives ............................................. 414
G. Paaß and H. de Vries

Part-of-Speech Induction by Singular Value Decomposition and Hierarchical Clustering ......................................... 422
R. Rapp

Near Similarity Search and Plagiarism Analysis ....................... 430
B. Stein and S.M. zu Eissen

Fuzzy Data Analysis

Objective Function-based Discretization ................................. 438
F. Höppner

Understanding and Controlling the Membership Degrees in Fuzzy Clustering .................................................. 446
F. Klawonn

Autonomous Sensor-based Landing Systems: Fusion of Vague and Incomplete Information by Application of Fuzzy Clustering Techniques 454
B. Korn

Outlier Preserving Clustering for Structured Data Through Kernels ... 462
M.-J. Lesot
Economics and Mining in Business Processes

Classification-relevant Importance Measures for the West German Business Cycle ........................................ 470
D. Enache, C. Weihs, and U. Garczarek

The Classification of Local and Branch Labour Markets in the Upper Silesia .................................................. 478
W. Hantke

An Overview of Artificial Life Approaches for Clustering .......................................................... 486
D. Kämpf and A. Ultsch

Design Problems of Complex Economic Experiments .................. 494
J. Kunze

Traffic Sensitivity of Long-term Regional Growth Forecasts............. 502
W. Polasek and H. Berrer

Spiralling in BTA Deep-hole Drilling: Models of Varying Frequencies ......................................................... 510
N. Raabe, O. Webber, W. Theis, and C. Weihs

Analysis of the Economic Development of Districts in Poland as a Basis for the Framing of Regional Policies ............ 518
M. Rozkrut and D. Rozkrut

Banking and Finance

The Classification of Candlestick Charts: Laying the Foundation for Further Empirical Research ................. 526
S. Etschberger, H. Fock, C. Klein, and B. Zwergel

Modeling and Estimating the Credit Cycle by a Probit-AR(1)-Process ......................................................... 534
S. Höse and K. Vogl

Comparing and Selecting SVM-Kernels for Credit Scoring ........ 542
R. Stecking and K.B. Schebesch

Value at Risk Using the Principal Components Analysis on the Polish Power Exchange .................................. 550
G. Trzpiot and A. Ganczarek
Marketing

A Market Basket Analysis Conducted with a Multivariate Logit Model .......................... 558
Y. Boztğu and L. Hildebrandt

Solving and Interpreting Binary Classification Problems in Marketing with SVMs .................................................... 566
G. Nalbantov, J.C. Bioch, and P.J.F. Groenen

Modeling the Nonlinear Relationship Between Satisfaction and Loyalty with Structural Equation Models ......................... 574
M. Paulssen and A. Sommerfeld

Job Choice Model to Measure Behavior in a Multi-stage Decision Process .......................... 582
T. Spengler and J. Malmendier

Semiparametric Stepwise Regression to Estimate Sales Promotion Effects ............................ 590
W.J. Steiner, C. Belitz, and S. Lang

Adaptivity and Personalization

Implications of Probabilistic Data Modeling for Mining Association Rules .......................... 598
M. Hahsler, K. Hornik, and T. Reutterer

Copula Functions in Model Based Clustering ................. 606
K. Jajuga and D. Papla

Attribute-aware Collaborative Filtering ......................... 614
K. Tso and L. Schmidt-Thieme

User and Data Authentication in IT Security

Towards a Flexible Framework for Open Source Software for Handwritten Signature Analysis .......................... 622
R. Guest, M. Fairhurst, and C. Vielhauer

Multimodal Biometric Authentication System Based on Hand Features .................. 630
N. Pavešić, T. Savič, and S. Ribarić
Labelling and Authentication for Medical Imaging Through Data Hiding ........................................... 638
A. De Rosa, R. Caldelli, and A. Piva

Hand-geometry Recognition Based on Contour Landmarks ........... 646
R. Veldhuis, A. Bazen, W. Booij, and A. Hendrikse

A Cross-cultural Evaluation Framework for Behavioral Biometric User Authentication .................. 654

Bioinformatics and Biostatistics

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

Tests for Multiple Change Points in Binary Markov Sequences ....... 670
J. Krauth

UnitExpressions: A Rational Normalization Scheme for DNA Microarray Data ........................................ 678
A. Ultsch

Classification of High-dimensional Biological and Medical Data

A Ridge Classification Method for High-dimensional Observations .... 684
M. Grünning and S. Kropf

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

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

Medical and Health Sciences

The Assessment of Second Primary Cancers (SPCs) in a Series of Splenic Marginal Zone Lymphoma (SMZL) Patients ...... 708
S. De Cantis and A.M. Taormina
Heart Rate Classification Using Support Vector Machines ............ 716
M. Vogt, U. Moissl, and J. Schaab

Music Analysis

Visual Mining in Music Collections ........................................ 724
F. Mörchen, A. Ultsch, M. Nöcker, and C. Stamm

Modeling Memory for Melodies ............................................ 732
D. Müllensiefen and C. Hennig

Parameter Optimization in Automatic Transcription of Music ........ 740
C. Weihs and U. Ligges

Data Mining Competition

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

Author Index ........................................................................... 759