

Contents

Part I. (Semi-) Plenary Presentations

Classification and Data Mining in Musicology	3
<i>Jan Beran</i>	
Bayesian Mixed Membership Models for Soft Clustering and Classification	11
<i>Elena A. Erosheva, Stephen E. Fienberg</i>	
Predicting Protein Secondary Structure with Markov Models. . . .	27
<i>Paul Fischer, Simon Larsen, Claus Thomsen</i>	
Milestones in the History of Data Visualization: A Case Study in Statistical Historiography	34
<i>Michael Friendly</i>	
Quantitative Text Typology: The Impact of Word Length	53
<i>Peter Grzybek, Ernst Stadlober, Emmerich Kelih, Gordana Antić</i>	
Cluster Ensembles	65
<i>Kurt Hornik</i>	
Bootstrap Confidence Intervals for Three-way Component Methods	73
<i>Henk A.L. Kiers</i>	
Organising the Knowledge Space for Software Components ...	85
<i>Claus Pahl</i>	
Multimedia Pattern Recognition in Soccer Video Using Time Intervals	97
<i>Cees G.M. Snoek, Marcel Worring</i>	
Quantitative Assessment of the Responsibility for the Disease Load in a Population	109
<i>Wolfgang Uter, Olaf Gefeller</i>	

Part II. Classification and Data Analysis

Classification

Bootstrapping Latent Class Models	121
<i>José G. Dias</i>	
Dimensionality of Random Subspaces	129
<i>Eugeniusz Gatnar</i>	
Two-stage Classification with Automatic Feature Selection for an Industrial Application	137
<i>Sören Hader, Fred A. Hamprecht</i>	
Bagging, Boosting and Ordinal Classification	145
<i>Klaus Hechenbichler, Gerhard Tutz</i>	
A Method for Visual Cluster Validation	153
<i>Christian Hennig</i>	
Empirical Comparison of Boosting Algorithms	161
<i>Riadh Khanchel, Mohamed Limam</i>	
Iterative Majorization Approach to the Distance-based Discriminant Analysis	168
<i>Serhiy Kosinov, Stéphane Marchand-Maillet, Thierry Pun</i>	
An Extension of the CHAID Tree-based Segmentation Algorithm to Multiple Dependent Variables	176
<i>Jay Magidson, Jeroen K. Vermunt</i>	
Expectation of Random Sets and the ‘Mean Values’ of Interval Data	184
<i>Ole Nordhoff</i>	
Experimental Design for Variable Selection in Data Bases	192
<i>Constanze Pumplün, Claus Weihs, Andrea Preusser</i>	
KMC/EDAM: A New Approach for the Visualization of K-Means Clustering Results	200
<i>Nils Raabe, Karsten Luebke, Claus Weihs</i>	

Clustering of Variables with Missing Data: Application to Preference Studies	208
<i>Karin Sahmer, Evelyne Vigneau, El Mostafa Qannari, Joachim Kunert</i>	
Binary On-line Classification Based on Temporally Integrated Information	216
<i>Christin Schäfer, Steven Lemm, Gabriel Curio</i>	
Different Subspace Classification	224
<i>Gero Szepannek, Karsten Luebke</i>	
Density Estimation and Visualization for Data Containing Clusters of Unknown Structure	232
<i>Alfred Ultsch</i>	
Hierarchical Mixture Models for Nested Data Structures	240
<i>Jeroen K. Vermunt, Jay Magidson</i>	
Data Analysis	
<hr/>	
Iterative Proportional Scaling Based on a Robust Start Estimator	248
<i>Claudia Becker</i>	
Exploring Multivariate Data Structures with Local Principal Curves	256
<i>Jochen Einbeck, Gerhard Tutz, Ludger Evers</i>	
A Three-way Multidimensional Scaling Approach to the Analysis of Judgments About Persons	264
<i>Sabine Krolak-Schwerdt</i>	
Discovering Temporal Knowledge in Multivariate Time Series	272
<i>Fabian Mörchen, Alfred Ultsch</i>	
A New Framework for Multidimensional Data Analysis	280
<i>Shizuhiko Nishisato</i>	
External Analysis of Two-mode Three-way Asymmetric Multidimensional Scaling	288
<i>Akinori Okada, Tadashi Imaizumi</i>	
The Relevance Vector Machine Under Covariate Measurement Error	296
<i>David Rummel</i>	

Part III. Applications

Archaeology

- A Contribution to the History of Seriation in Archaeology 307
Peter Ihm

- Model-based Cluster Analysis of Roman Bricks and Tiles from
 Worms and Rheinzabern 317
Hans-Joachim Mucha, Hans-Georg Bartel, Jens Dolata

Astronomy

- Astronomical Object Classification and Parameter Estimation
 with the Gaia Galactic Survey Satellite 325
Coryn A.L. Bailer-Jones

- Design of Astronomical Filter Systems for Stellar
 Classification Using Evolutionary Algorithms 330
Coryn A.L. Bailer-Jones

Bio-Sciences

- Analyzing Microarray Data with the Generative Topographic
 Mapping Approach 338
Isabelle M. Grimmerstein, Karsten Quast, Wolfgang Urfer

- Test for a Change Point in Bernoulli Trials with Dependence . 346
Joachim Krauth

- Data Mining in Protein Binding Cavities 354
Katrin Kupas, Alfred Ultsch

- Classification of *In Vivo* Magnetic Resonance Spectra 362
*Björn H. Menze, Michael Wormit, Peter Bachert, Matthias Lichy,
 Heinz-Peter Schlemmer, Fred A. Hamprecht*

- Modifying Microarray Analysis Methods for Categorical Data
 – SAM and PAM for SNPs 370
Holger Schwender

- Improving the Identification of Differentially Expressed Genes
 in cDNA Microarray Experiments 378
Alfred Ultsch

PhyNav: A Novel Approach to Reconstruct Large Phylogenies	386
<i>Le Sy Vinh, Heiko A. Schmidt, Arndt von Haeseler</i>	

Electronic Data and Web

NewsRec, a Personal Recommendation System for News Websites	394
<i>Christian Bomhardt, Wolfgang Gaul</i>	

Clustering of Large Document Sets with Restricted Random Walks on Usage Histories	402
<i>Markus Franke, Anke Thede</i>	

Fuzzy Two-mode Clustering vs. Collaborative Filtering	410
<i>Volker Schlecht, Wolfgang Gaul</i>	

Web Mining and Online Visibility	418
<i>Nadine Schmidt-Mänz, Wolfgang Gaul</i>	

Analysis of Recommender System Usage by Multidimensional Scaling	426
<i>Patrick Thoma, Wolfgang Gaul</i>	

Finance and Insurance

On a Combination of Convex Risk Minimization Methods	434
<i>Andreas Christmann</i>	

Credit Scoring Using Global and Local Statistical Models	442
<i>Alexandra Schwarz, Gerhard Armingier</i>	

Informative Patterns for Credit Scoring Using Linear SVM...	450
<i>Ralf Stecking, Klaus B. Schebesch</i>	

Application of Support Vector Machines in a Life Assurance Environment	458
<i>Sarel J. Steel, Gertrud K. Hechter</i>	

Continuous Market Risk Budgeting in Financial Institutions..	466
<i>Mario Straßberger</i>	

Smooth Correlation Estimation with Application to Portfolio Credit Risk	474
<i>Rafael Weißbach and Bernd Rosenow</i>	

Library Science and Linguistics

How Many Lexical-semantic Relations are Necessary?	482
<i>Dariusch Bagheri</i>	

Automated Detection of Morphemes Using Distributional Measurements	490
<i>Christoph Benden</i>	

Classification of Author and/or Genre? The Impact of Word Length	498
<i>Emmerich Kelih, Gordana Antić, Peter Grzybek, Ernst Stadlober</i>	

Some Historical Remarks on Library Classification – a Short Introduction to the Science of Library Classification	506
<i>Bernd Lorenz</i>	

Automatic Validation of Hierarchical Cluster Analysis with Application in Dialectometry	513
<i>Hans-Joachim Mucha, Edgar Haimerl</i>	

Discovering the Senses of an Ambiguous Word by Clustering its Local Contexts	521
<i>Reinhard Rapp</i>	

Document Management and the Development of Information Spaces	529
<i>Ulfert Rist</i>	

Macro-Economics

Stochastic Ranking and the Volatility “Croissant”: A Sensitivity Analysis of Economic Rankings	537
<i>Helmut Berrer, Christian Helmenstein, Wolfgang Polasek</i>	

Importance Assessment of Correlated Predictors in Business Cycles Classification	545
<i>Daniel Enache, Claus Weihs</i>	

Economic Freedom in the 25-Member European Union: Insights Using Classification Tools	553
<i>Clifford W. Sell</i>	

Marketing

Intercultural Consumer Classifications in E-Commerce	561
<i>Hans H. Bauer, Marcus M. Neumann, Frank Huber</i>	

Reservation Price Estimation by Adaptive Conjoint Analysis .	569
<i>Christoph Breidert, Michael Hahsler, Lars Schmidt-Thieme</i>	

Estimating Reservation Prices for Product Bundles Based on Paired Comparison Data	577
<i>Bernd Stauß, Wolfgang Gaul</i>	

Music Science

Classification of Perceived Musical Intervals	585
<i>Jobst P. Fricke</i>	

In Search of Variables Distinguishing Low and High Achievers in a Music Sight Reading Task.....	593
<i>Reinhard Kopiez, Claus Weihs, Uwe Ligges, Ji In Lee</i>	

Automatic Feature Extraction from Large Time Series.....	600
<i>Ingo Mierswa</i>	

Identification of Musical Instruments by Means of the Hough-Transformation.....	608
<i>Christian Röver, Frank Klefenz, Claus Weihs</i>	

Support Vector Machines for Bass and Snare Drum Recognition.....	616
<i>Dirk Van Steelant, Koen Tanghe, Sven Degroeve, Bernard De Baets, Marc Leman, Jean-Pierre Martens</i>	

Register Classification by Timbre.....	624
<i>Claus Weihs, Christoph Reuter, Uwe Ligges</i>	

Quality Assurance

Classification of Processes by the Lyapunov Exponent	632
<i>Anja M. Busse</i>	

Desirability to Characterize Process Capability	640
<i>Jutta Jessenberger, Claus Weihs</i>	

Application and Use of Multivariate Control Charts in a BTA Deep Hole Drilling Process	648
<i>Amor Messaoud, Winfried Theis, Claus Weihs, Franz Hering</i>	

Determination of Relevant Frequencies and Modeling Varying Amplitudes of Harmonic Processes	656
<i>Winfried Theis, Claus Weihs</i>	

Part IV. Contest: Social Milieus in Dortmund

Introduction to the Contest “Social Milieus in Dortmund” ...	667
<i>Ernst-Otto Sommerer, Claus Weihs</i>	
Application of a Genetic Algorithm to Variable Selection in Fuzzy Clustering	674
<i>Christian Röver, Gero Szepannek</i>	
Annealed k-Means Clustering and Decision Trees	682
<i>Christin Schäfer, Julian Laub</i>	
Correspondence Clustering of Dortmund City Districts	690
<i>Stefanie Scheid</i>	
Keywords	698
Authors	703

Classification - the Ubiquitous Challenge
Proceedings of the 28th Annual Conference of the
Gesellschaft für Klassifikation e.V., University of
Dortmund, March 9-11, 2004
Weihs, C.; Gaul, W. (Eds.)
2005, XX, 704 p. 181 illus., Softcover
ISBN: 978-3-540-25677-9