

Contents

**Part I Classification, Cluster Analysis, and Multidimensional
Scaling**

**Fuzzification of Agglomerative Hierarchical
Crisp Clustering Algorithms** 3
Mathias Bank and Friedhelm Schwenker

An EM Algorithm for the Student-*t* Cluster-Weighted Modeling 13
Salvatore Ingrassia, Simona C. Minotti, and Giuseppe Incarbone

Analysis of Distribution Valued Dissimilarity Data 23
Masahiro Mizuta and Hiroyuki Minami

An Overall Index for Comparing Hierarchical Clusterings 29
I. Morlini and S. Zani

An Accelerated *K*-Means Algorithm Based on Adaptive Distances 37
Hans-Joachim Mucha and Hans-Georg Bartel

Bias-Variance Analysis of Local Classification Methods 49
Julia Schiffner, Bernd Bischl, and Claus Weihs

Effect of Data Standardization on the Result of *k*-Means Clustering 59
Kensuke Tanioka and Hiroshi Yadohisa

**A Case Study on the Use of Statistical Classification Methods
in Particle Physics** 69
Claus Weihs, Olaf Mersmann, Bernd Bischl, Arno Fritsch,
Heike Trautmann, Till Moritz Karbach, and Bernhard Spaan

**Problems of Fuzzy c-Means Clustering and Similar Algorithms
with High Dimensional Data Sets** 79
Roland Winkler, Frank Klawonn, and Rudolf Kruse

Part II Quantification Theory

Reduced Versus Complete Space Configurations in Total Information Analysis	91
José G. Clavel and Shizuhiko Nishisato	
A Geometrical Interpretation of the Horseshoe Effect in Multiple Correspondence Analysis of Binary Data	101
Takashi Murakami	
Quantification Theory: Reminiscence and a Step Forward	109
Shizuhiko Nishisato	

Part III Analysis of m-Mode n-Way and Asymmetric Data

Modelling Rater Differences in the Analysis of Three-Way Three-Mode Binary Data.....	123
Michel Meulders	
Reconstructing One-Mode Three-way Asymmetric Data for Multidimensional Scaling	133
Atsuhō Nakayama and Akinori Okada	
Analysis of Car Switching Data by Asymmetric Multidimensional Scaling Based on Singular Value Decomposition	143
Akinori Okada	
Visualization of Asymmetric Clustering Result with Digraph and Dendrogram	151
Yuichi Saito and Hiroshi Yadohisa	

Part IV Analysis of Visual, Spatial, and Temporal Data

Clustering Temporal Population Patterns in Switzerland (1850–2000)	163
Martin Behnisch and Alfred Ultsch	
p-adic Methods in Stereo Vision.....	173
Patrick Erik Bradley	
Individualized Error Estimation for Classification and Regression Models	183
Krisztian Buza, Alexandros Nanopoulos, and Lars Schmidt-Thieme	
Evaluation of Spatial Cluster Detection Algorithms for Crime Locations.....	193
Marco Helbich and Michael Leitner	

Checking Serial Independence of Residuals from a Nonlinear Model	203
Luca Bagnato and Antonio Punzo	

Part V Network Data, Graphs, and Social Relationships

Analysis of Network Data Based on Probability	
Neighborhood Cliques	215
Andreas Baumgart and Ulrich Müller-Funk	
A Comparison of Agglomerative Hierarchical Algorithms for	
Modularity Clustering	225
Michael Ovelgönne and Andreas Geyer-Schulz	
Network Data as Contiguity Constraints in Modeling	
Preference Data	233
Giuseppe Giordano and Germana Scepi	
Clustering Coefficients of Random Intersection Graphs	
243	
Erhard Godehardt, Jerzy Jaworski, and Katarzyna Rybarczyk	
Immersive Dynamic Visualization of Interactions in a Social Network . . .	
255	
Nicolas Greffard, Fabien Picarougne, and Pascale Kuntz	
Fuzzy Boolean Network Reconstruction	
263	
Martin Hopfensitz, Markus Maucher, and Hans A. Kestler	
GIRAN: A Dynamic Graph Interface to Neighborhoods of	
Related Articles	271
Andreas W. Neumann and Kiryl Batsiukov	
Power Tags as Tools for Social Knowledge Organization Systems	
281	
Isabella Peters	
The Degree Distribution in Random Intersection Graphs	
291	
Katarzyna Rybarczyk	
Application of a Community Membership Life Cycle Model on	
Tag-Based Communities in Twitter	301
Andreas C. Sonnenbichler and Christopher Bazant	
Measuring the Influence of Network Structures on Social	
Interaction over Time	311
Christoph Stadtfeld	
Identifying Artificial Actors in E-Dating: A Probabilistic	
Segmentation Based on Interactional Pattern Analysis	319
Andreas Schmitz, Olga Yanenko, and Marcel Hebing	

Part VI Text Mining

Calculating a Distributional Similarity Kernel using the Nystrom Extension	331
Markus Arndt and Ulrich Arndt	

Text Categorization in R: A Reduced N-Gram Approach	341
Wilhelm M. Geiger, Johannes Rauch, Patrick Mair, and Kurt Hornik	

Part VII Dimension Reduction

HOMALS for Dimension Reduction in Information Retrieval	353
Kay F. Hildebrand and Ulrich Müller-Funk	

Feature Reduction and Nearest Neighbours	363
Ludwig Lausser, Christoph Müssel, Markus Maucher, and Hans A. Kestler	

Part VIII Statistical Musicology

Musical Instrument Recognition by High-Level Features	373
Markus Eichhoff and Claus Weihs	

The Recognition of Consonance Is not Impaired by Intonation Deviations: A Revised Theory	383
Jobst Peter Fricke	

Applying Multiple Kernel Learning to Automatic Genre Classification ...	393
Hanna Lukashevich	

Multi-Objective Evaluation of Music Classification.....	401
Igor Vatolkin	

Partition Based Feature Processing for Improved Music Classification ...	411
Igor Vatolkin, Wolfgang Theimer, and Martin Botteck	

Software in Music Information Retrieval	421
Claus Weihs, Klaus Friedrichs, Markus Eichhoff, and Igor Vatolkin	

Part IX Data Analysis in Banking and Finance

Conditional Factor Models for European Banks	435
Wolfgang Bessler and Philipp Kurmann	

Classification of Large Imbalanced Credit Client Data with Cluster Based SVM	443
Ralf Stecking and Klaus B. Schebesch	

Fault Mining Using Peer Group Analysis	453
David J. Weston, Niall M. Adams, Yoonseong Kim, and David J. Hand	

Part X Data Analysis in Health and Environment

Discovering Possible Patterns Associations Among Drug Prescriptions ...	465
Joana Fernandes and Orlando Belo	

Cluster Analytic Strategy for Identification of Metagenes Relevant for Prognosis of Node Negative Breast Cancer	475
Evgenia Freis, Silvia Selinski, Jan G. Hengstler and Katja Ickstadt	

Part XI Analysis of Marketing, Conjoint, and Multigroup Data

Image Clustering for Marketing Purposes	487
Daniel Baier and Ines Daniel	

PLS-MGA: A Non-Parametric Approach to Partial Least Squares-based Multi-Group Analysis.....	495
Jörg Henseler	

Love and Loyalty in Car Brands: Segmentation Using Finite Mixture Partial Least Squares	503
Sandra Loureiro	

Endogeneity and Exogeneity in Sales Response Functions	511
Wolfgang Polasek	

Lead User Identification in Conjoint Analysis Based Product Design	521
Alexander Sänn and Daniel Baier	

Improving the Validity of Conjoint Analysis by Additional Data Collection and Analysis Steps	529
Sebastian Selka, Daniel Baier, and Michael Brusch	

The Impact of Missing Values on PLS Model Fitting	537
Moritz Parwoll and Ralf Wagner	

Part XII Data Analysis in Education and Psychology

Teachers' Typology of Student Categories: A Cluster Analytic Study	547
Thomas Hörstermann and Sabine Krolak-Schwerdt	

ECTS Grades: Combination of Norm- and Criterion- Referenced Grading	557
Jonas Kunze and Andreas Geyer-Schulz	

Students Reading Motivation: A Multilevel Mixture Factor Analysis	567
Daniele Riggi and Jeroen K. Vermunt	

Part XIII Analysis of Tourism Data

Short Term Dynamics of Tourist Arrivals: What Do Italian Destinations Have in Common?	577
Anna Maria Parroco and Raffaele Scuderi	

A Measure of Polarization for Tourism: Evidence from Italian Destinations	585
Raffaele Scuderi	

Index	595
--------------------	------------

Challenges at the Interface of Data Analysis, Computer
Science, and Optimization

Proceedings of the 34th Annual Conference of the
Gesellschaft für Klassifikation e. V., Karlsruhe, July 21 -
23, 2010

Gaul, W.; Geyer-Schulz, A.; Schmidt-Thieme, B.; Kunze, J.
(Eds.)

2012, XIV, 598 p. 163 illus., Softcover

ISBN: 978-3-642-24465-0