
Contents

Part I New Methods in Cluster Analysis

Thinking Ultrametrically <i>Fionn Murtagh</i>	3
Clustering by Vertex Density in a Graph <i>Alain Guénoche</i>	15
Clustering by Ant Colony Optimization <i>Javier Trejos, Alex Murillo, Eduardo Piza</i>	25
A Dynamic Cluster Algorithm Based on L_r Distances for Quantitative Data <i>Francisco de A. T. de Carvalho, Yves Lechevallier, Renata M.C.R. de Souza</i>	33
The Last Step of a New Divisive Monothetic Clustering Method: the Gluing-Back Criterion <i>Jean-Yves Pirçon, Jean-Paul Rasson</i>	43
Standardizing Variables in K-means Clustering <i>Douglas Steinley</i>	53
A Self-Organizing Map for Dissimilarity Data <i>Aïcha El Golli, Briec Conan-Guez, Fabrice Rossi</i>	61
Another Version of the Block EM Algorithm <i>Mohamed Nadif, Gérard Govaert</i>	69
Controlling the Level of Separation of Components in Monte Carlo Studies of Latent Class Models <i>José G. Dias</i>	77

Fixing Parameters in the Constrained Hierarchical Classification Method: Application to Digital Image Segmentation	
<i>Kaddour Bachar, Israël-César Lerman</i>	85
New Approaches for Sum-of-Diameters Clustering	
<i>Sarnath Ramnath, Mynul H. Khan,, Zubair Shams</i>	95
Spatial Pyramidal Clustering Based on a Tessellation	
<i>Edwin Diday</i>	105

Part II Modern Nonparametrics

Relative Projection Pursuit and its Application	
<i>Masahiro Mizuta and Shintaro Hiro</i>	123
Priors for Neural Networks	
<i>Herbert K. H. Lee</i>	141
Combining Models in Discrete Discriminant Analysis Through a Committee of Methods	
<i>Ana Sousa Ferreira</i>	151
Phoneme Discrimination with Functional Multi-Layer Perceptrons	
<i>Brieuc Conan-Guez, Fabrice Rossi</i>	157
PLS Approach for Clusterwise Linear Regression on Functional Data	
<i>Cristian Preda, Gilbert Saporta</i>	167
On Classification and Regression Trees for Multiple Responses	
<i>Seong Keon Lee</i>	177
Subsetting Kernel Regression Models Using Genetic Algorithm and the Information Measure of Complexity	
<i>Xinli Bao, Hamparsum Bozdogan</i>	185
Cherry-Picking as a Robustness Tool	
<i>Leanna L. House, David Banks</i>	197

Part III Classification and Dimension Reduction

Academic Obsessions and Classification Realities: Ignoring Practicalities in Supervised Classification	
<i>David J. Hand</i>	209

Modified Biplots for Enhancing Two-Class Discriminant Analysis	
<i>Sugnet Gardner, Niel le Roux</i>	233
Weighted Likelihood Estimation of Person Locations in an Unfolding Model for Polytomous Responses	
<i>Guanzhong Luo, David Andrich</i>	241
Classification of Geospatial Lattice Data and their Graphical Representation	
<i>Koji Kurihara</i>	251
Degenerate Expectation-Maximization Algorithm for Local Dimension Reduction	
<i>Xiaodong Lin, Yu Zhu</i>	259
A Dimension Reduction Technique for Local Linear Regression	
<i>Joseph Lucas</i>	269
Reducing the Number of Variables Using Implicative Analysis	
<i>Raphaël Couturier, Régis Gras, Fabrice Guillet</i>	277
Optimal Discretization of Quantitative Attributes for Association Rules	
<i>Stefan Born, Lars Schmidt-Thieme</i>	287
<hr/>	
Part IV Symbolic Data Analysis	
<hr/>	
Clustering Methods in Symbolic Data Analysis	
<i>Rosanna Verde</i>	299
Dependencies in Bivariate Interval-Valued Symbolic Data	
<i>L. Billard</i>	319
Clustering of Symbolic Objects Described by Multi-Valued and Modal Variables	
<i>André Hardy, Pascale Lallemand</i>	325
A Hausdorff Distance Between Hyper-Rectangles for Clustering Interval Data	
<i>Marie Chavent</i>	333
Kolmogorov-Smirnov for Decision Trees on Interval and Histogram Variables	
<i>Chérif Mballo and Edwin Diday</i>	341

Dynamic Cluster Methods for Interval Data Based on Mahalanobis Distances <i>Renata M.C.R. de Souza, Francisco de A.T. de Carvalho, Camilo P. Tenório, Yves Lechevallier</i>	351
A Symbolic Model-Based Approach for Making Collaborative Group Recommendations <i>Sérgio R. de M. Queiroz, Francisco de A. T. de Carvalho</i>	361
Probabilistic Allocation of Aggregated Statistical Units in Classification Trees for Symbolic Class Description <i>Mohamed Mehdi Limam, Edwin Diday, Suzanne Winsberg</i>	371
Building Small Scale Models of Multi-Entity Databases by Clustering <i>Georges Hébraïl, Yves Lechevallier</i>	381
<hr/>	
Part V Taxonomy and Medicine	
<hr/>	
Phylogenetic Closure Operations and Homoplasy-Free Evolution <i>Tobias Dezulian, Mike Steel</i>	395
Consensus of Classification Systems, with Adams' Results Revisited <i>Florent Domenach, Bruno Leclerc</i>	417
Symbolic Linear Regression with Taxonomies <i>F. Afonso, L. Billard, and E. Diday</i>	429
Determining Horizontal Gene Transfers in Species Classification: Unique Scenario <i>Vladimir Makarenkov, Alix Boc, Abdoulaye Baniré Diallo</i>	439
Active and Passive Learning to Explore a Complex Metabolism Data Set <i>Susan J. Simmons, Xiaodong Lin, Chris Beecher, Young Truong and S. Stanley Young</i>	447
Mathematical and Statistical Modeling of Acute Inflammation <i>Gilles Clermont, Carson C. Chow, Gregory M. Constantine, Yoram Vodovotz, John Bartels</i>	457
Combining Functional MRI Data on Multiple Subjects <i>Mark G. Vangel</i>	469

Classifying the State of Parkinsonism by Using Electronic Force Platform Measures of Balance
Nicholas I. Bohnen, Marius G. Buliga, Gregory M. Constantine 477

Subject Filtering for Passive Biometric Monitoring
Vahan Grigoryan, Donald Chiarulli, Milos Hauskrecht 485

Part VI Text Mining

Mining Massive Text Data and Developing Tracking Statistics
Daniel R. Jeske, Regina Y. Liu 495

Contributions of Textual Data Analysis to Text Retrieval
Simona Balbi, Emilio Di Meglio 511

Automated Resolution of Noisy Bibliographic References
Markus Demleitner, Michael Kurtz, Alberto Accomazzi, Günther Eichhorn, Carolyn S. Grant, Steven S. Murray 521

Choosing the Right Bigrams for Information Retrieval
Maojin Jiang, Eric Jensen, Steve Beitzel, Shlomo Argamon 531

A Mixture Clustering Model for Pseudo Feedback in Information Retrieval
Tao Tao, ChengXiang Zhai 541

Analysis of Cross-Language Open-Ended Questions Through MFACT
Mónica Bécue, Jérôme Pagès, Campo-Elías Pardo 553

Inferring User’s Information Context from User Profiles and Concept Hierarchies
Ahu Sieg, Bamshad Mobasher, Robin Burke 563

Database Selection for Longer Queries
Wensheng Wu, Clement Yu, Weiyi Meng 575

Part VII Contingency Tables and Missing Data

An Overview of Collapsibility
Stefano De Cantis and Antonino M. Oliveri 587

Generalized Factor Analyses for Contingency Tables
François Bavaud 597

A PLS Approach to Multiple Table Analysis <i>Michel Tenenhaus</i>	607
Simultaneous Row and Column Partitioning in Several Contingency Tables <i>Vincent Loonis</i>	621
Missing Data and Imputation Methods in Partition of Variables <i>Ana Lorga da Silva, Gilbert Saporta, Helena Bacelar-Nicolau</i>	631
The Treatment of Missing Values and its Effect on Classifier Accuracy <i>Edgar Acuña, Caroline Rodriguez</i>	639
Clustering with Missing Values: No Imputation Required <i>Kiri Wagstaff</i>	649