Studies in Classification, Data Analysis, and Knowledge Organization

Springer-Verlag Berlin Heidelberg GmbH

Titles in the Series

H.-H. Bock and P. Ihm (Eds.)
Classification, Data Analysis, and Knowledge Organization. 1991 (out of print)

M. Schader (Ed.)
Analyzing and Modeling Data and Knowledge. 1992

O. Opitz, B. Lausen, and R. Klar (Eds.)
Information and Classification. 1993 (out of print)

H.-H. Bock, W. Lenski, and M. M. Richter (Eds.)
Information Systems and Data Analysis. 1994 (out of print)

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
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

Krzysztof Jajuga · Andrzej Sokołowski
Hans-Hermann Bock (Eds.)

Classification, Clustering, and Data Analysis

Recent Advances and Applications

With 84 Figures and 65 Tables

Springer

Prof. Krzysztof Jajuga
Wroclaw University
of Economics
ul. Komandorska 118/120
53-345 Wroclaw
Poland
jajuga@manager.ae.wroc.pl

Prof. Andrzej Sokołowski
Department of Statistics
Cracow University
of Economics
ul. Rakowicka 27
31-510 Cracow
Poland
sokolows@ae.krakow.pl

Prof. Hans-Hermann Bock
Technical University of Aachen
Institute of Statistics
Wuellnerstrasse 3
52056 Aachen
Germany
bock@stochastik.rwth-
aachen.de

ISSN 1431-8814
ISBN 978-3-540-43691-1

Cataloging-in-Publication Data applied for
Die Deutsche Bibliothek - CIP-Einheitsaufnahme
Classification, clustering and data analysis: recent advances and applications / Krzysztof Jajuga ... (ed.). - Berlin; Heidelberg; New York; Barcelona; Hong Kong; London; Milan; Paris; Tokyo: Springer, 2002
(Studies in classification, data analysis, and knowledge organization)
ISBN 978-3-540-43691-1 ISBN 978-3-642-56181-8 (eBook)
DOI 10.1007/978-3-642-56181-8

http://www.springer.de

Originally published by Springer-Verlag Berlin · Heidelberg in 2002

Softcover-Design: Erich Kirchner, Heidelberg

43/3111 - 5 4 3 2 1 – Printed on acid-free paper

Preface

The present volume contains a selection of papers presented at the Eighth Conference of the International Federation of Classification Societies (IFCS) which was held in Cracow, Poland, July 16-19, 2002. All originally submitted papers were subject to a reviewing process by two independent referees, a procedure which resulted in the selection of the 53 articles presented in this volume.

These articles relate to theoretical investigations as well as to practical applications and cover a wide range of topics in the broad domain of classification, data analysis and related methods. If we try to classify the wealth of problems, methods and approaches into some representative (partially overlapping) groups, we find in particular the following areas:

- Clustering
- Cluster validation
- Discrimination
- Multivariate data analysis
- Statistical methods
- Symbolic data analysis
- Consensus trees and phylogeny
- Regression trees
- Neural networks and genetic algorithms
- Applications in economics, medicine, biology, and psychology.

Given the international orientation of IFCS conferences and the leading role of IFCS in the scientific world of classification, clustering and data analysis, this volume collects a representative selection of current research and modern applications in this field and serves as an up-to-date information source for statisticians, data analysts, data mining specialists and computer scientists.

This is well in the mainstream of the activities of the International Federation of Classification Societies which considers as one of its main purposes the dissemination of technical and scientific information concerning data analysis, classification, related methods and their applications. Note that the IFCS comprizes, as its members, the following twelve regional or national classification societies:

- British Classification Society (BCS)
- Associaao Portuguesa de Classificação e Análise de Dados (CLAD)
- Classification Society of North America (CSNA)
- Gesellschaft für Klassifikation (GfKl)
- Japanese Classification Society (JCS)
- Korean Classification Society (KCS)

- Société Francophone de Classification (SFC)
- Società Italiana di Statistica (SIS)
- Sekcja Klasyfikacji i Analizy Danych PTS (SKAD)
- Vereniging voor Ordinatie en Classificatie (VOC)
- Irish Pattern Recognition and Classification Society (IPRCS)
- Central American and Carribean Society of Classification and Data Analysis (CASCCDA)

Previous conferences of IFCS took place in Aachen (1987), Charlottesville (1989), Edinburgh (1991), Paris (1993), Kobe (1996), Rome (1998), and Namur (2000).

The editors of this volume would like express their gratitude towards the authors of the papers contained in this volume for their valuable contributions. In particular, they thank the reviewers for their professional, careful and often time-consuming work when reviewing the submitted papers. Our special thanks and gratitude go to Dr. Katarzyna Kuziak of Wrocław University of Economics for her outstanding work and great help in the process of reviewing and preparing final version of the book. We express our gratitude to Dr. Andrzej Bąk of Wrocław University of Economics for his professional contribution in the process of the formatting and technical editing of this volume. Thanks also to Dr. Martina Bihn from Springer Verlag (Heidelberg) for the smooth and timely production and dispatch of these Proceedings.

Aachen, Wrocław, and Cracow
July 2002

Hans-Hermann Bock
Krzysztof Jajuga
Andrzej Sokołowski

Contents

Part III. Applications