Intelligent Systems Reference Library, Volume 24

**Editors-in-Chief**

Prof. Janusz Kacprzyk
Systems Research Institute
Polish Academy of Sciences
ul. Newelska 6
01-447 Warsaw
Poland
*E-mail:* kacprzyk@ibspan.waw.pl

Prof. Lakhmi C. Jain
University of South Australia
Adelaide
Mawson Lakes Campus
South Australia 5095
Australia
*E-mail:* Lakhmi.jain@unisa.edu.au

---

Further volumes of this series can be found on our homepage:
springer.com

Vol. 1. Christine L. Mumford and Lakhmi C. Jain (Eds.)
*Computational Intelligence: Collaboration, Fusion and Emergence,* 2009
ISBN 978-3-642-01798-8

Vol. 2. Yuehui Chen and Ajith Abraham
*Tree-Structure Based Hybrid Computational Intelligence,* 2009
ISBN 978-3-642-04738-1

Vol. 3. Anthony Finn and Steve Scheding
*Developments and Challenges for Autonomous Unmanned Vehicles,* 2010
ISBN 978-3-642-10703-0

Vol. 4. Lakhmi C. Jain and Chee Peng Lim (Eds.)
*Handbook on Decision Making: Techniques and Applications,* 2010
ISBN 978-3-642-13638-2

Vol. 5. George A. Anastassiou
*Intelligent Mathematics: Computational Analysis,* 2010
ISBN 978-3-642-17097-3

Vol. 6. Ludmila Dymowa
*Soft Computing in Economics and Finance,* 2011
ISBN 978-3-642-17718-7

Vol. 7. Gerasimos G. Rigatos
*Modelling and Control for Intelligent Industrial Systems,* 2011
ISBN 978-3-642-17874-0

Vol. 8. Edward H.Y. Lim, James N.K. Liu, and Raymond S.T. Lee
*Knowledge Seeker – Ontology Modelling for Information Search and Management,* 2011
ISBN 978-3-642-17915-0

Vol. 9. Menahem Friedman and Abraham Kandel
*Calculus Light,* 2011
ISBN 978-3-642-17847-4

Vol. 10. Andreas Tolk and Lakhmi C. Jain
*Intelligence-Based Systems Engineering,* 2011
ISBN 978-3-642-17930-3

Vol. 11. Samuli Niiranen and Andre Ribeiro (Eds.)
*Information Processing and Biological Systems,* 2011
ISBN 978-3-642-19620-1

Vol. 12. Florin Gorunescu
*Data Mining,* 2011
ISBN 978-3-642-19720-8

Vol. 13. Witold Pedrycz and Shyi-Ming Chen (Eds.)
*Granular Computing and Intelligent Systems,* 2011
ISBN 978-3-642-19819-9

Vol. 14. George A. Anastassiou and Oktay Duman
*Towards Intelligent Modeling: Statistical Approximation Theory,* 2011
ISBN 978-3-642-19825-0

Vol. 15. Antonino Freno and Edmondo Trentin
*Hybrid Random Fields,* 2011
ISBN 978-3-642-20307-7

Vol. 16. Alexiei Dingli
*Knowledge Annotation: Making Implicit Knowledge Explicit,* 2011
ISBN 978-3-642-20322-0

Vol. 17. Crina Grosan and Ajith Abraham
*Intelligent Systems,* 2011
ISBN 978-3-642-21003-7

Vol. 18. Achim Zielesny
*From Curve Fitting to Machine Learning,* 2011
ISBN 978-3-642-21279-6

Vol. 19. George A. Anastassiou
*Intelligent Systems: Approximation by Artificial Neural Networks,* 2011
ISBN 978-3-642-21430-1

Vol. 20. Lech Polkowski
*Approximate Reasoning by Parts,* 2011
ISBN 978-3-642-22278-8

Vol. 21. Igor Chikalov
*Average Time Complexity of Decision Trees,* 2011
ISBN 978-3-642-22660-1

Vol. 22. Przemysław Różewski,
Emma Kusztina, Ryszard Tadeusiewicz,
and Oleg Zaikin
*Intelligent Open Learning Systems,* 2011
ISBN 978-3-642-22666-3

Vol. 23. Dawn E. Holmes and Lakhmi C. Jain (Eds.)
*Data Mining: Foundations and Intelligent Paradigms,* 2012
ISBN 978-3-642-23165-0

Vol. 24. Dawn E. Holmes and Lakhmi C. Jain (Eds.)
*Data Mining: Foundations and Intelligent Paradigms,* 2012
ISBN 978-3-642-23240-4

Dawn E. Holmes and Lakhmi C. Jain (Eds.)

# Data Mining: Foundations and Intelligent Paradigms

## Volume 2: Statistical, Bayesian, Time Series and other Theoretical Aspects

Springer

Prof. Dawn E. Holmes
Department of Statistics and Applied Probability
University of California
Santa Barbara,
CA 93106
USA
E-mail: holmes@pstat.ucsb.edu

Prof. Lakhmi C. Jain
Professor of Knowledge-Based Engineering
University of South Australia
Adelaide
Mawson Lakes, SA 5095
Australia
E-mail: Lakhmi.jain@unisa.edu.au

ISBN 978-3-642-23240-4 e-ISBN 978-3-642-23241-1

DOI 10.1007/978-3-642-23242-8

Intelligent Systems Reference Library ISSN 1868-4394

Library of Congress Control Number: 2011936705



*Typeset* & *Cover Design:* Scientific Publishing Services Pvt. Ltd., Chennai, India.

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

# Preface

There are many invaluable books available on data mining theory and applications. However, in compiling a volume titled "DATA MINING: Foundations and Intelligent Paradigms: Volume 2: Core Topics including Statistical, Time-Series and Bayesian Analysis" we wish to introduce some of the latest developments to a broad audience of both specialists and non-specialists in this field.

The term 'data mining' was introduced in the 1990's to describe an emerging field based on classical statistics, artificial intelligence and machine learning. Important core areas of data mining such as support vector machines, a kernel based learning method, have been very productive in recent years as attested by the rapidly increasing number of papers published each year. Time series analysis and prediction have been enhanced by methods in neural networks, particularly in the area of financial forecasting. Bayesian analysis is of primary importance in data mining research, with ongoing work in prior probability distribution estimation.

In compiling this volume we have sought to present innovative research from prestigious contributors in these particular areas of data mining. Each chapter is self-contained and is described briefly in Chapter 1.

This book will prove valuable to theoreticians as well as application scientists/ engineers in the area of Data Mining. Postgraduate students will also find this a useful sourcebook since it shows the direction of current research.

We have been fortunate in attracting top class researchers as contributors and wish to offer our thanks for their support in this project. We also acknowledge the expertise and time of the reviewers. Finally, we also wish to thank Springer for their support.

Dr. Dawn E. Holmes
University of California
Santa Barbara, USA

Dr. Lakhmi C. Jain
University of South Australia
Adelaide, Australia

# Contents

# Editors

Dr. Dawn E. Holmes serves as Senior Lecturer in the Department of Statistics and Applied Probability and Senior Associate Dean in the Division of Undergraduate Education at UCSB. Her main research area, Bayesian Networks with Maximum Entropy, has resulted in numerous journal articles and conference presentations. Her other research interests include Machine Learning, Data Mining, Foundations of Bayesianism and Intuitionistic Mathematics. Dr. Holmes has co-edited, with Professor Lakhmi C. Jain, volumes 'Innovations in Bayesian Networks' and 'Innovations in Machine Learning'. Dr. Holmes teaches a broad range of courses, including SAS programming, Bayesian Networks and Data Mining. She was awarded the Distinguished Teaching Award by Academic Senate, UCSB in 2008.

As well as being Associate Editor of the International Journal of Knowledge-Based and Intelligent Information Systems, Dr. Holmes reviews extensively and is on the editorial board of several journals, including the Journal of Neurocomputing. She serves as Program Scientific Committee Member for numerous conferences; including the International Conference on Artificial Intelligence and the International Conference on Machine Learning. In 2009 Dr. Holmes accepted an invitation to join Center for Research in Financial Mathematics and Statistics (CRFMS), UCSB. She was made a Senior Member of the IEEE in 2011.

Professor Lakhmi C. Jain is a Director/Founder of the Knowledge-Based Intelligent Engineering Systems (KES) Centre, located in the University of South Australia. He is a fellow of the Institution of Engineers Australia.

His interests focus on the artificial intelligence paradigms and their applications in complex systems, art-science fusion, e-education, e-healthcare, unmanned air vehicles and intelligent agents.