Advanced Information and Knowledge Processing

Also in this series

Gregoris Mentzas, Dimitris Apostolou, Andreas Abecker and Ron Young
Knowledge Asset Management
1-85233-583-1

Michalis Vazirgiannis, Maria Halkidi and Dimitrios Gunopulos
Uncertainty Handling and Quality Assessment in Data Mining
1-85233-655-2

Asunción Gómez-Pérez, Mariano Fernández-López and Oscar Corcho
Ontological Engineering
1-85233-551-3

Amo Scharil (Ed.)
Environmental Online Communication
1-85233-783-4

Shichao Zhang, Chengqi Zhang and Xindong Wu
Knowledge Discovery in Multiple Databases
1-85233-703-6

Jason T.L. Wang, Mohammed J. Zaki,
Hannu T.T. Toivonen and Dennis Shasha (Eds)

Data Mining in Bioinformatics

With 110 Figures

Springer

Jason T.L. Wang, PhD
New Jersey Institute of Technology, USA

Mohammed J. Zaki, PhD
Computer Science Department, Rensselaer Polytechnic Institute, USA

Hannu T.T. Toivonen, PhD
University of Helsinki and Nokia Research Center

Dennis Shasha, PhD
New York University, USA

Series Editors
Xindong Wu
Lakhmi Jain

British Library Cataloguing in Publication Data
Data mining in bioinformatics. — (Advanced information and
knowledge processing)
1. Data mining 2. Bioinformatics — Data processing
I. Wang, Jason T. L.
006.3′12
ISBN 1852336714

Library of Congress Cataloging-in-Publication Data
A catalogue record for this book is available from the American Library of Congress.

AI&KP ISSN 1610-3947
ISBN 1-85233-671-4 Springer London Berlin Heidelberg
Springer Science+Business Media
springeronline.com

The publisher makes no representation, express or implied, with regard to the accuracy of the information contained in this book and cannot accept any legal responsibility or liability for any errors or omissions that may be made.

Typesetting: Electronic text files prepared by authors
Printed and bound in the United States of America
34/3830-543210 Printed on acid-free paper SPIN 10886107

Contents

Contributors

Peter Bajcsy
Center for Supercomputing Applications
University of Illinois at Urbana-Champaign
USA

Deb Bardhan
Department of Computer Science
Rensselaer Polytechnic Institute
USA

Chris Bystroff
Department of Biology
Rensselaer Polytechnic Institute
USA

Mukund Deshpande
Oracle Corporation
USA

Cinzia Di Pietro
School of Medicine
University of Catania
Italy

Alfredo Ferro
Department of Mathematics and Computer Science
University of Catania
Italy

Laurie Jane Hammel
Department of Defense
USA

Jiawei Han
Department of Computer Science
University of Illinois at Urbana-Champaign
USA

Kai Huang
Department of Biological Sciences
Carnegie Mellon University
USA

Donald P. Huddler
Biophysics Research Division
University of Michigan
USA

George Karypis
Department of Computer Science and Engineering
University of Minnesota
USA

Michihiro Kuramochi
Department of Computer Science and Engineering
University of Minnesota
USA

Lei Liu
Center for Comparative and Functional Genomics
University of Illinois at Urbana-Champaign
USA

Heikki Mannila
Department of Computer Science
Helsinki University of Technology
Finland

Robert F. Murphy
Departments of Biological Sciences and Biomedical Engineering
Carnegie Mellon University
USA

Vinay Nadimpally
Department of Computer Science
Rensselaer Polytechnic Institute
USA

Päivi Onkamo
Department of Computer Science
University of Helsinki
Finland

Roderic D. M. Page
Division of Environmental and Evolutionary Biology
Institute of Biomedical and Life Sciences
University of Glasgow
United Kingdom

Jignesh M. Patel
Electrical Engineering and Computer Science Department
University of Michigan
USA

Giuseppe Pigola
Department of Mathematics and Computer Science
University of Catania
Italy

Alfredo Pulvirenti
Department of Mathematics and Computer Science
University of Catania
Italy

Michele Purrello
School of Medicine
University of Catania
Italy

Marco Ragusa
School of Medicine
University of Catania
Italy

Marko Salmenkivi
Department of Computer Science
University of Helsinki
Finland

Petteri Sevon
Department of Computer Science
University of Helsinki
Finland

Dennis Shasha
Courant Institute of Mathematical Sciences
New York University
USA

Ambuj K. Singh
Department of Computer Science
University of California at Santa Barbara
USA

Hannu T. T. Toivonen
Department of Computer Science
University of Helsinki
Finland

Jason T. L. Wang
Department of Computer Science
New Jersey Institute of Technology
USA

Jiong Yang
Department of Computer Science
University of Illinois at
Urbana-Champaign
USA

Mohammed J. Zaki
Department of Computer Science
Rensselaer Polytechnic Institute
USA

Kaizhong Zhang
Department of Computer Science
University of Western Ontario
Canada

Part I

Overview