Contents

Foreword vii

Preface xiii

1 From Data Mining to Knowledge Discovery: An Overview

Usama M. Fayyad, Gregory Piatetsky-Shapiro, and Padhraic Smyth 1

I FOUNDATIONS

2 The Process of Knowledge Discovery in Databases: A Human-Centered Approach

Ronald J. Brachman and Tej Anand 37

3 Graphical Models for Discovering Knowledge

Wray Buntine 59

4 A Statistical Perspective on Knowledge Discovery in Databases

John Elder IV and Daryl Pregibon 83

II CLASSIFICATION AND CLUSTERING

5 Inductive Logic Programming and Knowledge Discovery in Databases

Sašo Džeroski 117

6 Bayesian Classification (AutoClass): Theory and Results

Peter Cheeseman and John Stutz 153

7 Discovering Informative Patterns and Data Cleaning

Isabelle Guyon, Nada Matić, and Vladimir Vapnik 181
8 Transforming Rules and Trees into Comprehensible Knowledge Structures  
*Brian R. Gaines* 205

III **TREND AND DEVIATION ANALYSIS**

9 Finding Patterns in Time Series: A Dynamic Programming Approach  
*Donald J. Berndt and James Clifford* 229

10 Explora: A Multipattern and Multistrategy Discovery Assistant  
*Willi Klösgen* 249

IV **DEPENDENCY DERIVATION**

11 Bayesian Networks for Knowledge Discovery  
*David Heckerman* 273

12 Fast Discovery of Association Rules  
*Rakesh Agrawal, Heikki Mannila, Ramakrishnan Srikant, Hannu Toivonen, and A. Inkeri Verkamo* 307

13 From Contingency Tables to Various Forms of Knowledge in Databases  
*Robert Zembowicz and Jan M. Żytkowski* 329

V **INTEGRATED DISCOVERY SYSTEMS**

14 Integrating Inductive and Deductive Reasoning for Data Mining  
*Evangelos Simoudis, Brian Livezey, and Randy Kerber* 353

15 Metaqueries for Data Mining  
*Wei-Min Shen, KayLiang Ong, Bharat Mitbander, and Carlo Zaniolo* 375
16 Exploration of the Power of Attribute-Oriented Induction in Data Mining  
Jiawei Han and Yongjian Fu

VI  NEXT GENERATION DATABASE SYSTEMS

17 Using Inductive Learning To Generate Rules for Semantic Query Optimization  
Chun-Nan Hsu and Craig A. Knoblock

18 Data Surveyor: Searching the Nuggets in Parallel  
Marcel Holsheimer, Martin L. Kersten, and Arno P.J.M. Siebes

VII  KDD APPLICATIONS

19 Automating the Analysis and Cataloging of Sky Surveys  
Usama M. Fayyad, S. George Djorgovski, and Nicholas Weir

20 Selecting and Reporting What is Interesting: The KEFIR Application to Healthcare Data  
Christopher J. Matheus, Gregory Piatetsky-Shapiro, and Dwight McNeill

21 Modeling Subjective Uncertainty in Image Annotation  
Padhraic Smyth, Michael C. Burl, Usama M. Fayyad, and Pietro Perona

22 Predicting Equity Returns from Securities Data with Minimal Rule Generation  
Chidanand Apte and Se June Hong
23 From Data Mining to Knowledge Discovery: Current Challenges and Future Directions
Ramasamy Uthurusamy 561

VIII Appendices

A Knowledge Discovery in Databases Terminology
Willi Klösgen and Jan M. Zytkow 573

B Data Mining and Knowledge Discovery
Internet Resources
Gregory Piatetsky-Shapiro 593

About The Editors 597

Index 601