Contents

Foreword vii
Preface xiii

1 From Data Mining to Knowledge Discovery: An Overview 1

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

I FOUNDATIONS

2 The Process of Knowledge Discovery in Databases: A Human-Centered Approach 37
Ronald J. Brachman and Tej Anand

3 Graphical Models for Discovering Knowledge 59
Wray Buntine

4 A Statistical Perspective on Knowledge Discovery in Databases 83
John Elder IV and Daryl Pregibon

II CLASSIFICATION AND CLUSTERING

5 Inductive Logic Programming and Knowledge Discovery in Databases 117
Sašo Džeroski

6 Bayesian Classification (AutoClass): Theory and Results 153
Peter Cheeseman and John Stutz

7 Discovering Informative Patterns and Data Cleaning 181
Isabelle Guyon, Nada Matic, and Vladimir Vapnik
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Author(s)</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>IV</td>
<td>Transforming Rules and Trees into Comprehensible Knowledge Structures</td>
<td>Brian R. Gaines</td>
<td>205</td>
</tr>
<tr>
<td>III</td>
<td>Trend and Deviation Analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Finding Patterns in Time Series: A Dynamic Programming Approach</td>
<td>Donald J. Berndt and James Clifford</td>
<td>229</td>
</tr>
<tr>
<td>10</td>
<td>Explora: A Multipattern and Multistrategy Discovery Assistant</td>
<td>Willi Klösgen</td>
<td>249</td>
</tr>
<tr>
<td>IV</td>
<td>Dependency Derivation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Bayesian Networks for Knowledge Discovery</td>
<td>David Heckerman</td>
<td>273</td>
</tr>
<tr>
<td>12</td>
<td>Fast Discovery of Association Rules</td>
<td>Rakesh Agrawal, Heikki Mannila, Ramakrishnan Srikant, Hannu Toivonen, and A. Inkeri Verkamo</td>
<td>307</td>
</tr>
<tr>
<td>13</td>
<td>From Contingency Tables to Various Forms of Knowledge in Databases</td>
<td>Robert Zembowicz and Jan M. Żytkow</td>
<td>329</td>
</tr>
<tr>
<td>V</td>
<td>Integrated Discovery Systems</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Integrating Inductive and Deductive Reasoning for Data Mining</td>
<td>Evangelos Simoudis, Brian Livezey, and Randy Kerber</td>
<td>353</td>
</tr>
<tr>
<td>15</td>
<td>Metaqueries for Data Mining</td>
<td>Wei-Min Shen, KayLiang Ong, Bharat Mitbander, and Carlo Zaniolo</td>
<td>375</td>
</tr>
</tbody>
</table>
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