

# Advances in Knowledge Discovery and Data Mining

**Edited by**

**Usama M. Fayyad**

*Jet Propulsion Laboratory, California Institute of Technology*

**Gregory Piatetsky-Shapiro**

*GTE Laboratories Incorporated*

**Padhraic Smyth**

*Jet Propulsion Laboratory, California Institute of Technology*

**& Ramasamy Uthurusamy**

*General Motors R & D Center*

AAAI Press / The MIT Press

Menlo Park, California • Cambridge, Massachusetts • London, England

# Contents

	<i>Foreword</i>	vii	
	<i>Preface</i>	xiii	
1	From Data Mining to Knowledge Discovery: An Overview  <i>Usama M. Fayyad, Gregory Piatetsky-Shapiro, and Padhraic Smyth</i>	1	X
I	FOUNDATIONS		
2	The Process of Knowledge Discovery in Databases: A Human-Centered Approach <i>Ronald J. Brachman and Tej Anand</i>	37	X
3	Graphical Models for Discovering Knowledge <i>Wray Buntine</i>	59	X
4	A Statistical Perspective on Knowledge Discovery in Databases <i>John Elder IV and Daryl Pregibon</i>	83	X
II	CLASSIFICATION AND CLUSTERING		
5	Inductive Logic Programming and Knowledge Discovery in Databases <i>Sašo Džeroski</i>	117	X
6	Bayesian Classification (AutoClass): Theory and Results <i>Peter Cheeseman and John Stutz</i>	153	X
7	Discovering Informative Patterns and Data Cleaning <i>Isabelle Guyon, Nada Matić, and Vladimir Vapnik</i>	181	X

8	Transforming Rules and Trees into Comprehensible Knowledge Structures <i>Brian R. Gaines</i>	205
III TREND AND DEVIATION ANALYSIS		
9	Finding Patterns in Time Series: A Dynamic Programming Approach <i>Donald J. Berndt and James Clifford</i>	229
10	Explora: A Multipattern and Multistrategy Discovery Assistant <i>Willi Klösgen</i>	249
IV DEPENDENCY DERIVATION		
11	Bayesian Networks for Knowledge Discovery <i>David Heckerman</i>	273
12	Fast Discovery of Association Rules <i>Rakesh Agrawal, Heikki Mannila, Ramakrishnan Srikant, Hannu Toivonen, and A. Inkeri Verkamo</i>	307
13	From Contingency Tables to Various Forms of Knowledge in Databases <i>Robert Zembowicz and Jan M. Żytkow</i>	329
V INTEGRATED DISCOVERY SYSTEMS		
14	Integrating Inductive and Deductive Reasoning for Data Mining <i>Evangelos Simoudis, Brian Livezey, and Randy Kerber</i>	353
15	Metaqueries for Data Mining <i>Wei-Min Shen, KayLiang Ong, Bharat Mitbander, and Carlo Zaniolo</i>	375

X	16	Exploration of the Power of Attribute-Oriented Induction in Data Mining <i>Jiawei Han and Yongjian Fu</i>	399	
VI		NEXT GENERATION DATABASE SYSTEMS		
	17	Using Inductive Learning To Generate Rules for Semantic Query Optimization <i>Chun-Nan Hsu and Craig A. Knoblock</i>	425	✓
	18	Data Surveyor: Searching the Nuggets in Parallel <i>Marcel Holsheimer, Martin L. Kersten, and Arno P.J.M. Siebes</i>	447	X
VII		KDD APPLICATIONS		
	19	Automating the Analysis and Cataloging of Sky Surveys <i>Usama M. Fayyad, S. George Djorgovski, and Nicholas Weir</i>	471	
	20	Selecting and Reporting What is Interesting: The KEFIR Application to Healthcare Data <i>Christopher J. Matheus, Gregory Piatetsky-Shapiro, and Dwight McNeill</i>	495	
	21	Modeling Subjective Uncertainty in Image Annotation <i>Padhraic Smyth, Michael C. Burl, Usama M. Fayyad, and Pietro Perona</i>	517	
	22	Predicting Equity Returns from Securities Data with Minimal Rule Generation <i>Chidanand Apte and Se June Hong</i>	541	

23	From Data Mining to Knowledge Discovery: Current Challenges and Future Directions <i>Ramasamy Uthurusamy</i>	561
----	--	-----

## VIII APPENDICES

A	Knowledge Discovery in Databases Terminology <i>Willi Klösgen and Jan M. Zytkow</i>	573
B	Data Mining and Knowledge Discovery Internet Resources <i>Gregory Piatetsky-Shapiro</i>	593
	About The Editors	597
	Index	601