

LIST OF FIGURES

Chapter 1

A COMMON LOGIC APPROACH TO DATA MINING AND PATTERN RECOGNITION, by A. Zakrevskij

	1
<i>Figure 1.</i> Using a Karnaugh Map to Find a Decision Boolean Function	3
<i>Figure 2.</i> Illustrating the Screening Effect	19
<i>Figure 3.</i> A Search Tree	25
<i>Figure 4.</i> The Energy Distribution of the Pronunciation of the Russian Word “ <i>nool</i> ” (meaning “zero”)	39

Chapter 2

THE ONE CLAUSE AT A TIME (OCAT) APPROACH TO DATA MINING AND

KNOWLEDGE DISCOVERY, by E. Triantaphyllou

	45
<i>Figure 1.</i> The One Clause At a Time Approach (for the CNF case)	59
<i>Figure 2.</i> Continuous Data for Illustrative Example and Extracted Sets of Classification Rules	63
<i>Figure 3.</i> The RA1 Heuristic [Deshpande and Triantaphyllou, 1998]	67
<i>Figure 4.</i> The Rejectability Graph for E^+ and E^-	74
<i>Figure 5.</i> The Rejectability Graph for the Second Illustrative Example	75
<i>Figure 6.</i> The Rejectability Graph for the new Sets E^+ and E^-	80

Chapter 3

AN INCREMENTAL LEARNING ALGORITHM FOR INFERRING LOGICAL RULES FROM EXAMPLES IN THE FRAMEWORK OF THE COMMON REASONING

PROCESS, by X. Naidenova

	89
<i>Figure 1.</i> Model of Reasoning. a) Under Pattern Recognition b) Under Learning	93
<i>Figure 2.</i> The Beginning of the Procedure for Inferring GMRTs	116
<i>Figure 3.</i> The Procedure for Determining the Set of Indices for Extending s	117
<i>Figure 4.</i> The Procedure for Generating All Possible Extensions of s	118
<i>Figure 5.</i> The Procedure for Analyzing the Set of Extensions of s	119
<i>Figure 6.</i> The Main Procedure NIAGaRa for inferring GMRTs	120
<i>Figure 7.</i> The Algorithm DIAGaRa	130
<i>Figure 8.</i> The Procedure for Generalizing the Existing GMRTs	133

<i>Figure 9.</i>	The Procedure for Preparing the Data for Inferring the GMRTs Contained in a New Example	134
<i>Figure 10.</i>	The Incremental Procedure INGOMAR	135

Chapter 4

DISCOVERING RULES THAT GOVERN MONOTONE PHENOMENA, by V.I. Torvik and E. Triantaphyllou		149
<i>Figure 1.</i>	Hierarchical Decomposition of the Breast Cancer Diagnosis Variables	156
<i>Figure 2.</i>	The Poset Formed by $\{0,1\}^4$ and the Relation \sim	157
<i>Figure 3.</i>	The Average Query Complexities for Problem 1	175
<i>Figure 4.</i>	The Average Query Complexities for Problem 2	177
<i>Figure 5.</i>	Increase in Query Complexities Due to Restricted Access to the Oracles	178
<i>Figure 6.</i>	Reduction in Query Complexity Due to the Nestedness Assumption	178
<i>Figure 7.</i>	Average Case Behavior of Various Selection Criteria for Problem 3	181
<i>Figure 8.</i>	The Restricted and Regular Maximum Likelihood Ratios Simulated with Expected $q = 0.2$ and $n = 3$	183

Chapter 5

LEARNING LOGIC FORMULAS AND RELATED ERROR DISTRIBUTIONS, by G. Felici, F. Sun, and K. Truemper		193
<i>Figure 1.</i>	Distributions for $Z=Z_A$ and $Z = Z_B$ and related Z'	217
<i>Figure 2.</i>	Estimated and verified F_A and G_B for Breast Cancer	218
<i>Figure 3.</i>	Estimated and verified F_A and G_B for Australian Credit Card	219
<i>Figure 4.</i>	Estimated and verified F_A and G_B for Congressional Voting	220
<i>Figure 5.</i>	Estimated and verified F_A and G_B for Diabetes	220
<i>Figure 6.</i>	Estimated and verified F_A and G_B for Heart Disease	221
<i>Figure 7.</i>	Estimated and verified F_A and G_B for Boston Housing	221

Chapter 6

FEATURE SELECTION FOR DATA MINING by V. de Angelis, G. Felici, and G. Mancinelli		227
<i>Figure 1.</i>	Wrappers and Filters	231

Chapter 7

TRANSFORMATION OF RATIONAL AND SET DATA TO LOGIC DATA, by S. Bartnikowski, M. Granberry, J. Mugan,	
---	--

Chapter 8

DATA FARMING: CONCEPTS AND METHODS, by A. Kusiak 279

<i>Figure 1.</i>	A Data Set with Five Features	284
<i>Figure 2.</i>	Rule Set Obtained from the Data Set in Figure 1	284
<i>Figure 3.</i>	Modified Data Set with Five Features	285
<i>Figure 4.</i>	Two Rules Generated from the Data Set of Figure 3	285
<i>Figure 5.</i>	(Part 1). Cross-validation Results: (a) Confusion Matrix for the Data Set in Figure 1, (b) Confusion Matrix for the Modified Data Set of Figure 3	286
<i>Figure 5.</i>	(Part 2). Cross-validation Results; (c) Classification Accuracy for the Data Set of Figure 1, (d) Classification Accuracy for the Data Set in Figure 3	287
<i>Figure 6.</i>	A Data Set with Four Features	288
<i>Figure 7.</i>	Transformed Data Set of Figure 6	288
<i>Figure 8.</i>	Cross Validation Results: (a) Average Classification Accuracy for the Data Set in Figure 6, (b) Average Classification Accuracy for the Transformed Data Set of Figure 7	289
<i>Figure 9.</i>	Data Set and the Corresponding Statistical Distributions	290
<i>Figure 10.</i>	Rule-Feature Matrix with Eight Rules	291
<i>Figure 11.</i>	Structured Rule-Feature Matrix	291
<i>Figure 12.</i>	Visual Representation of a Cluster of Two Rules	294
<i>Figure 13.</i>	A Data Set with Five Features	296
<i>Figure 14.</i>	Rules from the Data Set of Figure 13	296
<i>Figure 15.</i>	Rules Extracted from the Transformed Data Set of Figure 13	296
<i>Figure 16.</i>	Cross-validation Results: (a) Average Classification Accuracy for the Modified Data Set in Figure 13, (b) Average Classification Accuracy of the Data Set with Modified Outcome	297
<i>Figure 17.</i>	Average Classification Accuracy for the 599-Object Data Set	300
<i>Figure 18.</i>	Average Classification Accuracy for the 525-Object Data Set	301
<i>Figure 19.</i>	Average Classification Accuracy for the 525-Object Data Set with the Feature Sequence	301

Chapter 9

RULE INDUCTION THROUGH DISCRETE SUPPORT

VECTOR DECISION TREES, by C. Orsenigo and C. Vercellis	305
<i>Figure 1.</i> Margin Maximization for Linearly non Separable Sets	310
<i>Figure 2.</i> Axis Parallel Versus Oblique Splits	316

Chapter 10

MULTI-ATTRIBUTE DECISION TREES AND DECISION RULES, by J.-Y. Lee and S. Olafsson	327
<i>Figure 1.</i> The SODI Decision Tree Construction Algorithm	341
<i>Figure 2.</i> The SODI Rules for Pre-Pruning	343
<i>Figure 3.</i> Decision Trees Built by (a) ID3, and (b) SODI	345
<i>Figure 4.</i> Improvement of Accuracy Over ID3 for SODI, C4.5, and PART	348
<i>Figure 5.</i> Reduction in the Number of Decision Rules over ID3 for SODI, C4.5, and PART	349

Chapter 11

KNOWLEDGE ACQUISITION AND UNCERTAINTY IN FAULT DIAGNOSIS: A ROUGH SETS PERSPECTIVE, by L.-Y. Zhai, L.-P. Khoo, and S.-C. Fok	359
<i>Figure 1.</i> Knowledge Acquisition Techniques	362
<i>Figure 2.</i> Machine Learning Taxonomy	363
<i>Figure 3.</i> Processes for Knowledge Extraction	364
<i>Figure 4.</i> Basic Notions of Rough Set Theory for Illustrative Example	381
<i>Figure 5.</i> Framework of the RClass System	383

Chapter 12

DISCOVERING KNOWLEDGE NUGGETS WITH A GENETIC ALGORITHM, by E. Noda and A.A. Freitas	395
<i>Figure 1.</i> Pseudocode for a Genetic Algorithm at a High Level of Abstraction	400
<i>Figure 2.</i> An Example of Uniform Crossover in Genetic Algorithms	402
<i>Figure 3.</i> The Basic Idea of a Greedy Rule Induction Procedure	402
<i>Figure 4.</i> Attribute Interaction in a XOR (eXclusive OR) Function	403
<i>Figure 5.</i> Individual Representation	406
<i>Figure 6.</i> Examples of Condition Insertion/Removal Operations	411

Chapter 13

DIVERSITY MECHANISMS IN PITT-STYLE EVOLUTIONARY CLASSIFIER SYSTEMS, by M. Kirley,	
--	--

H.A. Abbass and R.I. McKay	433
<i>Figure 1.</i> Outline of a Simple Genetic Algorithm	437
<i>Figure 2.</i> The Island Model.	446

Chapter 14

FUZZY LOGIC IN DISCOVERING ASSOCIATION

RULES: AN OVERVIEW, by G. Chen, Q. Wei and E.E. Kerre	459
<i>Figure 1.</i> Fuzzy Sets Young(Y), Middle(M) and Old(O) with Y(20, 65), M(25, 32, 53, 60), O(20, 65)	468
<i>Figure 2.</i> Exact Taxonomies and Fuzzy Taxonomies	470
<i>Figure 3.</i> Part of a Linguistically Modified Fuzzy Taxonomic Structure	473
<i>Figure 4.</i> Static Matching Schemes	485

Chapter 15

MINING HUMAN INTERPRETABLE KNOWLEDGE WITH FUZZY MODELING METHODS: AN OVERVIEW,

by T.W. Liao	495
--------------	------------

Chapter 16

DATA MINING FROM MULTIMEDIA PATIENT RECORDS,

by A.S. Elmaghraby, M.M. Kantardzic, and M.P. Wachowiak	551
<i>Figure 1.</i> Phases of the Data Mining Process	555
<i>Figure 2.</i> Multimedia Components of the Patient Record	558
<i>Figure 3.</i> Phases in Labels and Noise Elimination for Digitized Mammography Images	567
<i>Figure 4.</i> The Difference Between PCA and ICA Transforms	570
<i>Figure 5.</i> Three EMG/ECG Mixtures (left) Separated into EMG and ECG Signals by ICA (right). Cardiac Artifacts in the EMG are Circled in Gray (upper left)	572
<i>Figure 6.</i> Sample of an Image of Size 5 x 5	577
<i>Figure 7.</i> Feature Extraction for the Image in Figure 6 by Using the Association Rules Method	578
<i>Figure 8.</i> Shoulder Scan	585
<i>Figure 9.</i> Parameter Maps: (a) INV (Nakagami Distribution); (b) TP (Nakagami Distribution); (c) SNR Values (K Distribution); (d) Fractional SNR (K Distribution)	587

Chapter 17

LEARNING TO FIND CONTEXT-BASED SPELLING

ERRORS, by H. Al-Mubaid and K. Truemper	597
---	------------

Chapter 18**INDUCTION AND INFERENCE WITH FUZZY RULES
FOR TEXTUAL INFORMATION RETRIEVAL, by J. Chen,
D.H. Kraft, M.J. Martin-Bautista, and M. –A., Vila****629****Chapter 19****STATISTICAL RULE INDUCTION IN THE PRESENCE OF
PRIOR INFORMATION: THE BAYESIAN RECORD
LINKAGE PROBLEM, by D.H. Judson****655***Figure 1.* File Processing Flowchart to Implement the Bayesian
Record Linkage

680

Figure 2. Posterior Kernel for $Y[1]$

683

Figure 3. Posterior Kernel for $Y[2]$

683

Figure 4. Posterior Kernel for $\mu[1,12]$

683

Figure 5. Posterior Kernel for $\mu[1,15]$

684

Chapter 20**FUTURE TRENDS IN SOME DATA MINING AREAS,
by X. Wang, P. Zhu, G. Felici, and E. Triantaphyllou****695***Figure 1.* Parallel Coordinate Visualization

704

Figure 2. Dense Pixel Displays

704

Figure 3. Dimensional Stacking Visualization

705