

# Table of Contents

|  |    |
|--|----|
| <b>1. The Alchemy of Intelligent IT (iIT): A Blueprint for Future Information Technology</b> |    |
| Ning Zhong and Jiming Liu .....  | 1  |
| 1.1 What Is iIT?.....  | 1  |
| 1.2 Why iIT? .....   | 1  |
| 1.3 iIT as the New Generation of Information Technology .....                                | 2  |
| 1.4 iIT for e-Business Intelligence .....  | 5  |
| 1.4.1 Virtual Industry Park: An Example of Enterprise Portals                                | 5  |
| 1.4.2 Web Mining and Farming .....   | 6  |
| 1.4.3 Semantic Social Networks for Intelligent Enterprise Portals .....                      | 7  |
| 1.4.4 Data Mining Grids for Web-Based Targeted Marketing                                     | 8  |
| 1.4.5 Wisdom Web-Based Computing .....   | 9  |
| 1.5 Dimensions of iIT Research .....   | 10 |
| 1.6 An Overview of This Book.....  | 11 |
| References .....   | 14 |

---

|  |  |
|--|--|
| <b>Part I. Emerging Data Mining Technology</b> |  |
|--|--|

---

|   |    |
|---|----|
| <b>2. Grid-Based Data Mining and Knowledge Discovery</b>  |    |
| Mario Cannataro, Antonio Congiusta, Carlo Mastroianni, Andrea Pugliese, Domenico Talia, and Paolo Trunfio ..... | 19 |
| 2.1 Introduction .....  | 19 |
| 2.2 Knowledge Discovery on Grids .....  | 21 |
| 2.3 The <i>Knowledge Grid</i> Architecture .....  | 26 |
| 2.3.1 <i>Knowledge Grid</i> Services .....  | 26 |
| 2.4 An XML-Based Metadata Model for the <i>Knowledge Grid</i> ....  | 28 |
| 2.4.1 Data Mining Software.....   | 29 |
| 2.4.2 Data Sources .....  | 30 |
| 2.4.3 Abstract and Concrete Resources.....  | 32 |
| 2.4.4 Execution Plans.....  | 32 |
| 2.4.5 Data Mining Models .....  | 33 |
| 2.4.6 Metadata Management .....   | 33 |

|           |  |     |
|-----------|--|-----|
| 2.5       | Design of a PDKD Computation .....                                 | 34  |
| 2.5.1     | Task Composition .....   | 36  |
| 2.5.2     | Task Consistency Checking .....                                    | 38  |
| 2.5.3     | Execution Plan Generation .....                                    | 39  |
| 2.6       | Execution of a PDKD Computation .....                              | 42  |
| 2.7       | Conclusions .....  | 43  |
|           | References .....   | 44  |
| <b>3.</b> | <b>The MiningMart Approach to Knowledge Discovery in Databases</b> |     |
|           | Katharina Morik and Martin Scholz .....                            | 47  |
| 3.1       | Introduction: Acquiring Knowledge from Existing Databases .        | 47  |
| 3.2       | The MiningMart Approach .....                                      | 51  |
| 3.2.1     | The Meta-Model of Metadata M4 .....                                | 52  |
| 3.2.2     | Editing the Conceptual Data Model .....                            | 54  |
| 3.2.3     | Editing the Relational Model .....                                 | 55  |
| 3.2.4     | The Case and Its Compiler .....                                    | 56  |
| 3.3       | The Case Base .....  | 59  |
| 3.4       | Conclusions .....  | 61  |
|           | References .....   | 64  |
| <b>4.</b> | <b>Ensemble Methods and Rule Generation</b>                        |     |
|           | Yongdai Kim, Jinseog Kim, and Jongwoo Jeon.....                    | 67  |
| 4.1       | Introduction .....   | 67  |
| 4.2       | Ensemble Algorithms: A Review.....                                 | 69  |
| 4.2.1     | Bagging .....  | 69  |
| 4.2.2     | Boosting .....   | 70  |
| 4.2.3     | Convex Hull Ensemble Machine: CHEM.....                            | 75  |
| 4.2.4     | Comments on Roles of Base Learners .....                           | 80  |
| 4.3       | Rule Generation.....   | 82  |
| 4.4       | Illustration .....   | 84  |
| 4.5       | Conclusions.....   | 86  |
|           | References .....   | 86  |
| <b>5.</b> | <b>Evaluation Scheme for Exception Rule/Group Discovery</b>        |     |
|           | Einoshin Suzuki .....  | 89  |
| 5.1       | Introduction .....   | 89  |
| 5.2       | Exception Rule/Group Discovery .....                               | 90  |
| 5.2.1     | Notice .....   | 90  |
| 5.2.2     | Use of Domain Knowledge .....                                      | 91  |
| 5.2.3     | Hypothesis-Driven Discovery .....                                  | 95  |
| 5.3       | Evaluation Scheme and Classification of Existing Methods ...       | 102 |
| 5.3.1     | Generality .....   | 102 |

|       |  |     |
|-------|--|-----|
| 5.3.2 | Break of Monotonicity in Discovered Patterns . . . . . | 103 |
| 5.3.3 | Evaluation of Reliability . . . . .                    | 103 |
| 5.3.4 | Search Range . . . . .                                 | 104 |
| 5.3.5 | Interpretation of the Evaluation Measure . . . . .     | 104 |
| 5.3.6 | Use of Domain Knowledge . . . . .                      | 105 |
| 5.3.7 | Successes in Real Applications . . . . .               | 105 |
| 5.4   | Conclusions . . . . .                                  | 106 |
|       | References . . . . .                                   | 106 |

## **6. Data Mining for Targeted Marketing**

|  |  |     |
|--|--|-----|
| Ning Zhong, Yiyu Yao, Chunnian Liu, Jiajin Huang, and Chuangxin Ou . . . . . |  | 109 |
| 6.1  | Introduction . . . . .                                   | 109 |
| 6.2  | The Process of Targeted Marketing . . . . .              | 110 |
| 6.3  | Problems in Targeted Marketing . . . . .                 | 110 |
| 6.4  | Target Selection Algorithms . . . . .                    | 112 |
| 6.4.1  | Segmentation Models . . . . .                            | 112 |
| 6.4.2  | Response Models . . . . .                                | 112 |
| 6.5  | Mining Market Value Functions . . . . .                  | 118 |
| 6.5.1  | Utility Functions . . . . .                              | 119 |
| 6.5.2  | Attribute Weighting . . . . .                            | 121 |
| 6.6  | Evaluation of the Learning Algorithms . . . . .          | 123 |
| 6.7  | Experimental Results . . . . .                           | 124 |
| 6.8  | New Directions . . . . .                                 | 125 |
| 6.8.1  | Web-Based Targeted Marketing . . . . .                   | 125 |
| 6.8.2  | Multi-Aspect Analysis in Multiple Data Sources . . . . . | 127 |
| 6.8.3  | Building a Data Mining Grid . . . . .                    | 127 |
| 6.9  | Conclusions . . . . .                                    | 128 |
|  | References . . . . .                                     | 129 |

---

## **Part II. Data Mining for Web Intelligence**

---

## **7. Mining for Information Discovery on the Web: Overview and Illustrative Research**

|   |  |     |
|---|--|-----|
| Hwanjo Yu, AnHai Doan, and Jiawei Han . . . . . |  | 135 |
| 7.1   | Introduction . . . . .                                   | 135 |
| 7.2   | Finding Information on the Web . . . . .                 | 136 |
| 7.2.1   | Exploring and Navigating the Web . . . . .               | 136 |
| 7.2.2   | Querying with Information Processing Systems . . . . .   | 138 |
| 7.3   | Web Page Classification from Positive Examples . . . . . | 141 |
| 7.3.1   | Related Work . . . . .                                   | 142 |
| 7.3.2   | SVM Margin-Maximization Property . . . . .               | 143 |
| 7.3.3   | The Mapping-Convergence (M-C) Algorithm . . . . .        | 144 |

|           |   |     |
|-----------|---|-----|
| 7.3.4     | Experimental Results . . . . .  | 148 |
| 7.4       | Object Matching Across Disparate Data Sources . . . . .                     | 153 |
| 7.4.1     | Problem Definition . . . . .  | 156 |
| 7.4.2     | The PROM Approach . . . . .   | 156 |
| 7.4.3     | Empirical Evaluation . . . . .  | 158 |
| 7.4.4     | Related Work . . . . .  | 161 |
| 7.4.5     | Summary . . . . .   | 162 |
| 7.5       | Conclusions . . . . .   | 163 |
|           | References . . . . .  | 163 |
| <b>8.</b> | <b>Mining Web Logs for Actionable Knowledge</b>                             |     |
|           | Qiang Yang, Charles X. Ling, and Jianfeng Gao . . . . .                     | 169 |
| 8.1       | Introduction . . . . .  | 169 |
| 8.2       | Web Log Mining for Prefetching . . . . .                                    | 170 |
| 8.2.1     | Data Cleaning on Web Log Data . . . . .                                     | 170 |
| 8.2.2     | Mining Web Logs for Path Profiles . . . . .                                 | 171 |
| 8.2.3     | Web Object Prediction . . . . .   | 171 |
| 8.2.4     | Learning to Prefetch Web Documents . . . . .                                | 173 |
| 8.3       | Web Page Clustering for Intelligent User Interfaces . . . . .               | 175 |
| 8.4       | Web Query Log Mining . . . . .  | 178 |
| 8.4.1     | Web Query Logs . . . . .  | 178 |
| 8.4.2     | Mining Generalized Query Patterns . . . . .                                 | 181 |
| 8.4.3     | A Bottom-Up Generalization Algorithm . . . . .                              | 182 |
| 8.4.4     | Improvement 1: A Hierarchy over Keywords . . . . .                          | 183 |
| 8.4.5     | Improvement 2: Flexible Generalizations . . . . .                           | 184 |
| 8.4.6     | Improvement 3: Morphology Conversion . . . . .                              | 185 |
| 8.4.7     | Improvement 4: Synonym Conversion . . . . .                                 | 185 |
| 8.4.8     | Implementations . . . . .   | 186 |
| 8.4.9     | Simulation Experiments . . . . .  | 187 |
| 8.4.10    | Analyses of the Results . . . . .   | 187 |
| 8.4.11    | Relation to Previous Work . . . . .   | 188 |
| 8.5       | Conclusions . . . . .   | 189 |
|           | References . . . . .  | 190 |
| <b>9.</b> | <b>Discovery of Web Robot Sessions Based on Their Navigational Patterns</b> |     |
|           | Pang-Ning Tan and Vipin Kumar . . . . .                                     | 193 |
| 9.1       | Introduction . . . . .  | 193 |
| 9.2       | Web Robot Detection: Overview . . . . .                                     | 197 |
| 9.2.1     | Limitations of Current Robot Detection Techniques . . . . .                 | 197 |
| 9.2.2     | Motivation for Proposed Robot Detection Technique . . . . .                 | 200 |
| 9.3       | Methodology . . . . .   | 202 |
| 9.3.1     | Data Source and Preprocessing . . . . .                                     | 202 |
| 9.3.2     | Feature Vector Construction . . . . .                                       | 204 |

|            |  |     |
|------------|--|-----|
| 9.3.3      | Session Labeling .....   | 207 |
| 9.3.4      | Classification .....   | 209 |
| 9.3.5      | Identifying Mislabeled Sessions .....  | 210 |
| 9.4        | Experimental Evaluation .....  | 211 |
| 9.4.1      | Experimental Data Set .....  | 211 |
| 9.4.2      | Correlation Analysis .....   | 212 |
| 9.4.3      | Classifier Performance .....   | 215 |
| 9.4.4      | Finding Mislabeled Data .....  | 217 |
| 9.5        | Conclusions .....  | 219 |
|            | References .....   | 221 |
| <b>10.</b> | <b>Web Ontology Learning and Engineering: An Integrated Approach</b>           |     |
|            | Roberto Navigli, Paola Velardi, and Michele Missikoff .....                    | 223 |
| 10.1       | Introduction .....   | 223 |
| 10.2       | The OntoLearn System .....   | 227 |
| 10.2.1     | Identification of Relevant Domain Terminology .....                            | 228 |
| 10.2.2     | Semantic Interpretation of Terms .....   | 230 |
| 10.2.3     | Creating a Specific Domain Ontology .....                                      | 234 |
| 10.2.4     | Creating a Domain Ontology from WordNet .....                                  | 237 |
| 10.3       | Evaluation of the OntoLearn System .....                                       | 237 |
| 10.4       | Conclusions .....  | 239 |
|            | References .....   | 241 |
| <b>11.</b> | <b>Browsing Semi-Structured Texts on the Web Using Formal Concept Analysis</b> |     |
|            | Richard Cole, Florence Amardeilh, and Peter Eklund .....                       | 243 |
| 11.1       | Introduction: Information Extraction and the Web .....                         | 243 |
| 11.1.1     | Definitions – Information Extraction .....                                     | 244 |
| 11.1.2     | Web Documents and Text Diversity .....   | 244 |
| 11.1.3     | Architecture and Components .....  | 245 |
| 11.1.4     | Related Work .....   | 246 |
| 11.1.5     | The Interaction Paradigm and Learning Context .....                            | 247 |
| 11.2       | Formal Concept Analysis and RDBMSs .....                                       | 248 |
| 11.3       | Web-Robot for Extracting Structured Data from Unstructured Sources .....       | 252 |
| 11.4       | RFCA – the Web-Based FCA Interface .....                                       | 255 |
| 11.5       | Reusing CEM for Nesting and Zooming .....                                      | 259 |
| 11.6       | Conclusions .....  | 262 |
|            | References .....   | 263 |

|  |     |
|--|-----|
| <b>12. Graph Discovery and Visualization from Textual Data</b> |     |
| Vincent Dubois and Mohamed Quafafou .....                      | 265 |
| 12.1 Introduction .....  | 265 |
| 12.2 Graph Structure Discovery .....                           | 266 |
| 12.2.1 Graph Structure Semantic .....                          | 266 |
| 12.2.2 Network Evaluation .....                                | 267 |
| 12.2.3 Methods and Algorithms .....                            | 268 |
| 12.3 Visualization Constrains Discovery .....                  | 269 |
| 12.3.1 Graph Structure Visualization .....                     | 269 |
| 12.3.2 Incremental and Dynamic Properties .....                | 271 |
| 12.4 Experimental Results .....                                | 272 |
| 12.4.1 Results on Corpus .....                                 | 272 |
| 12.4.2 Convergence Problem .....                               | 274 |
| 12.4.3 Effect of Initial Position on Result .....              | 275 |
| 12.4.4 Effect of New Data Insertion .....                      | 275 |
| 12.4.5 Influence of Data Order .....                           | 276 |
| 12.5 Application: Web Content Mining .....                     | 277 |
| 12.5.1 Introduction .....                                      | 277 |
| 12.5.2 General Architecture .....                              | 281 |
| 12.5.3 Related Work .....                                      | 282 |
| 12.5.4 Dynamic Search and Crawling .....                       | 283 |
| 12.5.5 Integration and Implementation Issues .....             | 284 |
| 12.5.6 Experimental Results .....                              | 285 |
| 12.6 Conclusions .....   | 287 |
| References .....   | 287 |

---

## Part III. Emerging Agent Technology

---

|  |     |
|--|-----|
| <b>13. Agent Networks: Topological and Clustering Characterization</b> |     |
| Xiaolong Jin and Jiming Liu .....                                      | 291 |
| 13.1 Introduction .....  | 291 |
| 13.1.1 Small World Phenomena .....                                     | 291 |
| 13.1.2 Agent Networks .....  | 291 |
| 13.1.3 Satisfiability Problems .....                                   | 292 |
| 13.1.4 Problem Statements .....  | 293 |
| 13.1.5 Organization of This Chapter .....                              | 293 |
| 13.2 Topologies of Agent Networks .....                                | 293 |
| 13.2.1 Representation I .....  | 294 |
| 13.2.2 Representation II .....   | 297 |
| 13.3 Discussions .....   | 299 |
| 13.3.1 Complexities in Different Representations .....                 | 299 |

|            |  |     |
|------------|--|-----|
| 13.3.2     | Balanced Complexities in Intra- and Inter-Agent Computations . . . . .                   | 300 |
| 13.3.3     | A Guiding Principle . . . . .  | 300 |
| 13.4       | Average Value of the Clustering Coefficient . . . . .                                    | 301 |
| 13.5       | Lower Bounds of the Clustering Coefficient . . . . .                                     | 303 |
| 13.6       | Conclusions . . . . .  | 306 |
|            | References . . . . .   | 308 |
| <b>14.</b> | <b>Finding the Best Agents for Cooperation</b>   |     |
|            | Francesco Buccafurri, Luigi Palopoli, Domenico Rosaci, and Giuseppe M.L. Sarnè . . . . . | 311 |
| 14.1       | Introduction . . . . .   | 311 |
| 14.2       | Related Work . . . . .   | 312 |
| 14.3       | The Knowledge Bases . . . . .  | 314 |
| 14.3.1     | An Ontology for Describing the Domain of Interest . . . . .                              | 314 |
| 14.3.2     | The Local Knowledge Base . . . . .   | 315 |
| 14.4       | Extraction of the Semantic Properties . . . . .  | 315 |
| 14.4.1     | Local Properties: Similarity . . . . .   | 316 |
| 14.4.2     | Global Properties: Interest and Attractiveness . . . . .                                 | 317 |
| 14.4.3     | Choice Lists . . . . .   | 321 |
| 14.4.4     | Reactive Properties . . . . .  | 322 |
| 14.5       | Local Knowledge Base Integration . . . . .   | 324 |
| 14.6       | The SPY System Architecture . . . . .  | 325 |
| 14.7       | Experiments . . . . .  | 329 |
| 14.8       | Conclusions . . . . .  | 330 |
|            | References . . . . .   | 330 |
| <b>15.</b> | <b>Constructing Hybrid Intelligent Systems for Data Mining from Agent Perspectives</b>   |     |
|            | Zili Zhang and Chengqi Zhang . . . . .   | 333 |
| 15.1       | Introduction . . . . .   | 333 |
| 15.2       | Data Mining Needs Hybrid Solutions . . . . .   | 335 |
| 15.3       | Agent Perspectives Are Suitable for Modeling Hybrid Intelligent Systems . . . . .        | 337 |
| 15.4       | Agent-Oriented Methodologies . . . . .   | 338 |
| 15.4.1     | The Role Model . . . . .   | 339 |
| 15.4.2     | The Interaction Model . . . . .  | 342 |
| 15.4.3     | Organizational Rules . . . . .   | 343 |
| 15.4.4     | The Agent Model . . . . .  | 343 |
| 15.4.5     | The Skill Model . . . . .  | 344 |
| 15.4.6     | The Knowledge Model . . . . .  | 344 |
| 15.4.7     | Organizational Structures and Patterns . . . . .   | 345 |
| 15.5       | Agent-Based Hybrid Systems for Data Mining . . . . .                                     | 346 |
| 15.5.1     | Analysis and Design . . . . .  | 346 |

|  |     |
|--|-----|
| 15.5.2 Implementation .....  | 352 |
| 15.5.3 Case Study .....  | 355 |
| 15.6 Evaluation .....  | 357 |
| 15.7 Conclusions .....   | 357 |
| References .....   | 358 |
| <br><b>16. Making Agents Acceptable to People</b>  |     |
| Jeffrey M. Bradshaw, Patrick Beaument, Maggie R. Breedy, Larry Bunch, Sergey V. Drakunov, Paul J. Feltovich, Robert R. Hoffman, Renia Jeffers, Matthew Johnson, Shriniwas Kulkarni, James Lott, Anil K. Raj, Niranjani Suri, and Andrzej Uszok ..... | 361 |
| 16.1 Introduction .....  | 361 |
| 16.2 Addressing Agent Acceptability Through the Use of Policy ..   | 367 |
| 16.2.1 What Is Policy? .....   | 368 |
| 16.2.2 Distinguishing Policy from Related Concepts .....   | 369 |
| 16.2.3 Types of Policy .....   | 370 |
| 16.2.4 Autonomy and Policy .....   | 371 |
| 16.2.5 Benefits of Policy Management .....   | 375 |
| 16.2.6 Applications of Policy Using KAoS and Nomads .....  | 377 |
| 16.3 Technical Aspects of Agent Acceptability .....  | 379 |
| 16.3.1 Examples of Policy Types Relating to Technical Aspects of Agent Acceptability .....   | 380 |
| 16.4 Social Aspects of Agent Acceptability .....   | 386 |
| 16.4.1 Examples of Policy Types Relating to Social Aspects of Agent Acceptability .....  | 387 |
| 16.4.2 Cognitive and Robotic Prostheses .....  | 394 |
| 16.5 Conclusions .....   | 398 |
| References .....   | 398 |

---

**Part IV. Emerging Soft Computing Technology**

---

|   |     |
|---|-----|
| <b>17. Constraint-Based Neural Network Learning for Time Series Predictions</b>   |     |
| Benjamin W. Wah and Minglun Qian .....  | 409 |
| 17.1 Introduction .....   | 409 |
| 17.2 Previous Work in Time Series Modeling .....                                  | 410 |
| 17.2.1 Linearity .....  | 411 |
| 17.2.2 Piecewise Chaos .....  | 412 |
| 17.2.3 Random Noise .....   | 412 |
| 17.2.4 Artificial Neural Networks .....   | 415 |
| 17.3 Predictions of Noise-Free Stationary and Piecewise Chaotic Time Series ..... | 417 |
| 17.3.1 Recurrent FIR Neural Networks .....  | 417 |



|            |  |     |
|------------|--|-----|
| 17.3.2     | Constrained Formulations for ANN Learning . . . . .            | 418 |
| 17.3.3     | Violation-Guided Backpropagation Algorithm . . . . .           | 420 |
| 17.3.4     | Experimental Results . . . . .                                 | 422 |
| 17.4       | Predictions of Noisy Time Series with High Frequency           |     |
|            | Random Noise . . . . .   | 423 |
| 17.4.1     | Review on Financial Time Series Predictions . . . . .          | 424 |
| 17.4.2     | Constraint in the Lag Period . . . . .                         | 424 |
| 17.4.3     | Experimental Results . . . . .                                 | 425 |
| 17.5       | Conclusions . . . . .  | 428 |
|            | References . . . . .   | 429 |
| <b>18.</b> | <b>Approximate Reasoning in Distributed Environments</b>       |     |
|            | Andrzej Skowron . . . . .                                      | 433 |
| 18.1       | Introduction . . . . .   | 433 |
| 18.2       | Information Granules . . . . .                                 | 439 |
| 18.2.1     | Rough Sets and Approximation Spaces . . . . .                  | 439 |
| 18.2.2     | Syntax and Semantics of Information Granules . . . . .         | 441 |
| 18.2.3     | Granule Inclusion and Closeness . . . . .                      | 447 |
| 18.2.4     | Rough-Fuzzy Granules . . . . .                                 | 451 |
| 18.2.5     | Classifiers as Information Granules . . . . .                  | 451 |
| 18.3       | Rough-Neural Computing: Weights Defined by                     |     |
|            | Approximation Spaces . . . . .                                 | 452 |
| 18.4       | Rough-Neural Computing: Rough Mereological Approach . . . . .  | 457 |
| 18.4.1     | Distributed Systems of Agents . . . . .                        | 458 |
| 18.4.2     | Approximate Synthesis of Complex Objects . . . . .             | 461 |
| 18.5       | Extracting <i>AR</i> -Schemes from Data and Background         |     |
|            | Knowledge . . . . .  | 464 |
| 18.5.1     | Granule Decomposition . . . . .                                | 465 |
| 18.6       | Conclusions . . . . .  | 469 |
|            | References . . . . .   | 472 |
| <b>19.</b> | <b>Soft Computing Pattern Recognition, Data Mining and</b>     |     |
|            | <b>Web Intelligence</b>  |     |
|            | Sankar K. Pal, Sushmita Mitra, and Pabitra Mitra . . . . .     | 475 |
| 19.1       | Introduction . . . . .   | 475 |
| 19.2       | Soft Computing Pattern Recognition . . . . .                   | 478 |
| 19.2.1     | Relevance of Fuzzy Set Theory in Pattern Recognition . . . . . | 478 |
| 19.2.2     | Relevance of Neural Network Approaches . . . . .               | 479 |
| 19.2.3     | Genetic Algorithms for Pattern Recognition . . . . .           | 481 |
| 19.2.4     | Relevance of Rough Sets . . . . .                              | 481 |
| 19.2.5     | Integration and Hybrid Systems . . . . .                       | 482 |
| 19.3       | Knowledge Discovery and Data Mining . . . . .                  | 485 |
| 19.3.1     | Data Mining . . . . .  | 485 |
| 19.4       | Soft Computing for Data Mining . . . . .                       | 486 |

|  |     |
|--|-----|
| 19.4.1 Fuzzy Sets .....  | 486 |
| 19.4.2 Neural Networks .....                                     | 489 |
| 19.4.3 Neuro-Fuzzy Computing .....                               | 490 |
| 19.4.4 Genetic Algorithms .....                                  | 491 |
| 19.4.5 Rough Sets .....  | 492 |
| 19.5 Web Mining .....  | 492 |
| 19.5.1 Web Mining Components and the Methodologies .....         | 493 |
| 19.6 Soft Computing for Web Mining .....                         | 494 |
| 19.6.1 Fuzzy Logic for Web Mining .....                          | 494 |
| 19.6.2 Neural Networks and Learning Systems for Web Mining ..... | 497 |
| 19.6.3 Genetic Algorithms for Web Mining .....                   | 502 |
| 19.6.4 Rough Sets for Web Mining .....                           | 503 |
| 19.7 Conclusions .....   | 505 |
| References .....   | 506 |

## **20. Dominance-Based Rough Set Approach to Knowledge Discovery (I): General Perspective**

|   |     |
|---|-----|
| Salvatore Greco, Benedetto Matarazzo, and Roman Slowinski .....   | 513 |
| 20.1 Introduction: Three Types of Prior Knowledge to Be Included in Knowledge Discovery .....                 | 513 |
| 20.2 The Influence of Preference Order in Data on Granular Computing .....                                    | 516 |
| 20.3 Dominance-Based Rough Set Approach (DRSA) .....  | 518 |
| 20.3.1 Granular Computing with Dominance Cones .....  | 518 |
| 20.3.2 Induction of Decision Rules .....  | 523 |
| 20.3.3 Illustrative Example .....   | 525 |
| 20.4 DRSA for Multicriteria Choice and Ranking .....  | 529 |
| 20.4.1 Pairwise Comparison Table (PCT) as a Preferential Information and a Learning Sample .....              | 530 |
| 20.4.2 Rough Approximation of Outranking and Non-outranking Relations Specified in PCT .....                  | 532 |
| 20.4.3 Induction of Decision Rules from Rough Approximations of Outranking and Non-outranking Relations ..... | 534 |
| 20.4.4 Use of Decision Rules for Decision Support .....   | 536 |
| 20.4.5 Illustrative Example .....   | 537 |
| 20.4.6 Summary .....  | 539 |
| 20.5 DRSA with Missing Values of Attributes and Criteria .....  | 540 |
| 20.5.1 Generalized Indiscernibility Relation .....  | 540 |
| 20.5.2 Illustrative Example .....   | 542 |
| 20.5.3 Generalized Dominance Relation .....   | 543 |
| 20.5.4 Illustrative Example .....   | 546 |
| 20.6 Conclusions .....  | 547 |
| References .....  | 548 |

|   |     |
|---|-----|
| <b>21. Dominance-Based Rough Set Approach to Knowledge Discovery (II): Extensions and Applications</b>              |     |
| Salvatore Greco, Benedetto Matarazzo, and Roman Slowinski . . . . .   | 553 |
| 21.1 Introduction . . . . .   | 553 |
| 21.2 Fuzzy Set Extensions of DRSA . . . . .   | 553 |
| 21.2.1 Fuzzy DRSA for Multicriteria Classification . . . . .  | 554 |
| 21.2.2 Fuzzy DRSA for Multicriteria Choice and Ranking . . .  | 560 |
| 21.2.3 Gradual Rules and Fuzzy Rough Approximations<br>Without Fuzzy Logical Connectives . . . . .                  | 564 |
| 21.3 DRSA for Decision Under Risk . . . . .   | 575 |
| 21.3.1 DRSA Based on Stochastic Dominance . . . . .   | 576 |
| 21.3.2 Illustrative Example . . . . .   | 577 |
| 21.4 DRSA for Hierarchical Decision Making . . . . .  | 580 |
| 21.4.1 Rough Set Approach for Attribute Subset Values and<br>Interval Order . . . . .                               | 582 |
| 21.4.2 Propagation of Inconsistencies and Application of<br>Decision Rules . . . . .                                | 587 |
| 21.4.3 Illustrative Example . . . . .   | 589 |
| 21.5 Comparison of DRSA with Other Paradigms . . . . .  | 596 |
| 21.5.1 Axiomatic Foundations of Multicriteria Classification<br>Problems and Associated Preference Models . . . . . | 596 |
| 21.5.2 Conjoint Measurement for Multicriteria Classification<br>Problems with Inconsistencies . . . . .             | 605 |
| 21.5.3 Summary . . . . .  | 606 |
| 21.6 Conclusions . . . . .  | 606 |
| References . . . . .  | 607 |

---

## Part V. Statistical Learning Theory

---

|  |     |
|--|-----|
| <b>22. Bayesian Ying Yang Learning (I): A Unified Perspective for Statistical Modeling</b> |     |
| Lei Xu . . . . .   | 615 |
| 22.1 Introduction: Basic Issues of Statistical Learning . . . . .                          | 615 |
| 22.2 Dependence Among Samples from One-Object World . . . . .                              | 616 |
| 22.3 Dependence Among Samples from a Multi-Object World . . . .                            | 620 |
| 22.3.1 Dependence Among Samples from a Multi-Object<br>World . . . . .                     | 620 |
| 22.3.2 Mining Dependence Structure Across Invisible<br>Multi-Object . . . . .              | 621 |
| 22.4 A Systemic View on Various Dependence Structures . . . . .                            | 623 |
| 22.5 Bayesian Ying Yang System . . . . .   | 627 |
| 22.6 BYY Harmony Learning . . . . .  | 631 |

|        |  |     |
|--------|--|-----|
| 22.6.1 | Kullback Divergence, Harmony Measure, and<br>Z-Regularization .....                  | 631 |
| 22.6.2 | BY Y Harmony Learning .....  | 634 |
| 22.6.3 | A Further Extension: From $\ln(r)$ to Convex Function .                              | 637 |
| 22.7   | Ying-Yang Alternative Procedure for Parameter Learning ...                           | 640 |
| 22.8   | Learning Implementation: From Optimization Search to<br>Accumulation Consensus ..... | 643 |
| 22.9   | Main Results and Bibliographic Remarks .....   | 647 |
| 22.9.1 | Main Results on Typical Learning Problems .....                                      | 647 |
| 22.9.2 | Bibliographic Remarks on BY Y System with KL<br>Learning .....                       | 650 |
| 22.9.3 | Bibliographic Remarks on Computing Techniques ...                                    | 654 |
| 22.10  | Conclusions .....  | 654 |
|        | References .....   | 655 |

## **23. Bayesian Ying Yang Learning (II): A New Mechanism for Model Selection and Regularization**

|   |     |
|---|-----|
| Lei Xu . . . . .  | 661 |
| 23.1 Introduction: A Key Challenge and Existing Solutions . . . . .                                     | 661 |
| 23.2 Existing Solutions . . . . .   | 663 |
| 23.2.1 Efforts in the First Stream . . . . .  | 663 |
| 23.2.2 Efforts in the Second Stream . . . . .   | 664 |
| 23.3 Bayesian Ying Yang Harmony Learning . . . . .  | 667 |
| 23.3.1 Bayesian Ying Yang Harmony Learning . . . . .  | 668 |
| 23.3.2 Structural Inner Representations . . . . .   | 672 |
| 23.4 Regularization Versus Model Selection . . . . .  | 675 |
| 23.4.1 ML, HL, and Z-Regularization . . . . .   | 675 |
| 23.4.2 KL- $\lambda$ -HL Spectrum . . . . .   | 678 |
| 23.5 An Information Transfer Perspective . . . . .  | 681 |
| 23.6 BYY Harmony Learning Versus Related Approaches . . . . .   | 684 |
| 23.6.1 Relation and Difference to the Bits-Back Based MDL<br>and Bayesian Approaches . . . . .          | 684 |
| 23.6.2 Relations to Information Geometry, Helmholtz<br>Machine and Variational Approximation . . . . .  | 686 |
| 23.6.3 A Projection Geometry Perspective . . . . .  | 688 |
| 23.7 Bibliographic Remarks . . . . .  | 693 |
| 23.7.1 On BYY Harmony Learning (I): Model Selection<br>Criteria vs. Automatic Model Selection . . . . . | 693 |
| 23.7.2 On BYY Harmony Learning (II): Model Selection<br>Criteria . . . . .                              | 694 |
| 23.7.3 On BYY Harmony Learning (III): Automatic Model<br>Selection . . . . .                            | 696 |
| 23.7.4 On Regularization Methods . . . . .  | 698 |
| 23.8 Conclusions . . . . .  | 700 |

References ..... 700

**Author Index** ..... 707

**Subject Index** ..... 709

Intelligent Technologies for Information Analysis

Zhong, N.; Liu, J. (Eds.)

2004, XXVII, 711 p., Hardcover

ISBN: 978-3-540-40677-8