Emerging Paradigms in Machine Learning
Contents

1 Emerging Paradigms in Machine Learning: An Introduction
   Sheela Ramanna, Lakhmi C. Jain, Robert J. Howlett
   1.1 Introduction ............................................. 1
   1.2 Chapters of the Book .................................... 4
   1.3 Concluding Remarks .................................... 7

Part A: Foundations

2 Extensions of Dynamic Programming as a New Tool for Decision
   Tree Optimization ............................................ 11
   Abdulaziz Alkhalid, Igor Chikalov, Shahid Hussain, Mikhail Moshkov
   2.1 Introduction .............................................. 11
   2.2 Basic Notions ............................................. 12
      2.2.1 Decision Tables and Trees .......................... 13
      2.2.2 Cost Functions ..................................... 14
   2.3 Representation of Sets of α-Decision Trees and Decision Trees .... 15
   2.4 Optimization of α-Decision Trees ........................ 18
      2.4.1 Proper Subgraphs of Graph \( \Delta_\alpha(T) \) ............. 18
      2.4.2 Procedure of Optimization .......................... 18
      2.4.3 Possibilities of Sequential Optimization .............. 20
      2.4.4 Experimental Results ................................ 21
   2.5 Relationships between Depth and Number of Misclassifications ... 24
      2.5.1 Computing the Relationships ......................... 24
      2.5.2 Experimental Results ................................ 26
   2.6 Conclusions ............................................... 27

References .................................................... 28
3 Optimised Information Abstraction in Granular Min/Max Clustering
Andrzej Bargiela, Witold Pedrycz
3.1 Introductory Comments ............................................. 31
3.2 Granular Information in Systems Modeling ..................... 35
3.3 Information Density Based Granulation ......................... 36
3.4 Granular Representatives of Data ................................. 40
3.5 Granular Refinement of Prototypes ............................... 44
3.6 Conclusions .......................................................... 47
References ............................................................... 47

4 Mining Incomplete Data—A Rough Set Approach ............... 49
Jerzy W. Grzymala-Busse, Zdzislaw S. Hippe
4.1 Introduction .......................................................... 49
4.2 Blocks of Attribute-Value Pairs ................................. 51
4.3 Approximations ...................................................... 54
4.4 Two Algorithms ...................................................... 57
4.5 Global MLEM2 ...................................................... 62
4.6 Local MLEM2 ......................................................... 63
4.7 Incomplete Data Sets with Numerical Attributes .............. 66
4.8 Experiments ......................................................... 70
4.9 Conclusions .......................................................... 71
References ............................................................... 72

5 Roles Played by Bayesian Networks in Machine Learning:
An Empirical Investigation ............................................. 75
Estevam R. Hruschka Jr., Maria do Carmo Nicoletti
5.1 Introduction .......................................................... 75
5.2 Relevant Concepts Related to Bayesian Networks and Bayesian Classifiers ................................................. 76
5.3 Learning Bayesian Networks and Bayesian Classifiers from Data ................................................................. 81
5.3.1 The Naive Bayes Classifier ........................................ 81
5.3.2 The PC Algorithm .................................................. 82
5.3.3 The K2 Algorithm .................................................. 84
5.4 Bayesian Classifiers in Feature Subset Selection ............... 85
5.4.1 Considerations about the Feature Subset Selection (FSS) Problem ...................................................... 85
5.4.2 Feature Subset Selection by Bayesian Networks - The K2×2 Method .................................................. 89
5.5 Bayesian Classifiers in Imputation Processes .................... 94
5.5.1 Considerations about Imputation Processes ................. 94
5.5.2 Commonly Used Imputation Methods ......................... 95
5.5.3 Imputation by Bayesian Networks and the K2×2 Method ................................................................. 96
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.3.5 Some Extended Aspects to NISs</td>
<td>224</td>
</tr>
<tr>
<td>9.4 An Aspect of Question-Answering and Decision Making in NISs</td>
<td>229</td>
</tr>
<tr>
<td>9.5 Rule Generation in NISs</td>
<td>230</td>
</tr>
<tr>
<td>9.5.1 Rule Generation Tasks in a NIS</td>
<td>230</td>
</tr>
<tr>
<td>9.5.2 Stability Factor of Rules in the Upper System</td>
<td>232</td>
</tr>
<tr>
<td>9.5.3 Current State of a Rule Generator in Prolog</td>
<td>232</td>
</tr>
<tr>
<td>9.5.4 An Example of Execution by a Rule Generator</td>
<td>232</td>
</tr>
<tr>
<td>9.5.5 An Application to Other Types of Rule Generation</td>
<td>235</td>
</tr>
<tr>
<td>9.6 Perspective of RNIA in Machine Learning</td>
<td>238</td>
</tr>
<tr>
<td>9.6.1 Handling of Inexact Data</td>
<td>238</td>
</tr>
<tr>
<td>9.6.2 Learning a DIS from a NIS by Constraints</td>
<td>239</td>
</tr>
<tr>
<td>9.6.3 Table Data and Logical Data in Machine Learning</td>
<td>240</td>
</tr>
<tr>
<td>9.7 Concluding Remarks</td>
<td>241</td>
</tr>
<tr>
<td>References</td>
<td>241</td>
</tr>
<tr>
<td>10 Introduction to Perception Based Computing</td>
<td>249</td>
</tr>
<tr>
<td>Andrzej Skowron, Piotr Wasilewski</td>
<td></td>
</tr>
<tr>
<td>10.1 Introduction</td>
<td>249</td>
</tr>
<tr>
<td>10.2 Motivation for Perception Based Computing</td>
<td>251</td>
</tr>
<tr>
<td>10.3 Perception [15, 3]</td>
<td>253</td>
</tr>
<tr>
<td>10.4 Interactive Information Systems</td>
<td>255</td>
</tr>
<tr>
<td>10.5 Interactive Computing</td>
<td>259</td>
</tr>
<tr>
<td>10.6 Action Attributes and Plans</td>
<td>261</td>
</tr>
<tr>
<td>10.7 Towards Granule Semantics</td>
<td>265</td>
</tr>
<tr>
<td>10.8 Conclusions</td>
<td>270</td>
</tr>
<tr>
<td>References</td>
<td>271</td>
</tr>
<tr>
<td>11 Overlapping, Rare Examples and Class Decomposition in Learning Classifiers from Imbalanced Data</td>
<td>277</td>
</tr>
<tr>
<td>Jerzy Stefanowski</td>
<td></td>
</tr>
<tr>
<td>11.1 Introduction</td>
<td>278</td>
</tr>
<tr>
<td>11.2 Evaluation Measures for Learning Classifiers from Imbalanced Data</td>
<td>280</td>
</tr>
<tr>
<td>11.3 Earlier Studies with Data Factors in Class Imbalance</td>
<td>281</td>
</tr>
<tr>
<td>11.4 Generation of New Artificial Data Sets</td>
<td>285</td>
</tr>
<tr>
<td>11.5 Experimental Analysis of Influence of Critical Factors on Classifiers</td>
<td>289</td>
</tr>
<tr>
<td>11.6 Improving Classifiers by Focused Re-sampling Methods</td>
<td>293</td>
</tr>
<tr>
<td>11.6.1 Informed Undersampling</td>
<td>294</td>
</tr>
<tr>
<td>11.6.2 Informed Oversampling Methods</td>
<td>295</td>
</tr>
<tr>
<td>11.6.3 SPIDER Method</td>
<td>295</td>
</tr>
<tr>
<td>11.7 Experiments with Focused Re-sampling Methods</td>
<td>297</td>
</tr>
<tr>
<td>11.8 Final Remarks</td>
<td>301</td>
</tr>
<tr>
<td>References</td>
<td>302</td>
</tr>
</tbody>
</table>
12 A Granular Computing Paradigm for Concept Learning

Yiyu Yao, Xiaofei Deng

12.1 Introduction .................................................. 307
12.2 A Triarchic Theory of Granular Computing ................. 308
  12.2.1 Multilevel, Multiview Granular Structures ............ 309
  12.2.2 Philosophy: Structured Thinking ..................... 312
  12.2.3 Methodology: Structured Problem Solving .......... 313
  12.2.4 Computation: Structured Information Processing ...... 314
12.3 Granular Computing and Concept Learning .................. 315
  12.3.1 Granules and Concepts ................................ 315
  12.3.2 Granulation and Classification ...................... 316
  12.3.3 Concept Learning as Searching ..................... 318
12.4 A Model for Learning a Classification ...................... 319
  12.4.1 A Decision Logic Language in an Information Table .... 320
  12.4.2 Conjunctively Definable Concepts .................... 321
  12.4.3 Attribute-Oriented Search Strategies in a Space of
         Partitions Defined by Subsets of Attributes .......... 321
  12.4.4 Attribute-Value-Oriented Search Strategies in a Space
         of Coverings Defined by Families of Sets of
         Attribute-Value Pairs ................................ 323
12.5 Conclusion .................................................. 324
References .......................................................... 325

Part B: Applications

13 Identifying Calendar-Based Periodic Patterns .................. 329

Jhimli Adhikari, P.R. Rao

13.1 Introduction .................................................. 329
13.2 Related Work .................................................. 332
13.3 Calendar-Based Periodic Patterns .......................... 333
  13.3.1 Extending Certainty Factor ............................ 334
  13.3.2 Extending Certainty Factor with Respect to Other
         Intervals .............................................. 337
13.4 Mining Calendar-Based Periodic Patterns ..................... 339
  13.4.1 Improving Mining Calendar-Based Periodic Patterns .... 339
  13.4.2 Data Structure ........................................ 339
  13.4.2 A Modified Algorithm ................................. 341
13.5 Experimental Studies ....................................... 344
  13.5.1 Selection of Mininterval and Maxgap ................. 348
  13.5.2 Selection of Minsupp ................................. 351
  13.5.3 Performance Analysis ................................. 352
13.6 Conclusions .................................................. 355
References .......................................................... 356
14 The Mamdani Expert-System with Parametric Families of Fuzzy Constraints in Evaluation of Cancer Patient Survival Length

Elisabeth Rakus-Andersson

14.1 Introduction 359
14.2 Making Fuzzification of Input and Output Variable Entries by Parametric s-Functions 361
14.3 The Rule Based Processing Part of Surviving Length Model 370
14.4 Defuzzification of the Output Variable 372
14.5 The Survival Length Prognosis for a Selected Patient 372
14.6 Conclusions 376
References 377

15 Support Vector Machines in Biomedical and Biometrical Applications

Krzysztof A. Cyran, Jolanta Kawulok, Michal Kawulok, Magdalena Stawarz, Marcin Michalak, Monika Pietrowska, Piotr Widlak, Joanna Polarska

15.1 Introduction 380
15.2 Support Vector Machines Applied in the Classification of Mass Spectra 384
15.2.1 MS Spectra Preprocessing 384
15.2.2 Preparing Spectra to Classification 388
15.2.3 Classification 389
15.3 Support Vector Machines Applied to Human Face Recognition 396
15.3.1 Face Recognition Process 397
15.3.2 Evaluation Protocol 398
15.3.3 Selecting SVM Training Set 399
15.3.4 Face Detection 402
15.3.5 Feature Vectors Comparison 407
15.3.6 Multi-method Fusion 411
15.4 Conclusions 413
References 413

16 Workload Modeling for Multimedia Surveillance Systems

Mukesh Saini, Pradeep K. Atrey, Mohan S. Kankanhalli

16.1 Introduction 419
16.1.1 Issues in Workload Characterization 421
16.1.2 Contributions Summary 421
16.1.3 Chapter Organization 422
16.2 Surveillance System 422
16.3 Previous Work 423
16.4 Proposed Model 424
16.4.1 Target Flow Graph (TFG) 425
16.4.2 Markov Chain Construction 426
16.4.3 Task Arrival 428
16.4.4 Processing Demand .............................................. 429
16.4.5 Memory Demand ................................................. 430

16.5 Performance Evaluation ........................................... 430
16.5.1 System Response Time ......................................... 431
16.5.2 Frame Drop Probability ....................................... 432

16.6 Experiments .......................................................... 433
16.6.1 Implementation .................................................. 433
16.6.2 Hypothesis Testing: Normal Distributed Processing Time .............................................. 435
16.6.3 Response Time ................................................... 435
16.6.4 Frame Drop Probability ....................................... 437
16.6.5 Implications ..................................................... 437

16.7 Conclusions and Future Work .................................... 438

References ................................................................. 439

17 Rough Set and Artificial Neural Network Approach to Computational Stylistics .............................................. 441

Urszula Stanczyk

17.1 Introduction .......................................................... 441

17.2 Basics of Computational Stylistics .................................... 442
17.2.1 Objectives of Textual Analysis .................................. 443
17.2.2 Short Historical Overview ..................................... 444
17.2.3 Methodologies Employed ...................................... 445

17.3 Connectionist and Rule-Based Classification .................................... 447
17.3.1 Artificial Neural Networks ..................................... 447
17.3.2 Rough Set Theory .............................................. 449

17.4 Experimental Setup .................................................. 452
17.4.1 Input Datasets .................................................. 452
17.4.2 Connectionist Classification ................................... 455
17.4.3 Rule-Based Classification ..................................... 455
17.4.4 Analysis of Characteristic Features .......................... 458
17.4.5 Performance for Feature Reduction ........................... 462

17.5 Conclusions and Future Research .................................. 468

References ................................................................. 469

18 Application of Learning Algorithms to Image Spam Evolution ............ 471

Shruti Wakade, Kathy J. Liszka, Chien-Chung Chan

18.1 Introduction .......................................................... 471

18.2 Related Work ........................................................ 473

18.3 Spam Images Evolution and Datasets ................................... 474
18.3.1 Types and Trends of Image Spam ................................ 474
18.3.2 The Corpus ..................................................... 477