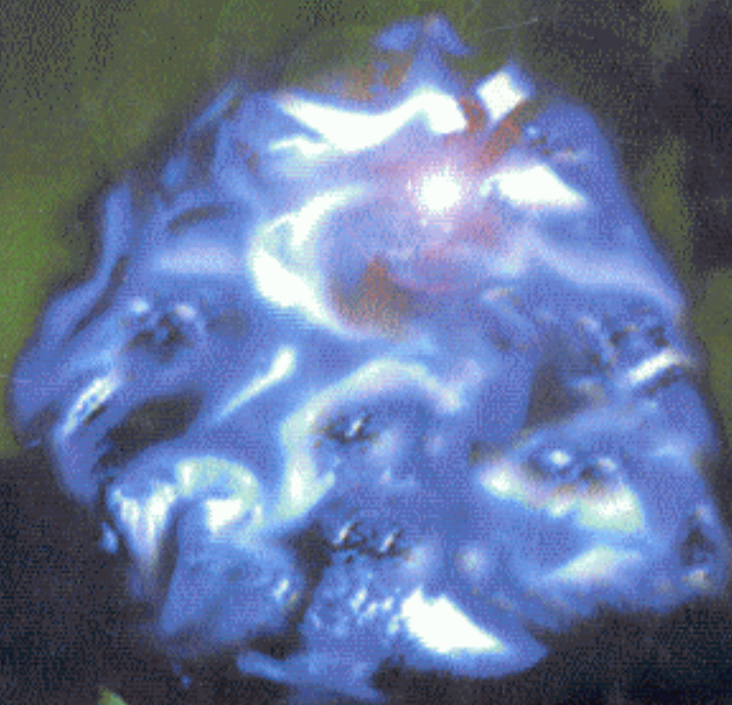
FEATURE EXTRACTION CONSTRUCTION and SELECTION

A DATA MINING PERSPECTIVE



HUAN LIU HIROSHI MOTODA

Contents

Preface xiii

Acknowledgments xv

Contributing Authors xvii

Part I Background and Foundation

1
Less is More 3
Huan Liu and Hiroshi Motoda
1.1 Background 3
1.2 Current Status of Feature Transformation and Subset Selection 5
1.3 Future Work 9

References 11

2
Feature Weighting for Lazy Learning Algorithms 13
David W. Aha
2.1 Introduction 13
2.2 Feature Weighting in A Lazy Learning Algorithm 15
2.3 A Categorization of Feature Weighting in Lazy Learners 16
2.4 Additional Contributions 27
2.5 Summary 28

References 29

3
The Wrapper Approach 33
Ron Kohavi and George H. John
3.1 Introduction 33
3.2 Relevance of Features 35
3.3 The Filter Approach 36

3.4 The Wrapper Approach 38
3.5 Related Work 45
3.6 Future Work 46
3.7 Conclusion 47

References 48

4
Data-driven Constructive Induction: Methodology and Applications 51
Eric Bloedorn, and Ryszard S. Michalski
4.1 Introduction 52
4.2 An Illustration of the Importance of the Representation Space 53
4.3 The Need for Constructive Induction 54
4.4 A General Schema for Constructive Induction 56
4.5 Experimental Applications 59
4.6 Conclusion 66

References 67

Part II Subset Selection

5
Selecting Features by Vertical Compactness of Data 71
Ke Wang and Suman Sundaresh
5.1 Introduction 71
5.2 The Vertical Compactness Criterion 73
5.3 Search Methods 75
5.4 Hybrid Search Algorithm 77
5.5 Experiments 80
5.6 Conclusions 83

References 84

6
Relevance Approach to Feature Subset Selection 85
Hui Wang, David Bell, and Fionn Murtagh
6.1 Introduction 85
6.2 Characterisation of Feature Subset Selection 87
6.3 A Relevance-based Algorithm for Feature Selection 90
6.4 Experiment and Evaluation 91
6.5 Comparison with Related Work 92
6.6 Conclusion 95
Appendix: A unified framework for relevance 96

References 97

7
Novel Methods for Feature Subset Selection with Respect to Problem Knowledge 101
Pavel Pudil and Jana Novovičová
7.1 Introduction 101

7.2 Feature Subset Selection Problem in Statistical Pattern Recognition 103
7.3 Basic Situation with Respect to Problem Knowledge 105
7.4 Floating Search Methods 106
7.5 Feature Selection by Modified Gaussian Mixtures 108
7.6 Experimental Results 112
7.7 Subset Selection Guide 114

References 115

8
Feature Subset Selection Using A Genetic Algorithm 117
Jihoon Yang and Vasant Honavar
8.1 Introduction 117
8.2 Related Work 118
8.3 Feature Selection Using a Genetic Algorithm for Neural Network Pattern Classifiers 122
8.4 Implementation Details 125
8.5 Experiments 127
8.6 Summary and Discussion 131

References 133

9
A Relevancy Filter for Constructive Induction 137
Nada Lavrač, Dragan Gamberger, and Peter Turney
9.1 Introduction 137
9.2 Relevance of Literals and Features 139
9.3 Utility Study: The East-West Challenge 144
9.4 Handling Noisy Data 150
9.5 Noise and Relevance 151
9.6 Related Work 151
9.7 Conclusion 152

References 153

Part III Feature Extraction

10
Lexical Contextual Relations for the Unsupervised Discovery of Texts Features 157
Patrick Perrin and Fred Petry
10.1 Introduction 157
10.2 Collocational Expressions as Text Features 160
10.3 Discovery of Contextual Text Features 163
10.4 Results and Discussion 164
10.5 Conclusion 170

References 172

11
Integrated Feature Extraction Using Adaptive Wavelets 175

Yvette Mallet, Olivier de Vel, and Danny Coomans

11.1 Introduction 175
11.2 Wavelet Analysis of Spectra 179
11.3 Multivariate Prediction Models 183
11.4 The Data Set 185
11.5 Classification Applications 185
11.6 Regression Applications 187
11.7 Future Directions 188

References 189

12
Feature Extraction via Neural Networks 191
Rudy Setiono and Huan Liu

12.1 Introduction 191
12.2 Feature Extraction via Neural Networks 193
12.3 Illustrative Examples 195
12.4 Empirical Study and Analysis 199
12.5 CNF9a Revisited for Decision Tree Induction 200
12.6 Summary 202
Appendix: Definitions for artificial problems 202

References 202

13
Using Lattice-based Framework as a Tool for Feature Extraction 205
E. Mephu Nguifo, P. Njiwoua

13.1 Introduction 205
13.2 Galois Lattice 207
13.3 Feature Extraction Algorithm 209
13.4 The IGLUE System 212
13.5 Experimental Results 213
13.6 Conclusion 216

References 217

14
Constructive Function Approximation 219
Paul E. Utgoff and Doina Precup

14.1 Introduction 219
14.2 The Need for Features 220
14.3 Feature Spaces 221
14.4 Comparison of Three Function Approximation Algorithms 223
14.5 Related Work 233

References 235

Part IV Feature Construction

15

A Comparison of Constructing Different Types of New Feature for Decision Tree Learning 239
Zijian Zheng
15.1 Introduction 239
15.2 Four Types of New Feature 241
15.3 A Single Algorithm for Creating Four Different Types of New Feature 243
15.4 Experiments 244
15.5 Discussion 251
15.6 Related Work 252
15.7 Conclusions and Future Work 253

References 254

16
Constructive Induction: Covering Attribute Spectrum 257
Yuh-Jyh Hu
16.1 Introduction 257
16.2 Important Issues for Constructive Induction 258
16.3 GALA2.0: Combining Neural Network with Iterative Attribute Construction 261
16.4 Discussion 269

References 271

17
Feature Construction Using Fragmentary Knowledge 273
Steve Donoho and Larry Rendell
17.1 Introduction 273
17.2 Types of Fragmentary Knowledge 274
17.3 Incorporating Fragmentary Knowledge into Search 278
17.4 Bankruptcy Experiments 280
17.5 Turfgrass Management Experiments 283
17.6 Discussion 286

References 287

18
Constructive Induction on Continuous Spaces 289
João Gama and Pavel Brazdil
18.1 Constructive Induction 289
18.2 An Illustrative Example 291
18.3 Growing the Tree 294
18.4 Related Work 298
18.5 Experiments 299
18.6 Conclusions 301

References 302

Part V Combined Approaches

19

Evolutionary Feature Space Transformation 307
Haleh Vafaie and Kenneth De Jong
19.1 Introduction 307
19.2 Approaches to Feature Selection and Construction 308
19.3 General System Architecture 310
19.4 Experimental Setup 315
19.5 Comparative studies 320
19.6 Summary and Conclusions 321

References 322

20
Feature Transformation by Function Decomposition 325
Blaž Zupan, Marko Bohanec, Janez Demšar, and Ivan Bratko
20.1 Introduction 325
20.2 Single-step Function Decomposition 327
20.3 Finding the Best Feature Partition 329
20.4 Redundancy Discovery and Removal 330
20.5 Discovering Feature Hierarchies 331
20.6 Feature Construction 334
20.7 Related Work 336
20.8 Summary 338

References 338

21
Constructive Induction of Cartesian Product Attributes 341
Michael J. Pazzani
21.1 Introduction 341
21.2 A Wrapper Approach for Creating Cartesian Product Attributes 342
21.3 Experiments 344
21.4 Attribute deletion 349
21.5 Related Work 350
21.6 Conclusions 351

References 352

Part VI Applications of Feature Transformation

22
Towards Automatic Fractal Feature Extraction for Image Recognition 357
Matteo Baldoni, Cristina Baroglio, Davide Cavagnino and Lorenza Saitta
22.1 Introduction 357
22.2 Fractals and IFS 359
22.3 Extraction of Fractal Features 361
22.4 Experiments 363
22.5 Discussion and future work 370

References 372

23
Feature Transformation Strategies for a Robot Learning Problem 375
Luis Seabra Lopes, Luis M. Camarinha-Matos
23.1 Introduction 375
23.2 Robot Programming by Demonstration 376
23.3 Learning Structured Concepts 377
23.4 Constructive Induction 379
23.5 Experimental Situation: the Pick and Place Task 380
23.6 Feature Transformation Strategies 381
23.7 Experimental Results 386
23.8 Conclusions 390

References 391

24
Interactive Genetic Algorithm Based Feature Selection and Its Application to Marketing Data Analysis 393
Takao Terano and Yoko Ishino
24.1 Introduction 394
24.2 Interactive GAs, Feature Selection, and Knowledge Extraction 395
24.3 Algorithm of SIBILE 396
24.4 Experiments 399
24.5 Concluding Remarks 404

References 405

Index 407