

Contents

1	Feature Selection for Data and Pattern Recognition: An Introduction	1
	Urszula Stańczyk and Lakhmi C. Jain	
1.1	Introduction	1
1.2	Chapters of the Book	3
1.3	Concluding Remarks	5
	References	5
 Part I Estimation of Feature Importance		
2	All Relevant Feature Selection Methods and Applications	11
	Witold R. Rudnicki, Mariusz Wrzesień and Wiesław Paja	
2.1	Introduction	12
2.1.1	Definitions	14
2.1.2	Algorithms for All-Relevant Feature Selection	16
2.1.3	Random Forest	18
2.2	Testing Procedure	18
2.2.1	Data Sets	18
2.2.2	Classification	19
2.2.3	Feature Selection	20
2.3	Results and Discussion	21
2.3.1	Classification	22
2.3.2	Feature Selection	23
2.4	Conclusions	27
	References	27

3	Feature Evaluation by Filter, Wrapper, and Embedded Approaches	29
	Urszula Stańczyk	
3.1	Introduction	29
3.2	Characteristic Features for Stylometric Analysis of Texts	30
3.3	Approaches to Feature Selection	31
3.3.1	Filters	32
3.3.2	Wrappers	32
3.3.3	Embedded Solutions	33
3.3.4	Ranking of Features	33
3.4	Details of Research Framework	34
3.4.1	Input Data Sets	34
3.4.2	Machine Learning Techniques Used in Research	35
3.4.3	Search Parameters	36
3.5	Feature Evaluation by Ranking	36
3.6	Feature Evaluation by Backward Reduction	38
3.6.1	Relief Ranking	38
3.6.2	Embedded DRSA Ranking	40
3.6.3	Comparison of Feature Reduction Results	41
3.7	Conclusions	42
	References	43
4	A Geometric Approach to Feature Ranking Based Upon Results of Effective Decision Boundary Feature Matrix	45
	Claudia Diamantini, Alberto Gemelli and Domenico Potena	
4.1	Introduction	45
4.2	Feature Ranking for Classification: The Background Picture	46
4.2.1	Intrinsic Discriminant Dimension of a Classification Task	46
4.2.2	Classical Feature Selection Strategies	47
4.2.3	A Multiple-Challenge Case Study for Feature Ranking	50
4.3	Focus on Feature Extraction Based Ranking	50
4.3.1	Linear Models	50
4.3.2	Feature Extraction Based on Decision Boundary	51
4.4	Feature Ranking Based on Effective Decision Boundary Feature Matrix	52
4.4.1	Geometric Considerations	52
4.4.2	The Algorithm	54
4.5	Experiments	56
4.5.1	Experimental Setting	56
4.5.2	Benchmarking the EDBFM Ranking Method	59
4.6	Conclusions	67
	References	67

5	Weighting of Features by Sequential Selection	71
	Urszula Stańczyk	
5.1	Introduction	71
5.2	Background	72
5.2.1	Algorithms for Feature Selection.	73
5.2.2	Connectionist Classifier	75
5.2.3	Rule-Based Classification.	75
5.2.4	Textual Analysis.	77
5.3	Experimental Setting	77
5.4	Sequential Forward Selection	78
5.5	Sequential Backward Selection	83
5.6	Concluding Remarks	87
	References.	89

Part II Rough Set Approach to Attribute Reduction

6	Dependency Analysis and Attribute Reduction in the Probabilistic Approach to Rough Sets	93
	Wojciech Ziarko	
6.1	Introduction	93
6.2	Variable Precision Rough Sets	95
6.2.1	Set Approximations in the VPRS Approach	96
6.2.2	Absolute Set Approximation Regions	98
6.3	Dependencies in Approximation Spaces.	99
6.3.1	Absolute Certainty Gain	99
6.3.2	Absolute Dependency Gain	100
6.3.3	Average Dependency Gain.	100
6.4	Probabilistic Decision Tables	101
6.4.1	Attributes.	101
6.4.2	Decision Tables	102
6.4.3	Classification Tables	103
6.5	Dependencies in Decision Tables	104
6.5.1	Functional and Partial Functional Dependencies	104
6.5.2	λ — <i>Dependency Measure</i>	105
6.6	λ — <i>Dependency-Based Reduct</i>	106
6.7	Probabilistic Decision Rules.	108
6.8	Significance of λ — <i>Reduct</i> Attributes.	108
6.9	λ — <i>Core</i> Collection of Attributes	109
6.10	Final Remarks	109
	References.	110

7	Structure-Based Attribute Reduction: A Rough Set Approach	113
	Yoshifumi Kusunoki and Masahiro Inuiguchi	
7.1	Introduction	113
7.2	Structure-Based Attribute Reduction in Rough Set Models	115
	7.2.1 Decision Tables	115
	7.2.2 Rough Set Models	117
	7.2.3 Reducts in Rough Set Models	121
	7.2.4 Boolean Functions Representing Reducts	124
7.3	Structure-Based Attribute Reduction in Variable Precision Rough Set Models	127
	7.3.1 Rough Membership Function	127
	7.3.2 Variable Precision Rough Set Models	128
	7.3.3 Structure-Based Reducts in Variable Precision Rough Set Models	133
	7.3.4 Boolean Functions Representing Reducts	136
7.4	Structure-Based Attribute Reduction in Dominance-Based Rough Set Models	144
	7.4.1 Decision Tables Under Dominance Principle and Dominance-Based Rough Set Models	144
	7.4.2 Structure-Based Reducts in Dominance-Based Rough Set Models	150
	7.4.3 Boolean Functions Representing Reducts	154
7.5	Concluding Remarks	157
	References	158

Part III Rule Discovery and Evaluation

8	A Comparison of Rule Induction Using Feature Selection and the LEM2 Algorithm	163
	Jerzy W. Grzymala-Busse	
8.1	Introduction	163
8.2	Rule Induction Based on Feature Selection.	165
8.3	LEM2	166
8.4	Inconsistent Data	168
8.5	LEERS Classification System	171
8.6	Experiments	171
8.7	Conclusions	175
	References	175

9	Meta-actions as a Tool for Action Rules Evaluation	177
	Hakim Touati, Zbigniew W. Raś and James Studnicki	
9.1	Introduction	178
9.2	Decision Systems	179
9.3	Action-Rules	179
9.3.1	Action Rules Evaluation	180
9.4	Meta-actions.	181
9.4.1	Meta-actions Influence Matrix	182
9.4.2	Mining Meta-actions Effects.	183
9.4.3	Meta-action Evaluation	185
9.5	Meta-actions Versus Action Rules.	186
9.5.1	Action Rules Selection by Meta-actions.	188
9.6	Side Effects	190
9.6.1	Meta-actions Side Effects.	190
9.6.2	Action Rules Side Effects	190
9.7	Experiments	191
9.7.1	Dataset Description	191
9.7.2	Meta-action Extraction.	192
9.7.3	Action Rules Extraction and Evaluation.	195
9.8	Conclusion.	196
	References.	196
10	Irrelevant Feature and Rule Removal for Structural Associative Classification Using Structure-Preserving Flat Representation.	199
	Izwan Nizal Mohd Shaharanee and Fedja Hadzic	
10.1	Introduction	200
10.2	Related Work	201
10.2.1	Relationship Between Feature Subset Selection and Frequent Subtree Interestingness	203
10.3	Problem Background	204
10.3.1	Feature Subset Selection	205
10.3.2	Modeling Tree-Structured Data	207
10.3.3	Database Structure Model (DSM)	209
10.3.4	Tree to Flat Conversion Example Using Academic Institution WebLogs Data	211
10.3.5	Representing Disconnected Trees w.r.t. DSM	212
10.4	Method and Experimental Setup	214
10.5	Experimental Evaluation	218
10.5.1	Experiment Set 1—CRM Data	219
10.5.2	Experiment Set 2—CSLOGS Data	221

10.5.3	Experiment Set 3—Academic Institution	
	Web Log Data	223
10.6	Conclusion and Future Work	225
	References.	226

Part IV Data- and Domain-Oriented Methodologies

11	Hubness-Aware Classification, Instance Selection and Feature Construction: Survey and Extensions to Time-Series	231
	Nenad Tomašev, Krisztian Buza, Kristóf Marussy and Piroska B. Kis	
11.1	Introduction	232
11.2	Problem Formulation and Basic Notations	233
11.3	Dynamic Time Warping	234
11.4	Hubs in Time-Series Data	238
11.5	Hubness-Aware Classification of Time-Series	240
11.5.1	hw- k NN: Hubness-Aware Weighting	241
11.5.2	h-FNN: Hubness-Based Fuzzy Nearest Neighbor	243
11.5.3	NHBNN: Naive Hubness Bayesian k -Nearest Neighbor	244
11.5.4	HIKNN: Hubness Information k -Nearest Neighbor	246
11.5.5	Experimental Evaluation of Hubness-Aware Classifiers	247
11.6	Instance Selection and Feature Construction for Time-Series Classification	254
11.6.1	Instance Selection for Speeding-Up Time-Series Classification	254
11.6.2	Feature Construction	256
11.7	Conclusions and Outlook	258
	References.	259
12	Selection of Visual Descriptors for the Purpose of Multi-camera Object Re-identification.	263
	Piotr Dalka, Damian Ellwart, Grzegorz Szwoch, Karol Lisowski, Piotr Szczuko and Andrzej Czyżewski	
12.1	Introduction	263
12.2	Video Preprocessing	266
12.3	Multi-camera Object Tracking	267
12.4	Visual Object Descriptors	268
12.4.1	Colour Histogram	269
12.4.2	Vertical Trace.	271

12.4.3	Moment Invariants	272
12.4.4	Colour Layout Descriptors	272
12.4.5	Co-occurrence Matrices Statistical Parameters	273
12.4.6	Edge Histogram Descriptor	275
12.4.7	Local Binary Pattern	276
12.4.8	Local Image Features	277
12.5	Descriptor Evaluation Methods	278
12.5.1	RS Index	278
12.5.2	SD Index	279
12.5.3	Dissimilarity Measure	280
12.5.4	Result Aggregation	281
12.6	Object Identification	283
12.7	Experiments and Results	288
12.7.1	Datasets	288
12.7.2	Descriptor Evaluation	289
12.7.3	Object Identification	296
12.8	Conclusions	300
	References	301

13 Improving the Recognition Performance of Moment Features by Selection

	George A. Papakostas	305
13.1	Introduction	305
13.2	Image Moment Features	306
13.2.1	Continuous Orthogonal Moments	307
13.2.2	Discrete Orthogonal Moments	311
13.2.3	Image Reconstruction by the Method of Moments	314
13.2.4	Moments Interpretation	315
13.3	Is There a Need for Moments' Selection?	318
13.4	Moment Features Selection	319
13.4.1	GA-Based Selection	320
13.4.2	Relief Algorithm	320
13.5	Experimental Study	321
13.5.1	Benchmark Datasets	321
13.5.2	Datasets Pre-processing	321
13.5.3	Simulations	323
13.6	Conclusions	325
	References	326

14 Signature Selection for Grouped Features with a Case Study on Exon Microarrays	329
Sangkyun Lee	
14.1 Introduction	329
14.1.1 Regularized Regression	330
14.2 Regularized Regression Methods for Grouped Features	332
14.2.1 Group Lasso	333
14.2.2 Overlapping Group Lasso	336
14.2.3 Sparse Group Lasso	338
14.3 A Case Study on Exon Microarray Data	340
14.3.1 Data	341
14.3.2 Algorithms for Comparison	342
14.3.3 Comparison of Performance	343
14.4 Discussion	347
14.5 Conclusion	347
References	348
Index	351