
Contents

Foreword	xiii
Preface	xxi
List of Tables	xxv
List of Figures	xxvii
1 Introduction	1
1.1 Introduction	1
1.2 Pattern Recognition in Brief	3
1.2.1 Data acquisition	4
1.2.2 Feature selection/extraction	4
1.2.3 Classification	5
1.3 Knowledge Discovery in Databases (KDD)	7
1.4 Data Mining	10
1.4.1 Data mining tasks	10
1.4.2 Data mining tools	12
1.4.3 Applications of data mining	12
1.5 Different Perspectives of Data Mining	14
1.5.1 Database perspective	14
1.5.2 Statistical perspective	15
1.5.3 Pattern recognition perspective	15
1.5.4 Research issues and challenges	16
1.6 Scaling Pattern Recognition Algorithms to Large Data Sets	17
1.6.1 Data reduction	17
1.6.2 Dimensionality reduction	18
1.6.3 Active learning	19
1.6.4 Data partitioning	19
1.6.5 Granular computing	20
1.6.6 Efficient search algorithms	20
1.7 Significance of Soft Computing in KDD	21
1.8 Scope of the Book	22

2	Multiscale Data Condensation	29
2.1	Introduction	29
2.2	Data Condensation Algorithms	32
2.2.1	Condensed nearest neighbor rule	32
2.2.2	Learning vector quantization	33
2.2.3	Astrahan's density-based method	34
2.3	Multiscale Representation of Data	34
2.4	Nearest Neighbor Density Estimate	37
2.5	Multiscale Data Condensation Algorithm	38
2.6	Experimental Results and Comparisons	40
2.6.1	Density estimation	41
2.6.2	Test of statistical significance	41
2.6.3	Classification: Forest cover data	47
2.6.4	Clustering: Satellite image data	48
2.6.5	Rule generation: Census data	49
2.6.6	Study on scalability	52
2.6.7	Choice of scale parameter	52
2.7	Summary	52
3	Unsupervised Feature Selection	59
3.1	Introduction	59
3.2	Feature Extraction	60
3.3	Feature Selection	62
3.3.1	Filter approach	63
3.3.2	Wrapper approach	64
3.4	Feature Selection Using Feature Similarity (FSFS)	64
3.4.1	Feature similarity measures	65
3.4.2	Feature selection through clustering	68
3.5	Feature Evaluation Indices	71
3.5.1	Supervised indices	71
3.5.2	Unsupervised indices	72
3.5.3	Representation entropy	73
3.6	Experimental Results and Comparisons	74
3.6.1	Comparison: Classification and clustering performance	74
3.6.2	Redundancy reduction: Quantitative study	79
3.6.3	Effect of cluster size	80
3.7	Summary	82
4	Active Learning Using Support Vector Machine	83
4.1	Introduction	83
4.2	Support Vector Machine	86
4.3	Incremental Support Vector Learning with Multiple Points	88
4.4	Statistical Query Model of Learning	89
4.4.1	Query strategy	90
4.4.2	Confidence factor of support vector set	90

4.5	Learning Support Vectors with Statistical Queries	91
4.6	Experimental Results and Comparison	94
4.6.1	Classification accuracy and training time	94
4.6.2	Effectiveness of the confidence factor	97
4.6.3	Margin distribution	97
4.7	Summary	101
5	Rough-fuzzy Case Generation	103
5.1	Introduction	103
5.2	Soft Granular Computing	105
5.3	Rough Sets	106
5.3.1	Information systems	107
5.3.2	Indiscernibility and set approximation	107
5.3.3	Reducts	108
5.3.4	Dependency rule generation	110
5.4	Linguistic Representation of Patterns and Fuzzy Granulation	111
5.5	Rough-fuzzy Case Generation Methodology	114
5.5.1	Thresholding and rule generation	115
5.5.2	Mapping dependency rules to cases	117
5.5.3	Case retrieval	118
5.6	Experimental Results and Comparison	120
5.7	Summary	121
6	Rough-fuzzy Clustering	123
6.1	Introduction	123
6.2	Clustering Methodologies	124
6.3	Algorithms for Clustering Large Data Sets	126
6.3.1	CLARANS: Clustering large applications based upon randomized search	126
6.3.2	BIRCH: Balanced iterative reducing and clustering us- ing hierarchies	126
6.3.3	DBSCAN: Density-based spatial clustering of applica- tions with noise	127
6.3.4	STING: Statistical information grid	128
6.4	CEMMiSTRI: Clustering using EM, Minimal Spanning Tree and Rough-fuzzy Initialization	129
6.4.1	Mixture model estimation via EM algorithm	130
6.4.2	Rough set initialization of mixture parameters	131
6.4.3	Mapping reducts to mixture parameters	132
6.4.4	Graph-theoretic clustering of Gaussian components	133
6.5	Experimental Results and Comparison	135
6.6	Multispectral Image Segmentation	139
6.6.1	Discretization of image bands	141
6.6.2	Integration of EM, MST and rough sets	141
6.6.3	Index for segmentation quality	141

6.6.4	Experimental results and comparison	141
6.7	Summary	147
7	Rough Self-Organizing Map	149
7.1	Introduction	149
7.2	Self-Organizing Maps (SOM)	150
7.2.1	Learning	151
7.2.2	Effect of neighborhood	152
7.3	Incorporation of Rough Sets in SOM (RSOM)	152
7.3.1	Unsupervised rough set rule generation	153
7.3.2	Mapping rough set rules to network weights	153
7.4	Rule Generation and Evaluation	154
7.4.1	Extraction methodology	154
7.4.2	Evaluation indices	155
7.5	Experimental Results and Comparison	156
7.5.1	Clustering and quantization error	157
7.5.2	Performance of rules	162
7.6	Summary	163
8	Classification, Rule Generation and Evaluation using Modular Rough-fuzzy MLP	165
8.1	Introduction	165
8.2	Ensemble Classifiers	167
8.3	Association Rules	170
8.3.1	Rule generation algorithms	170
8.3.2	Rule interestingness	173
8.4	Classification Rules	173
8.5	Rough-fuzzy MLP	175
8.5.1	Fuzzy MLP	175
8.5.2	Rough set knowledge encoding	176
8.6	Modular Evolution of Rough-fuzzy MLP	178
8.6.1	Algorithm	178
8.6.2	Evolutionary design	182
8.7	Rule Extraction and Quantitative Evaluation	184
8.7.1	Rule extraction methodology	184
8.7.2	Quantitative measures	188
8.8	Experimental Results and Comparison	189
8.8.1	Classification	190
8.8.2	Rule extraction	192
8.9	Summary	199

A Role of Soft-Computing Tools in KDD	201
A.1 Fuzzy Sets	201
A.1.1 Clustering	202
A.1.2 Association rules	203
A.1.3 Functional dependencies	204
A.1.4 Data summarization	204
A.1.5 Web application	205
A.1.6 Image retrieval	205
A.2 Neural Networks	206
A.2.1 Rule extraction	206
A.2.2 Clustering and self organization	206
A.2.3 Regression	207
A.3 Neuro-fuzzy Computing	207
A.4 Genetic Algorithms	208
A.5 Rough Sets	209
A.6 Other Hybridizations	210
B Data Sets Used in Experiments	211
References	215
Index	237
About the Authors	243