

# Contents

<b>1</b>	<b>Machine Learning and Data Mining</b>	<b>1</b>
1.1	Introduction	1
1.2	Historical Background	3
1.3	Knowledge Discovery Process and Standardization	4
1.4	Terminology and Categorization of Learning Tasks	6
1.5	Predictive Data Mining: Induction of Models	8
1.6	Descriptive Data Mining: Induction of Patterns	13
1.7	Relational Data Mining	16
1.8	Conclusion	17
<b>2</b>	<b>Rule Learning in a Nutshell</b>	<b>19</b>
2.1	Problem Definition	20
2.2	Data Representation	23
2.3	Rule Representation	25
2.4	Rule Learning Process	28
2.5	Learning a Single Rule	29
2.6	Learning a Rule-Based Model	40
2.7	Overfitting Avoidance	42
2.8	Making a Prediction	44
2.9	Estimating the Predictive Accuracy of Rules	45
2.10	A Brief History of Predictive Rule Learning Algorithms	48
2.11	Conclusion	55
<b>3</b>	<b>Formal Framework for Rule Analysis</b>	<b>57</b>
3.1	Introduction	57
3.2	ROC Space	58
3.3	Coverage Space	60
3.4	Rule Set Learning in Coverage Space	62
3.5	Conclusion	63
<b>4</b>	<b>Features</b>	<b>65</b>
4.1	Introduction	66
4.2	Feature Types	68

4.3	Feature Construction .....	71
4.4	Feature Relevancy .....	76
4.5	Handling Missing Attribute Values .....	85
4.6	Handling Imprecise Attribute Values .....	90
4.7	Conclusion .....	93
<b>5</b>	<b>Relational Features .....</b>	<b>95</b>
5.1	Introduction .....	95
5.2	Examples of Relational Learning Tasks .....	97
5.3	Examples of Relational Features .....	99
5.4	Relational Feature Construction .....	101
5.5	Alternative Knowledge Representation Formalisms .....	104
5.6	Conclusion .....	111
<b>6</b>	<b>Learning Single Rules .....</b>	<b>113</b>
6.1	Introduction .....	113
6.2	Single Rule Construction .....	114
6.3	Search Algorithms .....	116
6.4	Search Strategies .....	123
6.5	Relational Rule Construction .....	128
6.6	Conclusion .....	133
<b>7</b>	<b>Rule Evaluation Measures .....</b>	<b>135</b>
7.1	Introduction .....	135
7.2	Basic Properties of Heuristics .....	137
7.3	Heuristics with Linear Isometrics .....	139
7.4	Heuristics with Nonlinear Isometrics .....	152
7.5	Gain Heuristics .....	155
7.6	Complexity-Based Heuristics .....	161
7.7	Composite Heuristics .....	163
7.8	Practical Considerations and Recommendations .....	164
7.9	Conclusion .....	169
<b>8</b>	<b>Learning Rule Sets .....</b>	<b>171</b>
8.1	Disjunctive Rules and Rule Sets .....	171
8.2	The Covering Strategy .....	174
8.3	Weighted Covering .....	178
8.4	Nonreductive Rule Learning Algorithms .....	181
8.5	Combining Rules to Arbitrary Global Models .....	184
8.6	Conclusion .....	185
<b>9</b>	<b>Pruning of Rules and Rule Sets .....</b>	<b>187</b>
9.1	Overfitting and Pruning .....	187
9.2	Prepruning .....	190
9.3	Post-pruning .....	199
9.4	Integrating Pre- and Post-pruning .....	207
9.5	Optimizing Rule Sets .....	212
9.6	Conclusion .....	216

<b>10</b>	<b>Beyond Concept Learning</b>	217
10.1	Introduction	217
10.2	Binary Classification	218
10.3	Prediction with Rule Bases	220
10.4	Multiclass Classification	223
10.5	Learning in Structured Output Spaces	236
10.6	Regression	243
10.7	Unsupervised Learning	245
10.8	Conclusion	246
<b>11</b>	<b>Supervised Descriptive Rule Learning</b>	247
11.1	Introduction	247
11.2	Subgroup Discovery	249
11.3	Contrast Set Mining	251
11.4	Emerging Pattern Mining	253
11.5	A Unifying View	255
11.6	Other Approaches to Supervised Descriptive Rule Learning	259
11.7	Subgroup Visualization	261
11.8	Conclusion	265
<b>12</b>	<b>Selected Applications</b>	267
12.1	Introduction	267
12.2	Demographic Domain: Dealing with Sequential Data	268
12.3	Multiclass Gene Expression Domain: Feature Relevancy for Scientific Discovery Applications	274
12.4	Coronary Heart Disease Domain: Expert-Guided Data Analysis	281
12.5	Brain Ischaemia Domain: Insightful Data Analysis	288
12.6	Conclusion	298
	<b>References</b>	299
	<b>Index</b>	323