

Table of Contents

1	Introduction	1
1.1	Data Mining and Knowledge Discovery	1
1.1.1	Desirable Properties of Discovered Knowledge	2
1.2	Knowledge Representation	3
1.2.1	Prediction Rules	4
1.3	An Overview of Data Mining Paradigms	5
1.3.1	Rule Induction.....	7
1.3.2	Evolutionary Algorithms.....	9
	References	10
2	Data Mining Tasks and Concepts	13
2.1	Classification.....	13
2.1.1	The Non-determinism of the Classification Task.....	17
2.1.2	Overfitting and Underfitting	17
2.1.3	Estimating Accuracy via Training and Testing.....	19
2.2	Dependence Modeling.....	22
2.2.1	Dependence Modeling vs Association-Rule Discovery	23
2.3	The Challenge of Measuring Prediction-Rule Quality	24
2.3.1	Measuring Predictive Accuracy	24
2.3.2	Measuring Rule Comprehensibility	26
2.3.3	Measuring Rule Interestingness	27
2.4	Clustering	31
2.4.1	Hierarchical Agglomerative Clustering	33
2.4.2	The <i>K</i> -Means Algorithm	34
2.4.3	The Challenge of Measuring Clustering Quality	35
2.5	Inductive Bias.....	36
2.5.1	The Domain-Dependent Effectiveness of a Bias	37
2.5.2	The Simplicity Bias of Occam's Razor.....	38
2.5.3	The Minimum Description Length (MDL) Principle.....	39
	References	40

3	Data Mining Paradigms	45
3.1	Decision-Tree Building Algorithms	45
3.1.1	Decision-Tree Building	47
3.1.2	Decision-Tree Pruning	49
3.1.3	Pros and Cons of Decision-Tree Algorithms	50
3.1.4	The Inductive Bias of Decision-Tree Algorithms	52
3.1.5	Linear (Oblique) Decision Trees.....	53
3.2	Rule Induction Algorithms.....	54
3.2.1	The Pitfall of Attribute Interaction for Greedy Algorithms.....	55
3.3	Instance-Based Learning (Nearest Neighbor) Algorithms	58
	References	60
4	Data Preparation	65
4.1	Attribute Selection.....	65
4.1.1	The Motivation for Attribute Selection	66
4.1.2	Filter and Wrapper Approaches	66
4.1.3	Attribute Selection as a Particular Case of Attribute Weighting	71
4.2	Discretization of Continuous Attributes.....	71
4.2.1	An Overview of Discretization Methods	72
4.2.2	The Pros and Cons of Discretization.....	73
4.3	Attribute Construction.....	74
4.3.1	Categorizing Attribute Construction Methods	75
	References	76
5	Basic Concepts of Evolutionary Algorithms	79
5.1	An Overview of Evolutionary Algorithms (EAs).....	79
5.1.1	Key Issues in the Design of EAs.....	81
5.2	Selection Methods	83
5.3	Genetic Algorithms (GA).....	84
5.3.1	Individual (“Chromosome”) Representation.....	84
5.3.2	Crossover Operators.....	86
5.3.3	Mutation Operators	88
5.4	Genetic Programming (GP).....	89
5.4.1	Function Set	90
5.4.2	Crossover Operators.....	92
5.4.3	Mutation Operators	95

5.4.4	The Standard GP Approach for Classification	99
5.4.5	Code Bloat	100
5.5	Niching	101
References	103
6	Genetic Algorithms for Rule Discovery	107
6.1	Individual Representation	107
6.1.1	Pittsburgh vs Michigan Approach.....	107
6.1.2	Encoding a Rule Antecedent.....	110
6.1.3	Encoding a Rule Consequent	116
6.2	Task-Specific Generalizing/Specializing Operators.....	119
6.2.1	Generalizing/Specializing Crossover	119
6.2.2	Generalizing/Specializing Mutation.....	122
6.2.3	Condition Removal/Insertion.....	123
6.2.4	Rule Insertion/Removal	124
6.2.5	Discussion	124
6.3	Task-Specific Population Initialization and Seeding.....	126
6.4	Task-Specific Rule-Selection Methods	127
6.5	Fitness Evaluation	129
References	134
7	Genetic Programming for Rule Discovery	139
7.1	The Problem of Closure in GP for Rule Discovery.....	139
7.2	Booleanizing All Terminals	140
7.2.1	Methods for Booleanizing Attributes.....	142
7.2.2	Examples of GP Systems Booleanizing Terminals.....	145
7.2.3	A Note on Semantic Constraints	145
7.3	Constrained-Syntax and Strongly-Typed GP	146
7.4	Grammar-Based GP for Rule Discovery	150
7.5	GP for Decision-Tree Building	153
7.5.1	Evolving Decision Trees with First-Order-Logic Conditions	154
7.5.2	Evolving Linear Decision Trees	155
7.5.3	Hybrid GP, Cellular Automaton, and Simulated Annealing	157
7.5.4	Using GP for Implementing Decision-Tree Nodes	157
7.6	On the Quality of Rules Discovered by GP.....	158
References	161

8	Evolutionary Algorithms for Clustering	165
8.1	Cluster Description-Based Individual Representation	165
8.1.1	Individual Representation	166
8.1.2	Genetic Operators	167
8.2	Centroid/Medoid-Based Individual Representation	169
8.2.1	Centroid-Based Individual Representation	169
8.2.2	Medoid-Based Individual Representation	171
8.3	Instance-Based Individual Representation	172
8.3.1	Hybrid Genetic Algorithm/ <i>K</i> -Means	172
8.3.2	Graph-Based Individual Representation	173
8.4	Fitness Evaluation	174
8.4.1	Coping with Degenerate Partitions	176
8.5	EAs vs Conventional Clustering Techniques	177
	References	177
9	Evolutionary Algorithms for Data Preparation	179
9.1	EAs for Attribute Selection	179
9.1.1	Individual Encoding and Genetic Operators	180
9.1.2	Fitness Evaluation	182
9.1.3	Attribute Selection via Search for Partially Defined Instances	187
9.1.4	Joint Attribute Selection and Instance Selection	187
9.1.5	Attribute Selection for Clustering	189
9.1.6	Discussion	190
9.2	EAs for Attribute Weighting	191
9.2.1	Attribute Weighting	192
9.2.2	Towards Constructive Induction	193
9.3	Combining Attribute Selection and Attribute Weighting	194
9.4	GP for Attribute Construction	196
9.4.1	Constructing Combinations of Boolean Attributes	197
9.4.2	Discovering Specialized Functions	199
9.5	Combining Attribute Selection and Construction with a Hybrid GA/GP	200
	References	201
10	Evolutionary Algorithms for Discovering Fuzzy Rules	205
10.1	Basic Concepts of Fuzzy Sets	205
10.1.1	Fuzzy Sets	205
10.1.2	Operations on Fuzzy Sets	207

10.1.3 Membership Functions.....	209
10.2 Fuzzy Prediction Rules vs Crisp Prediction Rules	213
10.3 A Simple Taxonomy of EAs for Fuzzy-Rule Discovery.....	215
10.4 Using EAs for Generating Fuzzy Rules	215
10.4.1 Individual Encoding.....	216
10.4.2 Determining the Degree of Matching Between a Fuzzy Rule Antecedent and a Data Instance	220
10.4.3 Using Fuzzy Rules to Classify a Data Instance.....	221
10.4.4 Specifying the Shape and the Number of Membership Functions	222
10.5 Using EAs for Tuning Membership Functions.....	224
10.6 Using EAs for Both Generating Fuzzy Rules and Tuning Membership Functions.....	225
10.7 Fuzzy Fitness Evaluation	228
References	230
11 Scaling up Evolutionary Algorithms for Large Data Sets	233
11.1 Using Data Subsets in Fitness Evaluation	233
11.1.1 Random Training-Subset Selection.....	234
11.1.2 Adaptive Training-Subset Selection	235
11.2 An Overview of Parallel Processing.....	237
11.2.1 Basic Concepts of Parallel Processing	237
11.2.2 Load Balancing	239
11.2.3 Data Parallelism vs Control Parallelism.....	240
11.2.4 Speed up and Efficiency Measures	242
11.3 Parallel EAs for Data Mining.....	243
11.3.1 Exploiting Control Parallelism.....	246
11.3.2 Exploiting Data Parallelism	248
11.3.3 Exploiting Hybrid Control/Data-Parallelism	250
References	253
12 Conclusions and Research Directions.....	255
12.1 General Remarks on Data Mining with EAs	255
12.1.1 On Predictive Accuracy	255
12.1.2 On Knowledge Comprehensibility.....	256
12.1.3 On Computational Time	256
12.2 Research Directions.....	257
12.2.1 Developing EAs for Data Preparation.....	257

12.2.2 Developing Multi-objective EAs for Data Mining	258
12.2.3 Developing a GP System for Algorithm Induction	258
References	260
Index	263



<http://www.springer.com/978-3-540-43331-6>

Data Mining and Knowledge Discovery with Evolutionary Algorithms

Freitas, A.A.

2002, XIV, 265 p., Hardcover

ISBN: 978-3-540-43331-6