

Contents

1	Introduction	1
1.1	Learning Machines and Meta-Learning	2
1.2	Basic Definitions and Notations	4
2	Techniques of Decision Tree Induction	11
2.1	Recursive Top-Down Splits	12
2.2	Univariate Decision Trees	14
2.2.1	ID3	14
2.2.2	CART	16
2.2.3	C4.5	18
2.2.4	Cal5	20
2.2.5	FACT, QUEST and CRUISE	23
2.2.6	CTree	31
2.2.7	SSV	34
2.2.8	ROC-Based Trees	36
2.3	Multivariate Decision Trees	39
2.3.1	LMDT	40
2.3.2	OC1	43
2.3.3	LTree, QTree and LgTree	45
2.3.4	DT-SE, DT-SEP and DT-SEPIR	48
2.3.5	LDT	50
2.3.6	Dipolar Criteria for DT Induction	53
2.4	Generalization Capabilities of Decision Trees	55
2.4.1	Stop Criteria	56
2.4.2	Direct Pruning Methods	57
2.4.3	Validation Based Pruning	63
2.5	Search Methods for Decision Tree Induction	70
2.6	Decision Making with Tree Structures	77
2.7	Unbiased Feature Selection	79
2.8	Ensembles of Decision Trees	86
2.8.1	Option Decision Trees	88
2.8.2	Bagging and Wagging	89
2.8.3	Boosting	91

2.8.4	Random Forests.	97
2.9	Other Interesting Approaches Related to DT Induction.	98
2.10	Meta-Learning Germs.	106
	References	107
3	Unified View of Decision Tree Induction Algorithms	119
3.1	Decision Tree Construction	119
3.1.1	Search Strategies	121
3.1.2	Node Splitters	123
3.1.3	Stop Criteria	125
3.1.4	Split Acceptors	126
3.1.5	Split Prospects Estimators.	126
3.1.6	Decision Making Modules	127
3.1.7	Data Transformations	128
3.1.8	Some Implicit Details.	129
3.2	Decision Tree Refinement.	129
3.2.1	Stop Criteria	129
3.2.2	Direct Pruning.	130
3.2.3	DT Validation	130
3.2.4	Pruning Methods Parameters.	132
3.3	Well Known Algorithms as Instances of the Uniform Approach	135
3.4	Framework Facilities	135
	References	137
4	Intemi: Advanced Meta-Learning Framework	139
4.1	Machines and Models.	140
4.1.1	Feature Selection and Ranking	143
4.1.2	Schemes and Configuration Templates	144
4.1.3	Transform and Classify Machine	147
4.1.4	Repeater Machine	148
4.2	Machine Factory	150
4.2.1	Inputs Resolution.	152
4.2.2	Machine Unification System	155
4.2.3	Task Management	162
4.3	Results and Query System.	166
4.3.1	Results Repositories.	167
4.3.2	Query System	169
4.3.3	Series and Series Transformations	170
4.4	Meta-Learning Support.	173
4.4.1	Meta Parameter Search.	173
4.4.2	Meta Search Scenarios	180
	References	181

5	Meta-Level Analysis of Decision Tree Induction	183
5.1	Results Comparison Techniques	184
5.1.1	Bad Testing Practices	184
5.1.2	Reliable and Just Comparisons	188
5.2	Test Scenarios for DT Induction Analyses	190
5.3	Single Decision Tree Models	191
5.3.1	Algorithms	192
5.3.2	Experiment	193
5.3.3	General Results Visualization and Analysis	194
5.3.4	Analysis of Results Subgroups	200
5.3.5	Summary	204
5.4	Cross-Validation Committees	205
5.4.1	DTCV Committee Algorithm	206
5.4.2	Experiment	209
5.4.3	Win Counts	211
5.4.4	DTCV Committees Versus Single Validated Trees	213
5.4.5	DTCV Committees Versus Bagging and Boosting	216
5.4.6	Algorithm Parameters Analysis	220
5.4.7	Summary	228
	References	230
6	Meta-Learning	233
6.1	Meta-Learning Approaches	235
6.1.1	No Free Lunch Theorems	237
6.1.2	Ensembles of Decision Models	242
6.1.3	Meta-Level Regression	243
6.1.4	Rankings of Algorithms	244
6.1.5	Meta-Learning as Active Search	249
6.2	Meta-Learning as Search with Feedback from Validation	250
6.2.1	The Algorithm	252
6.2.2	Proper Meta-Learners	254
6.2.3	Task Requests and Task Running	257
6.3	Meta-Learning with Configuration Generators and Complexity Control	259
6.3.1	CDML as an Instance of GML	259
6.3.2	Machine Configuration Generators	260
6.3.3	Complexity Control	264
6.3.4	Analysis of Finished Tasks and the Quarantine	266
6.3.5	Machine Complexity Evaluation	269
6.3.6	Learning Evaluators	277
6.3.7	Example Experiment	279
6.4	Profile-Based Meta-Learning	293
6.4.1	The Algorithm	295
6.4.2	Profile Management	299

6.4.3	Ranking-Based Meta-Search	301
6.4.4	Comparing Rankings of Algorithms	305
6.4.5	Experiments, Results and Analyses	306
	References	313
7	Future Perspectives of Meta-Learning	319
	Appendix A.	325
	Appendix B.	333
	Index	337



<http://www.springer.com/978-3-319-00959-9>

Meta-Learning in Decision Tree Induction

Grąbczewski, K.

2014, XVI, 343 p. 33 illus., Hardcover

ISBN: 978-3-319-00959-9