

Contents

1	A Simple Machine-Learning Task	1
1.1	Training Sets and Classifiers	1
1.2	Minor Digression: Hill-Climbing Search	5
1.3	Hill Climbing in Machine Learning	8
1.4	The Induced Classifier’s Performance.....	11
1.5	Some Difficulties with Available Data	13
1.6	Summary and Historical Remarks	15
1.7	Solidify Your Knowledge	16
2	Probabilities: Bayesian Classifiers	19
2.1	The Single-Attribute Case	19
2.2	Vectors of Discrete Attributes	22
2.3	Probabilities of Rare Events: Exploiting the Expert’s Intuition	26
2.4	How to Handle Continuous Attributes	30
2.5	Gaussian “Bell” Function: A Standard <i>pdf</i>	33
2.6	Approximating PDFs with Sets of Gaussians	34
2.7	Summary and Historical Remarks	36
2.8	Solidify Your Knowledge	40
3	Similarities: Nearest-Neighbor Classifiers	43
3.1	The <i>k</i> -Nearest-Neighbor Rule	43
3.2	Measuring Similarity	46
3.3	Irrelevant Attributes and Scaling Problems	49
3.4	Performance Considerations	52
3.5	Weighted Nearest Neighbors	55
3.6	Removing Dangerous Examples.....	57
3.7	Removing Redundant Examples.....	59
3.8	Summary and Historical Remarks	61
3.9	Solidify Your Knowledge	62

4	Inter-Class Boundaries: Linear and Polynomial Classifiers	65
4.1	The Essence	65
4.2	The Additive Rule: Perceptron Learning	69
4.3	The Multiplicative Rule: WINNOW	73
4.4	Domains with More Than Two Classes	76
4.5	Polynomial Classifiers	79
4.6	Specific Aspects of Polynomial Classifiers	81
4.7	Numerical Domains and Support Vector Machines	84
4.8	Summary and Historical Remarks	86
4.9	Solidify Your Knowledge	87
5	Artificial Neural Networks	91
5.1	Multilayer Perceptrons as Classifiers	91
5.2	Neural Network's Error	95
5.3	Backpropagation of Error	97
5.4	Special Aspects of Multilayer Perceptrons	100
5.5	Architectural Issues	104
5.6	Radial-Basis Function Networks	106
5.7	Summary and Historical Remarks	109
5.8	Solidify Your Knowledge	110
6	Decision Trees	113
6.1	Decision Trees as Classifiers	113
6.2	Induction of Decision Trees	117
6.3	How Much Information Does an Attribute Convey?	119
6.4	Binary Split of a Numeric Attribute	122
6.5	Pruning	126
6.6	Converting the Decision Tree into Rules	130
6.7	Summary and Historical Remarks	132
6.8	Solidify Your Knowledge	133
7	Computational Learning Theory	137
7.1	PAC Learning	137
7.2	Examples of PAC Learnability	141
7.3	Some Practical and Theoretical Consequences	143
7.4	VC-Dimension and Learnability	145
7.5	Summary and Historical Remarks	148
7.6	Exercises and Thought Experiments	149
8	A Few Instructive Applications	151
8.1	Character Recognition	151
8.2	Oil-Spill Recognition	155
8.3	Sleep Classification	158
8.4	Brain-Computer Interface	161
8.5	Medical Diagnosis	165

8.6	Text Classification	167
8.7	Summary and Historical Remarks	169
8.8	Exercises and Thought Experiments	170
9	Induction of Voting Assemblies	173
9.1	Bagging	173
9.2	Schapire’s Boosting	176
9.3	Adaboost: Practical Version of Boosting	179
9.4	Variations on the Boosting Theme	183
9.5	Cost-Saving Benefits of the Approach	185
9.6	Summary and Historical Remarks	187
9.7	Solidify Your Knowledge	188
10	Some Practical Aspects to Know About	191
10.1	A Learner’s Bias	191
10.2	Imbalanced Training Sets	194
10.3	Context-Dependent Domains	199
10.4	Unknown Attribute Values	202
10.5	Attribute Selection	204
10.6	Miscellaneous	206
10.7	Summary and Historical Remarks	208
10.8	Solidify Your Knowledge	208
11	Performance Evaluation	211
11.1	Basic Performance Criteria	211
11.2	Precision and Recall	214
11.3	Other Ways to Measure Performance	219
11.4	Learning Curves and Computational Costs	222
11.5	Methodologies of Experimental Evaluation	224
11.6	Summary and Historical Remarks	227
11.7	Solidify Your Knowledge	228
12	Statistical Significance	231
12.1	Sampling a Population	231
12.2	Benefiting from the Normal Distribution	235
12.3	Confidence Intervals	239
12.4	Statistical Evaluation of a Classifier	241
12.5	Another Kind of Statistical Evaluation	244
12.6	Comparing Machine-Learning Techniques	245
12.7	Summary and Historical Remarks	247
12.8	Solidify Your Knowledge	248
13	Induction in Multi-Label Domains	251
13.1	Classical Machine Learning in Multi-Label Domains	251
13.2	Treating Each Class Separately: Binary Relevance	254
13.3	Classifier Chains	256



<http://www.springer.com/978-3-319-63912-3>

An Introduction to Machine Learning

Kubat, M.

2017, XIII, 348 p. 85 illus., 3 illus. in color., Hardcover

ISBN: 978-3-319-63912-3