Machine learning algorithms 2023: Contents

0. General 1

1. Introduction 6

1.1 About machine learning: Example 1 7

1.2 Tasks: about problems that can be solved with machine learning 21

1.3 Variables in machine learning 36

2. On classification and related tasks 46

2.1 Binary classification 49

2.2 Beyond binary classification 69

3. Probabilistic models 81
3.1 Turning data into probabilities 81
3.2 Minimising risk 95
3.3 Naive Bayes classifier 100
3.4 Some useful statistical formulas 108
3.5 The Gaussian 117
4. Nearest neighbor searching 127
4.1 Foundation of nearest neighbors 127
4.2. Weighted nearest neighbor searching 139

5. Discriminant analysis 147
5.1 Quadratic discriminant analysis 153
5.2 Linear discriminant analysis 161
5.3 Logistic discrimination 172
6. Support vector machines 179
6.1 Linearly separable data (classes) 180
6.2 Nonseparable classes 192
6.3 Multiclass classification 197

7. Decision or classification trees 208
7.1 Construction of decision trees 212
7.2 ID3 216
7.3 Dealing with continuous variables 230
7.4 Computational complexity 235
7.5 Classification and regression trees (CART) 236
8. Ensemble learning 249
8.1 Boosting 254
8.2 Bootstrap 262

8.3 Bagging 267
8.4 Random forests 268
9. Clustering 272
9.1 Partition-based or objective function-based clustering 274
9.2 Hierarchical clustering 282
10. Evolutionary learning 298
10.1. The genetic algorithm (GA) 299
10.2 Generating offspring: genetic operators 307
10.3 The basic genetic algorithm 314