Contents

Preface **xvii**

Notations **xxi**

1 ***Introduction*** **1**

1.1 What Is Machine Learning? 1
1.2 Examples of Machine Learning Applications 4
1.2.1 Learning Associations 4
1.2.2 Classification 5
1.2.3 Regression 9
1.2.4 Unsupervised Learning 11
1.2.5 Reinforcement Learning 13
1.3 Notes 14
1.4 Relevant Resources 17
1.5 Exercises 18
1.6 References 20

2 ***Supervised Learning*** **21**

2.1 Learning a Class from Examples 21
2.2 Vapnik-Chervonenkis Dimension 27
2.3 Probably Approximately Correct Learning 29
2.4 Noise 30
2.5 Learning Multiple Classes 32
2.6 Regression 34
2.7 Model Selection and Generalization 37
2.8 Dimensions of a Supervised Machine Learning Algorithm 41
2.9 Notes 42

Copyright © 2014. MIT Press. All rights reserved.

Alpaydin, Ethem. Introduction to Machine Learning, MIT Press, 2014. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/jiouniv/detail.action?docID=3339851.
Created from jiouniv on 2023-03-30 13:02:46.

2.10 Exercises 43
2.11 References 47

3 *Bayesian Decision Theory* 49
3.1 Introduction 49
3.2 Classification 51
3.3 Losses and Risks 53
3.4 Discriminant Functions 55
3.5 Association Rules 56
3.6 Notes 59
3.7 Exercises 60
3.8 References 64

4 *Parametric Methods* 65
4.1 Introduction 65
4.2 Maximum Likelihood Estimation 66
4.2.1 Bernoulli Density 67
4.2.2 Multinomial Density 68
4.2.3 Gaussian (Normal) Density 68
4.3 Evaluating an Estimator: Bias and Variance 69
4.4 The Bayes' Estimator 70
4.5 Parametric Classification 73
4.6 Regression 77
4.7 Tuning Model Complexity: Bias/Variance Dilemma 80
4.8 Model Selection Procedures 83
4.9 Notes 87
4.10 Exercises 88
4.11 References 90

5 *Multivariate Methods* 93
5.1 Multivariate Data 93
5.2 Parameter Estimation 94
5.3 Estimation of Missing Values 95
5.4 Multivariate Normal Distribution 96
5.5 Multivariate Classification 100
5.6 Tuning Complexity 106
5.7 Discrete Features 108
5.8 Multivariate Regression 109
5.9 Notes 111
5.10 Exercises 112

Copyright © 2014. MIT Press. All rights reserved.

Alpaydin, Ethem. Introduction to Machine Learning, MIT Press, 2014. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/jiouniv/detail.action?docID=3339851.
Created from jiouniv on 2023-03-30 13:02:46.

5.11 References 113

6 *Dimensionality Reduction* 115

6.1 Introduction 115
6.2 Subset Selection 116
6.3 Principal Component Analysis 120
6.4 Feature Embedding 127
6.5 Factor Analysis 130
6.6 Singular Value Decomposition and Matrix Factorization 135
6.7 Multidimensional Scaling 136
6.8 Linear Discriminant Analysis 140
6.9 Canonical Correlation Analysis 145
6.10 Isomap 148
6.11 Locally Linear Embedding 150
6.12 Laplacian Eigenmaps 153
6.13 Notes 155
6.14 Exercises 157
6.15 References 158

7 *Clustering* 161

7.1 Introduction 161
7.2 Mixture Densities 162
7.3 *k*-Means Clustering 163
7.4 Expectation-Maximization Algorithm 167
7.5 Mixtures of Latent Variable Models 172
7.6 Supervised Learning after Clustering 173
7.7 Spectral Clustering 175
7.8 Hierarchical Clustering 176
7.9 Choosing the Number of Clusters 178
7.10 Notes 179
7.11 Exercises 180
7.12 References 182

8 *Nonparametric Methods* 185

8.1 Introduction 185
8.2 Nonparametric Density Estimation 186
8.2.1 Histogram Estimator 187
8.2.2 Kernel Estimator 188
8.2.3 *k*-Nearest Neighbor Estimator 190
8.3 Generalization to Multivariate Data 192

Copyright © 2014. MIT Press. All rights reserved.

Alpaydin, Ethem. Introduction to Machine Learning, MIT Press, 2014. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/jiouniv/detail.action?docID=3339851.
Created from jiouniv on 2023-03-30 13:02:46.

8.4 Nonparametric Classification 193
8.5 Condensed Nearest Neighbor 194
8.6 Distance-Based Classification 196
8.7 Outlier Detection 199
8.8 Nonparametric Regression: Smoothing Models 201
8.8.1 Running Mean Smoother 201
8.8.2 Kernel Smoother 203
8.8.3 Running Line Smoother 204
8.9 How to Choose the Smoothing Parameter 204
8.10 Notes 205
8.11 Exercises 208
8.12 References 210

9 *Decision Trees* 213

9.1 Introduction 213
9.2 Univariate Trees 215
9.2.1 Classification Trees 216
9.2.2 Regression Trees 220
9.3 Pruning 222
9.4 Rule Extraction from Trees 225
9.5 Learning Rules from Data 226
9.6 Multivariate Trees 230
9.7 Notes 232
9.8 Exercises 235
9.9 References 237

10 *Linear Discrimination* 239

10.1 Introduction 239
10.2 Generalizing the Linear Model 241
10.3 Geometry of the Linear Discriminant 242
10.3.1 Two Classes 242
10.3.2 Multiple Classes 244
10.4 Pairwise Separation 246
10.5 Parametric Discrimination Revisited 247
10.6 Gradient Descent 248
10.7 Logistic Discrimination 250
10.7.1 Two Classes 250
10.7.2 Multiple Classes 254
10.8 Discrimination by Regression 257

Copyright © 2014. MIT Press. All rights reserved.

Alpaydin, Ethem. Introduction to Machine Learning, MIT Press, 2014. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/jiouniv/detail.action?docID=3339851.
Created from jiouniv on 2023-03-30 13:02:46.

10.9 Learning to Rank 260
10.10 Notes 263
10.11 Exercises 263
10.12 References 266

11 *Multilayer Perceptrons* 267

11.1 Introduction 267
11.1.1 Understanding the Brain 268
11.1.2 Neural Networks as a Paradigm for Parallel Processing 269
11.2 The Perceptron 271
11.3 Training a Perceptron 274
11.4 Learning Boolean Functions 277
11.5 Multilayer Perceptrons 279
11.6 MLP as a Universal Approximator 281
11.7 Backpropagation Algorithm 283
11.7.1 Nonlinear Regression 284
11.7.2 Two-Class Discrimination 286
11.7.3 Multiclass Discrimination 288
11.7.4 Multiple Hidden Layers 290
11.8 Training Procedures 290
11.8.1 Improving Convergence 290
11.8.2 Overtraining 291
11.8.3 Structuring the Network 292
11.8.4 Hints 295
11.9 Tuning the Network Size 297
11.10 Bayesian View of Learning 300
11.11 Dimensionality Reduction 301
11.12 Learning Time 304
11.12.1 Time Delay Neural Networks 304
11.12.2 Recurrent Networks 305
11.13 Deep Learning 306
11.14 Notes 309
11.15 Exercises 311
11.16 References 313

12 *Local Models* 317

12.1 Introduction 317
12.2 Competitive Learning 318

Copyright © 2014. MIT Press. All rights reserved.

Alpaydin, Ethem. Introduction to Machine Learning, MIT Press, 2014. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/jiouniv/detail.action?docID=3339851.
Created from jiouniv on 2023-03-30 13:02:46.

12.2.1 Online k-Means 318
12.2.2 Adaptive Resonance Theory 323
12.2.3 Self-Organizing Maps 324
12.3 Radial Basis Functions 326
12.4 Incorporating Rule-Based Knowledge 332
12.5 Normalized Basis Functions 333
12.6 Competitive Basis Functions 335
12.7 Learning Vector Quantization 338
12.8 The Mixture of Experts 338
12.8.1 Cooperative Experts 341
12.8.2 Competitive Experts 342
12.9 Hierarchical Mixture of Experts 342
12.10 Notes 343
12.11 Exercises 344
12.12 References 347

13 *Kernel Machines* 349

13.1 Introduction 349
13.2 Optimal Separating Hyperplane 351
13.3 The Nonseparable Case: Soft Margin Hyperplane 355
13.4 ν-SVM 358
13.5 Kernel Trick 359
13.6 Vectorial Kernels 361
13.7 Defining Kernels 364
13.8 Multiple Kernel Learning 365
13.9 Multiclass Kernel Machines 367
13.10 Kernel Machines for Regression 368
13.11 Kernel Machines for Ranking 373
13.12 One-Class Kernel Machines 374
13.13 Large Margin Nearest Neighbor Classifier 377
13.14 Kernel Dimensionality Reduction 379
13.15 Notes 380
13.16 Exercises 382
13.17 References 383

14 *Graphical Models* 387

14.1 Introduction 387
14.2 Canonical Cases for Conditional Independence 389
14.3 Generative Models 396

Copyright © 2014. MIT Press. All rights reserved.

Alpaydin, Ethem. Introduction to Machine Learning, MIT Press, 2014. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/jiouniv/detail.action?docID=3339851.
Created from jiouniv on 2023-03-30 13:02:46.

14.4 d-Separation 399
14.5 Belief Propagation 399
14.5.1 Chains 400
14.5.2 Trees 402
14.5.3 Polytrees 404
14.5.4 Junction Trees 406
14.6 Undirected Graphs: Markov Random Fields 407
14.7 Learning the Structure of a Graphical Model 410
14.8 Influence Diagrams 411
14.9 Notes 412
14.10 Exercises 413
14.11 References 415

15 *Hidden Markov Models* 417

15.1 Introduction 417
15.2 Discrete Markov Processes 418
15.3 Hidden Markov Models 421
15.4 Three Basic Problems of HMMs 423
15.5 Evaluation Problem 423
15.6 Finding the State Sequence 427
15.7 Learning Model Parameters 429
15.8 Continuous Observations 432
15.9 The HMM as a Graphical Model 433
15.10 Model Selection in HMMs 436
15.11 Notes 438
15.12 Exercises 440
15.13 References 443

16 *Bayesian Estimation* 445

16.1 Introduction 445
16.2 Bayesian Estimation of the Parameters of a Discrete Distribution 449
16.2.1 $K > 2$ States: Dirichlet Distribution 449
16.2.2 $K = 2$ States: Beta Distribution 450
16.3 Bayesian Estimation of the Parameters of a Gaussian Distribution 451
16.3.1 Univariate Case: Unknown Mean, Known Variance 451

Copyright © 2014. MIT Press. All rights reserved.

Alpaydin, Ethem. Introduction to Machine Learning, MIT Press, 2014. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/jiouniv/detail.action?docID=3339851.
Created from jiouniv on 2023-03-30 13:02:46.

16.3.2 Univariate Case: Unknown Mean, Unknown Variance 453
16.3.3 Multivariate Case: Unknown Mean, Unknown Covariance 455
16.4 Bayesian Estimation of the Parameters of a Function 456
16.4.1 Regression 456
16.4.2 Regression with Prior on Noise Precision 460
16.4.3 The Use of Basis/Kernel Functions 461
16.4.4 Bayesian Classification 463
16.5 Choosing a Prior 466
16.6 Bayesian Model Comparison 467
16.7 Bayesian Estimation of a Mixture Model 470
16.8 Nonparametric Bayesian Modeling 473
16.9 Gaussian Processes 474
16.10 Dirichlet Processes and Chinese Restaurants 478
16.11 Latent Dirichlet Allocation 480
16.12 Beta Processes and Indian Buffets 482
16.13 Notes 483
16.14 Exercises 484
16.15 References 485

17 *Combining Multiple Learners* 487

17.1 Rationale 487
17.2 Generating Diverse Learners 488
17.3 Model Combination Schemes 491
17.4 Voting 492
17.5 Error-Correcting Output Codes 496
17.6 Bagging 498
17.7 Boosting 499
17.8 The Mixture of Experts Revisited 502
17.9 Stacked Generalization 504
17.10 Fine-Tuning an Ensemble 505
17.10.1 Choosing a Subset of the Ensemble 506
17.10.2 Constructing Metalearners 506
17.11 Cascading 507
17.12 Notes 509
17.13 Exercises 511
17.14 References 513

Copyright © 2014. MIT Press. All rights reserved.

Alpaydin, Ethem. Introduction to Machine Learning, MIT Press, 2014. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/jiouniv/detail.action?docID=3339851.
Created from jiouniv on 2023-03-30 13:02:46.

18 *Reinforcement Learning* 517

18.1 Introduction 517
18.2 Single State Case: K-Armed Bandit 519
18.3 Elements of Reinforcement Learning 520
18.4 Model-Based Learning 523
18.4.1 Value Iteration 523
18.4.2 Policy Iteration 524
18.5 Temporal Difference Learning 525
18.5.1 Exploration Strategies 525
18.5.2 Deterministic Rewards and Actions 526
18.5.3 Nondeterministic Rewards and Actions 527
18.5.4 Eligibility Traces 530
18.6 Generalization 531
18.7 Partially Observable States 534
18.7.1 The Setting 534
18.7.2 Example: The Tiger Problem 536
18.8 Notes 541
18.9 Exercises 542
18.10 References 544

19 *Design and Analysis of Machine Learning Experiments* 547

19.1 Introduction 547
19.2 Factors, Response, and Strategy of Experimentation 550
19.3 Response Surface Design 553
19.4 Randomization, Replication, and Blocking 554
19.5 Guidelines for Machine Learning Experiments 555
19.6 Cross-Validation and Resampling Methods 558
19.6.1 K-Fold Cross-Validation 559
19.6.2 5×2 Cross-Validation 560
19.6.3 Bootstrapping 561
19.7 Measuring Classifier Performance 561
19.8 Interval Estimation 564
19.9 Hypothesis Testing 568
19.10 Assessing a Classification Algorithm's Performance 570
19.10.1 Binomial Test 571
19.10.2 Approximate Normal Test 572
19.10.3 t Test 572
19.11 Comparing Two Classification Algorithms 573
19.11.1 McNemar's Test 573

Copyright © 2014. MIT Press. All rights reserved.

Alpaydin, Ethem. Introduction to Machine Learning, MIT Press, 2014. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/jiouniv/detail.action?docID=3339851.
Created from jiouniv on 2023-03-30 13:02:46.

19.11.2 K-Fold Cross-Validated Paired t Test 573
19.11.3 5×2 cv Paired t Test 574
19.11.4 5×2 cv Paired F Test 575
19.12 Comparing Multiple Algorithms: Analysis of Variance 576
19.13 Comparison over Multiple Datasets 580
19.13.1 Comparing Two Algorithms 581
19.13.2 Multiple Algorithms 583
19.14 Multivariate Tests 584
19.14.1 Comparing Two Algorithms 585
19.14.2 Comparing Multiple Algorithms 586
19.15 Notes 587
19.16 Exercises 588
19.17 References 590

A *Probability* 593

A.1 Elements of Probability 593
A.1.1 Axioms of Probability 594
A.1.2 Conditional Probability 594
A.2 Random Variables 595
A.2.1 Probability Distribution and Density Functions 595
A.2.2 Joint Distribution and Density Functions 596
A.2.3 Conditional Distributions 596
A.2.4 Bayes' Rule 597
A.2.5 Expectation 597
A.2.6 Variance 598
A.2.7 Weak Law of Large Numbers 599
A.3 Special Random Variables 599
A.3.1 Bernoulli Distribution 599
A.3.2 Binomial Distribution 600
A.3.3 Multinomial Distribution 600
A.3.4 Uniform Distribution 600
A.3.5 Normal (Gaussian) Distribution 601
A.3.6 Chi-Square Distribution 602
A.3.7 t Distribution 603
A.3.8 F Distribution 603
A.4 References 603

***Index* 605**

Copyright © 2014. MIT Press. All rights reserved.

Alpaydin, Ethem. Introduction to Machine Learning, MIT Press, 2014. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/jiouniv/detail.action?docID=3339851.
Created from jiouniv on 2023-03-30 13:02:46.