

brief contents

PART 1	INTRODUCING ML ALGORITHMS.....	1
	1 ■ Machine learning algorithms	3
	2 ■ Markov chain Monte Carlo	14
	3 ■ Variational inference	43
	4 ■ Software implementation	58
PART 2	SUPERVISED LEARNING.....	71
	5 ■ Classification algorithms	73
	6 ■ Regression algorithms	106
	7 ■ Selected supervised learning algorithms	123
PART 3	UNSUPERVISED LEARNING.....	163
	8 ■ Fundamental unsupervised learning algorithms	165
	9 ■ Selected unsupervised learning algorithms	186
PART 4	DEEP LEARNING.....	217
	10 ■ Fundamental deep learning algorithms	219
	11 ■ Advanced deep learning algorithms	250

contents

<i>preface</i>	<i>xii</i>
<i>acknowledgments</i>	<i>xiii</i>
<i>about this book</i>	<i>xiv</i>
<i>about the author</i>	<i>xviii</i>
<i>about the cover illustration</i>	<i>xix</i>

PART 1 INTRODUCING ML ALGORITHMS.....1

<i>1</i>	<i>Machine learning algorithms</i>	<i>3</i>
1.1	Types of ML algorithms	4
1.2	Why learn algorithms from scratch?	7
1.3	Mathematical background	7
1.4	Bayesian inference and deep learning	8
	<i>Two main camps of Bayesian inference: MCMC and VI</i>	<i>9</i>
	<i>Modern deep learning algorithms</i>	<i>11</i>
1.5	Implementing algorithms	12
	<i>Data structures</i>	<i>12</i> ■ <i>Problem-solving paradigms</i> <i>12</i>

2 *Markov chain Monte Carlo* 14

- 2.1 Introduction to Markov chain Monte Carlo 15
Posterior distribution of coin flips 16 ▪ *Markov chain for page rank* 18
- 2.2 Estimating π 19
- 2.3 Binomial tree model 21
- 2.4 Self-avoiding random walk 24
- 2.5 Gibbs sampling 28
- 2.6 Metropolis-Hastings sampling 32
- 2.7 Importance sampling 35
- 2.8 Exercises 41

3 *Variational inference* 43

- 3.1 KL variational inference 44
- 3.2 Mean-field approximation 47
- 3.3 Image denoising in an Ising model 49
- 3.4 MI maximization 56
- 3.5 Exercises 57

4 *Software implementation* 58

- 4.1 Data structures 58
Linear 59 ▪ *Nonlinear* 60 ▪ *Probabilistic* 61
- 4.2 Problem-solving paradigms 61
Complete search 61 ▪ *Greedy* 62 ▪ *Divide and conquer* 64
Dynamic programming 65
- 4.3 ML research: Sampling methods and variational inference 67
- 4.4 Exercises 68

PART 2 SUPERVISED LEARNING71

5 *Classification algorithms* 73

- 5.1 Introduction to classification 74
- 5.2 Perceptron 74
- 5.3 Support vector machine 80

- 5.4 Logistic regression 86
- 5.5 Naive Bayes 93
- 5.6 Decision tree (CART) 98
- 5.7 Exercises 104

6 *Regression algorithms* 106

- 6.1 Introduction to regression 107
- 6.2 Bayesian linear regression 107
- 6.3 Hierarchical Bayesian regression 111
- 6.4 KNN regression 115
- 6.5 Gaussian process regression 117
- 6.6 Exercises 121

7 *Selected supervised learning algorithms* 123

- 7.1 Markov models 124
 - Page rank algorithm* 125 ▪ *Hidden Markov models* 128
- 7.2 Imbalanced learning 134
 - Undersampling strategies* 134 ▪ *Oversampling strategies* 136
- 7.3 Active learning 138
 - Query strategies* 139
- 7.4 Model selection: Hyperparameter tuning 147
 - Bayesian optimization* 148
- 7.5 Ensemble methods 150
 - Bagging* 151 ▪ *Boosting* 154 ▪ *Stacking* 157
- 7.6 ML research: Supervised learning algorithms 160
- 7.7 Exercises 161

PART 3 UNSUPERVISED LEARNING.....163

8 *Fundamental unsupervised learning algorithms* 165

- 8.1 Dirichlet process K-means 166
- 8.2 Gaussian mixture models 171
 - Expectation maximization (EM) algorithm* 171

- 8.3 Dimensionality reduction 179
 - Principal component analysis* 179 ■ *t-SNE manifold learning on images* 182
- 8.4 Exercises 184

9 *Selected unsupervised learning algorithms* 186

- 9.1 Latent Dirichlet allocation 187
 - Variational Bayes* 188
- 9.2 Density estimators 195
 - Kernel density estimator* 195 ■ *Tangent portfolio optimization* 198
- 9.3 Structure learning 201
 - Chow-Liu algorithm* 201 ■ *Inverse covariance estimation* 202
- 9.4 Simulated annealing 206
- 9.5 Genetic algorithm 210
- 9.6 ML research: Unsupervised learning 213
- 9.7 Exercises 214

PART 4 DEEP LEARNING217

10 *Fundamental deep learning algorithms* 219

- 10.1 Multilayer perceptron 220
- 10.2 Convolutional neural nets 225
 - LeNet on MNIST* 226 ■ *ResNet image search* 229
- 10.3 Recurrent neural nets 232
 - LSTM sequence classification* 233 ■ *Multi-input model* 237
- 10.4 Neural network optimizers 242
- 10.5 Exercises 248

11 *Advanced deep learning algorithms* 250

- 11.1 Autoencoders 251
 - VAE anomaly detection in time series* 253
- 11.2 Amortized variational inference 259
 - Mixture density networks* 259
- 11.3 Attention and transformers 265

11.4	Graph neural networks	273
11.5	ML research: Deep learning	279
11.6	Exercises	281
<i>appendix A</i>	<i>Further reading and resources</i>	<i>283</i>
<i>appendix B</i>	<i>Answers to exercises</i>	<i>286</i>
	<i>index</i>	<i>299</i>