

# Brief Contents

1	Introduction	1
<b>I</b>	<b>Foundations</b>	<b>19</b>
2	Probabilistic inference	21
3	Probabilistic models	41
4	Parameter estimation	77
5	Optimization algorithms	99
6	Information theory	145
7	Bayesian statistics	163
8	Bayesian decision theory	221
<b>II</b>	<b>Linear models</b>	<b>247</b>
9	Linear discriminant analysis	249
10	Logistic regression	265
11	Linear regression	305
12	Generalized linear models	353
<b>III</b>	<b>Deep neural networks</b>	<b>369</b>
13	Neural networks for unstructured data	371
14	Neural networks for images	407
15	Neural networks for sequences	443
<b>IV</b>	<b>Nonparametric models</b>	<b>469</b>
16	Exemplar-based methods	471
17	Kernel methods	491
18	Trees, forests, bagging and boosting	533

<b>V</b>	<b>Beyond supervised learning</b>	<b>553</b>
19	Learning with fewer labeled examples	555
20	Dimensionality reduction	591
21	Clustering	639
22	Recommender systems	663
23	Graph embeddings	675
<b>VI</b>	<b>Appendix: Mathematical background</b>	<b>697</b>
A	Some useful mathematics	699
B	Linear algebra	719
C	Probability	759
D	Frequentist statistics	779
E	Exercises	815