

CONTENTS IN DETAIL

FOREWORD	xix
-----------------	------------

ACKNOWLEDGMENTS	xxiii
------------------------	--------------

INTRODUCTION	xxv
---------------------	------------

Who Is This Book For?	xxvi
What Can You Expect to Learn?	xxvii
About This Book	xxvii

1	
GETTING STARTED	1

The Operating Environment	1
NumPy	2
scikit-learn	2
Keras with TensorFlow	2
Installing the Toolkits	3
Basic Linear Algebra	4
Vectors	4
Matrices	5
Multiplying Vectors and Matrices	5
Statistics and Probability	6
Descriptive Statistics	6
Probability Distributions	7
Statistical Tests	8
Graphics Processing Units	9
Summary	9

2	
USING PYTHON	11

The Python Interpreter	11
Statements and Whitespace	12

Variables and Basic Data Structures	13
Representing Numbers	13
Variables.....	14
Strings	14
Lists	15
Dictionaries	18
Control Structures	19
if-elif-else Statements.....	19
for Loops	19
while Loops	22
break and continue Statements	22
with Statement	23
Handling Errors with try-except Blocks	24
Functions	24
Modules	26
Summary	27

3 USING NUMPY 29

Why NumPy?	29
Arrays vs. Lists	30
Testing Array and List Speed	31
Basic Arrays	33
Defining an Array with np.array.....	33
Defining Arrays with 0s and 1s.....	36
Accessing Elements in an Array.....	37
Indexing into an Array	37
Slicing an Array.....	39
The Ellipsis	41
Operators and Broadcasting.....	42
Array Input and Output	45
Random Numbers	48
NumPy and Images.....	48
Summary	50

4 WORKING WITH DATA 51

Classes and Labels	51
Features and Feature Vectors	52
Types of Features.....	53
Feature Selection and the Curse of Dimensionality	55

Features of a Good Dataset	57
Interpolation and Extrapolation	58
The Parent Distribution	60
Prior Class Probabilities	60
Confusers	61
Dataset Size	62
Data Preparation	63
Scaling Features	63
Missing Features	67
Training, Validation, and Test Data	68
The Three Subsets	68
Partitioning the Dataset	69
k-Fold Cross Validation	74
Look at Your Data	76
Searching for Problems in the Data	76
Cautionary Tales	80
Summary	81

5 BUILDING DATASETS 83

Iris	84
Breast Cancer	86
MNIST Digits	88
CIFAR-10	90
Data Augmentation	92
Why Should You Augment Training Data?	93
Ways to Augment Training Data	94
Augmenting the Iris Dataset	95
Augmenting the CIFAR-10 Dataset	101
Summary	105

6 CLASSICAL MACHINE LEARNING 107

Nearest Centroid	108
k-Nearest Neighbors	112
Naïve Bayes	113
Decision Trees and Random Forests	117
Recursion Primer	120
Building Decision Trees	121
Random Forests	122

Support Vector Machines	124
Margins	124
Support Vectors	126
Optimization	126
Kernels	127
Summary	128

7 EXPERIMENTS WITH CLASSICAL MODELS 129

Experiments with the Iris Dataset	129
Testing the Classical Models	130
Implementing a Nearest Centroid Classifier	133
Experiments with the Breast Cancer Dataset	135
Two Initial Test Runs	135
The Effect of Random Splits	138
Adding k-fold Validation	140
Searching for Hyperparameters	145
Experiments with the MNIST Dataset	150
Testing the Classical Models	150
Analyzing Runtimes	156
Experimenting with PCA Components	158
Scrambling Our Dataset	161
Classical Model Summary	162
Nearest Centroid	162
k-Nearest Neighbors	163
Naïve Bayes	163
Decision Trees	164
Random Forests	164
Support Vector Machines	165
When to Use Classical Models	165
Handling Small Datasets	165
Dealing with Reduced Computational Requirements	165
Having Explainable Models	166
Working with Vector Inputs	166
Summary	167

8 INTRODUCTION TO NEURAL NETWORKS 169

Anatomy of a Neural Network	170
The Neuron	171
Activation Functions	172
Architecture of a Network	176
Output Layers	178
Representing Weights and Biases	180

Implementing a Simple Neural Network	182
Building the Dataset	182
Implementing the Neural Network	183
Training and Testing the Neural Network	185
Summary	188

9

TRAINING A NEURAL NETWORK 189

A High-Level Overview	190
Gradient Descent	190
Finding Minimums	192
Updating the Weights	193
Stochastic Gradient Descent	194
Batches and Minibatches	195
Convex vs. Nonconvex Functions	196
Ending Training	197
Updating the Learning Rate	198
Momentum	199
Backpropagation	200
Backprop, Take 1	200
Backprop, Take 2	204
Loss Functions	208
Absolute and Mean Squared Error Loss	209
Cross-Entropy Loss	210
Weight Initialization	211
Overfitting and Regularization	213
Understanding Overfitting	213
Understanding Regularization	215
L2 Regularization	216
Dropout	217
Summary	219

10

EXPERIMENTS WITH NEURAL NETWORKS 221

Our Dataset	222
The MLPClassifier Class	222
Architecture and Activation Functions	223
The Code	223
The Results	227
Batch Size	231
Base Learning Rate	235
Training Set Size	238
L2 Regularization	239
Momentum	242
Weight Initialization	243

Feature Ordering	247
Summary	249

11 EVALUATING MODELS 251

Definitions and Assumptions	251
Why Accuracy Is Not Enough	252
The 2 x 2 Confusion Matrix	254
Metrics Derived from the 2 x 2 Confusion Matrix	257
Deriving Metrics from the 2 x 2 Table	257
Using Our Metrics to Interpret Models	260
More Advanced Metrics	262
Informedness and Markedness	262
F1 Score	263
Cohen's Kappa	263
Matthews Correlation Coefficient	264
Implementing Our Metrics	264
The Receiver Operating Characteristics Curve	266
Gathering Our Models	266
Plotting Our Metrics	268
Exploring the ROC Curve	269
Comparing Models with ROC Analysis	271
Generating an ROC Curve	273
The Precision–Recall Curve	275
Handling Multiple Classes	276
Extending the Confusion Matrix	276
Calculating Weighted Accuracy	279
Multiclass Matthews Correlation Coefficient	281
Summary	282

12 INTRODUCTION TO CONVOLUTIONAL NEURAL NETWORKS 283

Why Convolutional Neural Networks?	284
Convolution	284
Scanning with the Kernel	285
Convolution for Image Processing	287
Anatomy of a Convolutional Neural Network	288
Different Types of Layers	289
Passing Data Through the CNN	291
Convolutional Layers	292
How a Convolution Layer Works	292
Using a Convolutional Layer	295
Multiple Convolutional Layers	298
Initializing a Convolutional Layer	299

Pooling Layers	299
Fully Connected Layers	301
Fully Convolutional Layers	302
Step by Step	304
Summary	308

13

EXPERIMENTS WITH KERAS AND MNIST 309

Building CNNs in Keras	310
Loading the MNIST Data	310
Building Our Model	312
Training and Evaluating the Model	314
Plotting the Error	317
Basic Experiments	319
Architecture Experiments	319
Training Set Size, Minibatches, and Epochs	323
Optimizers	326
Fully Convolutional Networks	328
Building and Training the Model	328
Making the Test Images	331
Testing the Model	333
Scrambled MNIST Digits	340
Summary	342

14

EXPERIMENTS WITH CIFAR-10 343

A CIFAR-10 Refresher	343
Working with the Full CIFAR-10 Dataset	344
Building the Models	345
Analyzing the Models	348
Animal or Vehicle?	352
Binary or Multiclass?	357
Transfer Learning	361
Fine-Tuning a Model	367
Building Our Datasets	368
Adapting Our Model for Fine-Tuning	371
Testing Our Model	373
Summary	375

15

A CASE STUDY: CLASSIFYING AUDIO SAMPLES 377

Building the Dataset	378
Augmenting the Dataset	379
Preprocessing Our Data	383

Classifying the Audio Features	385
Using Classical Models	385
Using a Traditional Neural Network	388
Using a Convolutional Neural Network	389
Spectrograms	394
Classifying Spectrograms	398
Initialization, Regularization, and Batch Normalization	402
Examining the Confusion Matrix	403
Ensembles	404
Summary	408

16
GOING FURTHER **411**

Going Further with CNNs	411
Reinforcement Learning and Unsupervised Learning	412
Generative Adversarial Networks	413
Recurrent Neural Networks	414
Online Resources	414
Conferences	415
The Book	416
So Long and Thanks for All the Fish	416

INDEX **417**