

CONTENTS IN DETAIL

FOREWORD TO THE FIRST EDITION	xxi
--------------------------------------	------------

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

INTRODUCTION	xxv
Who This Book Is For	xxvi
What You Can Expect to Learn	xxvii
About This Book	xxvii
Terminology	xxviii
What's New in the Second Edition	xxx
Synopsis	xxx

0	
ENVIRONMENT AND MATHEMATICAL PRELIMINARIES	1
The Operating Environment	1
NumPy	2
scikit-learn	2
TensorFlow with Keras	2
Installing the Toolkits	2
Basic Linear Algebra	4
Vectors	4
Matrices	5
Vector and Matrix Multiplication	5
Statistics and Probability	6
Descriptive Statistics	6
Probability Distributions	7
Statistical Tests	8
Graphics Processing Units	9
Summary	9

PART I

DATA IS EVERYTHING

1

IT'S ALL ABOUT THE DATA

13

Classes and Labels	13
Features and Feature Vectors	14
Types of Features	15
Feature Selection and the Curse of Dimensionality	17
Qualities of a Good Dataset	19
Interpolation and Extrapolation	20
The Parent Distribution	22
Prior Class Probabilities	22
Confusers	23
Dataset Size	24
Data Preparation	25
Scaling Features	25
Dealing with Missing Features	29
Training, Validation, and Test Data	30
The Three Subsets	30
Dataset Partitioning	31
k-Fold Cross-Validation	36
Data Analysis	37
How to Find Problems in the Data	38
Cautionary Tales	41
Summary	42

2

BUILDING THE DATASETS

43

Iris	43
Breast Cancer	45
MNIST Digits	47
CIFAR-10	50
Data Augmentation	52
Reasoning	53
Methods	54
The Iris Dataset	55
The CIFAR-10 Dataset	60
Summary	64

PART II CLASSICAL MACHINE LEARNING

3		
INTRODUCTION TO MACHINE LEARNING		67
Nearest Centroid		68
k-Nearest Neighbors		72
Naive Bayes		73
Tree Classifiers		77
Recursion Primer		80
Decision Trees		81
Random Forests		82
Support Vector Machines		83
Margins		84
Support Vectors		85
Optimization		86
Kernels		86
Summary		87
4		
EXPERIMENTS WITH CLASSICAL MODELS		89
Experiments with the Iris Dataset		89
Testing the Classical Models		90
Implementing a Nearest-Centroid Classifier		93
Experiments with the Breast Cancer Dataset		95
Running Two Initial Tests		95
Testing the Effect of Random Splits		98
Adding k-Fold Validation		100
Searching for Hyperparameters		104
Experiments with the MNIST Dataset		110
Testing the Classical Models		110
Analyzing Runtimes		116
Experimenting with PCA Components		117
Scrambling the Dataset		120
Classical Model Summary		121
Nearest Centroid		121
k-Nearest Neighbors		122
Naive Bayes		122
Decision Trees		123
Random Forests		123
Support Vector Machines		124

When to Use Classical Models	124
Handling Small Datasets	125
Dealing with Reduced Computational Requirements	125
Having Explainable Models	125
Working with Vector Inputs	125
Summary	126

PART III NEURAL NETWORKS

5 INTRODUCTION TO NEURAL NETWORKS 129

Anatomy of a Neural Network	130
The Neuron	131
Activation Functions	132
The Architecture of a Network	135
Output Layers	137
Weight and Bias Representation	139
A Simple Neural Network Implementation	140
Building the Dataset	140
Implementing the Neural Network	141
Training and Testing the Neural Network	143
Summary	145

6 TRAINING A NEURAL NETWORK 147

A High-Level Overview	148
Gradient Descent	148
Finding Minimums	150
Updating the Weights	151
Stochastic Gradient Descent	152
Batches and Minibatches	152
Convex vs. Nonconvex Functions	154
When to Stop Training	155
The Learning Rate	156
Momentum	157
Backpropagation	157
A Simple Example	158
An Abstract Example	161
Loss Functions	164
Absolute and Mean Squared Error Loss	165
Cross-Entropy Loss	166

Weight Initialization	167
Managing Model Complexity and Generalization	168
Overfitting	169
Regularization	171
L2 Regularization	171
Dropout	173
Summary	175

7

EXPERIMENTS WITH NEURAL NETWORKS 177

The Dataset	178
The MLPClassifier Class	178
Architecture and Activation Functions	179
The Code	179
The Results	183
Batch Size	186
Base Learning Rate	190
Training-Set Size	192
L2 Regularization	193
Momentum	195
Weight Initialization	197
Feature Ordering	201
Summary	202

8

EVALUATING MODELS 203

Definitions and Assumptions	203
Why Accuracy Is Not Enough	204
The 2×2 Confusion Matrix	206
Metrics Derived from the 2×2 Confusion Matrix	208
Deriving Metrics from the 2×2 Table	208
Using Metrics to Interpret Models	211
More-Advanced Metrics	213
Informedness and Markedness	213
F1 Score	214
Cohen's Kappa	215
Matthews Correlation Coefficient	215
Metric Implementation	215
The Receiver Operating Characteristics Curve	217
Gathering the Models	217
Plotting the Metrics	219
Exploring the ROC Curve	220
Comparing Models with ROC Analysis	222
Generating an ROC Curve	224

Handling Multiple Classes	227
Extending the Confusion Matrix	227
Calculating Weighted Accuracy	230
Considering the Multiclass Matthews Correlation Coefficient	232
Summary	233

PART IV CONVOLUTIONAL NEURAL NETWORKS

9 INTRODUCTION TO CONVOLUTIONAL NEURAL NETWORKS 237

Why Convolutional Neural Networks?	238
Convolution	239
Scanning with the Kernel	239
Using Convolution for Image Processing	242
Anatomy of a Convolutional Neural Network	243
Exploring the Types of Layers	244
Passing Data Through the CNN	246
Convolutional Layers	247
How They Work	247
In Action	250
Multiple Layers	252
Initialization	253
Pooling Layers	254
Fully Connected Layers	255
Fully Convolutional Layers	256
How the CNN Operates	258
Summary	262

10 EXPERIMENTS WITH KERAS AND MNIST 263

Building CNNs in Keras	264
Loading the MNIST Data	264
Building the Model	266
Training and Evaluating the Model	267
Plotting the Error	270
Basic Experiments	272
Architecture Experiments	272
Training-Set Size, Minibatches, and Epochs	275
Optimizers	278

Fully Convolutional Networks	280
Building and Training the Model.....	280
Making the Test Images.....	283
Testing the Model	284
Scrambled MNIST Digits	292
Summary	293

11 EXPERIMENTS WITH CIFAR-10 295

A CIFAR-10 Refresher	296
The Full CIFAR-10 Dataset	296
Building the Models	297
Analyzing the Models	300
Animal or Vehicle?	302
Binary or Multiclass?	307
Using the Augmented CIFAR-10 Dataset	310
Summary	314

12 A CASE STUDY: CLASSIFYING AUDIO SAMPLES 315

Building the Dataset	316
Augmenting the Dataset	317
Preprocessing the Data	321
Classifying the Audio Features	323
With Classical Models.....	323
With a Traditional Neural Network	326
With a Convolutional Neural Network	327
Spectrograms	332
Classifying Spectrograms	335
Ensembles	339
Summary	344

PART V ADVANCED NETWORKS AND GENERATIVE AI

13 ADVANCED CNN ARCHITECTURES 349

The Keras Functional API	350
VGG	354
Standardizing with Batch Normalization	354
Applying Dropout After Convolutional Layers	355
Building the VGG8 Model	356
Testing the VGG8 Model	359

ResNet	363
Mitigating the Vanishing Gradient Problem	364
Exploring ResNet Configurations	366
Building the ResNet-18 Model	367
Testing the ResNet-18 Model	369
MobileNet	371
Implementing Depthwise Convolutions	371
Building the MobileNet Model	373
Testing the MobileNet Model	374
Building an Ensemble	375
Summary	377

14

FINE-TUNING AND TRANSFER LEARNING 379

Fine-Tuning a Pretrained Model	380
VGG16 and MobileNet with CIFAR-10	380
Experiments	385
A Study in Fine-Tuning	388
Creating New Features with Transfer Learning	391
Extracting CIFAR-10 Features with VGG16	391
Training Classical Models with CIFAR-10 Features	396
Detecting Anomalies with CIFAR-10 Features	397
Retrieving Images with CIFAR-10 Features	399
Summary	405

15

FROM CLASSIFICATION TO LOCALIZATION 407

Detection Experiments with MNIST	408
Building the Dataset	409
Building the Model	410
Running the Model	412
Running an Advanced Detection Model	420
Semantic Segmentation with U-Net	422
Implementing the U-Net Model	424
Running the Model and Interpreting Its Output	427
Multilabel Classification	432
Dataset	432
Performance	434
Summary	437

16

SELF-SUPERVISED LEARNING

439

Building the Unlabeled Dataset	440
Rotation Prediction	441
Building and Testing RotNet	441
Exploring Variations on the RotNet Theme	446
Examining RotNet Features	448
Fine-Tuning Experiments	452
Siamese Networks	454
Building and Testing the Siamese Networks	454
Examining Siamese Network Features	459
Fine-Tuning Experiments	461
For Further Exploration	462
Summary	462

17

GENERATIVE ADVERSARIAL NETWORKS

463

How GANs Work	463
Unconditional GANs	464
Building a GAN with Multilayer Perceptrons	465
Building a GAN with Convolutional Layers	470
Experimenting with Unconditional GANs	473
Conditional GANs	480
Implementing a Conditional GAN	480
Experimenting with Conditional GANs	485
Exploring the Latent Space	486
Intrinsic Dimensionality	486
Latent Space Interpolation	488
Summary	490

18

LARGE LANGUAGE MODELS

493

Understanding Large Language Models	494
Evaluating the LLM Block Diagram	496
Tokenizing and Embedding	497
Exploring the Transformer Layer	499
Predicting the Next Token	502
In-Context Learning	504
Configuring the Experiment	504
Running the Experiment	505
Testing Another LLM	508

Running LLMs Locally	509
Representing Parameters with Quantization	510
Installing Ollama	511
Installing Open Source LLMs	512
Chatting with a Simple Chatbot	512
LLM Embeddings	515
Embeddings Encode Meaning	515
Semantic Search and Retrieval-Augmented Generation	518
Sentiment Analysis	523
LLMs and Images	527
Describing Images	527
Classifying Images with an LLM?	529
Are LLMs Creative?	531
The Divergent Association Task	532
DAT Scores as a Function of Temperature	533
Creative Writing as a Function of Temperature	535
Summary	538

AFTERWORD **539**

INDEX **543**