

# CONTENTS IN DETAIL

**FOREWORD BY MICHAEL C. MOZER, PHD.** **xix**

**ACKNOWLEDGEMENTS** **xxiii**

**INTRODUCTION** **xxv**

**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 .....	8
Summary .....	9

**2**  
**USING PYTHON** **11**

The Python Interpreter .....	11
Statements and White Space .....	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 .....	21
The break and continue Statements .....	22
The with Statement .....	23
Handling Errors with try-except Blocks .....	23
Functions .....	24
Modules .....	25
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 Zeros and Ones .....	36
Accessing Elements in an Array .....	38
Indexing into an Array .....	38
Slicing an Array .....	39
The Ellipsis .....	41
Operators and Broadcasting .....	43
Array Input and Output .....	46
Random Numbers .....	48
NumPy and Images .....	49
Summary .....	50

### **4 WORKING WITH DATA 53**

Classes and Labels .....	53
Features and Feature Vectors .....	54
Types of Features .....	55
Feature Selection and the Curse of Dimensionality .....	57
Features of a Good Dataset .....	58
Interpolation and Extrapolation .....	59
The Parent Distribution .....	61
Prior Class Probabilities .....	62
Confusers .....	62
Dataset Size .....	63
Preparing Your Dataset .....	64
Scaling Features .....	64
Missing Features .....	68
Train, Validate, and Test .....	69
The Three Subsets .....	69
Partitioning the Dataset .....	70

K-Fold Cross Validation .....	75
Look at Your Data .....	76
Summary .....	82

**5 BUILDING DATASETS 83**

Irises .....	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 .....	104

**6 CLASSICAL MACHINE LEARNING 107**

Nearest Centroid .....	108
k-Nearest Neighbors .....	112
Naïve Bayes .....	113
Decision Trees and Random Forests .....	117
Recursion Primer .....	119
Building Decision Trees .....	120
Random Forests .....	122
Support Vector Machines .....	123
Margins .....	123
Support Vectors .....	125
Optimization .....	125
Kernels .....	126
Summary .....	127

**7 EXPERIMENTS WITH CLASSICAL MODELS 129**

Experiments with the Iris Dataset .....	130
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 .....	144
MNIST .....	150

Testing the Classical Models .....	150
Analyzing Runtimes .....	156
An Experiment with PCA Components .....	157
Scrambling our Dataset .....	160
Classical Model Summary .....	161
Nearest Centroid .....	161
<i>k</i> -Nearest Neighbors .....	162
Naïve Bayes .....	162
Decision Trees .....	163
Random Forests .....	163
Support Vector Machines .....	164
When to use Classical Models .....	164
Handling Small Datasets .....	164
Dealing with Reduced Computational Requirements .....	165
Having Explainable Models .....	165
Working with Vector Inputs .....	165
Summary .....	166

## **8 INTRODUCTION TO NEURAL NETWORKS 167**

Anatomy of a Neural Network .....	168
The Neuron .....	169
Activation Functions .....	170
Architecture of a Network .....	173
Output Layers .....	175
Representing Weights and Biases .....	177
Implementing a Simple Neural Network .....	178
Building the Dataset .....	179
Implementing the Neural Network .....	180
Training and Testing the Neural Network .....	182
Summary .....	184

## **9 TRAINING A NEURAL NETWORK 185**

A High-Level Overview .....	186
Gradient Descent .....	186
Finding Minimums .....	188
Updating the Weights .....	189
Stochastic Gradient Descent .....	190
Batches and Minibatches .....	190
Convex vs. Nonconvex Functions .....	192
Ending Training .....	193
Updating the Learning Rate .....	194
Momentum .....	195
Backpropagation .....	195

Backprop, Take 1 .....	196
Backprop, Take 2 .....	199
Loss Functions .....	203
Absolute and Mean Squared Error Loss .....	203
Cross-Entropy Loss .....	204
Weight Initialization .....	205
Overfitting and Regularization .....	207
Understanding Overfitting .....	207
Understanding Regularization .....	210
L2 Regularization .....	210
Dropout .....	212
Summary .....	214

## 10

### **EXPERIMENTS WITH NEURAL NETWORKS 217**

Our Dataset .....	218
The MLPClassifier Class .....	218
Architecture and Activation Functions .....	218
The Code .....	219
The Results .....	222
Batch Size .....	226
Base Learning Rate .....	229
Training Set Size .....	232
L2 Regularization .....	233
Momentum .....	236
Weight Initialization .....	237
Feature Ordering .....	241
Summary .....	242

## 11

### **EVALUATING MODELS 245**

Definitions and Assumptions .....	245
Why Accuracy Is Not Enough .....	246
The 2x2 Confusion Matrix .....	248
Metrics Derived from the 2x2 Confusion Matrix .....	250
Deriving Metrics from the 2x2 Table .....	251
Using our Metrics to Interpret Models .....	253
More Advanced Metrics .....	255
Informedness and Markedness .....	255
F1 Score .....	256
Cohen's Kappa .....	257
Matthews Correlation Coefficient .....	257
Implementing our Metrics .....	258
The Receiver Operating Characteristics Curve .....	259
Gathering our Models .....	259

Plotting our Metrics .....	261
Exploring the ROC curve .....	262
Comparing Models with ROC Analysis .....	265
Generating an ROC Curve .....	267
The Precision–Recall Curve .....	269
Handling Multiple Classes .....	269
Extending the Confusion Matrix .....	270
Calculating Weighted Accuracy .....	273
Multiclass Matthews Correlation Coefficient .....	274
Summary .....	276

## 12

### **INTRODUCTION TO CONVOLUTIONAL NEURAL NETWORKS 277**

Why Convolutional Neural Networks? .....	278
Convolution .....	278
Scanning with the Kernel .....	279
Convolution for Image Processing .....	281
Anatomy of a Convolutional Neural Network .....	282
Different Types of Layers .....	283
Passing Data Through the CNN .....	285
Convolutional Layers .....	286
How a Convolution Layer Works .....	286
Using a Convolutional Layer .....	288
Multiple Convolutional Layers .....	291
Initializing a Convolutional Layer .....	292
Pooling Layers .....	292
Fully Connected Layers .....	294
Fully Convolutional Layers .....	295
Step by Step .....	297
Summary .....	301

## 13

### **EXPERIMENTS WITH MNIST 303**

Building CNNs in Keras .....	304
Loading the MNIST Data .....	304
Building our Model .....	306
Training and Evaluating the Model .....	307
Plotting the Error .....	310
Basic Experiments .....	313
Architecture Experiments .....	313
Training Set Size, Minibatches, and Epochs .....	317
Optimizers .....	320
Fully Convolutional Networks .....	321
Building and Training the Model .....	321
Making the Test Images .....	324

Testing the Model .....	325
Scrambled MNIST Digits.....	333
Summary .....	334

**14**  
**EXPERIMENTS WITH CIFAR-10** **337**

A CIFAR-10 Refresher .....	337
Working with the Full CIFAR-10 Dataset .....	338
Building the Models .....	339
Analyzing the Models .....	342
Animal or Vehicle? .....	344
Binary or Multiclass?.....	348
Transfer Learning .....	352
Fine-Tuning a Model .....	358
Building our Datasets .....	359
Adapting our Model for Fine-Tuning .....	361
Testing our Model .....	364
Chapter Summary .....	366

**15**  
**A CASE STUDY** **367**

Building the Dataset .....	368
Augmenting the Dataset .....	369
Preprocessing our Data.....	373
Classifying the Audio Features .....	375
Using Classical Models .....	375
Using a Traditional Neural Network.....	378
Using a Convolutional Neural Network .....	379
Spectrograms.....	383
Classifying Spectrograms .....	387
Initialization, Regularization, and Batch Normalization .....	389
Examining the Confusion Matrix.....	391
Ensembles .....	392
Average, Maximum, and Voting Ensembles .....	393
Summary .....	396

**16**  
**GOING FURTHER** **399**

Going Further with CNNs .....	399
Reinforcement Learning and Unsupervised Learning .....	400
Generative Adversarial Networks .....	401
Recurrent Neural Networks .....	402
Online Resources .....	402
Conferences .....	403

The Book ..... 404  
So Long and Thanks for All the Fish ..... 404

**INDEX**

**405**