

# contents

<i>preface</i>	<i>xiii</i>
<i>acknowledgments</i>	<i>xv</i>
<i>about this book</i>	<i>xvii</i>
<i>about the author</i>	<i>xix</i>
<i>about the cover</i>	<i>xx</i>

## PART 1 YOUR MACHINE-LEARNING RIG.....1

### 1 A machine-learning odyssey 3

1.1	Machine-learning fundamentals	5
	Parameters	7
	Learning and inference	8
1.2	Data representation and features	9
1.3	Distance metrics	15
1.4	Types of learning	17
	Supervised learning	17
	Unsupervised learning	19
	Reinforcement learning	19
1.5	TensorFlow	21
1.6	Overview of future chapters	22
1.7	Summary	24



## 2 *TensorFlow essentials* 25

- 2.1 Ensuring that TensorFlow works 27
- 2.2 Representing tensors 28
- 2.3 Creating operators 32
- 2.4 Executing operators with sessions 34
  - Understanding code as a graph* 35 ■ *Setting session configurations* 36
- 2.5 Writing code in Jupyter 38
- 2.6 Using variables 41
- 2.7 Saving and loading variables 43
- 2.8 Visualizing data using TensorBoard 44
  - Implementing a moving average* 44 ■ *Visualizing the moving average* 46
- 2.9 Summary 49

## PART 2 CORE LEARNING ALGORITHMS .....51

### 3 *Linear regression and beyond* 53

- 3.1 Formal notation 54
  - How do you know the regression algorithm is working?* 57
- 3.2 Linear regression 59
- 3.3 Polynomial model 62
- 3.4 Regularization 65
- 3.5 Application of linear regression 69
- 3.6 Summary 70

### 4 *A gentle introduction to classification* 71

- 4.1 Formal notation 73
- 4.2 Measuring performance 75
  - Accuracy* 75 ■ *Precision and recall* 76 ■ *Receiver operating characteristic curve* 77
- 4.3 Using linear regression for classification 78
- 4.4 Using logistic regression 83
  - Solving one-dimensional logistic regression* 84 ■ *Solving two-dimensional logistic regression* 87



- 4.5 Multiclass classifier 90
  - One-versus-all* 91 ■ *One-versus-one* 92
  - Softmax regression* 92

4.6 Application of classification 96

4.7 Summary 97

## 5 ***Automatically clustering data*** 99

5.1 Traversing files in TensorFlow 100

5.2 Extracting features from audio 102

5.3 K-means clustering 106

5.4 Audio segmentation 109

5.5 Clustering using a self-organizing map 112

5.6 Application of clustering 117

5.7 Summary 117

## 6 ***Hidden Markov models*** 119

6.1 Example of a not-so-interpretable model 121

6.2 Markov model 121

6.3 Hidden Markov model 124

6.4 Forward algorithm 125

6.5 Viterbi decoding 128

6.6 Uses of hidden Markov models 130

*Modeling a video* 130 ■ *Modeling DNA* 130

*Modeling an image* 130

6.7 Application of hidden Markov models 130

6.8 Summary 131

# PART 3 THE NEURAL NETWORK PARADIGM.....133

## 7 ***A peek into autoencoders*** 135

7.1 Neural networks 136

7.2 Autoencoders 140

7.3 Batch training 145

7.4 Working with images 146

7.5 Application of autoencoders 150

7.6 Summary 151



## 8 **Reinforcement learning** 153

- 8.1 Formal notions 155
  - Policy* 156 ■ *Utility* 157
- 8.2 Applying reinforcement learning 158
- 8.3 Implementing reinforcement learning 160
- 8.4 Exploring other applications of reinforcement learning 167
- 8.5 Summary 168

## 9 **Convolutional neural networks** 169

- 9.1 Drawback of neural networks 170
- 9.2 Convolutional neural networks 171
- 9.3 Preparing the image 173
  - Generating filters* 176 ■ *Convolving using filters* 178
  - Max pooling* 181
- 9.4 Implementing a convolutional neural network in TensorFlow 182
  - Measuring performance* 185 ■ *Training the classifier* 186
- 9.5 Tips and tricks to improve performance 187
- 9.6 Application of convolutional neural networks 188
- 9.7 Summary 188

## 10 **Recurrent neural networks** 189

- 10.1 Contextual information 190
- 10.2 Introduction to recurrent neural networks 190
- 10.3 Implementing a recurrent neural network 192
- 10.4 A predictive model for time-series data 195
- 10.5 Application of recurrent neural networks 198
- 10.6 Summary 199

## 11 **Sequence-to-sequence models for chatbots** 201

- 11.1 Building on classification and RNNs 202
- 11.2 Seq2seq architecture 205
- 11.3 Vector representation of symbols 210
- 11.4 Putting it all together 212



- 11.5 Gathering dialogue data 220
- 11.6 Summary 222

## 12 *Utility landscape* 223

- 12.1 Preference model 226
- 12.2 Image embedding 231
- 12.3 Ranking images 234
- 12.4 Summary 239
- 12.5 What's next? 239

*appendix Installation* 241

*index* 247

## *preface*

Like many people of my generation, I've always been addicted to the latest online trends. Around 2005, I remember reading everything FARK, YIMND, and Delicious for entertainment and news. Now, I shift between Medium and Hacker News, which led me to witness TensorFlow's first commit on November 9, 2015. The post appeared at the top of the front page of Hacker News and received hundreds of comments—that energy overshadowed anything I saw on the website.

At that time, machine-learning tools were already fragmented into a zoo of libraries; the ecosystem relied on experimental software packages from academic labs and proprietary solutions from industry giants. When Google revealed TensorFlow, the community's responses were mixed. Despite Google's history of retiring beloved services (such as Google Reader, Google Talk, and Google Wave), the company also had a history of nurturing open source projects (such as Android, Chromium, Go, and Protocol Buffers).

Buts had to be made right then and there about whether to adopt TensorFlow. Although many chose to wait until the library developed, a few dived right in. I sprinted through the official documentation, mastered the basics, and was ready to apply the technology to my doctoral research at UCLA. I accumulated notes diligently, having no idea that the pages I wrote for myself to navigate the TensorFlow documentation would develop into a book.

Around that time, an acquisitions editor at Manning Publications contacted me for a second opinion on a new Haskell book—part of their due diligence procedure, because I'm the author of *Haskell Data Analysis Cookbook* (Packt Publishing, 2014). The