Table of Contents

Preface ix

1. Introduction to Deep Learning 1
Machine Learning Eats Computer Science 1
Deep Learning Primitives 3
Fully Connected Layer 3
Convolutional Layer 4
Recurrent Neural Network Layers 4
Long Short-Term Memory Cells 5
Deep Learning Architectures 6
LeNet 6
AlexNet 6
ResNet 7
Neural Captioning Model 8
Google Neural Machine Translation 9
One-Shot Models 10
AlphaGo 12
Generative Adversarial Networks 13
Neural Turing Machines 14
Deep Learning Frameworks 15
Limitations of TensorFlow 16
Review 17

2. Introduction to TensorFlow Primitives 19
Introducing Tensors 19
Scalars, Vectors, and Matrices 20
Matrix Mathematics 24
Tensors 25

www.allitebooks.com

Tensors in Physics 27
Mathematical Asides 28
Basic Computations in TensorFlow 29
Installing TensorFlow and Getting Started 29
Initializing Constant Tensors 30
Sampling Random Tensors 31
Tensor Addition and Scaling 32
Matrix Operations 33
Tensor Types 35
Tensor Shape Manipulations 35
Introduction to Broadcasting 37
Imperative and Declarative Programming 37
TensorFlow Graphs 39
TensorFlow Sessions 39
TensorFlow Variables 40
Review 42

3. Linear and Logistic Regression with TensorFlow. 43
Mathematical Review 43
Functions and Differentiability 44
Loss Functions 45
Gradient Descent 50
Automatic Differentiation Systems 53
Learning with TensorFlow 55
Creating Toy Datasets 55
New TensorFlow Concepts 60
Training Linear and Logistic Models in TensorFlow 64
Linear Regression in TensorFlow 64
Logistic Regression in TensorFlow 73
Review 79

4. Fully Connected Deep Networks. 81
What Is a Fully Connected Deep Network? 81
"Neurons" in Fully Connected Networks 83
Learning Fully Connected Networks with Backpropagation 85
Universal Convergence Theorem 87
Why Deep Networks? 88
Training Fully Connected Neural Networks 89
Learnable Representations 89
Activations 89
Fully Connected Networks Memorize 90
Regularization 90

Training Fully Connected Networks 94
Implementation in TensorFlow 94
Installing DeepChem 94
Tox21 Dataset 95
Accepting Minibatches of Placeholders 96
Implementing a Hidden Layer 96
Adding Dropout to a Hidden Layer 97
Implementing Minibatching 98
Evaluating Model Accuracy 98
Using TensorBoard to Track Model Convergence 99
Review 101

5. Hyperparameter Optimization. 103
Model Evaluation and Hyperparameter Optimization 104
Metrics, Metrics, Metrics 105
Binary Classification Metrics 106
Multiclass Classification Metrics 108
Regression Metrics 110
Hyperparameter Optimization Algorithms 110
Setting Up a Baseline 111
Graduate Student Descent 113
Grid Search 114
Random Hyperparameter Search 115
Challenge for the Reader 116
Review 117

6. Convolutional Neural Networks. 119
Introduction to Convolutional Architectures 120
Local Receptive Fields 120
Convolutional Kernels 122
Pooling Layers 125
Constructing Convolutional Networks 125
Dilated Convolutions 126
Applications of Convolutional Networks 127
Object Detection and Localization 127
Image Segmentation 128
Graph Convolutions 129
Generating Images with Variational Autoencoders 131
Training a Convolutional Network in TensorFlow 134
The MNIST Dataset 134
Loading MNIST 135
TensorFlow Convolutional Primitives 138

The Convolutional Architecture 140
Evaluating Trained Models 144
Challenge for the Reader 146
Review 146

7. Recurrent Neural Networks. 149
Overview of Recurrent Architectures 150
Recurrent Cells 152
Long Short-Term Memory (LSTM) 152
Gated Recurrent Units (GRU) 154
Applications of Recurrent Models 154
Sampling from Recurrent Networks 154
Seq2seq Models 155
Neural Turing Machines 157
Working with Recurrent Neural Networks in Practice 159
Processing the Penn Treebank Corpus 159
Code for Preprocessing 160
Loading Data into TensorFlow 162
The Basic Recurrent Architecture 164
Challenge for the Reader 166
Review 166

8. Reinforcement Learning. 169
Markov Decision Processes 173
Reinforcement Learning Algorithms 175
Q-Learning 176
Policy Learning 177
Asynchronous Training 179
Limits of Reinforcement Learning 179
Playing Tic-Tac-Toe 181
Object Orientation 181
Abstract Environment 182
Tic-Tac-Toe Environment 182
The Layer Abstraction 185
Defining a Graph of Layers 188
The A3C Algorithm 192
The A3C Loss Function 196
Defining Workers 198
Training the Policy 201
Challenge for the Reader 203
Review 203

9. Training Large Deep Networks. 205
Custom Hardware for Deep Networks 205
CPU Training 206
GPU Training 207
Tensor Processing Units 209
Field Programmable Gate Arrays 211
Neuromorphic Chips 211
Distributed Deep Network Training 212
Data Parallelism 213
Model Parallelism 214
Data Parallel Training with Multiple GPUs on Cifar10 215
Downloading and Loading the DATA 216
Deep Dive on the Architecture 218
Training on Multiple GPUs 220
Challenge for the Reader 223
Review 223

10. The Future of Deep Learning. 225
Deep Learning Outside the Tech Industry 226
Deep Learning in the Pharmaceutical Industry 226
Deep Learning in Law 227
Deep Learning for Robotics 227
Deep Learning in Agriculture 228
Using Deep Learning Ethically 228
Is Artificial General Intelligence Imminent? 230
Where to Go from Here? 231

Index. 233