Table of Contents

Preface xvii

Foreword xxi

Part I. Deep Learning in Practice

1. Your Deep Learning Journey 3

Deep Learning Is for Everyone 3
Neural Networks: A Brief History 5
Who We Are 7
How to Learn Deep Learning 9
Your Projects and Your Mindset 11
The Software: PyTorch, fastai, and Jupyter (And Why It Doesn't Matter) 12
Your First Model 13
Getting a GPU Deep Learning Server 14
Running Your First Notebook 15
What Is Machine Learning? 20
What Is a Neural Network? 23
A Bit of Deep Learning Jargon 24
Limitations Inherent to Machine Learning 25
How Our Image Recognizer Works 26
What Our Image Recognizer Learned 33
Image Recognizers Can Tackle Non-Image Tasks 36
Jargon Recap 40
Deep Learning Is Not Just for Image Classification 41
Validation Sets and Test Sets 48

Use Judgment in Defining Test Sets 50
A *Choose Your Own Adventure* Moment 54
Questionnaire 54
Further Research 56

2. From Model to Production..... 57
The Practice of Deep Learning 57
Starting Your Project 58
The State of Deep Learning 60
The Drivetrain Approach 63
Gathering Data 65
From Data to DataLoaders 70
Data Augmentation 74
Training Your Model, and Using It to Clean Your Data 75
Turning Your Model into an Online Application 78
Using the Model for Inference 78
Creating a Notebook App from the Model 80
Turning Your Notebook into a Real App 82
Deploying Your App 83
How to Avoid Disaster 86
Unforeseen Consequences and Feedback Loops 89
Get Writing! 90
Questionnaire 91
Further Research 92

3. Data Ethics..... 93
Key Examples for Data Ethics 94
Bugs and Recourse: Buggy Algorithm Used for Healthcare Benefits 95
Feedback Loops: YouTube's Recommendation System 95
Bias: Professor Latanya Sweeney "Arrested" 95
Why Does This Matter? 96
Integrating Machine Learning with Product Design 99
Topics in Data Ethics 101
Recourse and Accountability 101
Feedback Loops 102
Bias 105
Disinformation 116
Identifying and Addressing Ethical Issues 118
Analyze a Project You Are Working On 118
Processes to Implement 119
The Power of Diversity 121

Fairness, Accountability, and Transparency 122
Role of Policy 123
The Effectiveness of Regulation 124
Rights and Policy 125
Cars: A Historical Precedent 125
Conclusion 126
Questionnaire 127
Further Research 128
Deep Learning in Practice: That's a Wrap! 128

Part II. Understanding fastai's Applications

4. Under the Hood: Training a Digit Classifier. 133
Pixels: The Foundations of Computer Vision 133
First Try: Pixel Similarity 137
NumPy Arrays and PyTorch Tensors 143
Computing Metrics Using Broadcasting 145
Stochastic Gradient Descent 149
Calculating Gradients 153
Stepping with a Learning Rate 156
An End-to-End SGD Example 157
Summarizing Gradient Descent 162
The MNIST Loss Function 163
Sigmoid 168
SGD and Mini-Batches 170
Putting It All Together 171
Creating an Optimizer 174
Adding a Nonlinearity 176
Going Deeper 180
Jargon Recap 181
Questionnaire 182
Further Research 184

5. Image Classification. 185
From Dogs and Cats to Pet Breeds 186
Presizing 189
Checking and Debugging a DataBlock 191
Cross-Entropy Loss 194
Viewing Activations and Labels 194
Softmax 195

Log Likelihood 198
Taking the log 200
Model Interpretation 203
Improving Our Model 205
The Learning Rate Finder 205
Unfreezing and Transfer Learning 207
Discriminative Learning Rates 210
Selecting the Number of Epochs 212
Deeper Architectures 213
Conclusion 215
Questionnaire 216
Further Research 217

6. Other Computer Vision Problems..... 219
Multi-Label Classification 219
The Data 220
Constructing a DataBlock 222
Binary Cross Entropy 226
Regression 231
Assembling the Data 232
Training a Model 235
Conclusion 237
Questionnaire 238
Further Research 238

7. Training a State-of-the-Art Model..... 239
Imagenette 239
Normalization 241
Progressive Resizing 243
Test Time Augmentation 245
Mixup 246
Label Smoothing 249
Conclusion 251
Questionnaire 252
Further Research 252

8. Collaborative Filtering Deep Dive..... 253
A First Look at the Data 254
Learning the Latent Factors 256
Creating the DataLoaders 257
Collaborative Filtering from Scratch 260

Weight Decay 264
Creating Our Own Embedding Module 265
Interpreting Embeddings and Biases 267
Using fastai.collab 269
Embedding Distance 270
Bootstrapping a Collaborative Filtering Model 270
Deep Learning for Collaborative Filtering 272
Conclusion 274
Questionnaire 274
Further Research 276

9. Tabular Modeling Deep Dive 277
Categorical Embeddings 277
Beyond Deep Learning 282
The Dataset 284
Kaggle Competitions 284
Look at the Data 285
Decision Trees 287
Handling Dates 289
Using TabularPandas and TabularProc 290
Creating the Decision Tree 292
Categorical Variables 297
Random Forests 298
Creating a Random Forest 299
Out-of-Bag Error 301
Model Interpretation 302
Tree Variance for Prediction Confidence 302
Feature Importance 303
Removing Low-Importance Variables 305
Removing Redundant Features 306
Partial Dependence 308
Data Leakage 311
Tree Interpreter 312
Extrapolation and Neural Networks 314
The Extrapolation Problem 315
Finding Out-of-Domain Data 316
Using a Neural Network 318
Ensembling 322
Boosting 323
Combining Embeddings with Other Methods 324
Conclusion 325

Questionnaire 326
Further Research 327

10. NLP Deep Dive: RNNs. 329
Text Preprocessing 331
Tokenization 332
Word Tokenization with fastai 333
Subword Tokenization 336
Numericalization with fastai 338
Putting Our Texts into Batches for a Language Model 339
Training a Text Classifier 342
Language Model Using DataBlock 342
Fine-Tuning the Language Model 343
Saving and Loading Models 345
Text Generation 346
Creating the Classifier DataLoaders 346
Fine-Tuning the Classifier 349
Disinformation and Language Models 350
Conclusion 352
Questionnaire 353
Further Research 354

11. Data Munging with fastai's Mid-Level API. 355
Going Deeper into fastai's Layered API 355
Transforms 356
Writing Your Own Transform 358
Pipeline 359
TfmdLists and Datasets: Transformed Collections 359
TfmdLists 360
Datasets 362
Applying the Mid-Level Data API: SiamesePair 364
Conclusion 368
Questionnaire 368
Further Research 369
Understanding fastai's Applications: Wrap Up 369

Part III. Foundations of Deep Learning

12. A Language Model from Scratch. 373
The Data 373

Our First Language Model from Scratch 375
Our Language Model in PyTorch 376
Our First Recurrent Neural Network 379
Improving the RNN 381
Maintaining the State of an RNN 381
Creating More Signal 384
Multilayer RNNs 386
The Model 388
Exploding or Disappearing Activations 389
LSTM 390
Building an LSTM from Scratch 390
Training a Language Model Using LSTMs 393
Regularizing an LSTM 394
Dropout 395
Activation Regularization and Temporal Activation Regularization 397
Training a Weight-Tied Regularized LSTM 398
Conclusion 399
Questionnaire 400
Further Research 402

13. Convolutional Neural Networks. 403
The Magic of Convolutions 403
Mapping a Convolutional Kernel 407
Convolutions in PyTorch 408
Strides and Padding 411
Understanding the Convolution Equations 412
Our First Convolutional Neural Network 414
Creating the CNN 415
Understanding Convolution Arithmetic 418
Receptive Fields 419
A Note About Twitter 421
Color Images 423
Improving Training Stability 426
A Simple Baseline 427
Increase Batch Size 429
1cycle Training 430
Batch Normalization 435
Conclusion 438
Questionnaire 439
Further Research 440

14. ResNets **441**
Going Back to Imagenette 441
Building a Modern CNN: ResNet 445
Skip Connections 445
A State-of-the-Art ResNet 451
Bottleneck Layers 454
Conclusion 456
Questionnaire 456
Further Research 457

15. Application Architectures Deep Dive **459**
Computer Vision 459
cnn_learner 459
unet_learner 461
A Siamese Network 463
Natural Language Processing 465
Tabular 466
Conclusion 467
Questionnaire 469
Further Research 470

16. The Training Process **471**
Establishing a Baseline 471
A Generic Optimizer 473
Momentum 474
RMSProp 477
Adam 479
Decoupled Weight Decay 480
Callbacks 480
Creating a Callback 483
Callback Ordering and Exceptions 487
Conclusion 488
Questionnaire 489
Further Research 490
Foundations of Deep Learning: Wrap Up 490

Part IV. Deep Learning from Scratch

17. A Neural Net from the Foundations **493**
Building a Neural Net Layer from Scratch 493

Modeling a Neuron 493
Matrix Multiplication from Scratch 495
Elementwise Arithmetic 496
Broadcasting 497
Einstein Summation 502
The Forward and Backward Passes 503
Defining and Initializing a Layer 503
Gradients and the Backward Pass 508
Refactoring the Model 511
Going to PyTorch 512
Conclusion 515
Questionnaire 515
Further Research 517

18. CNN Interpretation with CAM 519
CAM and Hooks 519
Gradient CAM 522
Conclusion 525
Questionnaire 525
Further Research 525

19. A fastai Learner from Scratch 527
Data 527
Dataset 529
Module and Parameter 531
Simple CNN 534
Loss 536
Learner 537
Callbacks 539
Scheduling the Learning Rate 540
Conclusion 542
Questionnaire 542
Further Research 544

20. Concluding Thoughts 545

A. Creating a Blog 549

B. Data Project Checklist 559

Index 567