

# Contents

<b>Preface</b>	<b>xi</b>
<b>1 Feed-Forward Neural Nets</b>	<b>1</b>
1.1 Perceptrons . . . . .	3
1.2 Cross-entropy Loss Functions for Neural Nets . . . . .	9
1.3 Derivatives and Stochastic Gradient Descent . . . . .	14
1.4 Writing Our Program . . . . .	18
1.5 Matrix Representation of Neural Nets . . . . .	21
1.6 Data Independence . . . . .	24
1.7 References and Further Readings . . . . .	25
1.8 Written Exercises . . . . .	26
<b>2 Tensorflow</b>	<b>29</b>
2.1 Tensorflow Preliminaries . . . . .	29
2.2 A TF Program . . . . .	33
2.3 Multilayered NNs . . . . .	38
2.4 Other Pieces . . . . .	42
2.4.1 Checkpointing . . . . .	42
2.4.2 <code>tf.nn.tensor_dot</code> . . . . .	43
2.4.3 Initialization of TF Variables . . . . .	44
2.4.4 Simplifying TF Graph Creation . . . . .	47
2.5 References and Further Readings . . . . .	48
2.6 Written Exercises . . . . .	49
<b>3 Convolutional Neural Networks</b>	<b>51</b>
3.1 Filters, Strides, and Padding . . . . .	52
3.2 A Simple TF Convolution Example . . . . .	57
3.3 Multilevel Convolution . . . . .	61
3.4 Convolution Details . . . . .	64

3.4.1	Biases . . . . .	64
3.4.2	Layers with Convolution . . . . .	65
3.4.3	Pooling . . . . .	66
3.5	References and Further Readings . . . . .	67
3.6	Written Exercises . . . . .	68
<b>4</b>	<b>Word Embeddings and Recurrent NNs</b>	<b>71</b>
4.1	Word Embeddings for Language Models . . . . .	71
4.2	Building Feed-Forward Language Models . . . . .	76
4.3	Improving Feed-Forward Language Models . . . . .	78
4.4	Overfitting . . . . .	79
4.5	Recurrent Networks . . . . .	82
4.6	Long Short-Term Memory . . . . .	88
4.7	References and Further Readings . . . . .	92
4.8	Written Exercises . . . . .	92
<b>5</b>	<b>Sequence-to-Sequence Learning</b>	<b>95</b>
5.1	The Seq2Seq Paradigm . . . . .	96
5.2	Writing a Seq2Seq MT program . . . . .	99
5.3	Attention in Seq2seq . . . . .	102
5.4	Multilength Seq2Seq . . . . .	107
5.5	Programming Exercise . . . . .	108
5.6	Written Exercises . . . . .	110
5.7	References and Further Readings . . . . .	111
<b>6</b>	<b>Deep Reinforcement Learning</b>	<b>113</b>
6.1	Value Iteration . . . . .	114
6.2	Q-learning . . . . .	117
6.3	Basic Deep-Q Learning . . . . .	119
6.4	Policy Gradient Methods . . . . .	124
6.5	Actor-Critic Methods . . . . .	130
6.6	Experience Replay . . . . .	133
6.7	References and Further Readings . . . . .	134
6.8	Written Exercises . . . . .	134
<b>7</b>	<b>Unsupervised Neural-Network Models</b>	<b>137</b>
7.1	Basic Autoencoding . . . . .	137
7.2	Convolutional Autoencoding . . . . .	140
7.3	Variational Autoencoding . . . . .	144
7.4	Generative Adversarial Networks . . . . .	152

7.5	References and Further Readings . . . . .	157
7.6	Written Exercises . . . . .	157
<b>A</b>	<b>Answers to Selected Exercises</b>	<b>159</b>
A.1	Chapter 1 . . . . .	159
A.2	Chapter 2 . . . . .	160
A.3	Chapter 3 . . . . .	160
A.4	Chapter 4 . . . . .	161
A.5	Chapter 5 . . . . .	161
A.6	Chapter 6 . . . . .	162
A.7	Chapter 7 . . . . .	162
	<b>Bibliography</b>	<b>165</b>
	<b>Index</b>	<b>169</b>