

9.10	The Neuroscientific Basis for Convolutional Networks . . . . .	364
9.11	Convolutional Networks and the History of Deep Learning . . . . .	371
<b>10</b>	<b>Sequence Modeling: Recurrent and Recursive Nets</b>	<b>373</b>
10.1	Unfolding Computational Graphs . . . . .	375
10.2	Recurrent Neural Networks . . . . .	378
10.3	Bidirectional RNNs . . . . .	394
10.4	Encoder-Decoder Sequence-to-Sequence Architectures . . . . .	396
10.5	Deep Recurrent Networks . . . . .	398
10.6	Recursive Neural Networks . . . . .	400
10.7	The Challenge of Long-Term Dependencies . . . . .	401
10.8	Echo State Networks . . . . .	404
10.9	Leaky Units and Other Strategies for Multiple Time Scales . . . . .	406
10.10	The Long Short-Term Memory and Other Gated RNNs . . . . .	408
10.11	Optimization for Long-Term Dependencies . . . . .	413
10.12	Explicit Memory . . . . .	416
<b>11</b>	<b>Practical Methodology</b>	<b>421</b>
11.1	Performance Metrics . . . . .	422
11.2	Default Baseline Models . . . . .	425
11.3	Determining Whether to Gather More Data . . . . .	426
11.4	Selecting Hyperparameters . . . . .	427
11.5	Debugging Strategies . . . . .	436
11.6	Example: Multi-Digit Number Recognition . . . . .	440
<b>12</b>	<b>Applications</b>	<b>443</b>
12.1	Large-Scale Deep Learning . . . . .	443
12.2	Computer Vision . . . . .	452
12.3	Speech Recognition . . . . .	458
12.4	Natural Language Processing . . . . .	461
12.5	Other Applications . . . . .	478
<b>III</b>	<b>Deep Learning Research</b>	<b>486</b>
<b>13</b>	<b>Linear Factor Models</b>	<b>489</b>
13.1	Probabilistic PCA and Factor Analysis . . . . .	490
13.2	Independent Component Analysis (ICA) . . . . .	491
13.3	Slow Feature Analysis . . . . .	493
13.4	Sparse Coding . . . . .	496