

Contents

	Preface	<i>page xi</i>
1	Introduction	1
	1.1 Carbon-Based and Silicon-Based Computing	1
	1.2 Early Beginnings Until the Late 1940s	4
	1.3 From 1950 to 1980	10
	1.4 From 1980 to Today	12
	1.5 Roadmap	14
	1.6 Exercises	15
2	Basic Concepts	16
	2.1 Synapses	16
	2.2 Units or Neurons	17
	2.3 Activations	17
	2.4 Transfer Functions	18
	2.5 Discrete versus Continuous Time	22
	2.6 Networks and Architectures	23
	2.7 Functional and Cardinal Capacity of Architectures	26
	2.8 The Bayesian Statistical Framework	28
	2.9 Information Theory	31
	2.10 Data and Learning Settings	33
	2.11 Learning Rules	35
	2.12 Computational Complexity Theory	36
	2.13 Exercises	37
3	Shallow Networks and Shallow Learning	41
	3.1 Supervised Shallow Networks and their Design	41
	3.2 Capacity of Shallow Networks	46
	3.3 Shallow Learning	52
	3.4 Extensions of Shallow Learning	56
	3.5 Exercises	58

viii	Contents	
4	Two-Layer Networks and Universal Approximation	63
	4.1 Functional Capacity	63
	4.2 Universal Approximation Properties	65
	4.3 The Capacity of $A(n, m, 1)$ Architectures	68
	4.4 Exercises	69
5	Autoencoders	71
	5.1 A General Autoencoder Framework	72
	5.2 General Autoencoder Properties	73
	5.3 Linear Autoencoders	75
	5.4 Non-Linear Autoencoders: Unrestricted Boolean Case	84
	5.5 Other Autoencoders and Autoencoder Properties	90
	5.6 Exercises	96
6	Deep Networks and Backpropagation	99
	6.1 Why Deep?	99
	6.2 Functional Capacity: Deep Linear Case	101
	6.3 Functional Capacity: Deep Unrestricted Boolean Case	102
	6.4 Cardinal Capacity: Deep Feedforward Architectures	103
	6.5 Other Notions of Capacity	104
	6.6 Learning by Backpropagation	105
	6.7 The Optimality of Backpropagation	109
	6.8 Architecture Design	111
	6.9 Practical Training Issues	114
	6.10 The Bias–Variance Decomposition	118
	6.11 Dropout	119
	6.12 Model Compression/Distillation and Dark Knowledge	124
	6.13 Multiplicative Interactions: Gating and Attention	125
	6.14 Unsupervised Learning and Generative Models	127
	6.15 Exercises	131
7	The Local Learning Principle	137
	7.1 Virtualization and Learning in the Machine	137
	7.2 The Neuronal View	138
	7.3 The Synaptic View: the Local Learning Principle	139
	7.4 Stratification of Learning Rules	141
	7.5 Deep Local Learning and its Fundamental Limitations	142
	7.6 Local Deep Learning: the Deep Learning Channel	144
	7.7 Local Deep Learning and Deep Targets Equivalence	147
	7.8 Exercises	149
8	The Deep Learning Channel	151
	8.1 Random Backpropagation (RBP) and its Variations	152
	8.2 Simulations of Random Backpropagation	154

	8.3 Understanding Random Backpropagation	155
	8.4 Mathematical Analysis of Random Backpropagation	157
	8.5 Further Remarks About Learning Channels	162
	8.6 Circular Autoencoders	164
	8.7 Recirculation: Locality in Both Space and Time	165
	8.8 Simulations of Recirculation	167
	8.9 Recirculation is Random Backpropagation	168
	8.10 Mathematical Analysis of Recirculation	170
	8.11 Exercises	173
9	Recurrent Networks	177
	9.1 Recurrent Networks	177
	9.2 Cardinal Capacity of Recurrent Networks	178
	9.3 Symmetric Connections: The Hopfield Model	179
	9.4 Symmetric Connections: Boltzmann Machines	182
	9.5 Exercises	185
10	Recursive Networks	189
	10.1 Variable-Size Structured Data	189
	10.2 Recursive Networks and Design	190
	10.3 Relationships between Inner and Outer Approaches	199
	10.4 Exercises	201
11	Applications in Physics	204
	11.1 Deep Learning in the Physical Sciences	204
	11.2 Antimatter Physics	208
	11.3 High Energy Collider Physics	214
	11.4 Neutrino Physics	224
	11.5 Dark Matter Physics	228
	11.6 Cosmology and Astrophysics	230
	11.7 Climate Physics	233
	11.8 Incorporating Physics Knowledge and Constraints	235
	11.9 Conclusion: Theoretical Physics	237
12	Applications in Chemistry	239
	12.1 Chemical Data and Chemical Space	240
	12.2 Prediction of Small Molecule Properties	242
	12.3 Prediction of Chemical Reactions	245
13	Applications in Biology and Medicine	257
	13.1 Biomedical Data	257
	13.2 Life in a Nutshell	258
	13.3 Deep Learning in Proteomics	261
	13.4 Deep Learning in Genomics and Transcriptomics	268

x	Contents	
	13.5 Deep Learning in Biomedical Imaging	270
	13.6 Deep Learning in Health Care	273
14	Conclusion	275
	14.1 Explainability and the Black-Box Question	275
	14.2 ANNs versus BNNs	277
Appendix A	Reinforcement Learning and Deep Reinforcement Learning	282
	A.1 Brief History and Background	282
	A.2 Main Algorithmic Approaches	287
	A.3 Limitations and Open Problems	298
	A.4 Other Directions of Research	302
	A.5 Deep Reinforcement Learning	303
	A.6 Exercises	306
Appendix B	Hints and Remarks for Selected Exercises	308
	References	313
	Index	365