

DEEP LEARNING FOR PHYSICS RESEARCH

Martin Erdmann

RWTH Aachen University, Germany

Jonas Glombitza

RWTH Aachen University, Germany

Gregor Kasieczka

University of Hamburg, Germany

Uwe Klemradt

RWTH Aachen University, Germany

 **World Scientific**

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI • TOKYO

Contents

<i>Preface</i>	v
Deep Learning Basics	1
1. Scope of this textbook	3
1.1 Data-driven methods	3
1.2 Physics data analysis	4
1.3 Machine learning methods	5
1.4 Deep learning	6
1.5 Statistical and systematic uncertainties	8
1.6 Causes for rapid advancements	9
1.7 Comprehension and hands-on	9
2. Models for data analysis	13
2.1 Data in physics research	13
2.2 Model building	16
2.3 Data-driven model optimization	18
3. Building blocks of neural networks	23
3.1 Linear mapping and displacement	23
3.2 Nonlinear mapping: activation function	24
3.3 Network prediction	25
3.4 Universal approximation theorem	28
3.5 Exercises	29

4.	Optimization of network parameters	33
4.1	Preprocessing of input data	33
4.2	Epoch and batch	35
4.3	Parameter initialization	36
4.4	Objective function	37
4.5	Gradients from backpropagation	42
4.6	Stochastic gradient descent	43
4.7	Learning rate	45
4.8	Learning strategies	46
4.9	Exercises	50
5.	Mastering model building	53
5.1	Criteria for model building	53
5.2	Data sets for training, validation, test	56
5.3	Monitoring	57
5.4	Regularization	58
5.5	Hyperparameters and activation function	61
5.6	Exercises	64
Standard Architectures of Deep Networks		67
6.	Revisiting the terminology	69
7.	Fully-connected networks: improving the classic all-rounder	71
7.1	N -layer fully-connected networks	71
7.2	Regression challenge: two-dimensional location	72
7.3	Classification challenge: two image categories	75
7.4	Challenges of training deep networks	76
7.5	Residual neural networks	79
7.6	Batch normalization	80
7.7	Self-normalizing neural networks	81
7.8	Exercises	84
8.	Convolutional neural networks and analysis of image-like data	87
8.1	Convolutional neural networks	87
8.2	Comparison to fully-connected networks	97
8.3	Reconstruction tasks	99

8.4	Advanced concepts	100
8.5	Convolutions beyond translational invariance	107
8.6	Physics applications	109
8.7	Exercises	112
9.	Recurrent neural networks: time series and variable input	115
9.1	Sequential relations and network architecture	115
9.2	Recurrent neural networks	116
9.3	Training of recurrent networks	119
9.4	Long short-term memory	119
9.5	Gated recurrent unit	123
9.6	Application areas	125
9.7	Physics applications	129
9.8	Exercises	130
10.	Graph networks and convolutions beyond Euclidean domains	133
10.1	Beyond Cartesian data structures	133
10.2	Graphs	135
10.3	Construction of graphs	138
10.4	Convolutions in the spatial domain	139
10.5	Convolutions in the spectral domain	149
10.6	Exercises	155
11.	Multi-task learning, hybrid architectures, and operational reality	157
11.1	Multi-task learning	157
11.2	Combination of network concepts	158
11.3	Operational reality and verification	159
11.4	Physics applications	159
11.5	Exercises	163
Introspection, Uncertainties, Objectives		165
12.	Interpretability	167
12.1	Interpretability and deep neural networks	167
12.2	Model interpretability and feature visualization	168
12.3	Interpretability of predictions	173
12.4	Exercises	181

13. Uncertainties and robustness	185
13.1 Measuring uncertainties	186
13.2 Decorrelation	189
13.3 Adversarial attacks	191
13.4 Exercises	192
14. Revisiting objective functions	195
14.1 Fitting function parameters to data	196
14.2 Reproducing probability distributions	196
14.3 Conditional probability in supervised learning	201
14.4 Adaptive objective function	204
14.5 Hamiltonian gradient objective function	205
14.6 Energy-based objective function	207
14.7 Learning without gradients	209
14.8 Different types of objectives	210
Deep Learning Advanced Concepts	215
15. Beyond supervised learning	217
16. Weakly-supervised classification	219
16.1 Incomplete supervision	220
16.2 Inexact supervision	221
16.3 Inaccurate supervision	222
16.4 Exercises	224
17. Autoencoders: finding and compressing structures in data	227
17.1 Autoencoder networks	227
17.2 Categories of autoencoders	229
17.3 Application areas	233
17.4 Physics applications	234
17.5 Exercises	236
18. Generative models: data from noise	239
18.1 Variational autoencoder	241
18.2 Generative adversarial networks	243
18.3 Normalizing flows	257

18.4	Physics applications	263
18.5	Exercises	267
19.	Domain adaptation, refinement, unfolding	271
19.1	Refinement for simulations to appear data-like	272
19.2	Scale factors	276
19.3	Common representation	277
19.4	Unfolding	278
20.	Model independent detection of outliers and anomalies	285
20.1	Basics	286
20.2	Low background density	288
20.3	Likelihood-ratio based detection	289
20.4	Exercises	291
21.	Beyond the scope of this textbook	293
21.1	Dedicated hardware and microsystems	293
21.2	Neuroscience inspired developments	294
21.3	Information field theory	295
21.4	Human readable physics concepts	295
Appendix A Notations		297
<i>Bibliography</i>		299
<i>Index</i>		317