

Inhaltsverzeichnis

1. Einführung in Machine Learning und Data Science	7
1.1 Was ist Machine Learning und wofür wird es angewendet?	7
1.2 Was ist Deep Learning?.....	10
1.3 Ansätze im Machine Learning	11
1.4 Vorgehensmodell in der Data Science.....	17
1.5 Data-Science-Programme	22
1.6 Definitionen im Machine Learning	23
1.7 Übungen.....	27
2. Einführung in Python	29
2.1 Installation mittels Anaconda	30
2.2 Basics: Datentypen, Schleifen, Funktionen.....	34
2.3 Mehrdimensionale Arrays mit NumPy	59
2.4 Daten einlesen und Arbeiten mit DataFrames in Pandas.....	70
2.5 Visualisierungen	86
3. Mathematische Grundlagen von Data Science	93
3.1 Lineare Algebra.....	93
3.2 Numerische Methoden.....	114
3.3 Stochastik	118
3.4 Übungen.....	134
4. Datenaufbereitung	138
4.1 Outlier Detection (Ermittlung von Ausreißern)	141
4.2 Variablenauswahl	142
4.3 Dummy Encoding	147
4.4 Standardisierung	149
4.5 Übungen.....	151
5. Data Science mit SciPy	154
5.1 Beschreibende Statistik	154
5.2 Parametrische Tests.....	156
5.3 Nichtparametrische Tests.....	159
5.4 ANOVA	160
5.5 Übungen.....	161
6. Regressionsmodelle	167
6.1 Einführung in die Regression	167
6.2 Einfache Lineare Regression mit Scikit-learn	169
6.3 Polynomiales Regressionsmodell	176
6.4 Multiple Lineare Regression	180
6.5 Annahmen bzw. Bedingungen der Regressionsanalyse	184
6.6 Regression mit Statsmodels	185
6.7 Ridge Regression.....	187
6.8 Bayes'sche Regression	192

6.9	Übungen.....	197
7.	Logistische Regression	200
7.1	Einführung zur logistischen Regression.....	200
7.2	Lineare Entscheidungsgrenzen	208
7.3	Regularisierung	216
7.4	Übungen.....	220
8.	Baumverfahren	224
8.1	Einführung Entscheidungsbäume	224
8.2	Klassifikationsbäume	225
8.3	Regressionsbäume	228
8.4	Aufteilungsregeln.....	230
8.5	Random Forests	231
8.6	Bagging- und Boosting-Baumverfahren.....	233
8.7	Übungen.....	236
9.	Support Vector Machine (SVM)	240
9.1	SVM-Klassifizierung	241
9.2	SVM-Regression (SVR)	245
9.3	Tuning-Parameter	248
9.4	Übungen.....	253
10.	Weitere Machine Learning-Modelle	257
10.1	k-Nearest-Neighbour	257
10.2	Naive Bayes	259
10.3	Übungen.....	263
11.	Unsupervised Learning	269
11.1	Clusteranalyse zur Segmentbildung.....	269
11.2	Hauptkomponentenanalyse zur Dimensionsreduktion	273
11.3	Übungen.....	276
12.	Aufbau eines Data Science-Modells	281
12.1	Über- und Unteranpassung.....	281
12.2	Bewertung eines Modells.....	283
12.3	Feature Wichtigkeit (Feature Importance)	286
12.4	Kreuzvalidierung	288
12.5	Ensemble-Ansätze in Python.....	290
12.6	Grid-Search.....	296
12.7	Übungen.....	299
13.	Deep Learning mit TensorFlow 2 und Keras	306
13.1	Grundlagen des Deep Learnings.....	307
13.2	Was ist ein neuronales Netz?.....	310
13.3	Grundkonzepte neuronaler Netze	310
13.4	Einführung in TensorFlow	312
13.5	Verschiedene Netztypen.....	325
13.6	Lernstrategien	342
13.7	Convolutional neuronale Netzwerke (CNNs).....	392
13.8	Rekurrente neuronale Netzwerke (RNNs).....	408

13.9	Restricted Boltzmann Machine	422
13.10	Generatives Deep Learning	435
13.11	Generative Adversarial Networks (GANs)	456
13.12	Übungen.....	466
14.	End-to-end-Beispiel	475
15.	Zusammenfassung und Ausblick	485
16.	Literaturverzeichnis	497
17.	Glossar	503