

Inhaltsverzeichnis

Vorwort zur deutschen Ausgabe	9
Vorwort	11
Einleitung	13
Wer sollte dieses Buch lesen?	14
Verwendung des Buchs	14
1 Einführung	15
1.1 Was ist Machine Learning?	15
1.2 Arten des Lernens	15
1.2.1 Überwachtes Lernen	15
1.2.2 Unüberwachtes Lernen	16
1.2.3 Teilüberwachtes Lernen	17
1.2.4 Reinforcement Learning	17
1.3 Funktionsweise des überwachten Lernens	18
1.4 Weshalb das Modell mit neuen Daten umgehen kann . . .	22
2 Notation und Definitionen	25
2.1 Notation	25
2.1.1 Datenstrukturen	25
2.1.2 Summenschreibweise	27
2.1.3 Produktschreibweise	27
2.1.4 Mengenoperationen	27
2.1.5 Vektoroperationen	27
2.1.6 Funktionen	29
2.1.7 Max und Arg Max	30
2.1.8 Zuweisungsoperator	30
2.1.9 Ableitung und Gradient	30
2.2 Zufallsvariable	32
2.3 Erwartungstreue Schätzer	34
2.4 Satz von Bayes	34
2.5 Parameterschätzung	35

2.6	Parameter und Hyperparameter	36
2.7	Klassifikation und Regression	36
2.8	Modellbasiertes und instanzbasiertes Lernen	37
2.9	Shallow Learning und Deep Learning	38
3	Grundlegende Algorithmen	39
3.1	Lineare Regression	39
3.1.1	Problemstellung	39
3.1.2	Lösung	41
3.2	Logistische Regression	43
3.2.1	Problemstellung	44
3.2.2	Lösung	45
3.3	Entscheidungsbaum-Lernen	46
3.3.1	Problemstellung	47
3.3.2	Lösung	47
3.4	Support Vector Machine	50
3.4.1	Handhabung von Rauschen	51
3.4.2	Handhabung inhärenter Nichtlinearität	52
3.5	k-Nearest-Neighbors	55
4	Aufbau eines Lernalgorithmus	57
4.1	Bausteine eines Lernalgorithmus	57
4.2	Gradientenabstieg	58
4.3	Wie Machine-Learning-Entwickler vorgehen	64
4.4	Besonderheiten von Lernalgorithmen	64
5	Grundlegende Techniken	67
5.1	Merkmalerstellung	67
5.1.1	One-hot-Codierung	68
5.1.2	Binning	69
5.1.3	Normalisierung	69
5.1.4	Standardisierung	70
5.1.5	Handhabung fehlender Merkmale	71
5.1.6	Datenimputationsverfahren	71
5.2	Auswahl von Lernalgorithmen	73
5.3	Drei Mengen	74
5.4	Unteranpassung und Überanpassung	75
5.5	Regularisierung	78
5.6	Beurteilung der Leistung eines Modells	80
5.6.1	Wahrheitsmatrix	81
5.6.2	Genauigkeit und Trefferquote	82

5.6.3	Korrektklassifikationsrate	83
5.6.4	Kostensensitive Korrektklassifikationsrate	83
5.6.5	Fläche unter der ROC-Kurve	84
5.7	Abstimmung der Hyperparameter	86
5.7.1	Kreuzvalidierung	87
6	Neuronale Netze und Deep Learning	89
6.1	Neuronale Netze	89
6.1.1	Beispiel: mehrschichtiges Perzeptron	90
6.1.2	Neuronale Feedforward-Netzarchitektur	92
6.2	Deep Learning	93
6.2.1	Convolutional Neural Networks (CNNs)	95
6.2.2	Rekurrente neuronale Netze (RNNs)	102
7	Aufgaben und Lösungen	109
7.1	Kernel-Regression	109
7.2	Multi-Class-Klassifikation	110
7.3	One-Class-Klassifikation	112
7.4	Multi-Label-Klassifikation	114
7.5	Ensemble Learning	117
7.5.1	Boosting und Bagging	117
7.5.2	Random Forest	118
7.5.3	Gradient Boosting	119
7.6	Kennzeichnung von Sequenzen erlernen	122
7.7	Sequenz-zu-Sequenz-Lernen	123
7.8	Aktives Lernen	125
7.9	Teilüberwachtes Lernen	128
7.10	One-Shot Learning	131
7.11	Zero-Shot Learning	133
8	Fortgeschrittene Techniken	137
8.1	Handhabung unausgewogener Datenmengen	137
8.2	Modelle kombinieren	139
8.3	Trainieren neuronaler Netze	140
8.4	Erweiterte Regularisierung	141
8.5	Handhabung mehrerer Eingaben	143
8.6	Handhabung mehrerer Ausgaben	144
8.7	Transfer Learning	145
8.8	Effizienz von Algorithmen	146

9	Unüberwachtes Lernen	151
9.1	Dichteschätzung	151
9.2	Clustering	154
9.2.1	k-Means-Clustering	154
9.2.2	DBSCAN und HDBSCAN	156
9.2.3	Anzahl der Cluster festlegen	157
9.2.4	Weitere Clustering-Algorithmen	160
9.3	Dimensionsreduktion	164
9.3.1	Hauptkomponentenanalyse	164
9.3.2	UMAP	166
9.4	Erkennung von Ausreißern	168
10	Weitere Formen des Lernens	169
10.1	Metric Learning	169
10.2	Ranking	171
10.3	Empfehlungen	175
10.3.1	Faktorisierungsmaschinen	176
10.3.2	Denoising Autoencoder	178
10.4	Selbstüberwachtes Lernen: Wort-Embeddings	179
11	Schlussbemerkungen	183
11.1	Was nicht behandelt wurde	183
11.1.1	Topic Modeling	183
11.1.2	Gaußprozesse	184
11.1.3	Generalisierte lineare Modelle	184
11.1.4	Probabilistische grafische Modelle	184
11.1.5	Markow-Ketten-Monte-Carlo-Algorithmen	185
11.1.6	Generative Adversarial Networks	185
11.1.7	Genetische Algorithmen	185
11.1.8	Reinforcement Learning	186
11.2	Danksagungen	187
Index		189