

Vorwort	XVII
----------------------	-------------

Teil I Die Grundlagen des Machine Learning

1 Die Machine-Learning-Umgebung	3
Was ist Machine Learning?	4
Warum wird Machine Learning verwendet?	4
Anwendungsbeispiel	7
Unterschiedliche Machine-Learning-Systeme	9
Überwachtes/unüberwachtes Lernen	9
Batch- und Online-Learning	16
Instanzbasiertes versus modellbasiertes Lernen	18
Die wichtigsten Herausforderungen beim Machine Learning	24
Unzureichende Menge an Trainingsdaten	24
Nicht repräsentative Trainingsdaten	26
Minderwertige Daten	27
Irrelevante Merkmale	28
Overfitting der Trainingsdaten	28
Underfitting der Trainingsdaten	30
Zusammenfassung	31
Testen und Validieren	31
Hyperparameter anpassen und Modellauswahl	32
Datendiskrepanz	33
Übungen	34
2 Ein Machine-Learning-Projekt von A bis Z	37
Der Umgang mit realen Daten	37
Betrachte das Gesamtbild	39
Die Aufgabe abstecken	39

Wähle ein Qualitätsmaß aus	41
Überprüfe die Annahmen	44
Beschaffe die Daten.	44
Erstelle eine Arbeitsumgebung.	44
Die Daten herunterladen	48
Wirf einen kurzen Blick auf die Datenstruktur	49
Erstelle einen Testdatensatz.	53
Erkunde und visualisiere die Daten, um Erkenntnisse zu gewinnen	57
Visualisieren geografischer Daten	58
Suche nach Korrelationen	60
Experimentieren mit Kombinationen von Merkmalen	63
Bereite die Daten für Machine-Learning-Algorithmen vor	64
Aufbereiten der Daten	64
Bearbeiten von Text und kategorischen Merkmalen	67
Eigene Transformer	70
Skalieren von Merkmalen	71
Pipelines zur Transformation.	72
Wähle ein Modell aus und trainiere es	74
Trainieren und Auswerten auf dem Trainingsdatensatz	74
Bessere Auswertung mittels Kreuzvalidierung	76
Optimiere das Modell.	78
Gittersuche.	78
Zufällige Suche.	80
Ensemble-Methoden	81
Analysiere die besten Modelle und ihre Fehler	81
Evaluere das System auf dem Testdatensatz.	82
Nimm das System in Betrieb, überwache und warte es	83
Probieren Sie es aus!	86
Übungen	87
3 Klassifikation	89
MNIST	89
Trainieren eines binären Klassifikators.	91
Qualitätsmaße.	92
Messen der Genauigkeit über Kreuzvalidierung	92
Konfusionsmatrix.	94
Relevanz und Sensitivität	96
Die Wechselbeziehung zwischen Relevanz und Sensitivität	97
Die ROC-Kurve	100
Klassifikatoren mit mehreren Kategorien	103
Fehleranalyse	106
Klassifikation mit mehreren Labels.	109

Klassifikation mit mehreren Ausgaben	110
Übungen	112
4 Trainieren von Modellen	115
Lineare Regression	116
Die Normalengleichung	118
Komplexität der Berechnung	120
Das Gradientenverfahren	121
Batch-Gradientenverfahren	124
Stochastisches Gradientenverfahren	127
Mini-Batch-Gradientenverfahren	130
Polynomielle Regression	131
Lernkurven	133
Regularisierte lineare Modelle	137
Ridge-Regression	137
Lasso-Regression	139
Elastic Net	142
Early Stopping	143
Logistische Regression	144
Abschätzen von Wahrscheinlichkeiten	145
Trainieren und Kostenfunktion	146
Entscheidungsgrenzen	147
Softmax-Regression	149
Übungen	153
5 Support Vector Machines	155
Lineare Klassifikation mit SVMs	155
Soft-Margin-Klassifikation	156
Nichtlineare SVM-Klassifikation	159
Polynomieller Kernel	160
Ähnlichkeitsbasierte Merkmale	161
Der gaußsche RBF-Kernel	162
Komplexität der Berechnung	163
SVM-Regression	164
Hinter den Kulissen	166
Entscheidungsfunktion und Vorhersagen	166
Zielfunktionen beim Trainieren	167
Quadratische Programme	169
Das duale Problem	170
Kernel-SVM	171
Online-SVMs	173
Übungen	175

6	Entscheidungsbäume	177
	Trainieren und Visualisieren eines Entscheidungsbaums	177
	Vorhersagen treffen	178
	Schätzen von Wahrscheinlichkeiten für Kategorien	181
	Der CART-Trainingsalgorithmus	181
	Komplexität der Berechnung	182
	Gini-Unreinheit oder Entropie?	182
	Hyperparameter zur Regularisierung	183
	Regression	185
	Instabilität	187
	Übungen	188
7	Ensemble Learning und Random Forests	191
	Abstimmverfahren unter Klassifikatoren	192
	Bagging und Pasting	195
	Bagging und Pasting in Scikit-Learn	196
	Out-of-Bag-Evaluation	197
	Zufällige Patches und Subräume	198
	Random Forests	199
	Extra-Trees	200
	Wichtigkeit von Merkmalen	200
	Boosting	202
	AdaBoost	202
	Gradient Boosting	205
	Stacking	210
	Übungen	213
8	Dimensionsreduktion	215
	Der Fluch der Dimensionalität	216
	Die wichtigsten Ansätze zur Dimensionsreduktion	217
	Projektion	217
	Manifold Learning	219
	Hauptkomponentenzerlegung (PCA)	221
	Erhalten der Varianz	221
	Hauptkomponenten	222
	Die Projektion auf d Dimensionen	223
	Verwenden von Scikit-Learn	224
	Der Anteil erklärter Varianz	224
	Auswählen der richtigen Anzahl Dimensionen	225
	PCA als Komprimierungsverfahren	226
	Randomisierte PCA	227
	Inkrementelle PCA	227

Kernel-PCA	228
Auswahl eines Kernels und Optimierung der Hyperparameter	229
LLE	231
Weitere Techniken zur Dimensionsreduktion	233
Übungen	234
9 Techniken des unüberwachten Lernens	237
Clustering	238
K-Means	240
Grenzen von K-Means	250
Bildsegmentierung per Clustering	251
Vorverarbeitung per Clustering	253
Clustering für teilüberwachtes Lernen einsetzen	254
DBSCAN	257
Andere Clustering-Algorithmen	260
Gaußsche Mischverteilung	262
Anomalieerkennung mit gaußschen Mischverteilungsmodellen ...	267
Die Anzahl an Clustern auswählen	269
Bayessche gaußsche Mischverteilungsmodelle	272
Andere Algorithmen zur Anomalie- und Novelty-Erkennung	276
Übungen	277

Teil II Neuronale Netze und Deep Learning

10 Einführung in künstliche neuronale Netze mit Keras	281
Von biologischen zu künstlichen Neuronen	282
Biologische Neuronen	283
Logische Berechnungen mit Neuronen	285
Das Perzeptron	286
Mehrschichtiges Perzeptron und Backpropagation	290
Regressions-MLPs	294
Klassifikations-MLPs	295
MLPs mit Keras implementieren	297
TensorFlow 2 installieren	298
Einen Bildklassifikator mit der Sequential API erstellen	299
Ein Regressions-MLP mit der Sequential API erstellen	309
Komplexe Modelle mit der Functional API bauen	310
Dynamische Modelle mit der Subclassing API bauen	315
Ein Modell sichern und wiederherstellen	316
Callbacks	317
TensorBoard zur Visualisierung verwenden	318

Feinabstimmung der Hyperparameter eines neuronalen Netzes	322
Anzahl verborgener Schichten	326
Anzahl Neuronen pro verborgene Schicht	327
Lernrate, Batchgröße und andere Hyperparameter	328
Übungen	330
11 Trainieren von Deep-Learning-Netzen	333
Das Problem schwindender/explodierender Gradienten	334
Initialisierung nach Glorot und He	335
Nicht sättigende Aktivierungsfunktionen	337
Batchnormalisierung	341
Gradient Clipping	347
Wiederverwenden vortrainierter Schichten	348
Transfer Learning mit Keras.	350
Unüberwachtes Vortrainieren	352
Vortrainieren anhand einer Hilfsaufgabe.	353
Schnellere Optimierer	354
Momentum Optimization	354
Beschleunigter Gradient nach Nesterov.	356
AdaGrad.	357
RMSProp	358
Adam-Optimierung	359
Scheduling der Lernrate	362
Vermeiden von Overfitting durch Regularisierung.	367
ℓ_1 - und ℓ_2 -Regularisierung	367
Drop-out	368
Monte-Carlo-(MC-)-Drop-out.	371
Max-Norm-Regularisierung.	374
Zusammenfassung und praktische Tipps	374
Übungen	376
12 Eigene Modelle und Training mit TensorFlow	379
Ein kurzer Überblick über TensorFlow	379
TensorFlow wie NumPy einsetzen	383
Tensoren und Operationen	383
Tensoren und NumPy	385
Typumwandlung	385
Variablen	386
Andere Datenstrukturen	387
Modelle und Trainingsalgorithmen anpassen.	388
Eigene Verlustfunktion	388
Modelle mit eigenen Komponenten sichern und laden	389

Eigene Aktivierungsfunktionen, Initialisierer, Regularisierer und Constraints	391
Eigene Metriken	392
Eigene Schichten	395
Eigene Modelle	398
Verlustfunktionen und Metriken auf Modell-Internen basieren lassen.	400
Gradienten per Autodiff berechnen	402
Eigene Trainingsschleifen	406
Funktionen und Graphen in TensorFlow	409
AutoGraph und Tracing	411
Regeln für TF Functions	412
Übungen	414
13 Daten mit TensorFlow laden und vorverarbeiten	417
Die Data-API	418
Transformationen verketteten	419
Daten durchmischen	420
Daten vorverarbeiten	423
Alles zusammenbringen	424
Prefetching	425
Datasets mit tf.keras verwenden	427
Das TFRecord-Format	428
Komprimierte TFRecord-Dateien	429
Eine kurze Einführung in Protocol Buffer	429
TensorFlow-Protobufs	431
Examples laden und parsen	432
Listen von Listen mit dem SequenceExample-Protobuf verarbeiten	433
Die Eingabemerkmale vorverarbeiten	434
Kategorische Merkmale mit One-Hot-Vektoren codieren	435
Kategorische Merkmale mit Embeddings codieren	437
Vorverarbeitungsschichten von Keras	441
TF Transform	443
Das TensorFlow-Datasets-(TFDS-)Projekt	445
Übungen	446
14 Deep Computer Vision mit Convolutional Neural Networks	449
Der Aufbau des visuellen Cortex	450
Convolutional Layers	451
Filter	453
Stapeln mehrerer Feature Maps	454

Implementierung in TensorFlow	456
Speicherbedarf	459
Pooling Layers	460
Implementierung in TensorFlow	462
Architekturen von CNNs	464
LeNet-5	466
AlexNet	467
GoogLeNet.	470
VGGNet	473
ResNet	474
Xception	477
SENet	479
Ein ResNet-34-CNN mit Keras implementieren.	481
Vortrainierte Modelle aus Keras einsetzen	482
Vortrainierte Modelle für das Transfer Learning	484
Klassifikation und Lokalisierung	487
Objekterkennung	488
Fully Convolutional Networks.	490
You Only Look Once (YOLO).	492
Semantische Segmentierung	495
Übungen	499
15 Verarbeiten von Sequenzen mit RNNs und CNNs	501
Rekurrente Neuronen und Schichten	502
Gedächtniszellen	504
Ein- und Ausgabesequenzen	505
RNNs trainieren	506
Eine Zeitserie vorhersagen	507
Grundlegende Metriken	508
Ein einfaches RNN implementieren.	509
Deep RNNs	510
Mehrere Zeitschritte vorhersagen	512
Arbeit mit langen Sequenzen	515
Gegen instabile Gradienten kämpfen.	516
Das Problem des Kurzzeitgedächtnisses	518
Übungen	527
16 Natürliche Sprachverarbeitung mit RNNs und Attention	529
Shakespearsche Texte mit einem Character-RNN erzeugen.	530
Den Trainingsdatensatz erstellen.	531
Wie ein sequenzieller Datensatz aufgeteilt wird	532
Den sequenziellen Datensatz in mehrere Fenster unterteilen	533

Das Char-RNN-Modell bauen und trainieren	535
Das Char-RNN-Modell verwenden	535
Einen gefälschten Shakespeare-Text erzeugen	536
Zustandsbehaftetes RNN	537
Sentimentanalyse	539
Maskieren	543
Vortrainierte Embeddings wiederverwenden	545
Ein Encoder-Decoder-Netzwerk für die neuronale maschinelle Übersetzung	547
Bidirektionale RNNs	550
Beam Search	551
Attention-Mechanismen	553
Visuelle Attention	556
Attention Is All You Need: Die Transformer-Architektur	558
Aktuelle Entwicklungen bei Sprachmodellen	566
Übungen	568
17 Representation Learning und Generative Learning mit Autoencodern und GANs	571
Effiziente Repräsentation von Daten	572
Hauptkomponentenzerlegung mit einem unvollständigen linearen Autoencoder	574
Stacked Autoencoder	575
Einen Stacked Autoencoder mit Keras implementieren	576
Visualisieren der Rekonstruktionen	577
Den Fashion-MNIST-Datensatz visualisieren	578
Unüberwachtes Vortrainieren mit Stacked Autoencoder	579
Kopplung von Gewichten	580
Trainieren mehrerer Autoencoder nacheinander	582
Convolutional Autoencoder	583
Rekurrente Autoencoder	584
Denoising Autoencoder	584
Sparse Autoencoder	586
Variational Autoencoder	589
Fashion-MNIST-Bilder erzeugen	593
Generative Adversarial Networks	595
Schwierigkeiten beim Trainieren von GANs	599
Deep Convolutional GANs	601
Progressive wachsende GANs	604
StyleGANs	607
Übungen	610

18 Reinforcement Learning	611
Lernen zum Optimieren von Belohnungen.	612
Suche nach Policies.	613
Einführung in OpenAI Gym	615
Neuronale Netze als Policies.	619
Auswerten von Aktionen: Das Credit-Assignment-Problem	621
Policy-Gradienten	622
Markov-Entscheidungsprozesse	627
Temporal Difference Learning	631
Q-Learning	632
Erkundungspolicies	634
Approximatives Q-Learning und Deep-Q-Learning	634
Deep-Q-Learning implementieren	636
Deep-Q-Learning-Varianten	640
Feste Q-Wert-Ziele.	640
Double DQN	641
Priorisiertes Experience Replay	642
Dueling DQN.	643
Die TF-Agents-Bibliothek.	644
TF-Agents installieren	645
TF-Agents-Umgebungen	645
Umgebungsspezifikationen	646
Umgebungswrapper und Atari-Vorverarbeitung.	647
Trainingsarchitektur	650
Deep-Q-Netz erstellen	652
DQN-Agenten erstellen	654
Replay Buffer und Beobachter erstellen.	655
Trainingsmetriken erstellen	657
Collect-Fahrer erstellen	658
Dataset erstellen	659
Trainingsschleife erstellen	662
Überblick über beliebte RL-Algorithmen	664
Übungen	666
19 TensorFlow-Modelle skalierbar trainieren und deployen	667
Ein TensorFlow-Modell ausführen.	668
TensorFlow Serving verwenden.	668
Einen Vorhersageservice auf der GCP AI Platform erstellen	677
Den Vorhersageservice verwenden.	682
Ein Modell auf ein Mobile oder Embedded Device deployen.	685
Mit GPUs die Berechnungen beschleunigen.	689
Sich eine eigene GPU zulegen	690

Eine mit GPU ausgestattete virtuelle Maschine einsetzen	693
Colaboratory	694
Das GPU-RAM verwalten	695
Operationen und Variablen auf Devices verteilen	698
Paralleles Ausführen auf mehreren Devices	700
Modelle auf mehreren Devices trainieren	702
Parallelisierte Modelle	703
Parallelisierte Daten	705
Mit der Distribution Strategies API auf mehreren Devices trainieren	710
Ein Modell in einem TensorFlow-Cluster trainieren	712
Große Trainingsjobs auf der Google Cloud AI Platform ausführen	715
Black Box Hyperparameter Tuning auf der AI Platform	717
Übungen	718
Vielen Dank!	719
A Lösungen zu den Übungsaufgaben	721
B Checkliste für Machine-Learning-Projekte	759
C Das duale Problem bei SVMs	765
D Autodiff	769
E Weitere verbreitete Architekturen neuronaler Netze	777
F Spezielle Datenstrukturen	787
G TensorFlow-Graphen	795
Index	805