

Vorwort	17
----------------------	-----------

Teil I Die Grundlagen des Machine Learning

1 Die Machine-Learning-Umgebung	31
Was ist Machine Learning?	32
Warum wird Machine Learning verwendet?	33
Anwendungsbeispiel	36
Unterschiedliche Machine-Learning-Systeme	37
Trainingsüberwachung	38
Batch- und Online-Learning	46
Instanzbasiertes versus modellbasiertes Lernen	49
Die wichtigsten Herausforderungen beim Machine Learning	55
Unzureichende Menge an Trainingsdaten	55
Nicht repräsentative Trainingsdaten	56
Minderwertige Daten	58
Irrelevante Merkmale	58
Overfitting der Trainingsdaten	59
Underfitting der Trainingsdaten	61
Zusammenfassung	62
Testen und Validieren	62
Hyperparameter anpassen und Modellauswahl	63
Datendiskrepanz	64
Übungen	66
2 Ein Machine-Learning-Projekt von A bis Z	69
Der Umgang mit realen Daten	69
Betrachte das Gesamtbild	71
Die Aufgabe abstecken	71

Wähle ein Qualitätsmaß aus	73
Überprüfe die Annahmen	76
Beschaffe die Daten.	76
Die Codebeispiele mit Google Colab ausführen	76
Ihre Codeänderungen und Daten sichern	79
Interaktivität – mächtig, aber gefährlich	80
Code im Buch versus Notebook-Code.	80
Die Daten herunterladen	81
Wirf einen kurzen Blick auf die Datenstruktur	82
Erstelle einen Testdatensatz.	86
Erkunde und visualisiere die Daten, um Erkenntnisse zu gewinnen . . .	91
Visualisieren geografischer Daten	91
Suche nach Korrelationen	93
Experimentieren mit Kombinationen von Merkmalen	96
Bereite die Daten für Machine-Learning-Algorithmen vor	97
Aufbereiten der Daten	98
Bearbeiten von Text und kategorischen Merkmalen	101
Skalieren und Transformieren von Merkmalen	105
Eigene Transformer	109
Pipelines zur Transformation.	113
Wähle ein Modell aus und trainiere es	118
Trainieren und auswerten auf dem Trainingsdatensatz.	118
Bessere Auswertung mittels Kreuzvalidierung	120
Optimiere das Modell.	122
Gittersuche.	122
Zufällige Suche.	124
Ensemble-Methoden	125
Analysiere die besten Modelle und ihre Fehler	126
Evaluieren das System auf dem Testdatensatz	127
Nimm das System in Betrieb, überwache und warte es	128
Probieren Sie es aus!	131
Übungen	132
3 Klassifikation	135
MNIST	135
Trainieren eines binären Klassifikators.	138
Qualitätsmaße.	139
Messen der Genauigkeit über Kreuzvalidierung	139
Konfusionsmatrix.	140
Relevanz und Sensitivität	142
Die Wechselbeziehung zwischen Relevanz und Sensitivität	143
Die ROC-Kurve	147

Klassifikatoren mit mehreren Kategorien	151
Fehleranalyse	154
Klassifikation mit mehreren Labels	158
Klassifikation mit mehreren Ausgaben	160
Übungen	161
4 Trainieren von Modellen	163
Lineare Regression	164
Die Normalengleichung	166
Komplexität der Berechnung	169
Das Gradientenverfahren	170
Batch-Gradientenverfahren	173
Stochastisches Gradientenverfahren	176
Mini-Batch-Gradientenverfahren	179
Polynomielle Regression	181
Lernkurven	183
Regularisierte lineare Modelle	187
Ridge-Regression	187
Lasso-Regression	189
Elastic-Net-Regression	192
Early Stopping	193
Logistische Regression	194
Abschätzen von Wahrscheinlichkeiten	195
Trainieren und Kostenfunktion	196
Entscheidungsgrenzen	197
Softmax-Regression	201
Übungen	204
5 Support Vector Machines	207
Lineare Klassifikation mit SVMs	207
Soft-Margin-Klassifikation	208
Nichtlineare SVM-Klassifikation	210
Polynomieller Kernel	212
Ähnlichkeitsbasierte Merkmale	213
Der gaußsche RBF-Kernel	213
SVM-Klassen und Komplexität der Berechnung	215
SVM-Regression	216
Hinter den Kulissen linearer SVM-Klassifikatoren	218
Das duale Problem	221
Kernel-SVM	222
Übungen	225

6	Entscheidungsbäume	227
	Trainieren und Visualisieren eines Entscheidungsbaums.	227
	Vorhersagen treffen.	229
	Schätzen von Wahrscheinlichkeiten für Kategorien.	231
	Der CART-Trainingsalgorithmus.	232
	Komplexität der Berechnung.	233
	Gini-Unreinheit oder Entropie?	233
	Hyperparameter zur Regularisierung.	234
	Regression.	236
	Empfindlichkeit für die Ausrichtung der Achsen.	238
	Entscheidungsbäume haben eine größere Varianz.	239
	Übungen.	240
7	Ensemble Learning und Random Forests	243
	Abstimmverfahren unter Klassifikatoren.	244
	Bagging und Pasting.	247
	Bagging und Pasting in Scikit-Learn.	249
	Out-of-Bag-Evaluation.	250
	Zufällige Patches und Subräume.	251
	Random Forests.	252
	Extra-Trees.	253
	Wichtigkeit von Merkmalen.	253
	Boosting.	255
	AdaBoost.	255
	Gradient Boosting.	258
	Histogrammbasiertes Gradient Boosting.	262
	Stacking.	263
	Übungen.	266
8	Dimensionsreduktion	269
	Der Fluch der Dimensionalität.	270
	Die wichtigsten Ansätze zur Dimensionsreduktion.	271
	Projektion.	271
	Manifold Learning.	273
	Hauptkomponentenzerlegung (PCA).	275
	Erhalten der Varianz.	276
	Hauptkomponenten.	276
	Die Projektion auf d Dimensionen.	278
	Verwenden von Scikit-Learn.	278
	Der Anteil erklärter Varianz.	279
	Auswählen der richtigen Anzahl Dimensionen.	279

PCA als Komprimierungsverfahren	281
Randomisierte PCA	282
Inkrementelle PCA	282
Zufallsprojektion	284
LLE	286
Weitere Techniken zur Dimensionsreduktion	289
Übungen	290
9 Techniken des unüberwachten Lernens	293
Clustering-Algorithmen: k -Means und DBSCAN	294
k -Means	297
Grenzen von k -Means	307
Bildsegmentierung per Clustering	308
Clustering für teilüberwachtes Lernen einsetzen	310
DBSCAN	313
Andere Clustering-Algorithmen	316
Gaußsche Mischverteilung	318
Anomalieerkennung mit gaußschen Mischverteilungsmodellen	322
Die Anzahl an Clustern auswählen	324
Bayessche gaußsche Mischverteilungsmodelle	326
Andere Algorithmen zur Anomalie- und Novelty-Erkennung	327
Übungen	329

Teil II Neuronale Netze und Deep Learning

10 Einführung in künstliche neuronale Netze mit Keras	333
Von biologischen zu künstlichen Neuronen	334
Biologische Neuronen	335
Logische Berechnungen mit Neuronen	337
Das Perzeptron	338
Mehrschichtiges Perzeptron und Backpropagation	342
Regressions-MLPs	347
Klassifikations-MLPs	349
MLPs mit Keras implementieren	350
Einen Bildklassifikator mit der Sequential API erstellen	351
Ein Regressions-MLP mit der Sequential API erstellen	362
Komplexe Modelle mit der Functional API bauen	363
Dynamische Modelle mit der Subclassing API bauen	369
Ein Modell sichern und wiederherstellen	371
Callbacks	372
TensorBoard zur Visualisierung verwenden	373

Feinabstimmung der Hyperparameter eines neuronalen Netzes	378
Anzahl verborgener Schichten	383
Anzahl Neuronen pro verborgene Schicht	384
Lernrate, Batchgröße und andere Hyperparameter	385
Übungen	387
11 Trainieren von Deep-Learning-Netzen	391
Das Problem schwindender/explodierender Gradienten	392
Initialisierung nach Glorot und He	393
Bessere Aktivierungsfunktionen	395
Batchnormalisierung	401
Gradient Clipping	407
Wiederverwenden vortrainierter Schichten	408
Transfer Learning mit Keras	410
Unüberwachtes Vortrainieren	412
Vortrainieren anhand einer Hilfsaufgabe	413
Schnellere Optimierer	414
Momentum	414
Beschleunigter Gradient nach Nesterov	416
AdaGrad	417
RMSProp	419
Adam	419
Scheduling der Lernrate	423
Vermeiden von Overfitting durch Regularisierung	428
ℓ_1 - und ℓ_2 -Regularisierung	428
Drop-out	429
Monte-Carlo-(MC-)Drop-out	432
Max-Norm-Regularisierung	435
Zusammenfassung und praktische Tipps	436
Übungen	438
12 Eigene Modelle und Training mit TensorFlow	441
Ein kurzer Überblick über TensorFlow	441
TensorFlow wie NumPy einsetzen	445
Tensoren und Operationen	445
Tensoren und NumPy	447
Typumwandlung	447
Variablen	448
Andere Datenstrukturen	449
Modelle und Trainingsalgorithmen anpassen	450
Eigene Verlustfunktion	450
Modelle mit eigenen Komponenten sichern und laden	451

Eigene Aktivierungsfunktionen, Initialisierer, Regularisierer und Constraints	453
Eigene Metriken	454
Eigene Schichten	457
Eigene Modelle	460
Verlustfunktionen und Metriken auf Modellinterna basieren lassen.	462
Gradienten per Autodiff berechnen	464
Eigene Trainingsschleifen	468
Funktionen und Graphen in TensorFlow	471
AutoGraph und Tracing	474
Regeln für TF Functions	475
Übungen	477
13 Daten mit TensorFlow laden und vorverarbeiten	479
Die tf.data-API.	480
Transformationen verketteten	481
Daten durchmischen	483
Daten vorverarbeiten	486
Alles zusammenbringen	487
Prefetching	488
Datasets mit tf.keras verwenden	490
Das TFRecord-Format	492
Komprimierte TFRecord-Dateien	492
Eine kurze Einführung in Protocol Buffer	493
TensorFlow-Protobufs	494
Examples laden und parsen	496
Listen von Listen mit dem SequenceExample-Protobuf verarbeiten	497
Vorverarbeitungsschichten von Keras	498
Die Normalization-Schicht	499
Die Discretization-Schicht	501
Die CategoryEncoding-Schicht	502
Die StringLookup-Schicht	503
Die Hashing-Schicht	504
Kategorische Merkmale mit Embeddings codieren	505
Vorverarbeitung von Text	509
Vortrainierte Sprachmodellkomponenten verwenden	511
Vorverarbeitungsschichten für Bilder	512
Das TensorFlow-Datasets-Projekt	513
Übungen	515

14	Deep Computer Vision mit Convolutional Neural Networks	519
	Der Aufbau des visuellen Cortex	520
	Convolutional Layers	521
	Filter	523
	Stapeln mehrerer Feature Maps	524
	Convolutional Layer mit Keras implementieren	526
	Speicherbedarf	530
	Pooling Layers	531
	Pooling Layer mit Keras implementieren	533
	Architekturen von CNNs	535
	LeNet-5	538
	AlexNet	539
	GoogLeNet	542
	VGGNet	545
	ResNet	545
	Xception	549
	SENet	550
	Weitere erwähnenswerte Architektur	552
	Die richtige CNN-Architektur wählen	554
	Ein ResNet-34-CNN mit Keras implementieren	555
	Vortrainierte Modelle aus Keras einsetzen	556
	Vortrainierte Modelle für das Transfer Learning	558
	Klassifikation und Lokalisierung	561
	Objekterkennung	563
	Fully Convolutional Networks	565
	You Only Look Once	567
	Objektverfolgung	570
	Semantische Segmentierung	572
	Übungen	575
15	Verarbeiten von Sequenzen mit RNNs und CNNs	577
	Rekurrente Neuronen und Schichten	578
	Gedächtniszellen	580
	Ein- und Ausgabesequenzen	581
	RNNs trainieren	582
	Eine Zeitserie vorhersagen	583
	Die ARMA-Modellfamilie	588
	Die Daten für Machine-Learning-Modelle vorbereiten	591
	Vorhersage mit einem linearen Modell	595
	Vorhersage mit einem einfachen RNN	595
	Vorhersage mit einem Deep RNN	597
	Multivariate Zeitserien vorhersagen	598

Mehrere Zeitschritte vorhersagen	600
Mit einem Sequence-to-Sequence-Modell vorhersagen	602
Arbeit mit langen Sequenzen	605
Gegen instabile Gradienten kämpfen	605
Das Problem des Kurzzeitgedächtnisses	608
Übungen	616
16 Verarbeitung natürlicher Sprache mit RNNs und Attention	619
Shakespearesche Texte mit einem Character-RNN erzeugen	620
Den Trainingsdatensatz erstellen	621
Das Char-RNN-Modell bauen und trainieren	623
Einen gefälschten Shakespeare-Text erzeugen	625
Zustandsbehaftetes RNN	626
Sentimentanalyse	629
Maskieren	632
Vortrainierte Embeddings wiederverwenden	635
Ein Encoder-Decoder-Netzwerk für die neuronale maschinelle Übersetzung	637
Bidirektionale RNNs	643
Beam Search	645
Attention-Mechanismen	647
Attention Is All You Need: die Transformer-Architektur	651
Eine Lawine an Transformer-Modellen	662
Vision Transformers	666
Die Transformers-Bibliothek von Hugging Face	672
Übungen	676
17 Autoencoder, GANs und Diffusionsmodelle	679
Effiziente Repräsentation von Daten	681
Hauptkomponentenzerlegung mit einem unvollständigen linearen Autoencoder	683
Stacked Autoencoder	684
Einen Stacked Autoencoder mit Keras implementieren	685
Visualisieren der Rekonstruktionen	686
Den Fashion-MNIST-Datensatz visualisieren	687
Unüberwachtes Vortrainieren mit Stacked Autoencoder	688
Kopplung von Gewichten	689
Trainieren mehrerer Autoencoder nacheinander	690
Convolutional Autoencoder	691
Denoising Autoencoders	692
Sparse Autoencoders	694
Variational Autoencoders	697
Fashion-MNIST-Bilder erzeugen	701

Generative Adversarial Networks	702
Schwierigkeiten beim Trainieren von GANs	706
Deep Convolutional GANs	708
Progressive wachsende GANs	712
StyleGANs	714
Diffusionsmodelle	717
Übungen	724
18 Reinforcement Learning	727
Lernen zum Optimieren von Belohnungen	728
Suche nach Policies	729
Einführung in OpenAI Gym	731
Neuronale Netze als Policies	735
Auswerten von Aktionen: das Credit-Assignment-Problem	737
Policy-Gradienten	739
Markov-Entscheidungsprozesse	743
Temporal Difference Learning	748
Q-Learning	749
Erkundungspolicies	750
Approximatives Q-Learning und Deep-Q-Learning	751
Deep-Q-Learning implementieren	752
Deep-Q-Learning-Varianten	757
Feste Q-Wert-Ziele	757
Double DQN	758
Priorisiertes Experience Replay	759
Dueling DQN	760
Überblick über beliebte RL-Algorithmen	761
Übungen	764
19 TensorFlow-Modelle skalierbar trainieren und deployen	767
Ein TensorFlow-Modell ausführen	768
TensorFlow Serving verwenden	769
Einen Vorhersageservice auf Vertex AI erstellen	777
Batch-Vorhersagejobs auf Vertex AI ausführen	785
Ein Modell auf ein Mobile oder Embedded Device deployen	787
Ein Modell auf einer Webseite laufen lassen	791
Mit GPUs die Berechnungen beschleunigen	793
Sich eine eigene GPU zulegen	794
Das GPU-RAM verwalten	796
Operationen und Variablen auf Devices verteilen	798
Paralleles Ausführen auf mehreren Devices	800

Modelle auf mehreren Devices trainieren	803
Parallelisierte Modelle	803
Parallelisierte Daten	805
Mit der Distribution Strategies API auf mehreren Devices trainieren	812
Ein Modell in einem TensorFlow-Cluster trainieren	813
Große Trainingsjobs auf Vertex AI ausführen	817
Hyperparameter auf Vertex AI optimieren	819
Hyperparameter mit Keras Tuner auf Vertex AI optimieren	822
Übungen	823
Vielen Dank!	823
A Checkliste für Machine-Learning-Projekte	825
B Autodiff	831
C Spezielle Datenstrukturen	839
D TensorFlow-Graphen	847
Index	857