

---

# Inhalt

<b>Danksagung</b> .....	<b>xv</b>
<b>Wie um alles in der Welt ist so etwas möglich?</b> .....	<b>1</b>
Über dieses Buch .....	2
Bevor wir beginnen .....	3
<b>Teil 1 Von null auf Bilderkennung</b> .....	<b>7</b>
<b>1 Einführung in Machine Learning</b> .....	<b>9</b>
Programmierung und Machine Learning im Vergleich .....	10
Überwachtes Lernen .....	12
Die Mathematik hinter dem Zaubertrick .....	14
Das System einrichten .....	18
<b>2 Ihr erstes ML-Programm</b> .....	<b>21</b>
Die Aufgabenstellung .....	22
Pizzavorhersage mit überwachtem Lernen .....	22
Zusammenhänge in den Daten erkennen .....	23
Eine lineare Regression programmieren .....	25
Das Modell definieren .....	26
Eine Vorhersage treffen .....	28
Das Training implementieren .....	29
Los geht's! .....	33
Bias hinzufügen .....	34
Zusammenfassung .....	37
Praktische Übung: Die Lernrate optimieren .....	39

<b>3 Am Gradienten entlang</b> .....	<b>41</b>
Unser Algorithmus bringt es nicht .....	42
Das Gradientenverfahren .....	43
Ein wenig Mathematik .....	45
Abwärts .....	46
Die dritte Dimension .....	48
Partielle Ableitung .....	49
Die Probe aufs Exempel .....	51
Probleme beim Gradientenverfahren .....	53
Zusammenfassung .....	55
Praktische Übung: Über das Ziel hinaus .....	56
<b>4 Hyperräume</b> .....	<b>57</b>
Noch mehr Dimensionen .....	58
Matrizenrechnung .....	61
Matrizen multiplizieren .....	61
Matrizen transponieren .....	65
Das ML-Programm erweitern .....	65
Die Daten aufbereiten .....	66
Die Vorhersagefunktion anpassen .....	68
Die Verlustfunktion anpassen .....	70
Die Gradientenfunktion anpassen .....	71
Der Code im Ganzen .....	71
Bye-bye, Bias! .....	73
Ein letzter Testlauf .....	74
Zusammenfassung .....	75
Praktische Übung: Statistik in der Praxis .....	75
<b>5 Ein binärer Klassifizierer</b> .....	<b>77</b>
Grenzen der linearen Regression .....	78
Invasion der Sigmoiden .....	81
Konfidenz .....	82
Glätten .....	84
Den Gradienten anpassen .....	86
Was ist mit der Modellfunktion geschehen? .....	87

---

Klassifizierung in Aktion . . . . .	89
Zusammenfassung . . . . .	91
Praktische Übung: Gewichtige Entscheidungen . . . . .	92
<b>6 Eine Aufgabe aus der Praxis . . . . .</b>	<b>93</b>
Die Daten. . . . .	94
MNIST . . . . .	94
Trainings- und Testdatensatz . . . . .	96
Unsere eigene MNIST-Bibliothek . . . . .	97
Die Eingabematrizen vorbereiten . . . . .	97
Die Daten aufbereiten . . . . .	100
Anwendung in der Praxis . . . . .	102
Zusammenfassung . . . . .	104
Praktische Übung: Knifflige Ziffern . . . . .	104
<b>7 Die große Herausforderung . . . . .</b>	<b>105</b>
Von zwei zu mehr Klassen . . . . .	105
1-aus-n-Codierung . . . . .	107
1-aus-n-Codierung in Aktion . . . . .	109
Die Antworten des Klassifizierers decodieren . . . . .	109
Mehr Gewichte . . . . .	110
Die Matrixdimensionen überprüfen . . . . .	111
Der Augenblick der Wahrheit. . . . .	112
Zusammenfassung . . . . .	115
Praktische Übung: Minensucher . . . . .	115
<b>8 Das Perzeptron . . . . .</b>	<b>119</b>
Gestatten, das Perzeptron! . . . . .	119
Perzeptrone kombinieren . . . . .	121
Die Grenzen von Perzeptronen. . . . .	122
Linear separierbare Daten . . . . .	123
Nicht linear separierbare Daten . . . . .	125
Die Geschichte des Perzeptrons . . . . .	126
Der entscheidende Schlag. . . . .	127
Nachwehen . . . . .	128

<b>Teil 2</b>	<b>Neuronale Netze</b>	<b>129</b>
<b>9</b>	<b>Das Netz entwerfen</b>	<b>131</b>
	Ein neuronales Netz aus Perzeptronen zusammenstellen	132
	Perzeptrone verketteten	133
	Wie viele Knoten?	136
	Die Softmax-Funktion	137
	Der Entwurf	139
	Zusammenfassung	140
	Praktische Übung: Auf eigene Faust	141
<b>10</b>	<b>Das Netz erstellen</b>	<b>143</b>
	Die Forward-Propagation programmieren	144
	Die Softmax-Funktion schreiben	145
	Die Klassifizierungsfunktionen schreiben	148
	Kreuzentropie	148
	Zusammenfassung	150
	Praktische Übung: Test durch Zeitreise	151
<b>11</b>	<b>Das Netz trainieren</b>	<b>153</b>
	Wozu Backpropagation?	154
	Von der Kettenregel zur Backpropagation	155
	Die Kettenregel in einem einfachen Netz	155
	Es wird komplizierter	157
	Backpropagation anwenden	160
	Auf Kurs bleiben	161
	Den Gradienten von $w_2$ berechnen	162
	Den Gradienten von $w_1$ berechnen	163
	Die Funktion back() erstellen	165
	Die Gewichte initialisieren	166
	Gefährliche Symmetrie	166
	Tote Neuronen	167
	Korrekte Gewichtsinitialisierung	169
	Das fertige neuronale Netz	170
	Zusammenfassung	173
	Praktische Übung: Fehlstart	173

<b>12 Funktionsweise von Klassifizierern</b> .....	<b>175</b>
Eine Entscheidungsgrenze einzeichnen .....	175
Heimspiel für das Perzeptron .....	177
Klassifizierung verstehen .....	178
Eine Gerade reicht nicht aus .....	181
Die Entscheidungsgrenze krümmen .....	182
Zusammenfassung .....	184
Praktische Übung: Albtraumdaten .....	185
<b>13 Das Mini-Batch-Verfahren</b> .....	<b>187</b>
Der Lernvorgang grafisch dargestellt .....	188
Batch für Batch .....	190
Batches erstellen .....	191
Training mit Batches .....	193
Was geschieht bei verschiedenen Batchgrößen? .....	194
Ein Zickzackpfad .....	195
Große und kleine Batches .....	197
Vor- und Nachteile von Batches .....	198
Zusammenfassung .....	199
Praktische Übung: Das kleinste Batch .....	199
<b>14 Die Kunst des Testens</b> .....	<b>201</b>
Die Gefahr der Überanpassung .....	201
Das Problem mit dem Testdatensatz .....	204
Zusammenfassung .....	206
Praktische Übung: Überlegungen zum Testen .....	207
<b>15 Entwicklung</b> .....	<b>209</b>
Daten aufbereiten .....	210
Den Wertebereich der Eingabevariablen prüfen .....	211
Eingabevariablen standardisieren .....	211
Standardisierung in der Praxis .....	213
Die Hyperparameter anpassen .....	214
Die Anzahl der Epochen festlegen .....	215
Die Anzahl der verdeckten Knoten einstellen .....	215
Die Lernrate einstellen .....	218
Die Batchgröße festlegen .....	221

Der Abschlusstest . . . . .	223
Auf dem Weg zu 99 % . . . . .	225
Praktische Übung: 99 % erreichen . . . . .	225
Zusammenfassung und Vorschau . . . . .	226

---

## **Teil 3    Deep Learning** **229**

---

### **16 Tiefere Netze** . . . . . **231**

Der Echidna-Datensatz . . . . .	232
Neuronale Netze mit Keras erstellen . . . . .	234
Den Aufbau des Netzes planen und der erste Code . . . . .	235
Die Daten laden . . . . .	237
Das Modell erstellen . . . . .	237
Das Modell kompilieren . . . . .	238
Das Netz trainieren . . . . .	239
Die Entscheidungsgrenze einzeichnen . . . . .	240
Das Netz ausführen . . . . .	240
Ein tieferes Netz . . . . .	242
Zusammenfassung . . . . .	243
Praktische Übung: Keras-Spielwiese . . . . .	244

### **17 Überanpassung vermeiden** . . . . . **247**

Was ist Überanpassung? . . . . .	248
Ursachen der Überanpassung . . . . .	248
Unteranpassung . . . . .	252
Das Modell regularisieren . . . . .	254
Eine Untersuchung unseres tiefen Netzes . . . . .	255
L1- und L2-Regularisierung . . . . .	258
Weitere Möglichkeiten zur Regularisierung . . . . .	262
Zusammenfassung . . . . .	264
Praktische Übung: Weitere Regularisierungstechniken . . . . .	265

### **18 Tiefe Netze zähmen** . . . . . **267**

Aktivierungsfunktionen . . . . .	268
Wozu Aktivierungsfunktionen gut sind . . . . .	269
Die sigmoide Aktivierungsfunktion und ihre Auswirkungen . . . . .	270
Verschwindender Gradient . . . . .	272

Alternativen zur Sigmoidfunktion . . . . .	274
Gestatten, die ReLU-Aktivierungsfunktion . . . . .	274
Die richtige Funktion auswählen . . . . .	277
Weitere Techniken . . . . .	279
Bessere Gewichtsinitialisierung . . . . .	279
Gradientenabstieg auf Speed . . . . .	280
Regularisierung für Fortgeschrittene . . . . .	282
Batchnormalisierung . . . . .	284
Zusammenfassung . . . . .	285
Praktische Übung: Die 10-Epochen-Aufgabe . . . . .	286
<b>19 Jenseits von Standardnetzen . . . . .</b>	<b>289</b>
Der CIFAR-10-Datensatz . . . . .	290
Was ist CIFAR-10? . . . . .	291
Das CIFAR-Waterloo . . . . .	292
Die Bausteine von Faltungsnetzen . . . . .	293
Ein Bild ist ein Bild . . . . .	293
Faltung . . . . .	294
Konvolutionsschichten . . . . .	298
Ein Faltungsnetz ausführen . . . . .	299
Zusammenfassung . . . . .	302
Praktische Übung: Massenweise Hyperparameter . . . . .	303
<b>20 Der Weg in die Tiefe . . . . .</b>	<b>305</b>
Der Aufstieg des Deep Learning . . . . .	306
Es braut sich etwas zusammen . . . . .	306
Der Wendepunkt . . . . .	307
Fortsetzung folgt . . . . .	308
Unverhältnismäßige Effektivität . . . . .	309
Was nun? . . . . .	312
Maschinelles Sehen . . . . .	313
Sprache . . . . .	313
Bildgenerierung . . . . .	315
Das Gesamtbild . . . . .	319
Praktischer Einstieg . . . . .	320
Sie sind am Zug . . . . .	321

---

<b>Anhang</b>	<b>323</b>
<b>A Grundlagen von Python</b>	<b>323</b>
Wie sieht Python-Code aus?	326
Dynamische Typisierung	327
Einrückungen	327
Die Bausteine von Python	328
Datentypen und Operatoren	329
Datenstrukturen	330
Strings	331
Schleifen	332
Funktionen definieren und aufrufen	333
Schlüsselwortargumente	334
Standardargumente	335
Module und Pakete	335
Module definieren und importieren	335
Das <code>__main__</code> -Idiom	337
Pakete verwalten	339
Objekte erstellen und verwenden	341
Das war's	342
<b>B Wörterbuch des Machine Learning</b>	<b>345</b>
<b>Stichwortverzeichnis</b>	<b>367</b>