

Inhaltsverzeichnis

1 Was ist intelligent an Künstlicher Intelligenz?	1
1.1 Menschliche Intelligenz hat viele Dimensionen	1
1.2 Woran erkennt man Künstliche Intelligenz	2
1.3 Computer lernen.	3
1.4 Tiefe neuronale Netze können Objekte erkennen	6
1.5 Wie kann man Künstliche Intelligenz verstehen?	8
1.6 Die Geschichte der Künstlichen Intelligenz.	10
1.7 Zusammenfassung	12
Literatur.	13
2 Was kann Künstliche Intelligenz?	15
2.1 Objekterkennung in Bildern.	15
2.1.1 Medizinische Diagnose	17
2.1.2 Vorhersage der 3D-Struktur von Proteinen	18
2.2 Spracherkennung	19
2.3 Maschinelle Übersetzung.	20
2.4 Die Beantwortung natürlichsprachiger Fragen	22
2.5 Dialoge und persönliche Assistenten	25
2.6 Brettspiele	27
2.6.1 Das Strategiespiel Go.	28
2.6.2 Künstliche Intelligenz gewinnt gegen fünf Poker-Professionals	29
2.7 Videospiele.	29
2.7.1 Atari 2600 Spielekonsole.	30
2.7.2 Capture the Flag	30
2.7.3 Das Echtzeit-Strategiespiel Dota2	32
2.8 Selbstfahrende Autos	33
2.8.1 Weiterentwicklung der selbstfahrenden Autos.	34
2.9 Der Computer als kreatives Medium	35
2.9.1 Neue Bilder komponieren	36
2.9.2 Geschichten erfinden	38
2.10 Allgemeine Künstliche Intelligenz	39

2.11 Zusammenfassung	40
Literatur	40
3 Einige Grundbegriffe des maschinellen Lernens	45
3.1 Die wichtigsten Arten des maschinellen Lernens	45
3.1.1 Überwachtes Lernen	45
3.1.2 Unüberwachtes Lernen	46
3.1.3 Bestärkungslernen	47
3.2 Programmieren und Lernen	48
3.2.1 Modelle transformieren eine Eingabe in eine Ausgabe	48
3.2.2 Algorithmen arbeiten schrittweise eine Liste von Anweisungen ab.	50
3.2.3 Ein Lernproblem: Die Erkennung von Ziffern.	50
3.2.4 Vektoren, Matrizen und Tensoren	51
3.3 Lernen eines Zusammenhangs.	53
3.3.1 Schema für das Lernen: Modell, Verlustfunktion und Optimierung	53
3.3.2 Detaillierter Ablauf des Lernens	53
3.4 Ein simples Modell: Die logistische Regression	56
3.4.1 Berechnung eines Punktwertes	56
3.4.2 Die gleichzeitige Berechnung aller Punktwerte.	58
3.4.3 Lineare Transformation	59
3.4.4 Die Softmaxfunktion erzeugt einen Wahrscheinlichkeitsvektor ...	60
3.4.5 Das logistische Regressionsmodell	61
3.5 Die Güte des Modells.	62
3.5.1 Maßstab für die Modellgüte: Die Wahrscheinlichkeit der kompletten Trainingsdaten.	62
3.5.2 Wie misst man den Lernerfolg: Die Verlustfunktion	63
3.5.3 Verdeutlichung für zwei Klassen und zwei Eingabemerkmale ...	65
3.6 Optimierung, oder wie findet man die besten Parameterwerte	66
3.6.1 Der Gradient zeigt in Richtung des steilsten Anstiegs.	67
3.6.2 Der Gradient für mehrere Dimensionen	67
3.6.3 Der Gradient der Verlustfunktion.	69
3.6.4 Schrittweise Minimierung durch Gradientenabstieg	71
3.6.5 Die Lernrate bestimmt die Länge eines Optimierungsschritts	72
3.6.6 Gradientenabstieg mit Minibatch benötigt viel weniger Rechenaufwand	72
3.7 Überprüfung des Lernerfolges.	74
3.7.1 Anwendung des Modells auf neue Daten	74
3.7.2 Überprüfung der Genauigkeit auf der Testmenge	75
3.7.3 Präzision und Recall für Klassen unterschiedlichen Umfangs	76
3.8 Zusammenfassung	77
Literatur	78

- 4 Tiefes Lernen kann komplexe Zusammenhänge erfassen.** 79
 - 4.1 Beim XOR-Problem gibt es Interaktionen zwischen den Merkmalen. 79
 - 4.2 Nichtlinearitäten erzeugen gebogene Trennflächen. 82
 - 4.3 Tiefe neuronale Netze sind Stapel von nichtlinearen Schichten 87
 - 4.3.1 Vektoren und Tensoren repräsentieren die umgeformten Inhalte. 88
 - 4.4 Training von TNN mit Backpropagation-Verfahren 90
 - 4.5 Toolkits erleichtern die Formulierung und das Training von TNN 93
 - 4.5.1 Parallele Berechnungen beschleunigen das Training von TNN. 93
 - 4.5.2 Toolkits erleichtern die Arbeit mit TNN 94
 - 4.6 Wie mache ich das Netz besser? 96
 - 4.6.1 Iterative Konstruktion eines guten Modells mit der Validationsmenge. 96
 - 4.6.2 Unteranpassung und Überanpassung führen zu höheren Fehlern . . . 97
 - 4.6.3 Ein Beispiel für Überanpassung. 98
 - 4.6.4 Regularisierungsverfahren reduzieren den Generalisierungsfehler. 100
 - 4.7 Unterschiedliche Anwendungen erfordern Netze unterschiedlichen Aufbaus 103
 - 4.7.1 Mehrschichtiges Feedforward-Netz. 104
 - 4.7.2 Convolutional Neural Network (CNN) 105
 - 4.7.3 Rekurrentes neuronales Netz (RNN) 105
 - 4.7.4 Bestärkungslernen-Netz. 105
 - 4.7.5 Generatives adversariales Netz (GAN) 106
 - 4.7.6 Autoencoder-Netze erzeugen eine komprimierte Darstellung. 106
 - 4.7.7 Architekturen für bestimmte Medien und Anwendungsbereiche. 106
 - 4.8 Die Konstruktion eines tiefen neuronalen Netzes ist ein Suchprozess 108
 - 4.8.1 Auswahl der Anzahl der Parameter und der Hyperparameter 108
 - 4.8.2 Der Standard-Prozess der Modellsuche führt zu besseren Modellen 109
 - 4.8.3 Automatische Suche von Modellarchitekturen und Hyperparametern 111
 - 4.9 Biologische neuronale Netze funktionieren anders 114
 - 4.10 Zusammenfassung und Trends. 116
 - Literatur. 117
- 5 Bilderkennung mit tiefen neuronalen Netzen.** 119
 - 5.1 Was bedeutet eigentlich Bilderkennung?. 119
 - 5.1.1 Arten der Objekterkennung in Bildern 119
 - 5.1.2 Inspirationen aus der Biologie 120
 - 5.1.3 Warum ist eine Bilderkennung schwierig? 123
 - 5.2 Die Bestandteile eines Convolutional Neural Networks 123
 - 5.2.1 Der Kernel einer Convolution-Schicht analysiert kleine Bildbereiche. 123

5.2.2	Convolution-Schicht mit vielen Kernen reagiert auf viele Merkmale	126
5.2.3	Die Pooling-Schicht wählt die wichtigsten Merkmale aus	127
5.3	Ein einfaches Convolutional Neural Network für die Ziffernerkennung . . .	128
5.4	Der ImageNet Wettbewerb befeuert die Methodenentwicklung	130
5.5	Fortschrittliche Convolutional Neural Networks	132
5.5.1	AlexNet nutzt erfolgreich GPUs zum Training	132
5.5.2	ResNet erleichtert die Optimierung durch Umgehungspfade	133
5.5.3	DenseNet verwendet zusätzliche Umgehungspfade	136
5.5.4	ResNeXt nutzt transformierte Bilder zum Training.	137
5.6	Analyse der CNN Ergebnisse	138
5.6.1	Einzelne Kernel reagieren auf Merkmale unterschiedlicher Art und Größe	138
5.6.2	Ähnlichen Bildern entsprechen benachbarte verdeckte Vektoren . . .	139
5.7	Transferlernen reduziert den Bedarf an Trainingsdaten.	140
5.8	Lokalisierung eines Objektes im Bild	143
5.8.1	Objektlokalisierung durch Rechtecke	143
5.8.2	Bildsegmentierung zur pixelgenaue Bestimmung von Klassen. . .	145
5.8.3	Max-Unpooling belegt ein vergrößertes Feld mit Werten	146
5.8.4	Das U-Net erkennt zunächst Objekte und findet dann die zugehörigen Pixel.	147
5.9	Die 3D-Rekonstruktion einer Szene	149
5.10	Gesichter können mit sehr hoher Genauigkeit zugeordnet werden	150
5.11	Abschätzung der Genauigkeit von Modellprognosen	152
5.11.1	Unsicherheit der Modellprognosen	153
5.11.2	Der Bootstrap erzeugt eine Menge plausibler Modelle	155
5.11.3	Bayessche neuronale Netze	155
5.12	Zuverlässigkeit der Bilderkennung	157
5.12.1	Der Einfluss von Bildverzerrungen	157
5.12.2	Gezielte Konstruktion von falschklassifizierten Bildern	160
5.13	Zusammenfassung und Trends.	163
	Literatur.	164
6	Erfassung der Bedeutung von geschriebenem Text	167
6.1	Wie kann man die Bedeutung von Wörtern durch Vektoren darstellen? . .	170
6.1.1	Das Konzept der Embeddingvektoren	172
6.1.2	Berechnung von Embeddingvektoren mit Word2vec	173
6.1.3	Die Approximation der Softmaxfunktion reduziert den Rechenaufwand	175
6.2	Eigenschaften der Embedding-Vektoren	176
6.2.1	Nächste Nachbarn von Embeddings haben ähnliche Bedeutungen . . .	176

6.2.2	Differenzen zwischen Embeddings lassen sich als Relationen interpretieren	178
6.2.3	FastText nutzt n-Gramme von Buchstaben	180
6.2.4	StarSpace erzeugen Embeddings für andere Objekte	182
6.3	Rekurrente neuronale Netze zur Erfassung von Sequenzen	183
6.3.1	Rekurrente neuronale Netze als Sprachmodelle	184
6.3.2	Training der rekurrenten neuronalen Netze	186
6.3.3	Die Eigenschaften der Gradienten beim RNN.	187
6.4	Das Long-Short Term Memory (LSTM) ist ein Langzeitspeicher	189
6.4.1	Gatter steuern die Speichervorgänge	189
6.4.2	LSTMs mit mehreren Schichten	192
6.4.3	Anwendungen des LSTM	193
6.4.4	Bidirektionale LSTM-Netze zur Prognose von Worteigenschaften	195
6.4.5	Visualisierung von rekurrenten neuronalen Netzen.	197
6.5	Übersetzung: Transformation einer Sequenz in eine andere Sequenz.	198
6.5.1	Sequence-to-Sequence Netze zur Übersetzung	199
6.5.2	Attention: Verbesserung der Übersetzung durch Rückgriff auf die Eingabeworte	203
6.5.3	Übersetzungsergebnisse mit Attention.	205
6.6	Transformer-Übersetzungsmodelle	208
6.6.1	Selbstattention analysiert die Abhängigkeiten der Worte eines Satzes	209
6.6.2	Kreuzattention analysiert die Abhängigkeiten zwischen Ein- und Ausgabe.	210
6.6.3	Transformer-Architektur nutzt Selbst- und Kreuzattention.	212
6.6.4	Training des Transformers für die Sprachübersetzung	214
6.6.5	Byte-Pair Kodierung zur Reduktion des Vokabulars und zur Repräsentation unbekannter Wörter.	216
6.6.6	Ergebnisse für das Transformer-modell.	216
6.6.7	Simultane Übersetzung erfordert Wartezeiten.	219
6.7	BERT: Sprachmodelle für die Repräsentation von Bedeutungen	221
6.7.1	BERT-Architektur	221
6.7.2	BERT-Prognoseaufgaben zum unüberwachten Vortraining.	222
6.8	Transferlernen mit umfangreichen BERT-Modellen der Sprache	224
6.8.1	Semantische Klassifikationsaufgaben	224
6.8.2	Die Beantwortung von Fragen.	226
6.8.3	Extraktion von Weltwissen	228
6.8.4	Transferlernen für Übersetzungsmodelle.	231
6.8.5	Anwendung von BERT in der Web-Suche.	233
6.9	Die Beschreibung von Bildern durch Text.	234
6.10	Die Erklärung der Prognosen von TNN.	237
6.10.1	Erklärungen sind notwendig	237

6.10.2 Globale Erklärungsmodelle	238
6.10.3 Lokale Erklärungsmodelle	239
6.11 Zuverlässigkeit der Texterkennung	241
6.11.1 Robustheit bei Textfehlern und Änderung der Domäne.	241
6.11.2 Anfälligkeit für böswillige Änderung von Eingaben	242
6.12 Zusammenfassung und Trends.	243
Literatur.	245
7 Gesprochene Sprache verstehen	249
7.1 Spracherkennung	249
7.1.1 Warum ist Spracherkennung schwierig?	249
7.1.2 Wie kann man Sprachsignale im Computer darstellen?	250
7.1.3 Messung der Genauigkeit der Spracherkennung	253
7.1.4 Die Geschichte der Spracherkennung	254
7.2 Tiefe Sequence-to-Sequence-Modelle	255
7.2.1 Listen-Attend-Spell erzeugt eine Folge von Buchstaben.	256
7.2.2 Sequence-To-Sequence Modell für Worte und Silben	259
7.3 Convolutional Neural Network zur Spracherkennung	259
7.3.1 CNN Modelle.	260
7.3.2 Kombinierte Modelle	263
7.4 Lippenlesen	264
7.5 Erzeugung von gesprochener Sprache aus Text.	265
7.5.1 WaveNet mit gedehnten Convolution-Schichten für lange Abhängigkeiten	266
7.5.2 Das Tacotron erzeugt ein Spektrogramm.	268
7.6 Dialoge und Sprachassistenten	270
7.7 Gunrock: Ein erweiterter Alexa-Sprachassistent	272
7.7.1 Sprachverstehen	273
7.7.2 Dialogmanagement	274
7.7.3 Antworterzeugung	275
7.7.4 Erprobung des Sprachassistenten.	275
7.8 Analyse der Inhalte von Videos	276
7.8.1 Aufgaben der Videoinhaltsanalyse.	277
7.8.2 Trainingsdaten zur Klassifikation von Videos nach Aktivitäten	277
7.8.3 Convolution-Schichten zur Erkennung von Videoinhalten	278
7.8.4 Genauigkeit der Videoklassifikation	281
7.8.5 Die Erzeugung von Untertiteln für Videos.	282
7.9 Zuverlässigkeit der Verarbeitung gesprochener Sprache	285
7.9.1 Der Effekt von Rauschen und anderen Verzerrungen auf die Spracherkennung	286
7.9.2 Adversariale Attacken auf die automatische Spracherkennung.	286
7.10 Zusammenfassung	287
Literatur.	289

8 Lernen von optimalen Strategien	293
8.1 Einige Grundbegriffe	295
8.2 Tiefes Q-Netz	298
8.2.1 Strategie zur Maximierung der Summe der Belohnungen	298
8.2.2 Eine kleine Navigationsaufgabe	298
8.2.3 Diskontierter zukünftiger Gewinn belohnt schnelle Lösungen	299
8.2.4 Die Q-Funktion bewertet Zustands-Aktionspaare	299
8.2.5 Die Bellmangleichung stellt eine Beziehung zwischen Q-Werten her	300
8.2.6 Approximation der Q-Funktion durch ein tiefes neuronales Netz	301
8.2.7 Q-Lernen: Training eines tiefen Q-Netzes	301
8.3 Anwendung von Q-Lernen auf Atari Videospiele	305
8.3.1 Definition des Spielzustands bei Atari-Spielen	305
8.3.2 Architektur des Atari-Netzes	306
8.3.3 Training	306
8.3.4 Auswertung der tiefen neuronalen Netze der Atari-Videospiele	307
8.4 Strategiegradienten zum Erlernen von stochastischen Strategien	309
8.4.1 Notwendigkeit von Strategien mit Zufallselementen	309
8.4.2 Direkte Optimierung einer Strategie durch Strategiegradienten	310
8.4.3 Erweiterungen des Strategiegradienten: Actor-Critic und Proximal Policy Optimization	312
8.4.4 Anwendung auf Robotik und Go	314
8.4.5 Anwendung in Dota2	316
8.5 Selbstfahrende Autos	317
8.5.1 Sensoren selbstfahrender Autos	318
8.5.2 Funktionalität eines Agenten fürs autonome Fahren	319
8.5.3 Feintuning durch Simulation	322
8.6 Zuverlässigkeit des Bestärkungslernens	325
8.6.1 Training in Simulationsumgebungen oft schwierig übertragbar	325
8.6.2 Adversariale Attacken auf Modelle zum Bestärkungslernen	326
8.7 Zusammenfassung und Trends	327
Literatur	328
9 Kreative Künstliche Intelligenz und Emotionen	331
9.1 Bilder erzeugen mit generativen adversarialen Netzen (GAN)	331
9.1.1 Fälscher und Kunstexperte	331
9.1.2 Generator und Diskriminator	332
9.1.3 Optimierungskriterium für Generator und Diskriminator	332
9.1.4 Die Ergebnisse generativer adversarialer Netze	333
9.1.5 Interpolation zwischen Bildern	337
9.1.6 Transformation von Bildern	338
9.1.7 Transformation von Bildern ohne Trainingspaare	339

9.1.8	Creative Adversarial Network	342
9.1.9	Erzeugung von Bildern aus Text	345
9.1.10	GAN-generierte Modelle von Personen in drei Dimensionen.	346
9.2	Texte verfassen.	347
9.2.1	Automatischer Reporter: Daten in Zeitungsmeldungen darstellen	347
9.2.2	Generierung von längeren Geschichten	347
9.2.3	GPT2 erfindet komplexe Geschichten	348
9.3	Musik automatisch komponieren.	355
9.3.1	MuseNet komponiert Mischungen von Klassik und Pop.	356
9.3.2	Der Music Transformer erfindet Klavierstücke	358
9.4	Emotionen und Persönlichkeit	359
9.4.1	Ein XiaoIce Dialog	359
9.4.2	Das Ziel: Personen zum Weiterreden animieren	361
9.4.3	Architektur von XiaoIce	362
9.4.4	Anzahl der Benutzerantworten als Optimierungskriterium	364
9.4.5	Emotionale Empathie und Unterstützung	366
9.5	Zusammenfassung und Trends.	369
	Literatur.	372
10	KI und ihre Chancen, Herausforderungen und Risiken	375
10.1	Chancen für Wirtschaft und Gesellschaft.	378
10.1.1	Smart Home, mein Haus kümmert sich um mich	378
10.1.2	Diagnose, Therapie, Pflege und Verwaltung in der Medizin	380
10.1.3	Maschine Learning in der industriellen Anwendung.	386
10.1.4	Weitere Einsatzgebiete der KI	389
10.2	Wirtschaftliche Auswirkungen und Zusammenhänge	394
10.2.1	Die Monetarisierung von Daten.	394
10.2.2	Die neue digitale Servicewelt – KI as a service.	398
10.2.3	Große Unternehmen als Treiber der KI	401
10.2.4	Die Auswirkungen auf den Arbeitsmarkt.	404
10.3	Gesellschaftliche Herausforderungen	410
10.3.1	Herausforderungen der KI in der Medizin.	412
10.3.2	1984 2.0: KI als Instrument der Überwachung	413
10.3.3	Krieg der Maschinen	416
10.3.4	Starke Künstliche Intelligenz.	418
10.4	Methodische Herausforderungen.	420
10.4.1	Kombination von Daten und unscharfem Schließen	421
10.4.2	Schnelles und Langsames Denken.	422
10.5	Vertrauen schaffen in die KI	426
10.5.1	Wie baut man vertrauenswürdige KI-Systeme?.	429
10.5.2	Wie kann man tiefe neuronale Netze testen?.	430

10.5.3	Ist eine selbstbestimmte, effektive Nutzung eines KI-Systems möglich?	431
10.5.4	Behandelt das KI-System alle Betroffenen fair?	433
10.5.5	Sind Funktionsweise und Entscheidungen der KI nachvollziehbar?	434
10.5.6	Ist das KI-System sicher gegenüber Angriffen, Unfällen und Fehlern?	435
10.5.7	Funktioniert die KI-Komponente zuverlässig und ist sie robust?	436
10.5.8	Schützt die KI die Privatsphäre und sonstige sensible Informationen?	437
10.5.9	Die Herausforderungen für ein KI-Gütesiegel.	438
10.6	Zusammenfassung	439
	Literatur.	441
A	Anhang	445
A.1	Glossar	447
A.2	Verzeichnis der Abbildungen und deren Quellen	467
A.3	Literatur zum Anhang	483
A.4	Stichwortverzeichnis	489