



Gerhard Paaß · Dirk Hecker

Künstliche Intelligenz

Was steckt hinter der
Technologie der Zukunft?

EBOOK INSIDE



Springer Vieweg

Künstliche Intelligenz

Gerhard Paaß • Dirk Hecker

Künstliche Intelligenz

Was steckt hinter der Technologie der Zukunft?



Springer Vieweg

Gerhard Paaß
Fraunhofer-Institut für Intelligente Analyse-
und Informationssysteme IAIS
Sankt Augustin, Deutschland

Dirk Hecker
Fraunhofer-Institut für Intelligente Analyse-
und Informationssysteme IAIS
Sankt Augustin, Deutschland

ISBN 978-3-658-30210-8 ISBN 978-3-658-30211-5 (eBook)
<https://doi.org/10.1007/978-3-658-30211-5>

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Springer Vieweg

© Springer Fachmedien Wiesbaden GmbH, ein Teil von Springer Nature 2020

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsgesetz zugelassen ist, bedarf der vorherigen Zustimmung des Verlags. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von allgemein beschreibenden Bezeichnungen, Marken, Unternehmensnamen etc. in diesem Werk bedeutet nicht, dass diese frei durch jedermann benutzt werden dürfen. Die Berechtigung zur Benutzung unterliegt, auch ohne gesonderten Hinweis hierzu, den Regeln des Markenrechts. Die Rechte des jeweiligen Zeicheninhabers sind zu beachten.

Der Verlag, die Autoren und die Herausgeber gehen davon aus, dass die Angaben und Informationen in diesem Werk zum Zeitpunkt der Veröffentlichung vollständig und korrekt sind. Weder der Verlag, noch die Autoren oder die Herausgeber übernehmen, ausdrücklich oder implizit, Gewähr für den Inhalt des Werkes, etwaige Fehler oder Äußerungen. Der Verlag bleibt im Hinblick auf geografische Zuordnungen und Gebietsbezeichnungen in veröffentlichten Karten und Institutionsadressen neutral.

Planung: Sybille Thelen Titelbild: sdecoret – stock.adobe.com

Springer Vieweg ist ein Imprint der eingetragenen Gesellschaft Springer Fachmedien Wiesbaden GmbH und ist ein Teil von Springer Nature.

Die Anschrift der Gesellschaft ist: Abraham-Lincoln-Str. 46, 65189 Wiesbaden, Germany

Grußwort von Minister Prof. Dr. Andreas Pinkwart

Liebe Leserinnen, liebe Leser,

getrieben von Neugier und Enthusiasmus für neue Technologien verfolgen Forscher und Entwickler seit vielen Jahrzehnten das Ziel, Menschen und Maschinen miteinander zu verbinden. Jenseits der Faszination, die die künstliche auf die menschliche Intelligenz ausübt, steht heute mit zunehmender Anwendungsbreite nicht mehr Utopie, sondern vielmehr die konkrete Rolle im Mittelpunkt, die Künstliche Intelligenz (KI) auf dem Weg zu einer digitalen Wirtschaft und Gesellschaft einnimmt. KI wird zu einem Treiber des Wandels.

Die KI-Forschung hat inzwischen eine Leistungsstärke erreicht, die selbst bei komplexen Aufgabenstellungen beeindruckende Ergebnisse erzielt. Und auch in unserem Alltag ist KI längst angekommen: Ob wir unser Smartphone per Spracheingabe nach dem Wetter fragen, Texte mit einem Mausklick in andere Sprachen übersetzen lassen oder automatisch über den Betrugsversuch mit unserer Kreditkarte informiert werden – all dies sind bereits etablierte Anwendungen Künstlicher Intelligenz. Aber die Potenziale und technologischen Möglichkeiten sind um ein Vielfaches höher.

Für Unternehmen bietet KI enorme Chancen, die es mit Mut und Kreativität beherzt zu ergreifen gilt. Wo Freude an Innovationen und die unternehmerische Bereitschaft, sich neuen Herausforderungen zu stellen, zusammentreffen, dort können diese Chancen für bessere Produkte, effiziente und sichere Prozesse sowie neue Geschäftsmodelle genutzt werden. Was hier möglich ist, zeigen uns unter anderem die zahlreichen agilen Start-ups, die mit kreativen Ideen die Technologie vorantreiben. Das autonome Fahren wird die Mobilität verändern und Sicherheit und Komfort auf ein neues Level heben. Auch im Gesundheitssektor und der E-Health-Branche verspricht KI konkreten Nutzen durch Verbesserungen in der Diagnostik, der Früherkennung von Krankheiten oder der Entwicklung neuer Medikamente.

Zu den enormen Chancen gehören aber auch elementare Fragen, die durch die Entwicklungen in der Künstlichen Intelligenz aufgeworfen werden: Was bedeutet es, wenn die Grenzen zwischen Mensch und Maschine verschwimmen? Welche Vorkehrungen sind für eine konsequente Bindung von KI-Anwendungen an ethische Standards zu treffen? Wer trägt die Verantwortung für vom Menschen an Maschinen delegierte Entscheidungen?

In Nordrhein-Westfalen wollen wir der Zukunftstechnologie KI den Weg ebnen und haben uns dabei ambitionierte Ziele gesetzt: Unser Bundesland soll zu den zehn stärksten europäischen Standorten für die Entwicklung und Anwendung Künstlicher Intelligenz gehören. Mit der Kompetenzplattform KI.NRW schaffen wir ein Netzwerk für Forschung und Technologietransfer, unterstützen die Unternehmen bei der Umsetzung von Anwendungen und Strategien Künstlicher Intelligenz und nehmen dabei auch ethische Fragen in den Blick, etwa mit der Entwicklung eines „KI-Gütesiegels“ zur Zertifizierung von KI-Lösungen.

Jede der Chancen, die mit Künstlicher Intelligenz verbunden ist, beginnt mit Verständnis und Expertise. Die Technologie zu begreifen, die unser Leben so stark prägen wird, ist von essentieller Bedeutung. Wir brauchen Expertinnen und Experten für die Entwicklung lernender Systeme und wir brauchen mündige Anwenderinnen und Anwender, die ihre Funktionsweisen verstehen, ihren Nutzen und ihre Grenzen abwägen können. Je mehr Wissen wir über die Technologie erwerben, desto besser können wir alle die digitale Zukunft mit Ideenreichtum, Innovationskraft und Verantwortung gestalten.

Das vorliegende Buch zielt genau auf diese Grundvoraussetzung für einen erfolgreichen Umgang mit Künstlicher Intelligenz und verbindet die in diesem Sinne „richtige“ Frage nach dem „Was steckt dahinter?“ mit dem erfreulichen Anspruch, einen verständlichen Einstieg in die Thematik zu schaffen. Bei der Lektüre der Antworten und ihrem Transfer in den praktischen Umgang mit KI wünsche ich viel Freude.

Prof. Dr. Andreas Pinkwart

Minister für Wirtschaft, Innovation, Digitalisierung und Energie des Landes Nordrhein-Westfalen

Vorwort

Über dieses Buch

Künstliche Intelligenz (KI) ist das Schlagwort unserer Zeit. Als zentraler Treiber der Digitalisierung verändert sie Gesellschaft, Wirtschaft und nahezu alle anderen Lebensbereiche in grundlegender Weise. Die Geschwindigkeit dieses Prozesses ist im Vergleich zu früheren gesellschaftlichen und technischen Veränderungen fast beispiellos. Durchbrüche in der Entwicklung sogenannter tiefer künstlicher neuronaler Netze auf hochleistungsfähigen Rechnern haben diese rasante technologische Entwicklung ausgelöst.

Schon Anfang der 50er-Jahre war dem KI-Pionier Alan Turing klar, dass man für viele Fragestellungen Computer nicht händisch bis in jedes Detail programmieren kann. Es müsste eine zügigere Methode geben, Computer zu programmieren: diese Methode ist das maschinelle Lernen. Die Techniken, mit denen Computer aus vorhandenen Daten lernen, ihr Verhalten zu verbessern, sind inzwischen so leistungsfähig, dass sie an vielen Stellen im Alltag und in unserem Berufsleben zum Einsatz kommen.

Der Vielfalt der Anwendungsgebiete sind kaum Grenzen gesetzt. Intelligente Maschinen nehmen die Umwelt wahr, machen Prognosen, geben Empfehlungen und treffen automatisiert Entscheidungen. Sie nehmen uns Routinearbeiten ab und unterstützen uns bei verantwortungsvollen Tätigkeiten. Die Mensch-Maschine-Beziehung ändert sich hin zu einem Partnerschaftsmodell. Intelligente Systeme entlasten uns, erweitern unsere Fähigkeiten und erhöhen unsere Lebensqualität, führen aber auch zu tiefgreifenden Änderungen der Gesellschaft.

So kann KI einen Beitrag zur Bewältigung von großen gesellschaftlichen Herausforderungen wie beispielsweise Mobilität und Gesundheit leisten. Um aber fundiert über Nutzen, Chancen und Risiken von KI diskutieren zu können, muss der Nutzer verstehen, wie intelligente Systeme im Prinzip funktionieren. Dafür muss er die wichtigsten Konzepte der dahinterstehenden Technologie des maschinellen Lernens begreifen. Zudem wird in immer mehr Einsatzgebieten deutlich, dass eine sorgfältige Gestaltung notwendig ist, damit KI im Einklang mit unseren gesellschaftlichen Werten und unserer Vorstellung von Souveränität steht. Hier müssen auch Nicht-Informatiker sachverständig und fundiert mitreden können.

Der Zweck des vorliegenden Buches ist es, die neuen Möglichkeiten des maschinellen Lernens in unterschiedlichen Einsatzgebieten, wie beispielsweise dem autonomen Fahren, der medizinischen Diagnose oder der Analyse der Bedeutung von Sprache anschaulich aufzuzeigen. Dabei werden das Fachvokabular aber auch Konzepte, Methoden und Netzarchitekturen mit vielen Grafiken und Bildern erklärt. Mathematische Zusammenhänge werden, wenn hilfreich, formuliert, und immer nachvollziehbar erläutert. Es stellt sich heraus, dass die verwendeten Verfahren aus sehr einfachen Operationen, wie Addition und Multiplikation, zusammengesetzt sind. Diese gewinnen ihre Leistungsfähigkeit, indem sie auf sehr große Zahlenpakete und mehrfach hintereinander angewendet werden. Möchte man die technischen Kap. 3, 4, 5, 6, 7, 8 und 9 detailliert nachvollziehen, so reicht ein mathematisches Verständnis auf Oberstufenniveau.

Das Buch befähigt Entscheider, aber auch interessierte Laien, bei der Gestaltung von intelligenten Systemen mitreden zu können und Anforderungen in ihren Auswirkungen besser einzuschätzen. Datenanalysten, Studierenden, Ingenieuren, Forschern, die sich neu in das Fachgebiet einarbeiten möchten, bietet das Buch den idealen Einstieg in weiterführende Fachliteratur.

Danksagung

Dieses Buch wurde erst ermöglicht durch die motivierende und fachlich anregende Umgebung des Fraunhofer-Instituts für Intelligente Analyse- und Informationssysteme IAIS in Sankt Augustin. Wir danken allen Kolleginnen und Kollegen sowie den Menschen aus unserem persönlichen Umfeld, die uns bei diesem Buchprojekt unterstützt haben – sei es durch fachliche Diskussionen, Lektorate einzelner Kapitel und hilfreiche Anmerkungen: Robert Babatz, Niklas Beck, Katharina Beckh, Sven Giesselbach, Harald Grund, Monika Hommes-Rüdiger, Birgit Kirsch, Franz Lutter, Dominik Paaß, Julia Paaß und Benjamin Schaarwächter. Mirco Lange und Dr. Henning Petzka haben darüber hinaus das gesamte Buch durchgearbeitet und wertvolle Hinweise für Verbesserungen der Darstellung und der Argumentation geliefert, herzlichen Dank dafür.

Ein ganz besonderer Dank gilt Frau Dr. Angi Voss. Sie fand nicht nur aufmunternde Worte in schwierigen Schreibphasen, sondern hat auch mit ihrer langjährigen Erfahrung als Autorin, ihren antreibenden Fragen und inspirierenden Hinweisen ganz wesentlich zu diesem Buch beigetragen.

Viel Aufwand wurde auch in die mehr als 400 Schaubilder, Grafiken und Fotos investiert. Von den ersten Ideen und Grafiken von Julia Paaß bis hin zur endgültigen und professionellen Gestaltung durch Svenja Niehues, die auch viele inhaltliche Vorschläge für das verständliche Design der Grafiken machte. Vielen Dank für Eure Ausdauer und die langen gemeinsamen Diskussionen am Bildschirm.

Der größte Dank gebührt aber unseren Familien, die uns während der langen Zeit des Schreibens die nötigen Freiräume gegeben haben. Insbesondere möchte ich, Gerhard Paaß, meiner Frau Margret Paaß danken, die mit viel Geduld und Zuspruch einen großen

Anteil am Gelingen dieses Buches hatte und von der Planung bis zur Korrektur der Druckfahnen eine unentbehrliche Hilfe war. Und ich, Dirk Hecker, danke meiner Familie Katrin Berkler und Lara Hecker, die selbst in Urlauben die Arbeit an diesem Buch unterstützt und auch in vielen anderen Situationen so viel Verständnis aufgebracht haben. Ohne Euren Zuspruch und Eure Unterstützung hätten wir dieses Buch nicht zustande gebracht.

Vielen Dank für all Eure Unterstützung!

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

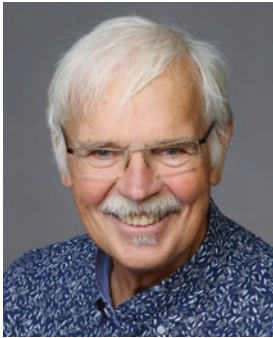
6.10.2 Globale Erklärungsmodelle	238
6.10.3 Lokale Erklärungsmodelle	239
6.11 Zuverlässigkeit der Texterkennung	241
6.11.1 Robustheit bei Textfehlern und Änderung der Domäne	241
6.11.2 Anfälligkeit für böswillige Änderung von Eingaben	242
6.12 Zusammenfassung und Trends	243
Literatur	245
7 Gesprochene 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 Answererzeugung	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

Über die Autoren



Dr. Gerhard Paaß wurde am 10.08.1949 in Langenfeld bei Köln geboren und hat seit 1976 als Mathematiker und Informatiker hochrangige wissenschaftliche Positionen im Themenfeld der Künstlichen Intelligenz bekleidet. Nach dem Abschluss seiner Promotion im Jahre 1985 zum Thema „Prognose und Asymptotik Bayesscher Modelle“ an der Universität Bonn arbeitete Dr. Paaß im Rahmen zahlreicher Forschungsaufenthalte an Universitäten im Ausland (China, USA, Australien, Japan). Darunter waren renommierte Einrichtungen wie die UC Berkeley in Kalifornien und die University of Technology in Brisbane. Als Mitarbeiter der Gesellschaft für Mathematik und Datenverarbeitung (GMD), dem heutigen Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS, war und ist Dr. Paaß bis heute ein gefragter Gutachter und Konferenzleiter auf internationalen Konferenzen, z. B. als Mitglied des Editorial Board der Zeitschrift „International Journal of Uncertainty, Fuzziness and Knowledge-based Systems“ oder als Workshopleiter auf der „Conference on Knowledge Discovery and Data Mining“ 2010 in Washington zum Thema Analyse von Texten für Sicherheitsanwendungen. Dr. Paaß ist Autor zahlreicher Publikationen und hat mehrere Best Paper Awards im Bereich der KI erhalten. Zudem ist er seit vielen Jahren als Dozent tätig und hat im Rahmen der Fraunhofer-Allianz Big Data und Künstliche Intelligenz ganz maßgeblich daran mitgearbeitet, das neue Berufsbild des Data Scientist zu definieren und auch in Deutschland erfolgreich zu etablieren. Als Lead Scientist des Fraunhofer IAIS hat Dr. Paaß an der Entwicklung zahlreicher Curricula in diesem Themenfeld mitgewirkt.



Dr. Dirk Hecker wurde am 29.08.1976 in Köln geboren. Er ist stellvertretender Institutsleiter des Fraunhofer-Instituts für Intelligente Analyse- und Informationssysteme IAIS und Geschäftsführer der Fraunhofer-Allianz Big Data und Künstliche Intelligenz, der größten Allianz der Fraunhofer-Gesellschaft mit mehr als 30 Mitgliedsinstituten. Zudem ist Dr. Hecker Mitglied des Direktoriums der Fraunhofer Academy, die sich als Weiterbildungseinrichtung der Fraunhofer-Gesellschaft an Fach- und Führungskräfte technologiegetriebener Unternehmen richtet. Dr. Hecker hat langjährige Erfahrung in der Leitung von Forschungs- und Industrieprojekten im Bereich Data Mining und Machine Learning. Seine aktuellen Arbeitsschwerpunkte liegen in den Bereichen Big Data Analytics, Predictive Analytics, Deep Learning und Explainable AI. Er studierte Geo-Informatik an den Universitäten Köln und Bonn und promovierte an der Universität zu Köln. Dr. Hecker ist Autor zahlreicher Publikationen zu den oben genannten Fachgebieten und als Fachexperte sowie Auditor zum Thema Künstliche Intelligenz in unterschiedlichsten Gremien aktiv.

Was ist intelligent an Künstlicher Intelligenz?

1

In der letzten Zeit ist der Begriff Künstliche Intelligenz (KI) in aller Munde. Presse, Parlamente und Regierungen betrachten KI als entscheidenden Motor für die weitere wirtschaftliche Entwicklung des Landes. Die Bundesregierung hat deswegen ein massives Programm zur Förderung der KI beschlossen (Álvarez 2018). Experten des Beratungsunternehmens McKinsey schätzen, dass KI weltweit ein Umsatzvolumen von ca. 12 Billionen Euro bis zum Jahre 2030 generieren wird (Tung 2018).

„*Künstliche Intelligenz* ist die Fähigkeit eines Computers oder computergesteuerten Roboters, Aufgaben zu lösen, die normalerweise von intelligenten Wesen erledigt werden“ (Copeland 2019). Das System soll in der Lage sein, sich ähnlich wie ein Mensch intelligent zu verhalten und selbstständig zu lernen. Allerdings ist diese Definition ungenau, da der Begriff „Intelligenz“ schwer abzugrenzen ist.

1.1 Menschliche Intelligenz hat viele Dimensionen

Es gibt eine Reihe von unterschiedlichen Beschreibungen der menschlichen Intelligenz. Gardner (1983) hat eine Theorie der multiplen Intelligenzen entwickelt, welche acht Dimensionen der Intelligenz aufführt (Abb. 1.1). Die Bewegungsintelligenz ist die Fähigkeit, den eigenen Körper zu fühlen und kontrolliert zu bewegen. Die bildlich-räumliche Intelligenz ermöglicht die Erkennung von Bildern und das Erfassen räumlicher Zusammenhänge. Zur Sprachintelligenz gehört das Verstehen von Sprache und die angemessene sprachliche Formulierung von Sachverhalten. Die logisch-mathematische Intelligenz ermöglicht die Analyse und Lösung logischer Probleme. Musikalische Intelligenz ist erforderlich zum verständnisvollen Hören von Musik und zum Musizieren. Die naturalistische Intelligenz umfasst die Fähigkeit, Natur zu beobachten, zu unterscheiden, zu erkennen, sowie eine Sensibilität für Naturphänomene zu entwickeln. Die zwischenmenschliche

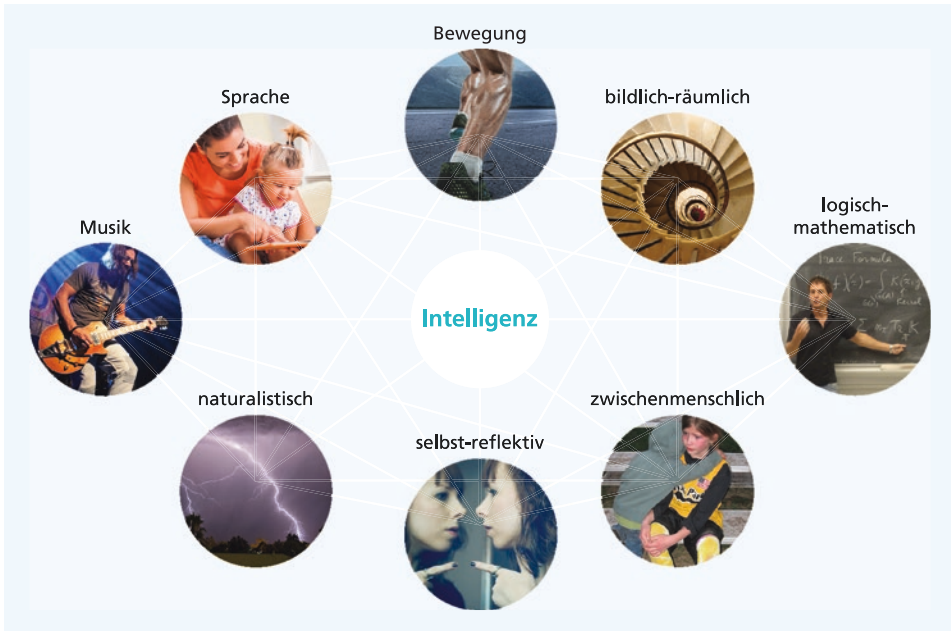


Abb. 1.1 Die Dimensionen der menschlichen Intelligenz nach Gardner (1983). Bildnachweis im Anhang

oder emotionale Intelligenz ist das Vermögen, die Absichten, Gefühle und Motive anderer Menschen zu verstehen und vorherzusagen. Die selbstreflektive Intelligenz umfasst die Begabung, die eigenen Stimmungen, Antriebe, Motive und Gefühle zu erkennen. Hierzu gehört auch ein Bewusstsein der eigenen Person und die Fähigkeit, das eigene Verhalten in neuen Situationen vorherzusagen und sich zu Handlungen zu motivieren. Wir werden sehen, dass die KI mittlerweile auf viele – aber nicht alle – dieser Dimensionen anwendbar ist.

1.2 Woran erkennt man Künstliche Intelligenz

Um zu beurteilen, ob ein Computersystem intelligent ist, hat der britische Mathematiker Alan Turing ein Testverfahren – den Turing-Test – vorgeschlagen (Turing 1950). Im Rahmen des Tests kann ein menschlicher Schiedsrichter mit zwei Partnern elektronisch kommunizieren und beliebige Fragen stellen: der eine Partner ist ein Mensch, der andere ein Computer (Abb. 1.2). Wenn der Schiedsrichter nach vielen Fragen anhand der Antworten nicht entscheiden kann, welcher der Partner der Computer ist, so gilt der Computer als intelligent.

In Hinblick auf die oben dargestellten Dimensionen der Intelligenz muss der Turing-Test allerdings erweitert werden, damit auch andere Dimensionen (Bild, Bewegung, Sprache) erfasst werden können.

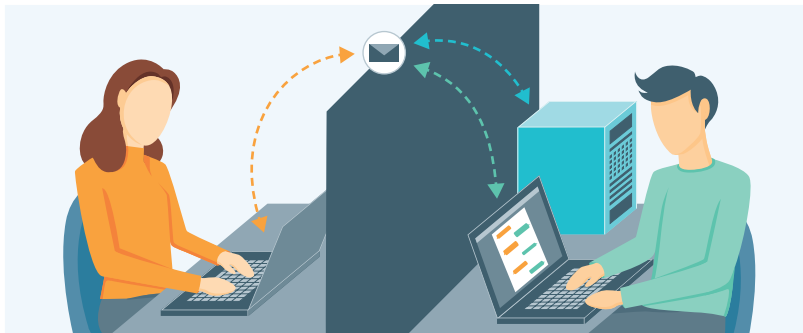


Abb. 1.2 Beim Turing-Test stellt der Schiedsrichter (links) Fragen an Partner, die er nicht sehen kann: einen Menschen und einen Computer (rechts). Von beiden erhält er Antworten. Wenn der Schiedsrichter den Computer durch seine Antworten nicht vom menschlichen Partner unterscheiden kann, so muss der Computer ebenfalls intelligent sein.

Viele Forscher favorisieren mittlerweile neue Testkriterien, die stärker prüfen, wie tief das Verständnis einer Situation reicht. Zum Beispiel könnte der Schiedsrichter mit den Partnern über ein gezeigtes Netflix-Video sprechen. „Warum ist diese Szene mit Bill Murray lustig?“ etwa wäre eine Frage, die für einen Computer schwerer zu beantworten ist als „Erzähl mir von deiner Mutter!“.

Schon früh wurde versucht, Computer manuell so zu programmieren, dass sie intelligentes Verhalten zeigen. Leider führten diese Ansätze nur mit Einschränkungen zu dem gewünschten Erfolg. Als Alternative setzte sich das Vorgehen durch, ein lernfähiges Computerprogramm zu entwickeln. Dieses Lernverfahren trainiert dann mit Hilfe von Beispieldaten die gewünschte Funktionalität. Hierdurch ist es heute möglich, Teilaufgaben der KI befriedigend zu lösen. Beispiele sind die Diagnose von Krankheiten anhand von Symptomen oder Röntgenaufnahmen, die Transkription von gesprochener Sprache in Text oder die Erkennung von Objekten in Bildern.

1.3 Computer lernen

Aber was heißt „Lernen“ bei einem Computersystem? Nehmen wir als Beispiel die Erkennung von Objekten in Bildern, z. B. einer Katze.

Als Eingabe erhält der Computer hier das Bild einer Katze (Abb. 1.3). Im rechten Teil von Abb. 1.3 sieht man einen vergrößerten Bildausschnitt, aus dem deutlich wird, dass das Bild der Katze ein Rechteck ist, das aus lauter kleinen quadratischen Farbflächen (Pixeln) besteht. Jedes dieser Pixel hat eine Farbe, welche man durch die Anteile der drei Grundfarben Rot, Grün und Blau kennzeichnen kann. Ein Pixel kann also durch ein Tripel von Zahlen beschrieben werden und das ganze Bild durch ein rechteckiges Schema von Zahlentripeln.

Der Computer erhält das Bild in Form eines rechteckigen Schemas von Zahlentripeln als Eingabe. Ziel ist es nun, dass der Computer das wichtigste Objekt in dem Bild benen-

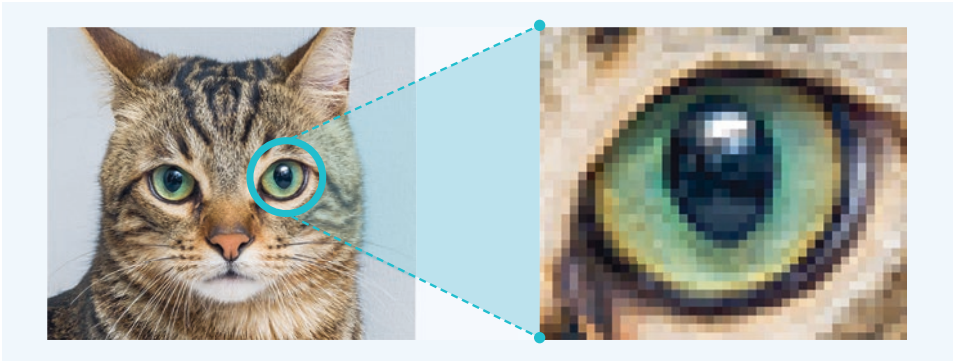


Abb. 1.3 Bild einer Katze und Bildausschnitt mit einzelnen Pixeln. Jedes Pixel wird durch drei Zahlen, die Farbwerte für Rot, Grün und Blau, beschrieben. Bildnachweis im Anhang

nen kann, in unserem Fall „Katze“. Diese Aufgabe nennt man Objektklassifikation in Bildern, eine Teilaufgabe der Bilderkennung. Es wird dem Computer also nicht mitgeteilt, wo im Bild das Objekt ist, das er benennen soll.

Frühe Lösungsansätze für diese Aufgabe versuchten, zunächst vorgegebene Teile der Bildobjekte zu erkennen, z. B. Ecken, Kanten, Linien und Flächen. Die größeren Objekte (z. B. Auge) wurden dann als Verbindungen der kleineren Teile rekonstruiert. Allerdings brachte dieses Vorgehen keine guten Resultate.

In letzter Zeit wurden Verfahren erprobt, bei denen der Computer nicht mehr die vom Menschen definierten Eigenschaften (Ecken, Kanten, Linien und Flächen) verwendet. Vielmehr wählt er selbstständig wichtige Merkmale aus, erkennt diese im Bild und nutzt sie dann zur Objektklassifikation. Dazu benötigt er allerdings eine große Menge von Beispielen, in denen das gesuchte Bildobjekt (z. B. Katze) vorkommt und auch Beispielsbilder, in denen es nicht vorkommt. Nur so kann der Computer die Gemeinsamkeiten und Unterschiede der Objekte erkennen.

Grundlage der Objektklassifikation ist also eine große Menge von Beispielen, welche aus der Eingabe (Bild) und der zugehörigen Ausgabe, der Objektklasse (z. B. Affe, Katze, ...) bestehen (Abb. 1.4). Die Menge der Beispiele nennt man „Trainingsmenge“ oder auch „Trainingsdaten“. Die Elemente der Trainingsmenge nennt man auch Trainingsbeispiele.

Aufgabe des Computers ist nun, die Menge der Beispiele und der zugehörigen Objektklasse zu analysieren. Anschließend soll er selbstständig eine Rechenvorschrift zu entwickeln, mit der die Objektklassen von neuen Objekten möglichst gut vorhergesagt werden können. Die Bestimmung einer solchen Rechenvorschrift nennt man „Lernen“. Die Situation ist vergleichbar mit der eines Kleinkindes, dem die Mutter, wie in Abb. 1.5, die Namen von Gegenständen im Bilderbuch sagt. Dabei lernt das Kind, wie es die verschiedenen Objekte voneinander unterscheiden und benennen kann.

Lernen wird als der Vorgang definiert, mit dem neue oder modifizierte Fähigkeiten, Wissensinhalte oder Verhaltensmuster erworben werden (Wikipedia-Learning 2019).

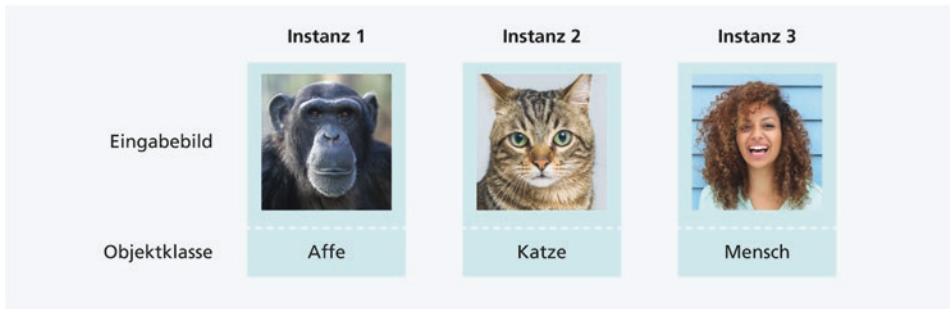


Abb. 1.4 Trainingsdaten aus unterschiedlichen Klassen für ein Bilderkennungssystem. Jedes Trainingsbeispiel besteht aus einem Eingabebild und der zugehörigen Objektklasse. Es sind pro Klasse eine hohe Anzahl von Trainingsbeispielen erforderlich. Bildnachweis im Anhang



Abb. 1.5 Eine Mutter zeigt ihrem Kind Objekte im Bilderbuch. Bildnachweis im Anhang

Gemeinhin wird Lernen als eine zutiefst menschliche Fähigkeit verstanden. Daher sind viele Leute nicht gewillt, einem Computerprogramm eine Lernfähigkeit zuzugestehen. Allerdings können auch Tiere lernen, wie viele Experimente aus der Biologie beweisen. Anders als bei lebenden Organismen, ist Lernen im Bereich der KI hingegen eher mit dem Begriff „trainieren“ gleichzusetzen: Hierbei kann das System die Fähigkeit erlangen, zu gegebenen Eingaben die passenden Ausgaben (z. B. Objektklasse) zu bestimmen. Dies bedeutet nicht, dass das System die Objekte der Trainingsmenge „auswendig“ lernt, son-

dern es kann auch neuen, noch nicht bearbeiteten Objekten die korrekte Klasse zuordnen. In diesem Sinne wird der Begriff „lernen“ in diesem Buch verwendet.

Es gibt eine Reihe anderer Verben, die normalerweise im Zusammenhang mit Menschen verwendet werden, aber auch im Bereich der KI auftauchen. Dazu gehören „erkennen“, „wissen“ usw. Übt der Mensch diese Tätigkeiten aus, so ist das immer verbunden mit menschlichem Bewusstsein und Emotionen. Im Bereich der KI werden diese Aspekte vollständig ausgeklammert. Dies muss man bei der Lektüre dieses Buch immer berücksichtigen.

1.4 Tiefe neuronale Netze können Objekte erkennen

Lernaufgaben, wie etwa die Objektklassifikation in Bildern, können heute von tiefen neuronalen Netzen durchgeführt werden. Wie in Abschn. 5.1.2 dargestellt wird, haben *tiefe neuronale Netze* (TNN) strukturelle Ähnlichkeiten mit der Informationsverarbeitung im Gehirn. Sie verarbeiten die Eingaben in einer Anzahl von aufeinanderfolgenden *Schichten* und transformieren die Eingabedaten in abstraktere Merkmale, die durch Pakete von Zahlen repräsentiert werden. In jeder Schicht werden ausgewählte Merkmale der Szene verarbeitet – je höher die Schicht, desto komplexer die Merkmale. Diese Merkmale wählt das System selbst aus. Abb. 1.6 zeigt die auf diese Weise von Lee et al. (2011) extrahierten Merkmale für die Klassifikation eines Objektes als Mensch. Schließlich lassen sich aus

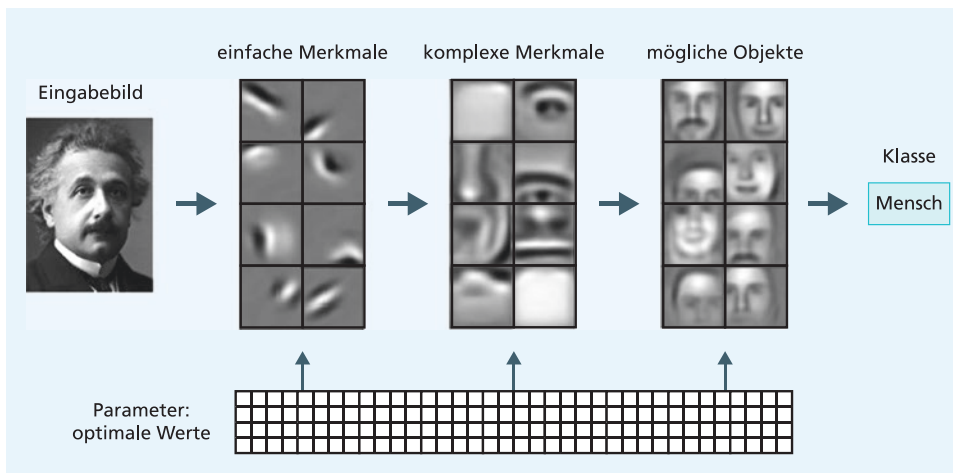


Abb. 1.6 Ein tiefes neuronales Netz (TNN) erhält eine Eingabe, z. B. das Bild einer Person. Aus diesem Bild werden in der unteren Schicht einfache Merkmale extrahiert, in den nachfolgenden Schichten komplexere Merkmale. Die Zuordnung zu einer Objektklasse findet in der letzten Schicht statt. Bildnachweis im Anhang

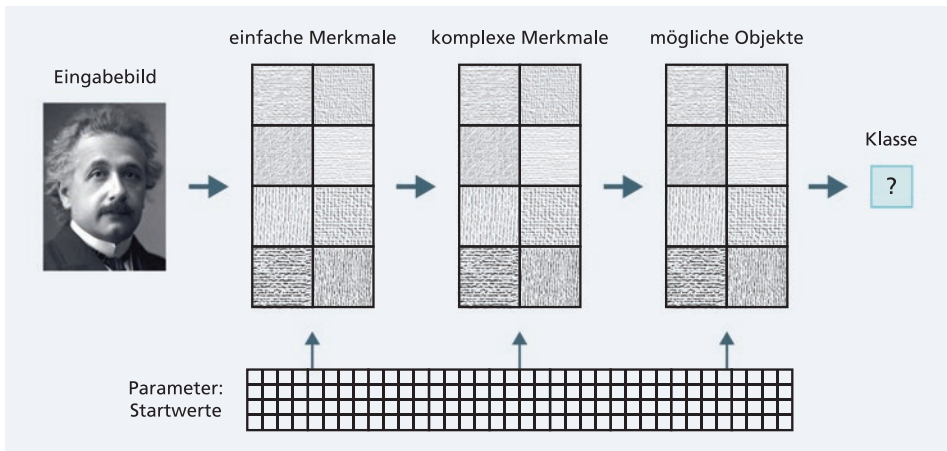


Abb. 1.7 Berechnete Zwischenmerkmale des TNN und ausgegebene Klassifikation zu Beginn des Trainings mit zufällig ausgewählten Startwerten für das Parameter-Zahlenpaket. Das TNN hat noch nichts gelernt. Bildnachweis im Anhang

den Merkmalen der letzten Schicht in einfacher Weise die gewünschten Ergebnisse, z. B. die Namen der Objekte, bestimmen.

Allerdings kann das tiefe neuronale Netz Bilder nur erkennen, wenn seine Parameter angepasst wurden. Die Parameter sind ebenfalls eine Menge von Zahlen – ein Zahlenpaket – welches die Eigenschaften des TNN steuert. Der Aufbau des TNN und die Anzahl der Zahlen in dem Parameter-Zahlenpaket werden dabei von dem Konstrukteur des Netzes festgelegt. Das Parameter-Zahlenpaket wird zu Beginn mit zufälligen Zahlenwerten gefüllt. Wie in Abb. 1.7 gezeigt, kann das TNN zu in diesem Zustand weder sinnvolle Zwischenmerkmale erkennen noch das gezeigte Objekt identifizieren.

Wie oben dargestellt, werden die Werte des Parameter-Zahlenpaketes an eine große Menge von Trainingsbeispielen angepasst. Diese bestehen in der Regel aus der Eingabe (Bild) und der zugehörigen Ausgabe, also der Klasse des Bildobjektes (z. B. Affe, Katze, ... Siehe Abb. 1.4). Meist sind für jede Klasse Hunderte solcher Trainingsbeispiele erforderlich. Der Computer passt nun die Werte des Parameter-Zahlenpaketes schrittweise so an, dass das TNN möglichst zu jedem Eingabebild die korrekte Klasse ausgibt. In den letzten Jahren ist es gelungen, auch Millionen von verschiedenen Parameterwerten durch sukzessive kleine Änderungen gleichzeitig so zu modifizieren, dass in einem hohen Prozentsatz der Fälle die korrekte Ausgabe erzeugt wird.

Dieses Vorgehen hat in letzter Zeit zu überraschend guten Ergebnissen in einer Vielzahl von intelligenten Erkennungsaufgaben geführt. Diesen Vorgang nennt man auch „*tiefes Lernen*“. Die Details dieses Lernvorgangs werden in den späteren Kapiteln dargelegt.

Tiefes Lernen ist eine spezielle Technik des *maschinellen Lernens*. Dieses umfasst alle Verfahren zur Suche von Mustern und Zusammenhängen in Daten (Abb. 1.8). Ein solches

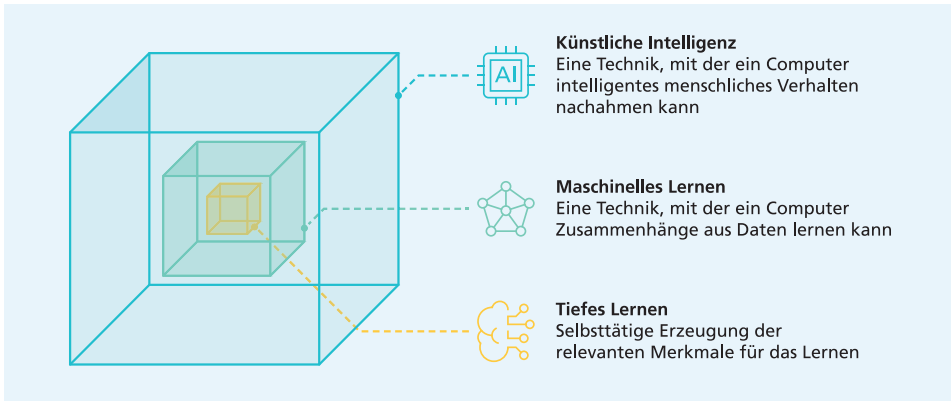


Abb. 1.8 Künstliche Intelligenz ist ein Oberbegriff von maschinellern Lernen, welches seinerseits das tiefe Lernen umfasst.

System kann beispielsweise die Niederschlagsmenge von morgen aus den Messwerten von Luftdruck, Temperatur und Windrichtung von heute vorhersagen.

Obwohl Künstliche Intelligenz und tiefe neuronale Netze in vielen Zeitschriftenbeiträgen und Talkshows diskutiert werden, liegt für die meisten Menschen die Funktionsweise dieser Computerprogramme im Dunkeln. Dieses Buch möchte daher für eine interessierte Öffentlichkeit klären, was Künstliche Intelligenz und tiefe neuronale Netze sind und wie sie funktionieren. Dabei sollen nicht nur die internen Mechanismen dargestellt, sondern auch die derzeitigen Möglichkeiten und Grenzen verdeutlicht werden.

1.5 Wie kann man Künstliche Intelligenz verstehen?

Die meisten Menschen werden von sich glauben, dass sie die Funktionsweise eines Autos verstehen. Abb. 1.9 zeigt das Funktionsschaubild eines Autos. Im Motor fangen Kolben den bei der Verbrennung entstehenden Druck auf und setzen ihn über die Kurbelwelle in eine Drehbewegung um. Das Getriebe in Zusammenspiel mit der Kupplung bestimmt die Geschwindigkeit der Drehbewegung, welche über das Differential auf die Räder übertragen wird. Dieser grobe Ablauf reicht den meisten Menschen für das Verständnis der Reaktionen des Autos bei der Steuerung durch den Fahrer aus. Dabei sind Details der elektronischen Motorsteuerung, des Getriebes mit seinen verwundenen Zahnrädern, der Servolenkung, des Bremskraftverstärkers usw. extrem kompliziert und können ohne eine Ingenieurausbildung nicht verstanden werden.

Künstliche Intelligenz kann man auf einem ähnlichen Abstraktionsniveau verstehen. Hier werden nicht Kräfte durch mechanische Bauteile übertragen, sondern Zahlenpakete werden durch Operatoren geschickt, die nach einem einfachen Schema ein

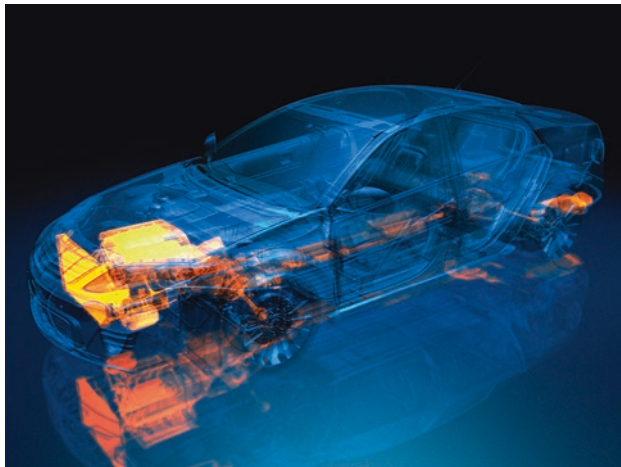


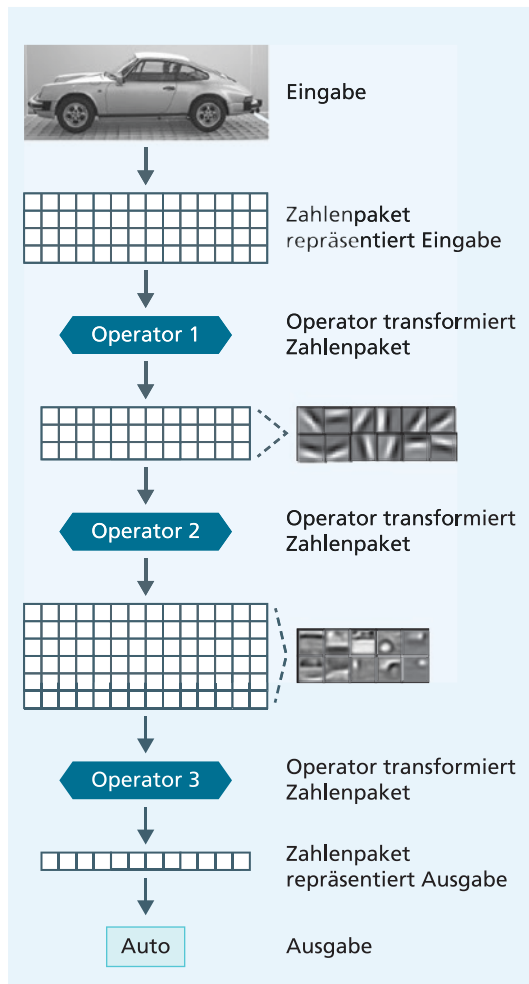
Abb. 1.9 Funktionsschaubild eines Autos mit Motor, Getriebe, Antriebswelle, Differential und Rädern. Bildnachweis im Anhang

Eingabezahlenpaket in ein Ausgabezahlenpaket transformieren (Abb. 1.10). Die Eingabezahlenpakete repräsentieren die Eingaben der Anwendung, z. B. Bilder, Tonaufnahmen, Texte, Videos. Jeder Operator erzeugt ein neues Zahlenpaket, welches meist als Eingabezahlenpaket des nächsten Operators verwendet wird. Die Gesamtheit der verbundenen Operatoren wird ein Modell genannt. Das letzte Ausgabe-Zahlenpaket des Modells repräsentiert die gewünschte Antwort, z. B. eine Bildkategorie, eine Übersetzung, oder ein neues Bild, welche von dem Modell erzeugt wird.

Das Verständnis der Künstlichen Intelligenz in diesem Buch wird auf diesem hohen Abstraktionsniveau vermittelt. Es wird die grobe Funktion der einzelnen Operatoren erklärt, ähnlich wie die Erklärung von Motor, Getriebe und Differential beim Auto. Es wird der Fluss der Zahlenpakete durch das Modell dargestellt, analog zur Kraftübertragung beim Auto. Und es wird sehr grob die Wirkungsweise der Optimierungsverfahren erklärt, welche das Modell an die Trainingsbeispiele anpassen. Diese Verfahren werden in der Regel von den vorhandenen Programmierwerkzeugen bereitgestellt.

Das in diesem Buch vermittelte Bild von Künstlicher Intelligenz bleibt auf dieser relativ abstrakten Ebene. Viele Details sind sehr komplex, aber auch nicht für ein grundlegendes Verständnis erforderlich.

Abb. 1.10 Funktions-schaubild eines Modells der künstlichen Intelligenz mit Eingabe, Zahlenpaketen, Operatoren und Ausgabe. Die Zahlenpakete repräsentieren je nach Schicht unterschiedliche Inhalte. Bildnachweis im Anhang



1.6 Die Geschichte der Künstlichen Intelligenz

Es ist aufschlussreich, sich die wechselvolle Geschichte der Künstlichen Intelligenz zu vergegenwärtigen (Abb. 1.11). Als Mitte des letzten Jahrhunderts die ersten programmierbaren Computer entwickelt wurden, dachten die Forscher schon bald darüber nach, ob diese Geräte auch intelligentes Verhalten an den Tag legen könnten. Zum Test eines Systems auf intelligentes Verhalten schlug Alan Turing 1950 den „Turing-Test“ vor. 1956 fand der von John McCarthy und Marvin Minsky veranstaltete Dartmouth Workshop statt, welcher den Begriff „Künstliche Intelligenz“ prägte. Ein Jahr später entwickelte Frank Rosenblatt ein neuronales Netz, das Perzeptron, welches trainiert werden konnte, einfache Muster zu unterscheiden. Etwa zur gleichen Zeit wurden erste Programme zum logischen