

IT kompakt

Daniel Sonnet

# Neuronale Netze kompakt

Vom Perceptron zum Deep Learning

 Springer Vieweg

---

# IT kompakt

Die Bücher der Reihe „IT kompakt“ zu wichtigen Konzepten und Technologien der IT:

- ermöglichen einen raschen Einstieg,
- bieten einen fundierten Überblick,
- eignen sich für Selbststudium und Lehre,
- sind praxisorientiert, aktuell und immer ihren Preis wert.

Weitere Bände in der Reihe <https://link.springer.com/bookseries/8297>

---

Daniel Sonnet

# Neuronale Netze kompakt

Vom Perceptron zum Deep  
Learning

 Springer Vieweg

Daniel Sonnet  
Hochschule Fresenius Hamburg  
Hamburg, Deutschland

ISSN 2195-3651                      ISSN 2195-366X (electronic)  
IT kompakt  
ISBN 978-3-658-29080-1              ISBN 978-3-658-29081-8 (eBook)  
<https://doi.org/10.1007/978-3-658-29081-8>

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

© Springer Fachmedien Wiesbaden GmbH, ein Teil von Springer Nature 2022  
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/Lektorat: David Imgrund  
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

---

# Inhaltsverzeichnis

<b>1</b>	<b>Einführung</b> . . . . .	1
<b>2</b>	<b>Neuronale Netze</b> . . . . .	17
2.1	Ein kompakter historischer Abriss . . . . .	18
2.1.1	Der Start: das Perceptron . . . . .	18
2.1.2	Die Weiterentwicklung: mehrschichtige Netze . . . . .	26
2.1.3	Heutiger Status Quo: Deep Learning . . . . .	37
2.2	Auswahl einiger Lern- bzw. Trainingsalgorithmen . . . . .	45
2.2.1	Hebb'sche Regel . . . . .	46
2.2.2	Deltaregel . . . . .	48
2.2.3	Backpropagation . . . . .	51
2.3	Vorstellung besonderer Netzwerktypen . . . . .	61
2.3.1	Feedforward networks versus recurrent networks . . . . .	62
2.3.2	Convolutional neural network . . . . .	65
<b>3</b>	<b>Best Practice: Möglichkeiten und Grenzen Neuronaler Netze</b> . . . . .	71
3.1	Vorteile und Möglichkeiten Neuronaler Netze . . . . .	71
3.2	Grenzen Neuronaler Netze . . . . .	84
<b>4</b>	<b>Ausblick: Mögliche Entwicklungen für Neuronale Netze</b> . . . . .	101
4.1	Entschlüsselung der Black Box . . . . .	101

4.2	Mehr Performance und weitere Einsatzmöglichkeiten.....	105
<b>5</b>	<b>Quick Start Guide Neuronale Netze und Fallstudien</b> .....	<b>111</b>
5.1	Quick Start Guide .....	111
5.2	Fallstudien .....	121
5.2.1	Vorhersage der Churn Rate (Kundenverlustrate) einer Bank .....	121
5.2.2	Der Klassiker MNIST sowie Fashion-MNIST .....	132
	<b>Literatur</b> .....	<b>143</b>



## Zusammenfassung

Daten sind das Gold des 21. Jahrhunderts, oder Daten sind das neue Öl. Diese markanten Sprüche erfreuen sich gerade enormer Beliebtheit. Im Kern soll ausgedrückt werden, dass gesammelte Daten für Institutionen wie Unternehmen und Forschungsstätten immense Werte haben können, die es zu erschließen gilt. In diesem Kontext fällt der Begriff der Neuronalen Netze, die sich zu Recht bereits auf sehr vielen Gebieten bewährt haben. Neuronale Netze sind ein Teil der künstlichen Intelligenz (KI), oder wie war das noch einmal genau? Wir beginnen dieses Kapitel darum, indem wir zunächst die Begriffe, KI, Machine Learning und Neuronale Netze sowie ihre Verbindung zu klassischen Disziplinen wie Mathematik, Statistik und Informatik sortieren. Wir werden anschließend beleuchten (in einer sehr vereinfachten Form) wie Menschen lernen und dies mit der klassischen Wissensakquise von Computerprogrammen vergleichen. Es werden die drei Begriffe Supervised, Unsupervised und Reinforcement Learning in diesem Kontext erarbeitet. In allen drei Bereichen sind Neuronale Netze eine große Bereicherung, wie wir feststellen werden. Diese flexible und universelle Einsatzfähigkeit motiviert die weitere Beschäftigung mit Neuronalen Netzen. Jedes Kapitel dieses Buches greift einen anderen Aspekt Neuronaler Netze auf. Am Ende des vorliegenden Einführungskapitels wird



zusammengefasst welcher Aspekt des faszinierenden Bereichs in den folgenden Kapiteln behandelt wird.

Daten sind wertvoll, da sie ggf. verwertbare Muster enthalten können. Die Idee ist simpel. Mittels Algorithmen wird das Muster extrahiert, um es anschließend auf neue Daten anzuwenden. Die Extraktion sowie die Anwendung von Mustern ist eine Paradisziplin Neuronaler Netze. Beispiele für den erfolgreichen Einsatz von Neuronalen Netzen sind:

- Bilderkennung, z. B. für das autonome Fahren,
- Spracherkennung, z. B. für die automatische Übersetzung eines gesprochenen Textes,
- Gesichtserkennung z. B. bei der Benutzeridentifikation von Smartphones,
- Schrifterkennung, z. B. für die Zuordnung von händisch geschriebenen Ziffern,
- Steuerung von technischen Prozesse, z. B. die Regelung von Kühlsystemen von Serverparks,
- Prognosen, z. B. wie viele Menschen werden am nächsten Tag abhängig vom Wetter auf einem Golfplatz spielen,
- Frühwarnsysteme, z. B. Predictive Maintenance - ein System warnt wenn bei einer Produktionsmaschine eine Wartungsarbeit ansteht,
- Zeitreihenanalysen, z. B. Prognosen, ob der DAX fallen oder steigen wird,
- Unterstützung von Ärztinnen und Ärzten im Rahmen von Krankheitsdiagnosen,
- biometrische Systeme, z. B. die Erkennung von Gesichtern während der Einreisekontrolle an Landesgrenzen,
- Wirtschaftsmodelle, z. B. Analyse, welche Auswirkung hat eine Leitzinserhöhung durch die EZB auf die Inflation im Eurowirtschaftsraum,
- und viele weitere.

Neuronale Netze betten sich in die Themen künstliche Intelligenz (engl. Artificial Intelligence) und Machine Learning ein. Teile

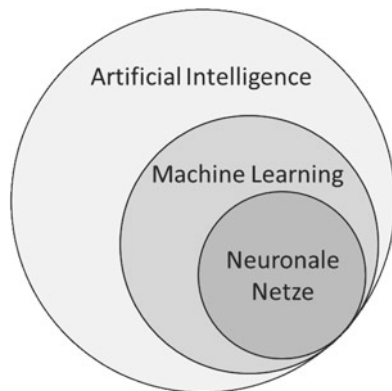
dieser Disziplinen sind noch recht neu, erfreuen sich jedoch wachsender Beliebtheit. Dieses Buch hat das Anliegen, eine kompakte Einführung in das Thema Neuronale Netze zu liefern. Darum sortieren wir zunächst, wie Neuronale Netze zu den anderen genannten Begriffen stehen. Abb. 1.1 illustriert, dass Neuronale Netze eine Teildisziplin des maschinellen Lernens sind, was wieder um ein Bereich der künstlichen Intelligenz ist.

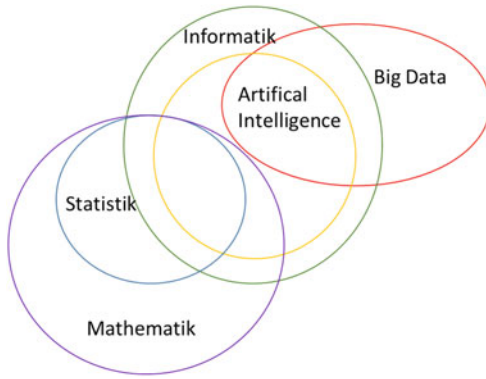
Es kommt sofort die Frage auf, wie die künstliche Intelligenz zu klassischen Disziplinen wie Mathematik und Informatik steht. In Abb. 1.2 ist der Versuch unternommen, eine solche Abgrenzung durchzuführen. Selbstverständlich könnte man diese Grafik beliebig ausdehnen und versuchen, weitere Disziplinen einzusortieren. Für den Moment erscheint es jedoch ausreichend festzuhalten, dass künstliche Neuronale Netze in der Informatik zu verordnen sind, die wiederum Schnittmengen mit der Mathematik und der Statistik hat.

Dieses Buch fokussiert auf Neuronale Netze und zieht Methoden aus angrenzenden Disziplinen nur dann heran, wenn es nötig und für das Verständnis hilfreich ist.

Warum sind Neuronale Netze derzeit so populär? Sie können leistungsstarke Partner sein, wenn es um die Analyse von (großen) Datenbeständen geht. Sie erfreuen sich derzeit großer Beliebtheit und sind in der breiten Diskussion angekommen. Darüber hinaus

**Abb. 1.1** Teilmengen der künstlichen Intelligenz





**Abb. 1.2** Teilmengen der künstlichen Intelligenz

sind sie mittlerweile ein fester Bestandteil in zahlreichen Businessanwendungen. Diverse Beispiele sind bereits oben in einer Liste aufgeführt. Die ersten Neuronale Netze wurden bereits am Ende der 1950er-Jahre konstruiert. Damals waren datengetriebene Geschäftsmodelle, wie wir es heute von Amazon, Google und Co. kennen, selbstverständlich noch kein Thema. Zu diesem Zeitpunkt wurden Neuronale Netze von Menschen mit wissenschaftlichen Anliegen kreiert und Big Data stand damals nicht im Fokus. Das erste dieser Netze geht auf den Psychologen Frank Rosenblatt zurück. Bis heute ist das erste Neuronale Netz vielfältig weiterentwickelt worden. Dabei waren Mathematiker und Informatiker stark beteiligt. Wenn viele unterschiedliche Menschen am gleichen Thema über einen längeren Zeitraum arbeiten, ist eine Begriffsvielfalt nicht verwunderlich. Es wurden Begriffe wie Perceptron, Feedforward Netze, Backpropagation, Convolutional Neural Network, Supervised und Unsupervised Training etc. geprägt. Als Einsteiger mit wenig Zeit in das Thema, kann dies zu Frust führen. Das Anliegen dieses Werkes ist darum, eine möglichst kompakte Einführung in das Thema zu geben, ohne auf die wichtigsten Praxisbegriffe zu verzichten. Das gerade Gesagte kann auf das Thema Darstellung von mathematischen Inhalten erweitert werden. Neuronale Netze wurden maßgeblich von Menschen mit guten mathematischen Kenntnissen entwickelt, und dementsprechend ist die

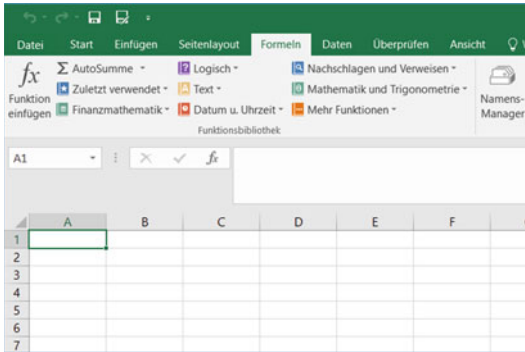
Literatur zu ihnen größtenteils mathematisch. Um weiterhin eine möglichst kompakte und intuitive Einführung in das Thema zu bieten, wird im vorliegenden Werk wenn möglich auf mathematische Schreibweisen verzichtet. Wenn dennoch einmal eine Formel auftaucht, ist sie eher als Ergänzung zu verstehen, sie ist nicht essentiell, um das grundsätzliche Themengebiet und das Vorgehen Neuronaler Netze zu verstehen.

Wann sollten Sie dieses Buch lesen? Wenn Sie auf der Suche nach einer kompakten Einführung in das Thema sind, die Sie schnellstmöglich mit Neuronalen Netzen vertraut macht. Oder anders formuliert, wenn es auf Sie zutrifft, dass Sie wenig Zeit besitzen, jedoch bald mit Kollegen oder Beratern im richtigen Vokabular zu Neuronalen Netzen fachsimpeln möchten, dann sind Sie hier richtig.

Wann sollten Sie dieses Buch nicht lesen? Wenn Sie sich bereits gut in das Themengebiet eingearbeitet haben, wird das vorliegende Werk Ihre speziellen Fragen ggf. nicht beantworten. Dafür fehlt in diesem kompakten Werk der Raum. Eine Literaturempfehlung wäre für Sie z. B. [28].

Neuronale Netze sind Werkzeuge des maschinellen Lernens, was bereits Abb. 1.1 nahelegt. Dies wirft die Frage auf, wann ein Computerprogramm als maschinell lernfähig gilt? Nicht jede Software die *smart* daherkommt ist aus diesem Bereich. Ist z. B. Microsoft Excel aus dem Bereich des maschinellen Lernens? Excel ist ein leistungsfähiges und sehr wertvolles Werkzeug im täglichen Umgang mit Daten. Es ermöglicht sogar einzelne statistische Analysen. Dennoch ist Excel ganz klar keine Software des Feldes „Machine Learning“. Excel ist eine Software die im klassischen Sinne programmiert ist. Dies bedeutet, dass ein von Menschen erdachtes Regelwerk, welches in vielen Tausend Zeilen Code in der Software steckt, dafür sorgt, dass Excel so reagiert, wie es von Excel erwartet wird. Sobald z. B. der Anwender auf die Schaltfläche „Formeln“ klickt, öffnet Excel das Formelmenü und bietet unterschiedliche Formeln zur Auswahl an. Der Nutzer des Programm sieht sich dann einer Situation wie in Abb. 1.3 dargestellt gegenüber.

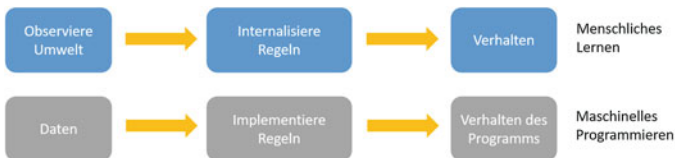
Dieses Verhalten wurde Excel mittels implementierter Regeln quasi in die DNA eingesetzt. Excel hat sich dieses Verhalten nicht selbst beigebracht, bzw. Excel hat nicht gelernt, dass dieses



**Abb. 1.3** Ein Klick auf den Reiter Formel öffnet in Excel den Formeleditor

Verhalten sinnvoll bzw. vom Nutzer gewünscht ist. Excel befolgt ausschließlich Regeln. Das Schreiben von Software auf der Basis von extrahierten Regeln (es wurde am Anfang auch von Muster gesprochen) ist angelehnt an den idealisierten Wunsch des menschlichen Lernens. Kinder zum Beispiel observieren die Welt und agieren stark auf der Basis von Regeln. Wenn die Ampel Rot zeigt, gehen Kinder nicht über die Ampel. Diesen regelbasierten Ansatz, um das gewünschte Verhalten von Software herzustellen, ist daran angelehnt. In Abb. 1.4 ist es dargestellt.

Ein Programm aus dem Bereich maschinellen Lernens kennt das Konzept von fest implementierten Regeln nicht. Statt Regeln wird eine Einheit implementiert, die in der Lage ist, aus bereitgestellten Daten ein relevantes Muster zu erkennen. Sobald das Muster extrahiert ist, kann es benutzt werden, um Einschätzungen zu dem Ergebnis neuer Daten zu bestimmen. Nehmen wir das Beispiel der



**Abb. 1.4** Die Programmierung klassischer Software basiert auf Regeln

Unterscheidung von Autos und Fahrrädern. Der Einfachheit halber werden lediglich die beiden Variablen Anzahl Räder und Gewicht benutzt. Natürlich kann es sinnvoll sein, andere oder weitere Variablen einzusetzen. Hier wird sich zunächst auf diese beschränkt. In Abb. 1.5 sind verschiedene Fahrräder und Autos gemäß der beiden Variablen Anzahl Räder und Gewicht beschrieben.

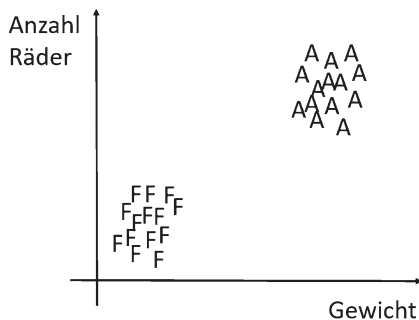
Aus Abb. 1.5 könnte man z. B. die Regel ableiten, wenn das Gewicht zwischen 500kg und 2000kg liegt und das Objekt 4 Räder hat, handelt es sich um ein Auto. Diese durch Menschen kreierte Regel wäre in eine Software implementierbar. Dieses Vorgehen entspräche dem in Abb. 1.4 gezeigten maschinellen Programmieren. Das maschinelle Lernen verfolgt einen alternativen Ansatz. Hierbei wird einem Lernalgorithmus Daten (Trainingsbeispiele) präsentiert und die Aufgabe gestellt, das dahinterliegende Pattern/Muster zu extrahieren, damit Fahrräder von Autos unterscheidbar werden. Es ist damit ähnlich zum Vorgehen beim menschlichen Lernen aus Erfahrungen, wie es in Abb. 1.4 gezeigt ist. Ein fiktiver Auszug solcher möglichen Lerndaten ist in Tab. 1.1 gezeigt.

Der Lernalgorithmus könnte durch simples Einziehen zweier Geraden in Abb. 1.5 einen Fahrradbereich sowie einen Autobereich festlegen. Diese Idee ist in Abb. 1.6 dargestellt.

Selbstverständlich hätte ein Lernalgorithmus statt Gerade auch andere Formen nutzen können. Hier wurde zur Illustration das einfachste Beispiel gewählt.

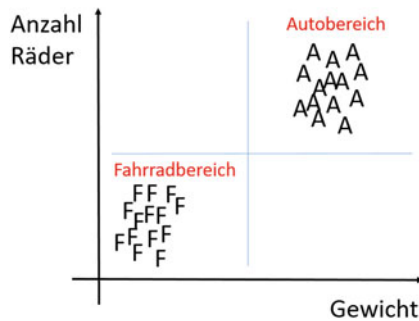
**Abb. 1.5**

Unterscheidung Autos  
und Fahrräder



**Tab. 1.1** Auszug möglicher Lerndaten zum Fahrrad-Auto-Problem

Gewicht in kg	Anzahl Räder	Auto/Fahrrad?
13	2	Fahrrad
7	2	Fahrrad
864	4	Auto
12	2	Fahrrad
1254	4	Auto
6	2	Fahrrad
...	...	...

**Abb. 1.6** Gelerntes Muster zur Unterscheidung von Autos und Fahrrädern

Für das Beispiel kann festgehalten werden: Nicht der Mensch hat Regeln zur Unterscheidung von Autos und Fahrrädern erdacht und in einem Computerprogramm implementiert. Ein Algorithmus hat selbstständig aus Daten gelernt, Autos und Fahrräder zu unterscheiden. Diese Eigenschaft bezeichnen wir als maschinelles Lernen. Allgemeiner gilt, dass Algorithmen beim maschinellen Lernen mittels Trainingsdaten ein (statistisches/mathematisches) Modell erlernen. Jedoch werden die Trainingsdaten nicht einfach nur auswendig gelernt, sondern das Muster bzw. die Gesetzmäßigkeiten in den Daten wird extrahiert.