



Verstehensprozesse modellieren und analysieren

Gregor Braun

F Frank & Timme

Verlag für wissenschaftliche Literatur

Gregor Braun
Verstehensprozesse modellieren und analysieren

Gregor Braun

Verstehensprozesse modellieren
und analysieren

F Frank & Timme

Verlag für wissenschaftliche Literatur

ISBN 978-3-86596-452-6

© Frank & Timme GmbH Verlag für wissenschaftliche Literatur
Berlin 2013. Alle Rechte vorbehalten.

Das Werk einschließlich aller Teile ist urheberrechtlich geschützt.
Jede Verwertung außerhalb der engen Grenzen des Urheberrechts-
gesetzes ist ohne Zustimmung des Verlags unzulässig und strafbar.
Das gilt insbesondere für Vervielfältigungen, Übersetzungen,
Mikroverfilmungen und die Einspeicherung und Verarbeitung in
elektronischen Systemen.

Herstellung durch das atelier eilenberger, Taucha bei Leipzig.

Printed in Germany.

Gedruckt auf säurefreiem, alterungsbeständigem Papier.

www.frank-timme.de

INHALTSVERZEICHNIS

GELEITWORT	7
DANKSAGUNG	8
1 Künstliche neuronale Netze	9
1.1 Einleitung.....	9
1.2 Biologisches Vorbild	10
1.3 Formales Modell des Neurons	12
1.4 Aufbau künstlicher neuronaler Netze	17
1.5 Die Gewichtsmatrix.....	21
1.6 Feedforward Netze	24
1.7 Feedback Netze	30
1.8 Semantische Netze.....	34
1.9 Lernregeln.....	37
1.10 Einsatzgebiete künstlicher neuronaler Netze	44
2 Das Experiment "Krimi-Leser".....	47
2.1 Einleitung.....	47
2.2 Erläuterung des Experiments.....	48
2.3 Fragebogen und Tabellen	53
2.4 Start des Experiments.....	58
2.5 Die spanische Truhe – Teil 1	60
2.6 Notizen und Tabellen für Teil 2	70
2.7 Die spanische Truhe – Teil 2	71
2.8 Notizen und Tabellen für Teil 3	81

2.9 Die spanische Truhe – Teil 3.....	82
2.10 Notizen und Tabellen für Teil 4.....	91
2.11 Die spanische Truhe – Teil 4.....	92
2.12 Abschließende Notizen	102
2.13 Die spanische Truhe – Letzter Teil.....	103
2.14 Überführen der Tabellen in ein künstliches neuronales Netz.....	105
3 Modellierung von Verstehensprozessen	109
3.1 Einleitung.....	109
3.2 Der "Krimi-Leser" und weitere Analysen	110
3.3 Eine erste Studie.....	114
3.4 Weitere Beispiele aus der Welt der Kriminalromane	118
3.5 Textinterpretation.....	130
3.6 Modellierung von Kommunikation mit semantischen Netzen / Verstehen von Nachrichten.....	132
3.7 Ausblick.....	135
LITERATURVERZEICHNIS	137

Zu diesem Buch

Dieses Buch entstand aus der von uns betreuten Masterarbeit von Gregor Braun und enthält seine Experimente auf der Basis unserer Konzeptionen. Wir waren und sind sehr angetan davon, wie Gregor Braun diese Ideen umgesetzt und in eigenen Programmen getestet hat. Außerdem schildert das Buch weitere Experimente, die wir mit verschiedenen Programmen durchgeführt haben. Es ging dabei immer um die Rekonstruktion verschiedener Kriminalgeschichten und um den Vergleich, wie menschliche Leser einerseits und die Programme aus dem Kontext der Künstlichen Intelligenz andererseits mit den Geschichten interpretatorisch umgingen. Es ist für uns sehr spannend, diese verschiedenen Experimente noch einmal zusammengefasst zu sehen. Auch dies ist Gregor Braun ausgesprochen gelungen.

Wir wünschen dem Buch den Erfolg, den es verdient; den Lesern wünschen wir, dass sie bei der Lektüre etwas von dem Vergnügen empfinden, das wir als Fans von Kriminalgeschichten sowohl bei der Konstruktion der Programme als auch bei den Experimenten mit zahlreichen Studierenden hatten.

Christina Klüver

Jürgen Klüver

Danksagung

Ich bedanke mich ganz herzlich bei Prof. Dr. Jürgen Klüver und PD Dr. Christina Klüver für die guten Ratschläge und die Unterstützung, die ich beim Schreiben dieses Buches erfahren habe.

Dank gilt auch dem Fischer Taschenbuch Verlag, mit dessen Erlaubnis die Geschichte "Die spanische Truhe" von Agatha Christie in diesem Buch abgedruckt werden durfte. Durch die Erlaubnis, diese Geschichte hier wiederzugeben, wird überhaupt erst ermöglicht, dass der Leser an einem in diesem Buch beschriebenen Experiment teilnehmen kann.

1 Künstliche neuronale Netze

1.1 Einleitung

Künstliche neuronale Netze gehören zum Gebiet der sog. Computational Intelligence und sind somit ein Teilgebiet der künstlichen Intelligenz. Es handelt sich dabei um selbstlernende Systeme, die u. a. überall dort eingesetzt werden können, wo ein System auf der Basis von Eingaben (Messwerte, historische Daten, etc.) fehlertolerante und intelligente Entscheidungen treffen soll. Neben diesen praktischen Anwendungen können künstliche neuronale Netze aber auch verwendet werden, um die Funktionen des Gehirns und die Verstehensprozesse des Menschen zu modellieren.

In der Informatik und den Ingenieurwissenschaften werden künstliche neuronale Netze vorwiegend zur Informationsverarbeitung verwendet. Die möglichen Aufgaben reichen hier von der fehlertoleranten Wiedererkennung von Mustern (wie z. B. Gesichtern auf einem digitalen Bild) über Diagnosen und das Finden von optimalen Lösungen (z. B. das Optimieren von Arbeitsplänen oder Reisewegen) bis hin zu Prognosen (Aktienkurse, Lastenprognose in Versorgungsnetzen, etc.).

Die Modellierung der Funktionen des Gehirns oder der Mechanismen der Nervensysteme des Menschen fällt in das Fachgebiet der Neurobiologie bzw. Psychologie. Hier werden künstliche neuronale Netze aufgrund ihrer Ähnlichkeit zu den biologischen Vertretern verwendet. Wie in späteren Kapiteln gezeigt wird, können künstliche neuronale Netze auch verwendet werden, um kognitive Leistungen des Menschen abzubilden und zu untersuchen.

Um die in diesem Buch vorgestellten Modelle besser verstehen zu können, ist es zunächst wichtig, die künstlichen neuronalen Netze als solche zu verstehen. Da es sich dabei um mathematische Modelle handelt, mit denen der eine oder andere Leser bisher noch keine Berührung hatte, wird dieses Kapitel auf einfache Weise erläutern, was künstliche neuronale Netze überhaupt sind und wie sie funktionieren.

Es sei an dieser Stelle darauf hingewiesen, dass es nicht darum geht, nach der Lektüre dieses Kapitels bereits in der Lage zu sein, eigene Modelle zu entwerfen oder gar zu programmieren. Es geht nur darum, zu verstehen, wie

ein künstliches neuronales Netz funktioniert, damit die nachfolgenden Kapitel auch für Leser, die bisher keine praktischen Erfahrungen mit künstlicher Intelligenz oder künstlichen neuronalen Netzen gemacht haben, einfach verständlich sind.

Die in diesem Kapitel vorgestellten Modelle sind dabei bei weitem nicht vollständig. Es werden hier nur die Modelle beschrieben, die auch für die weiteren Kapitel von Belang sind. Würde man alle bisher entworfenen und verwendeten Modelle künstlicher neuronaler Netze beschreiben wollen, so könnte man damit ein eigenes Buch füllen (ein dickes Buch!).

Um die Einführung in die Welt der künstlichen neuronalen Netze abzuschließen, wird in diesem Kapitel auch erläutert, wie künstliche neuronale Netze trainiert werden und warum es "selbstlernend" heißt. Für das Training eines neuronalen Netzes wird i. d. R. eine sog. Lernregel verwendet. In den weiteren Kapiteln wird aber auch gezeigt, dass der Trainingsvorgang eines Netzes auch durchaus von Menschen übernommen werden kann, um auf diese Weise das, was der Mensch gelernt hat, in einem mathematischen Modell abzubilden.

Da es sich bei diesem Kapitel um eine erste Einführung in die künstlichen neuronalen Netze handelt, möchte ich an diesem Punkt Lesern, die bereits mit solchen Systemen gearbeitet haben und mit ihnen vertraut sind, empfehlen, das erste Kapitel zu überspringen und mit Kapitel zwei fortzufahren.

1.2 Biologisches Vorbild

Künstliche neuronale Netze bestehen aus künstlichen Neuronen die untereinander verbunden sind. Der Aufbau und die Arbeitsweise eines künstlichen Neurons basiert dabei auf dem biologischen Neuron (oder auch Nervenzelle) des Gehirns. Analog basieren auch die Mechanismen eines künstlichen neuronalen Netzes auf den Funktionen des menschlichen Gehirns.

Ein biologisches Neuron besteht aus den sog. Dendriten, einer Art Empfänger für die Signale anderer Neuronen, dem Soma (oder auch Zellkern), dem Axon, das ein Signal weiterleitet, sowie den Synapsen, die als Sender des Neurons verstanden werden können. (vgl. Kruse et al., 2011) Eine vereinfachte Skizze ist in Abb. 1 dargestellt.

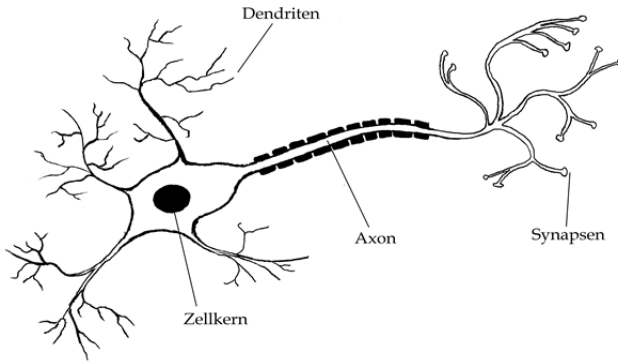


Abbildung 1: Aufbau eines Neurons

Ein Neuron erhält Signale über die Synapsen anderer Neuronen, die mit den Dendriten des empfangenden Neurons verbunden sind. Diese Signale können sich verstärkend oder hemmend auf die Aktivität des Zellkerns auswirken. Wird eine bestimmte Aktivität des Zellkerns erreicht oder überschritten, so wird ein Signal über das Axon des Neurons zu den Synapsen geleitet. Die Synapsen haben wiederum mit anderen Neuronen Verbindungen (zu deren Dendriten) und senden ein verstärkendes oder hemmendes Signal an diese.

Diese Eigenschaften alleine erklären aber noch nicht die Lernfähigkeit des menschlichen Gehirns. Wären sämtliche Verbindungen aller Neuronen untereinander bereits von Geburt an unveränderlich gegeben, so wäre auch jegliche Intelligenz bereits vorhanden und es könnte nichts Neues gelernt werden. Es muss also eine Möglichkeit geben, die Verbindungen der Neuronen untereinander zu verändern. In diesem Zusammenhang hat Donald Hebb 1949 eine Lernregel postuliert, die wie folgt lautet:

"When an Axon of a cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth or metabolic change takes place in one or both cells, so that A's efficiency as one of the cells firing B, is increased." (Hebb 1949, 50)

Frei übersetzt bedeutet das Folgendes: Ein Neuron B, das immer wieder Signale an ein Neuron A sendet, erfährt eine Veränderung der Verbindung zu Neuron A. Diese Veränderung bewirkt, dass die Signale, die von B gesendet werden, von A effizienter aufgenommen werden. Erhält ein Neuron A von einem Neuron B also öfter Signale, als von anderen Neuronen, so werden die Signale von Neuron A auch immer stärker aufgenommen. Man könnte auch sagen, sie werden stärker gewichtet, als die Signale von Neuronen, die weniger häufig senden.

Mit dieser Überlegung formulierte Donald Hebb bereits eine erste Lernregel, die für künstliche neuronale Netze verwendet wird, die sog. Hebb-Lernregel.

Um ein künstliches neuronales Netz zu erstellen, ist es nun zunächst nötig, ein formales Modell des Neurons zu haben, um künstliche Neuronen erzeugen zu können. Aus diesen kann dann später das Netz erstellt werden. Dieses formale Modell wird im folgenden Unterkapitel vorgestellt.

1.3 Formales Modell des Neurons

Nach dem Modell von McCulloch und Pitts (1943) ist ein Neuron eine Art Addierer, welches die über seine Dendriten empfangenen Signale anderer Neuronen gewichtet aufsummiert. (vgl. McCulloch, Pitts, 1943) Dabei hat jedes eingehende Signal eine eigene Gewichtung. Die Summe dieser Signale ist dann ausschlaggebend dafür, ob das Neuron selber ein Signal "feuert", oder ob es nichts sendet.

Das Neuron hat also eine Anzahl an Eingängen, über welche die Signale anderer Neuronen aufgenommen werden. Jedes dieser Signale hat eine Gewichtung, die angibt, wie effizient das Signal über diesen Eingang aufgenommen wird. Alle diese gewichteten Signale werden im Zellkern aufsummiert. Dieses gewichtete Aufsummieren wird als sog. Eingabefunktion des Neurons bezeichnet.

Um eine Aktivierung des Zellkerns in Form einer Zahl ausdrücken zu können, wird eine Aktivierungsfunktion benötigt, die aus den aufsummierten Eingaben des Neurons den Aktivierungswert des Zellkerns berechnet.

Über den Aktivierungswert wird anschließend bestimmt, ob das Neuron ein Signal sendet, oder nicht. Dazu wird eine Ausgabefunktion verwendet,

welche den Aktivierungswert erhält, um daraus die Ausgabe des Neurons zu berechnen.

Das künstliche Neuron besteht also aus drei Funktionen, der Eingabefunktion, der Aktivierungsfunktion und der Ausgabefunktion. Als Eingabefunktion wird im Allgemeinen immer die gewichtete Aufsummierung verwendet, während für die Aktivierungsfunktion und die Ausgabefunktion beliebige mathematische Funktionen verwendet werden können. Ein entsprechendes Modell ist in Abbildung 2 dargestellt.

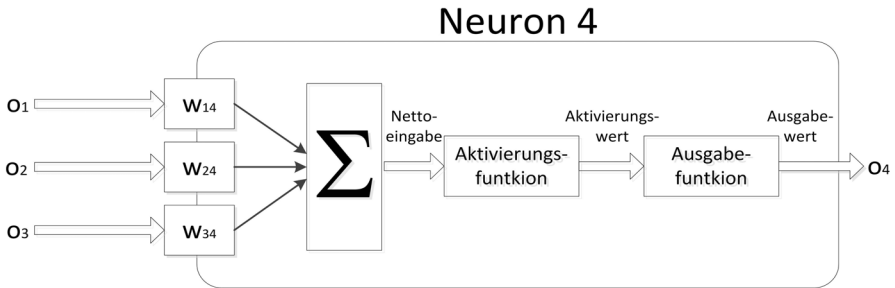


Abbildung 2: Formale Darstellung eines Neurons

In der Abbildung stellen die verschiedenen w_{x4} die Gewichtungen der eingehenden Signale dar. Die vier gibt ab, dass es sich beim empfangenden Neuron um Neuron 4 handelt, während das x angibt, welches das sendende Neuron ist. Das Gewicht w_{14} ist also das Gewicht, mit dem das Signal von Neuron 1 gewichtet (multipliziert) wird, bevor es mit den anderen eingehenden Signalen aufsummiert wird.

Die mathematischen Funktionen, die das künstliche Neuron verwendet, können beliebig festgelegt werden. Die Eingabefunktion ist aber, wie bereits erwähnt, in der Regel die beschriebene Aufsummierung der gewichteten, eingehenden Signale, die formal wie folgt aussieht:

$$net_j = \sum w_{ij} * o_i$$