Neuronale Netze und Matrixmultiplikation

Mathematik, Gymnasium, Qualifikationsphase

Das menschliche Nervensystem hat die Aufgabe, Reize aus der äußeren Umwelt und aus dem Körperinneren als Signale aufzunehmen, aufeinander zu beziehen und mit früheren zu vergleichen. Es besteht aus einer Vielzahl an einzelnen Nervenzellen, den sogenannten Neuronen, die zu einem neuronalen Netz verbunden sind.
Diesem biologischen Vorbild folgend, wurden *künstliche neuronale Netze* entwickelt, die Lernprozesse imitieren und so in vielfältigen Technologien durch maschinelles Lernen und weitere Aspekte der künstlichen Intelligenz Anwendung finden.

Ein künstliches neuronales Netzwerk besteht aus einzelnen künstlichen Neuronen, die auch als Knoten bezeichnet werden. Die Knoten sind miteinander verbunden und beeinflussen sich auf diese Weise wechselseitig.

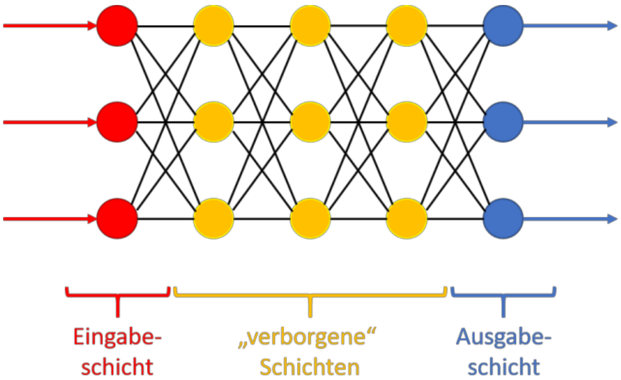

Abb. 1: Veranschaulichung eines künstlichen neuronalen Netzwerks. Die Knoten sind entlang der einzelnen Schichten jeweils miteinander verbunden. (in Anl. an Dontsov, 2021)

Die erste Schicht des neuronalen Netzes erhält äußere Informationen[1]. Im künstlichen neuronalen Netz sind dies Parameter in Form von reellen Zahlen, die dann in den Knoten weiterverarbeitet werden. Das vereinfachte Modell in Abb. 1 ist dabei von links nach rechts zu verstehen. Von links kommen die Eingangssignale, die dann in der ersten Schicht verarbeitet werden und ein Ausgangssignal erzeugen. Das Ausgangssignal eines solchen Knotens hat dabei wiederum die Funktion des Eingangssignals in einem Knoten der darauffolgenden Schicht. Allerdings gibt es in der sog. „verborgenen" Schicht nicht nur ein einzelnes Eingangssignal. Stattdessen erhält es Signale von allen Knoten aus der jeweils davorliegenden Schicht. Je größer das neuronale Netz, desto mehr Daten müssen verarbeitet werden[2].

[1] Im menschlichen Nervensystem zum Beispiel durch Sinnesreize von Augen oder Ohren.
[2] In Abb. 1 ist der Übersichtlichkeit halber ein recht kleines neuronales Netz dargestellt, in dem pro Schicht lediglich drei Knoten liegen. Für künstliche Intelligenz genutzte neuronale Netze haben eine weitaus höhere Anzahl an Knoten von teilweise mehreren Milliarden.

Die eingehenden Signale in Form von reellen Zahlen werden dabei zunächst addiert[3] und ergeben in Summe das neue auszugebende Signal: $b_1 = a_1 + a_2 + a_3$ (vgl. Abb. 2). Allerdings würde sich auf diese Weise die gleiche Summe auch für b_2 und b_3 ergeben, was eine Unterscheidung der Knoten überflüssig werden ließe. Daher werden nicht alle Eingangssignale gleich berücksichtigt – auch im menschlichen Nervensystem werden die Signale unterschiedlich stark weitergeleitet. Im künstlichen neuronalen Netz wird dies mithilfe von sogenannten Gewichtungsfaktoren berücksichtigt: Je stärker die Verbindung zwischen den Neuronen, desto größer die Gewichtungsfaktoren. Die eingehenden Parameter a_1 , a_2 bzw. a_3 werden jeweils noch mit einem solchen Gewichtungsfaktor $w_{1,1}$, $w_{1,2}$ bzw. $w_{1,3}$ multipliziert. Auf diese Weise kommt es zur Formel $b_1 = w_{1,1} \cdot a_1 + w_{1,2} \cdot a_2 + w_{1,3} \cdot a_3$, die für jeden Knoten in jeder Schicht ausgerechnet werden muss[4].

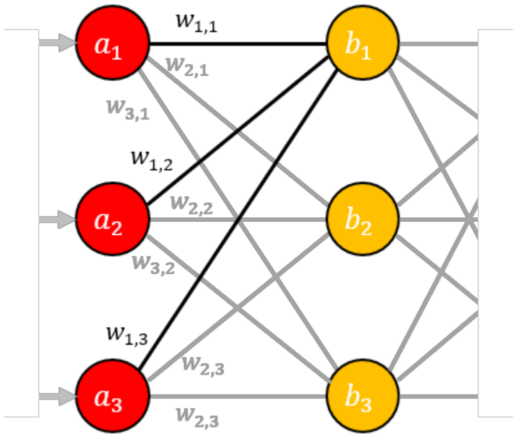

Abb. 2: Detaillierte Darstellung der Verknüpfungen eines künstlichen neuronalen Netzwerks. Die Werte (a_i) der Eingangsschicht werden unterschiedlich gewichtet ($w_{i,j}$), um die Werte der nächsten Schicht (b_i) zu berechnen. (gemäß Rashid, 2017)

Aufgabe 1: Bestimmen Sie b_1 für folgende Werte:

- $w_{1,1} = 0{,}8;\ w_{1,2} = 0{,}1;\ w_{1,3} = 0{,}3$
- $a_1 = 2;\ a_2 = 3;\ a_3 = 4$

Aufgabe 2: Geben Sie die Formeln für b_2 und b_3 an.

In diesem kleinen Beispiel mag das Erstellen der Formeln und deren Berechnung vielleicht noch eine schaffbare Aufgabe sein, doch mit einer wachsenden Zahl an Knoten und Schichten (teilweise mehrere Milliarden innerhalb eines neuronalen Netzes) wird es immer komplizierter. Dies führt irgendwann zu Rechnungen, die kaum noch fehlerfrei zu bewältigen sind – oder doch?

[3] In einem weiteren sehr wichtigen Schritt durchlaufen die resultierenden Summen anschließend eine sog. Aktivierungsfunktion, die im Rahmen dieses Arbeitsblattes nicht weiter thematisiert wird.
[4] Im neuronalen Netzwerk kommt in dieser Formel auch noch ein zusätzlicher sog. Bias-Term vor, der im Rahmen dieses Arbeitsblatt nicht weiter thematisiert wird.

Computer sind äußerst effiziente Rechenmaschinen, die fest vorgegebene Rechenstrukturen und wiederkehrende Rechenprinzipien schnell und fehlerfrei berechnen können. Ein weiterer Vorteil ist, dass sich die hier vorgenommenen Rechnungen über einen einzigen Rechenschritt zusammenfassen lassen, der von fast allen Computersprachen bereits beherrscht wird, ohne dass eine neue Rechenvorschrift vorgegeben werden muss: Die **Matrixmultiplikation**.

Die von der ersten Schicht a ausgehenden Werte lassen sich als Vektor zusammenfassen, wobei $\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$. Die Gewichtungsfaktoren hingegen lassen sich als Matrix darstellen:

$$\overrightarrow{W_a} = \begin{pmatrix} w_{1,1} & w_{1,2} & w_{1,2} \\ w_{2,1} & w_{2,2} & w_{2,3} \\ w_{3,1} & w_{3,2} & w_{3,3} \end{pmatrix}$$

Multipliziert man eine Matrix mit einem Vektor ($\overrightarrow{W_a} \cdot \vec{a}$), so wird jeweils eine Zeile der Matrix mit den Einträgen des Vektors multipliziert und die Ergebnisse addiert. Es ergibt sich also der folgende Ergebnisvektor:

$$\vec{b} = \begin{pmatrix} w_{1,1} \cdot a_1 + w_{1,2} \cdot a_2 + w_{1,3} \cdot a_3 \\ w_{2,1} \cdot a_1 + w_{2,2} \cdot a_2 + w_{2,3} \cdot a_3 \\ w_{3,1} \cdot a_1 + w_{3,2} \cdot a_2 + w_{3,3} \cdot a_3 \end{pmatrix}$$

Aufgabe 3: Stellen Sie einen Bezug zwischen den Zeilen des Vektors $\vec{b}$ und Ihren Ergebnissen aus der vorherigen Aufgabe her.

Die jeweiligen Zeilen des Vektors $\vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ entsprechen also exakt den gesuchten Knotensummen der Schicht b (vgl. Formeln für b_1, b_2 und b_3).

Aufgabe 4: Berechnen Sie die einzelnen Einträge des Vektors $\vec{b}$ mit folgender Gewichtungsmatrix $\overrightarrow{W_a} = \begin{pmatrix} 0{,}8 & 0{,}1 & 0{,}3 \\ 0{,}7 & 0{,}4 & 0{,}7 \\ 0{,}2 & 0{,}9 & 0{,}4 \end{pmatrix}$ und dem Eingangssignal $\vec{a} = \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}$.

An dieser Stelle kann die Rechenarbeit also getrost dem Computer überlassen werden, der die für ihn simple Matrixmultiplikation $\vec{b} = \overrightarrow{W_a} \cdot \vec{a}$ berechnet. Auch die nächste Schicht kann nun genauso berechnet werden mit neuen Gewichtungsfaktoren, also $\vec{c} = \overrightarrow{W_b} \cdot \vec{b}$, $\vec{d} = \overrightarrow{W_c} \cdot \vec{c}$, usw.

Aufgabe 5: Informieren Sie sich, wie Sie im GTR oder CAS Matrizen und/oder Vektoren eingeben und multiplizieren. Kontrollieren Sie Ihre Ergebnisse für $\vec{b}$.

Die Gewichtungsfaktoren sowie auch die Eingangssignale sind an dieser Stelle willkürlich gewählt worden. Allerdings ist nun deren Bestimmung genau der Wesenszug der künstlichen neuronalen Netze. Durch sogenannte Trainingsbeispiele, bei denen das

Endergebnis bereits bekannt ist, werden die Gewichtungsfaktoren durch Ausprobieren und Optimieren bestimmt. Dieser Schritt verwendet seinerseits wieder Matrixmultiplikation sowie lokale Extremstellen zur Minimierung von Abweichungen zum Endergebnis.

Zur weiteren Übung und zur Motivation, sich auf den GTR bzw. CAS zu verlassen, können nun weitere Schichten des neuronalen Netzes konstruiert werden. Die Eingabe in den Computer mag zwar ebenfalls als großer Aufwand wahrgenommen werden, dies lässt sich aber mit etwas Programmierarbeit automatisieren, sodass die Vorteile der Computergestützten Rechnung noch mehr überwiegen.

Tempoaufgabe: Bestimmen Sie die weiteren Schichtvektoren $\vec{c}$ und $\vec{d}$ mithilfe der Gewichtungsmatrizen $\overrightarrow{W_b} = \begin{pmatrix} 0{,}9 & 0{,}3 & 0{,}4 \\ 0{,}2 & 0{,}8 & 0{,}2 \\ 0{,}1 & 0{,}5 & 0{,}6 \end{pmatrix}$ und $\overrightarrow{W_c} = \begin{pmatrix} 0{,}3 & 0{,}7 & 0{,}5 \\ 0{,}6 & 0{,}5 & -0{,}2 \\ 0{,}8 & -0{,}1 & 0{,}9 \end{pmatrix}$.