KI 2: Neuronale Netze

Für ein neuronales Netz mit den Schichten $k = 0, 1, 2, ..., N$ legen wir folgende Bezeichnungen fest:

- Die k-te Schicht besitzt n_k viele Neuronen.
- Die Neuronen der k-ten Schicht werden mit $a_1^{(k)}$, $a_2^{(k)}$, ..., $a_{n_k}^{(k)}$ bezeichnet. Zusammengefasst können wir sie als Vektor darstellen:

$$a^{(k)} = \begin{pmatrix} a_1^{(k)} \\ \vdots \\ a_{n_k}^{(k)} \end{pmatrix}$$

- Die **Biasse** der Neuronen der k-ten Schicht ($k \geq 1$) werden mit $b_1^{(k)}$, $b_2^{(k)}$, ..., $b_{n_k}^{(k)}$ bezeichnet. Zusammengefasst können wir auch sie als Vektor darstellen:

$$b^{(k)} = \begin{pmatrix} b_1^{(k)} \\ \vdots \\ b_{n_k}^{(k)} \end{pmatrix}$$

- Die Verbindung zwischen dem Neuron $a_i^{(k)}$ und dem Neuron $a_j^{(k-1)}$ aus der Vorgänger-Schicht bezeichnen wir mit $w_{i,j}^{(k)}$. Zusammengefasst können wir alle Verbindungen der k-ten Schicht zu ihrer Vorgänger-Schicht als Matrix darstellen:

$$W^{(k)} = \begin{pmatrix} w_{1,1}^{(k)} & w_{1,2}^{(k)} & \dots & w_{1,n_{k-1}}^{(k)} \\ w_{2,1}^{(k)} & w_{2,2}^{(k)} & \dots & w_{2,n_{k-1}}^{(k)} \\ \dots & \dots & \dots & \dots \\ w_{n_k,1}^{(k)} & w_{n_k,2}^{(k)} & \dots & w_{n_k,n_{k-1}}^{(k)} \end{pmatrix}$$

  Diese Matrix hat so viele Zeilen, wie die k-te Schicht Neuronen besitzt ($= n_k$) und so viele Spalten wie die Vorgänger-Schicht Neuronen besitzt ($= n_{k-1}$).
  Die Zahlen $w_{i,j}^{(k)}$ werden auch als **Gewichte** des Netzes bezeichnet.

- Die **Sigmoid**-Funktion ist definiert als

$$\sigma(x) = \frac{1}{1+e^{-x}}$$

- Wenn wir die Sigmoid-Funktion auf einen Vektor anwenden, soll dies bedeuten, dass die Funktion auf jede einzelne Komponente angewendet wird:

$$\sigma(\vec{x}) = \begin{pmatrix} \sigma(x_1) \\ \vdots \\ \sigma(x_n) \end{pmatrix}$$

Mit diesen Bezeichnungen kann man formulieren, wie die Neuronen der k-ten Schicht aus denen der Vorgänger-Schicht bestimmt werden können:

$$a^{(k)} = \sigma\left(W^{(k)} \cdot a^{(k-1)} + b^{(k)}\right)$$

Dabei bezeichnet $W^{(k)} \cdot a^{(k-1)}$ das Produkt zwischen einer Matrix und einem Vektor.

1 In dieser Aufgabe untersuchen wir die Sigmoid-Funktion $\sigma(x) = \frac{1}{1+e^{-x}}$.

(a) Bestimmen Sie die Ableitungen $\sigma'(x)$ und $\sigma''(x)$.
[Zur Kontrolle: $\sigma''(x) = \frac{e^{-x} \cdot (e^{-x}-1)}{(1+e^{-x})^3}$

(b) Untersuchen Sie σ auf Extrema und Wendepunkte.

(c) Begründen Sie, dass $\lim\limits_{x \to -\infty} \sigma(x) = 0$ und $\lim\limits_{x \to \infty} \sigma(x) = 1$.

(d) Skizzieren Sie den Verlauf des Graphen in ein geeignetes Koordinatensystem.

2 Gegeben ist ein neuronales Netz mit folgenden Gewichten und Biassen:

$$W^{(1)} = \begin{pmatrix} 1 & 0 \\ -1 & 2 \\ 2 & 3 \end{pmatrix} \qquad W^{(2)} = \begin{pmatrix} 0 & 1 & 2 \\ -2 & 3 & 4 \end{pmatrix} \qquad b^{(1)} = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} \qquad b^{(2)} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

(a) Zeichnen Sie das neuronale Netz.

(b) Geben Sie die Werte der folgenden Variablen an: N, n_0, n_1, n_2, $w^{(2)}_{1,2}$ und $b^{(1)}_2$.

(c) Bestimmen Sie die Ausgabe des Netzes für die Eingabe $\begin{pmatrix} 5 \\ 2 \end{pmatrix}$.

3 Die Klassen `NeuronalesNetz` und `Schicht` wurden im Unterricht modelliert.

(a) Implementieren Sie diese beiden Klassen unter Beachtung dieser Modellierung.

(b) In der Hauptklasse soll in der `onStart`-Methode das neuronale Netz aus Aufgabe 2 erzeugt, mit der dort genannten Eingabe ausgeführt und das Ergebnis ausgegeben werden.
Implementieren Sie die `onStart`-Methode gemäß dieser Anforderungen und überprüfen Sie Ihr Ergebnis aus Aufgabe 2 (c).

4 Gegeben ist ein neuronales Netz mit zwei Schichten, die jeweils ein Neuron enthalten.

(a) Geben Sie mögliche Werte für die Gewichte und Biasse an. Berechnen Sie die Ausgabe es Netzes für eine selbstgewählte Eingabe.

(b) Beurteilen Sie die Aussage »Bei dieser einfachsten Form eines neuronalen Netzes handelt es sich im Wesentlichen um eine verzerrte lineare Funktion.«