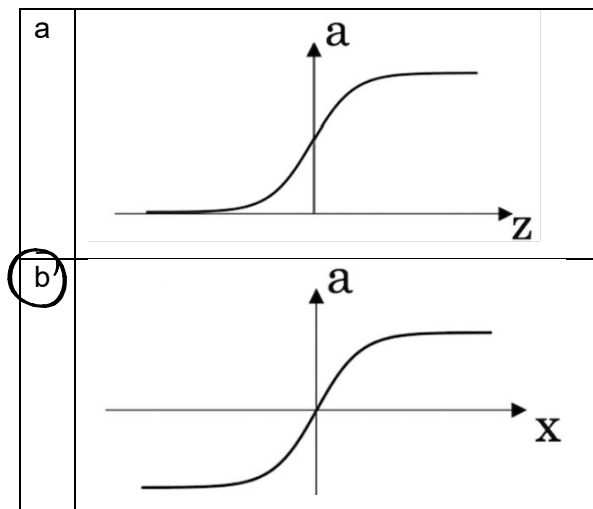


Neuronale Netze Quiz

Frage 1: Aktivierungsfunktionen

- a) Welche der beiden Graphen stellt die $\tanh()$ Aktivierungsfunktion dar?



$$= 1 - \tanh(x)^2$$

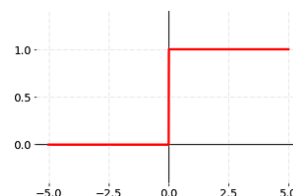
$$= \frac{(e^x + e^{-x})^2 - (e^x - e^{-x})^2}{(e^x + e^{-x})^2}$$

- b) Bestimmen Sie die Ableitung der Funktion

$$\varphi(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} = \frac{(e^x + e^{-x})(e^x + e^{-x}) - (e^x - e^{-x})(e^x - e^{-x})}{(e^x + e^{-x})^2}$$

Hinweis: Die Ableitung der \tanh -Funktion kann (wie beim logistischen Sigmoid) ebenfalls als Funktion von $\varphi(x)$ ausgedrückt werden.

- c) Warum eignet sich die Einheitsschrittfunktion nicht als Aktivierungsfunktion, wenn man das neuronale Netzwerk mit Gradient Descent (Backpropagation) trainieren möchte?



Frage 2: Künstliches Neuron

- a) Ein künstliches Neuron berechnet seinen Output (auch Aktivierung genannt) gemäss

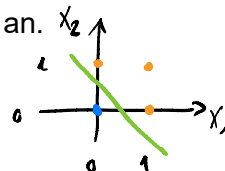
$$a = \varphi(\mathbf{w}^T \mathbf{x} + \mathbf{b}),$$

wobei \mathbf{w} , \mathbf{x} und \mathbf{b} Spaltenvektoren sind. Angenommen φ sei die Einheitsschrittfunktion. Sind die folgenden Aussagen wahr oder falsch.

- Ein einzelnes künstliches Neuron kann die **NOR**-Funktion lernen. Falls ja, geben ein mögliches Neuron (Gewichte und Bias) dafür an.

$$b = 1$$

$$w_1 = -2 \quad w_2 = -2$$

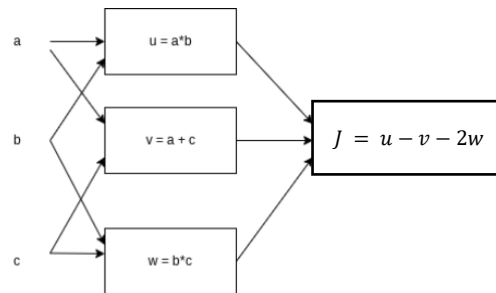


x_1	x_2	\checkmark	a
0	0	1	1
0	1	-1	0
1	0	-1	0
1	1	-3	0

- Ein einzelnes künstliches Neuron kann die **XOR**-Funktion lernen. Begründen Sie ihre Antwort.
- Ein Neuronales Netz mit einem hidden Layer (also mit insgesamt 3 Layern: Input, Hidden und Output Layer) kann die XOR-Funktion lernen.

Frage 3: Computational Graphs

Es sei der folgende Computational Graph gegeben.



- a) Berechnen Sie den Wert von J für $a = 2, b = 2$ und $c = 3$ durch Forward Propagation durch den Graphen.

$$J = 2 \cdot 2 - (2 + 3) - 2(2 \cdot 3) = 4 - 5 - 12 = \underline{\underline{-13}}$$

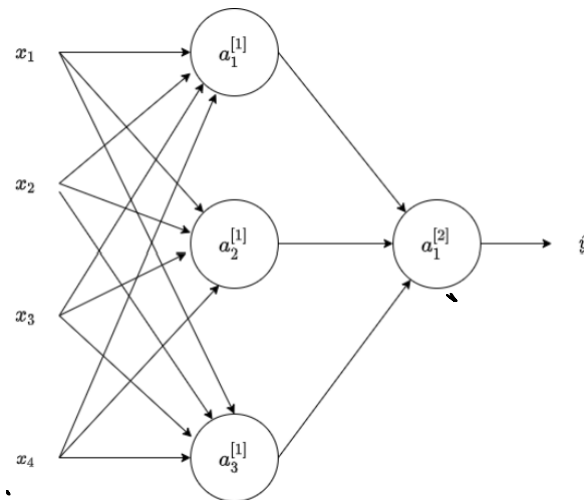
- b) Bestimmen Sie die Ableitung $\frac{\partial J}{\partial c}$ durch Backpropagation.

$$\frac{\partial J}{\partial v} = -1 \quad \frac{\partial J}{\partial w} = -2 \quad \frac{\partial v}{\partial c} = 1 \quad \frac{\partial w}{\partial c} = b \quad \frac{\partial J}{\partial c} = \frac{\partial J}{\partial w} \frac{\partial w}{\partial c} + \frac{\partial J}{\partial v} \frac{\partial v}{\partial c} = -2b - 1 \cdot 1 = -2b - 1 = \underline{\underline{-5}}$$

Frage 4: Binäre Klassifikation

- a) Angenommen Sie entwerfen einen binären Klassifikator, um anhand eines 2s langen Audiosamples zu erkennen, ob eine Person «Hey ZHAW» gesagt hat oder nicht. Welche Aktivierungsfunktion würden Sie für den Output Layer des Neuronalen Netzes einsetzen

- ReLU
- Sigmoid
- Leaky ReLU
- Linear

Frage 5: Dimensionen von Neuronalen Netzen

a) Aus wie vielen Layer L besteht dieses Netzwerk?

3

b) Bestimmen Sie die Dimension des Inputvektors x , der Aktivierungen $a^{[1]}, a^{[2]}$, des Outputs \hat{y} und der Gewichtsmatrizen $W^{[1]}, W^{[2]}$, deren Zeilen die Gewichtsvektoren der einzelnen Neuronen enthalten.

$$x \in \mathbb{R}^4$$

$$W^{[1]} \in \mathbb{R}^{3 \times 4}$$

$$a^{[1]} \in \mathbb{R}^3$$

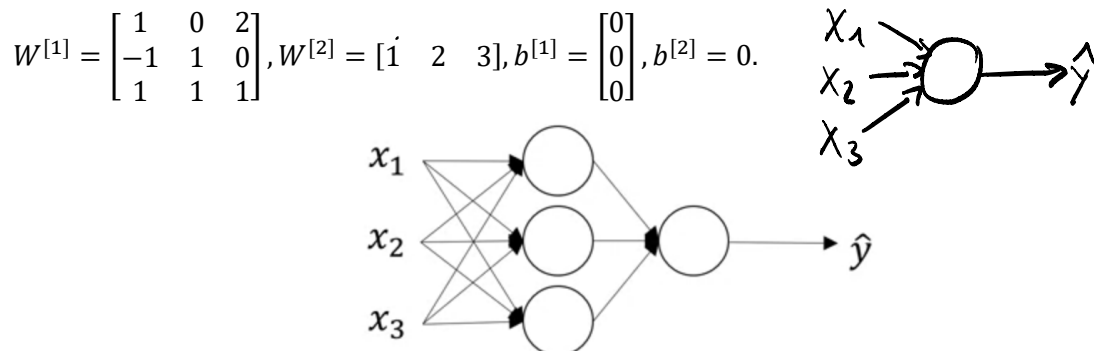
$$W^{[2]} \in \mathbb{R}^{1 \times 3}$$

$$a^{[2]} \in \mathbb{R}$$

$$\hat{y} \in \mathbb{R}$$

Frage 6: Dimensionen von neuronalen Netzen

Angenommen das unten dargestellte neuronale Netz besitzt nur lineare Aktivierungsfunktionen. Die Gewichtsmatrizen und Biase seien:



- a) Bestimmen Sie ein einstufiges neuronales Netz, welches das gleiche Input-Output Mapping (also die gleiche Funktion) darstellt. Geben Sie die Gewichtsmatrix und den Bias an.

$$W = W^{[2]} W^{[1]} = \begin{bmatrix} 2 & 5 & 5 \end{bmatrix} \quad b = 0$$

Frage 7: Convolutional Neural Networks

- Angenommen der Input sei ein 128×128 RGB-Bild. Dies entspricht einem Input der Dimension $(128, 128, 3)$. Des Weiteren soll ein fully-connected Feedforward Neuronales Netz (kein Convolutional Neuronales Netz) benutzt werden. Das erste hidden Layer besteht aus 64 Neuronen und jedes Neuron davon ist mit allen Inputs verbunden. Wie viele Parameter hat das dieses hidden Layer (inklusive Biase)?
- Nun wird derselbe Input mit einem CNN verarbeitet. Der erste Layer besteht aus Filtern der Dimension $(3, 3, 3)$ und im ersten Layer werden 16 davon eingesetzt. Wie viele Parameter hat dieses hidden Layer (inklusive Biase)?