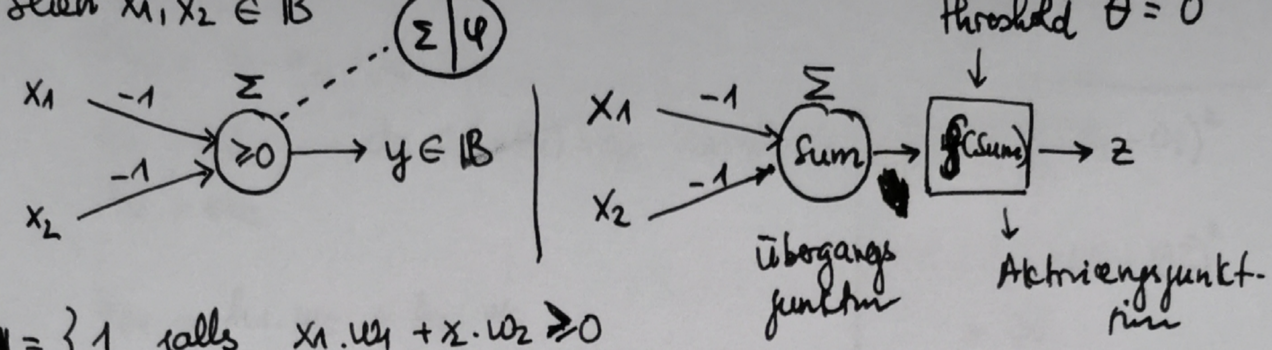
ML 10:

A1 Seien $x_1, x_2 \in \mathbb{B}$

1)



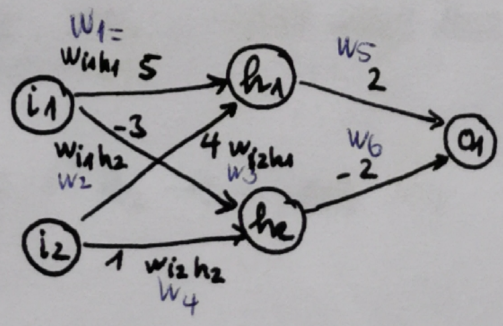
$$f(\text{sum}) = \begin{cases} 1 & \text{falls } x_1 \cdot w_1 + x_2 \cdot w_2 \geq 0 \\ 0 & \text{sonst.} \end{cases}$$

2) Eigenschaften von McCulloch-Pitts-Zelle / Neuron := Schwellenwertneuron

- besitzt nur 2 Zustände (binary)
- Es empfängt Eingaben von anderen Neuronen über Verbindungen, die einen Gewichtungswert aufweisen
- Es hat eine Aktivierungsfkt (:= Schwellenwertfunktion), die die gewichtete Summe seiner Eingänge mit einem $\underbrace{\text{Schwellenwert}}_{\text{(threshold)}}$ vergleicht. (~~threadhold~~)

⊛ inhibitory input := its contribution to weighted sum is negative

A2 Ausgangsfunktion: $\varphi(x) = x$ Identity



a) $i_1 = 5$, $i_2 = 2$

$$\overset{x}{h_1} = 5 \cdot 5 + 4 \cdot 2 = 33$$

$$\overset{x}{h_2} = -3 \cdot 5 + 1 \cdot 2 = -13$$

$\downarrow$ $(y = z)$

$$o_1' = 33 \cdot 2 + -13 \cdot -2$$

$$= 92 \xrightarrow{y=z} \boxed{o_1 = 92}$$

b) $MSE = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 = (100 - 92)^2 = 64$

c) $E = (o_1 - t_1)^2 \Rightarrow \frac{dE}{do_1} = (o_1 - t_1)$

Backpropagate:

<u>gegeben</u>:
$$z_{h1} = i_1 \cdot w_1 + i_2 \cdot w_3$$
$$z_{h2} = i_1 \cdot w_2 + i_2 \cdot w_4$$
$$h_1 = z_{h1} \quad \text{, da } f(z_{h1}) = h_1 \text{ Identität}$$
$$h_2 = z_{h2}$$
$$z_{o1} = h_1 \cdot w_5 + h_2 \cdot w_6$$
$$o_1 = z_{o1} \quad \text{, da Identität}$$

$$\sigma(x) = x$$
$$\Rightarrow \frac{d\sigma(x)}{dx} = 1$$

$$E = \frac{1}{2} \sum_{i=1}^{n} (t_i - o_i)^2$$
$$\hookrightarrow E = \frac{1}{2} (100 - 92)^2 = 32$$

① Für Gewichte in Output-Layer (w_5, w_6)

Ableitungen:
$$\frac{d_{o1}}{d_{z_{o1}}} = 1 \; ; \; \frac{d_E}{d_{o1}} = 2 \cdot \frac{1}{2} (t_1 - o_1) \cdot -1 = o_1 - t_1$$

⊕ Für w5:
Ableitung von Sum nach Weight. (Gewichte): $\frac{d_{z_{o1}}}{d_{w5}} = h_1$

$$\hookrightarrow \frac{d_E}{d_{w5}} = \frac{d_E}{d_{o1}} \cdot \frac{d_{o1}}{d_{z_{o1}}} \cdot \frac{d_{z_{o1}}}{d_{w5}}$$
$$= (o_1 - t_1) \cdot 1 \cdot h_1 = -8 \cdot 1 \cdot 33 = -264$$

$$\hookrightarrow new\_w_5 = w_5 - \alpha \cdot \frac{d_E}{d_{w5}} = 2 - 0{,}0005 \cdot -264 = 2{,}132$$
(3,32)

⊕ Für w6:
$$\frac{d_E}{d_{w6}} = \frac{d_E}{d_{o1}} \cdot \frac{d_{o1}}{d_{z_{o1}}} \cdot \frac{d_{z_{o1}}}{d_{w6}} = -8 \cdot 1 \cdot -13 = 104$$

$$\hookrightarrow new\_w_6 = w_6 - \alpha \cdot \frac{d_E}{d_{w6}} = -2 - 0{,}0005 \cdot (104) = -2{,}052$$
(−2,52)

⊛ Für Gewichte in Hidden-Layer (w_1, w_2, w_3, w_4)

⊕ Für w_1:

$$\frac{dE}{dw_1} = \frac{dE}{do_1} \cdot \frac{do_1}{dz_{o1}} \cdot \frac{dz_{o1}}{dh_1} \cdot \frac{dh_1}{dz_{h1}} \cdot \frac{dz_{h1}}{dw_1}$$

$$= (o_1 - t_1) \cdot 1 \cdot w_5 \cdot 1 \cdot i_1$$

$$= \underset{-8}{} \cdot 2 \cdot 5 = -80$$

↳ $\underline{\underline{new\text{-}w_1}} = w_1 - \alpha \cdot \frac{dE}{dw_1} = 5 - \underset{0{,}0005}{\cancel{0{,}05}} \cdot \overset{(-80)}{} = 5{,}04$

(5,4)

⊕ Für w_2:

$$\frac{dE}{dw_2} = \frac{dE}{do_1} \cdot \frac{do_1}{dz_{o1}} \cdot \underbrace{\frac{dz_{o1}}{dh_2}}_{(=w_6)} \cdot \frac{dh_2}{dz_{h2}} \cdot \underbrace{\frac{dz_{h2}}{dw_2}}_{(=i_1)}$$

$$= \underset{-8}{} \cdot 1 \cdot (-2) \cdot 1 \cdot 5 = 80$$

↳ $new\text{-}w_2 = w_2 - \alpha \frac{dE}{dw_2} = -3 - \underset{0{,}0005}{\cancel{0{,}05}} \cdot (\underset{80}{}) = -3{,}04$

(−3,4)

⊕ Für w_3

$\frac{dE}{dw_3} = \underset{-32}{} \Rightarrow new\text{-}w_3 = 4 - \underset{0{,}0005}{\cancel{0{,}05}} \cdot (\underset{-32}{}) = 4{,}016$

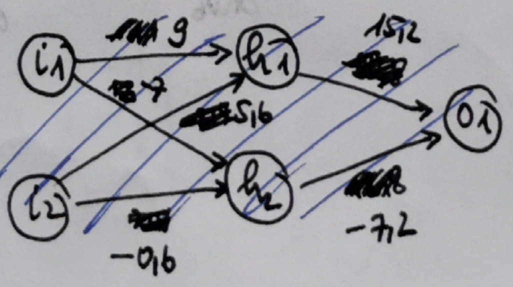
(4,16)

⊕ Für w_4

$\frac{dE}{dw_4} = \underset{32}{} \rightarrow new\text{-}w_4 = 1 - \underset{0{,}0005}{\cancel{0{,}05}} (\underset{32}{}) = 0{,}984$

(0,84)

Neuer graph:

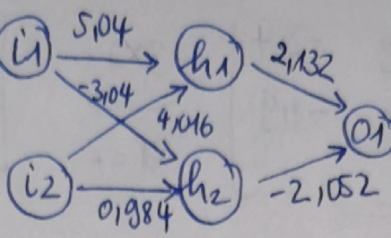


~~Neuer Ausgabe:~~

~~$h_1 = 56{,}2$~~ | ~~$o_1 =$~~

~~$h_2 = -36{,}2$~~

Neuer graph:



neu_h_1 = ~~29,216~~ 33,232 (~~20,2~~ 35,232)

neu_h_2 = ~~-14,216~~ -13,232 (-15,32)

neu_o_1 = 98,003

d) Das Ergebnis verschlechtert sich wenn α (Learnrate) = 0,005 bzw. 0,00001
Der Model wird überlernen bzw. unterlernen

e) $h_1, h_2, o_1 = 0$
⇒ es wird keine Verbesserung über Backpopagation gegeben!

$\boxed{A3}$

a) ohne padding $(5-3+1) = 3\times3$
i)

(VLF) gray-schwarz Bild.

$$\begin{bmatrix} 5 & 9 & 5 & 3 & 5 \\ 4 & 2 & 4 & 1 & 7 \\ 6 & 4 & 2 & 0 & 3 \\ 8 & 3 & 6 & 9 & 5 \\ 8 & 3 & 6 & 4 & 9 \end{bmatrix} * \begin{bmatrix} -1 & 1 & -1 \\ -1 & 1 & -1 \\ -1 & 1 & -1 \end{bmatrix} = \begin{bmatrix} -15/9 & -8/9 & -22/9 \\ -21/9 & -7/9 & -17/9 \\ -26/9 & -9/9 & -18/9 \end{bmatrix}_{3\times3}$$

$$= \begin{bmatrix} -1{,}7 & -0{,}9 & -2{,}4 \\ -2{,}3 & -0{,}7 & -1{,}9 \\ -2{,}9 & -1 & -2 \end{bmatrix}$$

with padding
ii)

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 5 & 9 & 5 & 3 & 5 & 0 \\ 0 & 4 & 2 & 4 & 1 & 7 & 0 \\ 0 & 6 & 4 & 2 & 0 & 3 & 0 \\ 0 & 8 & 3 & 6 & 9 & 5 & 0 \\ 0 & 8 & 3 & 6 & 4 & 9 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} * \begin{bmatrix} -1 & 1 & -1 \\ -1 & 1 & -1 \\ -1 & 1 & -1 \end{bmatrix} = \begin{bmatrix} -2/9 & -7/9 & -6/9 & -28/9 & -8/9 \\ 0 & -3/9 & -8/9 & -22/9 & 11/9 \\ 9/9 & -21/9 & -7/9 & -17/9 & 5/9 \\ 12/9 & -26/9 & -9/9 & -18/9 & 4/9 \\ 4/9 & -22/9 & -7/9 & -13/9 & 1/9 \end{bmatrix}$$

b) poolen ergebnis

i) $\begin{bmatrix} -1{,}7 & -0{,}9 & -2{,}4 \\ -4{,}3 & -0{,}7 & -1{,}9 \\ -2{,}9 & -1 & -2 \end{bmatrix} \xrightarrow[Sd=2]{3\times 3} \begin{bmatrix} -0{,}7 \end{bmatrix}$

ii) $\begin{bmatrix} -0{,}2 & -0{,}78 & -0{,}67 & -2{,}56 & -0{,}89 \\ 0 & -0{,}33 & -0{,}89 & -2{,}44 & 1{,}22 \\ 1 & -2{,}33 & -0{,}78 & -1{,}89 & 0{,}56 \\ 1{,}33 & -2{,}89 & -1 & -2 & 0{,}44 \\ 0{,}44 & -2{,}44 & -0{,}78 & -1{,}44 & 0{,}11 \end{bmatrix} \xrightarrow[Sd=2]{3\times 3} \begin{bmatrix} 1 & 0{,}56 \\ 1 & 0{,}56 \end{bmatrix}$