ENI-Hausi-07

Team AI

Aufgabe 1:

1a) $u_2 = w_2 \cdot y_1 + b_2 = 0{,}6 + 0{,}9 = 1{,}5$

$$f(u_2) = \frac{1}{1+e^{-1{,}5}} \approx 0{,}82$$

$$f'(u_2) = (1 - f(u_2))\,f(u_2) \approx 0{,}18 \cdot 0{,}82 \approx 0{,}148$$

$$\frac{\partial E}{\partial b_2} = -2(T - y_2) \cdot f'(u_2) = 2 f(u_2) f'(u_2)$$

$$\approx 2 \cdot 0{,}82 \cdot 0{,}148 \approx 0{,}244$$

---

b) $u_2 = 2 + 2 = 4$

$$f(u_2) = \frac{1}{1+e^{-4}} \approx 0{,}982$$

$$f'(u_2) = f(u_2)\,(1 - f(u_2)) \approx 0{,}018$$

$$\frac{\partial E}{\partial b_2} = 2 f(u_2) f'(u_2) \approx 2 \cdot 0{,}982 \cdot 0{,}018 \approx 0{,}035$$

---

c) Das Problem bei b) ist, dass die Startwerte unglücklich gewählt sind, so dass ~~man~~ in einem sehr flachen Areal in der Fehlerfunktion E startet.
Der Faktor, welcher hauptsächlich dafür verantwortlich ist, ist $f'(u_2) \approx 0{,}018$.

2a)
$$\frac{\partial E}{\partial b_1} = \frac{\partial E}{\partial y_2} \cancel{\frac{\partial y_1}{\partial u_2}} \cdot \frac{\partial y_2}{\partial u_2} \cdot \frac{\partial u_2}{\partial y_1} \cdot \frac{\partial y_1}{\partial u_1} \cdot \frac{\partial u_1}{\partial b_1}$$

$$= \underbrace{-2(T - y_2)}\,\underbrace{f'(u_2)}\,\underbrace{w_2}\,\underbrace{f'(u_1)} \cdot \underbrace{1}$$

Das Problem wird noch mehr verstärkt, da noch eine weitere Ableitung der Transferfunktion als Faktor vorkommt.
Und es gilt, dass $f'(x) < 1$, da $f(x)(1 - f(x)) < 1$
(zwischen 0 und 1)

b) Für jede neue Zwischenschicht würde eine weiter Ableitung der Transferfunktion hinzugefügt werden. d.h. das Problem würde sich noch weiter verstärken, da $0 < f'(x) < 1$.

c) Dadurch dass die Ableitungen der Fehlerfunktion so klein ist, sind die einzelnen Lernschritte ebenfalls sehr klein.

3a) $$\left(\frac{\partial E}{\partial b_2} = -2(T-y_2)f'(u_2)\right)$$

$$u_1 = w_1 x + b_1 = 0 \Rightarrow f(u_1) = y_1 = \frac{1}{1+e^0} = \frac{1}{2}$$

$$u_2 = w_2 y_1 + b_2 = 0 \Rightarrow y_2 = f(u_2) = \frac{1}{2}$$

$$f'(u_1) = f(u_1)(1-f(u_1)) = \frac{1}{2}\left(1-\frac{1}{2}\right) = \frac{1}{4}$$

$$f'(u_2) = \frac{1}{2}\left(1-\frac{1}{2}\right) = \frac{1}{4}$$

$$\Rightarrow \frac{\partial E}{\partial b_2} = -2(T-y_2)f'(u_2) = -2\left(0-\frac{1}{2}\right)\frac{1}{4} = +\frac{1}{4}$$

$$\frac{\partial E}{\partial b_1} = -2(T-y_2)f'(u_2)w_2 f'(u_1)$$
$$= -2\left(-\frac{1}{2}\right)\frac{1}{4}\cdot 100 \cdot \frac{1}{4} = \frac{25}{4}$$

b) Für jede weitere Schicht erhöht sich die Folge um Faktor $\frac{1}{4}\cdot w_i = 25$, da jedes mal eine $f'(0)\cdot w_i = 25$ zu der Ableitung hinzugefügt wird.

c) Durch dieses Verhalten entsteht ein sprunghaftes Verhalten beim Lernen des Netzwerks. Dadurch könnte das Lernverfahren seine Konvergenz verlieren, da um das lokale Minimum „herumgehüpft" wird.

4.) Die Ableitung von E_C hat eine Ableitung der Transferfunktion weniger, also wird der Lernschritt größer.