

3. Aufgabenblatt

$$\textcircled{1} \quad \nabla E \left( \begin{array}{c} \frac{\partial E}{\partial w} \\ \frac{\partial E}{\partial b} \end{array} \right)$$

$$\frac{\partial E}{\partial w} = \frac{1}{2} \cdot \sum_{\mu=1}^M \left[ 2 \cdot (T_{\mu} - f(w \cdot x_{\mu} + b)) \cdot (-f'(w \cdot x_{\mu} + b)) \cdot x_{\mu} \right]$$

$$\frac{\partial E}{\partial b} = \frac{1}{2} \sum_{\mu=1}^M \left[ 2 \cdot (T_{\mu} - f(w \cdot x_{\mu} + b)) \cdot (-f'(w \cdot x_{\mu} + b)) \cdot 1 \right]$$

$\textcircled{2}$

a) inkrementelle Version: S. 78

$$\begin{aligned} w(t+1) &= w(t) + 2 \ell(t) \cdot (T_{\mu} - y_{\mu}) \cdot f'(u_{\mu}) \cdot x_{\mu} \\ &= w(t) + 2 \eta \cdot (T_{\mu} - f(w x_{\mu} + b)) \cdot f'(w x_{\mu} + b) \cdot x_{\mu} \end{aligned}$$

$$b(t+1) = b(t) + 2 \underbrace{\ell(t)}_{=\eta} \cdot (T_{\mu} - f(w x_{\mu} + b)) \cdot f'(w x_{\mu} + b)$$

b) Batch-Version:

$$\begin{aligned} w(t+1) &= w(t) + 2 \ell(t) \sum_{\mu=1}^M (T_{\mu} - y_{\mu}) \cdot f'(u_{\mu}) \cdot x_{\mu} \\ &= w(t) + 2 \eta \sum_{\mu=1}^M (T_{\mu} - f(w x_{\mu} + b)) \cdot f'(w x_{\mu} + b) \cdot x_{\mu} \end{aligned}$$

$$b(t+1) = b(t) + 2 \underbrace{\ell(t)}_{=\eta} \sum_{\mu=1}^M (T_{\mu} - f(w x_{\mu} + b)) \cdot f'(w x_{\mu} + b)$$

④

a) Batch  $\hat{=}$  Pfad 1 : verläuft gerade

inkrementelle Lernregel  $\hat{=}$  Pfad 2: verläuft sprunghaft im Zickzack.

↑

nur 1 Datenpunkt verwenden

b) Batch besser als Inkrementelle Lernregel, da mehr Daten verwendet werden (also genauer) und somit weniger sprunghaft.

Mini-Batch besser, da wir weniger Datenpunkte verwenden & somit schnellere Rechenzeit erzielen.

⑤

Ist die Lernrate groß  $\Rightarrow$  Wir haben große Sprünge, welche schlecht sind, wenn das Ziel nah ist,

da wir dann immer „um das Ziel herumhüpfen“.

$\Rightarrow$  sprunghafte Fehlerrate mit vielen hohen Extremen.