NLP_übung 2, Y. Cui, 2758172

Aufgabe 1:

① 1) $\text{Sigmoid}(z) = \frac{1}{1+e^{-z}} \Rightarrow \text{Sig}'(z) = \frac{1}{(1+e^{-z})} \cdot \frac{e^{-z}}{(1+e^{-z})} = \text{Sig}(z) \cdot (1-\text{Sig}(z))$

2) $\tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} \Rightarrow \tanh'(z) = 1 - \tanh^2(z) = \frac{(e^z+e^{-z})^2 - (e^z - e^{-z})}{(e^z+e^{-z})^2}$

3) $\text{ReLU}(z) = \max(0, z) \Rightarrow \text{ReLU}'(z) = \begin{cases} 1, & z > 0 \\ 0, & z \leq 0 \end{cases}$

② Feed Forward:

1) $\gamma_1 = \text{ReLU}(x_1 \cdot T_{1,1} + x_2 \cdot T_{2,1}) = \text{ReLU}(1 \cdot 0,19 + 1 \cdot -0,42) = 0$

$\gamma_2 = \text{ReLU}(x_1 \cdot T_{1,2} + x_2 \cdot T_{2,2}) = \text{ReLU}(1 \cdot -0,92 + 1 \cdot -0,28) = 0,64$

2) $q_1 = \tanh(\gamma_1 \cdot U_{1,1} + \gamma_2 \cdot U_{2,1}) = \tanh(0 \cdot 0,61 + \underbrace{0,64 \cdot -1,5}_{-0,96}) = -0,7443$

3) $p_1 = \text{Sig}(q_1 \cdot V_{1,1}) = \text{sig}(-0,7433 \cdot 1,5) = \text{sig}(-1,1165) = 0,2467$

$p_2 = \text{Sig}(q_1 \cdot V_{1,2}) = \text{sig}(-0,7433 \cdot -0,81) = \text{Sig}(0,6029) = 0,6463$

$p_3 = \text{Sig}(q_1 \cdot V_{1,3}) = \text{sig}(-0,7433 \cdot -0,24) = \text{Sig}(0,1786) = 0,5445$

4) $y_1 = \text{sig}(p_1 \cdot W_{1,1} + p_2 \cdot W_{2,1} + p_3 \cdot W_{3,1})$

$= \text{Sig}(0,2467 \cdot -1,4 + 0,6463 \cdot -2,2 + 0,5445 \cdot -0,27) = \text{Sig}(-1,9143) = 0,1285$

$y_2 = \text{Sig}(p_1 \cdot W_{1,2} + p_2 \cdot W_{2,2} + p_3 \cdot W_{3,2})$

$= \text{Sig}(0,2467 \cdot -0,81 + 0,6463 \cdot -1,7 + 0,5445 \cdot 0,73) = \text{Sig}(-1,6960) = 0,1550$

③ Back Propagation

1) $\frac{\partial E}{\partial y_1} = 2 \cdot (y_1 - t_1) = 2 \cdot (0,1285 - 0) = 0,2570$

$\frac{\partial E}{\partial y_2} = 2 \cdot (y_2 - t_2) = 2 \cdot (0,1550 - 1) = -1,690$

2) $\frac{\partial E}{\partial p_1} = \frac{\partial E}{\partial y_1} \cdot \frac{\partial y_1}{\partial p_1} + \frac{\partial E}{\partial y_2} \cdot \frac{\partial y_2}{\partial p_1} = \sum_{i=1}^{2} \frac{\partial E}{\partial y_i} \cdot \text{Sig}'(z_i^y) \cdot W_{1,i}$

$= 0,257 \cdot \text{Sig}'(-1,9143) \cdot -1,4 + -1,69 \cdot \text{Sig}'(-1,696) \cdot -0,8$

$= 0,1389$

$$\frac{\partial E}{\partial p_2} = \frac{\partial E}{\partial y_1}\cdot\frac{\partial y_1}{\partial p_2} + \frac{\partial E}{\partial y_2}\cdot\frac{\partial y_2}{\partial p_2} = \sum_{i=1}^{2}\frac{\partial E}{\partial y_i}\cdot sig'(z_i^y)\cdot w_{2,i}$$

$$= 0.257 \cdot sig(-1.9143)\cdot -1.4 + -1.69 \cdot sig'(1.696)\cdot -0.81$$

$$= 0.3129$$

$$\frac{\partial E}{\partial p_3} = \frac{\partial E}{\partial y_1}\cdot\frac{\partial y_1}{\partial p_3} + \frac{\partial E}{\partial y_2}\cdot\frac{\partial y_2}{\partial p_3} = \sum_{i=1}^{2}\frac{\partial E}{\partial y_i}\cdot sig'(z_i^y)\cdot w_{3,i}$$

$$= 0.257 \cdot sig(-1.9143)\cdot -0.27 + -1.69 \cdot sig'(1.696)\cdot 0.73$$

$$= 0.1538$$

3) $$\frac{\partial E}{\partial q_1} = \frac{\partial E}{\partial p_1}\frac{\partial p_1}{\partial q_1} + \frac{\partial E}{\partial p_2}\cdot\frac{\partial p_2}{\partial q_1} + \frac{\partial E}{\partial p_3}\cdot\frac{\partial p_3}{\partial q_1} = \sum_{i=1}^{3}\frac{\partial E}{\partial p_i}\cdot sig'(z_i^p)\cdot v_i$$

$$= 0.1389\cdot sig'(-1.1165)\cdot 1.5 + 0.3129\cdot sig'(0.6029)\cdot -0.81 + 0.1538\cdot sig'(0.1786)\cdot -0.24$$

$$= -0.0284$$

4) $$\frac{\partial E}{\partial r_1} = \frac{\partial E}{\partial q_1}\cdot\frac{\partial q_1}{\partial r_1} = \frac{\partial E}{\partial q_1}\cdot \tanh'(z_1^q)\cdot v_{1,1}$$

$$= -0.0284\cdot \tanh'(-0.96)\cdot 0.61$$

$$= -0.0077$$

$$\frac{\partial E}{\partial r_2} = \frac{\partial E}{\partial q_1}\cdot\frac{\partial q_1}{\partial r_2} = \frac{\partial E}{\partial q_1}\cdot \tanh'(z_1^q)\cdot v_{1,2}$$

$$= -0.0284\cdot \tanh'(0.96)\cdot -1.5$$

$$= 0.0190$$

④ $$\frac{\partial E}{\partial w_{1,1}} = \frac{\partial E}{\partial y_1}\cdot sig'(z_1^y)\cdot p_1$$

$$= 0.2570\cdot sig'(-1.9143)\cdot 0.2467$$

$$= 0.0071$$

$$\frac{\partial E}{\partial v_{1,3}} = \frac{\partial E}{\partial p_3} \cdot Sig'(z_3^p) \cdot q_1$$

$$= 0,1538 \cdot Sig'(0.1786) \cdot -0,7443$$

$$= -0,0284$$

$$\frac{\partial E}{\partial t_{1,1}} = \frac{\partial E}{\partial r_1} \cdot ReLU'(z_1^r) \cdot x_1 = 0$$