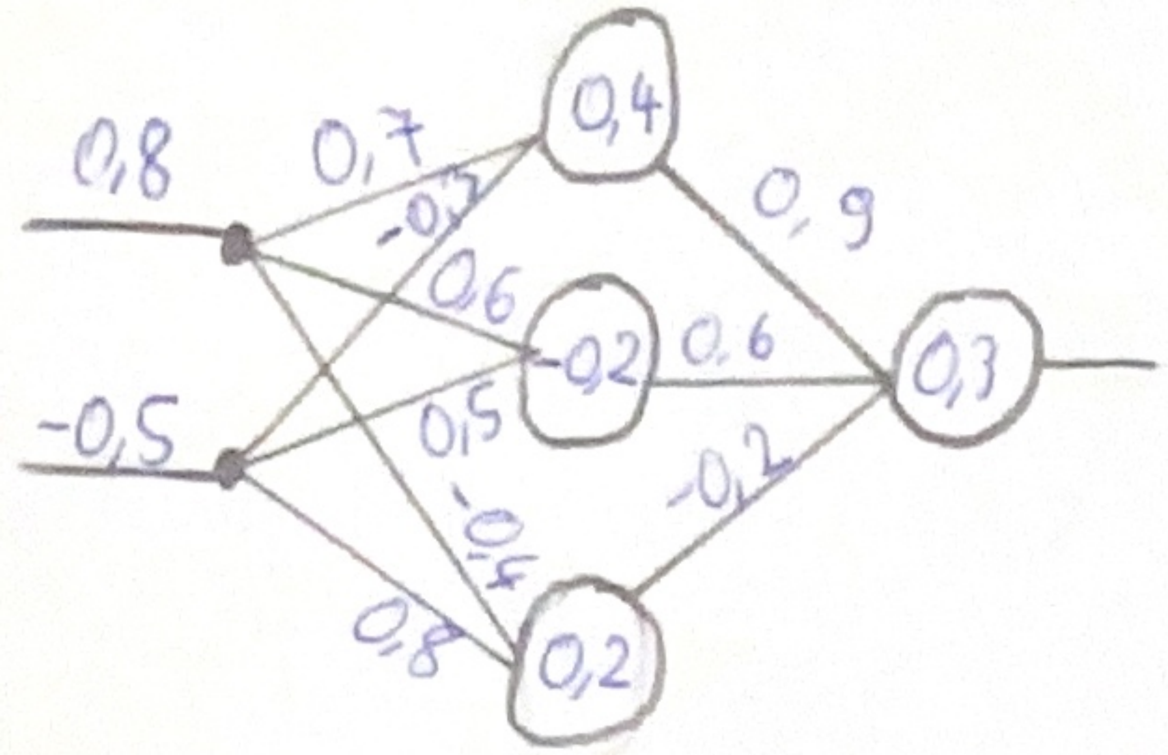
Neural Network 4 : 2-3-1

$y = 0,5$
$\sigma = \text{ReLU}$



Forwardpass

① Compute Forwards

$$z_0^1 = w_{00}^1 \cdot a_0^0 + w_{01}^1 \cdot a_1^0 + b_0^1$$
$$= 0,7 \cdot 0,8 - 0,3 \cdot -0,5 + 0,4$$
$$= 1,11 \to 1$$
$$\sigma = 1,11 \;//$$

$$z_1^1 = w_{10}^1 \cdot a_0^0 + w_{11}^1 \cdot a_1^0 + b_1^1$$
$$= 0,8 \cdot 0,6 - 0,5 \cdot 0,5 + (-0,2)$$
$$= 0,03 \to 1$$
$$\sigma = 0,03 \;//$$

$$z_2^1 = w_{20}^1 \cdot a_0^0 + w_{21}^1 \cdot a_1^0 + b_2^1$$
$$= 0,8 \cdot -0,4 - 0,5 \cdot 0,8 + 0,2$$
$$= -0,52 \to 0$$
$$\sigma = 0 \;//$$

$$z_0^2 = w_{00}^2 \cdot a_0^1 + w_{01}^2 \cdot a_1^1 + w_{02}^2 \cdot a_2^1 + b_0^2$$
$$= 0,9 \cdot \underset{1,11}{0,4} + 0,6 \cdot \underset{0,03}{-0,2} + -0,2 \cdot \underset{0}{0,2} + 0,3$$
$$= 1,317 \;//$$
$$\sigma = 1,317$$

② Calculate Loss (MSE)

$$L = \frac{1}{2}(\hat{y} - y)^2$$
$$= \frac{(1,317 - 0,5)^2}{2}$$
$$= 0,3337445 \;//$$

④ Calculate Values of L2

$$\delta^L = (1,317 - 0,5) \cdot 1$$
$$= 0,817 \;//$$

$$\frac{\partial L}{\partial w_{00}^2} = 0,817 \cdot 1,11 = 0,90687 \;//$$
$$\frac{\partial L}{\partial w_{01}^2} = 0,03 \cdot 0,817 = 0,02451 \;//$$
$$\frac{\partial L}{\partial w_{02}^2} = 0 \;//$$
$$\frac{\partial L}{\partial b_0^2} = 0,817 \;//$$

$$a^{L-1} = [\,1,11 \quad 0,03 \quad 0$$

Backpropagation

$$L = \frac{1}{2}(a^L - y)^2$$
$$a^L = \sigma(z_k^L)$$
$$z_k^L = \sum_k w_{jk}^L \cdot a^{L-1} + b_k^L$$
$$a^{L-1} = \sigma(z_k^{L-1})$$
$$z_k^{L-1} = \sum_k w_{jk}^{L-1} \cdot a^{L-2} + b_k^{L-1}$$

$$\text{ReLU} = \begin{cases} 0 & \text{if } z \leq 0 \\ z & \text{if } z > 0 \end{cases}$$

$$\text{ReLU}' = \begin{cases} 0 & \text{if } z \leq 0 \\ 1 & \text{if } z > 0 \end{cases} \;//$$

③ Derive and find formulas

$$\delta^L = \frac{\partial L}{\partial z_k^L}$$
$$= \frac{\partial L}{\partial a^L} \frac{\partial a^L}{\partial z_k^L}$$
$$= \frac{\partial}{\partial a^L} \frac{1}{2}(a^L - y)^2 \cdot \frac{\partial}{\partial z_k^L} \sigma(z_k^L)$$
$$= (a^L - y) \cdot \text{ReLU}'(z_k^L) \;//$$

$$\delta^{L-1} = \frac{\partial L}{\partial z_k^{L-1}}$$
$$= \delta^L \cdot \frac{\partial z_k^L}{\partial a_k^{L-1}} \cdot \frac{\partial a_k^{L-1}}{\partial z_k^{L-1}}$$
$$= \delta^L \cdot w_{jk}^L \cdot \text{ReLU}'(z_k^{L-1})$$
$$= \sum_k (w_{jk}^L \cdot \delta^L) \cdot \text{ReLU}'(z_k^{L-1})$$
$$= ((W^L)^T \delta^L) \cdot \text{ReLU}'(z_k^{L-1}) \;//$$

$$\frac{\partial L}{\partial w_{jk}^L} = \delta^L \cdot \frac{\partial z_k^L}{\partial w_{jk}^L} = \delta^L \cdot a^{L-1} \;//$$

$$\frac{\partial L}{\partial b_k^L} = \delta^L \cdot \frac{\partial z_k^L}{\partial b_k^L} = \delta^L \;//$$

$$\frac{\partial L}{\partial w_{jk}^{L-1}} = \delta^{L-1} \cdot \frac{\partial z_k^{L-1}}{\partial w_{jk}^{L-1}} = \delta^{L-1} \cdot a^{L-2} \;//$$

$$\frac{\partial L}{\partial b_k^{L-1}} = \delta^{L-1} \cdot \frac{\partial z_k^{L-1}}{\partial b_k^{L-1}} = \delta^{L-1} \;//$$

⑤ Calculating values of L1

$$\delta^{L-1} = ((W^L)^T \delta^L) \odot ReLU'(z_k^{L-1})$$

$$W^L = \begin{bmatrix} 0,9 \\ 0,6 \\ -0,2 \end{bmatrix} \quad \delta^L = 0,817 \quad ReLU'(z_k^{L-1}) = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

$$\delta^{L-1} = \begin{bmatrix} 0,7353 \\ 0,4902 \\ -0,1634 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0,7353 \\ 0,4902 \\ 0 \end{bmatrix}$$

$$\frac{\partial L}{\partial w_{00}^1} = 0,7353 \cdot 0,8 = 0,58824 \qquad \frac{\partial L}{\partial w_{01}^1} = 0,7353 \cdot -0,5 = -0,36765 \qquad b \Rightarrow \frac{\partial L}{\partial b_0^1} = 0,7353$$

$$\frac{\partial L}{\partial w_{10}^1} = 0,4902 \cdot 0,8 = 0,39216 \qquad \frac{\partial L}{\partial w_{11}^1} = 0,4902 \cdot -0,5 = -0,2451 \qquad \frac{\partial L}{\partial b_1^1} = 0,4902$$

$$\frac{\partial L}{\partial w_{20}^1} = 0 \qquad \frac{\partial L}{\partial w_{21}^1} = 0 \qquad \frac{\partial L}{\partial b_2^1} = 0$$

Optimization ($\alpha = 0,1$)

$$w_{jk}^L = w_{jk}^L - \alpha \cdot \frac{\partial L}{\partial w_{jk}^L}$$

$$b_k^L = b_k^L - \alpha \cdot \frac{\partial L}{\partial b_k^L}$$

$$w_{00}^1 = 0,7 - 0,1 \cdot 0,58824 = 0,641176$$

$$w_{01}^1 = -0,3 - 0,1 \cdot -0,36765 = -0,263235$$

$$b_0^1 = 0,4 - 0,1 \cdot 0,7353 = 0,32647$$

$$w_{10}^1 = 0,6 - 0,1 \cdot 0,39216 = 0,560784$$

$$w_{11}^1 = 0,5 - 0,1 \cdot -0,2451 = 0,52451$$

$$b_1^1 = -0,2 - 0,1 \cdot 0,4902 = -0,24902$$

$$w_{20}^1 = -0,4 - 0,8 \cdot 0 = -0,4$$

$$w_{21}^1 = 0,8$$

$$b_2^1 = 0,2$$

$$w_{00}^2 = 0,9 - 0,1 \cdot 0,90687 = 0,809313$$

$$w_{01}^2 = 0,6 - 0,1 \cdot 0,02451 = 0,597549$$

$$w_{02}^2 = -0,2 - 0,1 \cdot 0 = -0,2$$

$$b_0^2 = 0,3 - 0,1 \cdot 0,817 = 0,2183$$

⑦ Forward Pass

$$z_0^1 = 0,641176 \cdot 0,8 + (-0,263235) \cdot -0,5 + 0,32647 = 0,9660283$$

$$\sigma = 0,9660283$$

$$z_1^1 = 0,560784 \cdot 0,8 + 0,52451 \cdot -0,5 + (-0,24902) = -0,0626478$$

$$\sigma = 0$$

$$z_2^1 = -0,4 \cdot 0,8 + 0,8 \cdot -0,5 + 0,2 = -0,52$$

$$\sigma = 0$$

$$z_0^2 = 0,809313 \cdot 0,9660283 + 0,2183 = 1,000119262$$

⑧ Calculate Cost

$$L = \frac{1}{2}(\hat{y} - y)^2 = \frac{(1,000119262 - 0,5)^2}{2} = 0,1250596381$$

Virvi Huta