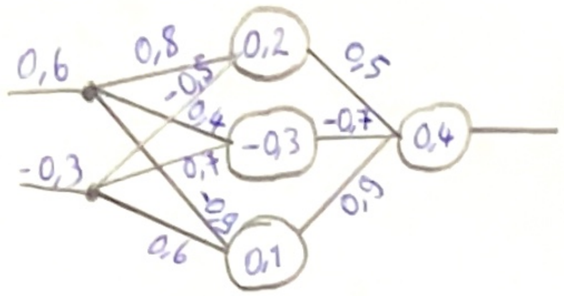
Neural Network 5: 2-3-1



$$z_0^2 = w_{00}^2 \cdot a_0^1 + w_{01}^2 \cdot a_1^1 + w_{02}^2 \cdot a_2^1 + b_0^2$$
$$= 0,5 \cdot 0,83 + 0,4$$
$$= 0,815 //$$

Forward Pass

① Compute Forwards

$y = 2$
$\sigma = \text{ReLU}$

$$z_0^1 = w_{00}^1 \cdot a_0^0 + w_{01}^1 \cdot a_1^0 + b_0^1$$
$$= 0,8 \cdot 0,6 + (-0,3) \cdot (-0,5) + 0,2$$
$$= 0,83$$
$\sigma = 0,83 // \rightarrow 1$

$$z_1^1 = w_{10}^1 \cdot a_0^0 + w_{11}^1 \cdot a_1^0 + b_1^1$$
$$= 0,6 \cdot 0,4 + 0,7 \cdot (-0,3) - 0,3$$
$$= -0,27$$
$\sigma = 0 // \rightarrow 0$

$$z_2^1 = w_{20}^1 \cdot a_0^0 + w_{21}^1 \cdot a_1^0 + b_2^1$$
$$= 0,6 \cdot (-0,9) + (-0,3) \cdot 0,6 + 0,1$$
$$= -0,62$$
$\sigma = 0 // \rightarrow 0$

② Calculate Cost (MSE)

$$L = \frac{1}{2}(\hat{y} - y)^2$$
$$= \frac{(0,815 - 2)^2}{2}$$
$$= 0,702 //$$

③ Derive/Find Formulas

Backpropagation

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

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

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

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

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

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

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

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

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

④ Calculate Values of L2

$$\delta^L = (0,815 - 2) \cdot 1$$
$$= -1,185 //$$

$$a^{L-1} = [0,83 \; 0 \; 0]$$

$$\frac{\partial L}{\partial b_0^2} = -1,185 //$$

$$\frac{\partial L}{\partial w_{00}^2} = -1,185 \cdot 0,83$$
$$= -0,98355 //$$

$$\frac{\partial L}{\partial w_{01}^2} = 0 //$$

$$\frac{\partial L}{\partial w_{02}^2} = 0 //$$

⑤ Calculate Values of L^1

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

$$W^L = \begin{bmatrix} 0,5 \\ -0,7 \\ 0,9 \end{bmatrix} \cdot \delta^L = -1,185 \cdot \text{ReLU}'(z_k^{L-1}) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\delta^{L-1} = \begin{bmatrix} -0,5925 \\ 0 \\ 0 \end{bmatrix} //$$

$\frac{\partial L}{\partial w_{00}^1} = -0,5925 \cdot 0,6$ $\frac{\partial L}{\partial w_{01}^1} = -0,5925 \cdot (-0,3)$ $\frac{\partial L}{\partial b_0^1} = -0,5925$

$\frac{\partial L}{\partial w_{00}^1} = -0,3555$ //

$\frac{\partial L}{\partial w_{10}^1} = 0$ // $\frac{\partial L}{\partial w_{11}^1} = 0$ // $\frac{\partial L}{\partial b_1^1} = 0$ //

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

Optimization ($\alpha = 0,1$)

$$w_{jk}^L = w_{jk}^L - 0,1 \cdot \frac{\partial L}{\partial w_{jk}^L}$$

$$b_k^L = b_k^L - 0,1 \cdot \frac{\partial L}{\partial b_k^L}$$

⑥ Update Values

$w_{jk}^L \Rightarrow$ $w_{00}^1 = 0,8 - 0,1 \cdot (-0,3555)$ $w_{01}^1 = -0,5 - 0,1 \cdot 0,17775$ $b_k^L \Rightarrow$

$= 0,83555$ // $= -0,517775$ //

$b_0^1 = 0,2 - 0,1 \cdot (-0,592$

$= 0,25925$ //

$w_{10}^1 = 0,4$ // $w_{11}^1 = 0,7$ // $b_1^1 = -0,3$ //

$w_{20}^1 = -0,9$ // $w_{21}^1 = 0,6$ // $b_2^1 = 0,1$ //

$w_{00}^2 = 0,5 - 0,1 \cdot (-0,98355)$ $w_{01}^2 = -0,7 - 0,1 \cdot 0$ $w_{02}^2 = 0,9$ //

$= 0,598355$ // $= -0,7$ //

$b_0^2 = 0,4 - 0,1 \cdot (-1,185)$

$= 0,5185$ //

⑦ Forward Pass

$z_0^1 = 0,83555 \cdot 0,6 + (-0,517775) \cdot -0,3 + 0,25925$

$= 0,9159125$

$\sigma = 0,9159125$ //

$z_1^1 = 0,4 \cdot 0,6 + 0,7 \cdot -0,3 + (-0,3)$

$= -0,27$

$\sigma = 0$ //

$z_2^1 = -0,9 \cdot 0,6 + 0,6 \cdot -0,3 + 0,1$

$= -0,62$

$\sigma = 0$ //

$z_0^2 = 0,598355 \cdot 0,9159125 + (-0,7) \cdot 0 + 0,9 \cdot 0 + 0,5185$

$= 0,956383706$

$\sigma = 0,956383706$ //

⑧ Calculate Cost

$$L = \frac{1}{2}(\hat{y} - y)^2$$

$$= \frac{(0,956383706 - 2)^2}{2}$$

$= 0,544567484\,6$ //

Virui Huta Vluta