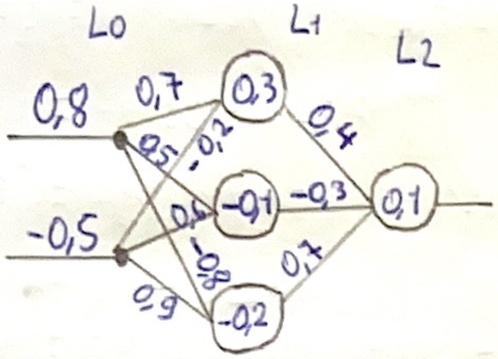
Neural Network 3: 2-3-1



$y = 0,5$
σ = ReLU

Forward Pass

① Compute Forwards

$z_0^1 = w_{00}^1 \cdot a_0^0 + w_{01}^1 \cdot a_1^0 + b_0^1$
$= 0,7 \cdot 0,8 + -0,2 \cdot -0,5 + 0,3$
$= 0,96$
$\sigma = 0,96$ // $\rightarrow 1$

$z_1^1 = w_{10}^1 \cdot a_0^0 + w_{11}^1 \cdot a_1^0 + b_1^1$
$= 0,8 \cdot 0,5 + -0,5 \cdot 0,6 - 0,1$
$= 0$
$\sigma = 0$ // $\rightarrow 0$

$z_2^1 = w_{20}^1 \cdot a_0^0 + w_{21}^1 \cdot a_1^0 + b_2^1$
$= -0,8 \cdot 0,8 + 0,9 \cdot -0,5 + (-0,2)$
$= -1,29$
$\sigma = 0$ // $\rightarrow 0$

$z_0^2 = w_{00}^2 \cdot a_0^1 + w_{01}^2 \cdot a_1^1 + w_{02}^2 \cdot a_2^1 + b_0^2$
$= 0,4 \cdot 0,96 + 0 \cdot (-3) + 0 \cdot 0,7 + 0,1$
$= 0,484$
$\sigma = 0,484$ //

② Calculate Cost (MSE)

$L = \frac{1}{2}(\hat{y} - y)^2$
$= \frac{(0,484 - 0,5)^2}{2}$
$= 0,000128$ //

③ Derive / Find Formulas

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

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

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

④ Calculate values of L2

$\delta^L = 0,484 - 0,5 \cdot 1$
$= -0,016$ //
$a^{L-1} = [0\ 0\ 0]$

$\frac{\partial L}{\partial w_{00}^2} = -0,016 \cdot 0$ // $= -0,016$ //
$= -0,01536$ //

$\frac{\partial L}{\partial w_{01}^2} = -0,016 \cdot 0 = 0$ //

$\frac{\partial L}{\partial w_{02}^2} = -0,016 \cdot 0 = 0$ //

$\frac{\partial L}{\partial b_k^L} = -0,016$ //

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

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

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

Backpropagation

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

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

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

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

Calculate Values of L1 ⑤

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

$$W^L = \begin{bmatrix} 0,4 \\ -0,3 \\ 0,7 \end{bmatrix} \quad ReLU'(z_k^{L-1}) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\delta^L = -0,016$$

$$\delta^{L-1} = \begin{bmatrix} 0,4 \\ -0,3 \\ 0,7 \end{bmatrix} (-0,016) \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} -0,0064 \\ 0,0048 \\ -0,0112 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} -0,0064 \\ 0 \\ 0 \end{bmatrix} //$$

$$\frac{\partial L}{\partial w_{00}^1} = -0,0064 \cdot 0,8 = -0,00512 // \qquad \frac{\partial L}{\partial w_{01}^1} = -0,0064 \cdot -0,5 = 0,0032 // \qquad \frac{\partial L}{\partial b_0^1} = -0,0064 //$$

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

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

Optimization ($\alpha = 0,1$)

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

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

⑥ Update Values

$$w_{00}^1 = 0,7 - 0,1 \cdot -0,00512 = 0,700512 //$$

$$b_0^1 = 0,3 - 0,1 \cdot 0,0064 = 0,29936 //$$

$$w_{01}^1 = -0,2 - 0,1 \cdot 0,0032 = -0,20032 //$$

$$w_{10}^1 = 0,5 // \qquad w_{11}^1 = 0,6 //$$

$$b_1^1 = -0,1 // \qquad w_{10}^1 = -0,8 //$$

$$w_{21}^1 = 0,9 // \qquad b_2^1 = -0,2 // \qquad w_{00}^2 = 0,4 - 0,1 \cdot -0,01536 = 0,401536 //$$

$$w_{01}^2 = -0,3 // \qquad w_{02}^2 = 0,7 //$$

$$b_0^2 = 0,1 - 0,1 \cdot -0,016 = 0,1016 //$$

⑦ Forward Pass

$$z_0^1 = 0,700512 \cdot 0,8 + (-0,20032) \cdot -0,5 + 0,29936 = 0,9599296 //$$

$$z_1^1 = 0,5 \cdot 0,8 + 0,6 \cdot -0,5 + (-0,1) = 0 \rightarrow \sigma = 0 //$$

$$z_2^1 = -0,8 \cdot 0,8 + 0,9 \cdot (-0,5) + (-0,2) = -1,29 \rightarrow \sigma = 0 //$$

$$z_0^2 = 0,401536 \cdot 0,9599296 + (-0,3) \cdot 0 + 0,7 \cdot 0 + 0,1016 = 0,4870462919 //$$

⑧ Calculate Cost

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

$$= \frac{(0,4870462919 - 0,5)^2}{2}$$

$$= 0,0000838992 //$$

Virvi Huta