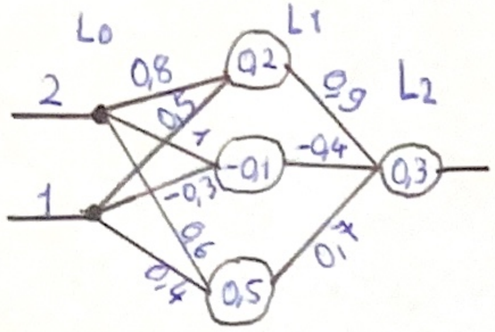
Neural Network 2: 231

$y = 0,9$
$\sigma = \text{ReLU}$



Forward Pass

① Compute Forwards

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

$z_1^1 = w_{10}^1 \cdot a_0^0 + w_{11}^1 \cdot a_1^0 + b_1^1$
$= 2 + (-0,3) + (-0,1)$
$= 1,6 \rightarrow 1$
$\sigma = 1,6$ //

$z_2^1 = w_{20}^1 \cdot a_0^0 + w_{21}^1 \cdot a_1^0 + b_2^1$
$= 1,2 + 0,4 + 0,5$
$= 2,1 \rightarrow 1$
$\sigma = 2,1$ //

$z_0^2 = w_{00}^2 \cdot a_0^1 + w_{01}^2 \cdot a_1^1 + w_{02}^2 \cdot a_2^1 + b_0^2$
$= 0,9 \cdot 2,3 + (-0,4) \cdot 1,6 + 0,7 \cdot 2,1 + 0,3$
$= 3,2$
$\sigma = 3,2$ //

② Calculate Cost (MSE)

$L = \frac{1}{2}(y^1 - y)^2$
$= \frac{(3,2 - 0,9)^2}{2}$
$= 2,645$ //

④ Calculate values of L2

$\delta^L = (3,2 - 0,9) \cdot 1$
$= 2,3$ //

$\frac{\partial L}{\partial w_{00}^2} = 2,3 \cdot 2,3$
$= 5,29$ //

$\frac{\partial L}{\partial w_{04}^2} = 2,3 \cdot 1,6$
$= 3,68$ //

$\frac{\partial L}{\partial w_{02}^2} = 2,3 \cdot 2,1$
$= 4,83$ //

$\frac{\partial L}{\partial b_0^2} = \delta^L \cdot 1$
$= 2,3$ //

③ Derive / Find Formulas

Backpropagation

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

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

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

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

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

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

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

$a^{L-1} = [2,3 \quad 1,6 \quad 2,1]$ //

(5) Calculate Values of L1

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

$$w^L = \begin{bmatrix} 0,9 \\ -0,4 \\ 0,7 \end{bmatrix} \cdot 2,3 \rightarrow \delta^L$$

$$\delta^{L-1} = \begin{bmatrix} 2,07 \\ -0,92 \\ 1,61 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2,07 \\ -0,92 \\ 1,61 \end{bmatrix} //$$

$\frac{\partial L}{\partial w_{00}^1} = 2,07 \cdot 2$ $\quad \frac{\partial L}{\partial w_{01}^1} = 2,07$ // $\quad \frac{\partial L}{\partial b_0^1} = 2,07$ //

$\frac{\partial L}{\partial w_{00}^1} = 4,14$ //

$\frac{\partial L}{\partial w_{10}^1} = -0,92 \cdot 2$ $\quad \frac{\partial L}{\partial w_{11}^1} = -0,92$ // $\quad \frac{\partial L}{\partial b_1^1} = -0,92$ //

$\frac{\partial L}{\partial w_{10}^1} = -1,84$ //

$\frac{\partial L}{\partial w_{20}^1} = 1,61 \cdot 2$ $\quad \frac{\partial L}{\partial w_{21}^1} = 1,61$ // $\quad \frac{\partial L}{\partial b_2^1} = 1,61$ //

$\frac{\partial L}{\partial w_{20}^1} = 3,22$ //

(6) Update Values

Optimization ($\alpha = 0,1$)

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

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

$w_{00}^1 = 0,8 - 0,1 \cdot 4,14$
$= 0,386$ //

$w_{01}^1 = 0,5 - 0,1 \cdot 2,07$
$= 0,293$ //

$b_0^1 = 0,2 - 0,1 \cdot 2,07$
$= -0,007$ //

$w_{10}^1 = 1 - 0,1 \cdot (-1,84)$
$= 1,184$ //

$w_{11}^1 = -0,3 - 0,1 \cdot (-0,92)$
$= -0,208$ //

$b_1^1 = -0,1 - 0,1 \cdot (-0,92)$
$= -0,008$ //

$w_{20}^1 = 0,6 - 0,1 \cdot 3,22$
$= 0,278$ //

$w_{21}^1 = 0,4 - 0,1 \cdot 1,61$
$= 0,239$ //

$b_2^1 = 0,5 - 0,1 \cdot 1,61$
$= 0,339$

$w_{00}^2 = 0,9 - 0,1 \cdot 5,29$
$= 0,371$ //

$w_{01}^2 = -0,4 - 0,1 \cdot 3,68$
$= -0,768$ //

$w_{02}^2 = 0,7 - 0,1 \cdot 4,83$
$= 0,217$ //

$b_0^2 = 0,3 - 0,1 \cdot 2,3$
$= 0,07$ //

(7) Forward Pass

$z_0^1 = 0,386 \cdot 2 + 0,293 + (-0,007)$
$= 1,058$
$\sigma = 1,058$ //

$z_1^1 = 1,184 \cdot 2 + (-0,208) - 0,008$
$= 2,152$
$\sigma = 2,152$ //

$z_2^1 = 0,278 \cdot 2 + 0,239 + 0,339$
$= 1,134$
$\sigma = 1,134$ //

$z_0^2 = 1,058 \cdot 0,371 + 2,152 \cdot -0,768 + 1,134 \cdot 0,217 + 0,07$
$= -0,94414$
$\sigma = 0$ //

(8) Calculate Cost

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

$$= \frac{(0,9 - 0)^2}{2}$$

$$= 0,405$$ //

Virvi Huta