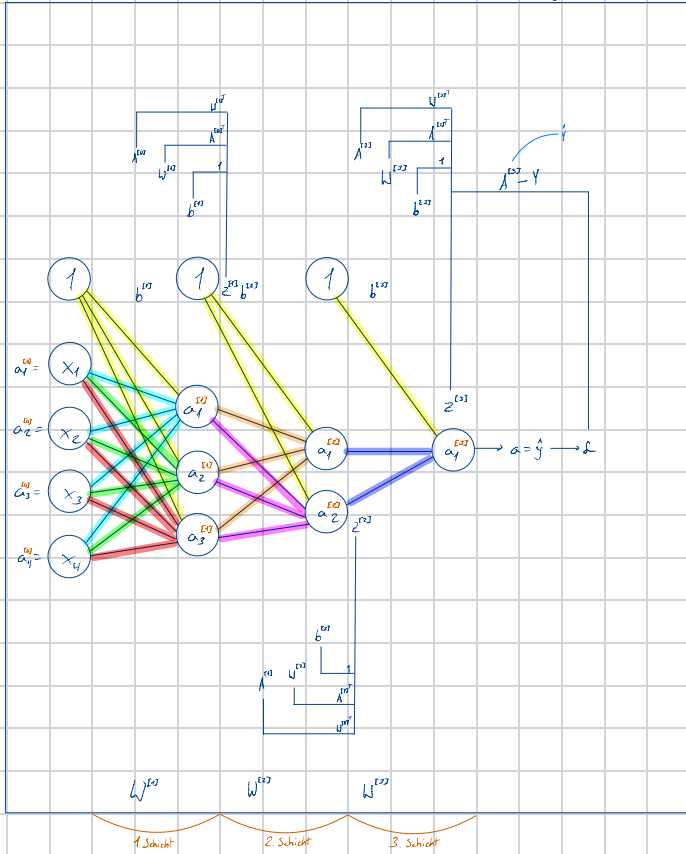


$$z^{[l]} = W^{[l]} \cdot a^{[l-1]} + b^{[l]}$$

$$a^{[l]} = \sigma(z^{[l]})$$

$$\longrightarrow \frac{1}{1+e^{-z}} \xrightarrow{\text{Aktiv.}} a(1-a) \rightarrow A \cdot (1-A)$$

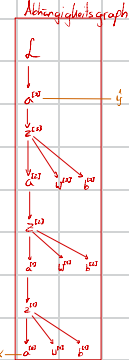


$$dW^{[l]} = \frac{\partial f}{\partial z^{[l]}} \cdot \frac{\partial z^{[l]}}{\partial a^{[l-1]}} \cdot \frac{\partial a^{[l-1]}}{\partial z^{[l-1]}} \cdot \frac{\partial z^{[l-1]}}{\partial a^{[l-2]}} \cdot \frac{\partial a^{[l-2]}}{\partial z^{[l-2]}} \cdot \frac{\partial z^{[l-2]}}{\partial W^{[l-1]}}$$

$$= (A^{[l-1]} - y) \cdot W^{[l-1]T} \cdot A^{[l-1]} \cdot (1 - A^{[l-1]}) \cdot W^{[l-1]T} \cdot A^{[l-2]} \cdot (1 - A^{[l-2]}) \cdot A^{[l-2]T}$$

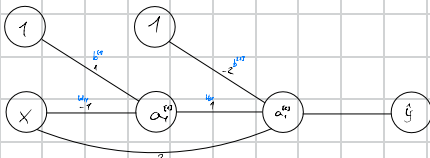
$$\frac{1}{3} \cdot d z^{[l]} \cdot A^{[l]T} = dW^{[l]}$$

m=3



2)

$$(x, y) = (0, 0.5)$$



$$-1 \cdot x_1 + 1 = 0$$

$$-b_0 = -1$$

$$x_1 = 1$$

$$a_1^{[1]} = w_{11}^{[1]} \cdot x_1 + b_1 = -1 + 1 = 0$$

$$a_1^{[2]} = w_{21}^{[2]} \cdot x_1 + w_{22}^{[2]} \cdot x_2 + b_2 = 1 \cdot 0.19 + 2 \cdot 0 - 2 = -1.81$$

$$\frac{1}{\sigma'(1.81)} = 0.22$$

$$L(0.19, 0.5) = 0.881$$

$$\frac{\partial L}{\partial a^{[1]}} = \frac{\partial L}{\partial a^{[2]}} \cdot \frac{\partial a^{[2]}}{\partial z^{[2]}} \cdot \frac{\partial z^{[2]}}{\partial a^{[1]}} \cdot \frac{\partial a^{[1]}}{\partial z^{[1]}} \cdot \frac{\partial z^{[1]}}{\partial W^{[1]}}$$

$$\frac{\partial L}{\partial a_1^{[1]}} = (a - y) \cdot a_1^{[1]} = (0.22 - 0.5) \cdot 0.731 = -0.205$$

$$\frac{\partial L}{\partial a_2^{[1]}} = (a - y) \cdot x = -0.28 \cdot 0 = 0$$

$$\frac{\partial L}{\partial b_1^{[1]}} = (a - y) \cdot 1 = -0.28$$

$$d a^{[1]} = (a - y) \cdot W^{[1]} \rightarrow W^{[1]} \cdot (a - y) = \begin{bmatrix} 2 \cdot (-0.28) \\ 1 \cdot (-0.28) \end{bmatrix} = \begin{bmatrix} -0.56 \\ -0.28 \end{bmatrix}$$

$$\frac{dW^{[1]}}{da^{[1]}} = \sigma'(a^{[1]}) = \begin{bmatrix} 0.344 \\ 0.430 \end{bmatrix}$$

$$\frac{\partial L}{\partial u_1} = \frac{1}{1} \cdot d z^{[1]} \cdot 0 = 0$$

$$d b^{[1]} = \frac{1}{1} \cdot \text{softmax}(d z^{[1]}) = 0.794$$

$$W_{\text{neu}} = W_{\text{alt}} - \alpha \cdot \nabla J = \begin{bmatrix} w_{11}^{[1]} \\ w_{12}^{[1]} \\ w_{21}^{[1]} \\ w_{22}^{[1]} \end{bmatrix} - 0.01 \cdot \begin{bmatrix} \frac{dw_{11}^{[1]}}{da^{[1]}} \\ \frac{dw_{12}^{[1]}}{da^{[1]}} \\ \frac{dw_{21}^{[1]}}{da^{[1]}} \\ \frac{dw_{22}^{[1]}}{da^{[1]}} \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 1 \\ 2 \end{bmatrix} - 0.01 \cdot \begin{bmatrix} 0.344 \\ 0.430 \\ -0.28 \\ -0.28 \end{bmatrix} = \begin{bmatrix} -1.00344 \\ 0.99566 \\ 0.9972 \\ 1.9972 \end{bmatrix}$$