

OZEKI

B. Muneer Nurkassat

$$X = [20, 3, 4]^T$$

$$Y = 18$$

$$W_1 = \begin{bmatrix} 0,1 & 0,2 & 0,3 \\ 0,4 & 0,5 & 0,6 \\ 0,7 & 0,8 & 0,9 \end{bmatrix} \quad b_1 = \begin{bmatrix} 0,1 \\ 0,2 \\ 0,3 \end{bmatrix}$$

$$W_2 = [0,2, 0,4, 0,6] \quad b_2 = 0,5$$

① Forward propagation.

$$Z_1 = W_1 \cdot X + b_1$$

$$Z_{1,1} = 2 + 0,6 + 1,2 + 0,1 = 3,9$$

$$Z_{1,2} = 8 + 1,5 + 2,4 + 0,2 = 12,1$$

2

Müner Nurekassar

$$z_{1,2} = 1,4 + 2,4 + 5,6 + 0,3 = 20,3$$

$$z_1 = \begin{bmatrix} 3,9 \\ 12,1 \\ 20,3 \end{bmatrix}$$

$$A_1 = \text{ReLU}(z_1) = \max(0, z_1)$$

$$A_1 = \begin{bmatrix} 3,9 \\ 12,1 \\ 20,3 \end{bmatrix}$$

$$z_2 = W_2 \cdot A_1 + b_2 =$$

$$= 0,2 \cdot 3,9 + 0,4 \cdot 12,1 + 0,6 \cdot 20,3 + 0,5 = 18,3$$

$$z_2 = 18,3$$

ÖZKEİ

Müner Nurekassar

$$A_2 = \sigma(z_2) = \frac{1}{1 + e^{-z_2}} \approx 1$$

$$A_2 \approx 1 = 0,999999988$$

$$\text{Loss} : (A_2 - Y)^2 = (1 - 18)^2 = 289$$

2) Backward Propagation:

$$1. \frac{dL}{dA_2} = 2 \cdot (A_2 - Y) = 2 \cdot (1 - 18) = -34$$

$$2. \frac{dL}{dz_2} = \sigma'(z_2) \cdot (1 - \sigma(z_2)) \approx 1 \cdot (1 - 1) = 0$$

$$3. \frac{dL}{dW_2} = \frac{dL}{dz_2} \cdot A_1^T = 0 \cdot A_1^T = [0, 0, 0]$$

$$4. \frac{dL}{db_2} = \frac{dL}{dz_2} = 0$$

ÖRNEKİ

② Backward Propagation.

$$1) \frac{dL}{dA_2} = 2(A_2 - Y) = 2(1 - 18) = -34$$

$$2) \frac{dL}{dz_2} = \left(\frac{dL}{dA_2} \right) \cdot g'(z_2) = -34 \cdot 1,2 \cdot 10^{-8} = -4,08 \cdot 10^{-7}$$

$$g'(z_2) = g(z_2) (1 - g(z_2)) \approx 1,2 \cdot 10^{-8}$$

$$3) \frac{dL}{dw_2} = \frac{dL}{dz_2} \cdot A_1^T = -4,08 \cdot 10^{-7} \cdot \begin{bmatrix} 3,9 & 12,1 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} -1,59 \cdot 10^{-6} & -4,94 \cdot 10^{-6} & -8,28 \cdot 10^{-6} \end{bmatrix}$$

$$4) \frac{dL}{db_2} = \frac{dL}{dz_2} = -4,08 \cdot 10^{-7}$$

72

$$5) \frac{dL}{dA_1} = W_2^T \cdot \frac{dL}{dz_2} = \begin{bmatrix} 0,2 \\ 0,4 \\ 0,6 \end{bmatrix} \cdot (1 - 4,08 \cdot 10^{-4})_2$$

$$= \begin{bmatrix} -8,16 \cdot 10^{-8} \\ -1,63 \cdot 10^{-7} \\ -2,45 \cdot 10^{-7} \end{bmatrix}$$

$$6) \frac{dL}{dz_1} = \frac{dL}{dA_1} \cdot \text{ReLU}'(z_1) = \frac{dL}{dA_1}$$

because $z_1 > 0$, $\text{ReLU}' = 1$.

$$7) \frac{dL}{dW_1} = \frac{dL}{dz_1} \cdot X^T = \begin{bmatrix} -8,16 \cdot 10^{-8} \\ -1,63 \cdot 10^{-7} \\ -2,45 \cdot 10^{-7} \end{bmatrix} \cdot \begin{bmatrix} 20 & 3 & 4 \end{bmatrix}$$

$$= \begin{bmatrix} -1,63 \cdot 10^{-6} & -2,45 \cdot 10^{-7} & -3,26 \cdot 10^{-7} \\ -3,26 \cdot 10^{-6} & -4,89 \cdot 10^{-7} & -6,52 \cdot 10^{-7} \\ -4,90 \cdot 10^{-6} & -7,35 \cdot 10^{-7} & -9,80 \cdot 10^{-7} \end{bmatrix}$$

$$8) \frac{dL}{db_1} = \frac{dL}{dz_1} = \begin{bmatrix} -8,16 \cdot 10^{-8} \\ -1,63 \cdot 10^{-7} \\ -2,45 \cdot 10^{-7} \end{bmatrix}$$