

Part B:

$$X = \begin{bmatrix} 20 \\ 3 \\ 4 \end{bmatrix}$$

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

$$b_1 = \begin{bmatrix} 0,1 \\ 0,2 \\ 0,3 \end{bmatrix}$$

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

$$b_2 = [0,5]$$

Forward Propagation

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

$$Z_1^1 = 0,1 \cdot 20 + 0,2 \cdot 3 + 0,3 \cdot 4 + 0,1 = 3,9$$

$$Z_2^1 = 0,4 \cdot 20 + 0,5 \cdot 3 + 0,6 \cdot 4 + 0,2 = 12,1$$

$$Z_3^1 = 0,7 \cdot 20 + 0,8 \cdot 3 + 0,9 \cdot 4 + 0,3 = 20,3$$

$$Z_1 = \begin{bmatrix} 3,9 \\ 12,1 \\ 20,3 \end{bmatrix} = A_1$$

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

$$Z_2 = W_2 \cdot A_1 + b_2$$

$$Z_2 = 0,2 \cdot 3,9 + 0,4 \cdot 12,1 + 0,6 \cdot 20,3 + 0,5 = 18,3$$

$$A_2 = \text{Sigmoid}(Z_2) \text{ where } \sigma(x) = 1 / (1 + e^{(-x)})$$

$$A_2 = \sigma(Z_2) = \frac{1}{1 + e^{-18,3}} \approx 0,999$$

$$L = (A_2 - 4)^2 = (0,999 - 18)^2 \approx (-17)^2 = 289$$

Backward Propagation

$$1. \frac{dL}{dA_2} = 2(A_2 - y) = 2(0,999 - 1) \approx -34$$

$$2. \frac{dL}{dz_2} = \frac{dL}{dA_2} \cdot \text{sigmoid}'(z_2) \text{ where } \text{sigmoid}'(z) = g'(z)(1 - g(z))$$

$$\times g'(z_2) = g(z_2)(1 - g(z_2)) \approx 0,999 \cdot (1 - 0,999) \approx 1,1 \cdot 10^{-8}$$

$$\frac{dL}{dz_2} = \frac{dL}{dA_2} \cdot g'(z_2) \approx -34 \cdot 1,1 \cdot 10^{-8} \approx -3,7 \cdot 10^{-7}$$

$$3. \frac{dL}{dw_2} = dz_2 \cdot A_1^T = -3,7 \cdot 10^{-7} \cdot [3,9; 12,1; 20,3]$$

$$dw_2 \approx [-1,4 \cdot 10^{-6}, -4,5 \cdot 10^{-6}, -7,5 \cdot 10^{-6}]$$

$$4. \frac{dL}{db_2} = \frac{dL}{dz_2} = -3,7 \cdot 10^{-7}$$

$$5. \frac{dL}{dA_1} = w_2^T \cdot \frac{dL}{dz_2} = [0,2; 0,4; 0,6]^T \cdot (-3,7 \cdot 10^{-7})$$

$$dA_1 \approx [-7,4 \cdot 10^{-8}, -1,5 \cdot 10^{-7}, -2,2 \cdot 10^{-7}]$$

$$6. \frac{dL}{dz_1} = \frac{dL}{dA_1} \cdot \text{ReLU}'(z_1) \text{ where } \text{ReLU}'(x) = 1 \text{ if } x > 0, \text{ else } 0$$

$$\frac{dL}{dz_1} = dA_1 \cdot \text{ReLU}'(z_1), \text{ all } z_1 > 0, \text{ ReLU}' = 1, \text{ then } dz_1 = dA_1$$

$$7. \frac{dL}{dw_1} = \frac{dL}{dz_1} \cdot x^T$$

$$dW_{1, \text{row}_1} = -7,4 \cdot 10^{-8} \cdot [20, 3, 4] = [-1,5 \cdot 10^{-6}, -2,2 \cdot 10^{-7}, -3 \cdot 10^{-7}]$$

$$dW_{1, \text{row}_2} = -1,5 \cdot 10^{-7} \cdot [20, 3, 4] = [-3 \cdot 10^{-6}, -4,5 \cdot 10^{-7}, -6 \cdot 10^{-7}]$$

$$dW_{1, \text{row}_3} = -2,2 \cdot 10^{-7} \cdot [20, 3, 4] = [-4,4 \cdot 10^{-6}, -6,6 \cdot 10^{-7}, -8,8 \cdot 10^{-7}]$$

$$8. \frac{dL}{db_1} = \frac{dL}{dz_1} = [-7,4 \cdot 10^{-8}, -1,5 \cdot 10^{-7}, -2,2 \cdot 10^{-7}]$$