Iliyos Nurlybek

Assignment 01. Training a Simple Two-Layer Neural Network with Numpy.

Part B.

$X = \begin{bmatrix} 20 \\ 3 \\ 4 \end{bmatrix}$ $Y = 18$ $W_1 = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \\ 0.7 & 0.8 & 0.9 \end{bmatrix}$ $b_1 = \begin{bmatrix} 0.1 \\ 0.2 \\ 0.3 \end{bmatrix}$ W - weigh

b - bias

$W_2 = [0.2 \ 0.4 \ 0.6]$ $b_2 = [0.5]$

Forward Propagation - Прямое распространение

1. $Z_1 = W_1 X + b_1$
2. $A_1 = ReLU(Z_1)$ where $ReLU(x) = \max(0, x)$
3. $Z_2 = W_2 A_1 + b_2$
4. $A_2 = Sigmoid(Z_2)$ where $\sigma(x) = \frac{1}{1+e^{-x}}$
5. $Loss = (A_2 - Y)^2$ where $Y = 18$.

1. $Z_1 = W_1 X + b_1$

$W_1 = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \\ 0.7 & 0.8 & 0.9 \end{bmatrix}$ $X = \begin{bmatrix} 20 \\ 3 \\ 4 \end{bmatrix}$ $b_1 = \begin{bmatrix} 0.1 \\ 0.2 \\ 0.3 \end{bmatrix}$ $Z_1 = \begin{bmatrix} 3.9 \\ 12.1 \\ 20.3 \end{bmatrix}$

$Z_1 \Rightarrow 0.1 \cdot 20 + 0.2 \cdot 3 + 0.3 \cdot 4 + 0.1 = 2 + 0.6 + 1.2 + 0.1 = 3.9$

$0.4 \cdot 20 + 0.5 \cdot 3 + 0.6 \cdot 4 + 0.2 = 8 + 1.5 + 2.4 + 0.2 = 12.1$

$0.7 \cdot 20 + 0.8 \cdot 3 + 0.9 \cdot 4 + 0.3 = 14 + 2.4 + 3.6 + 0.3 = 20.3$.

2. $ReLU(x) = \max(0, x)$ $x = Z_1$

if numbers are positive → don't change.

if numbers are negative → changes to 0

$A_1 = ReLU(Z_1) \Rightarrow Z_1 > 0 \Rightarrow Z_1 = A_1 \Rightarrow A_1 = \begin{bmatrix} 3.9 \\ 12.1 \\ 20.3 \end{bmatrix}$ $A_1 = \begin{bmatrix} 3.9 \\ 12.1 \\ 20.3 \end{bmatrix}$

$ReLU'(x_1) = \begin{cases} 1 & x > 0 \\ 0 & x \leq 0 \end{cases}$ $Z_1 > 0 \Rightarrow ReLU'(Z_1) = 1$

3. $Z_2 = W_2 A_1 + b_2$

$W_2 = [0.2 \ 0.4 \ 0.6]$ $b_2 = [0.5]$ $A_1 = Z_1 = \begin{bmatrix} 3.9 \\ 12.1 \\ 20.3 \end{bmatrix}$ $Z_2 = [18.3]$

$Z_2 = 0.2(3.9) + 0.4(12.1) + 0.6(20.3) + 0.5 = 18.3$.

4. $A_2 = sigmoid(Z_2) = \sigma(Z_2)$

$A_2 = \sigma(18.3)$

$\sigma = \frac{1}{1+e^{(-18.3)}} \approx 0.999 \simeq 1$ $A_2 = 1$

Iligas Nurlybek

5. Loss $= (A_2 - Y)^2$ $\quad Y = 18$

Loss $= (1-18)^2 = (-17)^2 = 289$ $\quad L = 289$

Backward Propagation

1. $\frac{dL}{dA_2} = 2(A2 - Y)$

2. $\frac{dL}{dZ_2} = \frac{dL}{dA_2} \cdot \sigma(Z_2)$ where $\sigma'(z) = \sigma(z)(1-\sigma(z))$

3. $\frac{dL}{dW_2} = \frac{dL}{dZ_2} \cdot A_1^T$

4. $\frac{dL}{db_2} = \frac{dL}{dZ_2}$

5. $\frac{dL}{dA_1} = W_2^T \cdot \frac{dL}{dZ_2}$

6. $\frac{dL}{dZ_1} = \frac{dL}{dA_1} \cdot ReLU'(Z_1)$ where $ReLU'(x) = 1$ if $x > 0$ else 0

7. $\frac{dL}{dW_1} = \frac{dL}{dZ_1} \cdot X^T$

8. $\frac{dL}{db_1} = \frac{dL}{dZ_1}$

1. $\frac{dL}{dA_2} = 2(A_2 - Y) = 2(1-18) = 2(-17) = -34$

$A_2 = 1 \quad Y = 18$

$\frac{dL}{dA_2} = -34$

2. $\frac{dL}{dZ_2} = \frac{dL}{dA_2} \cdot \sigma(Z_2) = -34 \cdot 0 = 0$

$\frac{dL}{dZ_2} = 0$

$\sigma'(Z_2) = \sigma(Z_2)(1-\sigma(Z_2)) \Rightarrow \sigma'(18.3) = \sigma(18.3)(1-\sigma(18.3)) = 1 \cdot (1-1) = 0$

$\frac{dL}{dA_2} = -34$

3. $\frac{dL}{dW_2} = \frac{dL}{dZ_2} \cdot A_1^T = 0 \cdot [3.9 \; 12.1 \; 20.3] = 0$

$\frac{dL}{dZ_2} = 0 \qquad A_1^T = \begin{bmatrix} 3.9 \\ 12.1 \\ 20.3 \end{bmatrix}^T = [3.9 \; 12.1 \; 20.3] \qquad \frac{dL}{dW_2} = 0$

$\frac{dL}{db_2} = 0$

4. $\frac{dL}{db_2} = \frac{dL}{dZ_2} = 0.$

5. $\frac{dL}{dA_1} = W_2^T \cdot \frac{dL}{dz_2} \cdot \begin{bmatrix} 0.2 \\ 0.4 \\ 0.6 \end{bmatrix} \cdot 0 = 0$ $\quad \frac{dL}{dA_1} = 0$

$\frac{dL}{dz_2} = 0 \quad W_2^T \cdot [0.2 \ 0.4 \ 0.6]^T = \begin{bmatrix} 0.2 \\ 0.4 \\ 0.6 \end{bmatrix}$ İlişki Nöronlar

6. $\frac{dL}{dz_1} = \frac{dL}{dA_1} \cdot ReLU'(z_1) = 0 \cdot 1 = 0$ $\quad \frac{dL}{dz_1} = 0$

$ReLU'(z_1) = 1 \quad z_1 > 0$

7. $\frac{dL}{dW_1} = \frac{dL}{dz_1} \cdot x^T = 0$ $\quad \frac{dL}{dW_1} = 0$

$\frac{dL}{dz_1} = 0 \quad x^T = \begin{bmatrix} 20 \\ 8 \\ 4 \end{bmatrix} \cdot [20, 3, 4]$

8. $\frac{dL}{db_1} = \frac{dL}{z_1} = 0$ $\quad \frac{dL}{db_1} = 0.$

$\frac{dL}{z_1} = 0$