No
Date

Assignment 1 - Part B : Forward & Backward Propagation

Data Point : $x : [20\ ,\ 3\ ,\ 4]$

$y : 18$

Initial weights and biases : $W1 : \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \\ 0.7 & 0.8 & 0.9 \end{bmatrix}$ $b1 : \begin{bmatrix} 0.1 \\ 0.2 \\ 0.3 \end{bmatrix}$

$W2 : [0.2 \quad 0.4 \quad 0.6]$ $b2 : [0.5]$

Forward Propagation

- Step 1 : Compute $Z1 : W1 \times X + b1$

$Z1 : (0.1 \times 20) + (0.2 \times 3) + (0.3 \times 4) + 0.1 = 3.9$

$(0.4 \times 20) + (0.5 \times 3) + (0.6 \times 4) + 0.2 : 12.1$

$(0.7 \times 20) + (0.8 \times 3) + (0.9 \times 4) + 0.3 : 20.3$

$\rightarrow Z1 : \begin{bmatrix} 3.9 \\ 12.1 \\ 20.3 \end{bmatrix}$

- Step 2 : A1 : ReLU (Z1)

ReLU (z) = max (0, z)

All value > 0 → A1 : Z1

$\rightarrow A1 : \begin{bmatrix} 3.9 \\ 12.1 \\ 20.3 \end{bmatrix}$

- Step 3 : $Z2 : W2 \times A1 + b2$

$Z2 : (0.2 \times 3.9) + (0.4 \times 12.1) + (0.6 \times 20.3) + 0.5 : 18.3$

$\rightarrow Z2 : 18.3$

- Step 4 : A2 = Sigmoid (Z2) where $\sigma(x) : 1/(1+e^{(-x)})$

$\sigma(18.3) : 1/(1+e^{\wedge}(-18.3))$ → its almost 0

So : $1/(1+0) \approx 1$

$\rightarrow A2 \approx 1$

- Step 5 : Calculate Loss $= (A2 - Y)^2$ where $Y : 18$

$: (1 - 18)^2 = (-17)^2 : 289$

No
Date

Backward Propagation

1. $dL/dA2 = 2(A2 - y)$

   $= 2(1 - 18) = 2(-17) = -34$

2. $dL/dZ2 = dL/dA2 \times \text{Sigmoid}'(Z2)$ where $\text{sigmoid}'(z) = \sigma(z)(1-\sigma(z))$

   $\sigma(Z2) \approx 1 \rightarrow \text{Sigmoid}'(18.3) \approx 1 \times (1-1) = 0$

   $= dL/dZ2 \approx -34 \times 0 = \approx 0$

3. $dL/dW2 = dL/dZ2 \times A1^T$

   $\rightarrow dL/dZ2 \approx 0$

   $dL/dW2 \approx [0, 0, 0]$

4. $dL/db2 = dL/dZ2$

   $\rightarrow \approx 0$

5. $dL/dA1 = W2^T \times dL/dZ2$

   $W2^T = \begin{bmatrix} 0.2 \\ 0.4 \\ 0.6 \end{bmatrix}$ because $dL/dZ2 \approx 0$ $\rightarrow dL/dA1 \approx \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

6. $dL/dZ1 = dL/dA1 \times \text{ReLU}'(Z1)$ where $\text{ReLU}'(x) = 1$ if $x > 0$, else 0

   $\text{ReLU}'(Z1) = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ Since all $Z1 > 0$

   $dL/dA1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \rightarrow dL/dZ1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

7. $dL/dW1 = dL/dZ1 \times X^T$

   $\rightarrow dL/dZ1 \approx 0 \rightarrow dL/dW1 \approx \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

8. $dL/db1 = dL/dZ1$

   $\rightarrow dL/db1 \approx \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$