Date _____

- Question #2 :-

= $X = \begin{bmatrix} 0.34 \\ 0.55 \end{bmatrix}$

= $W_1 = \begin{bmatrix} 0.42 & 0.72 \\ 0 & 0.30 \\ 0.15 & 0.09 \end{bmatrix}$, $W_2 = [0.44 \quad 0.03 \quad 0.55]$

= $W_3 = \begin{bmatrix} 0.55 \\ 0.71 \\ 0.29 \end{bmatrix}$, $W_4 = \begin{bmatrix} 0.97 & 0.55 & 0.9 \\ 0.71 & 0.70 & 0.22 \end{bmatrix}$

- Forward propagation :

- L_1 :

= $Z_1 = \begin{bmatrix} 0.42 & 0.72 \\ 0 & 0.30 \\ 0.15 & 0.09 \end{bmatrix} \begin{bmatrix} 0.34 \\ 0.55 \end{bmatrix}$

= $Z_1 = \begin{bmatrix} 0.53 \\ 0.16 \\ 0.1005 \end{bmatrix} = a_1$

- L_2

= $Z_2 = [0.44 \quad 0.03 \quad 0.55] \begin{bmatrix} 0.53 \\ 0.16 \\ 0.1005 \end{bmatrix}$

= $Z_2 = 0.297 = a_2$

Date

- L3 :

$$z_3 = \begin{bmatrix} 0.55 \\ 0.71 \\ 0.29 \end{bmatrix} [0.297]$$

$$= z_3 = \begin{bmatrix} 0.163 \\ 0.211 \\ 0.086 \end{bmatrix} = a_3$$

- L4 :

$$= z_4 = \begin{bmatrix} 0.97 & 0.55 & 0.97 \\ 0.71 & 0.70 & 0.22 \end{bmatrix} \begin{bmatrix} 0.163 \\ 0.211 \\ 0.086 \end{bmatrix}$$

$$= z_4 = \begin{bmatrix} 0.358 \\ 0.282 \end{bmatrix} : a_4 = \hat{x}$$

- Error Reconstruction :-

$$= J = -\sum x_i \log(\hat{x}_i)$$

$$J = -(0.34 \ln(0.358) + 0.55 \ln(0.282))$$

$J = 1.043$ is the reconstructed error

Backpropagation :-

$$d = \begin{bmatrix} 0.358 \\ 0.282 \end{bmatrix} - \begin{bmatrix} 0.34 \\ 0.55 \end{bmatrix}$$

$$= \begin{bmatrix} 0.018 \\ -0.267 \end{bmatrix}$$

unics

Date _____

- W_4 gradient:

$$\partial W_4 = \begin{bmatrix} 0.002 & 0.003 & 0.001 \\ -0.043 & -0.05 & -0.023 \end{bmatrix}$$

$$\delta_2 = (W_4)^T \delta_4 \odot g'(z_3)$$

Here, $g'(z_3) = 1$

$$(W_4)^T \delta_4 = \begin{bmatrix} 0.97 & 0.71 \\ 0.55 & 0.70 \\ 0.97 & 0.22 \end{bmatrix} \begin{bmatrix} 0.018 \\ -0.267 \end{bmatrix}$$

$$\delta_3 = \begin{bmatrix} -0.171 \\ -0.176 \\ -0.04 \end{bmatrix}$$

- W_3 gradient:

$$\partial W_3 = \delta_3 (a_2)^T$$

$$\partial W_3 = \begin{bmatrix} -0.051 \\ -0.052 \\ -0.012 \end{bmatrix}$$

$$\delta_2 = (W_3)^T \delta_3 \cdot g'(z_2)$$

$$\delta_2 = \begin{bmatrix} 0.55 & 0.71 & 0.29 \end{bmatrix} \begin{bmatrix} -0.171 \\ -0.176 \\ -0.040 \end{bmatrix}$$

$$\delta_2 = [-0.232]$$

- W_2 gradient:

$$\partial W_2 = \delta_2 (a_1)^T$$

Date

$dW_2 = [-0.125 \quad +0.038 \quad -0.023]$

$\delta_1 = (w_2)^T \delta_2 \odot g'(z_1)$

$\delta_1 = \begin{bmatrix} 0.44 \\ 0.03 \\ 0.55 \end{bmatrix} [-0.232]$

$\delta_1 = \begin{bmatrix} -0.102 \\ -0.006 \\ -0.127 \end{bmatrix}$

- w_1 gradient:

$dw_1 = \delta_1 (x)^T$

$dw_1 = \begin{bmatrix} -0.03 & -0.056 \\ -0.002 & -0.003 \\ -0.043 & -0.07 \end{bmatrix}$

- Weight updates:

$\therefore \alpha =$

- w_4:

$w_4' = \begin{bmatrix} 0.968 & 0.548 & 0.969 \\ 0.731 & 0.728 & 0.231 \end{bmatrix}$

- w_3:

$w_3' = \begin{bmatrix} 0.575 \\ 0.736 \\ 0.296 \end{bmatrix}$

Date _____

- w_2:

$w_2' = [0.502 \quad 0.049 \quad 0.561]$

- w_1:

$$w_1' = \begin{bmatrix} 0.937 & 0.748 \\ 0.001 & 0.301 \\ 0.171 & 0.125 \end{bmatrix}$$