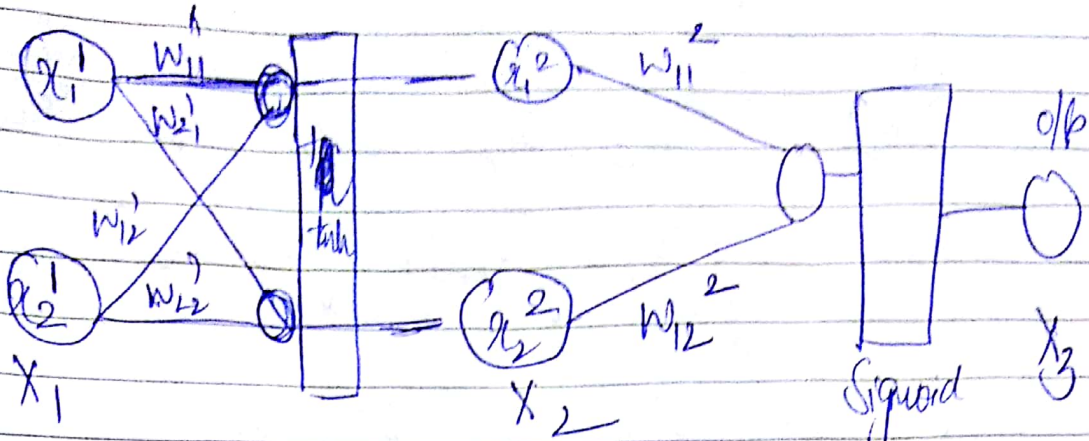
Ques - 18

DATE

Ans=1)



$$X_1 = \begin{bmatrix} x_1^1 \\ x_2^1 \end{bmatrix} \quad W_1 = \begin{bmatrix} w_{11}^1 & w_{12}^1 \\ w_{21}^1 & w_{22}^1 \end{bmatrix} \quad X_2 = \begin{bmatrix} x_1^2 \\ x_2^2 \end{bmatrix}$$

$$W_2 = \begin{bmatrix} w_{11}^2 \\ w_{12}^2 \end{bmatrix} \qquad X_3 = [x^3]$$

$X_2 = \tanh(W_1 X_1)$ Say $W_1 X_1 = Z_1$

So $X_2 = \tanh(Z_1)$ Also $Z_2 = W_2 X_2$

So $X_3 = \dfrac{1}{1+e^{-Z_2}} = \text{Sigmoid}(W_2 X_2)$

The loss is calculated as $L = \frac{1}{2}(y - X_3)^2$

Now we calculate the gradients as:

$$\frac{\partial L}{\partial w_{11}^2} = \frac{\partial L}{\partial X_3} \cdot \frac{\partial X_3}{\partial Z_2} \cdot \frac{\partial Z_2}{\partial w_{11}^2} = -(y - X_3) \cdot Sig(Z_2)(1 - Sig(Z_2)) \cdot x_1^2$$

$$\& \quad \frac{\partial L}{\partial w_{12}^2} = \frac{\partial L}{\partial X_3} \cdot \frac{\partial X_3}{\partial Z_2} \cdot \frac{\partial Z_2}{\partial w_{12}^2} = -(y - X_3) \cdot Sig(Z_2)(1 - Sig(Z_2)) \cdot x_2^2$$

$$\frac{\partial L}{\partial w_{11}^1} = \frac{\partial L}{\partial X_3} \cdot \frac{\partial X_3}{\partial Z_2} \cdot \frac{\partial Z_2}{\partial X_2} \cdot \frac{\partial X_2}{\partial Z_1} \cdot \frac{\partial Z_1}{\partial w_{11}^1}$$

$$= -(y - X_3) \cdot Sig(Z_2)(1 - Sig(Z_2)) \cdot W_2 \cdot (1 - \tanh^2(Z_1)) \cdot x_1^1$$

PAGE

$$\frac{\partial L}{\partial w_{11}^1} = \frac{\partial L}{\partial x_3} \cdot \frac{\partial x_3}{\partial z_2} \cdot \frac{\partial z_2}{\partial x_2} \cdot \frac{\partial x_2}{\partial z_1} \cdot \frac{\partial z_1}{\partial w_{11}^1}$$

$$= -(y - x_3) \cdot sig(z_2)(1 - sig(z_2)) w_2 (1 - \tanh^2(z_1)) \cdot x_1$$

$$\frac{\partial L}{\partial w_{21}^1} = \frac{\partial L}{\partial x_3} \cdot \frac{\partial x_3}{\partial z_2} \cdot \frac{\partial z_2}{\partial x_2} \cdot \frac{\partial x_2}{\partial z_1} \cdot \frac{\partial z_1}{\partial w_{21}^1}$$

$$= -(y - x_3) sig(z_2) [1 - sig(z_2)) \cdot w_2 \cdot (1 - \tanh^2(z_1)) \cdot x_1$$

$$\frac{\partial L}{\partial w_{12}^2} = \frac{\partial L}{\partial x_3} \cdot \frac{\partial x_3}{\partial z_2} \cdot \frac{\partial z_2}{\partial x_2} \cdot \frac{\partial x_2}{\partial z_1} \cdot \frac{\partial z_1}{\partial w_{12}^2}$$

$$= -(y - x_3) sig(z_2) (1 - sig(z_2)) \cdot w_2 (1 - \tanh^2(z_1)) \cdot x_1$$

So above we have derived the gradients. Now, to compute parameters we can use the rule -

$$w_{ij}^{k+1} \leftarrow w_{ij}^{k} - \eta \frac{\partial L}{\partial w_{ij}}$$

Here η is the learning rate.