1) a) First iteration considering sigmoid as activation fn

$$W^1 = \begin{bmatrix} W_{01}^1 & W_{11}^1 & W_{21}^1 \\ W_{02}^1 & W_{12}^1 & W_{22}^1 \end{bmatrix} = \begin{bmatrix} 0.3 & 0.1 & 0.2 \\ 0.2 & 0.1 & 0.2 \end{bmatrix}$$

$$W^2 = \begin{bmatrix} W_{01}^2 & W_{11}^2 & W_{21}^2 \\ W_{02}^2 & W_{12}^2 & W_{22}^2 \end{bmatrix} = \begin{bmatrix} 0.5 & 0.1 & 0.2 \\ 0.4 & 0.1 & 0.2 \end{bmatrix}$$

$$X = \begin{bmatrix} 1 \\ x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0.1 \\ 0.2 \end{bmatrix}, \quad t = \begin{bmatrix} 0.4 \\ 0.3 \end{bmatrix}$$

i) Feed forward pass [layer 0 to 1]

$$W^1 X = \begin{bmatrix} 0.3 & 0.1 & 0.2 \\ 0.2 & 0.1 & 0.2 \end{bmatrix} \begin{bmatrix} 1 \\ 0.1 \\ 0.2 \end{bmatrix} = \begin{bmatrix} 0.35 \\ 0.25 \end{bmatrix}$$

$$\sigma(W^1 X) = \sigma\left(\begin{bmatrix} 0.35 \\ 0.25 \end{bmatrix}\right) = \begin{bmatrix} 0.54 \\ 0.56 \end{bmatrix} = \begin{bmatrix} X_1^1 \\ X_2^1 \end{bmatrix}$$

ii) Feed forward pass from layer1 to Layer 2

$$W^2 X_i = \begin{bmatrix} 0.5 & 0.1 & 0.2 \\ 0.4 & 0.1 & 0.2 \end{bmatrix} \begin{bmatrix} 1 \\ 0.54 \\ 0.56 \end{bmatrix} = \begin{bmatrix} 0.67 \\ 0.57 \end{bmatrix}$$

$$\sigma(W^2 X_i^1) = \sigma\left(\begin{bmatrix} 0.67 \\ 0.57 \end{bmatrix}\right) = \begin{bmatrix} 0.66 \\ 0.64 \end{bmatrix} = \begin{bmatrix} X_1^2 \\ X_2^2 \end{bmatrix} = \begin{bmatrix} O_1^2 \\ O_2^2 \end{bmatrix}$$

Actual o/p = $\begin{bmatrix} 0.66 \\ 0.64 \end{bmatrix}$, Target o/p = $\begin{bmatrix} 0.4 \\ 0.3 \end{bmatrix}$

iii) Back propogation from o/p layer to layer 1

We know that

$$\delta_j^2 = (x_j^2 - t_j) \cdot x_j^2 \cdot (1 - x_j)^2$$

Substituting it :

$$\delta_1^2 = (0.66 - 0.4)(0.66)(1 - 0.66) = 0.058$$

$$\delta_2^2 = (0.64 - 0.3)(0.64)(1 - 0.64) = 0.078$$

Hence

$$\frac{\partial E}{\partial w^2} = \begin{bmatrix} 0.058 \\ 0.078 \end{bmatrix} \begin{bmatrix} 1 & 0.59 & 0.56 \end{bmatrix} = \begin{bmatrix} 0.058 & 0.034 & 0.032 \\ 0.078 & 0.043 & 0.044 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{\partial E}{\partial w_{01}^2} & \frac{\partial E}{\partial w_{11}^2} & \frac{\partial E}{\partial w_{21}^2} \\ \frac{\partial E}{\partial w_{02}^2} & \frac{\partial E}{\partial w_{12}^2} & \frac{\partial E}{\partial w_{22}^2} \end{bmatrix}$$

iv) Back Propagation from Layer 1 to Layer 0

Consider

$$\partial_i = x_i'(1 - x_i') \sum_{j=1}^{2} \partial_j w_{ij}^2$$

$$\frac{\partial E}{\partial w_1} = \begin{bmatrix} \frac{\partial E}{\partial w_{01}^1} & \frac{\partial E}{\partial w_{11}^1} & \frac{\partial E}{\partial w_{21}^1} \\ \frac{\partial E}{\partial w_{02}^1} & \frac{\partial E}{\partial w_{12}^1} & \frac{\partial E}{\partial w_{22}^1} \end{bmatrix} = \begin{bmatrix} 0.003 & 0.0003 & 0.0006 \\ 0.007 & 0.0007 & 0.0013 \end{bmatrix}$$

Hence

~~w_{01}~~ $w_{01}' \leftarrow w_{01}' - 0.5 \frac{\partial E_1}{\partial w_{01}'} = 0.296$ $\quad w_{12}' \leftarrow w_{12}' - 0.5 \frac{\partial E}{\partial w_{12}'} = 0.0993$

$w_{02}' \leftarrow w_{02}' - 0.5 \frac{\partial E}{\partial w_{02}'} = 0.193$ $\quad w_{21}' \leftarrow w_{21}' - 0.5 \frac{\partial E}{\partial w_{21}'} = 0.1993$

$w_{11}' \leftarrow w_{11}' - 0.5 \frac{\partial E}{\partial w_{11}'} = 0.0996$ $\quad w_{22}' \leftarrow w_{22}' - 0.5 \frac{\partial E}{\partial w_{22}'} = 0.1986$

Parallely,

$$W_{01}^2 \leftarrow W_{01}^2 - 0.5\frac{\partial E}{\partial W_{01}^2} = 0.4414$$

$$W_{12}^2 \leftarrow W_{12}^2 - 0.5\frac{\partial E}{\partial W_{12}^2} = 0.0541$$

$$W_{02}^2 \leftarrow W_{02}^2 - 0.5\frac{\partial E}{\partial W_{02}^2} = 0.3218$$

$$W_{21}^2 \leftarrow W_{21}^2 - 0.5\frac{\partial E}{\partial W_{21}^2} = 0.1671$$

$$W_{11}^2 \leftarrow W_{11}^2 - 0.5\frac{\partial E}{W_{11}^2} = 0.0656$$

$$W_{22}^2 \leftarrow W_{22}^2 - 0.5\frac{\partial E}{\partial W_{22}^2} = 0.1560$$