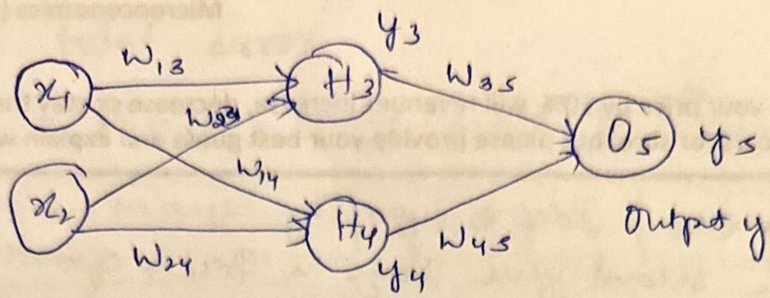
Step 8 :



Given inputs :

$x_1 = 0.35$

$x_2 = 0.9$

$\eta = 1$

Output / Target values :

$y = 0.5$

Weights :

$w_{13} = 0.1$

$w_{14} = 0.4$

$w_{23} = 0.8$

$w_{24} = 0.6$

$w_{35} = 0.3$

$w_{45} = 0.9$

Activation function is logistic function
Given by sigmoid curve $= \dfrac{1}{1+e^{-x}}$

Forward pass : compute output y_3, y_4 and y_5

$$a_j = \sum_j (w_{i,j} * x_i) \qquad y_j = F(a_j) = \frac{1}{1+e^{-a_j}}$$

$$a_1 = (w_{13} * x_1) + (w_{23} * x_2)$$
$$= (0.1 * 0.35) + (0.8 * 0.9) = 0.755$$

$$y_3 = f(a_1) = \frac{1}{1+e^{-0.755}} = \underline{0.68}$$

$$a_2 = (w_{14} * x_1) + (w_{24} * x_2)$$
$$= (0.4 * 0.35) + (0.6 * 0.9) = 0.68$$

$$y_4 = f(a_2) = \frac{1}{1+e^{-0.68}} = \underline{0.6637}$$

$$a_3 = (w_{35} * y_3) + (w_{45} * y_4)$$
$$= (0.3 * 0.68) + (0.9 * 0.6637) = 0.801$$

$$y_5 = f(a_3) = \frac{1}{1+e^{-0.801}} = \underline{0.69} \rightarrow \text{output of network}$$

Calculating total error

$$\text{Error} = MSE = \frac{1}{N} \sum (y - \hat{y})^2$$

or $\text{Error} = MAE = |y - \hat{y}| = |0.5 - 0.69| = \underline{0.19}$

$$\boxed{E_{total} = 0.19}$$

<u>Backward pass</u>

$$\Delta w_{ji} = \eta \, \delta_j \, O_i$$

δ_j → partial derivative

O_i → output at that node

Δw_{ji} → weight update

η → learning rate

$$\delta_j = O_j (1 - O_j)(t_j - O_j)$$

If j is an output unit

$$\delta_j = O_j (1 - O_j) \sum_k \delta_k w_{kj}$$

If j is an hidden unit

Lets calculate partial derivatives for H_3, H_4 & O_5

i.e $\delta_3, \delta_4, \delta_5$

For output unit O_5:

$$\delta_5 = y(1-y)(y_{target} - y)$$

$$= 0.69 * (1 - 0.69) * (0.5 - 0.69)$$

$$\boxed{\delta_5 = -0.0406}$$

For hidden unit:

$$\delta_3 = y_3(1-y_3) \, w_{35} * \delta_5$$

$$= 0.68 * (1 - 0.68) * (0.3 * -0.0406)$$

$$\boxed{\delta_3 = 0.00265}$$

$\delta_4 = y_4 (1-y_4) W_{45} * \delta_5$

$= 0.6637 * (1-0.6637) * (0.9 * -0.0406)$

$\boxed{\delta_4 = -0.0082}$

Since we have partial derivatives

Let us compute new weights

$\Delta W_{ji} = \eta \delta_j O_i$

$\Delta W_{45} = \eta \delta_5 y_4$

$= 1 * -0.0406 * 0.6637$

$\Delta W_{45} = -0.0269$

$W_{45}(new) = \Delta W_{45} + W_{45}(old)$

$= -0.0269 + 0.9$

$\boxed{W_{45}(new) = 0.8731}$

$\Delta W_{14} = \eta \delta_4 x_1 = 1 * -0.0082 * 0.35$

$\Delta W_{14} = -0.00287$

$W_{14}(new) = \Delta W_{14} + W_{14}(old)$

$= -0.00287 + 0.4$

$W_{14}(new) = -0.00287 + 0.4 = 0.3971$

$\boxed{W_{14}(new) = 0.3971}$

$\Delta W_{13} = \eta \delta_3 x_1 = 1 * (-0.00265) * 0.35$

$W_{13}(new) = \Delta W_{13} + W_{13}(old) = [1 * (-0.00265) * 0.35] + 0.1 = 0.0991$

$W_{23}(new) = \Delta W_{23} + W_{23}(old) = [1 * (-0.00265) * 0.9] + 0.8 = 0.7976$

$$W_{24(new)} = \Delta W_{24} + W_{24(old)} = (1 * (-0.0082) * 0.4) + 0.6$$

$$\boxed{W_{24(new)} = 0.5926}$$

$$W_{35(new)} = \Delta W_{35} + W_{35(old)} = (1 * (-0.0406) * 0.68) + 0.3$$

$$\boxed{W_{35(new)} = 0.2724}$$

Thus, weight are updated after backward pass and this loop goes on until a minima is reached