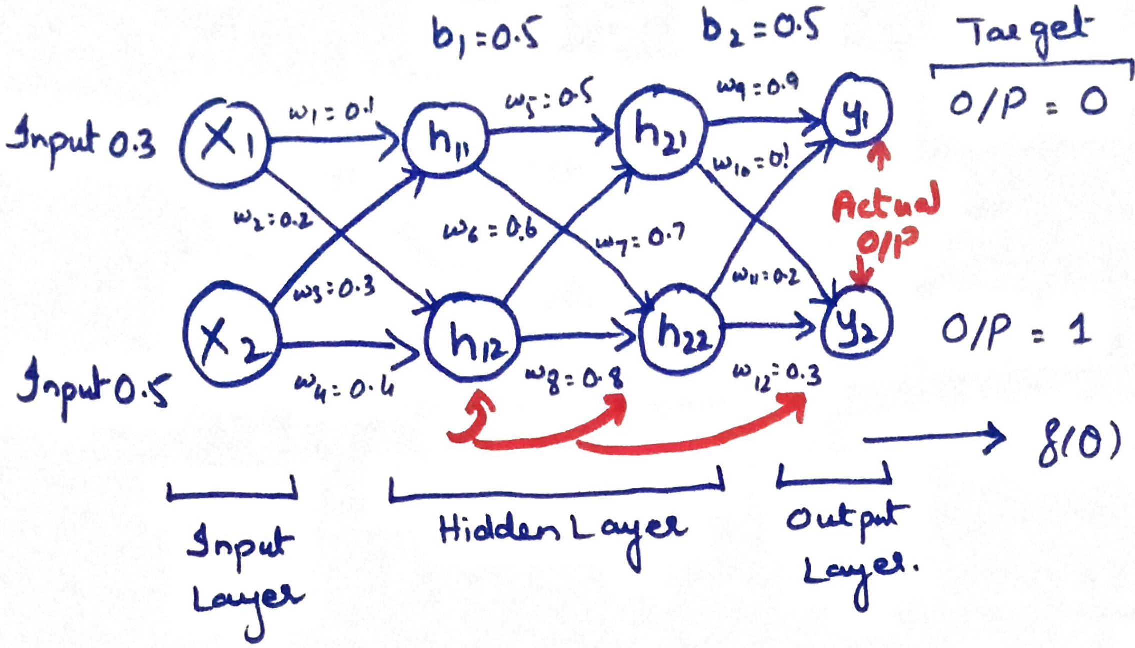
FEED FORWARD NEURAL NETWORK



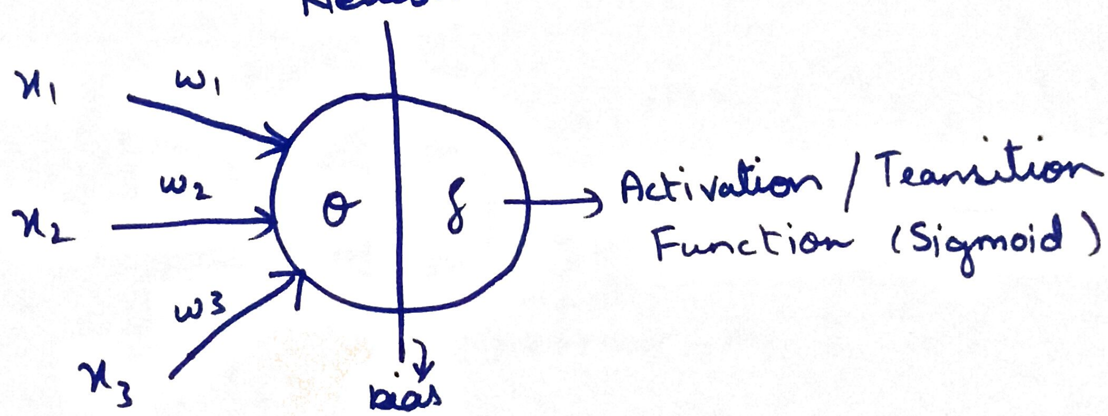
$$f(\theta) = \frac{1}{1+e^{-\theta}}$$

Overall I/P of Neuron can be calculated by the dot product of Input & Weight

$\boxed{+}$ **bias** b_1

$$\theta = w_1 x_1 + w_2 x_2 + w_3 x_3 + b \qquad \text{(Input)}$$

Neuron



$$f = \frac{1}{1+e^{-\theta}} \qquad \text{(Output)}$$

$$h_{11} = w_1 \cdot x_1 + w_3 x_2 + b_1$$
$$= 0.1 \times 0.3 + 0.3 \times 0.5 + 0.5$$
$$= 0.03 + 0.15 + 0.5$$
$$= 0.68$$

$$\delta(h_{11}) = \delta(0.68) = \frac{1}{1+e^{-0.68}} = 0.66$$

$$h_{12} = w_2 x_1 + w_4 x_2 + b_1$$
$$= 0.2 \times 0.3 + 0.4 \times 0.5 + 0.5$$
$$= 0.06 + 0.2 + 0.5$$
$$= 0.76$$

$$\delta(h_{12}) = \delta(0.78) = \frac{1}{1+e^{-0.78}} = 0.68$$

Input h_{21}

$$h_{21} = w_5 \cdot \delta(h_{11}) + w_7 \cdot \delta(h_{12}) + b_2$$
$$= 0.5 \times 0.66 + 0.7 \times 0.68 + 0.5$$
$$= 0.33 + 0.47 + 0.5$$
$$= 1.306$$

$$\delta(h_{21}) = 0.786$$

$$h_{22} = w_6 \cdot \delta(h_{11}) + w_8 \cdot \delta(h_{12}) + b_2$$
$$= 0.6 \times 0.66 + 0.8 \times 0.76 + 0.5$$
$$= 1.504$$

$$\delta(h_{22}) = 0.818$$

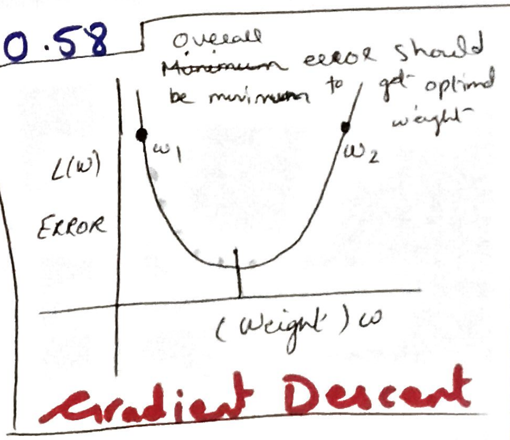
Calculate For Output Layer

$$y_1 = w_9 \cdot \delta(h_{22}) + w_{11} \cdot (h_{22})$$
$$= 0.3 \times 0.786 + 0.2 \times 0.818$$
$$= 0.1636$$

$$\delta(y_1) = 0.54$$

$$y_2 = w_{10}\, \delta(h_{21}) + w_{12}(h_{22})$$
$$= 0.1 \times 0.786 + 0.3 \times 0.818$$
$$= 0.324$$

$$\delta(y_2) = 0.58$$



ERROR

E/L = MEAN SQUARE ERROR.

$$\frac{1}{2}\left[(Y_A^1 - Y_T^1)^2 + (Y_A^2 - Y_T^2)^2\right]$$

$$= \frac{1}{2}\left[(0.54-0)^2 + (0.58-1)^2\right] = \frac{1}{2}(0.2196 + 0.1764)$$

$$= \boxed{0.234} \text{ (loss)}$$

BACK PROPAGATION NEURAL NETWORK

LEARNING RATE (η) = 0.1

TOTAL ERROR/LOSS (L) = 0.234

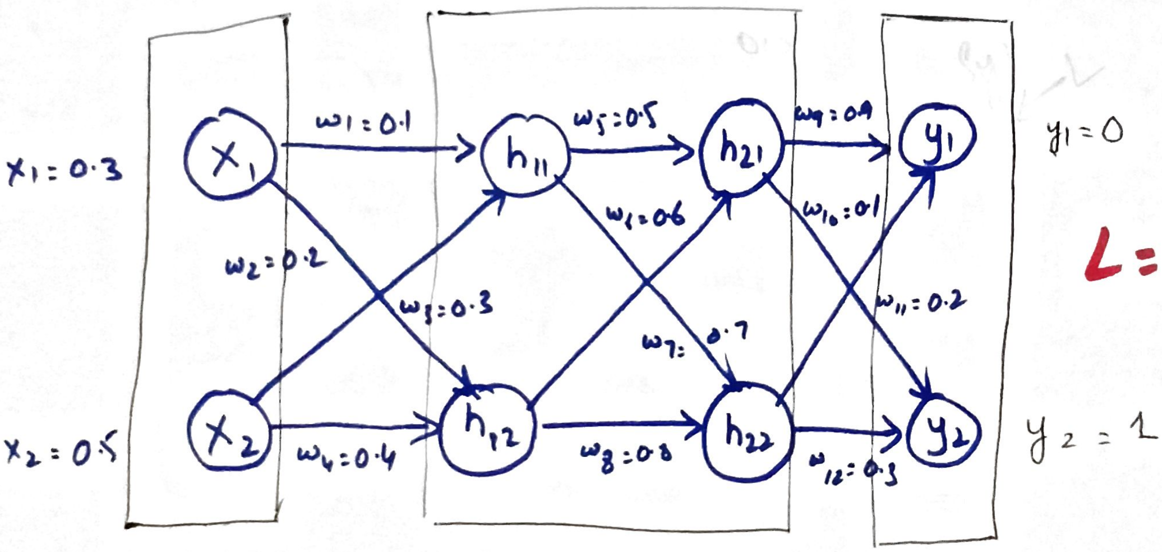
Re-Calculate the weights

$$w_{new} = w_{old} - \text{step size}$$

$$= w_{old} - \text{learning rate} \times \text{slope}$$

$$= w_{old} - \eta \times \frac{\partial L(w)}{\partial w} \rightarrow \text{ERROR}$$

Learning Rate   Partial Derivative



L = 0.23

L1 = $w_9, w_{10}, w_{11}, w_{12}$

L2 = $w_5 - w_8$

L3 = $w_1 - w_4$

$h_{11} = 0.68$ $\delta(h_{11}) = 0.66$

$h_{12} = 0.76$ $\delta(h_{12}) = 0.68$

$h_{21} = 1.306$ $\delta(h_{21}) = 0.786$

$h_{22} = 1.504$ $\delta(h_{22}) = 0.818$

BACK PROPAGATING (Thinner NN)

$$\frac{\partial L}{\partial w_{12}} = w_{12} \rightarrow y_2 \rightarrow \sigma(y_1) \rightarrow L$$

Chain Rule

$$\frac{\partial L}{\partial \sigma(y_2)} \cdot \frac{\partial \sigma(y_2)}{\partial (y_2)} \cdot \frac{\partial y_2}{\partial w_{12}} \Rightarrow -0.42 \times 0.58 \times 0.42 \times 0.818$$

$$= \boxed{-0.0836} \; \frac{\partial L}{\partial w_{12}}$$

$$\frac{\partial L}{\partial \sigma(y_2)} = \cancel{2} \cdot \frac{1}{\cancel{2}} (\sigma(y_2) - O_2) = \sigma(y_2) - O_2 = 0.58 - 1 = -0.42$$

$$\frac{\partial \sigma(y_2)}{\partial (y_2)} = \sigma(y_2) \cdot (1 - \sigma(y_2)) \quad \boxed{\sigma(x) \cdot (1 - \sigma(x))}$$

$$0.58 \cdot (1 - 0.58) = 0.58 \times 0.42 =$$

$$\frac{\partial (y_2)}{\partial w_{12}} = \sigma(h_{22}) = \boxed{\frac{\partial \left(w_{12} \cdot \sigma(h_{12}) + w_{10} \cdot \sigma(h_{21}) + b_2 \right)}{\partial w_{12}}}$$

$$= 0.818$$

$$w_{NEW} = w_{old} - \eta \times \frac{\partial L(w)}{\partial w}$$

$$w_{12}' = w_{12} - \eta \cdot \frac{\partial L}{\partial w_{12}}$$

$$= 0.3 - 0.1 \times (-0.0836)$$

$$= 0.3 + 0.00836$$

$$w_{12}' \quad \boxed{= 0.30836}$$

$$\frac{\partial L}{\partial w_{11}} = w_{11} \rightarrow y_1 \rightarrow \sigma y_1 \rightarrow L$$

$$\underset{0.54}{\frac{\partial L}{\partial \sigma(y_1)}} \cdot \underset{0.54 \times 0.46}{\frac{\partial \sigma(y_1)}{\partial (y_1)}} \cdot \frac{\partial y_1}{\partial w_{11}} = 0.1097 \Rightarrow w_{11}' = \boxed{0.189} \; w_{11}'$$

$$w_{10}' = \underset{P.D}{-0.084} \Rightarrow \boxed{0.108} \; w_{10}'$$

$$w_9' = \underset{P.D}{0.1054} \Rightarrow \boxed{0.89} \; w_9'$$

$$\frac{\partial L}{\partial w_5} = \begin{tikzcd} w_6 \arrow[r] & h_{21} \arrow[r] & \sigma h_{21} \arrow[d, dash] \arrow[dr, dash] & \\ & & y_1 \arrow[d, dash] & y_2 \arrow[d, dash] \\ & & \sigma y_1 \arrow[d, dash] & \sigma y_2 \arrow[d, dash] \\ & & L & L \end{tikzcd}$$

$$= \frac{\partial L}{\partial \sigma(h_{21})} \cdot \frac{\partial \sigma(h_{21})}{\partial(h_{21})} \cdot \frac{\partial h_{21}}{\partial w_5}$$

$0.1105 \cdot 0.1682 \cdot$

$\boxed{0.0122}$

$$\frac{\partial L}{\partial \sigma(h_{21})} = \left[\frac{\partial L}{\partial \sigma(y_1)} \cdot \frac{\partial \sigma(y_1)}{\partial(y_1)} \cdot \frac{\partial y_1}{\partial \sigma h_{21}}\right] + \left[\frac{\partial L}{\partial \sigma(y_2)} \cdot \frac{\partial \sigma(y_2)}{\partial(y_2)} \cdot \frac{\partial y_2}{\partial \sigma h_{21}}\right]$$

$$= 0.54 \times 0.54 \times 0.46 \times 0.9 + (-0.42) \times 0.58 \times 0.42 \times 0.1$$

$$= 0.1207 + (-0.0102) = \boxed{0.1105} \ \frac{\partial L}{\partial \sigma(h_{21})}$$

$$\frac{\partial y_1}{\partial \sigma h_{21}} = \frac{\partial^{\text{delta}}}{\partial \sigma(h_{21})} = \left[w_9 \cdot \sigma(h_{21}) + w_{11} \cdot \sigma(h_{22}) + b_2\right]$$

$$= w_9 = \boxed{0.9}$$

$$\frac{\partial y_2}{\partial \sigma(h_{21})} = \frac{\partial}{\partial \sigma(h_{21})} = \left[w_{10} \cdot \sigma(h_{21}) + w_{12}\, \sigma(h_{22}) + b_2\right]$$

$$w_{10} = \boxed{0.1}$$

$$w_5' = w_5 - \eta \frac{\partial L}{\partial w_5} = 0.5 - 0.1 \times 0.0122 = \boxed{0.499} \ w_5'$$

$$w_6' = -\underset{\text{PD}}{0.0003} = \boxed{0.60003} \ w_6'$$

$$\frac{\partial L}{\partial w_1} = w_1 \rightarrow h_{11} \rightarrow \sigma(h_{11})$$

$$\frac{\partial L}{\partial \sigma(h_{11})} \cdot \frac{\partial \sigma(h_{11})}{\partial (h_{11})} \cdot \frac{\partial h_{11}}{\partial w_1} = 0.0089 \times 0.66 \times 0.34 \times 0.3$$

$$\boxed{0.00059} \quad \frac{\partial L}{\partial w_1}$$

$$\frac{\partial L}{\partial \sigma(h_{11})} = \left[\frac{\partial L}{\partial \sigma(h_{21})} \cdot \frac{\partial \sigma(h_{21})}{\partial (h_{21})} \cdot \frac{\partial h_{21}}{\partial \sigma(h_{11})}\right] + \left[\frac{\partial L}{\partial \sigma(h_{22})} \cdot \frac{\partial \sigma(h_{22})}{\partial h_{22}} \cdot \frac{\partial h_{22}}{\partial \sigma(h_{22}}\right]$$

$$= (0.115 \times 0.1682 \times 0.5) + (-0.0038 \times 0.1488 \times 0.6)$$

$$= \boxed{0.0089} \quad \frac{\partial L}{\partial \sigma h_{11}}$$

$$\frac{\partial h_{21}}{\partial \sigma h_{11}} = w_5 = 0.5 \quad \Big| \quad \frac{\partial}{\partial \sigma(h_{11})}\left[w_5 \cdot \sigma(h_{11}) + w_7 \cdot \sigma(h_{12}) + b_1\right]$$

$$\frac{\partial h_{22}}{\partial \sigma(h_{22})} = w_6 = 0.6 \quad \Big| \quad \frac{\partial}{\partial \sigma(h_{11})}\left[w_6 \cdot \sigma(h_{11}) + w_8\, \sigma(h_{12}) + b_1\right]$$

$$\frac{\partial \sigma(h_{11})}{\partial (h_{11})} = \sigma(h_{11}) \cdot (1 - \sigma(h_{11})) = 0.66(1 - 0.66) = 0.66 \times 0.34$$

$$\frac{\partial h_{11}}{\partial w_1} = \frac{\partial}{\partial w_1}(w_1 \cdot x_1 + w_3 x_2] = x_1 = 0.3$$

$$w_1' = w_1 - \eta \frac{\partial L}{\partial w_1} = 0.1 - 0.1 \times 0.00059$$

$$= \boxed{0.099} \quad w_1'$$