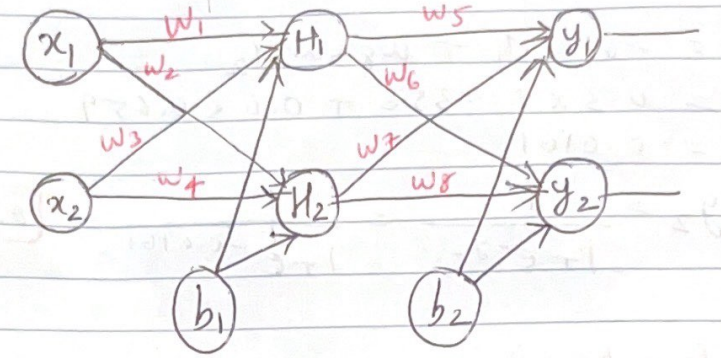
ASSIGNMENT 2 : (STEP 8.)

Name: SHRUTI BADRINARAYANAN



Weight Initialisation:

$\alpha = 0.5$

$w_1 = 0.5$
$w_2 = 0.7$
$w_3 = 0.95$
$w_4 = 0.1$
$w_5 = 0.25$
$w_6 = 0.3$
$w_7 = 0.5$
$w_8 = 0.6$

$x_1 = 0.85$
$x_2 = 0.63$

$b_1 = 0$
$b_2 = 0$

(Activation Fn = sigmoid $= \frac{1}{1+e^{-x}}$)

$y_1 = 0.8$
$y_2 = 0.3$

Forward Pass:

$$H_1 = w_1 x_1 + w_3 x_2 = 0.5 \times 0.85 + 0.95 \times 0.63 = 1.0235$$

$$\text{Out}\, H_1 = \frac{1}{1+e^{-H_1}} \overset{+b_1}{=} \frac{1}{1+e^{-1.0235}} = 0.7356$$

$$H_2 = w_2 x_1 + w_4 x_2 = 0.7 \times 0.85 + 0.1 \times 0.63 = 0.658$$

$$\text{Out}\, H_2 = \frac{1}{1+e^{-H_2}} \overset{+b_1}{=} \frac{1}{1+e^{-0.658}} = 0.659$$

$$y_1 = w_5\, \text{Out}\, H_1 + w_7\, \text{Out}\, H_2$$
$$= 0.25 \times 0.7356 + 0.5 \times 0.659 = 0.5134$$

$$\text{Out}\, y_1 = \frac{1}{1+e^{-y_1}} = \frac{1}{1+e^{-0.5134}} = \boxed{0.6256}$$

$$y_2 = w_6 \text{out}H_1 + w_8 \text{out}H_2$$
$$= 0.3 \times 0.7356 + 0.6 \times 0.659$$
$$= 0.6161$$
$$\text{out}\,y_2 = \frac{1}{1+e^{-y_2}} = \frac{1}{1+e^{-0.6161}} = \boxed{0.6493}$$

Loss Function:

Using MSE loss :

$$L = \text{MSE Loss} = \frac{1}{2}\Sigma(y-\hat{y})^2$$
$$= \frac{1}{2}(0.8-0.6256)^2 + \frac{1}{2}(0.3-0.6493)^2$$
$$= 0.0762$$

Backward Pass :

(Note $\Rightarrow \sigma'(x) = \sigma(x)(1-\sigma(x))$)

(i) $$w_5 = w_5 - \alpha \cdot \frac{\partial L}{\partial w_5}$$

$$\frac{\partial L}{\partial w_5} = \frac{\partial L}{\partial \text{out}\,y_1} \times \frac{\partial \text{out}\,y_1}{\partial y_1} \times \frac{\partial y_1}{\partial w_5}$$
$$= (0.8-0.6256) \times (0.6256)(1-0.6256) \times (0.7356)$$
$$= 0.03$$

$$\Rightarrow w_5 = 0.25 - (0.03) \times (0.5) = 0.235$$

(ii) $$w_6 = w_6 - \alpha \frac{\partial L}{\partial w_6}$$

$$\frac{\partial L}{\partial w_6} = \frac{\partial L}{\partial \text{out}\,y_2} \times \frac{\partial \text{out}\,y_2}{\partial y_2} \times \frac{\partial y_2}{\partial w_6}$$

$= (0.3 - 0.6493) \times (0.6493)(1 - 0.6493) \times (0.7356)$

$= -0.058$

$\Rightarrow w_6 = 0.3 - 0.5 \times (-0.058)$

$= 0.3293$

(iii) $w_7 = w_7 - \alpha \frac{\partial L}{\partial w_7}$

$\frac{\partial L}{\partial w_7} = \frac{\partial L}{\partial out_{y_1}} \times \frac{\partial out_{y_1}}{\partial y_1} \times \frac{\partial y_1}{\partial w_7}$

$= (0.8 - 0.6256) \times (0.6256)(1 - 0.6256) \times 0.659$

$= 0.0269$

$\Rightarrow w_7 = 0.5 - 0.5 \times 0.0269 = 0.4865$

(iv) $w_8 = w_8 - \alpha \frac{\partial L}{\partial w_8}$

$\frac{\partial L}{\partial w_8} = \frac{\partial L}{\partial out_{y_2}} \times \frac{\partial out_{y_2}}{\partial y_2} \times \frac{\partial y_2}{\partial w_8}$

$= (0.3 - 0.6493) \times (0.6493)(1 - 0.6493) \times 0.659$

$= -0.0524$

$\Rightarrow w_8 = 0.6 - 0.5 \times (-0.0524) = 0.6262$

(v) $w_1 = w_1 - \alpha \frac{\partial L}{\partial w_1}$

$\frac{\partial L}{\partial w_1} = \frac{\partial L}{\partial out_{y_1}} \times \frac{\partial out_{y_1}}{\partial y_1} \frac{\partial y_1}{\partial out_{H_1}} \frac{\partial out_{H_1}}{\partial H_1} \frac{\partial H_1}{\partial w_1}$

$+ \frac{\partial L}{\partial out_{y_2}} \frac{\partial out_{y_2}}{\partial y_2} \frac{\partial y_2}{\partial out_{H_1}} \frac{\partial out_{H_1}}{\partial H_1} \frac{\partial H_1}{\partial w_1}$

$$\frac{\partial L}{\partial w_1} = \Big[(0.8-0.6256)\big[(0.6256)(1-0.6256)\big](0.25) + (0.3-0.6493)\big[(0.6493)(1-0.6493)\big](0.3)\Big] \times (0.7356)(1-0.7356)(0.85)$$

$$= -0.0022$$

$$\Rightarrow w_1 = 0.5 - 0.5 \times (-0.0022) = \underline{0.5011}$$

(vi) $$w_2 = w_2 - \alpha \frac{\partial L}{\partial w_2}$$

$$\frac{\partial L}{\partial w_2} = \left[\frac{\partial L}{\partial out_{y_1}} \frac{\partial out_{y_1}}{\partial y_1} \frac{\partial y_1}{\partial out_{H_2}} + \frac{\partial L}{\partial out_{y_2}} \frac{\partial out_{y_2}}{\partial y_2} \frac{\partial y_2}{\partial out_{H_2}}\right] \times \frac{\partial out_{H_2}}{\partial H_2} \frac{\partial H_2}{\partial w_2}$$

$$= \Big[(0.8-0.6256)\big[(0.6256)(1-0.6256)\big](0.5) + (0.3-0.6493)\big[(0.6493)(1-0.6493)\big](0.6)\Big] \times (0.659)(1-0.659)(0.85)$$

$$= -0.0052$$

$$\Rightarrow w_2 = 0.7 - 0.5 \times (-0.0052) = \underline{0.7026}$$

(vii) $$w_3 = w_3 - \alpha \frac{\partial L}{\partial w_3}$$

$$\frac{\partial L}{\partial w_3} = \left[\frac{\partial L}{\partial out_{y_1}} \frac{\partial out_{y_1}}{\partial y_1} \frac{\partial y_1}{\partial out_{H_1}} + \frac{\partial L}{\partial out_{y_2}} \frac{\partial out_{y_2}}{\partial y_2} \frac{\partial y_2}{\partial out_{H_1}}\right] \times \frac{\partial out_{H_1}}{\partial H_1} \times \frac{\partial H_1}{\partial w_3}$$

$$= [(0.8-0.6256)[(0.6256)(1-0.6256)](0.25)$$
$$+(0.3-0.6493)[(0.6493)(1-0.6493)](0.3)]$$
$$\times (0.7356)(1-0.7356)(0.63)$$
$$= -0.0017$$

$$\Rightarrow w_3 = 0.95 - 0.5(-0.0017) = \underline{0.9508}$$

(viii) $$w_4 = w_4 - \alpha \frac{\partial L}{\partial w_4}$$

$$\frac{\partial L}{\partial w_4} = \left[\frac{\partial L}{\partial out_{y_1}} \frac{\partial out_{y_1}}{\partial y_1} \frac{\partial y_1}{\partial out_{H_2}} + \frac{\partial L}{\partial out_{y_2}} \frac{\partial out_{y_2}}{\partial y_2} \frac{\partial y_2}{\partial out_{H_2}}\right]$$
$$\times \frac{\partial out_{H_2}}{\partial H_2} \times \frac{\partial H_2}{\partial w_4}$$
$$= [(0.8-0.6256)[(0.6256)(1-0.6256)](0.5)$$
$$+(0.3-0.6493)[(0.6493)(1-0.6493)](0.6)]$$
$$\times (0.659)(1-0.659) \times (0.63)$$
$$= -0.0039$$

$$\Rightarrow w_4 = 0.1 - 0.5(-0.0039) = \underline{0.1019}$$