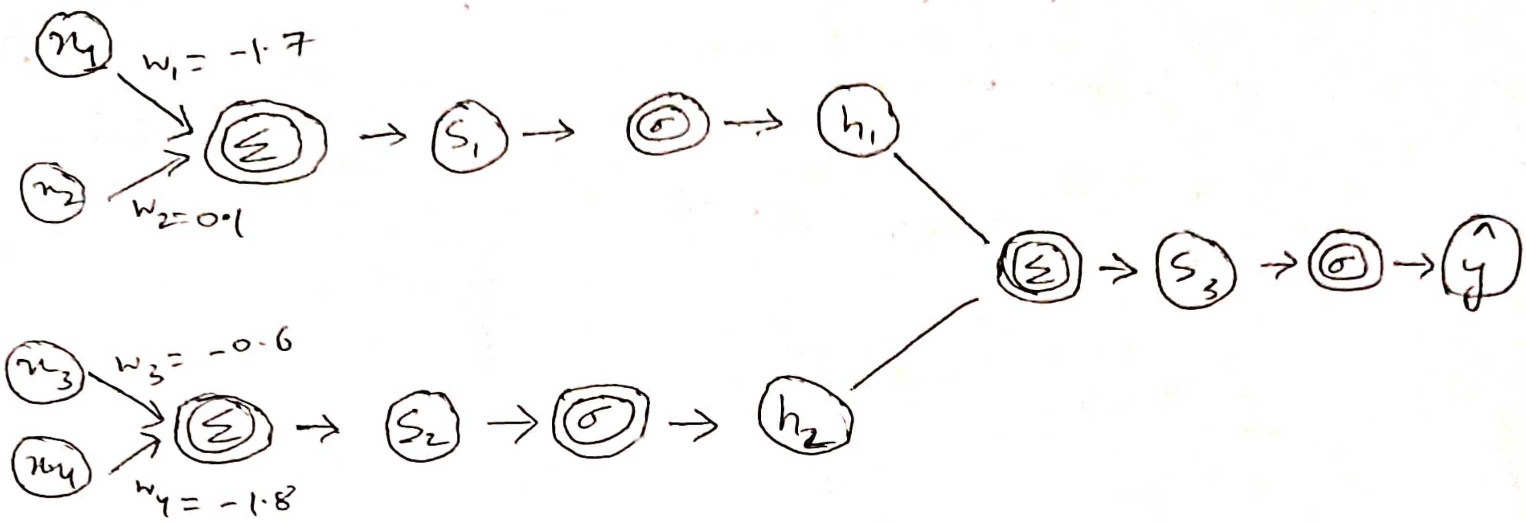
DEEP LEARNING

Q1)



Answer

Inputs $= (x_1, x_2, x_3, x_4) = (0.7, 1.2, 1.1, 2)$

weights $=$ $w_1 = -1.7$ $\quad$ $w_4 = -1.8$

$w_2 = 0.1$ $\quad$ $w_5 = -0.2$

$w_3 = -0.6$ $\quad$ $w_6 = 0.5$

Activation function $\sigma(z) = \dfrac{1}{1+e^{-z}}$

Hidden layer, $h_1 = \dfrac{1}{1+e^{-w_1 \cdot x_1 - w_2 x_2}}$

Loss function $L(y, \hat{y}) = \|\hat{y} - y\|^2$

$\dfrac{\partial L}{\partial w_1} = ?$

$y = 0.5$

$$s_1 = x_1 \cdot w_1 + x_2 w_2$$
$$= 0.7(-1.7) + (1.2)(0.1)$$
$$= (-1.19) + (0.12)$$
$$= -1.07$$
$$\boxed{\therefore s_1 = -1.07}$$

$$s_2 = x_3 w_3 + x_4 w_4$$
$$= (1.1)(-0.6) + (2)(-1.8)$$
$$= (-0.66) + (-3.6)$$
$$\boxed{\therefore s_2 = -4.26}$$

Now, $h_1 = \frac{1}{1+e^{-wx_1 - x_2 w_2}}$

$$h_1 = \frac{1}{1+e^{-(0.7)(-1.7)-(1.2)(0.1)}}$$

$$h_1 = \frac{1}{1+e^{(1.19 - 0.12)}}$$

$$h_1 = \frac{1}{1+e^{1.07}} = h_1 = \frac{1}{3.915}$$

$$\boxed{h_1 = 0.255}$$

$$h_2 = \frac{1}{1+e^{-x_3 w_3 - x_4 w_4}}$$

$$= \frac{1}{1+e^{-(0.1)(-0.6)-(2)(-1.8)}}$$

$$= \frac{1}{1+e^{0.66+3.6}}$$

$$= \frac{1}{1+e^{4.26}} = \frac{1}{1+70.5}$$

$$\boxed{h_2 = 0.0139}$$

$$S_3 = h_1 w_5 + h_2 w_6$$

$$= (0.255)(-0.2) + (0.0139)(0.5)$$

$$= -0.051 + 0.00695$$

$$\cancel{= -0.04} \quad \boxed{= -0.04405}$$

Now $\hat{y} = \dfrac{1}{1+e^{-h_1 w_5 - h_2 w_6}}$

$$= \frac{1}{1+e^{(-0.255)(-0.2)-(0.0139)(0.5)}}$$

$$= \frac{1}{1+e^{(0.051)-(0.00695)}}$$

$$= \frac{1}{1+e^{0.04405}}$$

$$= \frac{1}{1+1.04503}$$

$$\boxed{\hat{y} = 0.4889}$$

Gradient of L_2 loss function

$(\|\hat{y}-y\|^2$ is $2\|\hat{y}-y\| = \dfrac{\partial \epsilon}{\partial y}$

Backward propagation

$$\frac{\partial \epsilon}{\partial w_1} = \frac{\partial \epsilon}{\partial \hat{y}} \times \frac{\partial \hat{y}}{\partial s_3} \times \frac{\partial s_3}{\partial h_1} \times \frac{\partial h_1}{\partial s_1} \times \frac{\partial s_1}{\partial w_1}$$

$$\frac{\partial \epsilon}{\partial w_1} = 2\|\hat{y}-y\| \times \sigma(s_3) \times w_5 \times \sigma(s_1) \times (x_1)$$

$$\frac{\partial s_3}{\partial h_1} = w_5 \quad / \quad \frac{\partial s_1}{\partial w_1} = x_1$$

Substituting the value in $\frac{\partial E}{\partial w_1}$

$$\frac{\partial E}{\partial w_1} = \left[2 \cdot \| 0.4899 - 0.5 \|\right] \times \left[\sigma(s_3)\right)\left(1 - \sigma(s_3)\right)\Big] \times \left[(-0.2)\right] \times \sigma(s_1)\left(1 - \sigma(s_1)\right] \times (0.7)$$

$$\sigma(s_3) = \frac{1}{1 + e^{-(-0.04405)}} = 0.4884$$

$$\sigma(s_1) = \frac{1}{1 + e^{-(-1.07)}} = 0.2554$$

$$\frac{\partial E}{\partial w_1} = \left[2 \| 0.4899 - 05 \|\right] \times \left[(0.4884) \, (1 - 0.4884)\right] \times 0.2 \times \left[(0.2554) \, (1 - 0.2554)\right] \times 0.7$$

$$= \left[2(0.0107)\right] \times \left[(0.4889) \, 0.5116\right)$$

$$\Rightarrow \frac{\partial E}{\partial w_1} = 0.0014$$