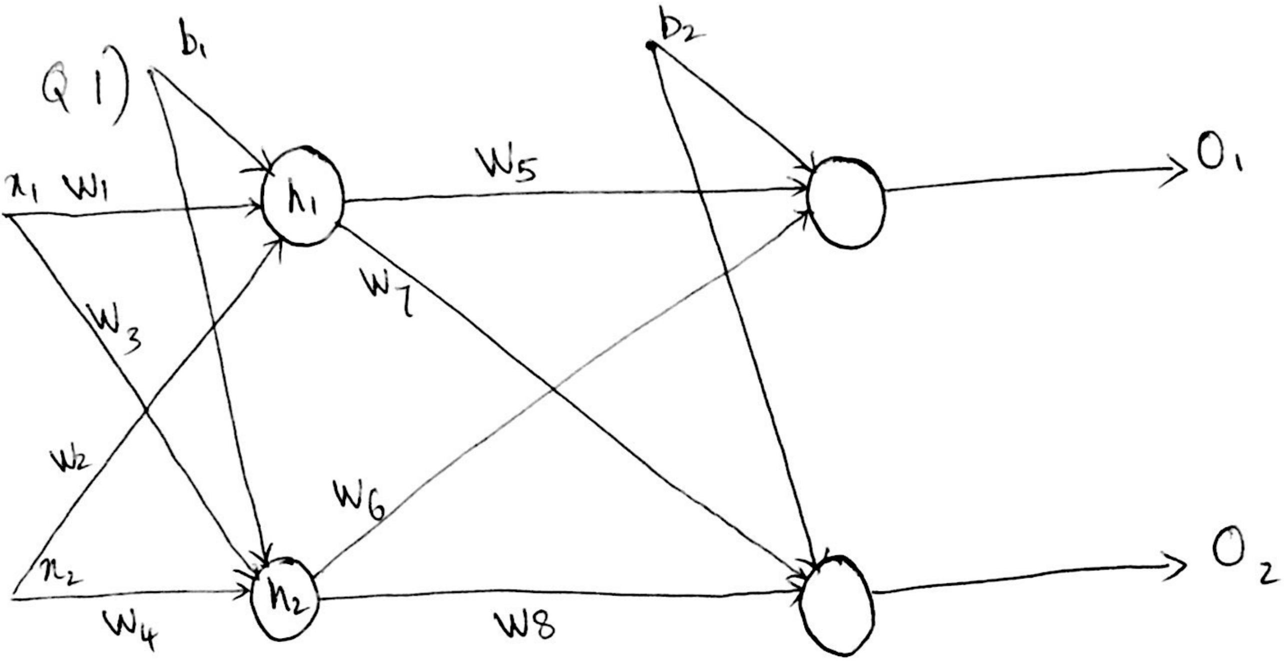
HomeWork - II - Neural Networks and Deep Learning

MEHUL SURESH KUMAR - U52982215

Q 1)



$x_1 = 1,\ x_2 = 2$

$w_1 = 0.1,\ w_2 = 0.2,\ w_3 = 0.3,\ w_4 = 0.4$

$w_5 = 0.2,\ w_6 = 0.1,\ w_7 = 0.4,\ w_8 = 0.3$

$b_1 = 0.1,\ b_2 = 0.2$

$$h_{1i} = w_1 x_1 + w_2 x_2 + b_1$$

$$= 0.1 \times 1 + 0.2 \times 2 + 0.1 = 0.6$$

$$h_{2i} = w_3 x_1 + w_4 x_2 + b_1$$

$$= 0.3 \times 1 + 0.4 \times 2 + 0.1 = 1.2$$

Scanned by CamScanner

$$h_{10} = \frac{1}{1+e^{-h_{1i}}}$$

$$= \frac{1}{1+e^{-0.6}} = 0.6456$$

$$h_{20} = \frac{1}{1+e^{-h_{2i}}}$$

$$= \frac{1}{1+e^{-1.2}} = 0.7685$$

$$O_{1i} = w_5 h_{10} + w_6 h_{20} + b_2$$

$$= 0.2 \times 0.6456 + 0.1 \times 0.7685 + 0.2 = 0.4059$$

$$O_{2i} = w_7 h_{10} + w_8 h_{20} + b_2$$

$$= 0.4 \times 0.6456 + 0.3 \times 0.7685 + 0.2 = 0.6887$$

$\therefore O_1 = \dfrac{1}{1+e^{-O_{1i}}}$

$= \dfrac{1}{1+e^{-0.4059}} = 0.6001$

$\therefore O_2 = \dfrac{1}{1+e^{-O_{2i}}}$

$= \dfrac{1}{1+e^{-0.6887}} = 0.6656$

Q2) $L = (O_1 + \hat{O}_1)^2 + (O_2 - \hat{O}_2)^2$

Lets find $\dfrac{\partial L}{\partial w_7}$ first

$$\frac{\partial L}{\partial w_7} = \frac{\partial L}{\partial O_2} \times \frac{\partial O_2}{\partial O_{2i}} \times \frac{\partial O_{2i}}{\partial w_7}$$

$$\frac{\partial L}{\partial O_2} = 0 + 2(O_2 - \hat{O}_2)^{2-1} \times -1$$

$$= -2(O_2 - \hat{O}_2)$$

$$= -2(1 - 0.6656)$$

$$= -0.6688$$

We know $O_2 = \dfrac{1}{1 + e^{-O_{2i}}}$

$$\frac{\partial O_2}{\partial O_{2i}} = O_2(1 - O_2)$$

$$= 0.6656(1 - 0.6656)$$

$$= 0.2225$$

$$O_{2i} = w_7 h_{10} + w_8 h_{20} + b_2$$

$$\partial O_{2i} / \partial w_7 = 1 \times h_{10} + 0 + 0 = h_{10} = 0.6456$$

$$\therefore \frac{\partial L}{\partial w_1} = -0.6688 \times 0.2225 \times 0.6456$$

$$= -0.0961$$

now lets find $\frac{\partial L}{\partial w_3}$

$$\frac{\partial L}{\partial w_3} = \frac{\partial L}{\partial h_{2o}} \times \frac{\partial h_{2o}}{\partial h_{2i}} * \frac{\partial h_{2i}}{\partial w_3}$$

$$\frac{\partial L}{\partial h_{2o}} = \frac{\partial (o_1 - \hat{o}_1)^2}{\partial_{h_{2o}}} + \frac{\partial (o_2 - \hat{o}_2)^2}{\partial h_{2o}}$$

$$\frac{\partial (o_1 - \hat{o}_1)^2}{\partial h_{2o}} = \frac{\partial (o_1 - \hat{o}_1)^2}{\partial o_1} \cdot \frac{\partial o_1}{\partial o_{1i}} \cdot \frac{\partial o_{1i}}{\partial h_{2o}}$$

Scanned by CamScanner

$= -2(O_1 - \hat{O}_1) \times O_1(1-O_1) \times W_6$

$= -2(0 - 0.6001) \times (0.24) \times 0.1$

$\cancel{= 0.0288} = 0.0288$ — 1

$$\frac{\partial(O_2 - \hat{O}_2)}{\partial h_{20}} = \frac{\partial(O_2 - \hat{O}_2)}{\partial O_2} \frac{\partial O_2}{\partial O_{2i}} \frac{\partial O_{2i}}{\partial h_{20}}$$

$= -0.6688 \times 0.2225 \times W8$

$= -0.6688 \times 0.2225 \times 0.3$

$= -0.0446$ — 2

$$\frac{\partial L}{\partial h_{20}} = 1 + 2 = -0.0158$$

$$\therefore \frac{\partial L}{\partial w_3} = \frac{\partial L}{\partial h_{2o}} \times \frac{\partial h_{2o}}{\partial h_{2i}} \times \frac{\partial h_{2i}}{\partial w_3}$$

$$= -0.0158 \times h_{2o}(1-h_{2o}) \times x_1$$

$$= -0.0158 \times 0.1779 \times 1$$

$$= -2.81 \times 10^{-3}$$

$$= 0.00281$$