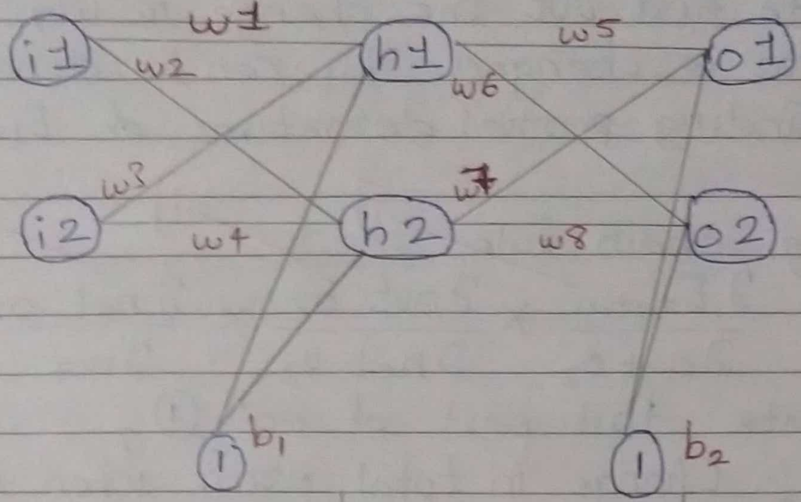
Shrita Gaonkar
100799307

Exercise #1



i1 = 0.05	w1 = 0.15	w_5 = 0.40
i2 = 0.10	w2 = 0.20	w_6 = 0.45
o1 = 0.01	w3 = 0.25	w_7 = 0.50
o2 = 0.99	w4 = 0.30	w_8 = 0.55
b1 = 0.35		
b2 = 0.60		

For finding the weights of w7 and w3, we use backward propogation & Forward propogation

① Forward propogation:

After completing forward propogation, we get:

net_{h1} = 0.3775	net_{o2} = 1.2249214039
out_{h1} = 0.593269992	net_{o1} = 1.105905967
net_{h2} = 0.39	out_{o1} = 0.75136507
out_{h2} = 0.596884378	out_{o2} = 0.772928465

② Backward propogation: We want to find out the updated weights on w7 first. Looking at the diagram it comes from h1 and goes to o1.

We want to find out the change in w_7 when the total error is changed. This can be done by gradient descent. Finding partial derivatives of E_{total} w.r.t w_7.

By applying chain rule,

i) ii) iii)

$$\frac{\partial E_{Total}}{\partial w_7} = \frac{\partial E_{Total}}{\partial out\, o_2} \times \frac{\partial out\, o_2}{\partial net\, o_2} \times \frac{\partial net\, o_2}{\partial w_7} \quad — (1)$$

let's calculate first part of eq^n (1),

i) $\frac{\partial E_{Total}}{\partial out\, o_2}$ = Change in total error when out o_2 is changed

$$E_{Total} = \frac{1}{2}(target\ o_1 - out\ o_1)^2 + \frac{1}{2}(target\ o_2 - out\ o_2)^2$$

target = ~~or~~ original
out = predicted

$$\frac{\partial E_{Total}}{\partial out\, o_2} = 0 + 2 \times \frac{1}{2}(target\ o_2 - out\ o_2)^{2-1} \times \frac{\partial}{\partial out\, o_2}\left[target\ o_2 - out\ o_2\right]$$

because o_1, not associated with eq^n

$$= (target\ o_2 - out\ o_2) \times [0-1] + 0$$

$$= -(target\ o_2 - out\ o_2)$$

$$= -(0.99 - 0.772928465)$$

$$\frac{\partial E_{TOTAL}}{\partial out\, o_2} = -0.217071535$$

ii) Solving the second eq^n of eq^n (1),

$\frac{\partial out\, o_2}{\partial net\, o_2}$ = Change in ~~total~~ out o_2 w.r.t net o_2

$$out\ o_2 = \frac{1}{1+e^{-net\, o_2}} \longrightarrow \text{Sigmoid function}$$

Finding the derivative of sigmoid we get,

~~out o_2~~ $$\frac{\partial out\ o_2}{\partial net\ o_2} = out\ o_2\ (1 - out\ o_2)$$

$$= 0.772928465(1-0.772928465)$$

$$\frac{\partial out\ o_2}{\partial net\ o_2} = 0.1755510053$$

iii) Solving the third part of equation 1,

$$\frac{\partial net\ o_2}{\partial w_7} = \text{Change in net } o_2 \text{ w.r.t } w_7$$

$$net\ o_2 = w_7 \times out\ h_1 + w_8 \times out\ h_2 + 1 \times b_2$$

$$\frac{\partial net\ o_2}{\partial w_7} = \frac{\partial}{\partial w_7}(w_7 \times out\ h_1) + \frac{\partial}{\partial w_7}(w_8 \times out\ h_2) + \frac{\partial}{\partial w_7}(1 \times b_2)$$

$$= out\ h_1 \times \frac{\partial w_7}{\partial w_7} + 0 + 0$$

$$= out\ h_1$$

$$\frac{\partial net\ o_2}{\partial 7} = 0.59326992$$

Putting all the values of i], ii], iii] of eq^n 1,

$$\frac{\partial E_{Total}}{\partial w_7} = \frac{\partial E_{Total}}{\partial out\ o_2} \times \frac{\partial out\ o_2}{\partial net\ o_2} \times \frac{\partial net\ o_2}{\partial w_7}$$

$$= -0.217071525 \times 0.175510053 \times 0.59326992$$

$$\frac{\partial E_{Total}}{\partial w_7} = -0.022602540$$

Now, we update the weight of w_7 using gradient descent,

$$W_{new} = W_{old} - \alpha \times \frac{\partial E_{Total}}{\partial W_{old}}$$

$$w_7^+ = w_7 - \alpha \times \frac{\partial E_{Total}}{\partial w_7}$$

$$w_7^+ = 0.50 - 0.5 \times (-0.022602540)$$

$$w_7^+ = 0.511301270$$

Calculate the updated weight on w_3,

$$\frac{\partial E_{Total}}{\partial w_3} = \overset{iii]}{\frac{\partial E_{Total}}{\partial out\,h_2}} \times \overset{iv]}{\frac{\partial out\,h_2}{\partial net\,h_2}} \times \overset{v]}{\frac{\partial net\,h_2}{\partial w_3}} \quad \text{--- } (2)$$

iii] Lets calulate first part (iii]) of eq^n 2,

$$\frac{\partial E_{Total}}{\partial out\,h_2} = \overset{vi]}{\frac{\partial E_{O_1}}{\partial out\,h_2}} + \overset{vii]}{\frac{\partial E_{O_2}}{\partial out\,h_2}} \quad \text{--- } (3)$$

w7 affect h_2. And h_2 inturn affect o1 and o2

vi] $$\frac{\partial E_{O_1}}{\partial out\,h_2} = \overset{viii]}{\frac{\partial E_{O_1}}{\partial net\,O_1}} \times \overset{ix]}{\frac{\partial net\,O_1}{\partial out\,h_2}} \quad \text{--- } (4)$$

viii] $$\frac{\partial E_{O_1}}{\partial net\,O_1} = \overset{x]}{\frac{\partial E_{O_1}}{\partial out\,O_1}} \times \overset{xi]}{\frac{\partial out\,O_1}{\partial net\,O_1}} \quad \text{--- } (5)$$

x] $$\frac{\partial E_{O_1}}{\partial out\,O_1} = \frac{\partial}{\partial out\,O_1}\left[\frac{1}{2}(target_{O_1} - out\,O_1)^2\right]^2$$

$$= 2 \times \frac{1}{2}(target_{O_1} - out\,O_1)^{2-1} \times \frac{\partial}{\partial out\,O_1}(target\,O_1 - out\,O_1)$$

$= -(target_{o_1} - out_{o_1})$

$= -(0.01 - 0.75136507)$

$\frac{\partial E_{o_1}}{\partial out_{o_1}} = +0.74136507.$

xi) $\frac{\partial out_{o_1}}{\partial net_{o_1}} = out_{o_1}(1 - out_{o_1})$

$= 0.75136507(1 - 0.75136507)$

$= 0.186815602$

Equation 5, $\frac{\partial E_{o1}}{\partial net_{o_1}} = \frac{\partial E_{o1}}{\partial out_{o_1}} \times \frac{\partial out_{o1}}{\partial net_{o_1}}$

$= +0.74136507 \times 0.186815602$

$\frac{\partial E_{o1}}{\partial net_{o_1}} = 0.138498561$

ix) $\frac{\partial net_{o1}}{\partial out_{h_2}} = \frac{\partial}{\partial out_{h_2}}(w_5 \times out_{h_1} + w_6 \cdot out_{h_2} + 1 \times b_2)$

$= w_6 = 0.45$

Equation 4, $\frac{\partial E_{o1}}{\partial out_{h_2}} = \frac{\partial E_{o1}}{\partial net_{o1}} \times \frac{\partial net_{o1}}{\partial out_{h_2}}$

$= 0.138498561 \times 0.45$

$= 0.062324352$

We now need to find vii), second part of equation 3,

vii) $\frac{\partial E_{o2}}{\partial out_{h_2}} = \underbrace{\frac{\partial E_{o2}}{\partial net_{o2}}}_{xii)} \times \underbrace{\frac{\partial net_{o2}}{\partial out_{h_2}}}_{xiii)}$ — (6)

xii) $\frac{\partial E_{O_2}}{\partial net_{O_2}} = \underbrace{\frac{\partial E_{O_2}}{\partial out_2}}_{xiv)} \times \underbrace{\frac{\partial out_{O_2}}{\partial net_{O_2}}}_{xv)}$ —— (7)

xiv) $\frac{\partial E_{O_2}}{\partial out_{O_2}} = \frac{\partial}{\partial out_{O_2}}\left[\frac{1}{2}(target_{O_2} - out_{O_2})^2\right]^2$

$= 2 \times \frac{1}{2}(target_{O_2} - out_{O_2})^{2-1}$

$\times \frac{\partial}{\partial out_{O_2}}(target_{O_2} - out_{O_2})$

$= -(target_{O_2} - out_{O_2})$

~~$= -(0.01 - 0.75136507)$~~

$= 0.99 - 0.772928465$

$= -0.217071535$

xv) $\frac{\partial out_{O_2}}{\partial net_{O_2}} = out_{O_2}(1 - out_{O_2})$

$= 0.772928465(1 - 0.772928465)$

$= 0.175510053$

Equation 7, $\frac{\partial E_{O_2}}{\partial net_{O_2}} = \frac{\partial E_{O_2}}{\partial out_2} \times \frac{\partial out_{O_2}}{\partial net_{O_2}}$

$= -0.217071535 \times 0.175510053$

$\frac{\partial E_{O_2}}{\partial net_{O_2}} = -0.038098236$

xiii) ~~Equation 6~~, $\frac{\partial net_{O_2}}{\partial out_{h_2}} = \frac{\partial}{\partial out_{h_2}}(w_7 \times out_{h_1} + w_8 \times out_{h_2} + 1 \times b_2)$

$= (w_8 \times 1) + 0 + 0$

$= 0.55$

From equation 6,

$$\frac{\partial EO2}{\partial out h_2} = \frac{\partial EO2}{\partial net o_2} \times \frac{\partial net o_2}{\partial out h_2}$$
$$= -0.03809826 \times 0.55$$
$$= -0.02095403$$

We got the value of viii] & ix] from equation 4, and equation 6.

Now we substitute that value in equation 3,

$$\frac{\partial E_{Total}}{\partial out h_2} = \frac{\partial E_{O1}}{\partial out h_2} + \frac{\partial E_{O2}}{\partial out h_2}$$
$$= 0.06232435 + -0.02095403$$
$$= 0.04137032$$

iv] Second part of equation 2,

$$\frac{\partial out h_2}{\partial net h_2} = \frac{\partial}{\partial net h_2}\left(\frac{1}{1+e^{-net h_2}}\right)$$
$$= out h_2 (1 - out h_2)$$
$$= 0.596884378 \times (1 - 0.596884378)$$
$$= 0.240613417.$$

v] Third part of equation 2,

$$\frac{\partial net h_2}{\partial w_3} = \frac{\partial}{\partial w_3}(w_3 \times i + w_4 \times i_2 + 1 \times b_1)$$
$$= i_1 + 0 + 0$$
$$= 0.05.$$

Equation 2,

$$\frac{\partial E_{total}}{\partial w_3} = \frac{\partial E_{total}}{\partial out_{h_2}} \times \frac{\partial out_{h_2}}{\partial net_{h_2}} \times \frac{\partial net_{h_2}}{\partial w_3}$$

$$= 0.04137032 \times 0.240612417 \times 0.05$$

$$= 0.0004977127$$

Now, update the weights using gradient descent,

$$w_{new} = w_{old} - \alpha \times \frac{\partial E_{Total}}{\partial w_{old}}$$

$$w_3^+ = w_3 - \alpha \times \frac{\partial E_{Total}}{\partial w_3}$$

$$= 0.25 - 0.50 \times 0.0004977127$$

$$w_3^+ = 0.24975144$$

The updated weight of w_7 is 0.511301270.
The updated weight of w_3 is 0.24975144