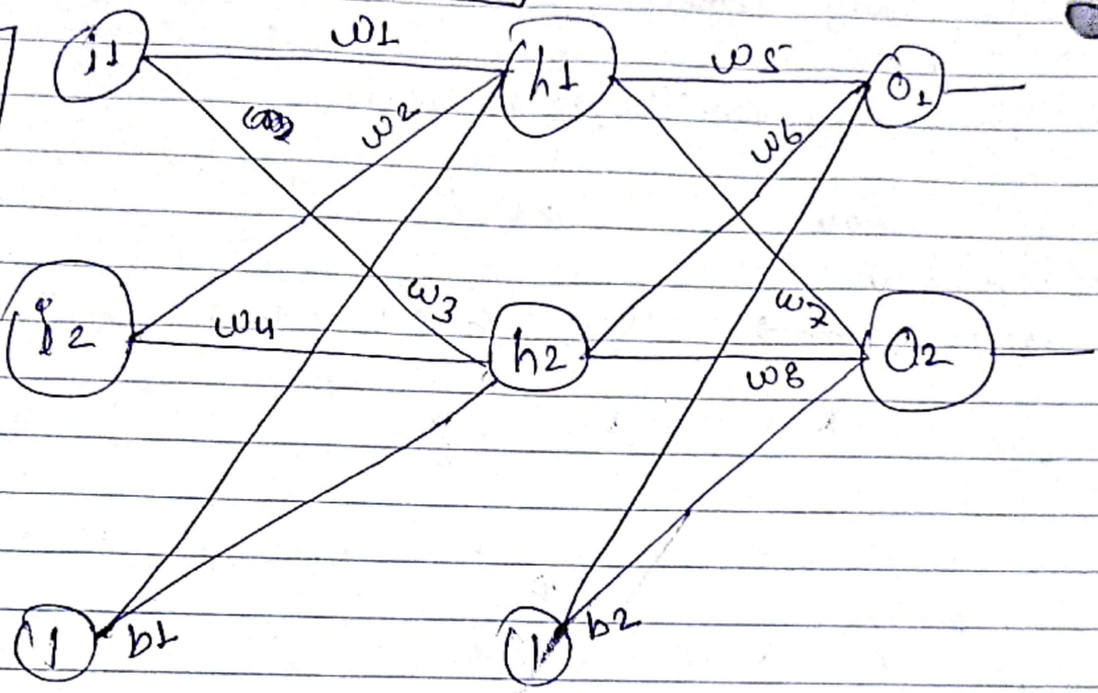
Name : Bhavna Panwar
Student ID : 100802837

Exercise 1



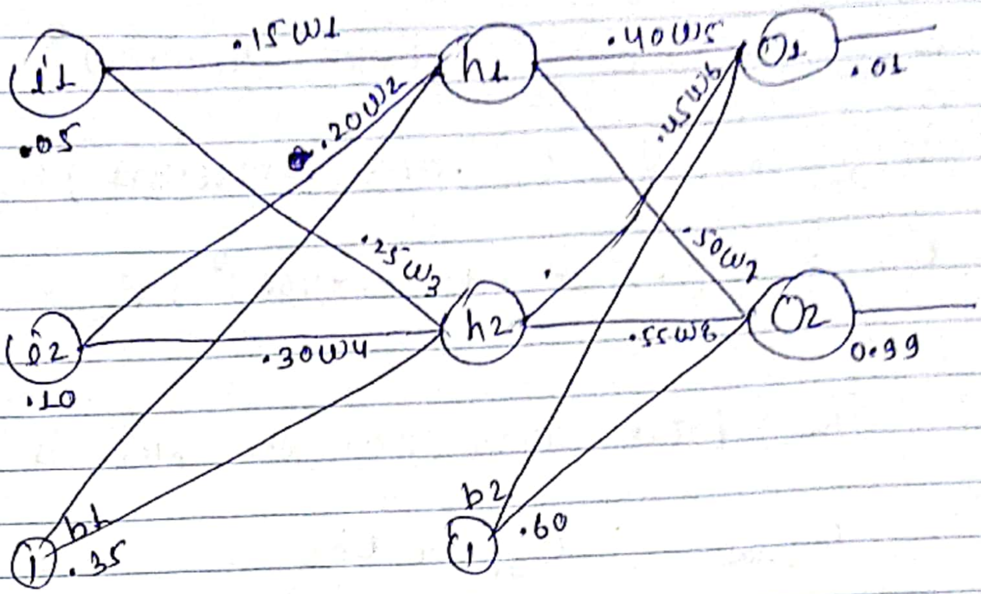
$i1 = .05$
$i2 = .10$
$O1 = .01$
$O2 = .99$
$b1 = .35$
$b2 = .60$
$w1 = .15$
$w2 = .20$
$w3 = .25$
$w4 = .30$
$w5 = .40$
$w6 = .45$
$w7 = .50$
$w8 = .55$

Find :

Updated weights of w7 & w3 after the first iteration of your NN training

→ Given, learning rate of for gradient descent as 0.5

Forward Propogation!



Step 1:

$\rightarrow$ $net_{h1} = w_1 * i_1 + w_2 * i_2 + b_1 * 1$

$= .15 * .05 + .20 * .10 + .35 * 1$

$net_{h1} = 0.0075 + 0.02 + 0.35 \Rightarrow 0.3775$

$\rightarrow$ $out_{h1} = \frac{1}{1+e^{-x}} = \frac{1}{1+e^{-0.3775}} \Rightarrow \frac{1}{1+0.68557}$

$out\ h1 = 0.593269993$

$\rightarrow$ $net\ h2 = w_3 * i_1 + w_4 * i_2 + 1 * b_1$

$= .25 * 0.05 + .30 * .10 + 1 * 0.35 \Rightarrow 0.0125 + 0.03 + 0.35$

$net\ h2 = 0.3925$

$\rightarrow$ $out\ h2 = \frac{1}{1+e^{-x}} \Rightarrow \frac{1}{1+e^{-0.3925}} \Rightarrow 0.596884378$

$\rightarrow$ $net\ O1 = w_5 * out\ h1 + w_6 * out\ h2 + 1 * b_2$

$\Rightarrow .40 * 0.593269993 + 0.45 * 0.596884378 + 1 * 0.60$

$\Rightarrow 0.237307997 + 0.268597970 + 0.60$

$net\ O1 = 1.105905967$

$\rightarrow$ $out\ O1 = \frac{1}{1+e^{-1.105905967}} \Rightarrow 0.75136507$

$\rightarrow$ $net\ O2 = w_7 * out\ h1 + w_8 * out\ h2 + 1 * b_2$

$\Rightarrow .50 * 0.593269993 + .55 * 0.596884378 + 1 * .60$

$\Rightarrow 0.2966349965 + 0.3282864079 + 0.60 \Rightarrow 1.2249214044$

$\rightarrow$ $out\ O2 = \frac{1}{1+e^{-1.2249214044}} \Rightarrow 0.772928465$

Step 2: Calculating the total errors (squared error funⁿ)

$\rightarrow$ $E_{total} = \sum \frac{1}{2}(\text{target} - \text{output})^2$

⊙ $E_{o1} = \frac{1}{2}(.01 - .751365071)^2 \Rightarrow E_{o1} = .274811084$

⊙ $E_{o2} = \frac{1}{2}(.99 - .772928465)^2 \Rightarrow E_{o2} = 0.023560257$

∴ The total error for the NN is:

$E_{total} = E_{o1} + E_{o2}$

$E_{total} = .274811084 + 0.02356026$

$\boxed{E_{Total} = 0.298371109}$

Backward Propagation!

$\Rightarrow$ Consider w_7; we want to know how much a change in w_7 affects the total error

This is represented as $\frac{\partial E_{total}}{\partial w_7}$

By applying the chain rule!

$$\frac{\partial E_{total}}{\partial w_7} = \frac{\partial E_{total}}{\partial out_{o2}} * \frac{\partial out_{o2}}{\partial net_{o2}} * \frac{\partial net_{o2}}{\partial w_7}$$

Now we need to figure out each pieces of this equation.

Scanned with CamScanner

First how much does the Total error change w.r.t to the output (out O2) !

$$E_{Total} = \frac{1}{2}(\text{target } O_2 - \text{out } O_2)^2 + \frac{1}{2}(\text{target } O_2 - \text{out } O_2)^2$$

① $$\frac{\partial E_{Total}}{\partial \text{out } O_2} = 0 + 2 * \frac{1}{2}(\text{Target } O_2 - \text{out } O_2) * \frac{\partial}{\partial \text{out}_{O_2}}(\text{Target } O_2 - \text{out } O_2)$$

$$\Rightarrow (\text{target } O_2 - \text{out } O_2) * (0 - 1)$$

$$\Rightarrow -(\text{target } O_2 - \text{out } O_2)$$

$$\Rightarrow -(0.99 - 0.772928465)$$

$$\frac{\partial E_{Total}}{\partial \text{out } O_2} = -0.217071535$$

② Second, how much does the output of O_2 change w.r.t to its total net input !

$$\left\{ f(x) = \frac{1}{1+e^{-x}} \qquad \frac{\partial}{\partial x} f(x) = f(x)(1 - f(x)) \right\}$$

$$\text{out } O_2 = \frac{1}{1+e^{-\text{net } O_2}}$$

$$\frac{\partial \text{ out } O_2}{\partial \text{ net } O_2} = \text{out}_{O_2}(1 - \text{out } O_2) \Rightarrow 0.772928465 * (1 - 0.772928465)$$

$$\Rightarrow 0.175510053$$

③ last ; how much does the total net input of O2 change w.r.t w7 :

$$\text{net } O_2 = w_7 * \text{out } h_1 + w_8 * \text{out } h_2 + 1 * b_2$$

$$\frac{\partial \text{ net } O_2}{\partial w_7} = \frac{\partial}{\partial w_7}(w_7 * \text{out } h_1) + \frac{\partial}{\partial w_7}(w_8 * \text{out } h_2) + \frac{\partial}{\partial w_7}(b_2)$$

$$\Rightarrow \text{out } h_1 * \frac{\partial w_7}{\partial w_7} + 0 + 0 \Rightarrow \text{out } h_1 * 1$$

$$\Rightarrow 0.593269993$$

Scanned with CamScanner

→ putting it all together,

$$\frac{\partial E_{Total}}{\partial w_7} = -0.217071535 * 0.175510053 * 0.593269993$$

$$\Rightarrow -0.022602541$$

→ Now we update the weight using G.D.

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

$$w_7^+ = w_7 - \alpha * \frac{\partial E_{total}}{\partial w_7} \Rightarrow$$

$$\Rightarrow .50 - .5 * (-0.022602541)$$

$$\Rightarrow .50 + (.5 * 0.022602541)$$

$$\boxed{w_7^+ = 0.511301271}$$

⇒ Hidden layer, Consider w_3 :

Here is what we need to figure out :

$$\frac{\partial E_{total}}{\partial w_3} = \frac{\partial E_{total}}{\partial out_{h2}} * \frac{\partial out_{h2}}{\partial net_{h2}} * \frac{\partial net_{h2}}{\partial w_3}$$

Now, we need to figure out each piece of eqn.

We know that out h2 affects both out o1 & out o2 therefore the $\frac{\partial E_{Total}}{\partial out_{h2}}$ needs to take into consideration its effect on the both neurons.

$$\therefore \quad \frac{\partial E_{total}}{\partial out_{h2}} = \frac{\partial E_{o1}}{\partial out_{h2}} + \frac{\partial E_{o2}}{\partial out_{h2}}$$

- Starting with $\frac{\partial E_{o1}}{\partial out_{h2}}$

$$\rightarrow \frac{\partial E_{o1}}{\partial out_{h2}} = \frac{\partial E_{o1}}{\partial net_{o1}} * \frac{\partial net_{o1}}{\partial out_{h2}}$$

- $\frac{\partial E_{o1}}{\partial net_{o1}} = \frac{\partial E_{o1}}{\partial out_{o1}} * \frac{\partial out_{o1}}{\partial net_{o1}}$

$$= \frac{\partial}{\partial out_{o1}} \left[\frac{1}{2} (Target_{o1} - out_{o1})^2 \right] * out_{o1}(1 - out_{o1})$$

$$= \left\{ \frac{1}{\cancel{2}} * \cancel{2} (Target_{o1} - out_{o1}) * \frac{\partial}{\partial out_{o1}} (Target_{o1} - out_{o1}) \right\}$$

$$* \left[0.751365071 * (1 - 0.751365071) \right]$$

$$\Rightarrow \left\{ (Target_{o1} - out_{o1}) * (0 - 1) \right\} * 0.186815601$$

$$\Rightarrow (0.01 - 0.751365071)(-1) * 0.186815601$$

$$\Rightarrow (-0.741365071) * (-1) * 0.186815601$$

$$\Rightarrow 0.138498561$$

- $\frac{\partial net_{o1}}{\partial out_{h2}} = \frac{\partial}{\partial out_{h2}} (w_5 * out_{h1} + w_6 * out_{h2} + 1 * b_2)$

$$= 0 + w_6 * 1 + 0$$

$$\Rightarrow 0.45$$

$$\therefore \quad \frac{\partial E_{o1}}{\partial out_{h2}} = 0.138498562 * 0.45 \Rightarrow 0.062324353$$

Next find $\frac{\partial E_{O2}}{\partial out_{h2}}$!

$$\frac{\partial E_{O2}}{\partial out_{h2}} = \frac{\partial E_{O2}}{\partial out_{O2}} * \frac{\partial out_{O2}}{\partial net_{O2}} * \frac{\partial net_{O2}}{\partial out_{h2}}$$

$$= \frac{\partial}{\partial out_{O2}} \left[\frac{1}{2} (Target_{O2} - out_{O2})^2 \right] * out_{O2} (1 - out_{O2}) * \frac{\partial}{\partial out_{h2}} (w_7 * out_{h1} + w_8 * out_{h2} + 1 * b_2)$$

$$\Rightarrow \left\{ \frac{1}{\cancel{2}} * \cancel{2} * (target_{O2} - out_{O2}) * \frac{\partial}{\partial out_{O2}} (target_{O2} - out_{O2}) \right\} * (0.175510053) * [0 + w_8 * 1 + 0]$$

$$\Rightarrow \{ (0.99 - 0.772928465) * (0 - 1) \} * 0.175510053 * .55$$

$$\Rightarrow -0.020954030$$

$$\therefore \frac{\partial E_{total}}{\partial out_{h2}} = 0.062384353 - 0.020954030$$

$$= 0.041370323$$

→ Second Nece :

$$\frac{\partial out_{h2}}{\partial net_{h2}}$$

$$\left\{ \begin{array}{l} f(x) = \frac{1}{1+e^{-x}} \\ \frac{\partial}{\partial x} f(x) = f(x)(1-f(x)) \end{array} \right.$$

$$\Rightarrow \frac{\partial}{\partial net_{h2}} \left(\frac{1}{1+e^{-net_{h2}}} \right)$$

$$\frac{\partial out_{h2}}{\partial net_{h2}} = out_{h2} (1 - out_{h2})$$

$$\Rightarrow 0.596884378 (1 - 0.596884378)$$

$$\Rightarrow 0.240613417$$

→ Third piece :

$$\frac{\partial net\ h_2}{\partial w_3} = \frac{\partial}{\partial w_3}(w_3 * i_1 + w_4 * i_2 + 1 * b_1)$$

$$= i_1 * 1 + 0 + 0$$

$$= 0.05$$

→ Putting it all Together :

$$\frac{\partial E_{Total}}{\partial w_3} = 0.041370323 * 0.240613417 * 0.05$$

$$= 0.000497713$$

→ Now we update the weight w3 using GD

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

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

$$w_3^+ = 0.25 - 0.5 * 0.000497713$$

$$w_3^+ = 0.25 - 0.000248857$$

$$\boxed{w_3^+ = 0.249751143}$$

Hence, updated weight ;

$$\boxed{w_7^+ = 0.511301271}$$

and

$$\boxed{w_3^+ = 0.24975143}$$