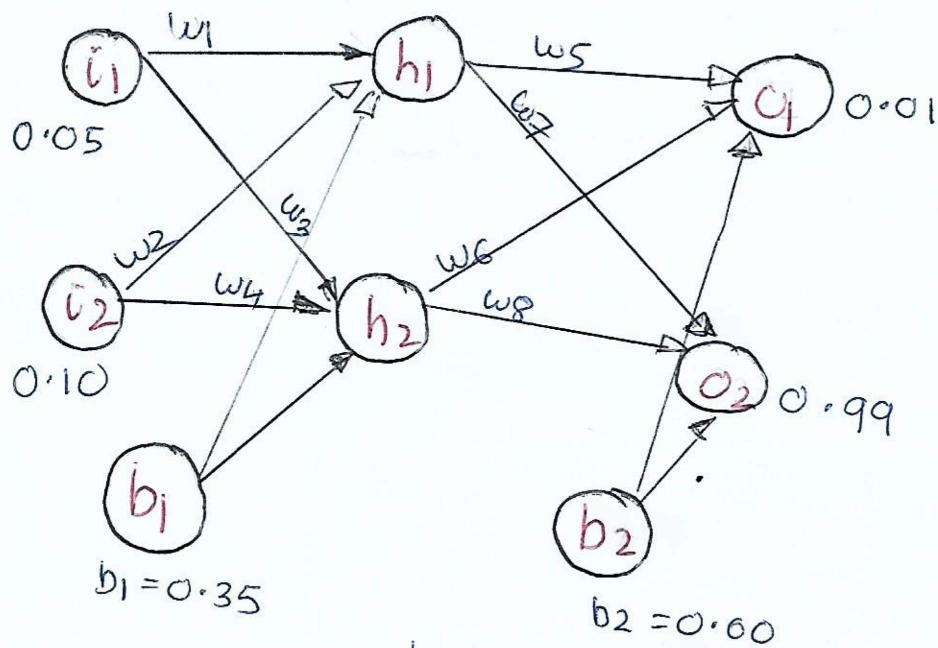
BACK PROPagation

Consider the Network shown below.



initial bias and Weights table

W	initial Value
W1	0.15
W2	0.2
W3	0.25
W4	0.30
W5	0.4
W6	0.45
W7	0.5
W8	0.55
b1	0.35
b2	0.60

[inputs]

i_1	= 0.05
i_2	= 0.10

[Outputs]

O_1	= 0.01
O_2	= 0.99

Forward Pass. [INNER LAYER]

- Compute net input for h1

$$h_{1\,Net} = w_1 i_1 + w_2 i_2 + b_1$$

$$= 0.15 \times 0.05 + 0.2 \times 0.1 + 0.35$$

or $h_1 net = 0.3775$

- Compute output for h_1 [Using Logistics function Sigmoid]

$$h_{1\,out} = \frac{1}{1+e^{-h_1 net}} = \frac{1}{1+e^{-0.3775}} = 0.593$$

○ Compute net input for h2

$h_{2_{Net}} = w_3 i_i + w_4 i_2 + b_1$

$= 0.25 \times 0.05 + 0.30 \times 0.1 + 0.35$

$= 0.3925$

○ Compute output for h2

$h_{2_{out}} = \frac{1}{1 + e^{-0.3925}} = 0.5968$

If Forward Pass [Outer Layer]

○ Compute net input for O1

$O_{1\,Net} = h_1 w_5 + h_2 w_6 + b_2$

$= 0.593 \times 0.4 + 0.597 \times 0.45 + 0.6$

or $O_{1\,Net} = 1.105$

○ Compute net input for O_2

$O_{2_{net}} = h_1 w_7 + h_2 w_8 + b_2$

$= 0.593 \times 0.5 + 0.597 \times 0.55 + 0.6$

$= 1.225$

○ Compute output for O1

$O_{1\,out} = \frac{1}{1 + e^{-1.105}} = 0.751$

○ Compute output for O_2

$O_{2_{out}} = \frac{1}{1 + e^{-1.225}} = 0.773.$

Error Computation

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

$$E_{O1} = \frac{1}{2}(target_{o1} - out_{o1})^2 \qquad E_{O2} = \frac{1}{2}(target_{O2} - out_{O2})^2$$

Compute Error at the first output node

$$E_{O1} = \frac{1}{2}(target_{o1} - out_{o1})^2 = \frac{1}{2}(0.01 - 0.751)^2$$

or $E_{O1} = 0.274$

Compute Error at the 2nd output node.

we have $E_{O_2} = \frac{1}{2}(target_{O2} - out_{O2})^2$

$$= \frac{1}{2}(0.99 - 0.773)^2 = 0.023$$

Computing Weights] Output Layer (Back Propagation)

Objective: We wish to Find How the weights need to be modified to reduce the Total Error.

Consider output layer.] Weights Feeding into output layer are w_5, w_6, w_7 and w_8.

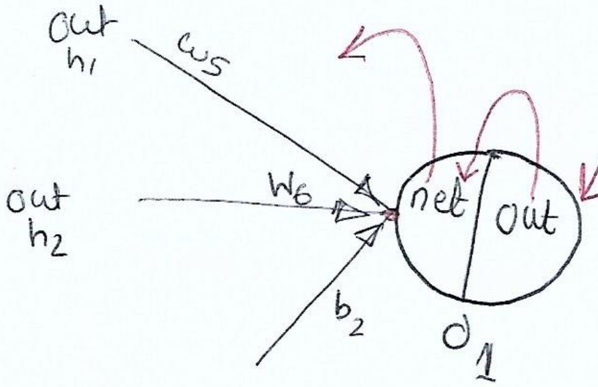
Consider w_5

We are interested to Find the term $\frac{\partial E_{total}}{\partial w_5}$

That is the Partial Derivative of E_{total} wrt 5

or we can consider the gradient of E_{total} wrt w_5

using this value the wts would be adjusted.



Note E_{o2} is not affected by w_5

$$E_{o1} = \frac{1}{2}(target_1 - out_{o1})^2$$

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

Thus we have.

$$\frac{\partial E_{total}}{\partial w_5} = \frac{\partial E_{total}}{\partial out_{o1}} \times \frac{\partial out_{o1}}{\partial net_{o1}} \times \frac{\partial net_{o1}}{\partial w_5}.$$

Let us look at each of the terms separately

1] $\frac{\partial E_{total}}{\partial out_{o1}}$: How much does the total error change with respect to the output?

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

$$= \frac{1}{2}(target_1 - out_{o1})^2 + \frac{1}{2}(target_2 - out_{o2})^2$$

Thus

$$\frac{\partial E_{total}}{\partial out_1} = (-1) \times 2 \times \frac{1}{2} \times (target_{o1} - out_{o1})$$

$$= (out - target) = 0.751 - 0.01 = 0.741$$

2] $\frac{\partial out_{o1}}{\partial net_{o1}}$: We have $out_{o1} = \frac{1}{1 - e^{-net_{o1}}}$

We know from the property of the Logistics Sigmoid that

$$\frac{\partial out_{o1}}{\partial net_{o1}} = out_{o1}(1 - out_{o1}).$$

$$= 0.751 * (1 - 0.751)$$

$$= 0.187$$

3] $\frac{\partial net_{o1}}{\partial w_5}$: we have $net_{o1} = w_5 \times out_{h_1} + w_6 \times out_{h_2} + b_2$

Thus $\frac{\partial net_{o1}}{\partial w_5} = out_{h_1} = 0.593$

Putting it all together we have

$$\frac{\partial E_{total}}{\partial w_5} = \frac{\partial E_{total}}{\partial out_{o1}} \times \frac{\partial out_{o1}}{\partial net_{o1}} \times \frac{\partial net_{o1}}{\partial w_5}$$

$$= 0.741 \times 0.187 \times 0.593$$

$$= 0.082$$

Now Consider the delta rule.

$$\delta_{o1} = \left[\frac{\partial E_{total}}{\partial out_{o1}} \times \frac{\partial out_{o1}}{\partial net_{o1}}\right] * out_{h_1}$$

$$\therefore \frac{\partial E_{total}}{\partial w_5} = \delta_{o1} * out_{h_1}$$

Now to decrease error

we Subtract this value from the current weight.

$$w_5^* = w_5 - \eta \times \frac{\partial E_{total}}{\partial w_5}$$

Learning Rate here we set it to 0.5

Thus we have

$$w_5^* = 0.4 - 0.5 \times 0.082 = 0.358$$

$$w_5^* = 0.358$$

Computing correction for w_6

We have.

$$\frac{\partial E_{Total}}{\partial w_6} = \frac{\partial E_{total}}{\partial out_1} \times \frac{\partial out_1}{\partial net_1} \times \frac{\partial net_1}{\partial w_6}$$

Now from Previous computation for w_5

we have ① $$\frac{\partial E_{Total}}{\partial out_1} = (out_1 - target_1) = 0.741$$

we also have ② $$\frac{\partial out_1}{\partial net_1} = out_{o1}(1 - out_{o1})$$

$$= 0.187.$$

③ Now $net_1 = w_5 \times out\,h_1 + w_6 \times out\,h_2 + b_2$

thus $$\frac{\partial net_{o1}}{\partial w_6} = out\,h_2 = 0.5968$$

Thus $$\frac{\partial E_{total}}{\partial w_6} = 0.741 \times 0.187 \times 0.5968$$

$$= 0.0827$$

Thus

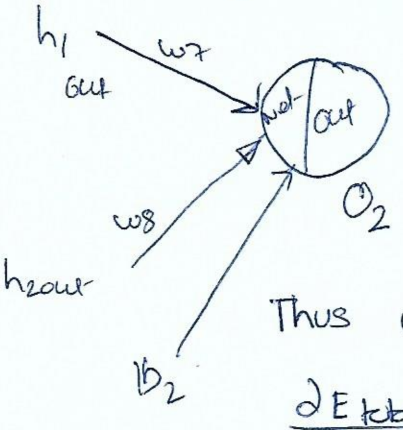
Corrected w_6

$$w_6^* = w_6 - \eta \times \frac{\partial E_{Total}}{\partial w_6}$$

$$= 0.45 - 0.5 \times 0.0827.$$

Thus $\boxed{w_6^* = 0.409}$.

Computing Correction for w_7.



$$E_{O2} = \frac{1}{2}(\text{target}_{O2} - \text{out}_{O2})^2$$

$$E_{total} = E_{O1} + E_{O2}$$

Thus we have.

$$\frac{\partial E_{total}}{\partial w_7} = \frac{\partial E_{total}}{\partial out_{O2}} \times \frac{\partial out_{O2}}{\partial net_{O2}} \times \frac{\partial net_{O2}}{\partial w_7}$$

Considering individual terms, we have:

① $$\frac{\partial E_{total}}{\partial out_{O2}} = (-1)(2) \times \frac{1}{2} \times (\text{target}_{O2} - \text{out}_{O2})$$

$$= out_2 - target_2$$

$$= (0.773 - 0.99)$$

$$= -0.217$$

② $\frac{\partial out_2}{\partial net_2}$: $$out_2 = \frac{1}{1+e^{-net_2}}$$

Thus $$\frac{\partial out_{O2}}{\partial net_{O2}} = out_{O2}(1 - out_{O2})$$

$$= 0.773(1 - 0.773)$$

$$= 0.175$$

③ We have $$net_{O2} = h_1out \times w_7 + h_2out \times w_8 + b_2$$

Thus $$\frac{\partial net_{O2}}{\partial w_7} = h_1out = 0.593$$

Thus $$\frac{\partial E_{Total}}{\partial w_7} = -0.217 \times 0.175 \times 0.593 = -0.0225$$

Thus $$w_7^* = w_7 - \eta \frac{\partial E_{total}}{\partial w_7} = 0.5 - 0.5(-0.0225)$$

Thus $\boxed{w_7^* = 0.5112}$

Compute Correction For w_8

We have.

$$\frac{\partial E_{Total}}{\partial w_8} = \frac{\partial E_{Total}}{\partial out_{o2}} \times \frac{\partial out_{o2}}{\partial net_{o2}} \times \frac{\partial net_{o2}}{\partial w_8}$$

Now from previous computation for w_8 we have

① $$\frac{\partial E_{total}}{\partial out_{o2}} = -0.217$$

② $$\frac{\partial out_{o2}}{\partial net_{o2}} = 0.175$$

③ and we have

$$net_{o2} = h_{1out} \times w_7 + h_{2out} \times w_8 + b_2.$$

Thus we have $$\frac{\partial net_{o2}}{\partial w_8} = h_{2out} = 0.5968$$

Thus we have.

$$\frac{\partial E_{Total}}{\partial w_8} = -0.217 \times 0.175 \times 0.5908$$
$$= -0.02266$$

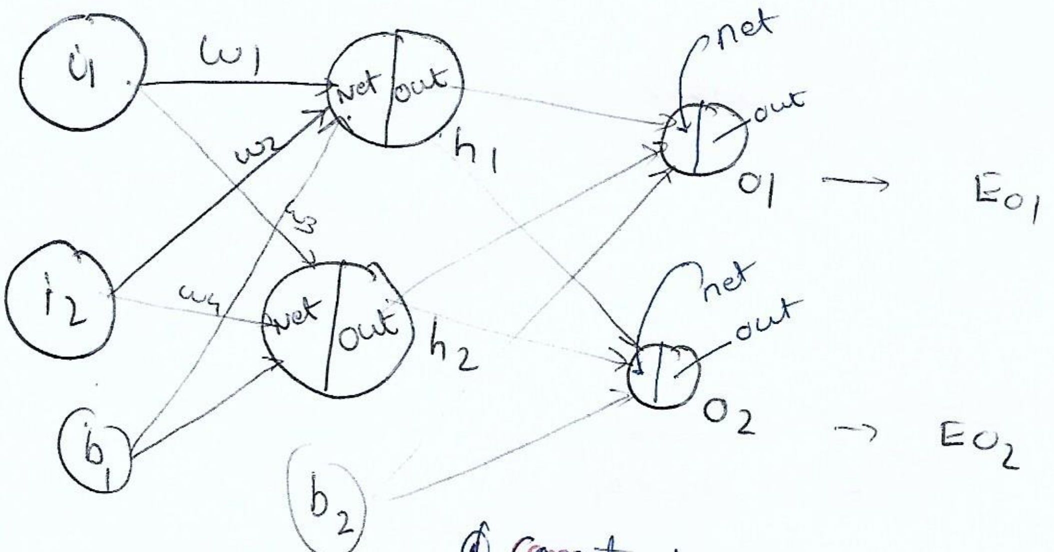
Correction to w_8

$$w_8^* = w_8 - \eta \times \frac{\partial E_{Total}}{\partial w_8}$$

$$= 0.55 - 0.5 \times (-0.02267)$$

Thus $w_8^* = 0.561$

W/s computation for hidden Layer



Compute weight correction for w_1

We wish to find:

Now $E_{Total} = E_{O1} + E_{O2}$.

$$\frac{\partial E_{Total}}{\partial w_1} = \overset{①}{\frac{\partial E_{Total}}{\partial out_{h1}}} \times \overset{②}{\frac{\partial out_{h1}}{\partial net_{h1}}} \times \overset{③}{\frac{\partial net_{h1}}{\partial w_1}}$$

(*) Note

$$\frac{\partial E_{Total}}{\partial w_1} = \frac{\partial E_1}{\partial w_1} + \frac{\partial E_2}{\partial w_1}$$

Consider each term separately:

We have.

① A

$$\frac{\partial E_{o1}}{\partial out_{h1}} = \overbrace{\frac{\partial E_{o1}}{\partial net_{o1}}}^{term1} \times \overbrace{\frac{\partial net_{o1}}{\partial out_{h1}}}^{term2}$$

Now $\overset{term2}{\frac{\partial net_{o1}}{\partial out_{h1}}}$ | $net_{o1} = w_5 \times h_{1out} + w_6 \times h_{2out} + b_2$

Thus $\overset{term2}{\frac{\partial net_{o1}}{\partial out_{h1}}} = w_5$

term 1

→ computed earlier for w_5

$$\frac{\partial E_{o1}}{\partial net_{o1}} = \left[\frac{\partial E_{o1}}{\partial out_1}\right] \times \left[\frac{\partial out_1}{\partial net_{o1}}\right] = 0.741 \times 0.187$$

$$= 0.138$$

Thus

(1) A

$$\frac{\partial E_{01}}{\partial outh_1} = 0.138 \times w_5 = 0.138 \times 0.4 = 0.055$$

Now compute for (1) B

$$\frac{\partial E_{O2}}{\partial outh_1} = \overbrace{\frac{\partial E_{O2}}{\partial net_{O2}}}^{term1} \times \overbrace{\frac{\partial net_{O2}}{\partial outh_1}}^{term2}$$

term 1

$$\frac{\partial E_{O2}}{\partial net_{O2}} = \frac{\partial E_{O2}}{\partial out_2} \times \frac{\partial out_2}{\partial Net_{O2}}.$$

Computed Earlier For w_7

$$= -0.217 \times 0.175$$

$$\frac{\partial E_{O2}}{\partial net_{O2}} = -0.0379$$

term2

$$\frac{\partial Net_{O2}}{\partial outh_1}$$

$$Net_{O2} = w_7 \times h_1out + w_8 \times h_2out + b_2$$

$$\therefore \frac{\partial Net_{O2}}{\partial outh_1} = w_7$$

Thus

$$\frac{\partial E_{O2}}{\partial outh_1} = -0.0379 \times 0.5$$
$$= -0.01895$$

Thus (1)

$$\frac{\partial E_{Total}}{\partial outh_1} = \frac{\partial E_1}{\partial outh_1} + \frac{\partial E_2}{\partial outh_1}$$

$$= 0.055 - 0.0189$$

$$= 0.036$$

Next consider the term

2

$$\frac{\partial out_{h1}}{\partial net_{h1}} \quad \Bigg| \quad out_{h1} = \frac{1}{1+e^{-net_{h1}}}$$

Thus $$\frac{\partial out_{h1}}{\partial net_{h1}} = out_{h1}(1-out_{h1})$$
$$= 0.593(1-0.593)$$
$$= 0.241$$

and Finally

3

$$\frac{\partial Net_{h1}}{\partial w_1} \quad \Bigg| \quad Net_{h1} = w_1 i_1 + w_2 * i_2 + b_1$$
$$\frac{\partial Net_{h1}}{\partial w_1} = i = 0.05.$$

Thus we have

$$\frac{\partial E_{Total}}{\partial w_1} = \frac{\partial E_{total}}{\partial out_{h1}} \times \frac{\partial out_{h1}}{\partial net_{h1}} \times \frac{\partial net_{h1}}{\partial w_1}$$
$$= 0.036 \times 0.241 \times 0.05$$
$$= 0.00043$$

Now updating w_1

we have $$w_1^* = w_1 - \eta \frac{\partial E_{total}}{\partial w_1}$$ Learning rate = α 0.5

or $$w_1^* = 0.15 - 0.5 \times 0.00043$$

Thus. $$w_1^* = 0.149$$

Exercise:] Repeat Procedure For w_2, w_3, w_4
To get $w_2^* = 0.199$, $w_3^* = 0.249$, $w_4^* = 0.299$.