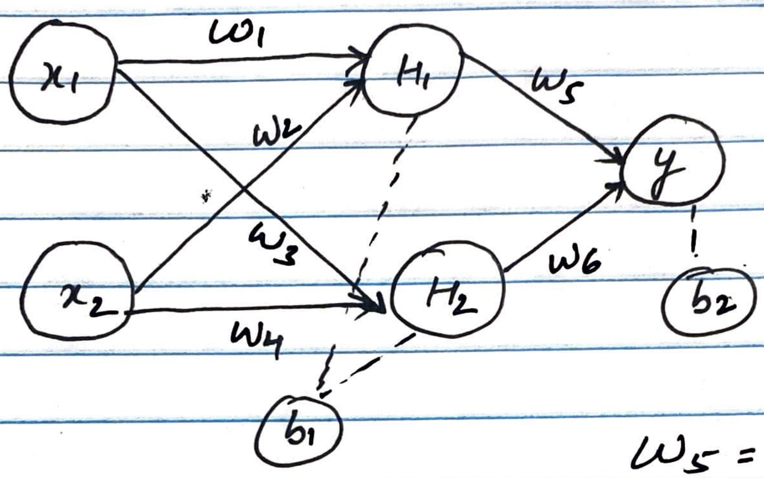
016978741

HW-2

Example: The figure shows 2 layers, feed forward neural network with 2 hidden layer nodes & 1 output node. x_1 & x_2 are 2 inputs.



$x_1 = 1$; $w_1 = 0.5$

$x_0 = 0$; $w_3 = 0.5$

$w_2 = -0.25$

$w_4 = -0.25$

$w_5 = 1$; $w_6 = -1$

$b_1 = b_2 = 0.1$

$y_{Target} = 1$

Forward Pass

$$H_1 = w_1 x_1 + w_2 x_2 + b_1$$
$$= 0.5 \times 1 + (-0.25 \times 0) + 0.1 = 0.5 + 0.1$$
$$\boxed{H_1 = 0.6}$$

Activation f^n is sigmoid $= 1/1+e^{-x}$

↳ on H_1 = out-H_1 $= \dfrac{1}{1+e^{-H_1}} = \dfrac{1}{1+e^{-0.6}} = 0.646$

$$H_2 = w_3 x_1 + w_4 x_2 + b_1$$
$$= 0.5 \times 1 + (-0.25 \times 0) + 0.1 = 0.5 + 0.1$$
$$H_2 = 0.6$$

out $H_2 = \dfrac{1}{1+e^{-0.6}} = 0.646$

out $H_2 = 0.646$

$y = out\, H_1 \cdot w_5 + out\, H_2 \cdot w_6 + b_2$

$\Rightarrow 0.646 \times 1 + 0.646 \times -1 + 0.1$

$\Rightarrow 0.646 - 0.646 + 0.1$

$\Rightarrow 0.1$

$out\, y = \frac{1}{1+e^{-y}} = \frac{1}{1+e^{-0.1}} \Rightarrow 0.525$

$y_{target} = 1$ where as $y_{out} = 0.525$

Calculating Error

Loss f^n = MSE $= \frac{1}{2}(y-\hat{y})^2$

$MSE = \frac{1}{2}(1-0.525)^2$

$\Rightarrow \frac{1}{2}(0.2256) = \underline{\underline{0.113}}$

So far, we have

$out\, H_1 = 0.646$ $\quad out\, y = 0.525$

$out\, H_2 = 0.646$ $\quad$ Error $= 0.113$

Back Propogation

Error at $w_5 = \frac{\partial E}{\partial w_5} = \frac{\partial E}{\partial w_5}$ (1)

To find $\frac{\partial E}{\partial w_5} = \frac{\partial E}{\partial out\,y} \times \frac{\partial out\,y}{\partial y} \times \frac{\partial y}{\partial w_5}$

$$\frac{\partial E}{\partial out_y} = \frac{\partial}{\partial out_y}\left(\tfrac{1}{2}(y-\hat{y})^2\right)$$

$$= -2 \times \frac{1}{2}(y-\hat{y})^{2-1} = -(y-\hat{y}) = -(1-0.525)$$

$$\boxed{\frac{\partial E}{\partial out_y} = -0.475}$$

$$\frac{\partial out_y}{\partial y} = \frac{\partial}{\partial y}\left(1/1+e^{-x}\right) = out_y(1-out_y)$$

$$= 0.525(1-0.525)$$

$$= 0.525 \times 0.475$$

$$\boxed{\frac{\partial out_y}{\partial y} = 0.250}$$

$$\frac{\partial y}{\partial w_5} = \frac{\partial}{\partial w_5}\left(out\,H_1 \cdot w_5 + out\,H_2 \cdot w_6 + b_2\right)$$

$$\boxed{\frac{\partial y}{\partial w_5} = 0.646}$$

Now, we have all values to put in eq (1)

$$\frac{\partial E}{\partial w_5} = -0.475 \times 0.250 \times 0.646$$

$$\boxed{\frac{\partial E}{\partial w_5} = -0.077} \rightarrow \text{change in } w_5$$

update w_5 now,

$$w_{5\,new} = w_5 - \eta \times \frac{\partial E}{\partial w_5}$$

Taking, η (learning rate) $= 0.1$

$$w_{5\,new} = 1 - (0.1(-0.077))$$

$$= 1.0077$$

$$\boxed{w_{5\,new} = 1.0077}$$

For w_6:

$$\frac{\partial E}{\partial w_6} = \frac{\partial E}{\partial out_y} \times \frac{\partial out_y}{\partial y} \times \frac{\partial y}{\partial w_6} \quad (2)$$

we have calculated $\frac{\partial E}{\partial out_y}$, $\frac{\partial out_y}{\partial y}$

$$\frac{\partial E}{\partial out_y} = -0.475 \quad ; \quad \frac{\partial out_y}{\partial y} = 0.250$$

$$\frac{\partial y}{\partial w_6} = \frac{\partial}{\partial w_6} (out H_1 \cdot w_5 + out H_2 \cdot w_6 + b_2)$$

$$= out H_2 = 0.646$$

Putting these in (2)

$$\frac{\partial E}{\partial w_6} = -0.475 \times 0.250 \times 0.646$$

$$= -0.077$$

Taking η (LR) $= 0.1$

updating the weight w_6

$$w_{6\,new} = w_6 - \eta \left(\frac{\partial E}{\partial w_8} \right)$$

$$= -1 - (0.1 \times (-0.077))$$

$$\boxed{w_{6\,new} = -0.992}$$

We have updated weights w_5 & w_6, now we got to update w_1, w_2, w_3, w_4 at Hidden layer.

For w_1:

$$\frac{\partial E}{\partial w_1} = \frac{\partial E}{\partial out H_1} \times \frac{\partial out H_1}{\partial H_1} \times \frac{\partial H_1}{\partial w_1}$$

$$\frac{\partial E}{\partial out\, H_1} = \boxed{\frac{\partial E}{\partial out_y} \times \frac{\partial out_y}{\partial y}} \times \frac{\partial y}{\partial out\, H_1}$$

$$\text{So } \frac{\partial E}{\partial w_1} = \frac{\partial E}{\partial out_y} \times \frac{\partial out_y}{\partial y} \times \frac{\partial y}{\partial out H_1} \times \frac{\partial out H_1}{\partial H_1} \times \frac{\partial H_1}{\partial w_1} \rightarrow (3)$$

We know,

$$\frac{\partial E}{\partial out_y} = -0.475, \quad \frac{\partial out_y}{\partial y} = 0.250$$

$$\frac{\partial y}{\partial out\, H_1} = \frac{\partial}{\partial out H_1}(out\, H_1 \cdot w_5 + out H_2 \cdot w_6 + b_2)$$

$$= w = 1$$

$$\boxed{\frac{\partial y}{\partial out H_1} = 1}$$

$$\frac{\partial out\, H_1}{\partial H_1} = \frac{\partial (1/1+e^{-z})}{\partial H_1} = out\, H_1 (1 - out\, H_1)$$

$$= 0.646(1 - 0.646)$$

$$= 0.229$$

$$\frac{\partial H_1}{\partial w_1} = \frac{\partial}{\partial w_1}(w_1 x_1 + w_2 x_2 + b_1)$$

$= x_1 = 1$ Now, put all values in (3)

$$\frac{\partial E}{\partial w_1} = -0.475 \times 0.250 \times 1 \times 0.229 \times 1$$

$$= -0.027$$

Taking $\eta = 0.1$

$$w_1\, new = w_1 - \eta \times \frac{\partial E}{\partial w_1} \qquad w_1 = 0.5$$

$$= 0.5 - (0.1 \times (-0.027))$$

$$\boxed{w_{1,new} = 0.50227}$$

For w_2 : $\frac{\partial E}{\partial w_2} = \frac{\partial E}{\partial out_{H_1}} \times \frac{\partial out_{H_1}}{\partial H_1} \times \frac{\partial H_1}{\partial w_2}$

(3) where $\frac{\partial E}{\partial out_{H_1}} = \frac{\partial E}{\partial out_y} \times \frac{\partial out_y}{\partial y} \times \frac{\partial y}{\partial out_{H_1}}$

So $\frac{\partial E}{\partial w_2} = \frac{\partial E}{\partial out_y} \times \frac{\partial out_y}{\partial y} \times \frac{\partial y}{\partial out_{H_1}} \times \frac{\partial out_{H_1}}{\partial H_1} \times \frac{\partial H_1}{\partial w_2}$ → (4)

we know,

$\frac{\partial E}{\partial out_y} = -0.475$, $\frac{\partial out_y}{\partial y} = 0.250$, $\frac{\partial y}{\partial out_{H_1}} = 1$

$\frac{\partial out_{H_1}}{\partial H_1} = 0.229$

$\frac{\partial H_1}{\partial w_2} = \frac{\partial}{\partial w_2}(w_1 x_1 + w_2 x_2 + b_1)$

$= x_2 = 0$

Putting these all values in eq. (4)

$\frac{\partial E}{\partial w_2} = 0$; Taking $\eta = 0.1$

$w_{2new} = w_2 - \eta \times \frac{\partial E}{\partial w_2}$

$= -0.25 - 0 = -0.25$

$$\boxed{w_{2new} = -0.25}$$

For w_3

$$\frac{\partial E}{\partial w_3} = \frac{\partial E}{\partial \, out\,H_2} \times \frac{\partial \, out\,H_2}{\partial H_2} \times \frac{\partial H_2}{\partial w_3}$$

where $$\frac{\partial E}{\partial \, out\,H_2} = \frac{\partial E}{\partial \, out\,y} \times \frac{\partial \, out\,y}{\partial y} \times \frac{\partial y}{\partial \, out\,H_2}$$

Hence,

$$\frac{\partial E}{\partial w_3} = \frac{\partial E}{\partial \, out\,y} \times \frac{\partial \, out\,y}{\partial y} \times \frac{\partial y}{\partial \, out\,H_2} \times \frac{\partial \, out\,H_2}{\partial H_2} \times \frac{\partial H_2}{\partial w_3} \quad \rightarrow (5)$$

We know,

$$\frac{\partial E}{\partial \, out\,y} = -0.475, \quad \frac{\partial \, out\,y}{\partial y} = 0.250$$

$$\frac{\partial y}{\partial \, out\,H_2} = \frac{\partial}{\partial \, out\,H_2} (out\,H_1 \cdot w_5 + out\,H_2 \cdot w_6 + b_2)$$

$$= w_6 = -1$$

$$\frac{\partial \, out\,H_2}{\partial H_2} = \frac{\partial}{\partial H_2} (1/1+e^{-x}) = out\,H_2 (1 - out\,H_2)$$

$$= 0.646 (1 - 0.646)$$

$$\frac{\partial \, out\,H_2}{\partial H_2} = 0.229$$

$$\frac{\partial H_2}{\partial w_3} = \frac{\partial}{\partial w_3} (w_3 x_1 + w_4 x_2 + b_1)$$

$= 1 \cdot x_1 = 1$ Putting all these values in (5)

$$\frac{\partial E}{\partial w_3} = -0.475 \times 0.250 \times -1 \times 0.229 \times 1$$

$$= 0.0271$$

Taking $\eta = 0.1$

$$w_{3\,new} = w_3 - \eta \left(\frac{\partial E}{\partial w_3}\right)$$

$$w_{3\,new} = 0.5 - (0.1 \times 0.0271)$$

$$= 0.5 - 0.00271$$

$$\boxed{w_{3\,new} = 0.4973}$$

For w_4: $\dfrac{\partial E}{\partial w_4} = \dfrac{\partial E}{\partial out_y} \times \dfrac{\partial out_y}{\partial y} \times \dfrac{\partial y}{\partial outH_2} \times \dfrac{\partial outH_2}{\partial H_2} \times \dfrac{\partial H_2}{\partial w_4} \rightarrow (6)$

We know,

$\dfrac{\partial E}{\partial out_y} = -0.475$, $\dfrac{\partial out_y}{\partial y} = 0.250$, $\dfrac{\partial y}{\partial outH_2} = -1$

$\dfrac{\partial outH_2}{\partial H_2} = 0.229$

$$\frac{\partial H_2}{\partial w_4} = \frac{\partial}{\partial w_4}(w_3 x_1 + w_4 x_2 + b_1)$$

$$= x_2 = 0$$ Putting all in eq. (6)

$$\therefore \frac{\partial E}{\partial w_4} = -0.475 \times 0.250 \times -1 \times 0.229 \times 0$$

$$\frac{\partial E}{\partial w_4} = 0$$

Taking $\eta = 0.1$

$$w_{4\,new} = w_4 - \eta\left(\frac{\partial E}{\partial w_4}\right)$$

$$= w_4 - 0.00 = -0.25 - 0$$

$$= -0.25$$

$$\boxed{w_{4\,new} = -0.25}$$

For b_1 :

$$\frac{\partial E}{\partial b_1} = \frac{\partial E}{\partial out_y} \times \frac{\partial out_y}{\partial y} \times \frac{\partial y}{\partial out_{H_1}} \times \frac{\partial out_{H_1}}{\partial H_1} \times \frac{\partial H_1}{\partial b_1} \rightarrow (7)$$

we know

$$\frac{\partial E}{\partial out_y} = -0.475 \; ; \; \frac{\partial out_y}{\partial y} = 0.250$$

$$\frac{\partial y}{\partial out_{H_1}} = 1 \; ; \; \frac{\partial out_{H_1}}{\partial H_1} = 0.229$$

$$\frac{\partial H_1}{\partial b_1} = \frac{\partial}{\partial b_1}(w_1 x_1 + w_2 x_2 + b_1)$$

$$= 1$$

Therefore, Putting all in (7)

$$\frac{\partial E}{\partial b_1} = \frac{\partial E}{\partial out_y} \times \frac{\partial out_y}{\partial y} \times \frac{\partial y}{\partial out_{H_1}} \times \frac{\partial out_{H_1}}{\partial H_1} \times \frac{\partial H_1}{\partial b_1}$$

$$= -0.475 \times 0.250 \times 1 \times 0.229 \times 1$$

$$= -0.027$$

$$b_1 new = b_1 - \eta\left(\frac{\partial E}{\partial b_1}\right) \qquad \eta = 0.1$$

$$= 0.1 - (0.1 \times (-0.027))$$

$$\boxed{b_1 new = 0.1027}$$

For b_2 :

$$\textcircled{1} \leftarrow \frac{\partial E}{\partial b_2} = \frac{\partial E}{\partial out_y} \times \frac{\partial out_y}{\partial y} \times \frac{\partial y}{\partial b_2} \rightarrow \textcircled{8}$$

We already know the value of,

$$\frac{\partial E}{\partial out_y} = -0.475 \; ; \; \frac{\partial out_y}{\partial y} = 0.250$$

$$\frac{\partial y}{\partial b_2} = \frac{\partial}{\partial b_2}(out\text{-}H_1 \cdot w_5 + out\text{-}H_2 \cdot w_6 + b_2)$$

$$= 1$$

Substituting all values in eq. 8

$$\frac{\partial E}{\partial b_2} = \frac{\partial E}{\partial out_y} \times \frac{\partial out_y}{\partial y} \times \frac{\partial y}{\partial b_2}$$

$$= -0.475 \times 0.250 \times 1$$

$$\boxed{\frac{\partial E}{\partial b_2} = -0.119}$$

Taking $\eta = 0.1$

$$b_{2\,new} = b_2 - \eta\left(\frac{\partial E}{\partial b_2}\right)$$

$$= 0.1 - (0.1 \times (-0.119))$$

$$\boxed{b_{2\,new} = 0.112}$$

We have calculated the new weights using back propogation.

Updated weights after one iteration are summarised –

old weight		new weight	
w_1 =	0.5	w_1 =	0.5027
w_2 =	-0.25	w_2 =	-0.25
w_3 =	+ ~~0.25~~ 0.5	w_3 =	0.4973
w_4 =	-0.25	w_4 =	-0.25
w_5 =	1	w_5 =	1.0077
w_6 =	-1	w_6 =	-0.992
b_1 =	0.1	b_1 =	0.1027
b_2 =	0.1	b_2 =	0.112

We see that after back propogation weights w_2 & w_4, did not change as $x_2 = 0$ i.e one of the input $x_2 = 0$ was zero.
We also, see that weights are updated very little for rest of the weights. Here we took learning rate $= 0.1$, however that can be adjusted.