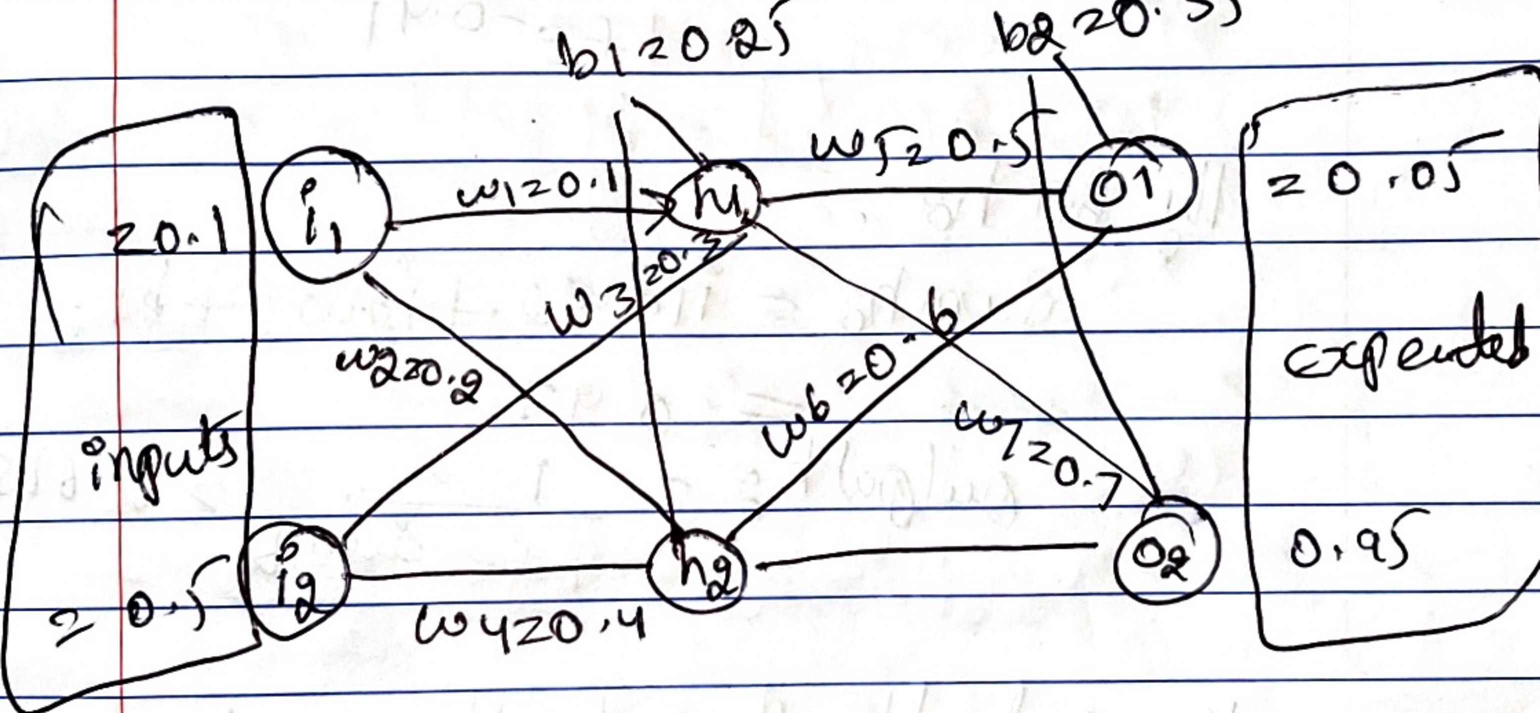
Rohith Reddy Vangala

016762107

→ Step by step forward pass and Backpropagation example.



Forward pass.

lets start with h_1

$\text{sum } h_1 = i_1 * w_1 + i_2 * w_3 + b_1$

$\text{sum } h_1 = 0.1 * 0.1 + 0.5 * 0.3 + 0.25$

$= 0.41$

No we pass this weighted sum through the logistic function (sigmoid) to squash the weighted sum into the range (0, 1).

$$\text{output } h_1 = \frac{1}{1+e^{-\text{sum}h_1}}$$

$$\text{output } h_1 = \frac{1}{1+e^{-0.41}} = 0.60108$$

Ily for h_2.

$$\text{sum}h_2 = i_1 * w_2 + i_2 w_4 + b_1$$

$$= 0.47$$

$$\text{output } h_2 = \frac{1}{1+e^{-\text{sum}h_2}} = 0.61538$$

Now outputh_1 & outputh_2 will be considered as i/p to next layer.

for O_1

$$\text{sum}O_1 = \text{output } h_1 * w_5 + \text{output}h_2 * w_6 + b_2$$

$$= 1.01977$$

$$\text{output}_{O_1} = \frac{1}{1+e^{-\text{sum}_{O_1}}} = 0.73492$$

lly for O_2

$\text{sum}_{O_2} = \text{output}_{h_1} * w_7 + \text{output}_{h_2} * w_8 + b_2 = 1.26306$

$$\text{output}_{O_2} = \frac{1}{1+e^{-\text{sum}_{O_2}}}$$

$$= 0.77955$$

computing the total error

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

$$E_1 = \frac{1}{2}(\text{target}_1 - \text{output}_1)^2$$

$$= \frac{1}{2}(0.05 - 0.73492)^2$$

$$= 0.23456.$$

lly for E_2.

$$E_2 = \frac{1}{2}(\text{target}_2 - \text{output}_{O_2})^2$$

$$E_2 = \frac{1}{2}(0.95 - 0.77955)^2$$

$$= 0.01452.$$

$\therefore E_{total} = E_1 + E_2$
$= 0.24908$.

Back propagation
(for weights in o/p layer)
(w_5, w_6, w_7, w_8)

for w_5

$\hookrightarrow \dfrac{\partial E_{total}}{\partial w_5} = \dfrac{\partial E_{total}}{\partial output_{o1}} * \dfrac{\partial output_{o1}}{\partial sum_{o1}} * \dfrac{\partial sum_{o1}}{\partial w_5}$ ← *

consider

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

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

$\therefore \dfrac{\partial E_{total}}{\partial output_{o1}} = 2 * \frac{1}{2}(target_1 - output_{o1}) * (-1)$

$= output_{o1} - target_1$

(2) $\rightarrow$ $\sigma(x) = \frac{1}{1+e^{-x}}$.

$\frac{\partial}{\partial x}(\sigma(x)) = \sigma(x)(1-\sigma(x))$.

$\frac{\partial\, output_{o1}}{\partial\, sum_{o1}} = output_{o1}(1 - output_{o1})$

(3) $\rightarrow$ $sum_{o1} = output_{h1} * w_5 + output_{h2} * w_6 + b_2$.

$\therefore \frac{\partial\, sum_{o1}}{\partial w_5} = output_{h1}$

Now putting (1), (2), (3) in the (*) eq

$\frac{\partial E_{total}}{\partial w_5} = [output - target_1] * [output_{o1}(1 - output_{o1}) * output_{h1}]$

$= 0.68492 * 0.19480 * 0.60108$

$\frac{\partial E_{total}}{\partial w_5} = 0.08020$

$new\_w_5 = w_5 - n * \frac{\partial E_{total}}{\partial w_5}$

η = learning Rate

$new\_w_5 = 0.5 - 0.6 * 0.08020$

$new\_w_5 = 0.45187$

Hly

for w_6

$$\frac{\partial E_{total}}{\partial w_6} = \frac{\partial E_{total}}{\partial output_{o1}} * \frac{\partial output_{o1}}{\partial sum_{o1}} * \frac{\partial sum_{o1}}{\partial w_6}$$

$$= 0.68492 * 0.19480 * 0.61538$$

$$= 0.08211$$

$$new\_w_6 = w_6 - \eta * \frac{\partial E_{total}}{\partial w_6}$$

$$= 0.6 - 0.6 * 0.08211$$

$new\_w_6 = 0.55073$

For w_7

$$\frac{\partial E_{total}}{\partial w_7} = \frac{\partial E_{total}}{\partial output_{o2}} * \frac{\partial output_{o2}}{\partial sum_{o2}} * \frac{\partial sum_{o2}}{\partial w_7}$$

$$\frac{\partial E_{total}}{\partial output_{o2}} = output_{o2} - target_{2}$$

for the second component

$$\frac{\partial output_{o2}}{\partial sum_{o2}} = output_{o2}(1 - output_{o2})$$

for the third component

$$\frac{\partial sum_{o2}}{\partial w_7} = output_{h1}$$

putting all together

$$\frac{\partial E_{total}}{\partial w_7} = [output_{o2} - target_2] * [output_{o2}(1 - output_{o2})] * [output_{h1}]$$

$$= -0.17044 * 0.17184 * 0.60108$$

$$\frac{\partial E_{total}}{\partial w_7} = -0.01760$$

$$new\_w_7 = w_7 - \eta * \frac{\partial E_{total}}{\partial w_7}$$

$$= 0.7 - 0.6 * -0.01760$$

$$new\_w_7 = 0.71056.$$

Ily for w_8
$new\ w_8 = 0.81081$ (with $\frac{\partial E_{total}}{\partial w_8} = -0.01802$.

for weights in Hidden layer (w_1, w_2, w_3, w_4)

for w_1
$$\frac{\partial E_1}{\partial w_1} = \frac{\partial E_1}{\partial output\ o_1} * \frac{\partial output\ o_1}{\partial sum\ o_1} * \frac{\partial output\ h_1}{\partial sum\ h_1} * \frac{\partial sum\ h_1}{\partial w_1}$$

$$\frac{\partial E_1}{\partial output\ o_1} = Output\ o_1 - target\ o_1$$

$Sum\ o_1 = Output\ h_1 * w_5 + Output\ h_2 * w_2 + w_6 + b_2$

$$\frac{\partial sum\ o_1}{\partial output\ h_1} = w_5.$$

$$\frac{\partial output\ h_1}{\partial sum\ h_1} = output\ h_1 * (1 - output\ h_1)$$

for fifth component

$sum\ h_1 = i_1 * w_1 + i_2 * w_3 + b_1$

$$\frac{\partial sum\ h_1}{\partial w_1} = i_1$$

Putting all together

$$\frac{\partial E_1}{\partial w_1} = \frac{\partial E_1}{\partial output\ o_1} * \frac{\partial output\ o_1}{\partial sum\ o_1} * \frac{\partial sum\ o_1}{\partial output} * \frac{\partial output\ h_1}{\partial sum\ h_1} * \frac{\partial sum\ h_1}{\partial w_1}$$

$= 0.68492 * 0.198180 * 0.5 * 0.239778 * 0.1$

$= 0.00159$.

lly for w_1 (w.r.t. to E_2)

$$\frac{\partial E_2}{\partial w_1} = \frac{\partial E_2}{\partial output\ o_2} * \frac{\partial output\ o_2}{\partial sum\ o_2} * \frac{\partial sum\ o_2}{\partial output\ h_1} * \frac{\partial output\ h_1}{\partial sum\ h_1} * \frac{\partial sum\ h_1}{\partial w_1}$$

$$\frac{\partial E_2}{\partial output\ o_2} = output\ o_2 - target_2.$$

$sum\ o_2 = output\ h_1 * w_7 + output\ h_2 * w_8 + b_2$.

$$\frac{\partial sum\ o_2}{\partial output\ h_1} = w_7,$$

Putting all together:

$$\frac{\partial E_2}{\partial w_1} = \frac{\partial E_2}{\partial output_{o2}} * \frac{\partial output_{o2}}{\partial sum_{o2}} * \frac{\partial sum_{o2}}{\partial output_{h1}} * \frac{\partial output_{h1}}{\partial sum_{h1}} * \frac{\partial sum_{h1}}{\partial w_1}$$

$$\frac{\partial E_2}{\partial w_1} = -0.17044 * 0.17184 * 0.7 * 0.23978 * 0.1$$

$$= -0.00049$$

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

$$\frac{\partial E_{total}}{\partial w_1} = 0.00159 + (-0.00049)$$

$$= 0.00110$$

The new_w_1 is

$$new\_w_1 = w_1 - \eta * \frac{\partial E_{total}}{\partial w_1}$$

$$= 0.1 - 0.6 * 0.00110$$

$$new\_w_1 = 0.09933$$

proceeding lly we can easily update the other weights (w_2, w_3 and w_4).

new_w2 = 0.19919
new_w3 = 0.299667
new_w4 = 0.39597

So we have computed all the weights now, we have updated these weights to the old weights ∴ one back propagation cycle is finished.
→ This process goes until the loss value converges to minima and finds the optimal value.