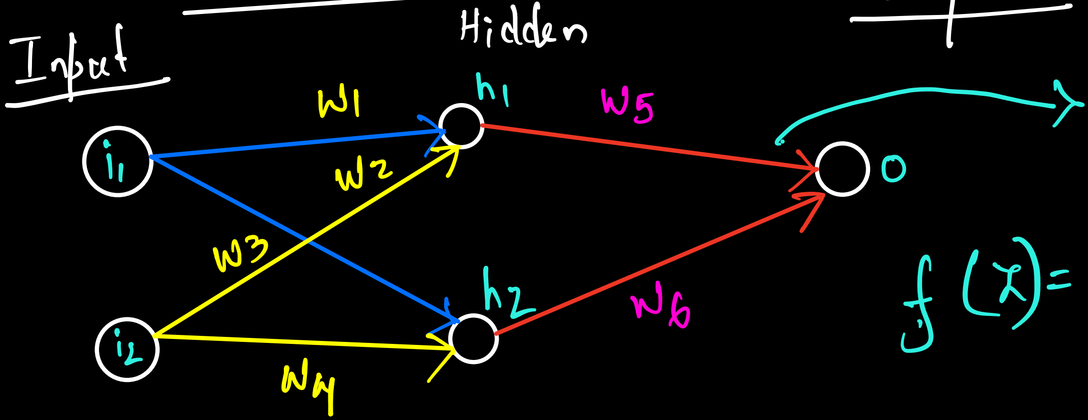
Today's Agenda

1) Recap :- Forward Propagation

2) Backpropagation

Network



Sigmoid

$$f(x) = \frac{1}{1 + e^{-x}}$$

$$f'(x) = f(x) * (1 - f(x))$$

Forward

$x_1 (i_1) = 0.05$

$x_2 (i_2) = 0.10$

$Y = 0.01$

I 2 H
$$\left\{ \begin{array}{l} w_1 = 0.15 \\ w_2 = 0.20 \\ w_3 = 0.25 \\ w_4 = 0.30 \end{array} \right.$$

H 2 O
$$\left\{ \begin{array}{l} w_5 = 0.40 \\ w_6 = 0.45 \end{array} \right.$$

$b_1 = 0.35$

o $b_2 = 0.60$

Forward

$\{wx+b\}$ $\{y = mx + c\}$

$$z\_h_1 = x_1 \cdot w_1 + x_2 \cdot w_2 + b_1$$
$$= 0.05 \times 0.15 + 0.1 \times 0.2 + 0.35$$
$$= 0.3775$$

$$z\_h_2 = x_1 \cdot w_3 + x_2 \cdot w_4 + b_1$$
$$= 0.05 \times .25 + 0.1 \times 0.3 + 0.35$$
$$= 0.3925$$

$$a\_h_1 = \text{Sigmoid}(0.3775) \approx 0.5933$$
$$a\_h_2 = \text{Sigmoid}(0.3925) \approx 0.5969$$

Hidden → Output Layer

$$z_{out} = a\_h_1 w_5 + a\_h_2 w_6 + b_2$$
$$= 0.5933 \times 0.40 + 0.5969 \times 0.45 + 0.60$$
$$= 1.1059$$

$a\_out = \text{sigmoid}(1.1059) \approx 0.7514$

Predicted Output $= 0.7514$

Compute Loss

$$\text{MSE Loss} = 0.5 \times (Y\_true - Y\_hat)^2$$

$$= 0.5 \times (0.01 - 0.7514)^2$$

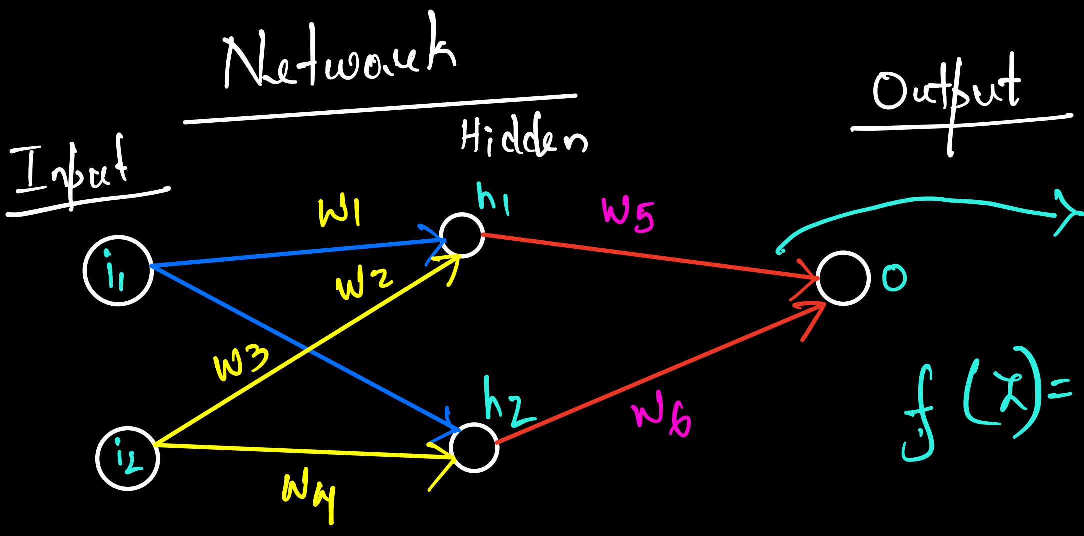
$$= 0.2748$$

Backpropagation

* update w & b, to reduce the error.

W_5

$$\frac{\partial L}{\partial w_5} = \frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z_{out}} \cdot \frac{\partial z_{out}}{\partial w_5}$$



Sigmoid

$$f(x) = \frac{1}{1 + e^{-z}}$$

$$f'(x) = f(x) * (1 - f(x))$$

$\frac{\partial L}{\partial y'}$ = Derivative of loss wrt predicted output

$$Loss = 0.5 \times (y - \hat{y})^2$$

$\frac{\partial y'}{\partial z_{out}}$ = Derivative of output activation wrt sum (z_{out})

$$Sigmoid = 1/1 + (1 + e^{-z})$$

Derivative

$= \text{Sigmoid}(2) * (1 - \text{Sigmoid}(2))$

$= \text{Y\_hat} * (1 - \text{Y\_hat})$

$= 0.7514 * (1 - 0.7514)$

$\dfrac{\partial y^{hat}}{\partial z_{out}} = 0.1868$

$\dfrac{\partial L}{\partial y^{hat}} = (\text{Y\_hat} - \text{Y\_true}) = 0.7514 - 0.01$

$= 0.7414$

$\dfrac{\partial z_{out}}{\partial w_5}$ = Derivative of the output neuron weighted sum

Output neuron $(z_{out}) = w_5 * a_{h_1} + w_6 * a_{h_2} + b_2$

For four constant

$$\frac{d}{dx}(x \cdot c) = c \quad (\text{four constant } c)$$

$$\frac{\partial z_{out}}{\partial w_5} = a\_h_1 = 0.5933$$

Combine,

$$\frac{\partial Loss}{\partial w_5} = \frac{\partial L}{\partial y^{hat}} \cdot \frac{\partial y^{hat}}{\partial z_{out}} \cdot \frac{\partial z_{out}}{\partial w_5}$$

$$= 0.7414 \times 0.1868 \times 0.5933$$

$$= 0.0822$$

Weight & Bias

$$\left\{ \begin{aligned} W_{new} &= W_{old} - \eta \frac{\partial L}{\partial W_{old}} \\ .0001 \leftarrow LR \\ B_{new} &= B_{old} - \eta \frac{\partial L}{\partial B_{old}} \end{aligned} \right\}$$

$$\text{New } W_5 = 0.40 - (.0001 \times 0.0822)$$

$$\text{New } W_5 = 0.40 - (.0000000)$$

$$= 0.399927$$

W_6

$$\frac{\partial Loss}{\partial w_6} = \frac{\partial Loss}{\partial y^{hat}} * \frac{\partial y^{hat}}{\partial z_{out}} \times \frac{\partial z_{out}}{\partial w_6}$$

$$\frac{\partial z_{out}}{\partial w_6} = a\_h_2 = 0.5969$$

Finally :-

$$\frac{\partial L}{\partial w_6} = 0.7414 \times 0.1868 \times 0.5969$$

$$= 0.0827$$

$$\text{New } w_6 = 0.45 - \eta \frac{\partial L}{\partial w_6}$$

$$= 0.45 - (.0001 \times 0.0827)$$

$$\text{New } w_6 = 0.4499$$

$\underline{\underline{W_1}}$

$$\frac{\partial Loss}{\partial w_1} = \frac{\partial Loss}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z_{out}} \cdot \frac{\partial z_{out}}{\partial a\_h_1} \cdot \frac{\partial a\_h_1}{\partial z\_h_1} \cdot \frac{\partial z\_h_1}{\partial w_1}$$

$z\_h_1 \rightarrow$ without activation

a_h1 → with activation

$$\frac{\partial z_{out}}{\partial a\_h_1} =$$

$$z_{out} = w_5 * ah_1 + w_6 \times a\_h_2 + b_2$$

$$\frac{\partial z_{out}}{\partial a\_h_1} = w_5 = 0.40$$

$$\frac{\partial\_ah_1}{\partial z_h} = \left\{ a\_h_1 * (1 - a\_h_1) \right\}$$

$$a\_h_1 = (0.5933 * (1 - 0.5933))$$

$$= 0.2413$$

$$\frac{\partial z h_1}{\partial w_1} =$$

$$z_{h_1} = w_1 \times x_1 + w_2 \times x_2 + b_1$$

$$z_{h_1} = x_1 = 0.05$$

<u>Finally w_1</u>

$$= 0.7414 \times 1868 = 0.1385$$

$$= 0.1385 \times 0.40 = 0.0554$$

$$= 0.0554 \times 0.2413 = 0.0134$$

$$= 0.0134 \times 0.05$$

$$= 0.00067$$

$$\text{New } w_1 = 0.15 - (.0001 \times .00067)$$

$$=$$

W2

$$\frac{\partial Loss}{\partial w_2} = \frac{\partial Loss}{\partial y^{hat}} * \frac{\partial y^{hat}}{\partial z_{out}} * \frac{\partial z_{out}}{\partial ah_1} * \frac{\partial ah_1}{\partial zh_1} * \frac{\partial zh_1}{\partial w_2}$$

$$\frac{\partial zh_1}{\partial w_2} = X2 = 0.10$$

$$\frac{\partial loss}{\partial w_2} = 0.7414 \times 0.1868 = 0.1385$$

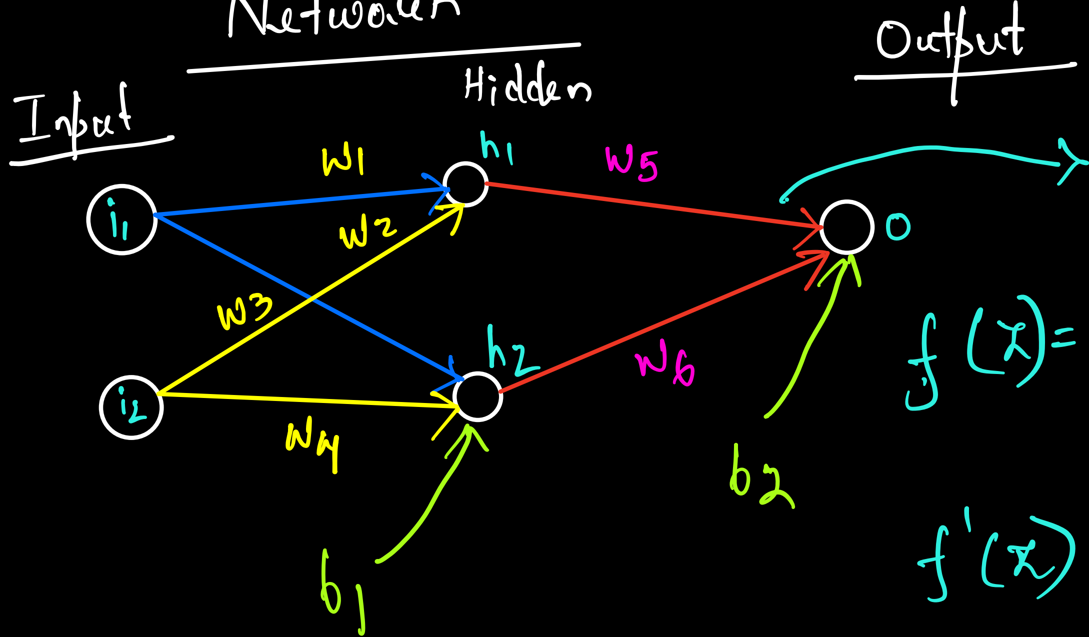
$$0.1385 \times 0.40 = 0.0554$$

$$0.554 \times 0.2413 = 0.0134$$

$$0.0134 \times 0.10 \approx 0.00135$$

$$\frac{\partial loss}{\partial w_2} = 0.00135$$

Network



Sigmoid

$$f(x) = \frac{1}{1 + e^{-x}}$$

$$f'(x) = f(x) * (1 - f(x))$$

w_4

$$\frac{\partial loss}{\partial w_4} = \frac{\partial loss}{\partial y^{hat}} * \frac{\partial y^{hat}}{\partial z_{out}} * \frac{\partial z_{out}}{\partial a\_h_2} * \frac{\partial a\_h_2}{\partial z h_2} * \frac{\partial z\_h_2}{w_4}$$

$$\frac{\partial a\_h_2}{\partial z h_2} = a\_h_2 \left(1 - a\_h_2\right)$$

$$= 0.5969 \times (1 - 0.5969)$$

$$= 0.2407$$

$$\frac{\partial z_{out}}{\partial a\_h_2} = W_6 = 0.45$$

$$\frac{\partial z h_2}{\partial w_4} = x_2 = 0.10$$

$$\frac{\partial Loss}{\partial w_4} = .7414 \times .1868 \times 0.45 \times .2407 \times 0.10$$

$$= .00150$$

Gradient for W_3

$$\frac{\partial loss}{\partial w_3} = \frac{\partial loss}{\partial y^{hat}} \times \frac{\partial y^{hat}}{\partial z_{out}} \times \frac{\partial z_{out}}{\partial a\_h_2} \times \frac{\partial a\_h_2}{\partial z h_2} \times \frac{\partial z h_2}{\partial w_3}$$

$$\frac{\partial z h_2}{\partial w_3} = x_1 = 0.05$$

$$\frac{\partial loss}{\partial w_3} = .7414 \times .1868 \times 0.45 \times 0.2407 \times 0.05$$

$$= .00075$$

All gradients

$W_1 = 0.00067$

$W_2 = 0.00135$

$W_3 = 0.00075$

$W_4 = .00150$

$W_5 = 0.0822$

$W_6 = 0.0827$

Bias Gradients

$$\frac{\partial loss}{\partial b_2} = \frac{\partial loss}{\partial y^{hat}} * \frac{\partial y^{hat}}{\partial Z_{out}} * \frac{\partial Z_{out}}{\partial b_2}$$

$$\frac{\partial Z_{out}}{\partial b_2} = 1$$

$$\frac{\partial loss}{\partial b_2} = .7414 \times .1868 \times 1$$

$$= 0.1385$$

B_1 gradient

$$\frac{\partial Loss}{\partial b_1}$$

$$z\_h_1 = w_1 \times x_1 + w_2 \times x_2 + b_1$$

$$z\_h_2 = w_3 \times x_1 + w_4 \times x_2 + b_1$$

$$\frac{\partial loss}{\partial b_1} = \frac{\partial loss}{\partial b_1 \text{ from } h_1} + \frac{\partial loss}{\partial b_1 \text{ from } h_2}$$

$$\frac{\partial loss}{\partial b_1 \text{ from } h_1} = \frac{\partial loss}{\partial y^{hat}} * \frac{\partial y^{hat}}{\partial Zout} \times \frac{\partial Zout}{\partial a.h_1} \times \frac{\partial a.h_1}{\partial z h_1} \times \frac{\partial z h_1}{\partial b_1}$$

$$\frac{\partial z h_1}{\partial b_1} = 1$$

$$\frac{\partial\, loss}{\partial b_1 \text{ from } h_1} = .7414 \times .1868 \times 0.40 \times .2413 \times 1$$

$$= 0.0134$$

$$\frac{\partial\, Loss}{\partial b_1 \text{ from } h_2} = \frac{\partial\, loss}{\partial y^{hat}} * \frac{\partial y^{hat}}{\partial Z_{out}} \times \frac{\partial Z_{out}}{\partial a_{h_2}} * \frac{\partial a_{h_2}}{\partial z_{h_2}} \times \frac{\partial z_{h_2}}{\partial b_1}$$

$$= .7414 \times .1868 \times .45 \times .2417 \times 1$$

$$\frac{\partial\, loss}{\partial b_1 \text{ from } h_2} = 0.0150$$

$$\frac{\partial Loss}{\partial b_1} = 0.0134 + 0.0150$$

$$= 0.0284$$

LR = .0001

LR = 1