Neural Network with computational Graph

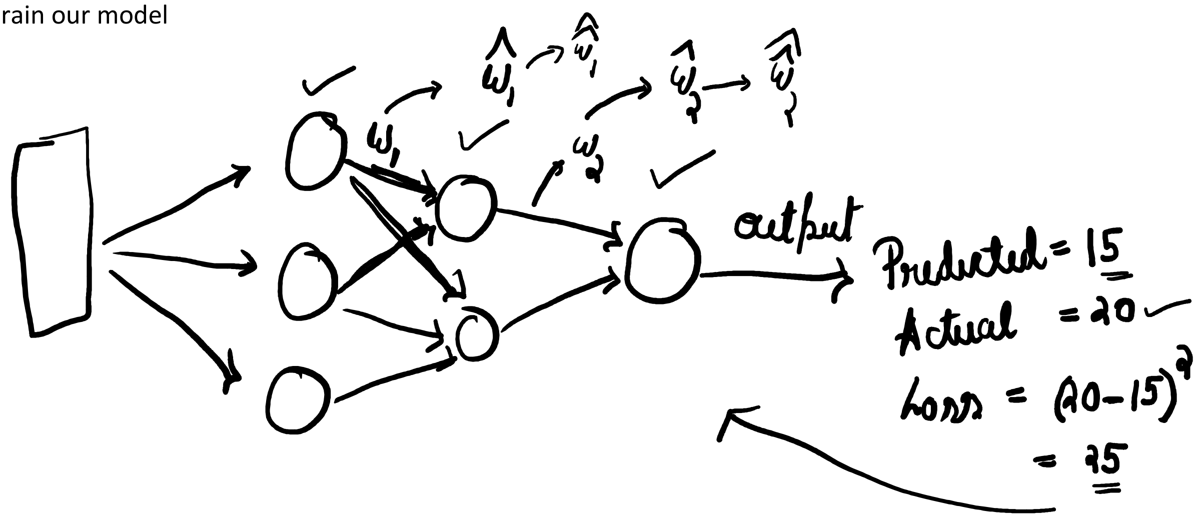
Forward Propagation

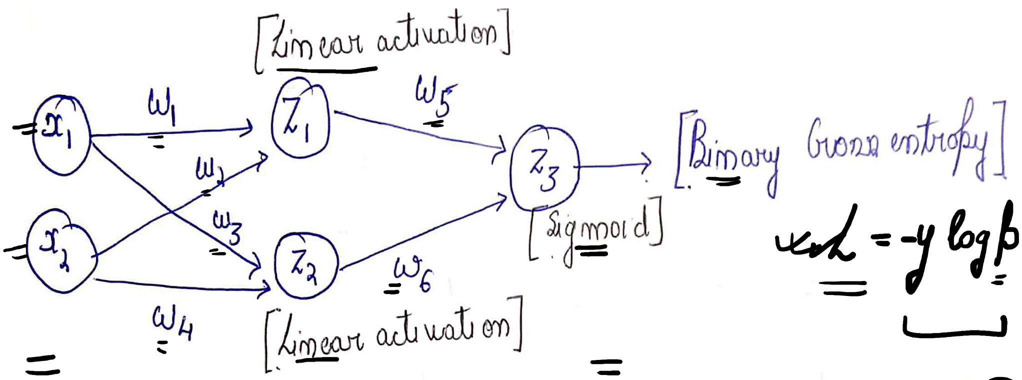
1. We use the input data given with the weights at any given point of time to get predictions from the model.
2. We propagate layer by layer and towards the end we get the predictions

After forward propagation step, we have a Loss function.

Backward Propagation

We use the feedback from the loss function to adjust the weights so that we get lesser and lesser loss every time we train our model





$$L = -y \log p - (1-y) \log(1-p)$$

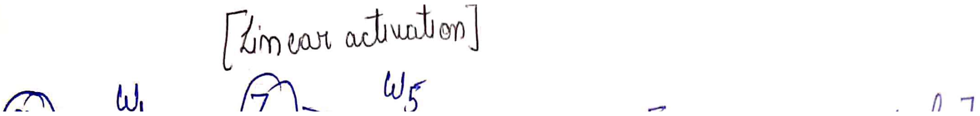
$$\frac{\partial L}{\partial w_1}$$

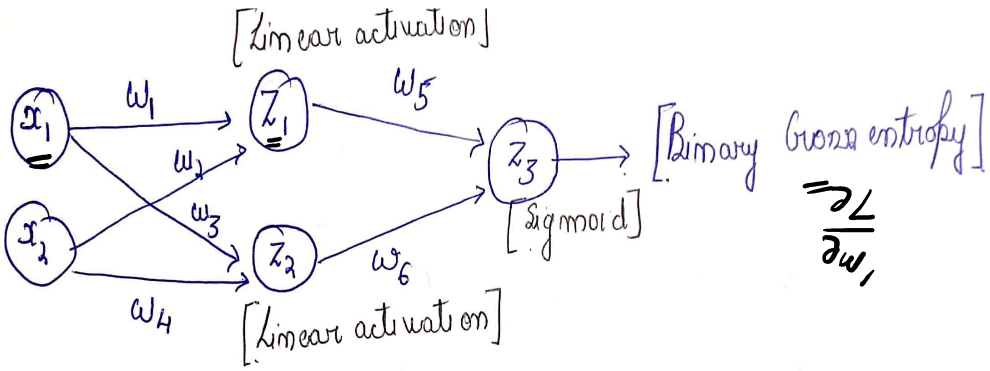
$$z_1 = x_1 w_1 + x_2 w_2$$

$$z_2 = x_1 w_3 + x_2 w_4$$

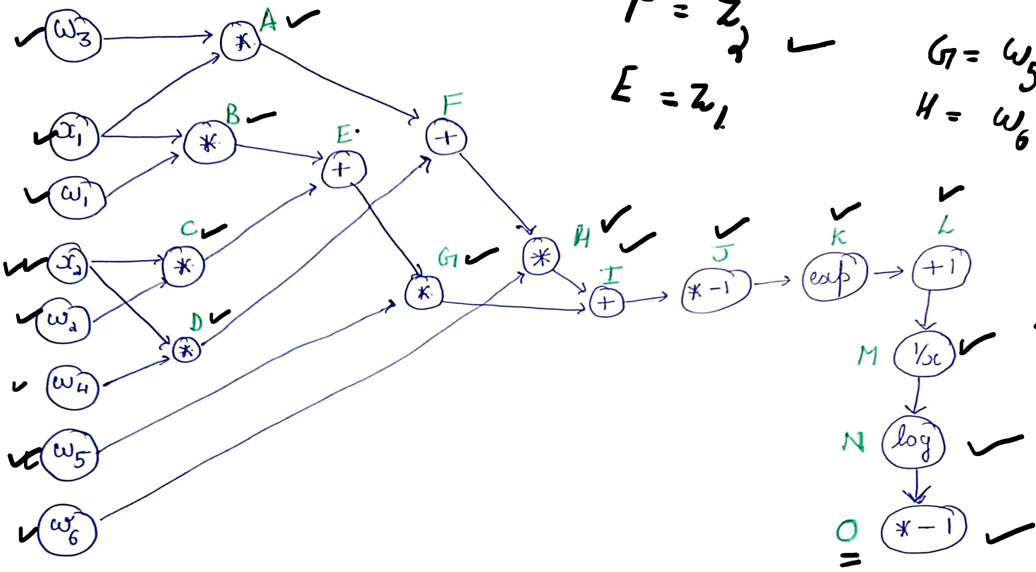
$$z_3 = \frac{1}{1 + \exp[-(w_5 z_1 + w_6 z_2)]} = \text{prob.} (0-1)$$

$$\text{Loss function} = L(w, x) = -y \log_e z_3 - (1-y) \log_e(1-z_3)$$





Computational Graph drawn below



$F = z_2$

$E = z_1$

$G = w_5 z_1$

$H = w_6 z_2$

$I = w_5 z_1 + w_6 z_2$

$z_3 = \left[\frac{1}{1+e^{-I}}\right]$

$-\log z_3$

While performing back propagation we need to find $\nabla_w L(w,x)$

$\frac{\partial L(w,x)}{\partial w_1}$, $\frac{\partial L(w,x)}{\partial w_2}$, $\frac{\partial L(w,x)}{\partial w_3}$, $\frac{\partial L(w,x)}{\partial w_4}$, $\frac{\partial L(w,x)}{\partial w_5}$

$\frac{\partial L(w,x)}{\partial w_6}$.

Let us suppose $x_1 = 1$

$x_2 = 2$

$w_1 = 1$

$w_2 = 1$ $w_5 = 1$

$w_3 = 2$ $w_6 = 2$

$w_4 = 1$

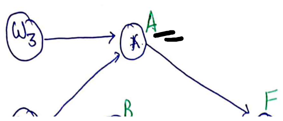
$y = 1$

$y = 1$ $z_3 \to 1$

$L = -1 \log p - (1-1)\log(1-p)$

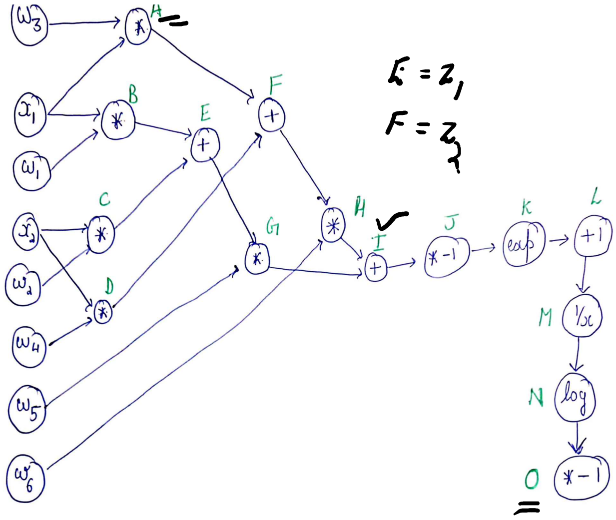
$L = -\log z_3$

Values	Actual Values	Derived values
A	2	$x_1 w_3$



$E = z_1$

Values	Actual Values	Derived values
A	2	$x_1 w_3$
B	1	$x_1 w_1$
C	2	$x_2 w_2$
D	2	$x_2 w_4$
E	$x_1 w_1 + x_2 w_2 = 3$	$B + C$
F	$x_1 w_3 + x_2 w_4 = 4$	$A + D$
G	$w_5 (x_1 w_1 + x_2 w_2) = 3$	$E * w_5$
H	$w_6 (x_1 w_3 + x_2 w_4) = 8$	$F * w_6$
I	11	$G + H$
J	-11	$-I$
K	$\exp(-11)$	$\exp(J)$
L	$1 + \exp(-11)$	$1 + K$
M	$\frac{1}{1+\exp(-11)}$	$1/L$
N	$\log_e\left[\frac{1}{1+\exp(-11)}\right]$	$\log M$
O	$-\log\left[\frac{1}{1+\exp(-11)}\right]$	$-N$

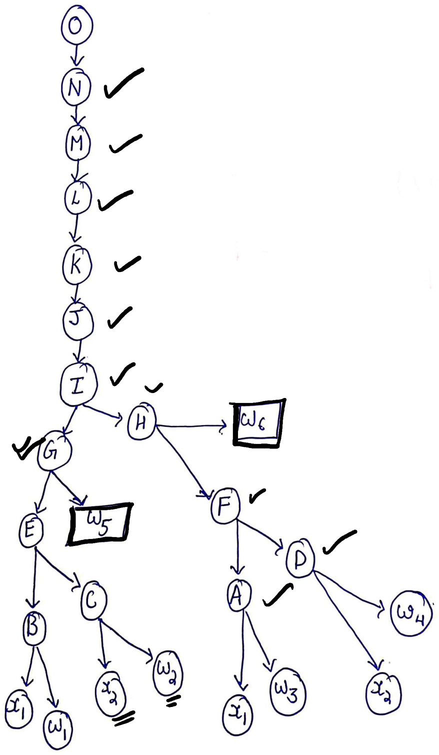


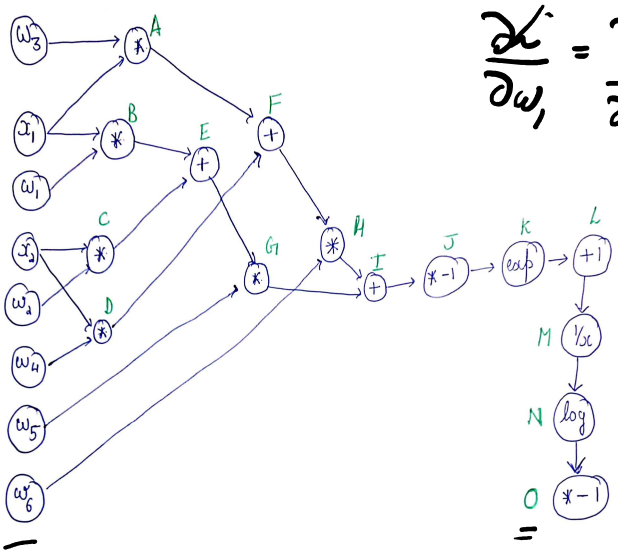
$E = 2,$

$F = 2$

x1=1, x2= 2,w1=1 , w2=1 , w3 =2 ,w4= 1, w5=1 and w6 = 2

y = 1





$$\frac{\partial L}{\partial w_1} = \frac{\partial O}{\partial w_1}$$

$$\frac{\partial O}{\partial w_1} = \frac{\partial O}{\partial N}\frac{\partial N}{\partial M}\frac{\partial M}{\partial L}\frac{\partial L}{\partial K}\frac{\partial K}{\partial J}\frac{\partial J}{\partial I}\frac{\partial I}{\partial G}\frac{\partial G}{\partial E}\frac{\partial E}{\partial B}\frac{\partial B}{\partial w_1}$$

$$\frac{\partial O}{\partial w_1} = \frac{\partial O}{\partial N}\frac{\partial N}{\partial M}\frac{\partial M}{\partial L}\frac{\partial L}{\partial K}\frac{\partial K}{\partial J}\frac{\partial J}{\partial I}\frac{\partial I}{\partial G}\frac{\partial G}{\partial E}\frac{\partial E}{\partial B}\frac{\partial B}{\partial w_1}$$

$$= (-1)\frac{1}{M}\left(\frac{-1}{L^2}\right)(1)\, e^J (-1)\,(1)\, w_5\,(1)\, x_1$$

$$= -\frac{1}{(1/L)}\left(\frac{1}{L^2}\right) e^J w_5 x_1$$

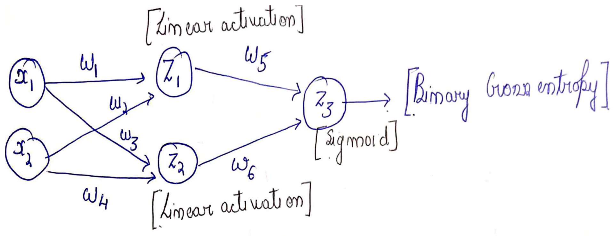
$$= -\frac{e^J w_5 x_1}{L} = \left[\frac{-e^{-11}}{1+e^{-11}}\right](1)(1)$$

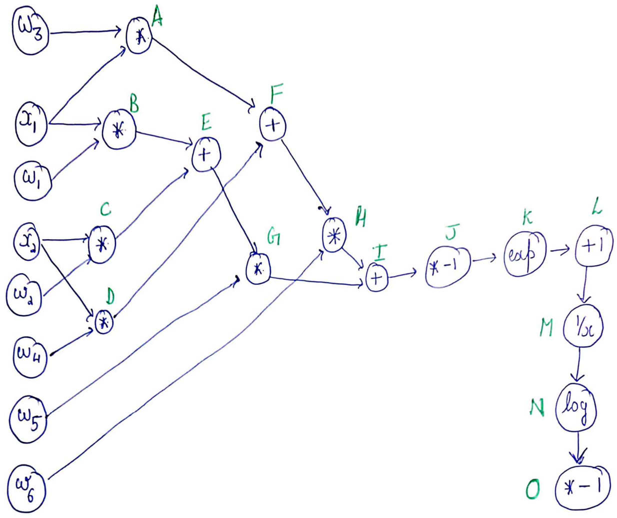
$$= \frac{-e^{-11}}{1+e^{-11}} = \frac{\partial L}{\partial w_1}$$

$$w_1^{(\text{updated})} = w_1 - \alpha\frac{\partial O}{\partial w_1}$$

$$= 1 - 0.05\left(\frac{-e^{-11}}{1+e^{-11}}\right)$$

$$= 1$$


[Linear activation]
w_1
w_5
w_2
w_3
w_6
w_4
[Binary Cross entropy]
[Sigmoid]
[Linear activation]


A
B
C
D
E
F
G
H
I
J
K
L
M
N
O