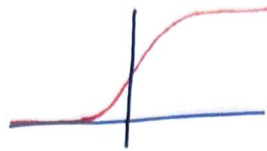


\*mLP

# Softmax Layer, Cross Entropy

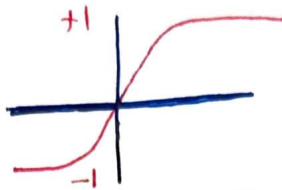
Sigmoid

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



tanh

$\tanh(x)$

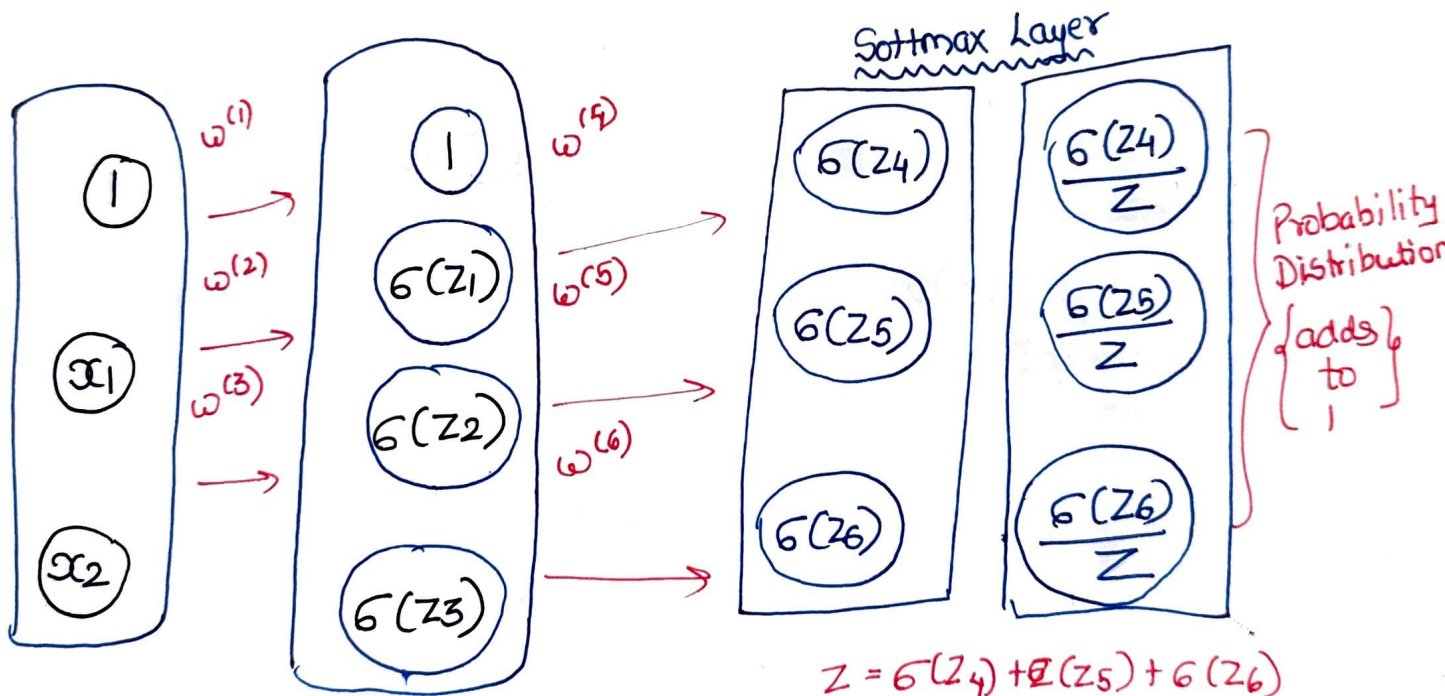


ReLU

$\max(0, x)$



Rectified Linear Unit  
(Non Linear Activation Function)



Softmax



find sum and divide by addition  
so it always adds to 1

Cross Entropy :  $\begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}$   $\rightarrow$  ground truth probability

The Softmax output

Desired output ?

class	
class 1	1
class 2	0
class 3	0

$\begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix}$  one hot  $\rightarrow$  predicted probability

$$\text{Cross entropy}(p, q) = - \sum_{i=1}^3 q_i \log p_i \quad \left. \vphantom{\sum_{i=1}^3} \right\} \text{non symmetric}$$

Its called Cross Entropy because



$$ce(p, p) = \text{entropy}(p)$$

$x \xrightarrow{m(x)} p(x), q(x)$   $\rightarrow c.e(x) = c.e(p(x), q(x))$   
Individual error/loss for the point

~~CEL(x, y)~~

$$* \text{CEL}(X, y) = \sum_{j=1}^n c.e(x_j)$$

Cross entropy loss

$\hookrightarrow$  Sum of all individual cross entropy losses

Cross Entropy  $\rightarrow$  difference in ground truth and predicted probability