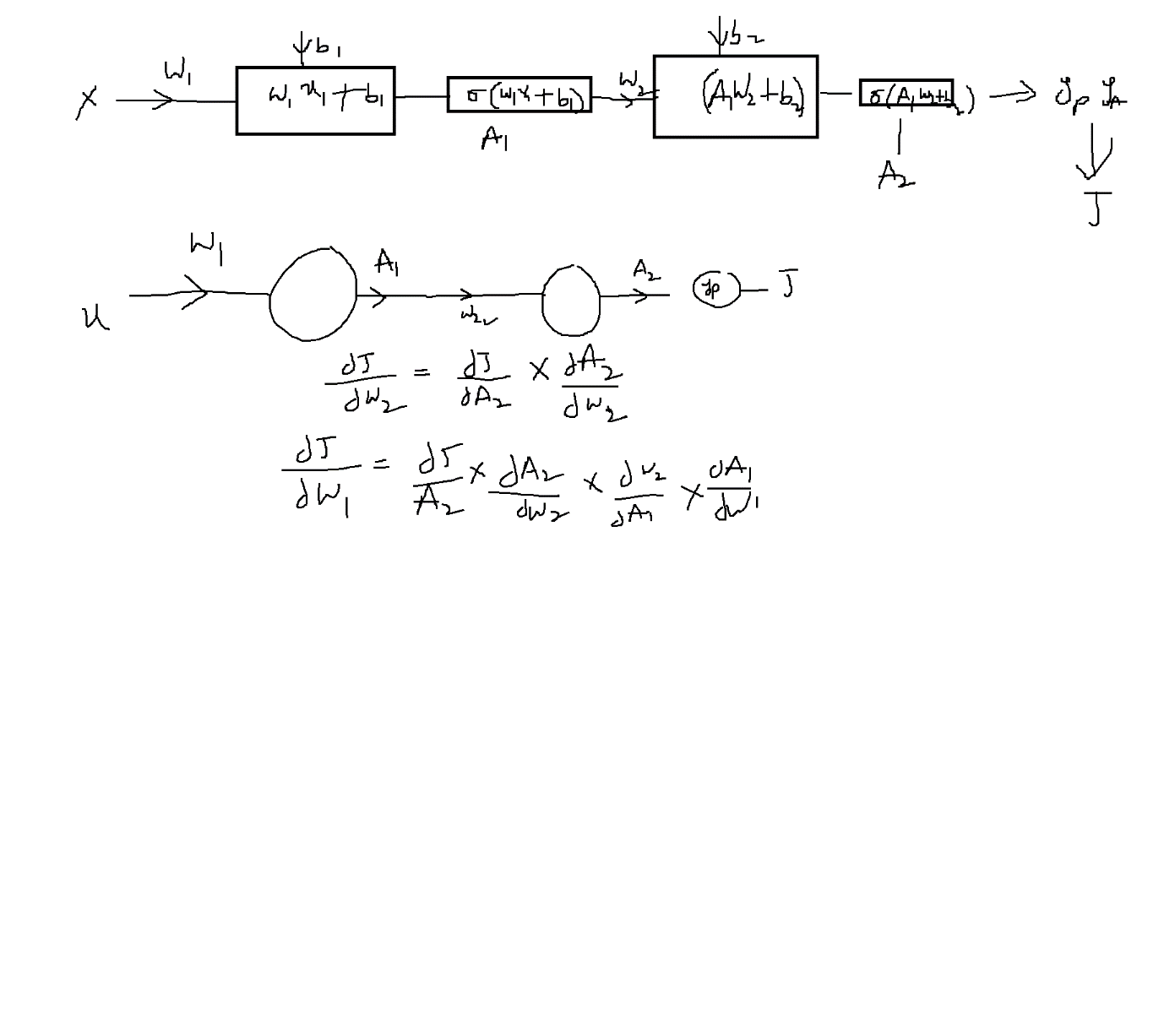
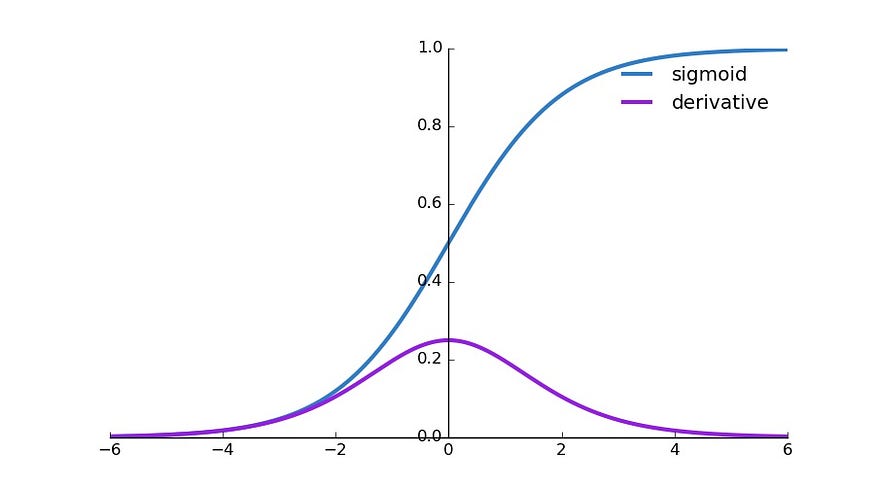
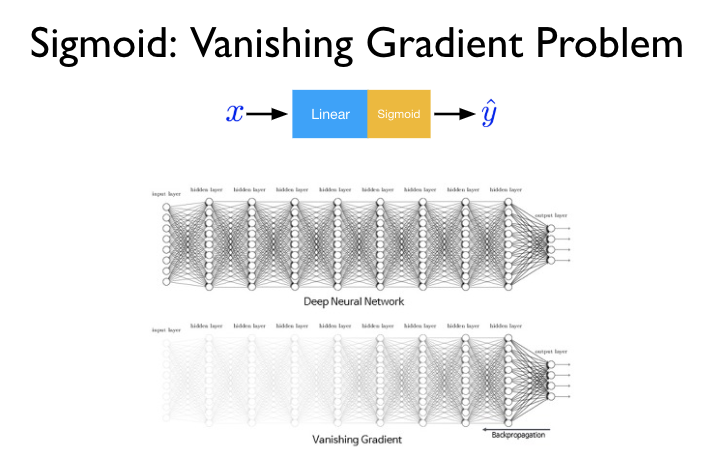
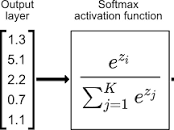
1. How to choose weights : Gradient Descent Back propagation
2. Activation Functions



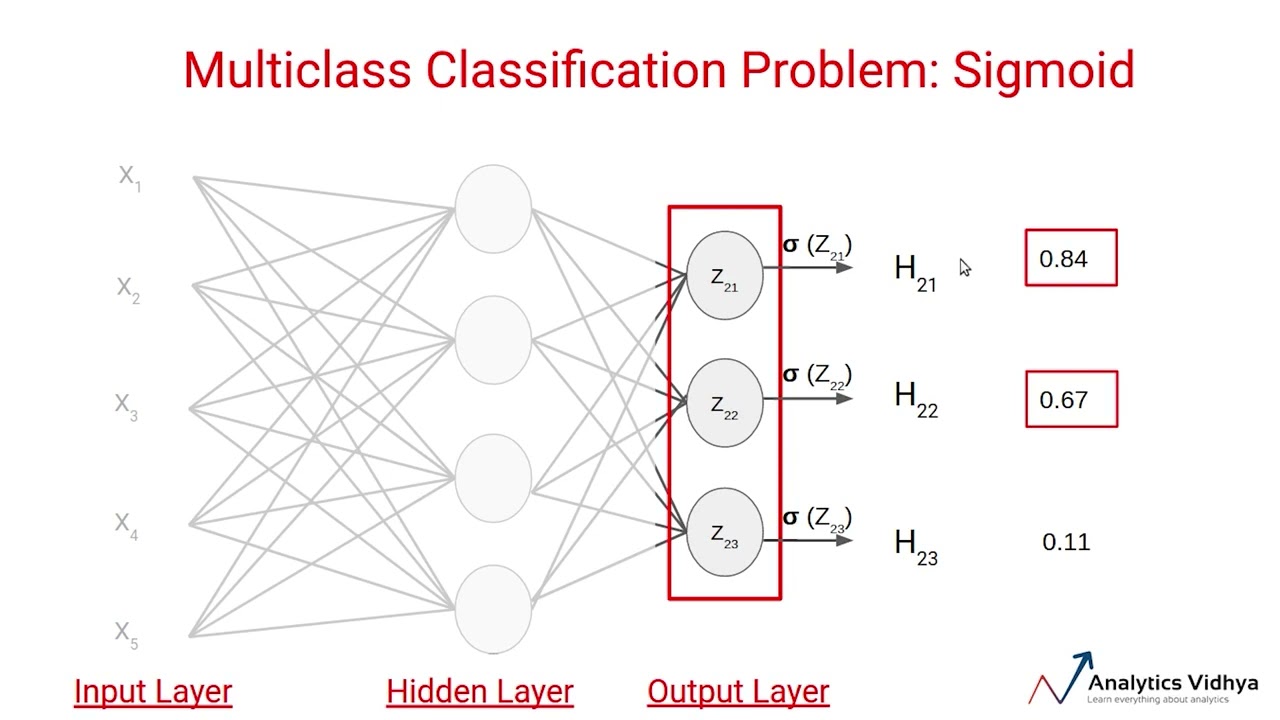
****

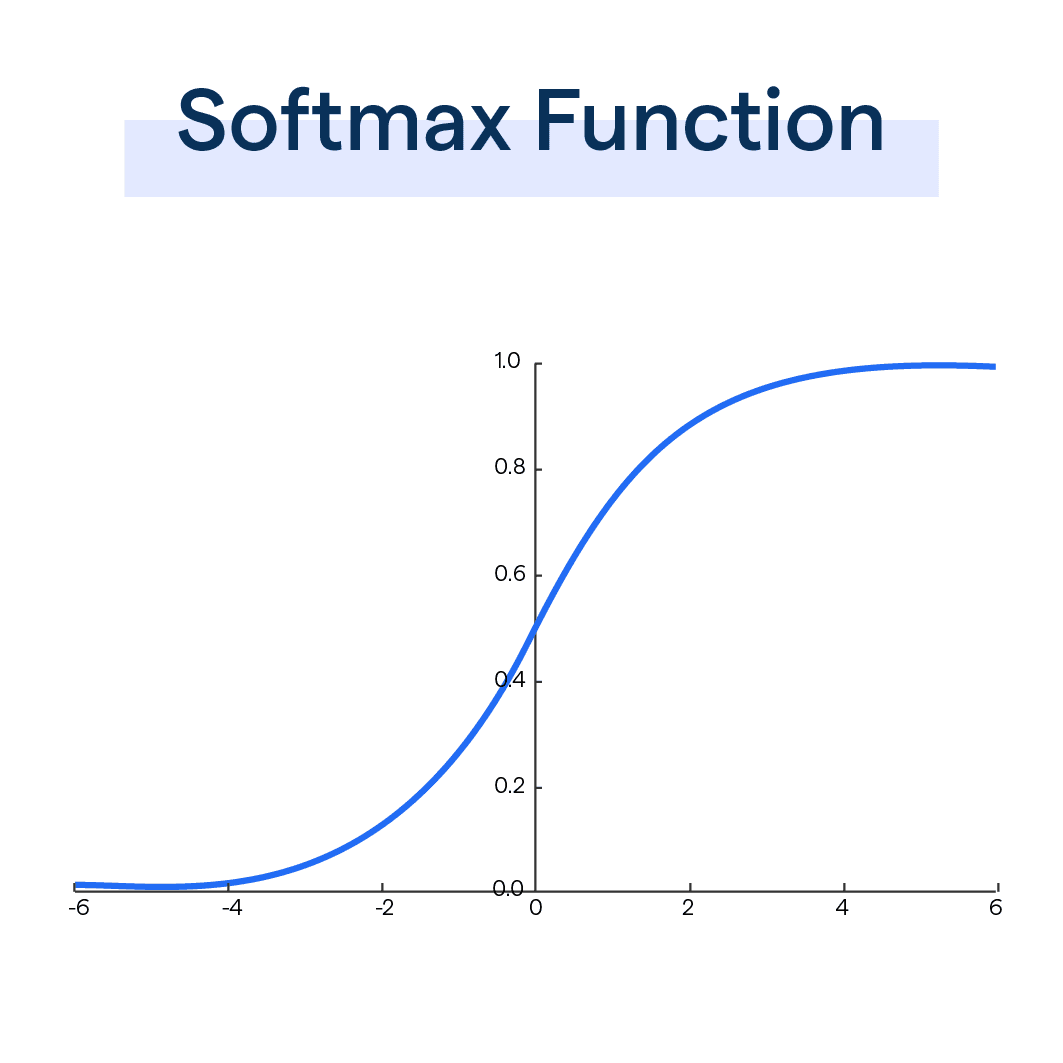






*Getting 1 out of 6= 1/6*



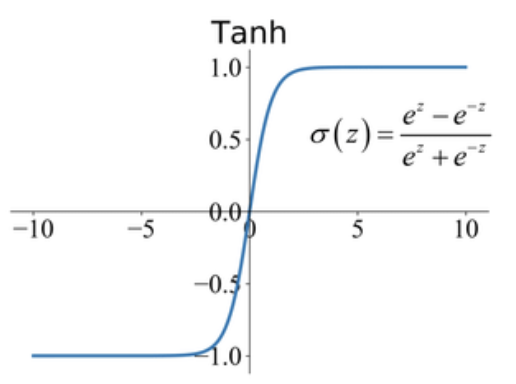


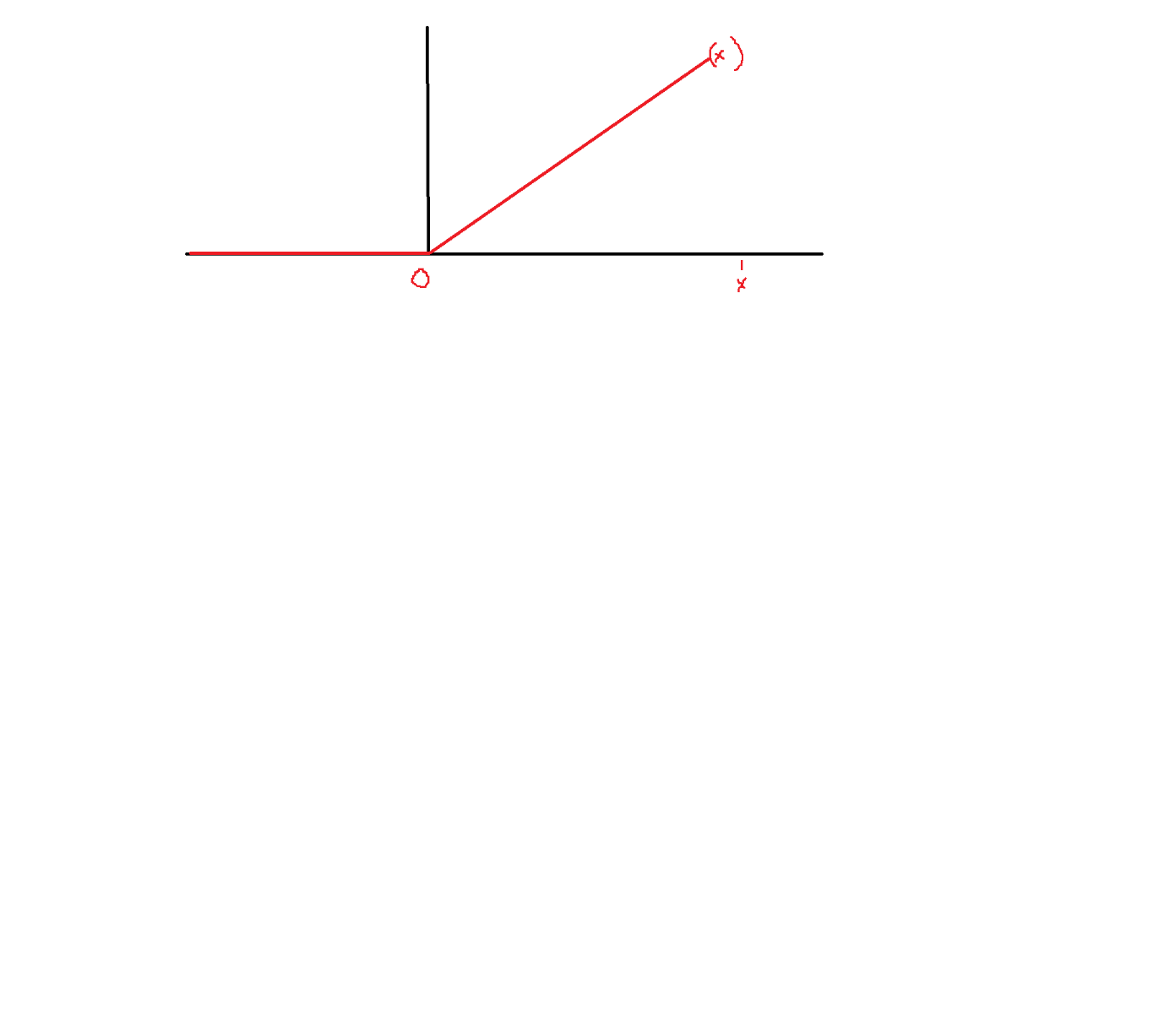
Softmax used for Multiclassification at output layer

Why not at Hidden layer:

Vanishing gradients problem

1. Equation
2. Which problem
3. Vanishing gradients happens in both cases
4. Graph also look like same

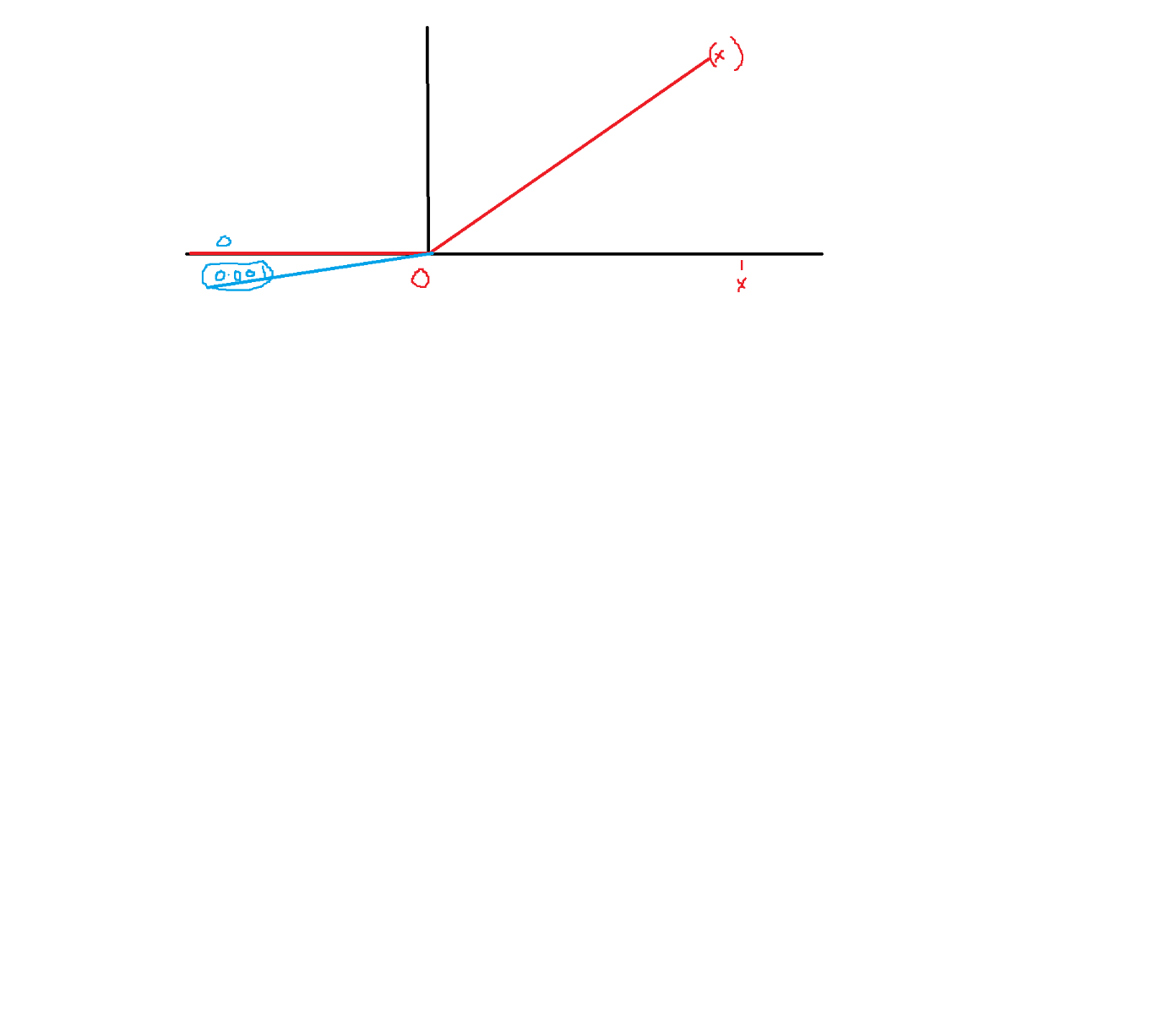


Activations is used to provide Non linearity in the Neurons

ReLU looks like Linear curve only

Linear curve will never

to calculate gradients, is reLU differentiable at 0?



EWA : Forecasting

Over optimisa

Wed === Python implement

CNN