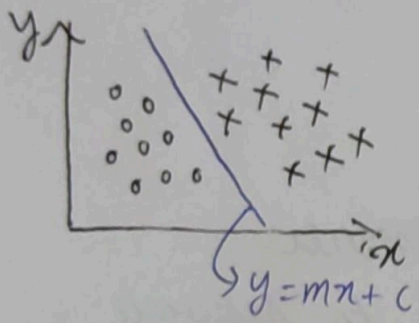
Activation Function

WHAT? ⇒ in NN, activation func^n of node defines the output of that node given an input or set of input

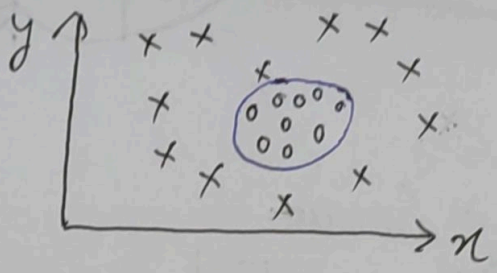
USE ⇒ In every NN during forward propagation.

★ Linealy Seperable



⇒ 2 classes can be Seperated by a Str. line

Linearly Unseperable



⇒ 2 classes can't be Seperated by a curve or a more complex function than a Str. line

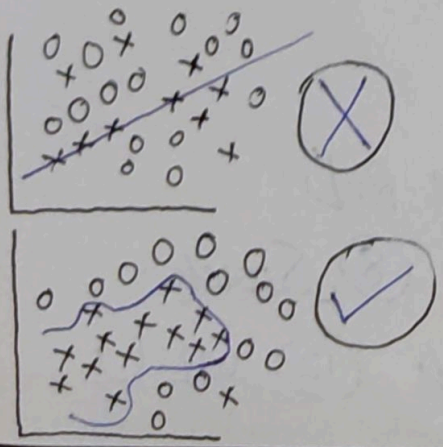
★ Why need activation?

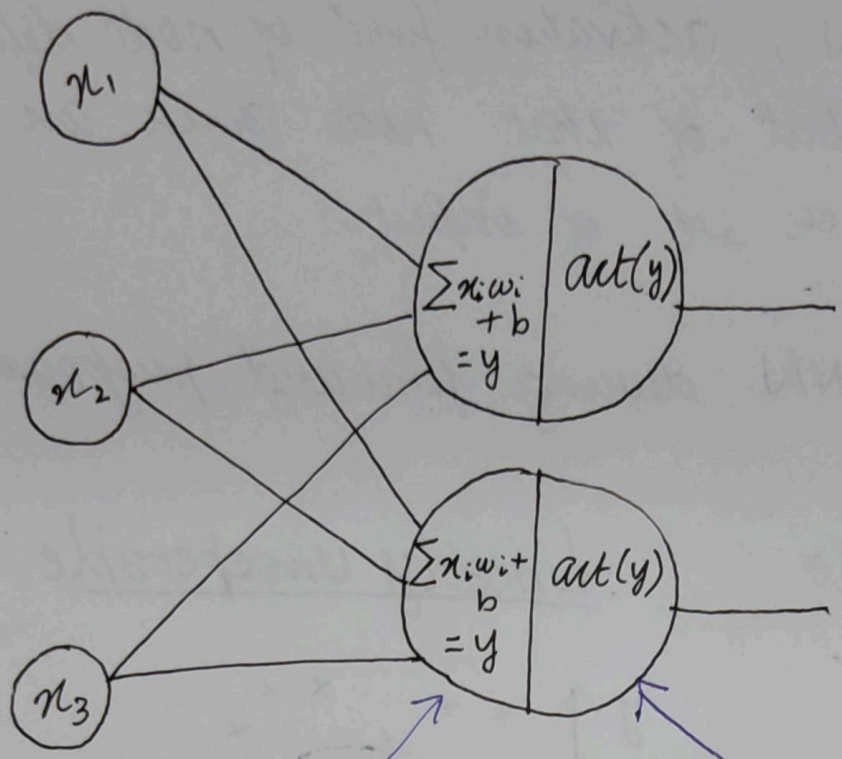
Real world data → non-linear

⇒ can't be seperated by Str. line

⇒ linearly unseperable

⇒ more complex func^{n}s needed to be able to predict / classify the data we are working with.





⇒ linear

⇒ $\sum_{i=1}^{N} [w_i x_i + b_i]$

⇒ $w^T x + b$

⇒ $w_1 x_1 + w_2 x_2 + w_3 x_3 = y$

⇒ Linear Regression (linear)

⇒ not able to seperate the data

act(y)

⇒ non linear

⇒ able to seperate data

eg

$act(y) = Y = \frac{1}{1 + e^{-y}}$

⇒ 

☆ Types of activation function

① Sigmoid ② Tanh ③ ReLU ④ Leaky ReLU

⑤ PReLU ⑥ ELU ⑦ Softmax ⑧ Swish

⑨ Maxout ⑩ Softplus ⑪ Mish ⑫ GLU