Neural Networks & Non-linearity

Recap:

a) MLP } NN
b) CNN }

→ Need for multi-layer neural networks.

⇓

XOR problem.

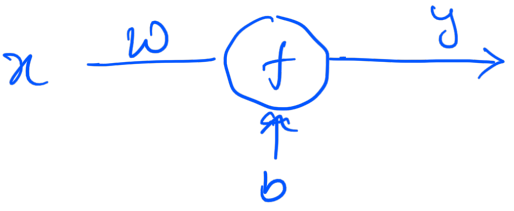
c) Advs of CNN over MLP's.

d) Activation functions.

→ Q) What are these ?

- ReLU → max(0, i/p)
- Sigmoid ∈ [0,1]
- Tanh ∈ [-1,1]
- Softmax ⇒ prob ∈ [0,1]
- P-ReLU
- Leaky ReLU
- ⋮

Perceptron:

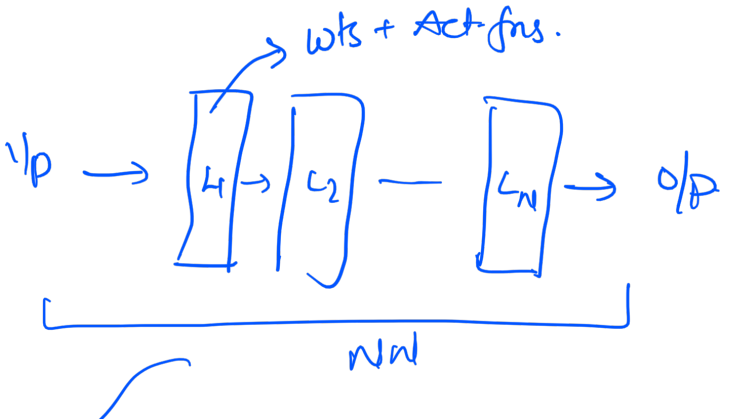


$y = f((xw + b))$

- Act. fn.
  - Un-confined (or) Unbounded
  - Confined or bounded

Q2) Are Act-fn. linear / non-linear ?

↳ functions.



Q) Are NNs Linear / Non-Linear / Maybe / Don't know ?

(a) (b) (c)

0 2 11

Q) What makes NNs non linear?

(2) a) Act. fns alone ?

0 b) Act fn. + something else?

11 c) Don't know.

(A) Linear/ Linearity

Eg: 1) $y = mx$

2) $y = mx + c$ ------> (x)

(B) Linear Systems ⇒ transforms input

Eg: $y = mx$

Q) Is NN a system?

Yes

Q) How do we verify linearity of a system?

2 properties: (P1) additive property, (P2) homogeneity } superposition principle //.

Additive

P_1

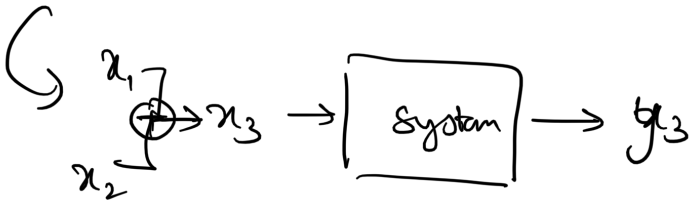
$y = mx$

- $y_1 = m x_1$
- $y_2 = m x_2$

- $(x_1 + x_2) = x_3$.

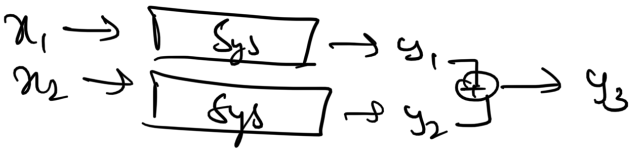
$y_3 = y_1 + y_2 = mx_1 + mx_2$

$mx_3 = m(x_1 + x_2) \equiv m(x_1 + x_2)$



(A)

$\Downarrow\Uparrow$



(B)

Homogeneity

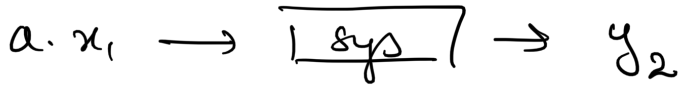
P_2

$y = mx$.

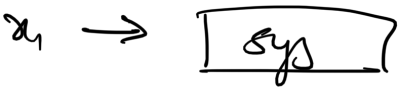
$y_1 = m(ax_1)$

$x_1 \Rightarrow a \cdot x_1 \Rightarrow m(ax_1) = a(mx_1)$

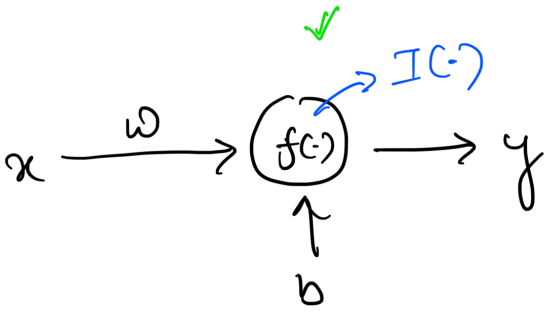
$= ay_1 = y_2$



$\Downarrow\Uparrow$



Basic NN:



$$y = f(wx+b)$$

$$y = wx+b$$

$f \Rightarrow$ Act. fn.

We know that $f \Rightarrow$ Non-linear and we want final if there are any other sources / causes for non-linearity.

$f \equiv$ Identity function.

P_1

$$y_1 = wx_1+b$$

$$y_2 = wx_2+b$$

- $y_3 = y_1+y_2 = wx_1+b+wx_2+b$

$$= w(x_1+x_2)+2b$$

- $x_1+x_2 \Rightarrow \hat{y}_3 = w(x_1+x_2)+b$

$\neq$

P_1 ✗

P_2

$$y_1 = wx_1+b$$

$x_1 \Rightarrow ax_1 \Rightarrow \hat{y}_2 = w(ax_1)+b$

$$\hat{y}_2 = a(wx_1)+b$$

$\neq$

$$y_2 = ay_1 = a(wx_1+b) = a(wx_1)+ab$$

P_2 ✗

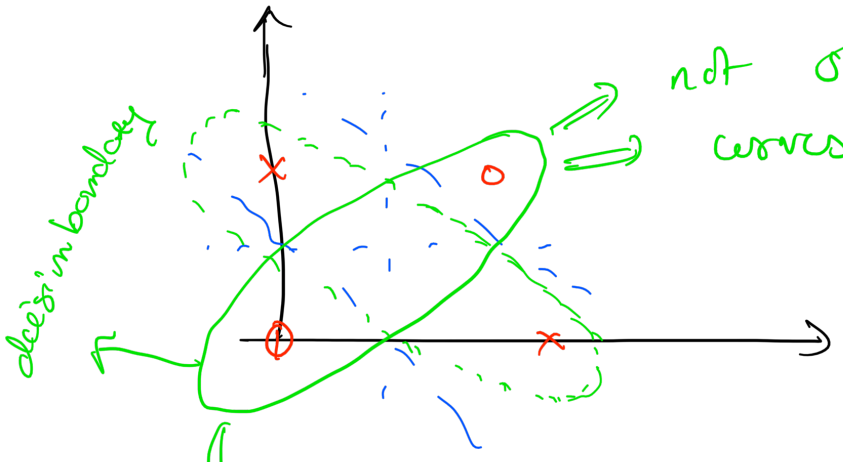
∴ Hence proved a perceptron / NN building block is a NN System.

↳ NN are NN Systems!

Q) What other components introduce NN into NN?

Q1) What is the use of AFs ⇒ $(-\infty,\infty) \Rightarrow \begin{matrix}(0,1)\\(-1,1)\end{matrix}$ } Why? Impact) (I)

Q2) Why do you need multiple layers / is a single layer good enough? (II)



not ∞ lines

curves ⇒ 2nd degree polynomials ✓

↓

non-linear.

What would be the NN for this? (I)

Logic Gates

a) NOT
b) AND
c) OR
d) XOR
e) XNOR