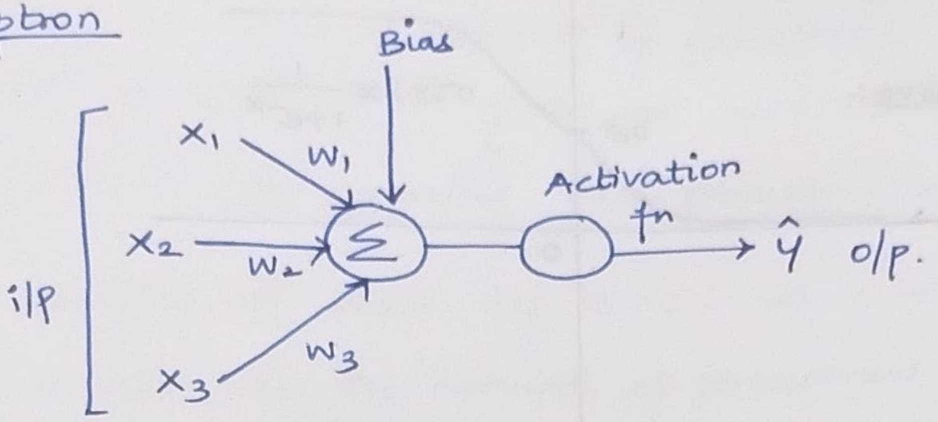
Perceptron



i.e, $\hat{y}$ = Activation_fn $\left(\sum_{i=1}^{n} w_i x_i + b\right)$

so, it gives weights to each of its input values and computes the o/p based on its sum and threshold value.

in case of sigmoid neuron, the activation function is a sigmoid function:-

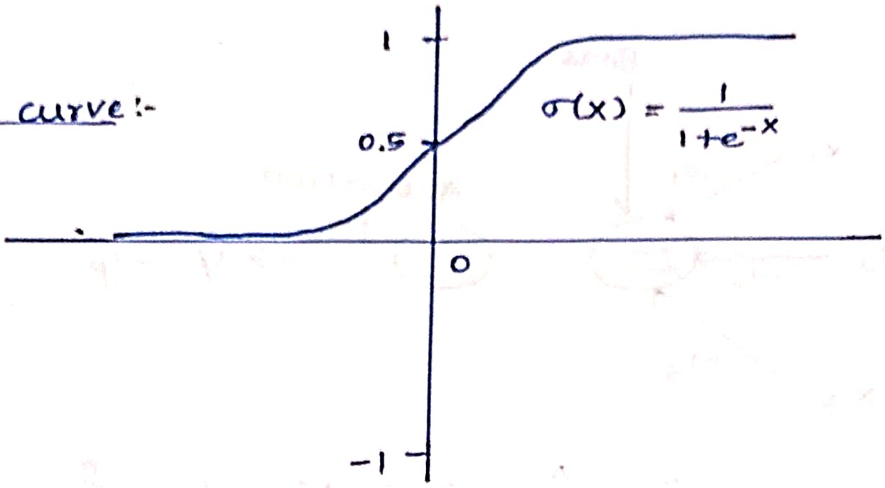
ie, $\sigma(z) = \dfrac{1}{1+e^{-z}}$

in case of perceptron, it is a step function

ie, $H(z) = \begin{cases} 0 & \text{if } z < a \\ 1 & \text{if } z \geq a \end{cases}$

so, can I say sigmoid neuron as a perceptron with sigmoid fn as its activation fn? $\boxed{\text{YES}}$

sigmoid curve:-



thus, $\sigma(x)\ \mathbb{R} \to [0,1]$

$$f = wx + b$$

$$\frac{\partial f}{\partial w} = x \qquad \frac{\partial f}{\partial b} = 1$$

$$\Delta f \approx \frac{\partial f}{\partial w}\Delta w + \frac{\partial f}{\partial b}\Delta b$$

$$\Rightarrow \Delta f \approx x\,\Delta w + \Delta b$$

Ex : 1

suppose we take all the w & b in a network of perceptrons, and multiply them by a +ve constant, show that the behaviour of the network doesn't change.

Given is a network of perceptron

so, $\boxed{f(x) = g\left(\sum w_i x_i + b\right)}$

here, $g(x)$ is a step function, ie, $g(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$

so, when we multiply the result by a +ve constant, it does not change the result. Because, it is based only on the sign (not magnitude).

part II

→ Now, we have a network of perceptrons.

→ ∀ perceptron, it is given as $w.x + b \neq 0$

→ Now, we are replacing each perceptron with a sigmoid neuron.

→ multiply all w & b by $c > 0$. as $c \to \infty$, the network behaviour is same as the network of perceptrons. (show)

$$z_c = c(w.x + b)$$

$$\sigma(z_c) = \frac{1}{1 + e^{-c(w.c + b)}}$$

taking $\lim_{c \to \infty}$

Then, case 1 $\quad w.c + b > 0 \Rightarrow -c(w.c + b) \to +\infty$

$$\Rightarrow \sigma(z_c) = 1$$

same as perceptron o/p.

case 2 $\quad w.c + b < 0 \Rightarrow -c(w.c + b) \to -\infty$

$$\Rightarrow \sigma(z_c) = 0$$

same as perceptron o/p.

But when $w.c + b = 0$,

$$\sigma(z_c) = \frac{1}{1+1} = \underline{0.5}$$

Not the behavior of perceptron.

Exercise #2

Find the set of weights & biases for the 4th layer

Assumption

First 3 layers are correct with 3rd layer has an activation of ≥ 0.99 for correct o/p neuron & < 0.01 activation for incorrect neuron.

i/p layer is $a = (a_0, a_1, a_2, \dots, a_9)$.

o/p layer is $O = (O_0, O_1, O_2, O_3)$.

Thus, we can have $W^{[1]} = [0,0,0,0,0,0,0,0,1,1]$

$W^{[2]} = [0,0,0,0,1,1,1,1,0,0]$

$W^{[3]} = [0,0,1,1,0,0,1,1,0,0]$

Binary representation $W^{[4]} = [0,1,0,1,0,1,0,1,0,1]$

Since each neuron in 3rd layer has correct o/p, we can directly use these weights ($W^{[1]} \dots W^{[4]}$).

Bias

Since for correct digits, o/p will be ≥ 0.99.

Thus, $f = \underset{(1)}{w}\,\underset{(0.99)}{x} + b > 0$

i.e, $\boxed{b > -0.99}$

for wrong digits, o/p will be < 0.01

Thus, $f = wx + b < 0$

(ie) $\boxed{b < -0.01}$

So, bias can be in anywhere in range $(-0.99, -0.01)$.