(1) o/p of the last fc layer

$$\begin{bmatrix} 2.0 \\ 1.0 \\ 0.1 \end{bmatrix} \xrightarrow{\text{apply softmax}}$$

softmax formulae;

$$S(x) = \frac{e^{x_i}}{\sum_{i=1}^{9} e^{x_i}} = \frac{7.39}{7.39 + 2.72 + 1.11}$$

$$= \frac{7.39}{11.22}$$

$$\boxed{P(\hat{y} = 0) = 0.65}$$

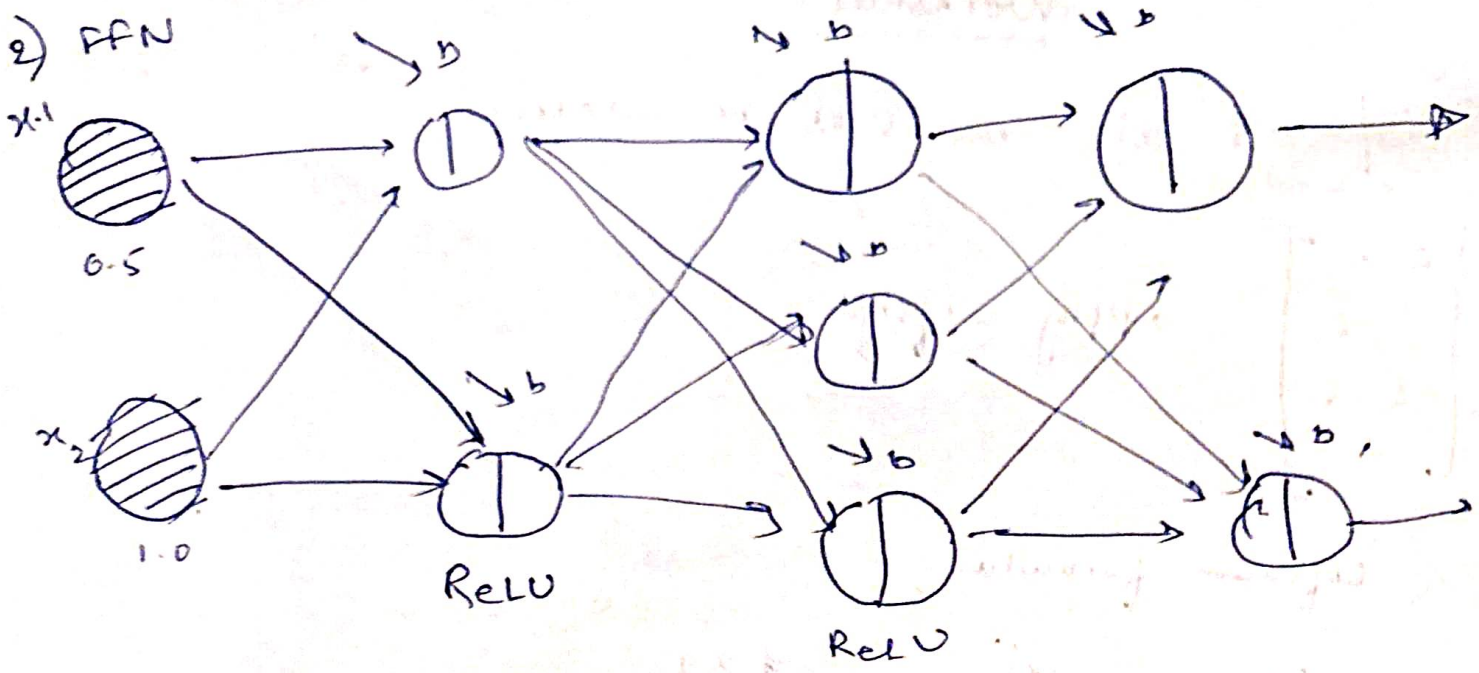
$$P(\hat{y} = 1) = \frac{2.72}{11.22}$$

$$(P\hat{y} = 1) \Rightarrow 0.24$$

$$(P\hat{y} = 2) = \frac{1.11}{11.22}$$

$$P(\hat{y} = 2) = 0.91$$

2) FFN



$$W^{[1]} = \begin{bmatrix} w_{11}^{[1]} & w_{12}^{[1]} \\ w_{21}^{[1]} & w_{22}^{[1]} \end{bmatrix} = \begin{bmatrix} 0.742 & 0.794 \\ 0.561 & 1.0 \end{bmatrix}$$

$$W^{[2]} = \begin{bmatrix} w_{11}^{[2]} & w_{21}^{[2]} \\ w_{21}^{[2]} & w_{22}^{[2]} \\ w_{31}^{[2]} & w_{32}^{[2]} \end{bmatrix} = \begin{bmatrix} 0.725 & 0.613 \\ 0.416 & 0.092 \\ 0.876 & 0.590 \end{bmatrix}$$

$$W^{[3]} = \begin{bmatrix} w_{11}^{[3]} & w_{12}^{[3]} & w_{13}^{[3]} \\ w_{21}^{[3]} & w_{22}^{[3]} & w_{23}^{[3]} \end{bmatrix}$$

$$= \begin{bmatrix} 0.383 & 0.107 & 0.233 \\ 0.112 & 0.132 & 0.536 \end{bmatrix}$$

$$z^{[1]} = W^{[1]} x \rightarrow \begin{bmatrix} 0.742 & 0.794 \\ 0.561 & 1.0 \end{bmatrix} \begin{bmatrix} 0.5 \\ -1.0 \end{bmatrix}$$

$$z^{[1]} = \begin{bmatrix} -0.423 \\ -0.7195 \end{bmatrix}.$$

$$A^{[1]} = \text{Relu}(z^{(1)}) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

$$z^{[2]} = w^{[2]}[1]$$

$$z^{[2]} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

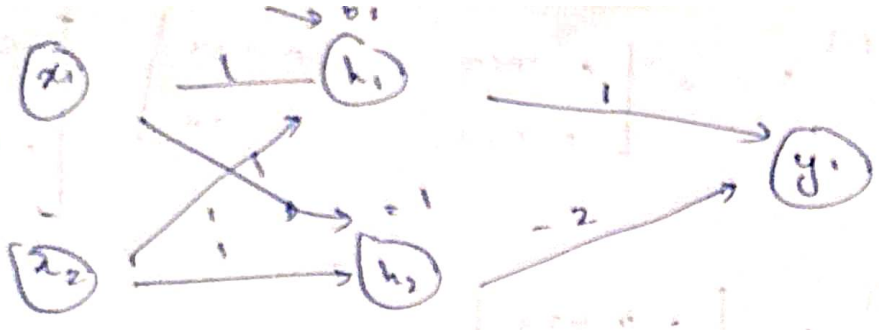
$$A^{[2]} = \text{Relu}(z^{[2]}).$$

$$z^{[3]} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

$$A^{[3]} = \text{Softmax}(z^{[3]}).$$

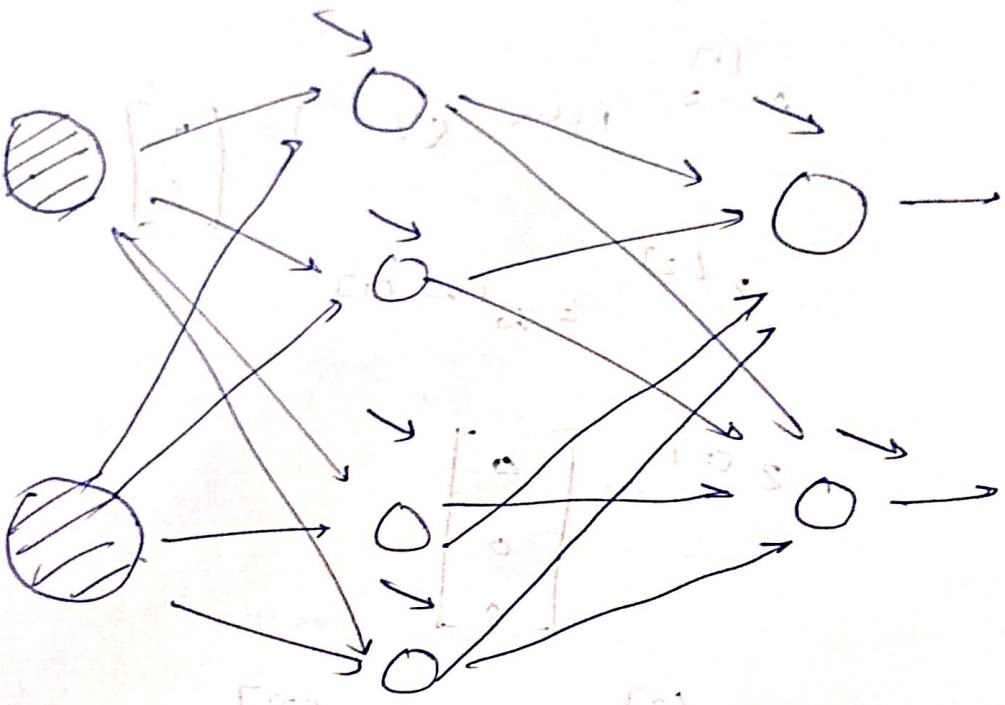
$$A^{[3]} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$$

(3)



total loss = 0

(4)

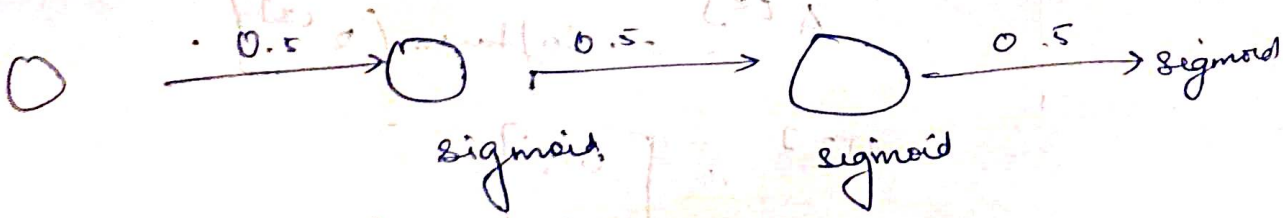


Total param.

$3 + 4 + 4 + 4 \times 2 + 2$

$= 26$ params used

(5)



$z_1 = 0.25$ | $z^{[2]} = 0.28$ | $z^{[3]} = 0.285$

$a^{[1]} = 0.562$ | $a^{[2]} = 0.569$ | $a^{[3]} = 0.570$

① $\frac{dy}{dz} = 1$

② $\frac{dy}{dz^{[3]}} = \frac{dy}{dy} \frac{dy}{dz^3}$

3) $1 \times z^{(3)} [1 - z^{[3]}$

$= 0.204$

③ 0.111

iv) 0.0205

v) 0.1152

vi) 0.000192402 .

vii) 0.0096 .

it is eventually decreasing.

5) $X = [2, 4, 6, 8, 10]$.

$\alpha = 2$

$\beta = 1$.

x	$(x - \bar{x})^\alpha$	$\frac{x - \bar{x}}{\sqrt{var(x)}}$	$x_{new} = x_{norm}$
2	16	-1.265	-1.53
4		-0.633	-0.26 .

6	0	0	4
8	4	0.633	2.266
10	16	1.265	3.53
Σx : 30	40		

$$\frac{\Sigma x}{n} = \frac{30}{5} = 6 = \bar{x}$$

$$Var(x) = \frac{1}{n_1} \Sigma (x - \bar{x})^2 .$$

$$\frac{40}{4} = 10$$

$$Var(x) = 10 .$$

(d). O/p of the conv operation :

$6 \times 6 * 3 \times 3 = 4 \times 4$.

30 + 0 - 30 = 0	30 + 0 + 0 = 30	30	0
0	30	30	0
0	30	30	0
0	30	30	0