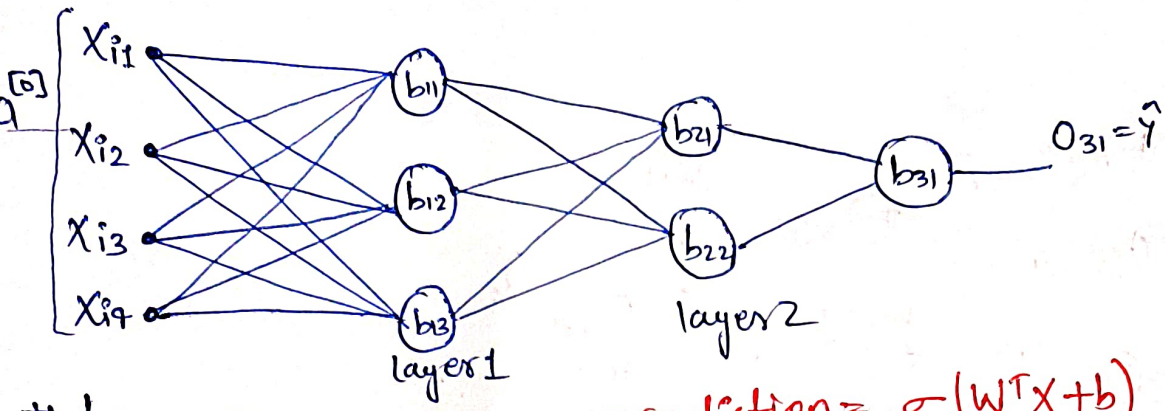
→ Forward Propogation



Layer I

$$\text{prediction} = \sigma(W^T X + b)$$

$$\begin{bmatrix} w'_{11} & w'_{12} & w'_{13} \\ w'_{21} & w'_{22} & w'_{23} \\ w'_{31} & w'_{32} & w'_{33} \\ w'_{41} & w'_{42} & w'_{43} \end{bmatrix}^T \begin{bmatrix} x_{i1} \\ x_{i2} \\ x_{i3} \\ x_{i4} \end{bmatrix} + \begin{bmatrix} b_{11} \\ b_{12} \\ b_{13} \end{bmatrix}$$

$4 \times 3 \xrightarrow{T} 3 \times 4$; 4×1 ; 3×1

$$\begin{bmatrix} w^1_{11} x_{i1} + w'_{21} x_{i2} + w^1_{31} x_{i3} + w^1_{41} x_{i4} \\ w^1_{12} x_{i1} + w^1_{22} x_{i2} + w'_{32} x_{i3} + w^1_{42} x_{i4} \\ w'_{13} x_{i1} + w^1_{23} x_{i2} + w'_{33} x_{i4} + w^1_{43} x_{i4} \end{bmatrix}_{3\times1} + \begin{bmatrix} b_{11} \\ b_{12} \\ b_{13} \end{bmatrix}$$

$$\sigma\left(\begin{bmatrix} w^1_{11} x_{i1} + w'_{21} x_{i2} + w^1_{31} x_{i3} + w^1_{41} x_{i4} + b_{11} \\ w'_{12} x_{i1} + w^1_{22} x_{i2} + w^1_{32} x_{i3} + w^1_{41} x_{i4} + b_{12} \\ w'_{13} x_{i1} + w^1_{23} x_{i2} + w'_{33} x_{i4} + w^1_{43} x_{i4} + b_{13} \end{bmatrix} \right) \rightarrow \begin{bmatrix} O_{11} \\ O_{12} \\ O_{13} \end{bmatrix} = a^{[1]}$$

Layer II

$$\begin{bmatrix} w^2_{11} & w^2_{12} \\ w^2_{21} & w^2_{22} \\ w^2_{31} & w^2_{32} \end{bmatrix}^T \begin{bmatrix} O_{11} \\ O_{12} \\ O_{13} \end{bmatrix} + \begin{bmatrix} b_{21} \\ b_{22} \end{bmatrix}_{2\times1}$$

$3 \times 2 \xrightarrow{T} 2 \times 3$; 3×1

$$\sigma\left(\begin{bmatrix} w^2_{11} O_{11} + w^2_{21} O_{12} + w^2_{31} O_{13} + b_{21} \\ w^2_{12} O_{11} + w^2_{22} O_{12} + w^2_{32} O_{13} + b_{22} \end{bmatrix} \right) = \begin{bmatrix} O_{21} \\ O_{22} \end{bmatrix} = a^{[2]}$$

#layer 3

$$\underset{2\times1 \to 1\times2}{\begin{bmatrix} w_{11}^3 \\ w_{21}^3 \end{bmatrix}^T} \underset{2\times1}{\begin{bmatrix} o_{21} \\ o_{22} \end{bmatrix}} + \begin{bmatrix} b_{31} \end{bmatrix}$$

$$\sigma\left(\left[w_{11}^3 o_{21} + w_{21}^3 o_{22} + b_{31}\right]\right) = \hat{y}_i = o_{31} = a^{[3]}$$

$$a^{[1]} = \sigma(a^{[0]} w^{[1]} + b^{[1]})$$

$$a^{[2]} = \sigma(a^{[1]} w^{[2]} + b^{[2]})$$

$$a^{[3]} = \sigma(a^{[2]} w^{[3]} + b^{[3]})$$

$$a^{[3]} = \sigma\Big(\underbrace{\sigma\big(\underbrace{\sigma(a^{[0]} w^{[1]} + b^{[1]})}_{a^1} w^{[2]} + b^{[2]}\big)}_{a^2} w^{[3]} + b^{[3]}\Big)$$