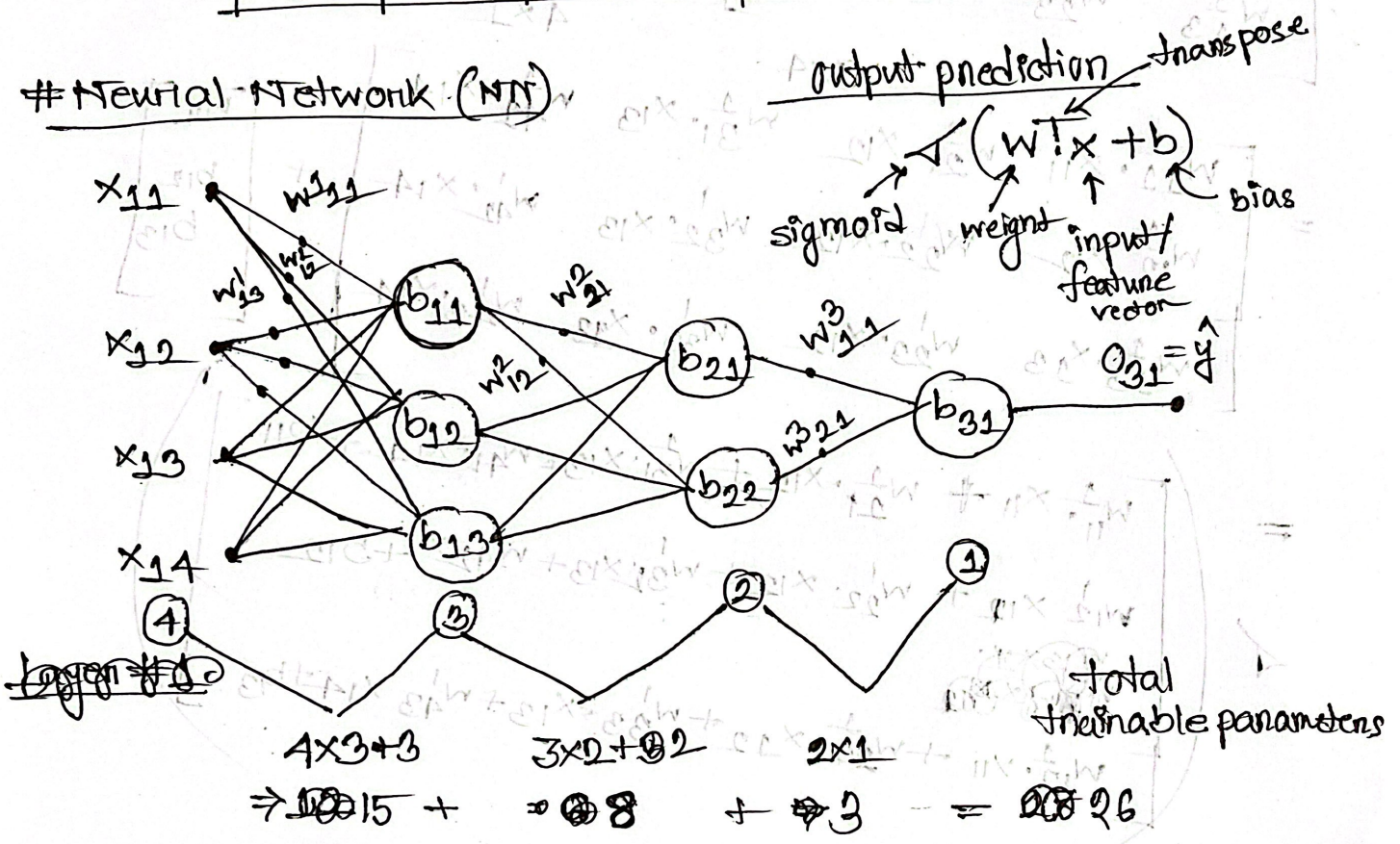


①

Number of trainable parameters.

Forward Propagation				
x_{11}	x_{12}	x_{13}	x_{14}	\hat{y}
CGPA	IQ	1000 HSC	SSC	Rank
3.9	92	785	1250	1
3.7	89	772	1235	2

Neural Network (NN)



Layer 1

$$W_{11}^T \cdot x_{11} + W_{12}^T \cdot x_{12} + W_{13}^T \cdot x_{13} + W_{14}^T \cdot x_{14} + b_{11}$$

$$\begin{bmatrix} w_{11}^1 & w_{12}^1 & w_{13}^1 \\ w_{21}^1 & w_{22}^1 & w_{23}^1 \\ w_{31}^1 & w_{32}^1 & w_{33}^1 \\ w_{41}^1 & w_{42}^1 & w_{43}^1 \end{bmatrix}^T \cdot \begin{bmatrix} x_{11} \\ x_{12} \\ x_{13} \\ x_{14} \end{bmatrix} + \begin{bmatrix} b_{11} \\ b_{12} \\ b_{13} \end{bmatrix} = \begin{bmatrix} 0_{11} \\ 0_{12} \\ 0_{13} \end{bmatrix}$$

$$4 \times 3 \rightarrow 3 \times 4 \cdot 4 \times 1 = 3 \times 1 + 3 \times 1 = 3 \times 1$$

$$\begin{bmatrix} w_{11}^1 & w_{21}^1 & w_{31}^1 & w_{41}^1 \\ w_{12}^1 & w_{22}^1 & w_{32}^1 & w_{42}^1 \\ w_{13}^1 & w_{23}^1 & w_{33}^1 & w_{43}^1 \end{bmatrix} \cdot \begin{bmatrix} x_{11} \\ x_{12} \\ x_{13} \\ x_{14} \end{bmatrix} + \begin{bmatrix} b_{11} \\ b_{12} \\ b_{13} \end{bmatrix}$$

3x4 4x1

$$= \begin{bmatrix} w_{11}^1 \cdot x_{11} & w_{21}^1 \cdot x_{12} & w_{31}^1 \cdot x_{13} & w_{41}^1 \cdot x_{14} \\ w_{12}^1 \cdot x_{11} & w_{22}^1 \cdot x_{12} & w_{32}^1 \cdot x_{13} & w_{42}^1 \cdot x_{14} \\ w_{13}^1 \cdot x_{11} & w_{23}^1 \cdot x_{12} & w_{33}^1 \cdot x_{13} & w_{43}^1 \cdot x_{14} \end{bmatrix} + \begin{bmatrix} b_{11} \\ b_{12} \\ b_{13} \end{bmatrix}$$

$$= \begin{bmatrix} w_{11}^1 \cdot x_{11} + w_{21}^1 \cdot x_{12} + w_{31}^1 \cdot x_{13} + w_{41}^1 \cdot x_{14} + b_{11} \\ w_{12}^1 \cdot x_{11} + w_{22}^1 \cdot x_{12} + w_{32}^1 \cdot x_{13} + w_{42}^1 \cdot x_{14} + b_{12} \\ w_{13}^1 \cdot x_{11} + w_{23}^1 \cdot x_{12} + w_{33}^1 \cdot x_{13} + w_{43}^1 \cdot x_{14} + b_{13} \end{bmatrix}$$

$$= \begin{bmatrix} 0_{11} \\ 0_{12} \\ 0_{13} \end{bmatrix}$$

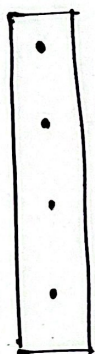
Layer 2

$$\begin{bmatrix} w_{11}^2 \\ w_{21}^2 \\ w_{31}^2 \\ w_{41}^2 \end{bmatrix} \cdot \begin{bmatrix} 0_{11} \\ 0_{12} \\ 0_{13} \end{bmatrix} + \begin{bmatrix} b_{21} \\ b_{22} \end{bmatrix} = \begin{bmatrix} 0_{21} \\ 0_{22} \end{bmatrix}$$

Layer 3

$$\begin{bmatrix} w_{11}^3 \\ w_{21}^3 \end{bmatrix}^T \cdot \begin{bmatrix} 0.21 \\ 0.22 \end{bmatrix} + \begin{bmatrix} b_{31} \end{bmatrix} = \begin{bmatrix} 0.31 \end{bmatrix} = \hat{y}^{(1)}$$

activation
function
 $a^{[0]} \leftarrow$



0

0

0

chain:

$$\sigma \left(\sigma \left(\sigma \left(a^{[0]}, w^{[1]} + b^{[1]} \right) \cdot w^{[2]} + b^{[2]} \right) \cdot w^{[3]} + b^{[3]} \right)$$

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

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

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