Artificial Neural Networks

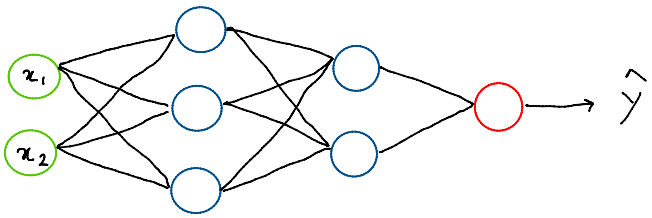
24 February 2024 10:50 AM

Type of Layer	Number Of Layers	Number Of Neurons
Input Layer	One	Number of Inputs
Output Layer	One	Number of Outputs (Task)
Hidden Layer	Zero or More	

Task	Activation (Hidden Layers)	Activation (Output Layer)	Loss Function	No. of Neurons(Output)
Regression	ReLU	Linear	Mean Squared Error (MSE)	1
Binary Classification	ReLU	Sigmoid	Binary Crossentropy (Log Loss)	1
Multiclass Classification	ReLU	Softmax	Categorical Crossentropy	N (N : Number of classes)

Ex: LOAN DEFAULT

Income	Spend	Default
-	-	1
		0



* HIDDEN LAYER 1 :

$$z_1 = w_{11}x_1 + w_{12}x_2 + b_1, \quad a_1 = ReLU(z_1)$$
$$z_2 = w_{21}x_1 + w_{22}x_2 + b_2, \quad a_2 = ReLU(z_2)$$
$$z_3 = w_{31}x_1 + w_{32}x_2 + b_3, \quad a_3 = ReLU(z_3)$$

Outputs : a_1, a_2, a_3

* HIDDEN LAYER 2 :-

$$z_1 = w_{11}a_1 + w_{12}a_2 + w_{13}a_3 + b_1, \quad a_1 = ReLU(z_1)$$
$$z_2 = w_{11}a_1 + w_{12}a_2 + w_{13}a_3 + b_2, \quad a_2 = ReLU(z_2)$$

Outputs : a_1, a_2

* OUTPUT LAYER :-

$z_1 = w_{11} a_1 + w_{12} a_2 + b_1$, $a_1 = \text{Sigmoid}(z_1)$

---

* BASIC LINEAR ALGEBRA :-

* MATRIX

$$\begin{bmatrix} 1 & 2 & 1 \\ 3 & 2 & 4 \\ 4 & 9 & 8 \end{bmatrix}$$

* VECTOR :-

> ROW VECTOR : $[2 \ 3 \ 1]$

> COLUMN VECTOR : $\begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix}$

* ADDITION :-

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 3 \\ 3 & 1 & 2 \end{bmatrix} + \begin{bmatrix} 2 & 1 & 4 \\ 3 & 2 & 1 \\ 2 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 3 & 5 \\ 5 & 3 & 4 \\ 5 & 2 & 4 \end{bmatrix}$$

* The shapes must be same

* MULTIPLICATION :-

* $(M, N) * (P, Q)$

> $N = P$ (Condition)

> (M, Q) (Shape of result)

> $\begin{bmatrix} 1 & 2 & 0 \\ 1 & 0 & 3 \\ 2 & 1 & 0 \end{bmatrix} * \begin{bmatrix} 2 & 1 & 0 \\ 3 & 1 & 2 \\ 2 & 1 & 3 \end{bmatrix} = \begin{bmatrix} 8 & 3 & 4 \\ 8 & 4 & 9 \\ 7 & 3 & 2 \end{bmatrix}$

$(3, 3)$ $(3, 3)$ $=$ (3×3)

> $\begin{bmatrix} 1 & 2 & 0 \\ 1 & 0 & 3 \\ 2 & 1 & 0 \end{bmatrix} * \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 5 & 4 \\ 1 & 11 \\ 4 & 5 \end{bmatrix}$

$(3, 3)$ $(3, 2)$ $(3, 2)$

> $[2 \ 1 \ 2] * \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = [6]$

$(1, 3)$ $(3, 1)$ $(1, 1)$

> $\begin{bmatrix}1\\2\\1\end{bmatrix} * \begin{bmatrix}3 & 4 & 2\end{bmatrix} = \begin{bmatrix}3 & 4 & 2\\6 & 8 & 4\\3 & 4 & 2\end{bmatrix}$

$(3,1) \quad (1,3) \quad (3,3)$

Q.) $\begin{bmatrix}2 & 7 & 3\\8 & 4 & 2\\9 & 0 & 3\end{bmatrix} * \begin{bmatrix}3 & 8 & 2\\9 & 2 & 0\\4 & 1 & 1\end{bmatrix} = \begin{bmatrix}81 & 33 & 7\\68 & 74 & 18\\39 & 75 & 21\end{bmatrix}$

---

* Input to a neural network is represented by a vector

$$X = \begin{bmatrix}x_1\\x_2\end{bmatrix}$$

* Weights of each layer are represented as a matrix
  > Weights of a neuron in a row
  > Number of rows are equal to number of neurons

$$W1 = \begin{bmatrix}w_{11} & w_{12}\\w_{21} & w_{22}\\w_{31} & w_{32}\end{bmatrix}, \quad W2 = \begin{bmatrix}w_{11} & w_{12} & w_{13}\\w_{21} & w_{22} & w_{23}\end{bmatrix}, \quad W3 = \begin{bmatrix}w_{11} & w_{12}\end{bmatrix}$$

* Biases associated with each layer are represented by a vector.

$$B1 = \begin{bmatrix}b_1\\b_2\\b_3\end{bmatrix}, \quad B2 = \begin{bmatrix}b_1\\b_2\end{bmatrix}, \quad B3 = \begin{bmatrix}b_1\end{bmatrix}$$

* Outputs associated with each layer are represented by a vector

$$A1 = \begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix}, \quad A2 = \begin{bmatrix}a_1\\a_2\end{bmatrix}, \quad A3 = \begin{bmatrix}a_1\end{bmatrix}$$

* HIDDEN LAYER 1 :-

$$Z1 = W1 . X + B1$$
$$A1 = ReLU(Z1)$$

* HIDDEN LAYER 2 :-

$$Z2 = W2 \cdot A1 + B2$$
$$A2 = ReLU(Z2)$$

* OUTPUT LAYER :-

$$Z3 = W3 \cdot A2 + B3$$
$$A3 = Sigmoid(Z3)$$

* COMPUTATION GRAPH :-