Artificial Neural Network

12 January 2024 08:55 PM

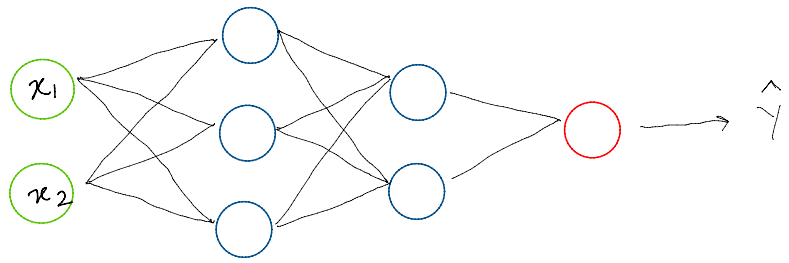
	#LAYERS	#NEURONS
INPUT LAYER	1	NO. of inputs
OUTPUT LAYER	1	Task
HIDDEN LAYER	0 or more	

	ACTIVATION (HIDDEN)	ACTIVATION (OUTPUT)	LOSS FUNCTION	#NEURONS (OUTPUT)
REGRESSION	RELU	LINEAR	MSE	1
B CLASSIFICATION	RELU	SIGMOID	LOG LOSS	1
M CLASSIFICATION	RELU	SOFTMAX	CATEGORICAL CROSSENTROPY	n

∴ LOG LOSS : BINARY CROSSENTROPY

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 :

* HIDDEN LAYER 2:

Input : a_1, a_2, a_3

$$z_1 = w_{11} a_1 + w_{12} a_2 + w_{13} a_3 + b_1 \;, \quad a_1 = RELU(z_1)$$
$$z_2 = w_{21} a_1 + w_{22} a_2 + w_{23} a_3 + b_2 \;, \quad a_2 = RELU(z_2)$$

Output : a_1, a_2

* OUTPUT LAYER:

Input = a_1, a_2

$$z_1 = w_{11} a_1 + w_{12} a_2 + b_1 \;, \quad a_1 = SIGMOID(z_1)$$

---

* MATRIX :-

$$\begin{bmatrix} 1 & 2 & 1 \\ 4 & 6 & 7 \\ 8 & 2 & 9 \end{bmatrix}$$

* VECTOR :-

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

> COLUMN VECTOR : $\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$

* ADDITION :-

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

$(3,3) \qquad (3,3) \qquad (3,3)$

> Shapes must be same.

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

* MULTIPLICATION :-

$(M, N) * (P, Q)$, $N = P$ (Condition for *)

(M, Q) (Shape of result)

i) $\begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 1 & 0 & 2 \end{bmatrix} * \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & 3 \\ 2 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 7 & 9 \\ 4 & 4 & 7 \\ 5 & 4 & 5 \end{bmatrix}$

$(3,3) \qquad (3,3) \qquad (3,3)$

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

$(3,3) \qquad (3,2) \qquad (3,2)$

iii) $\begin{bmatrix} 2 & 1 & 2 \end{bmatrix} * \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \end{bmatrix}$

$(1,3) \qquad (3,1) \qquad (1,1)$

iv) $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} * \begin{bmatrix} 2 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 2 \\ 4 & 2 & 4 \\ 2 & 1 & 2 \end{bmatrix}$

$(3,1) \qquad (1,3) \qquad (3,3)$

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

$(3,3) \qquad (3,1) \qquad (3,1)$

---

* Input to a neural network is represented by a vector.

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

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

* Weights of each layer are represented by a matrix.
  - > Weights of a neuron in a row.
  - > Number of rows 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 of each layer are represented as 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 of each layer are represented as 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 :-

$$Z1 = W1.X + B1, \qquad Z1(3,1) : \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$$

$$A1 = RELU(Z1), \qquad A1(3,1) : \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

* HIDDEN LAYER 2 :-

$$Z2 = W2.A1 + B2, \qquad Z2(2,1) : \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

$$A2 = RELU(Z2), \qquad A2(2,1) : \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

* OUTPUT LAYER :

$$Z3 = W3.A2 + B3, \qquad Z3(1,1) : [z_1]$$

$$A3 = Sigmoid(Z3), \qquad A3(1,1) : [a_1]$$

---

* COMPUTATION GRAPH :-

* COMPUTATION GRAPH :-