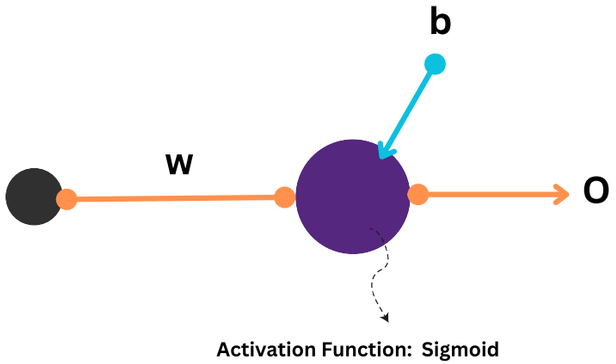
2. Obtain new value of weight and bias after Gradient descent first iteration, use MSE as a loss function.



Initial weight value(weight have been initialized randomly) : 1.5

Initial bias value (weight have been initialized randomly) : 0.5

learning rate (alpha) = 0.1

$x = 1,3$
$y = 2,7$
$w = 1,5$
$b = 0,5$
$\alpha = 0,1$

Forward pass : $z = wx + b \rightarrow (1,5 \times 1,3) + 0,5 = 2,45$

$\hat{y} = \text{Sigmoid}(z) \rightarrow \text{Sigmoid}(2,45) = 0,92$

Back propagation : $L = (\hat{y} - y)^2$

$$\frac{\partial L}{\partial w} = \frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z} \cdot \frac{\partial z}{\partial w} \rightarrow \left(2(\hat{y} - y)\right) \cdot \left(\hat{y}(1 - \hat{y})\right) \cdot x \approx -0,34$$

$$\frac{\partial L}{\partial b} = \frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z} \cdot \frac{\partial z}{\partial b} \rightarrow \left(2(\hat{y} - y)\right) \cdot \left(\hat{y}(1 - \hat{y})\right) \cdot 1 \approx -0,26$$

New w, b :

$$w_{new} = w - \alpha \frac{\partial L}{\partial w} \rightarrow 1,5 - 0,1(-0,34) \approx 1,534$$

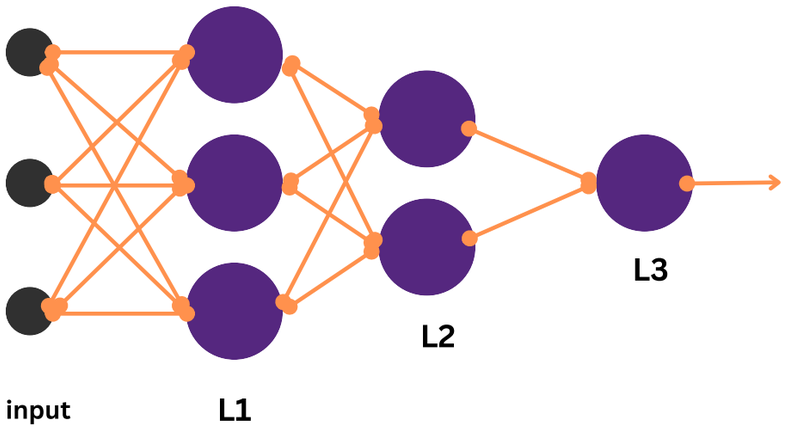
$$b_{new} = b - \alpha \frac{\partial L}{\partial b} \rightarrow 0,5 - 0,1(-0,26) \approx 0,526$$

3. For the below feed-forward neural network architecture obtain the following variables:

O_1^1 = ?

O_2^2 = ?

O_1^3 = y_predicted = ?



inputs

$$\begin{bmatrix} 0.3 \\ 0.5 \\ 0.8 \end{bmatrix}$$

L1-Weights

$$\begin{bmatrix} 0.1 & 0.2 & 0.2 \\ 0.4 & 0.1 & 0.2 \\ 0.2 & 0.3 & 0.1 \end{bmatrix}$$

L1-Biases

$$\begin{bmatrix} 0.2 \\ 0.2 \\ 0.1 \end{bmatrix}$$

L2-Weights

$$\begin{bmatrix} 0.1 & 0.2 & 0.2 \\ 0.4 & 0.1 & 0.2 \end{bmatrix}$$

L2-Biases

$$\begin{bmatrix} 0.1 \\ 0.4 \end{bmatrix}$$

L3-Weights

$$\begin{bmatrix} 0.5 & 0.1 \end{bmatrix}$$

L3-Biases

$$\begin{bmatrix} 0.5 \end{bmatrix}$$

Activation Functions:

$$\begin{bmatrix} \text{L1: Relu} \\ \text{L2: Tanh} \\ \text{L3: Sigmoid} \end{bmatrix}$$

L1 8 (Relu)

$$z_1^1 = (0,3 \times 0,1) + (0,5 \times 0,2) + (0,8 \times 0,2) + 0,2 = 0,49$$

$$\longrightarrow O_1^1 = \text{Relu}(0,49) = 0,49$$

$$z_2^1 = (0,3 \times 0,4) + (0,5 \times 0,1) + (0,8 \times 0,2) + 0,2 = 0,53$$

$$\longrightarrow O_2^1 = \text{Relu}(0,53) = 0,53$$

$$z_3^1 = (0,3 \times 0,2) + (0,5 \times 0,3) + (0,8 \times 0,1) + 0,1 = 0,39$$

$$\longrightarrow O_3^1 = \text{Relu}(0,39) = 0,39$$

L2 8 (tanh)

$$z_2^2 = (0,49 \times 0,4) + (0,53 \times 0,1) + (0,39 \times 0,2) + 0,4 = 0,73$$

$$\longrightarrow O_2^2 = \tanh(0,73) = 0,62$$

$$z_1^2 = (0,49 \times 0,1) + (0,53 \times 0,2) + (0,39 \times 0,2) + 0,1 = 0,3$$

$$\longrightarrow o_1^2 = \tanh(0,33) = 0,32$$

L3 8 $Z_1^3 = (0,32 \times 0,5) + (0,62 \times 0,1) + 0,5 = 0,72$

(Sigmoid) → $O_1^3 = \text{Sigmoid}(0,72) = 0,67$

$O_1^1 = 0,49$

$O_2^2 = 0,62$

$O_1^3 = 0,67$