1. **Consider a simple neural network with one input layer, one hidden layer with 2 neurons, and one output layer. The activation function used in both hidden and output layers is the sigmoid function.**
   Given the input values [2, 3] and the following weights and biases:

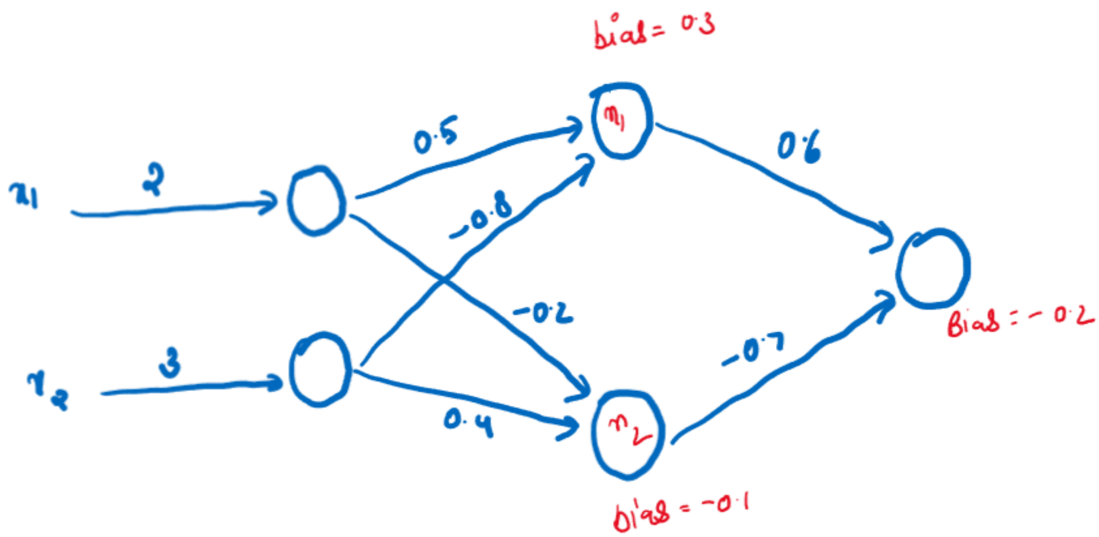
   Hidden Layer:
   Neuron 1: Weight1 = 0.5, Weight2 = -0.8, Bias = 0.3
   Neuron 2: Weight1 = -0.2, Weight2 = 0.4, Bias = -0.1

   Output Layer:
   Neuron 1: Weight1 = 0.6, Weight2 = -0.7, Bias = -0.2
   Calculate the output value of this neural network using forward propagation.

Solution :

Input Layer:

Given input values: [2, 3]



Hidden Layer:

For Neuron 1:

Weighted sum = (2 * 0.5) + (3 * -0.8) + 0.3 = 1 - 2.4 + 0.3 = -1.1

Activation = sigmoid = 1 / 1+e^-x = sigmoid(-1.1) =

0.24973989440481

≈ 0.25 (rounded to two decimal places)

For Neuron 2:

Weighted sum = (2 * -0.2) + (3 * 0.4) - 0.1 = -0.4 + 1.2 - 0.1 = 0.7

Activation = sigmoid(0.7) = 0.6681877721682206 ≈ 0.67 (rounded to two decimal places)

Output Layer:

For Neuron 1:

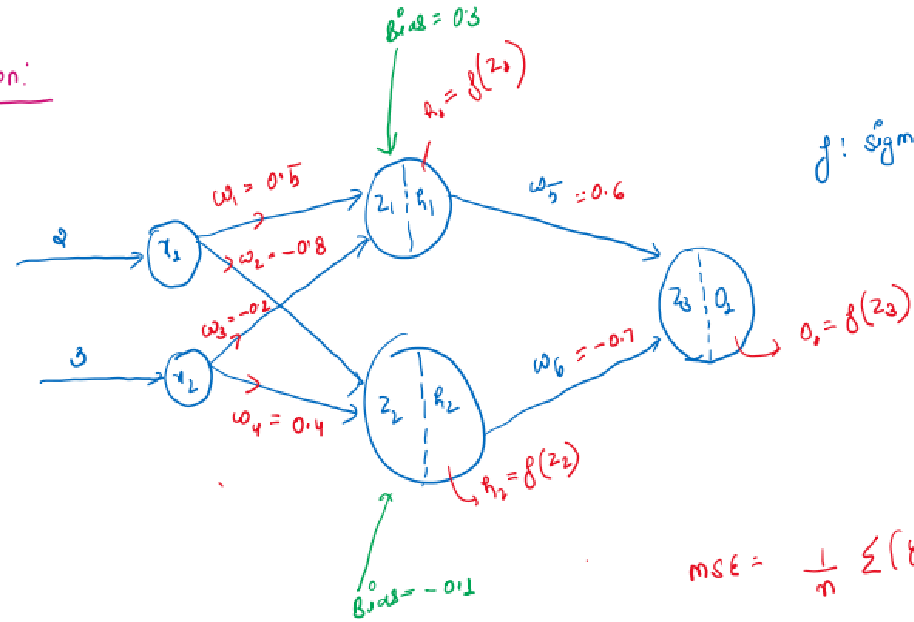
Weighted sum = (0.25 * 0.6) + (0.67 * -0.7) - 0.2 ≈ 0.15 - 0.47 - 0.2 ≈ -0.52

Activation = sigmoid(-0.52) = 0.3728522336867617 ≈ 0.37 (rounded to two decimal places)

So, the output value of this neural network using forward propagation is approximately **0.37**.

2. **Consider the same neural network as in question 1. Given that the expected output value is 0.8, calculate the weight updates needed for the output layer using backpropagation with a learning rate of 0.1.**
   Hint: Use the Mean Squared Error (MSE) loss function for this calculation.

Backpropagation:



f : sigmoid

$$MSE = \frac{1}{n} \sum (y_i - \hat{y}_i)^2$$

$$z_1 = x_1 w_1 + x_2 w_3 + Bias = -1.1$$

$$h_1 = \text{Sigmoid}(-1.1) = 0.25$$

$$z_2 = x_1 w_2 + x_2 w_4 = 0.7$$

$$h_2 = \text{Sigmoid}(0.7) = 0.67$$

$$z_3 = w_5 h_1 + w_6 h_2$$

$$= -0.52$$

$$O_1 = \text{Sigmoid}(-0.52)$$

$$= 0.37$$

$$mse = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

$$= (0.8 - 0.37)^2$$

$$= \underline{0.1849} \text{ — loss}$$

updated weights = $\boxed{w' = w - \eta \frac{\partial L}{\partial w}}$

η = learning rate = 0.1

$$w_5' = w_5 - \eta \frac{\partial L}{\partial w_5}$$

$$\frac{\partial L}{\partial w_5} = \frac{\partial L}{\partial o_1} \times \frac{\partial o_1}{\partial z_3} \times \frac{\partial z_3}{\partial w_5}$$

$o_1 = \hat{y}_1$ predicted output

$$\frac{\partial L}{\partial o_1} = \frac{\partial mse}{\partial o_1} = \frac{\partial (y_i - \hat{y}_i)^2}{\partial y_i}$$

$$= 2 \times (y_i - \hat{y}_i) = 2 \times (0.8 - 0.37) = 0.86$$

$$\frac{\partial o_1}{\partial z_3} = \frac{\partial\, sig(z_3)}{\partial z_3} = sig(z_3) \times (1 - sig(z_3)) =$$

$$= 0.37 \times (1 - 0.37)$$

$$\frac{\partial z_3}{\partial w_5} = \frac{\partial\, w_5 h_1 + w_6 h_2}{\partial w_5} = h_1$$

$$\frac{\partial L}{\partial \omega_5} = 2 \times (y_i - \hat{y}_i) \times Sig(z_3)(1 - Sig(z_3)) \times x_1$$

$$= 2 \times (0.8 - 0.37) \times 0.37 \times (1 - 0.37) \times 0.25$$

$$= 0.86 \times 0.2331 \times 0.25$$

$$= 0.050$$

$$\omega_5' = \omega_5 - \eta \frac{\partial L}{\partial \omega_5}$$

$$= 0.6 - 0.1 \times 0.050$$

$$= 0.6 - 0.005$$

$$= \underline{0.595}$$

$$\frac{\partial L}{\partial \omega_6} = 2 \times (y - \hat{y}_i) \times Sig(z_3)(1 - Sig(z_3) \times x_2$$

$$= 2 \times (0.8 - 0.37) \times 0.37 \times (1 - 0.37) \times 0.67$$

$$= 0.86 \times 0.2331 \times 0.67$$

$$= 0.13431222$$

$$\omega_6' = \omega_6 - \eta \frac{\partial L}{\partial \omega_6}$$

$$= -0.7 - 0.1 \times 0.1343$$

$$= -0.7 - 0.01343$$

$$= -0.71343$$

$$w_1' = w_1 - \eta \frac{\partial L}{\partial w_1}$$

$$\frac{\partial L}{\partial w_1} = \frac{\partial L}{\partial O_1} \times \frac{\partial O_1}{\partial z_3} \times \frac{\partial z_3}{\partial h_1} \times \frac{\partial h_1}{\partial z_1} \times \frac{\partial z_1}{\partial w_1}$$

$O_1 = \hat{y}_1$

$$\frac{\partial L}{\partial O_1} = \frac{\partial MSE}{\partial O_1} = \frac{(y_i - \hat{y}_i)^2}{\partial \hat{y}_1} = 2(y_i - \hat{y}_i) =$$

$$= 2(0.8 - 0.37) = 0.86$$

$$\frac{\partial O_1}{\partial z_3} = \frac{\partial\, sigmoid(z_3)}{\partial z_3} = sigmoid(z_3)\,(1 - sigmoid(z_3))$$

$$= 0.37 \times (1 - 0.37)$$

$$= 0.2331$$

$$\frac{\partial z_3}{\partial h_1} = \frac{\partial(w_5 h_1 + w_5 h_2)}{\partial h_1} = w_5 = 0.6$$

$$\frac{\partial h_1}{\partial z_1} = \frac{\partial\, sig(z_1)}{\partial z_1} = sigmoid(z_1) \times (1 - sigmoid(z_1))$$

$$= 0.25 \times (1 - 0.25)$$

$$= 0.25 \times 0.75 = 0.1875$$

$$\frac{\partial z_1}{\partial w_1} = \frac{\partial(x_1 w_1 + x_2 w_2)}{\partial w_1} = x_1 = 2$$

$$\frac{\partial L}{\partial w_1} = 0.86 \times 0.2331 \times 0.6 \times 0.1875 \times 2$$

$$= \underline{\underline{0.0451}}$$

$$w_1' = w_1 - \eta \frac{\partial L}{\partial w_1}$$

$$= 0.5 - 0.1 \times 0.0451$$

$$= 0.5 - 0.00451$$

$$= \underline{\underline{0.49}}$$

$$w_3' = w_3 - \eta \frac{\partial L}{\partial w_3}$$

$$\frac{\partial L}{\partial w_3} = \frac{\partial L}{\partial o_1} \times \frac{\partial o_1}{\partial z_3} \times \frac{\partial z_3}{\partial h_1} \times \frac{\partial h_1}{\partial z_1} \times \frac{\partial z_1}{\partial w_3}$$

this change

$$=$$

$$\frac{\partial z_1}{\partial w_3} = \frac{\partial (x_1 w_1 + x_2 w_3)}{\partial w_3} = x_2$$

$$\frac{\partial L}{\partial w_3} = 0.86 \times 0.2331 \times 0.6 \times 0.1875 \times 3$$

$$= 0.06765$$

$$w_3' = w_3 - \eta \frac{\partial L}{\partial w_3}$$

$$= -0.8 - 0.1 \times 0.06765$$

$$= -0.8 - 0.006765 = \underline{-0.806765}$$

$$w_2' = ?$$

$$w_4' = ?$$

3. **Consider a neural network with one input layer, one hidden layer with 3 neurons, and one output layer with 4 neurons. The activation function used in both hidden and output layers is the ReLU function.**

Given the input values [1, 2], and the following weights and biases:

Hidden Layer:

- **Neuron 1: Weight1 = 0.5, Weight2 = -0.8, Bias = 0.3**
- **Neuron 2: Weight1 = -0.2, Weight2 = 0.4, Bias = -0.1**
- **Neuron 3: Weight1 = 0.1, Weight2 = 0.3, Bias = 0.2**

Output Layer:

- **Neuron 1: Weight1 = 0.6, Weight2 = -0.7, Weight3 = 0.4, Bias = -0.2**
- **Neuron 2: Weight1 = -0.3, Weight2 = 0.2, Weight3 = -0.5, Bias = 0.1**
- **Neuron 3: Weight1 = 0.2, Weight2 = 0.1, Weight3 = -0.6, Bias = 0.3**
- **Neuron 4: Weight1 = -0.4, Weight2 = 0.5, Weight3 = -0.3, Bias = -0.1**

Given that the expected outputs are [0.8, 0.6, 0.4, 0.2], calculate the weight updates needed for the output layer using backward propagation with a learning rate of 0.01.

Hint: Use the Mean Squared Error (MSE) loss function for this calculation.