Neural Network Output Calculation Problem

Problem Statement

Consider a neural network with the following architecture:

- Input layer: 3 nodes (x_1, x_2, x_3)
- Hidden layer: 5 nodes (h_1, h_2, h_3, h_4, h_5)
- Output layer: 1 node (y)

Each node uses the sigmoid activation function:

$$\sigma(z) = \frac{1}{1+e^{-z}}$$

Random Weights and Biases

Weights:

- Weights from input to hidden layer:

$$W_1 = \begin{pmatrix} w_{11} & w_{12} & w_{13} & w_{14} & w_{15} \\ w_{21} & w_{22} & w_{23} & w_{24} & w_{25} \\ w_{31} & w_{32} & w_{33} & w_{34} & w_{35} \end{pmatrix} = \begin{pmatrix} 0.6 & -0.1 & 0.8 & -0.2 & 0.3 \\ -0.3 & 0.7 & -0.6 & 0.4 & -0.5 \\ 0.5 & 0.4 & 0.3 & 0.9 & 0.2 \end{pmatrix}$$

- Weights from hidden to output layer:

$$W_2 = \begin{pmatrix} w_{1y} \\ w_{2y} \\ w_{3y} \\ w_{4y} \\ w_{5y} \end{pmatrix} = \begin{pmatrix} 1.1 \\ -0.6 \\ 0.4 \\ 0.7 \\ -0.3 \end{pmatrix}$$

Biases:

$$b_{h1} = 0.2,\ b_{h2} = -0.1,\ b_{h3} = 0.3,\ b_{h4} = -0.4,\ b_{h5} = 0.1,\ b_y = 0.5$$

Input Values:

$$X = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} = \begin{pmatrix} 0.5 & -0.3 & 0.7 \end{pmatrix}$$

Task:

Compute the output y of the neural network.

Notes

The formula used for calculating the output of a particular layer in a neural network is generally based on the following two steps:

1. **Weighted Sum:** The first step is to compute the weighted sum of the inputs to the node (which is typically the previous layer's output or the input layer for the first hidden layer), including the bias term.
2. **Activation Function:** The second step is to pass the weighted sum through an activation function to obtain the output of the node.

A schematic of such a calculation can be seen in the figure below (figure 1).

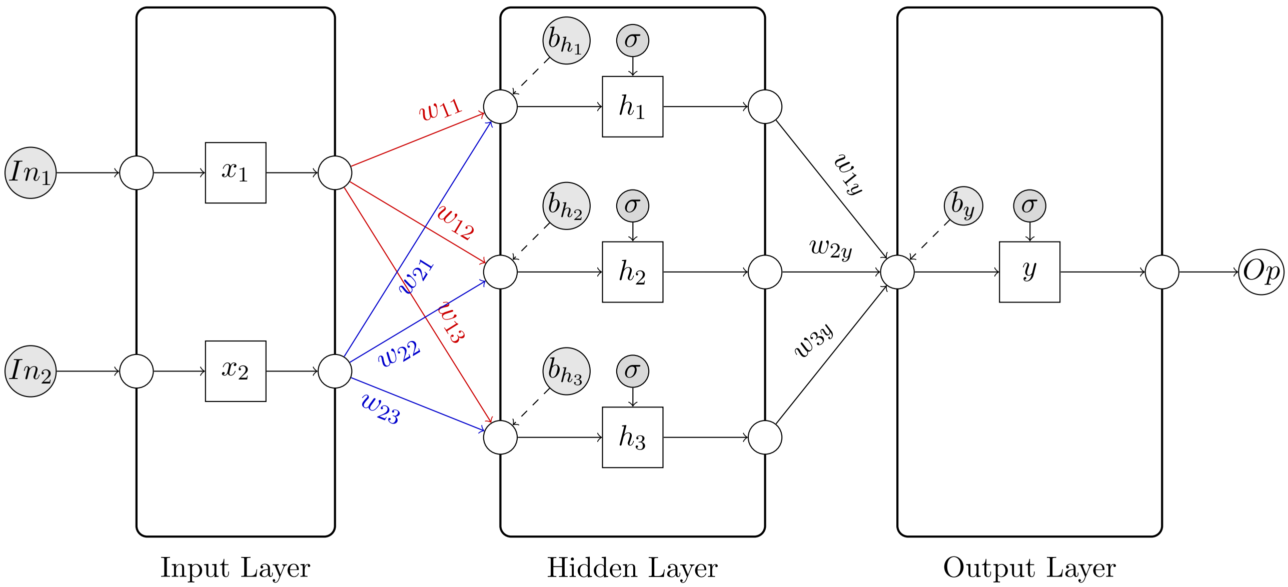

Figure 1: A schematic to represent the layers, connections and weights and biases.
Note: This schematic does not represent the network given in the problem.

Let's break this down more formally:
For a node j in layer L, the weighted sum of the inputs can be written as:

$$z_j = \sum_i w_{ij} \cdot a_i + b_j$$

In the above formula,

- w_{ij} is the weight from node i in the previous layer, $L - 1$ to node j in the current layer, L.

  For example: weight of the connection between x_1 and h_3 is w_{13}.

- a_i is the output of node i from the previous layer (or input value for the first layer),

  For example, a_2 is the input associated with node x_2,

- b_j is the bias associated with node j, For example, b_{h_2} is the bias associated with node h_2,

- The sum runs over all nodes j in the previous layer.

Once the weighted sum is computed, the result is passed through an activation function, typically a sigmoid, ReLU, or tanh function. For the sigmoid activation function, the output is:

$$o_i = \sigma(z_i) = \frac{1}{1 + e^{-z_i}}$$

where:

- $\sigma(z_i)$ is the sigmoid activation function,
- a_i is the output of node i in layer L.

For the hidden node h_1, the formula would look like:

$$z_{h1} = w_{11} \cdot x_1 + w_{21} \cdot x_2 + w_{31} \cdot x_3 + b_{h1}$$

where:

- x_1, x_2, x_3 are the inputs to the node (from the input layer),
- w_{11}, w_{21}, w_{31} are the weights connecting the input nodes to hidden node h_1,
- b_{h1} is the bias term for node h_1.

Then, the output of node h_1 would be:

$$h_1 = \sigma(z_{h1}) = \frac{1}{1 + e^{-z_{h1}}}$$

For the output layer, the output y is calculated by a similar process. Suppose the output layer has weights $w_{1y}, w_{2y}, w_{3y}, \ldots$ from each hidden node to the output node:

$$z_y = w_{1y} \cdot h_1 + w_{2y} \cdot h_2 + w_{3y} \cdot h_3 + \cdots + w_{5y} \cdot h_5 + b_y$$

where:

- $h_1, h_2, h_3, \ldots$ are the outputs from the hidden layer,
- $w_{1y}, w_{2y}, w_{3y}, \ldots$ are the weights from the hidden nodes to the output node,
- b_y is the bias term for the output node.

Finally, the output y would be:

$$y = \sigma(z_y) = \frac{1}{1 + e^{-z_y}}$$

Solution

Hidden Layer Calculations:

Subtask Draw the Neural network diagram similar to the one given above (see figure 1), and verify that the weights and calculation below are correct.

For each hidden node, calculate the weighted sum of inputs plus the bias, and apply the sigmoid function.

$$\begin{aligned} z_{h1} &= w_{11}x_1 + w_{21}x_2 + w_{31}x_3 + b_{h1} \\ &= (0.6 \times 0.5) + (-0.3 \times -0.3) + (0.5 \times 0.7) + 0.2 \\ &= 0.3 + 0.09 + 0.35 + 0.2 = 0.94 \end{aligned}$$

$$h_1 = \sigma(0.94) = \frac{1}{1 + e^{-0.94}} \approx 0.719$$

$$\begin{aligned} z_{h2} &= w_{12}x_1 + w_{22}x_2 + w_{32}x_3 + b_{h2} \\ &= (-0.1 \times 0.5) + (0.7 \times -0.3) + (0.4 \times 0.7) + (-0.1) \\ &= -0.05 - 0.21 + 0.28 - 0.1 = -0.08 \end{aligned}$$

$$h_2 = \sigma(-0.08) = \frac{1}{1 + e^{0.08}} \approx 0.480$$

$$\begin{aligned} z_{h3} &= w_{13}x_1 + w_{23}x_2 + w_{33}x_3 + b_{h3} \\ &= (0.8 \times 0.5) + (-0.6 \times -0.3) + (0.3 \times 0.7) + 0.3 \\ &= 0.40 + 0.18 + 0.21 + 0.3 = 1.09 \end{aligned}$$

$$h_3 = \sigma(1.09) = \frac{1}{1 + e^{-1.09}} \approx 0.748$$

$$\begin{aligned} z_{h4} &= w_{14}x_1 + w_{24}x_2 + w_{34}x_3 + b_{h4} \\ &= (-0.2 \times 0.5) + (0.4 \times -0.3) + (0.9 \times 0.7) + (-0.4) \\ &= -0.1 - 0.12 + 0.63 - 0.4 = 0.01 \end{aligned}$$

$$h_4 = \sigma(0.01) = \frac{1}{1 + e^{-0.01}} \approx 0.502$$

$$\begin{aligned} z_{h5} &= w_{15}x_1 + w_{25}x_2 + w_{35}x_3 + b_{h5} \\ &= (0.3 \times 0.5) + (-0.5 \times -0.3) + (0.2 \times 0.7) + 0.1 \\ &= 0.15 + 0.15 + 0.14 + 0.1 = 0.54 \end{aligned}$$

$$h_5 = \sigma(0.54) = \frac{1}{1 + e^{-0.54}} \approx 0.632$$

Output Layer Calculation:

Now, compute the output layer using the hidden layer outputs and the weights from the hidden layer to the output layer.

$$
\begin{aligned}
z_y &= w_{1y}h_1 + w_{2y}h_2 + w_{3y}h_3 + w_{4y}h_4 + w_{5y}h_5 + b_y \\
&= (1.1 \times 0.719) + (-0.6 \times 0.48) + (0.4 \times 0.748) + (0.7 \times 0.502) + (-0.3 \times 0.632) + 0.5 \\
&= 1.4639 \\
y &= \sigma(1.4639) = \frac{1}{1 + e^{-1.4639}} \approx 0.812
\end{aligned}
$$

Final Output:

$$\boxed{y \approx 0.812}$$

Question 1

Neural Network Architecture:

- **Input layer**: 2 nodes (x_1, x_2)
- **Hidden layer**: 2 nodes (h_1, h_2)
- **Output layer**: 1 node (y)

Each node uses the sigmoid activation function.
Weights and Biases:

$$W_1 = \begin{bmatrix} 0.3 & -0.5 \\ 0.6 & 0.2 \end{bmatrix}, \qquad b_h = \begin{bmatrix} -0.2 & 0.4 \end{bmatrix}$$

$$W_2 = \begin{bmatrix} 0.7 \\ -0.6 \end{bmatrix}, \qquad b_y = 0.3$$

Input Values:

$$X = (-0.3, 0.5) \tag{1}$$

Compute the output y of the neural network.

Question 2

Neural Network Architecture:

- **Input layer**: 3 nodes (x_1, x_2, x_3)
- **Hidden layer**: 4 nodes (h_1, h_2, h_3, h_4)
- **Output layer**: 2 nodes (y_1, y_2)

Each node uses the sigmoid activation function.
Weights and Biases:

$$W_1 = \begin{bmatrix} 0.3 & -0.2 & 0.5 & 0.1 \\ -0.4 & 0.7 & 0.1 & -0.6 \\ 0.2 & -0.5 & 0.8 & 0.4 \end{bmatrix}, \qquad b_h = \begin{bmatrix} 0.2 & -0.1 & 0.3 & 0.5 \end{bmatrix}$$

$$W_2 = \begin{bmatrix} 0.5 & -0.3 \\ -0.6 & 0.9 \\ 0.2 & -0.7 \\ 0.4 & 0.8 \end{bmatrix}, \qquad b_y = \begin{bmatrix} 0.1 & -0.2 \end{bmatrix}$$

Input Values:

$$X = (0.6, -0.3, 0.8) \tag{2}$$

Compute the outputs y_1, y_2 of the neural network.

Question 3

Neural Network Architecture:

- **Input layer**: 2 nodes (x_1, x_2)
- **Hidden layer**: 3 nodes (h_1, h_2, h_3)
- **Output layer**: 1 node (y)

Weights and Biases:

$$W_1 = \begin{bmatrix} 0.7 & -0.3 & 0.4 \\ -0.5 & 0.6 & -0.2 \end{bmatrix}, \qquad b_h = \begin{bmatrix} 0.1 \\ -0.2 \\ 0.3 \end{bmatrix}$$

$$W_2 = \begin{bmatrix} 0.9 \\ -0.8 \\ 0.5 \end{bmatrix}, \qquad b_y = 0.4$$

Input Values:

$$X = (-0.5, 0.7) \tag{3}$$

Compute the output y of the neural network.

Question 4

Neural Network Architecture:

- **Input layer**: 3 nodes (x_1, x_2, x_3)
- **Hidden layer**: 3 nodes (h_1, h_2, h_3)
- **Output layer**: 2 nodes (y_1, y_2)

Each node uses the sigmoid activation function.
Weights and Biases:

$$W_1 = \begin{bmatrix} 0.5 & -0.3 & 0.2 \\ -0.7 & 0.6 & 0.1 \\ 0.3 & -0.2 & 0.8 \end{bmatrix}, \qquad b_h = \begin{bmatrix} 0.1 \\ -0.4 \\ 0.2 \end{bmatrix}$$

$$W_2 = \begin{bmatrix} 0.6 & -0.5 \\ 0.9 & -0.3 \\ -0.4 & 0.7 \end{bmatrix}, \qquad b_y = \begin{bmatrix} 0.2 \\ -0.1 \end{bmatrix}$$

Input Values:

$$X = (0.5, -0.2, 0.9) \tag{4}$$

Compute the outputs y_1, y_2 of the neural network.

Question 5

Neural Network Architecture:

- **Input layer**: 2 nodes (x_1, x_2)
- **Hidden layer**: 4 nodes (h_1, h_2, h_3, h_4)
- **Output layer**: 1 node (y)

Each node uses the sigmoid activation function.
Weights and Biases:

$$W_1 = \begin{bmatrix} 0.2 & -0.4 & 0.5 & 0.1 \\ 0.6 & -0.3 & 0.8 & -0.2 \end{bmatrix}, \qquad b_h = \begin{bmatrix} -0.1 \\ 0.3 \\ 0.2 \\ 0.4 \end{bmatrix}$$

$$W_2 = \begin{bmatrix} 0.7 \\ -0.5 \\ 0.9 \\ -0.6 \end{bmatrix}, \qquad b_y = 0.3$$

Input Values:

$$X = (0.4, 0.7) \tag{5}$$

Compute the output y of the neural network.

Question 6

Neural Network Architecture:

- **Input layer**: 4 nodes (x_1, x_2, x_3, x_4)
- **Hidden layer**: 2 nodes (h_1, h_2)
- **Output layer**: 1 node (y)

Each node uses the sigmoid activation function.
Weights and Biases:

$$W_1 = \begin{bmatrix} 0.3 & -0.5 \\ 0.7 & 0.2 \\ -0.6 & 0.8 \\ 0.1 & -0.3 \end{bmatrix}, \qquad b_h = \begin{bmatrix} 0.4 \\ -0.2 \end{bmatrix}$$

$$W_2 = \begin{bmatrix} 0.6 \\ -0.7 \end{bmatrix}, \qquad b_y = 0.1$$

Input Values:

$$X = (0.2, -0.4, 0.5, 0.3) \tag{6}$$

Compute the output y of the neural network.

Question 7

Neural Network Architecture:

- **Input layer**: 3 nodes (x_1, x_2, x_3)
- **Hidden layer**: 3 nodes (h_1, h_2, h_3)
- **Output layer**: 2 nodes (y_1, y_2)

Each node uses the sigmoid activation function.
Weights and Biases:

$$W_1 = \begin{bmatrix} 0.1 & -0.4 & 0.3 \\ 0.7 & -0.2 & 0.5 \\ -0.6 & 0.8 & 0.2 \end{bmatrix}, \qquad b_h = \begin{bmatrix} 0.3 \\ -0.1 \\ 0.4 \end{bmatrix}$$

$$W_2 = \begin{bmatrix} 0.5 & -0.7 \\ 0.8 & 0.2 \\ -0.4 & 0.9 \end{bmatrix}, \qquad b_y = \begin{bmatrix} 0.2 \\ -0.3 \end{bmatrix}$$

Input Values:

$$X = (0.6, -0.1, 0.4) \tag{7}$$

Compute the outputs y_1, y_2 of the neural network.

Question 8

Neural Network Architecture:

- **Input layer**: 4 nodes (x_1, x_2, x_3, x_4)
- **Hidden layer**: 2 nodes (h_1, h_2)
- **Output layer**: 3 nodes (y_1, y_2, y_3)

Each node uses the sigmoid activation function:

$$\sigma(z) = \frac{1}{1 + e^{-z}} \tag{8}$$

Weights and Biases:

$$W_1 = \begin{bmatrix} 0.2 & -0.4 \\ -0.3 & 0.7 \\ 0.5 & -0.2 \\ 0.1 & 0.6 \end{bmatrix}, \qquad b_h = \begin{bmatrix} 0.1 \\ -0.2 \end{bmatrix}$$

$$W_2 = \begin{bmatrix} 0.8 & -0.5 & 0.3 \\ -0.6 & 0.9 & -0.4 \end{bmatrix}, \qquad b_y = \begin{bmatrix} 0.3 \\ -0.1 \\ 0.5 \end{bmatrix}$$

Input Values:

$$X = (0.4, -0.7, 0.2, 0.9) \tag{9}$$

Compute the outputs y_1, y_2, y_3 of the neural network.