Neural Network Assignment

# Part 1: Understanding Neural Networks

1. What is a Neural Network?

In your own words, describe what a neural network is and how it is used in machine learning.

2. What are Neurons in Neural Networks?

Explain what neurons are in the context of neural networks and how they are used to process information.

3. What is an Activation Function?

Define what an activation function is and explain why it is important in a neural network.

4. What is Backpropagation?

Describe the backpropagation process in neural networks and explain why it is used for training models.

5. What are Layers in Neural Networks?

Discuss the different types of layers in a neural network (input, hidden, and output) and their purpose.

6. What is the Role of Weights and Biases in Neural Networks?

Explain what weights and biases are, and how they affect the output of a neural network.

7. What is Overfitting in Neural Networks?

Define overfitting in the context of neural networks and explain how it can be prevented.

# Part 2: Activation Functions

Task:

Choose an activation function that was not explained in class (examples: Leaky ReLU, ELU, Swish, etc.). Write a detailed explanation of the function including the following:

1. Mathematical Formula:

Provide the formula for the activation function.

2. Behavior of the Activation Function:

Describe how the function behaves, i.e., how it transforms input values to output values. Include any specific characteristics like non-linearity, thresholding, etc.

3. Where and Why It's Used:

Explain why this activation function is useful and where it can be applied in a neural network architecture. For example, when is it better than other activation functions like Sigmoid or Tanh?

4. Advantages and Disadvantages:

Discuss the advantages and disadvantages of this activation function compared to others.

5. Real-World Application:

Provide an example of how this activation function might be used in a real-world machine learning problem.

ANSWERS TO THE QUESTION

Odukoya Samuel

**Part 1: Understanding Neural Networks**

#### 1. What is a Neural Network?

A neural network is a computational model inspired by the structure and function of the human brain. It consists of layers of interconnected nodes called neurons which process data by passing it through weighted connections.

In simple word Neural networks are used in machine learning for tasks such as classification, regression, and pattern recognition by learning from data and improving their performance over time.

#### 2. What are Neurons in Neural Networks?

Neurons are the fundamental units of a neural network, similar to biological neurons. Each neuron receives input data, processes it using a weighted sum and an activation function, and passes the result to the next layer of neurons. They are responsible for transforming inputs into meaningful outputs through multiple layers of computation.

#### 3. What is an Activation Function?

An activation function is a mathematical function applied to the output of a neuron to introduce non-linearity into the model. This non linearity allows neural networks to learn complex patterns and relationships in data. Without activation functions, the network would behave like a linear model and would not be able to solve complex problems.

#### 4. What is Backpropagation?

Backpropagation is an algorithm used for training neural networks. It calculates the error between the predicted and actual outputs and propagates this error backward through the network. By adjusting the weights and biases of the neurons based on the error, backpropagation optimizes the network to improve its accuracy. It typically involves two phases:

* **Forward Pass:** Computes the output and loss.
* **Backward Pass:** Updates weights using gradients calculated via the chain rule.

#### 5. What are Layers in Neural Networks?

Neural networks consist of three main types of layers:

* **Input Layer:** The first layer that receives raw data.
* **Hidden Layers:** Intermediate layers where computations are performed to extract features and patterns. These can be one or more layers depending on the complexity of the network.
* **Output Layer:** The final layer that provides the result, such as a classification or regression output. Each layer transforms the data to move it closer to the desired output.

#### 6. What is the Role of Weights and Biases in Neural Networks?

* **Weights:** These are the parameters that define the importance of each input to a neuron. They are multiplied by the input values to control the influence of each feature on the output.
* **Biases:** These are additional parameters that allow the network to shift the activation function, helping it fit the data better. Together, weights and biases determine how a neural network processes inputs to generate outputs.

#### 7. What is Overfitting in Neural Networks?

Overfitting occurs when a neural network learns the training data too well, including noise and irrelevant patterns, resulting in poor generalization to new data. It can be prevented by:

* Using regularization techniques like L1 or L2.
* Employing dropout to randomly deactivate neurons during training.
* Using more data or data augmentation.
* Simplifying the network by reducing the number of parameters.

### Part 2: Activation Functions

#### Chosen Activation Function: Swish

#### 1. Mathematical Formula:

The Swish activation function is defined as:

where is the sigmoid function:

#### 2. Behavior of the Activation Function:

* Swish is a smooth, non-linear activation function.
* It is similar to the ReLU function for large positive inputs but introduces a non-monotonic curve for negative inputs, allowing small negative values to propagate.
* For inputs , the function behaves like ReLU with a small enhancement.
* For inputs , the output is slightly negative, which helps in preventing neurons from dying (a problem in ReLU).

#### 3. Where and Why It's Used:

* Swish is useful in deep neural networks where gradient flow and non-linearity are crucial.
* It is particularly effective in architectures requiring high accuracy, such as image recognition and natural language processing.
* Compared to Sigmoid and Tanh, Swish avoids vanishing gradients for large input values.
* It is often better than ReLU in terms of model accuracy because it allows small negative gradients to flow through.

#### 4. Advantages and Disadvantages:

* **Advantages:**
  + Smooth and non-monotonic, leading to better gradient flow.
  + Prevents the "dying neuron" problem common in ReLU.
  + Can improve model performance in certain tasks compared to traditional activation functions.
* **Disadvantages:**
  + Computationally more expensive than ReLU due to the sigmoid component.
  + May not always outperform simpler activation functions in smaller or less complex models.

#### 5. Real-World Application:

Swish has been successfully used in state-of-the-art architectures like EfficientNet for image classification. For instance:

* In **medical imaging**, Swish can enhance model accuracy when analyzing X-rays or MRIs by improving feature extraction and gradient flow in deep networks.

Odukoya Samuel