Name: MD. Mainul Hasan Plabon

ID: 20-43451-1

Section: C

Report on Activation function

In a neural network, an activation function is a mathematical function that determine the output of a neuron based on the input in receives. The activation function is a key component of the artificial neural network, as it introduces non-linearity into the model, allowing it to learn more complex patterns and relationship in the data. I'll discuss six different activation functions, their mathematical formulas, and also advantages and disadvantages.

1. Step function

All inputs greater than or equal to zero are mapped to one by the step function, which is a straightforward activation function, while all inputs less than zero are mapped to zero.

Formula:

If x >= 0, then f(x) = 1.

f(x) = 0, if x < 0

Advantages:

- Simplicity
- Discrete output
- Robustness

Disadvantages:

- Non-differentiability
- Limited expressivity
- Gradient vanishing

2. Sigmoid Function:

A well-liked activation function is the sigmoid function, which converts any real-valued input to a number between 0 and 1.

Formula:

1 / (1 + e(-x))f(x)

Advantages:

- Sigmoid function outputs values in the range of 0 to 1, which makes it suitable for binary classification problems.
- It is a standard function and its properties are well known and understood, making it easy to use.

Disadvantages:

- It is expensive compared to other activation function.
- This function is not suited for regression problems.

3. Tanh Function:

Similar to the sigmoid function, the tanh function converts input into a number between -1 and 1.

Formula:

f(x) is calculated as (ex - e(-x)) / (ex + e(-x)).

Advantages:

- It can introduce non-linearity into the neural network, enabling it to learn complex patterns in the data.
- Tanh is symmetric around the origin, which means it can model negative inputs as well as positive inputs.

Disadvantages:

- The output of the tanh function is not centered around zero, which can make optimization more challenging.

4. Relu Function:

A well-liked activation function, the Rectified Linear Unit (ReLU) function transfers every negative input to zero and any positive input to itself.

Formula:

max(f(x) (0, x)

Advantages:

- Sparsity
- Non-linearity
- Computational efficiency

Disadvantages:

- Unbounded output
- Not suitable for negative inputs

5. PReLU Function:

The slope of the negative half of the function can be learned during training using the Parametric Rectified Linear Unit (PReLU) function, which is a version of the ReLU function.

Formula:

f(x) = max (0, x) + alpha * min (0, x)

Advantages:

- Non-linearity
- Avoids dead neurons
- Better performance

Disadvantages:

- Overfitting
- Computational cost
- Model complexity

6. EReLU Function:

The Exponential Linear Unit (ELU) function, a variant of the ReLU function, uses the exponential function to smooth out the negative portion of the function.

Formula:

If x > 0 then f(x) = alpha * (exp(x) - 1), otherwise f(x) = x.

Advantages:

- Avoids the ‘dying ReLU’ problems
- Better generalization
- Parameterized

Disadvantages:

- High computational cost
- Sensitive to initialization
- Limited resources.

The activation functions that are applied to the outputs of each layer of a neural network transform them into a set of values better suited for further processing. ReLU (Rectified Linear Unit) is a popular activation function for deep neural networks, as it is computationally efficient and helps speed up the convergence of the network during training. Different activation functions have unique advantages and disadvantages, making it important to choose the optimum activation function for a given task.

For binary classification tasks, sigmoid and tanh activation functions are used, and ReLU is used in the network's hidden layers. Softmax is employed for multiclass classification jobs, however its outputs are constrained to the ranges of 0 and 1 and -1 and 1 correspondingly.

ReLU activation function improves performance of deep neural networks by returning 0 for negative inputs and the input itself for positive inputs. By returning 0 for negative inputs and the input itself for positive inputs, the ReLU activation function brings non-linearity into the network and enables it to learn more complex functions.

PReLU and ELU have been found to enhance neural networks, with PReLU adding a learnable parameter and ELU providing an exponential function for negative inputs.