AMERICAN INTERNATIONAL UNIVERSITY-BANGLADESH

408/1, Kuratoli, Khilkhet, Dhaka 1229, Bangladesh

Report Title: Activation function	
Report No: 1	Date of Submission: 11-03-2023
Course Title: Computer vision and pattern recognition	
Course Code:	Section: A
Semester: 2022-2023	Course Teacher: DR. DEBAJYOTI KARMAKER

No	Name	ID	Program	Signature
1	TM Ragib Shahrier	19-39713-1	BSc CSE	Ragib

Faculty use only		
FACULTYCOMMENTS	**Marks Obtained**	
	Total Marks	

Contents

What is Activation function? 4
Some functions are: 4
Step function 4
Formula 4
Advantages 4
Disadvantages 4
Sigmoid Function 4
Formula 4
Advantages 4
Disadvantages 5
Tanh Function 5
Formula 5
Advantages 5
Disadvantages 5
Relu Function 5
Formula 5
Advantages 5
Disadvantages: 5
PReLU Function 6
Formula 6
Advantages 6
Disadvantages 6
EReLU Function 6
Formula 6
Advantages 6
Disadvantages 6
Discussion 7

What is Activation function?

In artificial neural networks, an activation function is a function that outputs a smaller value for tiny inputs and a higher value if its inputs are greater than a threshold. The activation function "fires" if the inputs are big enough; otherwise, nothing happens.

Some functions are:

- Step
- Sigmoid
- Tanh
- Relu
- Elu
- Selu

Step function

A step function, sometimes known as a staircase function, is a piecewise constant function with a limited number of parts, according to the definition given in mathematics. In other words, one can think of a function on the real numbers as a finite linear combination of indicator functions of given intervals.

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

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.

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.

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

PReLU Function

An activation function known as a Parametric Rectified Linear Unit (PReLU) is a traditional rectified unit with a slope for negative values.

Formula

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

Advantages

- Non-linearity
- Avoids dead neurons
- Better performance

Disadvantages

- Overfitting
- Computational cost
- Model complexity

EReLU Function

With no additional parameters and no overfitting risk, EReLU enhances model fitting. Moreover, by utilizing EReLU and parametric ReLU, we propose Elastic Parametric Rectified Linear Unit (EPReLU) (PReLU). EPReLU has the ability to enhance network performance even more.

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.

Discussion

A neural network's outputs from each layer are transformed into a set of values more suited for additional processing by activation functions that are applied to them. ReLU (Rectified Linear Unit) is a well-liked activation function for deep neural networks since it is computationally effective and aids in accelerating the network's convergence during training. It's crucial to select the best activation function for a given activity because different activation functions each have their own pros and drawbacks.

The network's hidden layers employ ReLU while the sigmoid and tanh activation functions are used for binary classification tasks. Although Softmax is used for multiclass classification tasks, its outputs are restricted to the ranges of 0 and 1 and accordingly -1 and 1. By returning 0 for negative inputs and the input itself for positive inputs, ReLU activation function enhances deep neural network performance. The ReLU activation function adds nonlinearity to the network and lets it to learn more complex functions by returning 0 for negative inputs and the input itself for positive inputs.

The addition of a learnable parameter by PReLU and the provision of an exponential function for negative inputs by ELU have both been found to improve neural networks.