Name: Md Shahadat Hossen

ID: 20-43083-1

Sub: Computer vision and pattern recognition

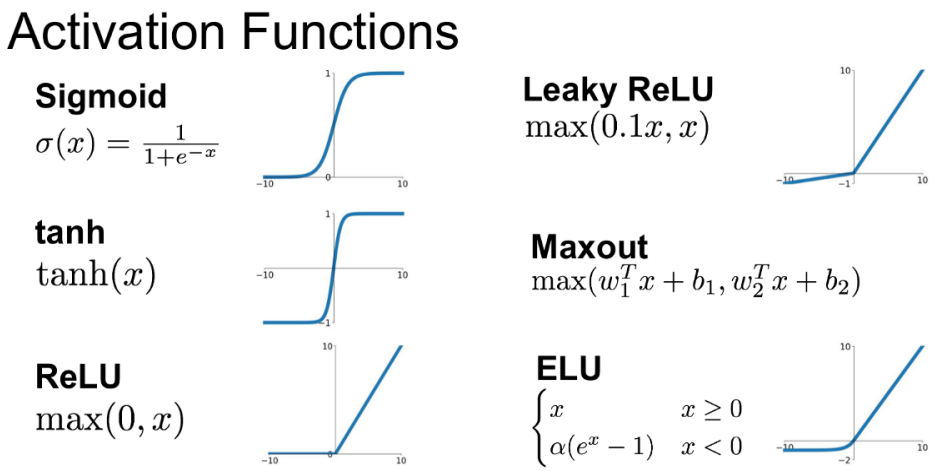
Sec: A

REPORT ON ACTIVATION FUNCTION

What is Activation Function:

Activation functions are the most crucial part of any neural network in **deep learning**. In deep learning, very complicated tasks are **image classification**, **language transformation**, **object detection**, etc which are needed to address with the help of neural networks and activation function. This means that the prediction process uses simpler mathematical operations to determine if a neuron's input to the network is significant. Activation functions are the **gates of neural networks** that decide how much output to pass on.

Common activation functions include **Linear**, **Sigmoid**, **Tanh**, and **ReLU** but there are many others:



1.Step function:

The step function, also known as the Heaviside step function, is a simple activation. Step Function is one of the simplest kind of activation functions. In this, we consider a threshold value and if the value of net input say y is greater than the threshold then the neuron is activated.

function commonly used in neural networks. It is defined mathematically as follows:

f(x) = 0 when x= 0

1 when x >= 0

The step function has a few advantages and disadvantages:

Advantages:

- Great for binary classification
- The step function is simple to implement and can be computed quickly.
- It is a binary function, which makes it useful for binary classification problems.
- It is easy to interpret and understand, which makes it useful for educational purposes.

Disadvantages:

- The slope of the step function is zero. This makes the step function less useful during backward propagation, where the gradient of the activation function is sent for error computation to improve and optimize the result.
- The step function is not differentiable at x=0, which can cause problems when training neural networks using gradient descent algorithms.
- Not available for multiclass classification.

2. Sigmoid function:

The sigmoid function is a mathematical function that maps any input to a value between 0

and 1. The most commonly used sigmoid function is the logistic function, which is defined

as:

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

where x is the input to the function

The sigmoid unit of the neural network. If a neuron's activation function is a sigmoid function, this ensures that the output of that unit is always between 0 and 1. Since the sigmoid function is a nonlinear function, the output of this unit will be a nonlinear function of the weighted sum of the inputs.

Advantage:

- It is a smooth function and continuously differentiable.
- Sigmoid functions are bounded, meaning that their outputs are always between 0 and 1, which can be useful in certain applications like probability calculations.
- 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:

The main advantage of the tanh function is that it aids the backpropagation process by

producing a zero-centered output. The tanh function is primarily used in recurrent neural

networks for natural language processing and speech recognition tasks. The mathematical

formula for the tanh function is follows:

f(x) = (e^x – e^(-x)) / (e^x + e^(-x))

Most tanh functions are commonly used in the hidden layers of neural networks. This is

because their values range from -1 to 1. This is why the hidden layer means 0 or very

close to 0. We center the data by bringing the mean closer to 0, allowing for further

training.

Advantages:

• It is continuous and differentiable at all points.

• It basically solves our problem of values all being of the same sign.

• The function as you can see is non-linear so we can easily backpropagate the errors.

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

• Vanishing gradient problem.

• The gradients are low.

• Computationally expensive function

4. Relu function:

The ReLU (Rectified Linear Unit) function is a commonly used activation function in neural

networks. It is defined as:

f(x) = max(0, x) where x is the input to the function

Advantages:

It has been shown to work well in many types of neural networks, including deep neural networks.

- Sparsity
- Non-linearity
- Computational efficiency

Disadvantages:

The ReLU function is not symmetric, which can make it difficult to use in certain

types of neural networks.

- Unbounded output
- Not suitable for negative inputs

5.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

Disadvantages

- Overfitting
- Computational cost
- Model complexity

6.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.