Name: Khondoker Hozaifa Ibne Shafiq

Id: 20-42883-1

Course: Computer Vision and Pattern Recognition

Sec: C

**Report on Activation Function**

An activation function is a mathematical function applied to a neuron to feed the output into artificial neural networks and machine learning nonlinearities in the output of the network based on information from other neurons. In a network, the activation function helps determine if it is a neuron fire. This study uses six different activation functions. Mathematics basics, strengths, and weaknesses.

1. **Step function:**

The Step function, also called the Heaviside Step function, is a simple activation. It is defined mathematically as follows:

0, 𝑖𝑓 𝑥 < 0

(𝑥) = &1, 𝑖𝑓 𝑥 ≥ 0

Advantages:

• Step functions are easy to implement and quick to compute.  
• This is a binomial function and is useful for binary classification problems.  
• Easy to interpret and understand, useful in classroom purposes.

Disadvantages:

• Problems can arise because the step function cannot be differentiated at x=0.  
Training a neural network using the gradient descent algorithm.  
• Not suitable for regression problems or other tasks that require output.  
Must be constant.  
• Step functions can suffer from the problem of vanishing gradients.  
The gradient of the function is too small to be used for training the network.

1. **Sigmoid function:**

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

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

as:

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

where x is the input to the function.

Advantages:

• The sigmoid function is tractable and easy to differentiate.  
Useful for optimization algorithms such as gradient descent.  
• Often used in neural networks to map neuron outputs probability distribution.  
• The sigmoid function is finite. That is, its output is always in the middle.  
0 and 1. Useful in certain applications such as probability calculation.

Disadvantages:

• The sigmoid function tends to saturate at large values.  
The output of the function on x will be very close to 1,  
Ability to keep learning.  
• The slope of the sigmoid function becomes very small for large values ​​of x,  
This can slow down the convergence of the optimization algorithm.  
• The sigmoid function is not symmetric about zero.  
Introduce a bias in the output of the model.

1. **Tanh function:**

A mathematical function called the hyperbolic tangent function (tanh) converts input values into

output values between -1 and 1. The equation is as follows:

tanh(x) = (e^x - e^(-x)) / (e^x + e^(-x))

where e is the mathematical constant approximately equal to 2.71828, and x is the input value.

Advantages:

• Like the sigmoid function, tanh is a smooth function that is easily differentiated, making it useful for training neural networks using backpropagation.  
• This is a zero-centered function. That is, its output is centered around zero.  
This helps prevent vanishing gradients during the training of deep neural networks.  
• Tanh is bounded between -1 and 1, making it useful for normalizing data with a wide range of values.

Disadvantages:

• Similar to the sigmoid function, Tanh can suffer from the vanishing gradient problem, which can make it difficult to train deep neural networks.  
• Tanh is not monotonic. That is, the derivative is not always positive or negative. This can make optimization difficult for some optimization techniques.  
• The output of tanh is not sparse. This means that it may be less efficient than other activation functions in terms of storage and computational requirements.  
Overall, the Tanh function is a valuable activation function in neural networks, offering advantages over other functions such as the sigmoid function. However, it has significant limitations that must be considered when applying it in various situations.

1. **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:

• Like the sigmoid function, tanh is a smooth function that is easily differentiated, making it useful for training neural networks using backpropagation.  
• This is a zero-centered function. That is, its output is centered at 0.  
This helps prevent vanishing gradients during the training of deep neural networks.  
• Tanh is bounded between -1 and 1, which is useful for normalizing data with a wide range of values.

Disadvantages:

• ReLU functions can suffer from the "dying ReLU" problem, where some neurons stop producing outputs because their inputs are negative. This makes the network less expressive and can lead to poor performance.  
• The ReLU function is not symmetrical, which can make it difficult to use in certain kinds of neural networks.  
• The ReLU function is not differentiable at x=0. This can cause problems with some optimization algorithms based on gradient information.  
In summary, the ReLU function is a popular activation function because it is computationally efficient and useful in many kinds of neural networks. However, it has certain drawbacks, such as the potential for "dying ReLU" problems and difficulty in optimization techniques.

1. **ELU function:**

The ELU (Exponential Linear Unit) step function is an activation function used in artificial neural networks. It is a variant of the ReLU (Rectified Linear Unit) function and is defined as:

ELU(x) = { x if x >= 0, alpha \* (exp(x) - 1) if x < 0 }where alpha is a hyperparameter that determines the negative saturation value.

Advantages:

• Smoother gradients: ELU function has a continuous derivative, which avoids the zero gradient problem that occurs with other activation functions such as the sigmoid function.  
• Better performance: Empirical evidence suggests that ELU activation functions may improve the performance of deep neural networks compared to other activation functions.  
• Faster convergence: The ELU function has been shown to help the network converge faster than other activation functions.

Disadvantages:

• Compute-intensive: The ELU function requires exponential calculations and can be computationally expensive.  
• Unstable for large negative inputs: The ELU function can become unstable for large negative inputs, which can lead to numerical instability in the network.  
• Less Popular: The ELU function is less well-known than other activation functions, such as the sigmoid and ReLU functions, which can make it more difficult to find resources and support implementations in neural networks.

1. **SELu function:**

The SELu (Scaled Exponential Linear Unit) activation function is a type of activation function used in neural networks. Unlike other activation functions, such as ReLU, which have a hard cutoff at zero, SELu has a smooth, continuous curve.

The formula for the SELu activation function is:

f(x) = 1.0507 \* (e^x - 1), x < 0

f(x) = x, x >= 0

Advantages:

・The SELu activation function is a self-normalization function.  
The output of each neuron has a mean of zero and a variance of one, regardless of the input distribution. This helps reduce gradient vanishing/exploding issues.  
This can happen in deep neural networks.  
• SELu has proven to be stronger than other activation features such as ReLU  
and tanh are used for image classification and  
voice recognition.

Disadvantages:

• SELu activation functions are more computationally intensive than other activations.  
It works like ReLU because it involves exponentiation.  
• SELu may not work well with all data types. Although it has been shown to work well,  
May not be the best choice for all types of data for many deep-learning tasks  
Some experimentation may be required to determine the optimal activation function for .