Computer Vision and Pattern Recognition

Assignment-2

Topic: Activation Function

Submitted By:

ISLAM MD SIRAJUL

ID:20-42564-1

SEC: B

Submitted To:

DR. DEBAJYOTI KARMAKER

American International University- Bangladesh

Activation functions play a crucial role in neural network models. They introduce non-linearity into the network, allowing it to learn complex relationships between input and output. Numpy is a popular numerical computation library in Python that provides several activation functions. This report compares the various activation functions available in Numpy, including the Step Activation Function, Sigmoid Activation Function, Tanh Activation Function, ReLU Activation Function, ELU Activation Function, and SELU Activation Function.

Activation Function Comparison:

1. Step Activation Function: The step activation function is a binary function that returns 1 if the input is greater than or equal to zero and 0 otherwise. It is the simplest activation function and can be used for binary classification problems. However, it is not suitable for multi-class classification or regression problems as it only produces two outputs.
2. Sigmoid Activation Function: The sigmoid activation function is a smooth and continuous function that maps any input to a value between 0 and 1. It is widely used in binary classification problems where the output needs to be probabilistic. However, the gradient of the sigmoid function is small for large input values, which can lead to slow convergence during training.
3. Tanh Activation Function: The tanh activation function is similar to the sigmoid function but maps the input to a value between -1 and 1. It is also smooth and continuous, making it suitable for use in neural networks. However, like the sigmoid function, the gradient of the tanh function is small for large input values.
4. ReLU Activation Function: The rectified linear unit (ReLU) activation function is a piecewise linear function that returns the input if it is positive and 0 otherwise. It is currently the most popular activation function and has been shown to work well in many deep learning models. The ReLU function is easy to compute and has a large gradient for positive input values, which leads to faster convergence during training. However, it can suffer from the "dying ReLU" problem, where neurons can become inactive and output zero for all inputs.
5. ELU Activation Function: The exponential linear unit (ELU) activation function is similar to the ReLU function but uses an exponential function for negative input values. This helps to avoid the "dying ReLU" problem by ensuring that all neurons have non-zero output. The ELU function also produces negative output values, which can be useful in some cases. However, it can be computationally more expensive than the ReLU function.
6. SELU Activation Function: The scaled exponential linear unit (SELU) activation function is a variation of the ELU function that uses a scaling parameter to ensure that the output has unit mean and variance. This can lead to better convergence and generalization in deep neural networks. However, the SELU function requires careful initialization of the network weights and biases to ensure that the mean and variance of the input data are preserved during forward propagation.

Numpy provides several activation functions that can be used in neural network models. The choice of activation function depends on the specific problem and the architecture of the network. The ReLU function is currently the most popular activation function due to its simplicity and good performance in many applications. However, the ELU and SELU functions can be useful in certain

situations, especially when dealing with deep neural networks. It is essential to experiment with different activation functions to determine which one works best for a particular problem.