NAME: ANNY SAHA

ID: 20-43417-1

REPORT

Activation functions are a crucial component in the architecture of deep learning models. They introduce non-linearity to the model and allow it to learn complex representations of the data. In this report, we will discuss the properties and characteristics of five popular activation functions - SELU, ReLU, ELU, step, sigmoid, and tanh.

STEP Function:

A common component of neural networks is a unit step activation function. For negative arguments, the result is set to 0 and for positive arguments, to 1. This is how the method works: The range is between (0,1) and the output is binary in character.

SELU (Scaled Exponential Linear Unit):

SELU is a self-normalizing activation function, which means that it maintains a stable distribution of activations throughout the layers of a deep neural network. SELU has been shown to improve the training of deep neural networks and achieve state-of-the-art performance on various tasks. It has the following properties:

1. It is a piecewise linear function with a negative slope for negative inputs.
2. It has an exponential increase for positive inputs.
3. It is computationally efficient and can be easily implemented.

ReLU (Rectified Linear Unit):

ReLU is the most widely used activation function in deep learning models. It is simple and computationally efficient. ReLU has the following properties:

1. It is a piecewise linear function that returns the input for positive values and 0 for negative values.
2. It is computationally efficient and can be easily implemented.
3. It can suffer from the "dying ReLU" problem, where some neurons may become permanently inactive during training.

ELU (Exponential Linear Unit):

ELU is similar to ReLU but has a non-zero output for negative inputs. It has the following properties:

1. It is a smooth function and its derivative is continuous everywhere.
2. It avoids the "dying ReLU" problem by having a non-zero output for negative inputs.
3. It is computationally more expensive than ReLU.

Sigmoid function:

Sigmoid is a popular activation function that is often used in binary classification tasks. It has the following properties:

1. It is a smooth function that maps the input to a value between 0 and 1.
2. It is computationally efficient and can be easily implemented.
3. It suffers from the "vanishing gradient" problem, where the gradient approaches 0 for large inputs.

Tanh function:

Tanh is similar to sigmoid but maps the input to a value between -1 and 1. It has the following properties:

1. It is a smooth function that is symmetric around 0.
2. It is computationally efficient and can be easily implemented.
3. It also suffers from the "vanishing gradient" problem for large inputs.

In conclusion, the choice of activation function depends on the specific task and architecture of the deep learning model. I think SELU and ELU are effective for deep neural networks, while ReLU is simple and efficient for shallower networks. Sigmoid and tanh are often used in binary classification tasks, but can suffer from the vanishing gradient problem.