Exponential Linear Unit (ELU)

The *Exponential Linear Unit (ELU)* is a type of activation function often used in deep learning models. It is designed to address the limitation of the *Rectified Linear Unit (ReLU)* function when dealing with negative values.

The ELU function is defined as:

$$f(x) = x \text{ if } x > 0$$

$$f(x) = \alpha * (\exp(x) - 1) \text{ if } x$$

where α is a hyperparameter that controls the saturation value when (x \leq 0).

- **Advantages of ELU:**

1. ***Smoothness:*** **ELU is a smooth function which helps in continuous gradient flow, especially around zero, unlike ReLU which has a gradient of zero for negative inputs.**

2. ***Robust to Vanishing Gradient:*** **ELU helps mitigate the vanishing gradient problem since it can have non-zero gradients for negative values.**

3. ***Capturing Negative Information:*** **The ELU function captures negative information more effectively than ReLU, leading to improved learning in certain scenarios.**

- **Hyperparameter α:**

The parameter α in the ELU function determines the output of the function for negative inputs. Common values for α are often chosen to be around 1.0. The selection of α affects the behavior of the ELU function and should be tuned based on the specific task and dataset.

- **Implementation:**

In neural network libraries like TensorFlow or PyTorch, the ELU activation function is readily available for use. When implementing neural networks, using ELU as an activation function can be beneficial, especially when the model needs to learn from data with negative values.

Overall, the ELU activation function is a useful alternative to ReLU and can be particularly effective in scenarios where handling negative values is crucial for learning accurate representations in *deep learning* models.

Producer: Elham Jafari

Computer Engineering