Tanh

- **Introduction:**

Tanh, short for hyperbolic tangent, is a commonly used activation function in neural networks. It is a scaled version of the logistic sigmoid function and is popular for its ability to squish input values to the range [-1, 1], allowing for easier training of deep neural networks.

The Tanh activation function is defined as:

Tanh (x) = e^{x} - e^{-x} /
e^{x} + e^{-x}

* Characteristics:

- Range: **Tanh squashes input values to the range [-1, 1], which helps in centering the activations around zero.**

- Symmetry: **Tanh function is symmetric around the origin, which can aid in training deep networks by making the outputs zero-centered.**

-Smoothness: **Tanh is a smooth, continuous function that is differentiable everywhere.**

- Sensitivity to Input: **Tanh is more sensitive to small changes in input compared to the sigmoid function.**

* Advantages:

1. ***Zero Centering:*** **Tanh helps center the output activations around zero, which can help in optimizing the neural network.**

2. ***Gradient Saturation:*** **Tanh mitigates the vanishing gradient problem better than sigmoid by mapping the input values to a wider range.**

* Disadvantages:

1. Vanishing Gradient: **Tanh is still susceptible to the vanishing gradient problem, especially for deeper networks.**

2. Computation Cost: **Tanh function involves exponential operations, which can be computationally expensive compared to functions like ReLU.**

- ## Conclusion:

Tanh activation function is a valuable tool in deep learning due to its ability to squash input values to a range that aids in optimizing neural networks. While it has advantages such as zero-centering and alleviating gradient saturation to some extent, it is important to consider its potential drawbacks like vanishing gradients and computational cost when choosing it for a model architecture.

Producer: Elham Jafari

Computer Engineering