Different Types of Activation Function (ANN)

What is an Activation Function?

An activation function is a mathematical function applied to the output of a neuron. It introduces non-linearity into the model, allowing the network to learn and represent complex patterns in the data. Without this non-linearity feature, a neural network would behave like a linear regression model, no matter how many layers it has.

Why is Non-Linearity Important in Neural Networks?

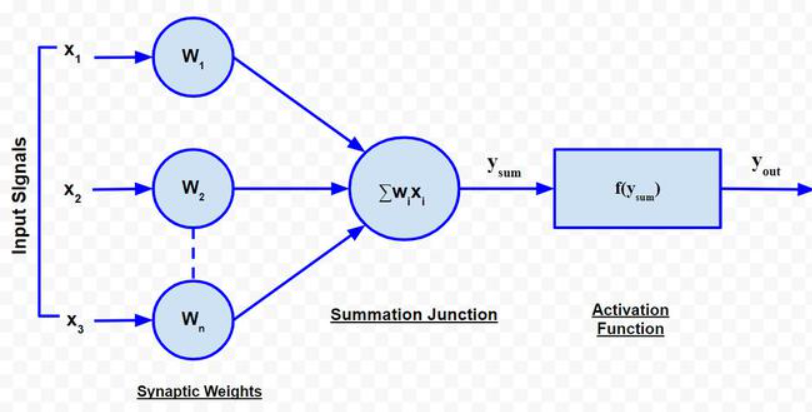
Neural networks consist of neurons that operate using weights, biases, and activation functions.

In the learning process, these weights and biases are updated based on the error produced at the output—a process known as backpropagation. Activation functions enable backpropagation by providing gradients that are essential for updating the weights and biases.

Without non-linearity, even deep networks would be limited to solving only simple, linearly separable problems.

Activation functions empower neural networks to model highly complex data distributions and solve advanced deep learning tasks. Adding non-linear activation functions introduce flexibility and enable the network to learn more complex and abstract patterns from data.

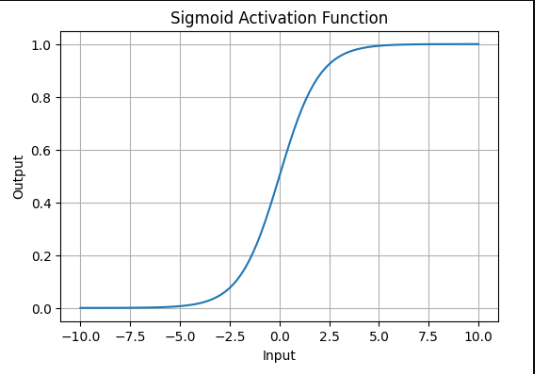
Some common types of activation functions used in Artificial Neural Networks (ANNs) include: Sigmoid, Tanh (Hyperbolic Tangent), ReLU (Rectified Linear Unit), Softmax, Linear Activation Function, Exponential Linear Unit (ELU), and Leaky ReLU, each with different characteristics and use cases depending on the problem at hand.



Key points about each activation function:

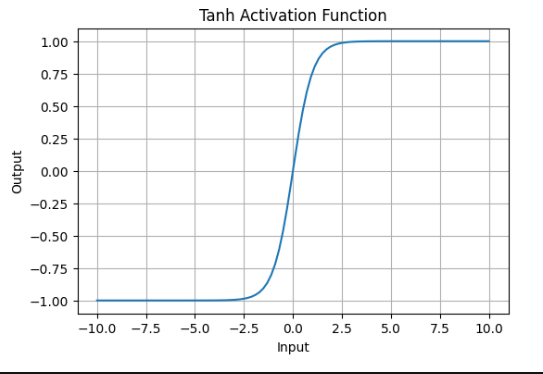
* **Sigmoid:**

Outputs values between 0 and 1, making it useful for binary classification problems where the output represents a probability.



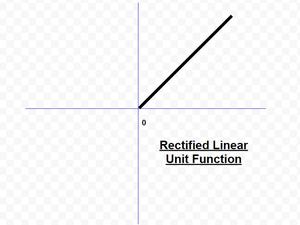
* **Tanh (Hyperbolic Tangent):**

Similar to sigmoid but outputs values between -1 and 1, often used in recurrent neural networks.



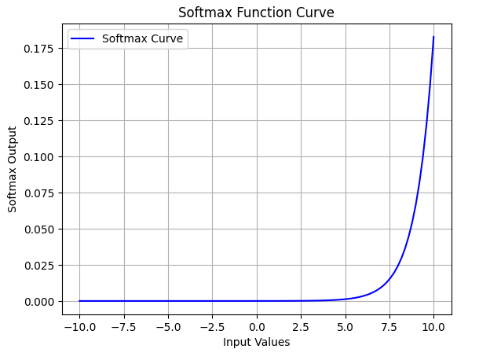
* **ReLU (Rectified Linear Unit):**

A popular choice due to its computational efficiency, where the output is simply the input if positive, otherwise 0.



* **Softmax:**

Generates a probability distribution over multiple classes, commonly used for multi-class classification problems.



* **Linear Activation Function:**

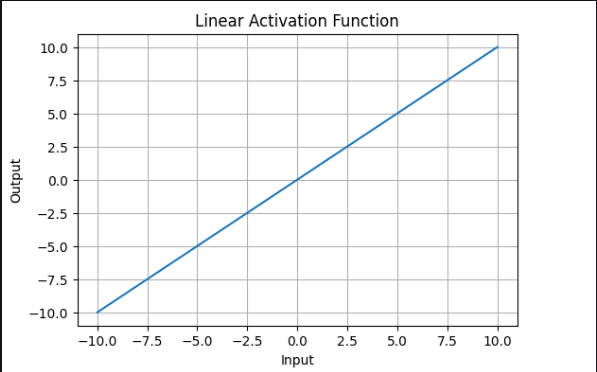
No non-linearity, simply returns the input value, rarely used in deep learning.

**Linear Activation Function**

**Linear Activation Function**resembles straight line define by y=x. No matter how many layers the neural network contains, if they all use linear activation functions, the output is a linear combination of the input.

* The range of the output spans from (−∞ to +∞)(−∞ to +∞).
* **Linear activation function** is used at just one place i.e. output layer.
* Using linear activation across all layers makes the network’s ability to learn complex patterns limited.

Linear activation functions are useful for specific tasks but must be combined with non-linear functions to enhance the neural network’s learning and predictive capabilities.



* **Exponential Linear Unit (ELU):**

Aims to alleviate the "dying ReLU" problem by introducing negative values, potentially improving learning speed.

* **Leaky ReLU:**

A variation of ReLU where a small slope is applied to negative values, preventing "dying neurons".

