

# What is activation function?

An activation function, also known as a transfer function, is a crucial component of a neural network in machine learning. It introduces non-linearity to the network, allowing it to learn complex patterns in data. The activation function operates on the weighted sum of inputs plus a bias, determining the output of a neuron or node in the network.

Common activation functions include:

## 1. Sigmoid Function (Logistic):

$$\sigma(x) = \frac{1}{1+e^{-x}}$$

Output range: (0, 1)

Used in the output layer for binary classification problems.

## 2. Hyperbolic Tangent (tanh):

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

Output range: (-1, 1)

Similar to the sigmoid but with an output range that includes negative values.

## 3. Rectified Linear Unit (ReLU):

$$f(x) = \max(0, x)$$

Output range:  $[0, +\infty)$

Popular for hidden layers due to its simplicity and efficiency in training.

4. **Leaky ReLU:**

$$f(x) = \max(\alpha x, x)$$

Introduces a small slope ( $\alpha$ ) for negative values, preventing dead neurons in ReLU.

5. **Softmax:**

$$\text{Softmax}(x)_i = \frac{e^{x_i}}{\sum_j e^{x_j}}$$

Converts a vector of real numbers into a probability distribution.

The choice of activation function depends on the specific task and the characteristics of the data.

Different activation functions have different properties and affect how well a neural network can learn and generalize from the data.