UNIT – III

NEURAL NETWORKS, DEEP LEARNING AND FOUNDATION MODELS

3.1 Artificial Neurons

Artificial neurons are computational components inspired by the biological neurons that make up the human brain. Just like real neurons receive, process, and transmit information, artificial neurons perform a similar role in artificial neural networks. They receive input data, apply a mathematical calculation — usually based on weights and activation functions — and produce an output that is transmitted to other neurons in the network.

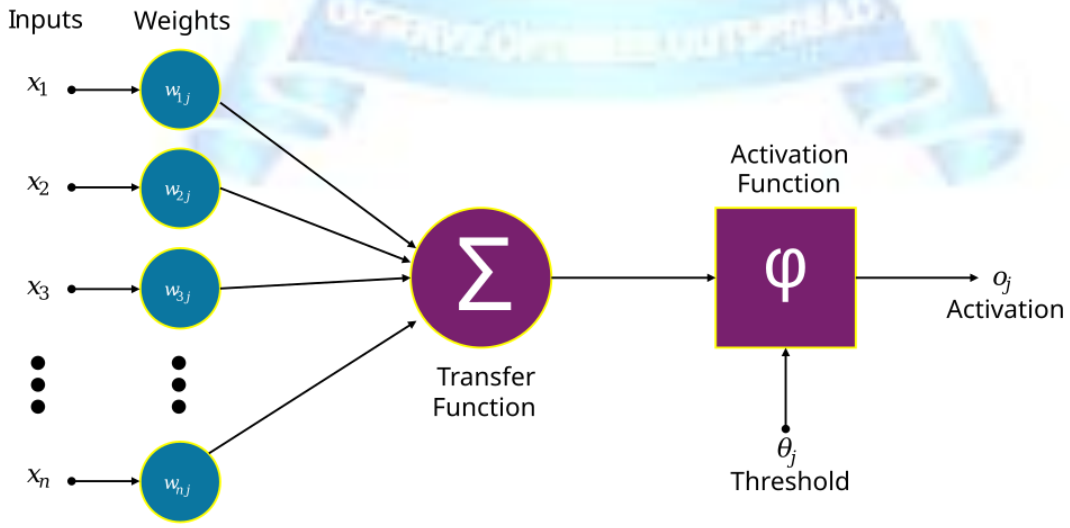
Unlike biological neurons, artificial neurons are not physical entities, but rather mathematical representations implemented in software or hardware. Their simplicity compared to the human brain is compensated by the collective strength of the networks, which can be trained to recognize patterns, predict outcomes, and even make decisions in various contexts.

Definition

A neuron is the basic building block of an artificial neural network, inspired by the human brain. It processes input data, applies weights, and produces an output based on an activation function. An artificial neuron takes in a set of input signals, performs a weighted sum, adds a bias term, and passes the result through an activation function to produce the final output signal.

Structure Of an Artificial Neuron

The structure of an artificial neuron is formed by mathematical components that simulate the basic functions of receiving, processing, and transmitting information. Let's explore the main elements of this structure:



1. Inputs

Inputs represent the data that arrive at the neuron. They can be numbers, variables, or results from other neurons in a previous layer. Each input is associated with a weight, which determines the relative importance of this input in the final calculation.

2. Weights

The weights are adjustable values that help the neuron "learn" during training. When a model is trained, the weights are adjusted to minimize the error in the output, allowing the neural network to adapt to the problem it is solving.

3. Weighted Sum

The neuron combines the inputs multiplied by their respective weights in a weighted sum operation. Mathematically, this is represented as:

$$z = \sum (x_i \cdot w_i) + b$$

Here:

- x_i são as entradas,
- w_i são os pesos,
- b é o **bias**, um valor adicional que ajuda a ajustar o resultado da soma.

4. Activation Function

The output of the above calculation goes through an activation function, which introduces non-linearity into the system. This is crucial, as it allows the neural network to solve complex problems, such as pattern recognition. Some of the most common activation functions include:

- **Sigmoid:** used to constrain the output between 0 and 1.
- **ReLU (Rectified Linear Unit):** activates positive values and sets negative values to 0.
- **Tanh:** compresses values into a range from -1 to 1.

5. Output

The output is the final result of the neuron after the activation function. This output can be sent to other neurons in subsequent layers or used directly to generate a result, depending on the position of the neuron in the network.

This structure allows artificial neurons to function as building blocks in neural networks. When connected in layers, they form powerful systems capable of solving highly complex problems.

Types of Activation Function :

There are different types of activation functions. The most commonly used activation function are listed below:

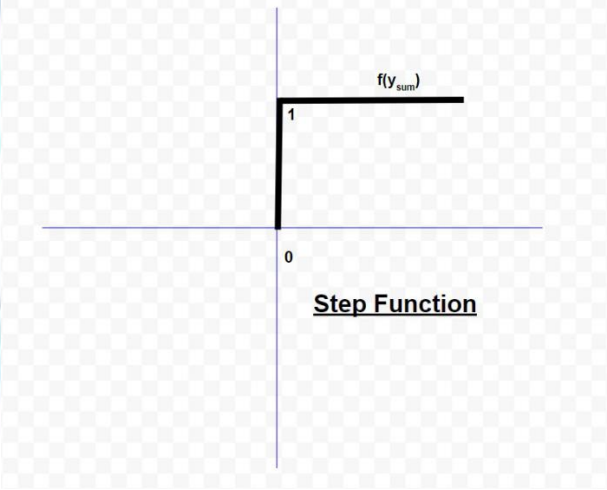
A. Identity Function: Identity function is used as an activation function for the input layer. It is a linear function having the form

$$y_{out} = f(x) = x, \forall x$$

As obvious, the output remains the same as the input.

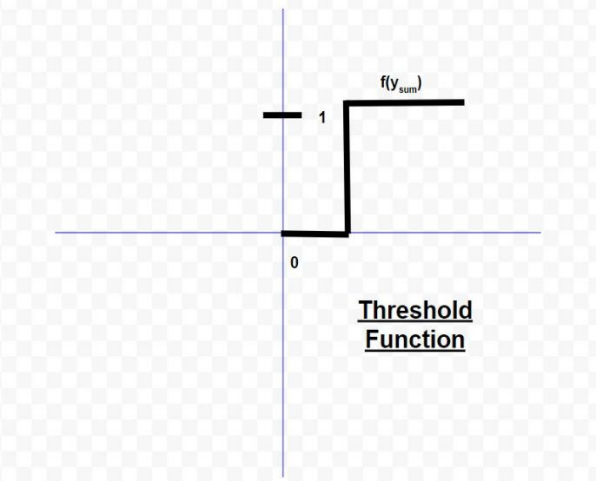
B. Threshold/step Function: It is a commonly used activation function. As depicted in the diagram, it gives 1 as output of the input is either 0 or positive. If the input is negative, it gives 0 as output. Expressing it mathematically,

$$y_{out} = f(y_{sum}) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases}$$



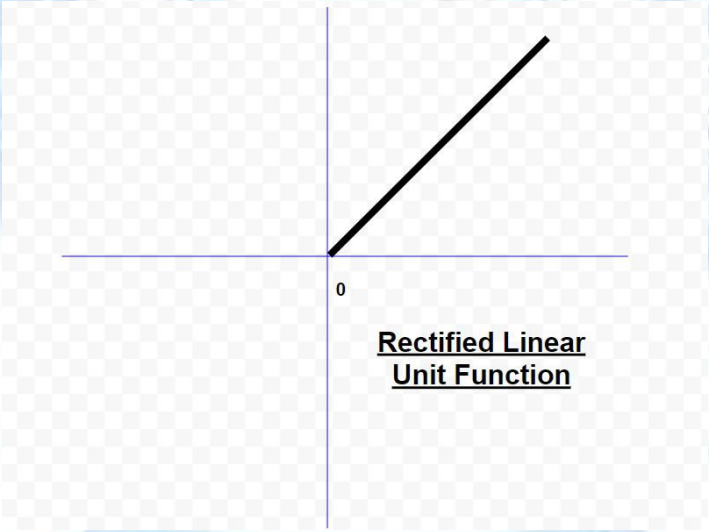
The threshold function is almost like the step function, with the only difference being a fact that θ is used as a threshold value instead of . Expressing mathematically,

$$y_{out} = f(y_{sum}) = \begin{cases} 1, x \geq \theta \\ 0, x < \theta \end{cases}$$



C. ReLU (Rectified Linear Unit) Function: It is the most popularly used activation function in the areas of convolutional neural networks and deep learning. It is of the form:

$$f(x) = \begin{cases} x, x \geq 0 \\ 0, x < 0 \end{cases}$$

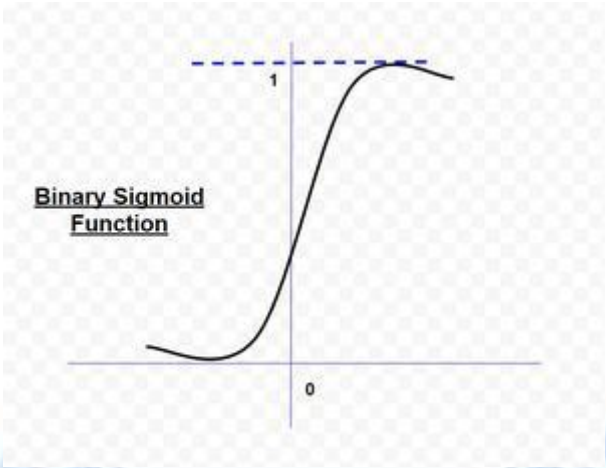


This means that f(x) is zero when x is less than zero and f(x) is equal to x when x is above or equal to zero. This function is differentiable, except at a single point x = 0. In that sense, the derivative of a ReLU is actually a sub-derivative.

D. Sigmoid Function: It is by far the most commonly used activation function in neural networks. The need for sigmoid function stems from the fact that many learning algorithms require the activation function to be differentiable and hence continuous.

There are two types of sigmoid function:

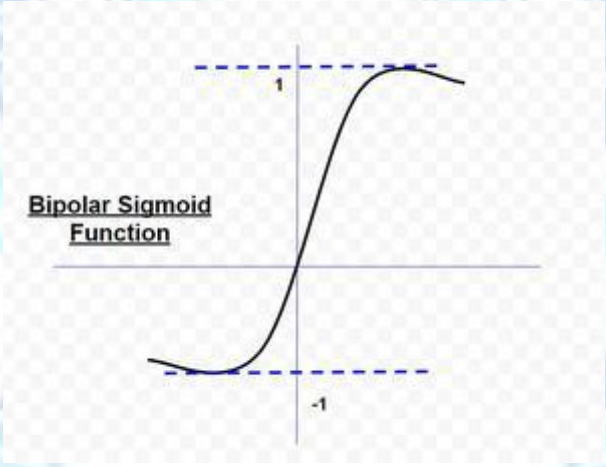
1. Binary Sigmoid Function



A binary sigmoid function is of the form: $y_{out} = f(x) = \frac{1}{1+e^{-kx}}$

, where k = steepness or slope parameter, By varying the value of k, sigmoid function with different slopes can be obtained. It has a range of (0,1). The slope of origin is k/4. As the value of k becomes very large, the sigmoid function becomes a threshold function.

2. Bipolar Sigmoid Function



A bipolar sigmoid function is of the form $y_{out} = f(x) = \frac{1-e^{-kx}}{1+e^{-kx}}$

The range of values of sigmoid functions can be varied depending on the application. However, the range of (-1,+1) is most commonly adopted.

E. Hyperbolic Tangent Function: It is bipolar in nature. It is a widely adopted activation function for a special type of neural network known as Backpropagation Network. The hyperbolic tangent function is of the form

$y_{out} = f(x) \frac{e^{x}-e^{-}x}{e^{x}+e^{-}x}$

This function is similar to the bipolar sigmoid function.

Applications Of Artificial Neurons

Artificial neurons are the basis of numerous innovations transforming various sectors of the economy. Here's how this technology is being applied in the healthcare, finance, and technology sectors:

1. Healthcare: Diagnosis and treatment personalization

- **Medical diagnostics:** artificial neural networks are used to analyze medical exams, such as MRIs and CT scans, identifying anomalies with high precision. For example, algorithms based on artificial neurons assist in the early diagnosis of diseases such as cancer and heart disease.
- **Personalized medicine:** using genetic data and medical histories, it is possible to create treatments tailored to the individual needs of each patient, optimizing outcomes and reducing side effects.
- **Drug development:** models based on artificial neurons accelerate the process of discovering new drugs by identifying potential chemical compounds and simulating their effects.

2. Finance: Predictions and process automation

- **Market forecasting:** artificial neurons are widely used in trading algorithms to predict market trends based on historical and real-time analyses. This allows for more accurate decisions for investors.
- **Risk analysis:** banks and insurance companies deploy neural networks to assess the probability of default or fraud, reducing losses and offering more secure solutions to clients.
- **Automated customer service:** chatbots and virtual assistants powered by neural networks enhance the customer experience by providing quick and accurate responses, even in complex operations.

3. Technology: Innovation and automation

- **Image and voice recognition:** artificial neurons are the foundation of facial recognition systems, such as those used in security, and voice assistants like Alexa, Siri, and Google Assistant.
- **Recommendation systems:** streaming platforms like Netflix and Spotify, and e-commerce sites like Amazon, use neural networks to understand user preferences and suggest relevant content or products.
- **Autonomous cars:** autonomous vehicles rely on neural networks to process data from sensors and cameras in real time, allowing safe decisions in traffic**.**

- **Cybersecurity:** models based on artificial neurons detect anomalous patterns in networks, helping to prevent attacks and digital fraud.

These applications show how artificial neurons are transforming critical areas, bringing efficiency, innovation, and personalization.