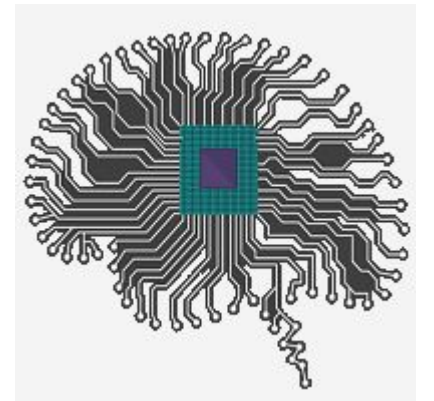


# DEEP LEARNING

UNIT I

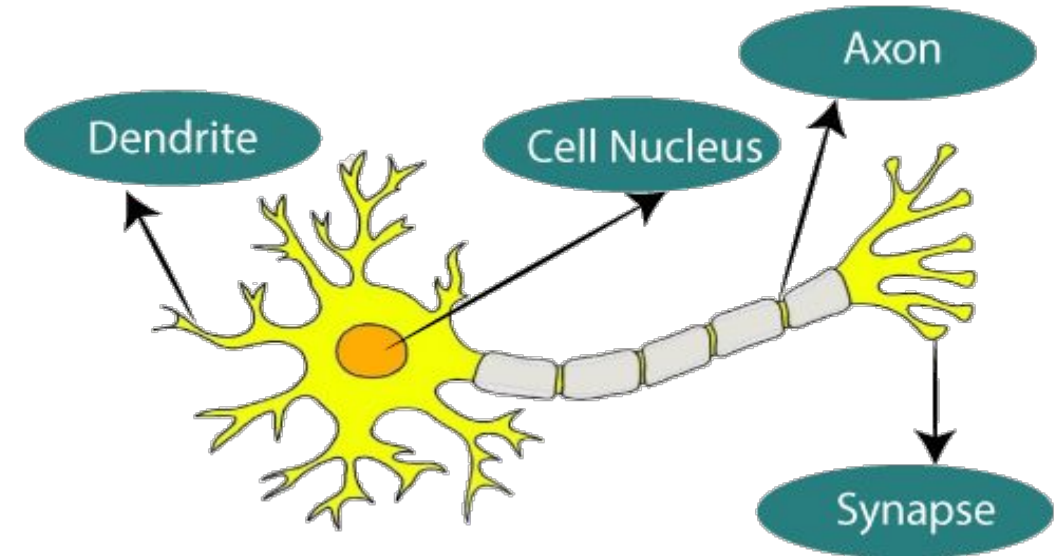
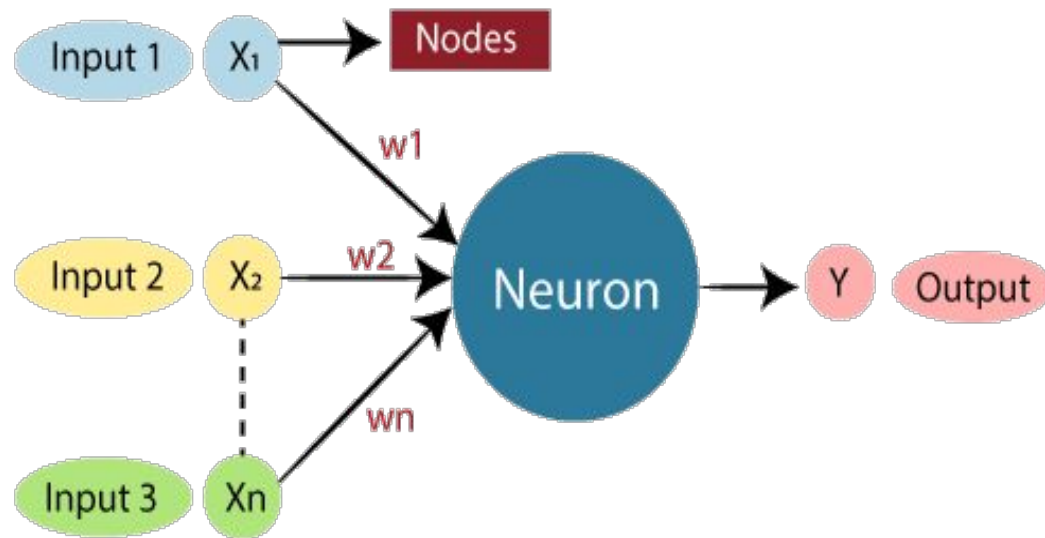
# Basics of Artificial Neural Network

- The term "Artificial neural network" refers to a biologically inspired sub-field of artificial intelligence modeled after the brain.
- An Artificial neural network is usually a computational network based on biological neural networks that construct the structure of the human brain.
- Similar to a human brain has neurons interconnected to each other, artificial neural networks also have neurons that are linked to each other in various layers of the networks.
- These neurons are known as nodes.



# What is Artificial Neural Network?

- The term "**Artificial Neural Network**" is derived from Biological neural networks that develop the structure of a human brain.
- Similar to the human brain that has neurons interconnected to one another, artificial neural networks also have neurons that are interconnected to one another in various layers of the networks.
- These neurons are known as nodes.



## Relationship between Biological neural network and artificial neural network:

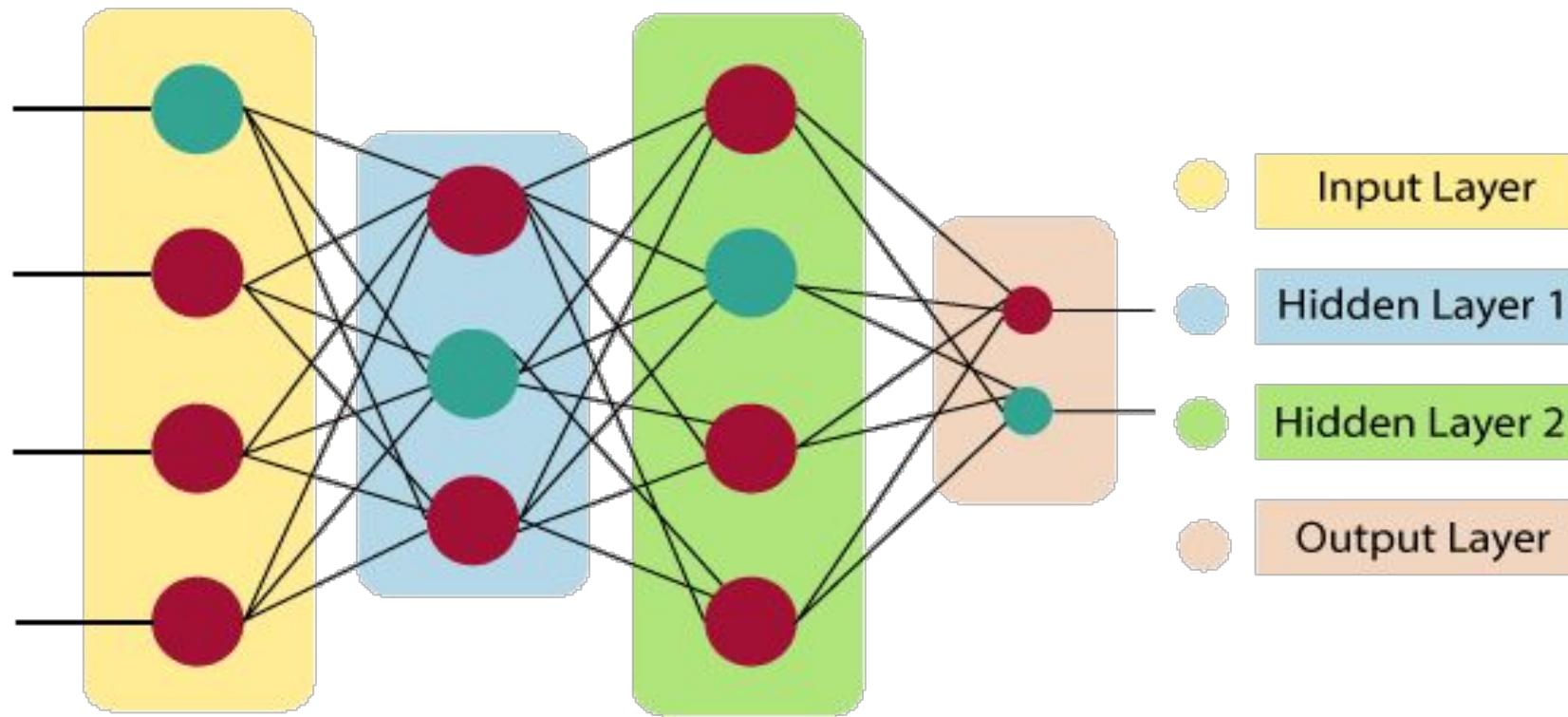
Biological Neural Network	Artificial Neural Network
Dendrites	Inputs
Cell nucleus	Nodes
Synapse	Weights
Axon	Output

- An **Artificial Neural Network** in the field of **Artificial intelligence** where it attempts to mimic the network of neurons makes up a human brain so that computers will have an option to understand things and make decisions in a human-like manner.
- The artificial neural network is designed by programming computers to behave simply like interconnected brain cells.
- There are around 1000 billion neurons in the human brain.
- Each neuron has an association point somewhere in the range of 1,000 and 100,000. In the human brain, data is stored in such a manner as to be distributed, and we can extract more than one piece of this data when necessary from our memory parallelly.
- We can say that the human brain is made up of incredibly amazing parallel processors.

- We can understand the artificial neural network with an example, consider an example of a digital logic gate that takes an input and gives an output.
- "OR" gate, which takes two inputs. If one or both the inputs are "On," then we get "On" in output.
- If both the inputs are "Off," then we get "Off" in output.
- Here the output depends upon input. Our brain does not perform the same task. The outputs to inputs relationship keep changing because of the neurons in our brain, which are "learning."

- The architecture of an artificial neural network:
- To understand the concept of the architecture of an artificial neural network, we have to understand what a neural network consists of.
- In order to define a neural network that consists of a large number of artificial neurons, which are termed units arranged in a sequence of layers.
- Lets us look at various types of layers available in an artificial neural network.

**Artificial Neural Network primarily consists of three layers:**





- **Input Layer:**

- As the name suggests, it accepts inputs in several different formats provided by the programmer.

- **Hidden Layer:**

- The hidden layer presents in-between input and output layers. It performs all the calculations to find hidden features and patterns.

- **Output Layer:**

- The input goes through a series of transformations using the hidden layer, which finally results in output that is conveyed using this layer.
- The artificial neural network takes input and computes the weighted sum of the inputs and includes a bias. This computation is represented in the form of a transfer function.

$$\sum_{i=1}^n W_i * X_i + b$$

- It determines weighted total is passed as an input to an activation function to produce the output.
- Activation functions choose whether a node should fire or not. Only those who are fired make it to the output layer.
- There are distinctive activation functions available that can be applied upon the sort of task we are performing.

# **Advantages of Artificial Neural Network (ANN)**

- **Parallel processing capability:**
- **Storing data on the entire network:**
- **Capability to work with incomplete knowledge:**
- **Having a memory distribution:**
- **Having fault tolerance:**

# **Disadvantages of Artificial Neural Network:**

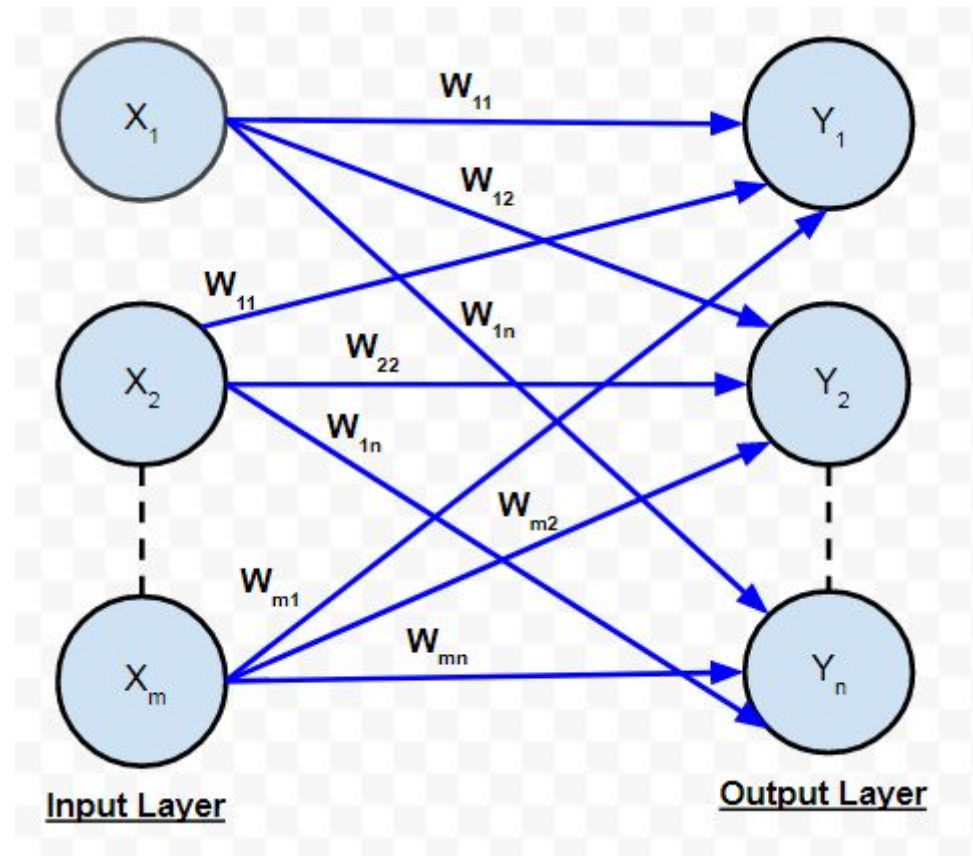
- **Assurance of proper network structure:**
- **Unrecognized behavior of the network:**
- **Hardware dependence:**
- **Difficulty of showing the issue to the network:**
- **The duration of the network is unknown:**

# Architecture and Learning process in neural network

- we first have to understand the architecture of the neural network and then the learning process in ANN.
- ANN is a computational system consisting of many interconnected units called **artificial neurons**.
- The connection between artificial neurons can transmit a signal from one neuron to another. So, there are multiple possibilities for connecting the neurons based on which the **architecture** we are going to adopt for a specific solution.

- Some permutations and combinations are as follows:
  - There may be just two layers of neuron in the network – the input and output layer.
  - There can be one or more intermediate '**hidden**' layers of a neuron.
  - The neurons may be connected with all neurons in the next layer and so on .....

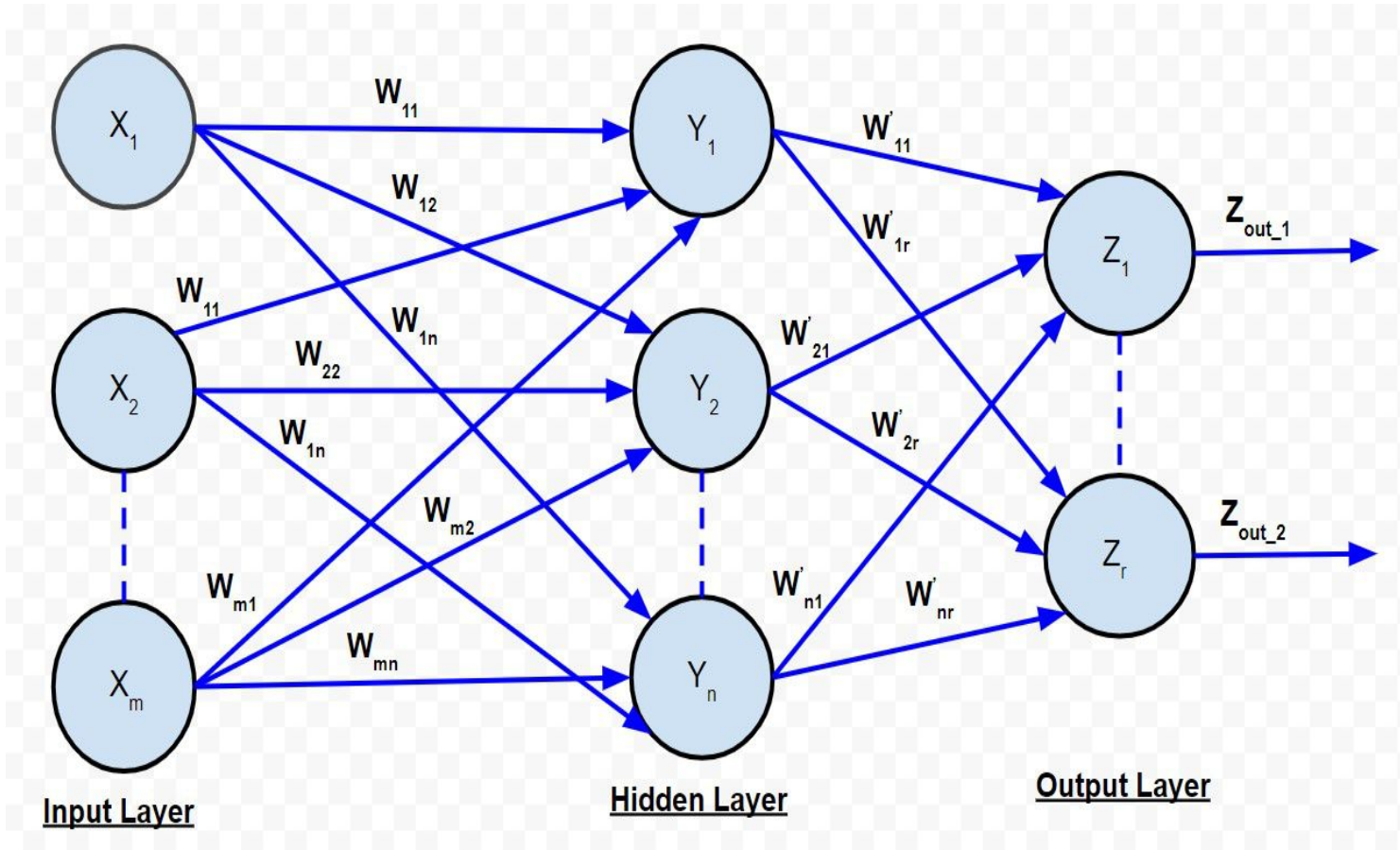
- So let's start talking about the various possible architectures:
- **Single-layer Feed Forward Network:**



- It is the simplest and most basic architecture of ANN's. It consists of only two layers- the input layer and the output layer.
- **The input layer** consists of 'm' input neurons connected to each of the 'n' output neurons.
- The connections carry weights  $w_{11}$  and so on. The input layer of the neurons doesn't conduct any processing – they pass the i/p signals to the o/p neurons.
- The computations are performed in the output layer. So, though it has 2 layers of neurons, only one layer is performing the computation.
- This is the reason why **the network is known as SINGLE** layer. Also, the signals always flow from the input layer to the output layer. Hence, **the network is known as FEED FORWARD**.
- The signal output from each output neuron will depend on the activation function used.



- **Multi-layer Feed Forward Network:**



- The multi-layer feed-forward network is quite similar to the single-layer feed-forward network, except for the fact that there are one or more intermediate layers of neurons between the input and output layer.
- Hence, the **network is termed as multi-layer**. Each of the layers may have a varying number of neurons.
- For example, the one shown in the above diagram has 'm' neurons in the input layer and 'r' neurons in the output layer and there is only one hidden layer with 'n' neurons

# Pattern Recognition

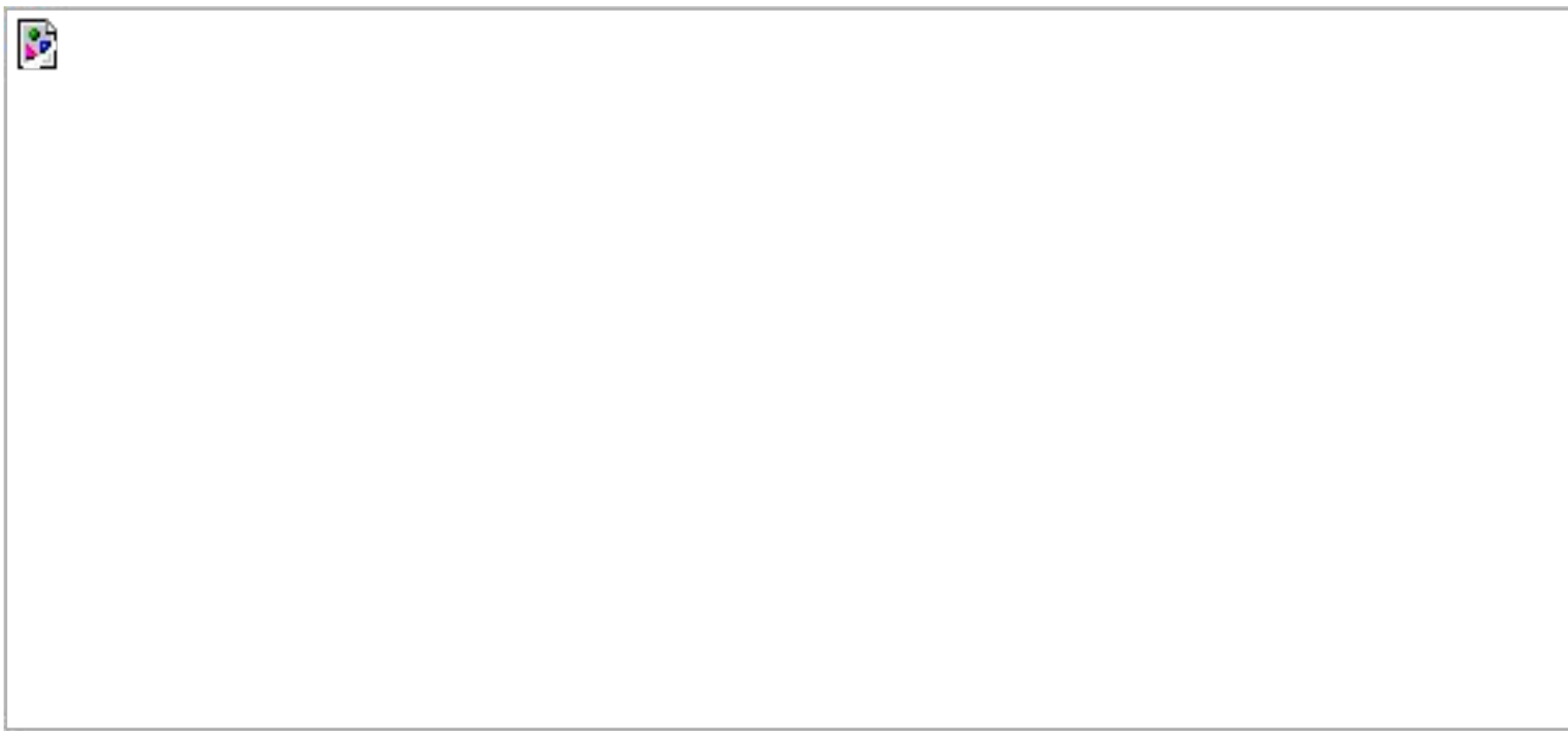
- **Pattern** is everything around in this digital world. A pattern can either be seen physically or it can be observed mathematically by applying algorithms.
- **Example:** The colors on the clothes, speech pattern, etc. In computer science, a pattern is represented using vector feature values.

## What is Pattern Recognition?

**Pattern recognition** is the process of recognizing patterns by using a machine learning algorithm.

- Pattern recognition can be defined as the classification of data based on knowledge already gained or on statistical information extracted from patterns and/or their representation.
- One of the important aspects of pattern recognition is its application potential.

- **Examples:** Speech recognition, speaker identification, multimedia document recognition (MDR), automatic medical diagnosis.  
In a typical pattern recognition application, the raw data is processed and converted into a form that is amenable for a machine to use. Pattern recognition involves the classification and cluster of patterns.
- In classification, an appropriate class label is assigned to a pattern based on an abstraction that is generated using a set of training patterns or domain knowledge. Classification is used in supervised learning.
- Clustering generated a partition of the data which helps decision making, the specific decision-making activity of interest to us. Clustering is used in unsupervised learning.



- **Advantages:**

- Pattern recognition solves classification problems
- Pattern recognition solves the problem of fake biometric detection.
- It is useful for cloth pattern recognition for visually impaired blind people.
- It helps in speaker diarization.
- We can recognize particular objects from different angles.

- **Disadvantages:**

- The syntactic pattern recognition approach is complex to implement and it is a very slow process.
- Sometimes to get better accuracy, a larger dataset is required.
- It cannot explain why a particular object is recognized.  
Example: my face vs my friend's face.

- **Applications:**

- **Image processing, segmentation, and analysis**

Pattern recognition is used to give human recognition intelligence to machines that are required in image processing.

- **Computer vision**

Pattern recognition is used to extract meaningful features from given image/video samples and is used in computer vision for various applications like biological and biomedical imaging.

- **Seismic analysis**

The pattern recognition approach is used for the discovery, imaging, and interpretation of temporal patterns in seismic array recordings. Statistical pattern recognition is implemented and used in different types of seismic analysis models.

- **Radar signal classification/analysis**

Pattern recognition and signal processing methods are used in various applications of radar signal classifications like AP mine detection and identification.

- **Speech recognition**

The greatest success in speech recognition has been obtained using pattern recognition paradigms. It is used in various algorithms of speech recognition which tries to avoid the problems of using a phoneme level of description and treats larger units such as words as pattern

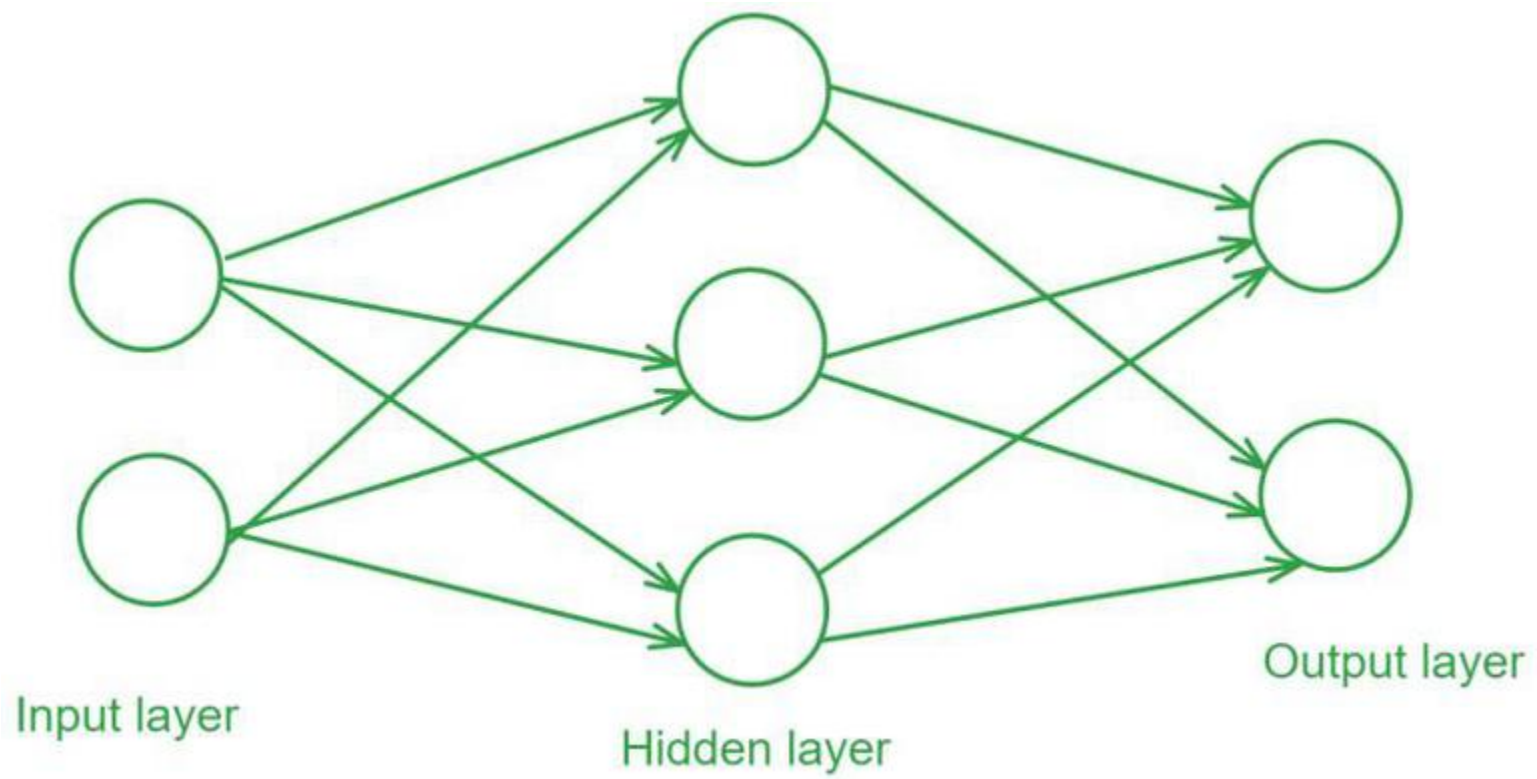
- **Fingerprint identification**

Fingerprint recognition technology is a dominant technology in the biometric market. A number of recognition methods have been used to perform fingerprint matching out of which pattern recognition approaches are widely used.

# What is a Feedforward Neural Network?

- A Feedforward Neural Network (FNN) is a type of artificial neural network where connections between the nodes do not form cycles.
- This characteristic differentiates it from recurrent neural networks (RNNs).
- The network consists of an input layer, one or more hidden layers, and an output layer. Information flows in one direction—from input to output—hence the name "feedforward."





# Activation Functions

- This function takes any real value as input and outputs values in the range of 0 to 1.
- The larger the input (more positive), the closer the output value will be to 1.0, whereas the smaller the input (more negative), the closer the output will be to 0.0, as shown below.
- Mathematically it can be represented as:

*Sigmoid / Logistic*

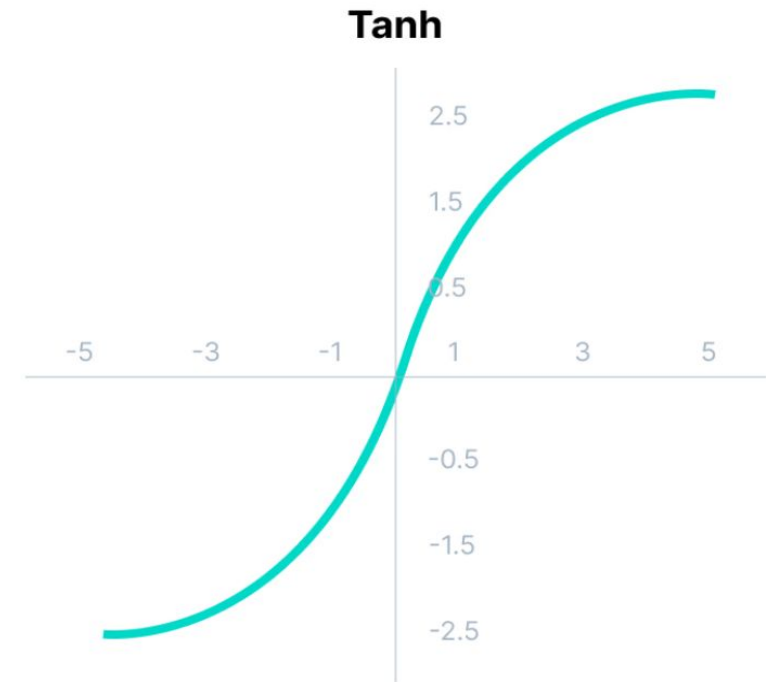
$$f(x) = \frac{1}{1 + e^{-x}}$$



# Tanh Function (Hyperbolic Tangent)

- Tanh function is very similar to the sigmoid/logistic activation function, and even has the same S-shape with the difference in output range of -1 to 1. In Tanh, the larger the input (more positive), the closer the output value will be to 1.0, whereas the smaller the input (more negative), the closer the output will be to -1.0.

$$\text{Tanh}$$
$$f(x) = \frac{(e^x - e^{-x})}{(e^x + e^{-x})}$$

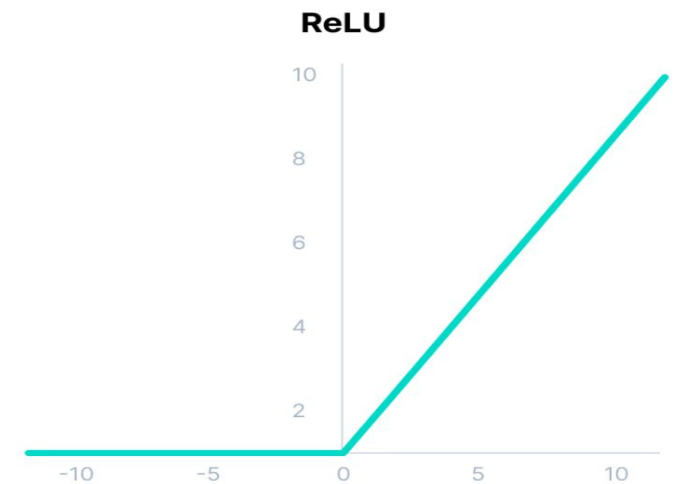


# ReLU Function

- ReLU stands for Rectified Linear Unit.
- Although it gives an impression of a linear function, ReLU has a derivative function and allows for backpropagation while simultaneously making it computationally efficient.
- The main catch here is that the ReLU function does not activate all the neurons at the same time.
- The neurons will only be deactivated if the output of the linear transformation is less than 0.

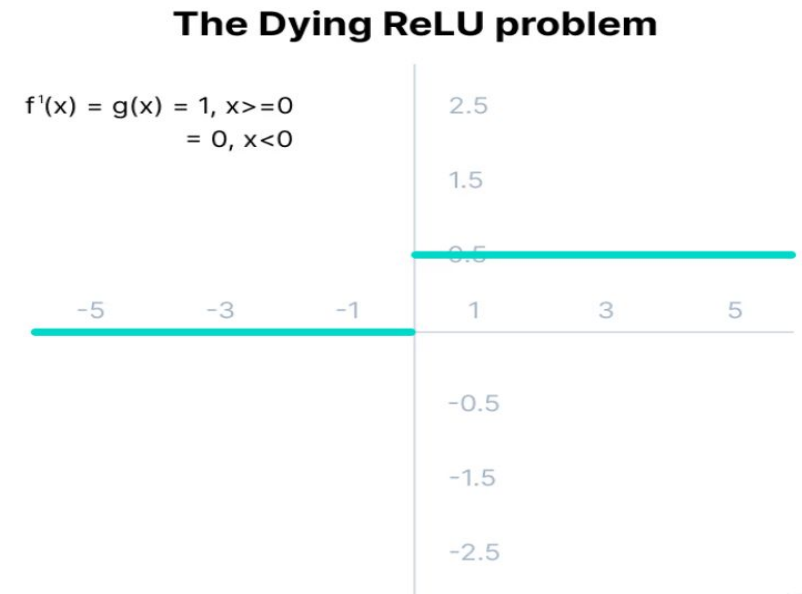
*ReLU*

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



# The Dying ReLU problem

- The negative side of the graph makes the gradient value zero. Due to this reason, during the backpropagation process, the weights and biases for some neurons are not updated. This can create dead neurons which never get activated.
- All the negative input values become zero immediately, which decreases the model's ability to fit or train from the data properly.
- **Note:** For building the most reliable ML models, split your data into train, validation, and test sets.



# Leaky ReLU Function

- Leaky ReLU is an improved version of ReLU function to solve the Dying ReLU problem as it has a small positive slope in the negative area.

*Leaky ReLU*

$$f(x) = \max(0.1x, x)$$

