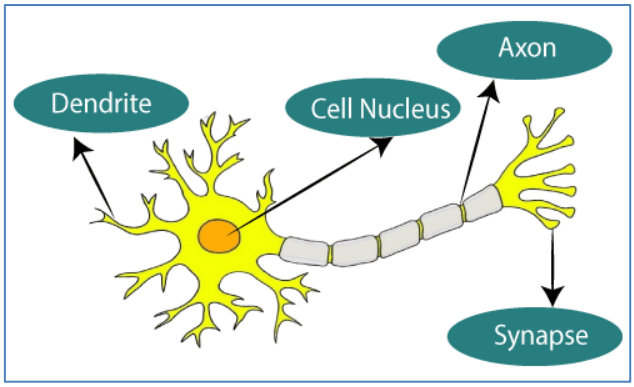
Chap 03 Introduction to Neural Network and Deep Learning

Artificial Neural Network Tutorial provides basic and advanced concepts of ANNs. Our Artificial Neural Network tutorial is developed for beginners as well as professions.

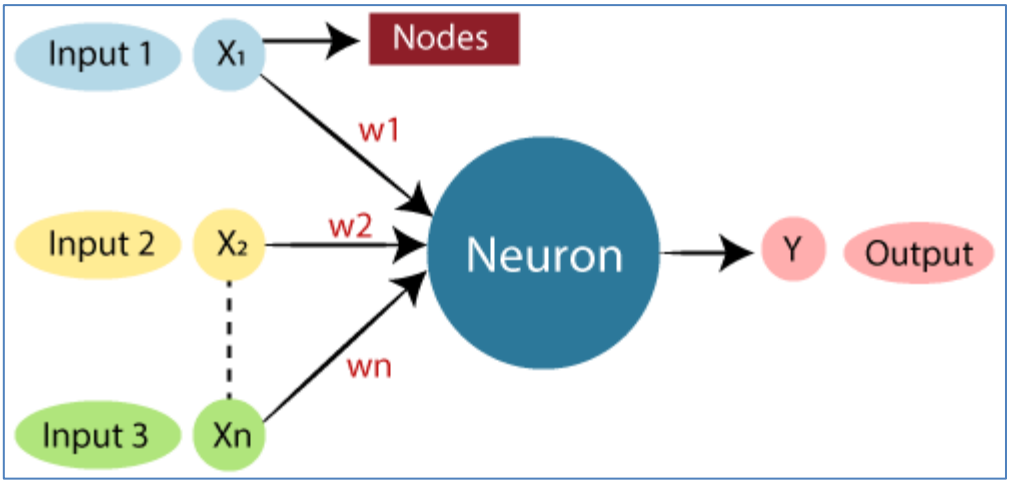
- The term "Artificial neural network" refers to a biologically inspired sub-field of artificial intelligence modeled after the brain.
- An Artificial neural network is usually a computational network based on biological neural networks that construct the structure of the human brain.
- Similar to a 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.



The given figure illustrates the typical diagram of Biological Neural Network.



Dendrites from Biological Neural Network represent inputs in Artificial Neural Networks, cell nucleus represents Nodes, synapse represents Weights, and Axon represents Output.

Relationship between Biological neural network and artificial neural network:

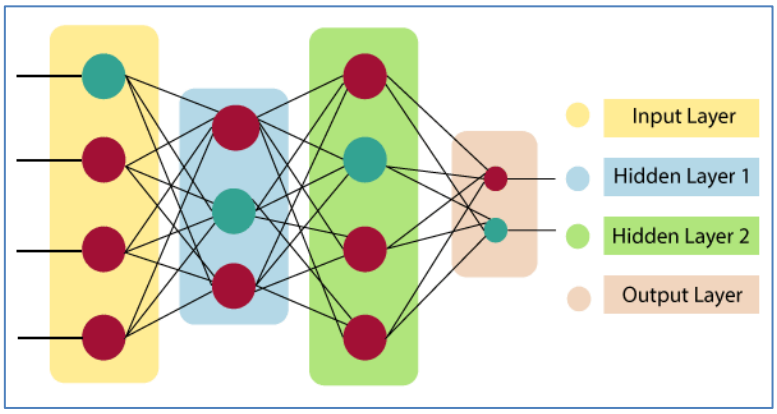
Biological Neural Network	Artificial Neural Network
Dendrites	Inputs
Cell nucleus	Nodes
Synapse	Weights
Axon	Output

☯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.

$$\sum_{i=1}^{n} Wi * Xi + b$$

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:** Artificial neural networks have a numerical value that can perform more than one task simultaneously.
- **Storing data on the entire network:** Data that is used in traditional programming is stored on the whole network, not on a database. The disappearance of a couple of pieces of data in one place doesn't prevent the network from working.
- **Capability to work with incomplete knowledge:** After ANN training, the information may produce output even with inadequate data. The loss of performance here relies upon the significance of missing data.
- **Having fault tolerance:** Extortion of one or more cells of ANN does not prohibit it from generating output, and this feature makes the network fault-tolerance.
- **Robustness to Noise:** ANNs are often robust to noise and can handle imperfect, incomplete, or noisy datasets. They can identify underlying patterns even in data with varying degrees of noise.

☸Disadvantages of Artificial Neural Network:

- **Need for Large Datasets:** ANNs often require large volumes of labeled training data to achieve good generalization and avoid overfitting. Insufficient or biased data might lead to poor performance.
- **Hardware dependence:** Artificial neural networks need processors with parallel processing power, as per their structure. Therefore, the realization of the equipment is dependent.
- **Vulnerability to Adversarial Attacks**: ANNs can be susceptible to adversarial attacks, where carefully crafted inputs can manipulate their decision-making process, leading to erroneous outputs. This poses security concerns in applications involving critical decision-making.
- **The duration of the network is unknown:** The network is reduced to a specific value of the error, and this value does not give us optimum results.

✡Types Basic Architecture of Neural Networks:

There are many types of neural networks available or that might be in the development stage. They can be classified depending on their: Structure, Data flow, Neurons used and their density, Layers and their depth activation filters etc. Also, learn about the Neural network in R to further your learning.

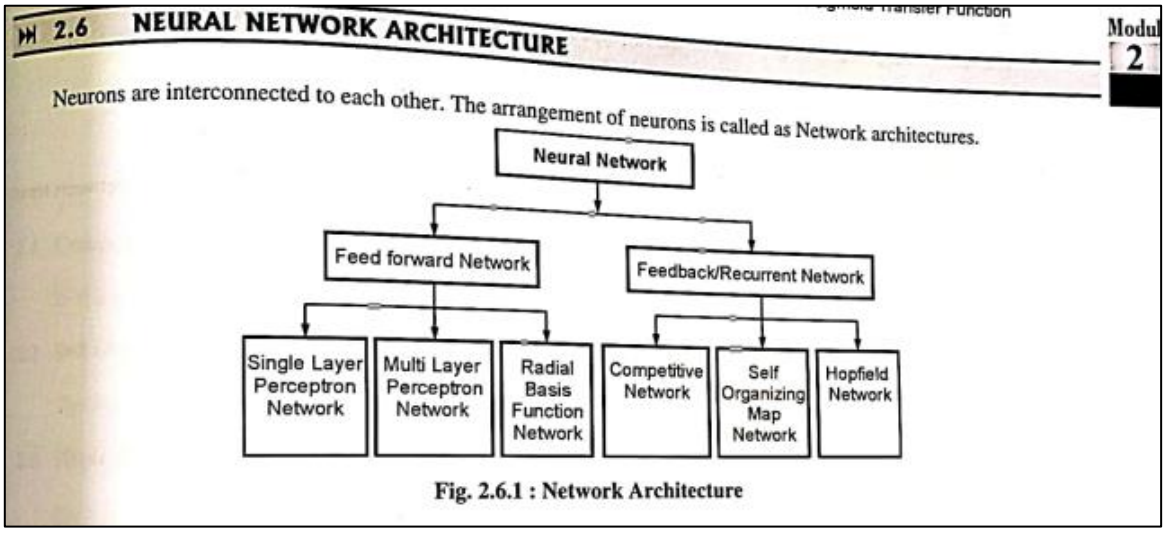
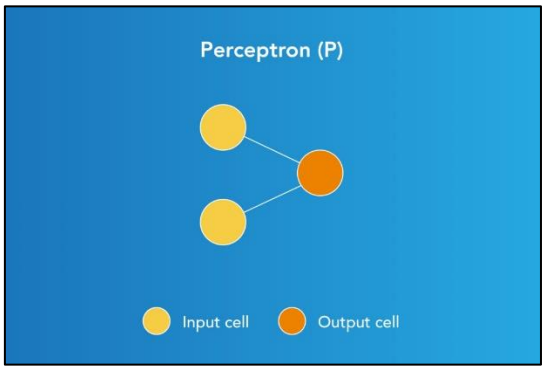


Fig. 2.6.1 : Network Architecture

A. Perceptron

Perceptron model, proposed by Minsky-Papert is one of the simplest and oldest models of Neuron. It is the smallest unit of neural network that does certain computations to detect features or business intelligence in the input data. It accepts weighted inputs, and apply the activation function to obtain the output as the final result. Perceptron is also known as TLU(threshold logic unit)



Advantages of Perceptron

Perceptrons can implement Logic Gates like AND, OR, or NAND.

Disadvantages of Perceptron

Perceptrons can only learn linearly separable problems such as boolean AND problem. For non-linear problems such as the boolean XOR problem, it does not work.

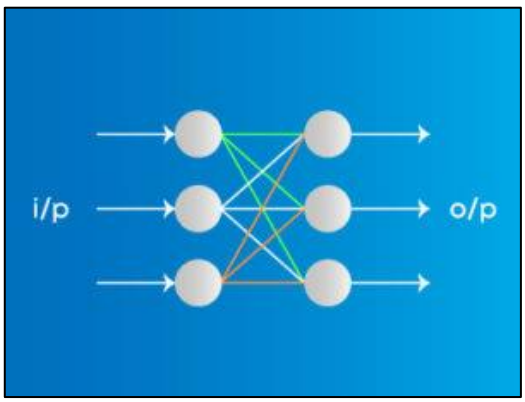
B Feed Forward Neural Networks

The simplest form of neural networks where input data travels in one direction only, passing through artificial neural nodes and exiting through output nodes. Where hidden layers may or may not be present, input and output layers are present there. Based on this, they can be further classified as a single-layered or multi-layered feed-forward neural network.

Number of layers depends on the complexity of the function. It has uni-directional forward propagation but no backward propagation. Weights are static here. An activation function is fed by inputs which are multiplied by weights. To do so, classifying activation function or step activation function is used. For example: The neuron is activated if it is above threshold (usually 0) and the neuron produces 1 as an output. The neuron is not activated if it is below threshold (usually 0) which is considered as -1. They are fairly simple to maintain and are equipped with to deal with data which contains a lot of noise.

Advantages of Feed Forward Neural Networks

1. Less complex, easy to design & maintain
2. Fast and speedy [One-way propagation]
3. Highly responsive to noisy data



Disadvantages of Feed Forward Neural Networks:

1. Cannot be used for deep learning [due to absence of dense layers and back propagation]

Applications on Feed Forward Neural Networks:

- Simple classification (where traditional Machine-learning based classification algorithms have limitations)
- Face recognition [Simple straight forward image processing]
- Speech Recognition

C Multilayer feed-forward network:

An entry point towards complex neural nets where input data travels through various layers of artificial neurons. Every single node is connected to all neurons in the next layer which makes it a fully connected neural network. Input and output layers are present having multiple hidden Layers i.e. at least three or more layers in total. It has a bi-directional propagation i.e. forward propagation and backward propagation.

Advantages on Multi-Layer Perceptron

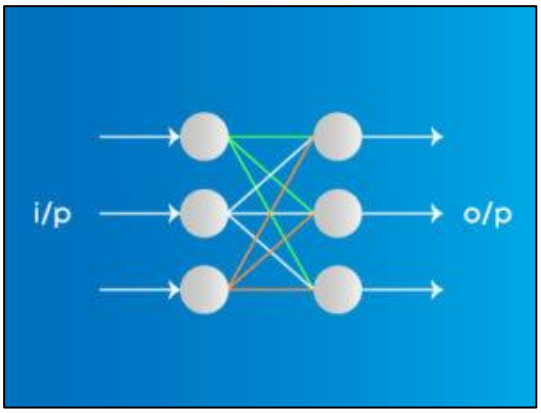
1. Used for deep learning [due to the presence of dense fully connected layers and back propagation]

Disadvantages on Multi-Layer Perceptron:

1. Comparatively complex to design and maintain

Applications on Multi-Layer Perceptron

- Speech Recognition
- Machine Translation
- Complex Classification

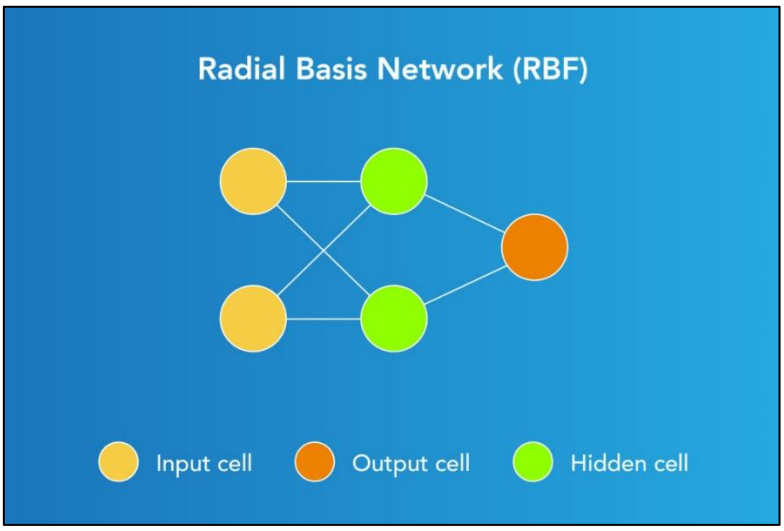


D)Radial Basis Function Neural Networks

Radial Basis Function Network consists of an input vector followed by a layer of RBF neurons and an output layer with one node per category. Classification is performed by measuring the input's similarity to data points from the training set where each neuron stores a prototype. This will be one of the examples from the training set.

When a new input vector [the n-dimensional vector that you are trying to classify] needs to be classified, each neuron calculates the Euclidean distance between the input and its prototype. For example, if we have two classes i.e. class A and Class B, then the new input to be classified is more close to class A prototypes than the class B prototypes. Hence, it could be tagged or classified as class A.

Each RBF neuron compares the input vector to its prototype and outputs a value ranging which is a measure of similarity from 0 to 1. As the input equals to the prototype, the output of that RBF neuron will be 1 and with the distance grows between the input and prototype the response falls off exponentially towards 0. The curve generated out of neuron's response tends towards a typical bell curve. The output layer consists of a set of neurons [one per category].



E) Recurrent Neural Networks:

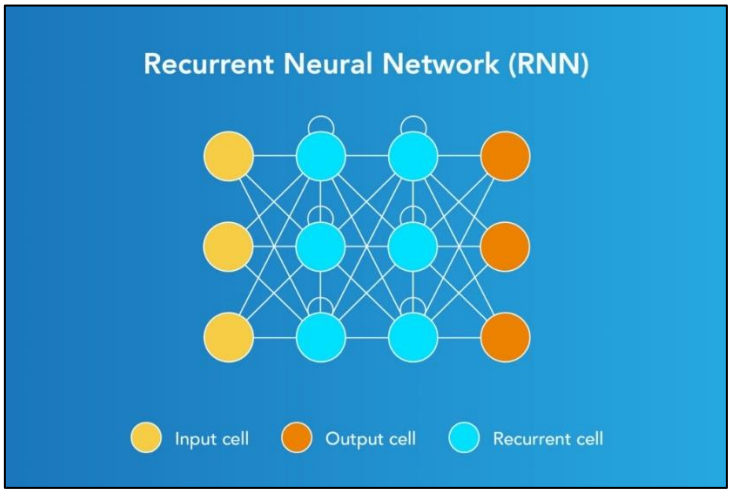
Recurrent Neural Network is fed back to the input to help in predicting the outcome of the layer. The first layer is typically a feed forward neural network followed by recurrent neural network layer where some information it had in the previous time-step is remembered by a memory function. Forward propagation is implemented in this case. It stores information required for it's future use. If the prediction is wrong, the learning rate is employed to make small changes. Hence, making it gradually increase towards making the right prediction during the backpropagation.

Advantages of Recurrent Neural Networks

- **Model sequential data where each sample can be assumed to be dependent on historical ones is one of the advantage.**
- **Used with convolution layers to extend the pixel effectiveness.**

Disadvantages of Recurrent Neural Networks

- **Gradient vanishing and exploding problems**
- **Training recurrent neural nets could be a difficult task**



Applications of Recurrent Neural Networks

- Text processing like auto suggest, grammar checks, etc.
- Text to speech processing
- Image tagger
- Sentiment Analysis

☮Issues in Neural Network Training

Neural networks are powerful tools for machine learning, but they can be challenging to train effectively. Here are some of the common issues that arise during neural network training:

1. **Overfitting**: This occurs when the neural network learns the training data too well and fails to generalize to new, unseen data. This can lead to poor performance on real-world tasks.

2. **Underfitting**: This occurs when the neural network does not learn the training data well enough, leading to poor performance on both training and testing data.

3**. Vanishing and exploding gradients**: This occurs when the gradients of the error function become very small or very large during training. This can make it difficult for the optimization algorithm to find the optimal weights for the neural network.

4. **Data dependency**: Neural networks rely heavily on large amounts of data for training and optimization. Without sufficient data, they may not learn effectively or generalize well to new scenarios.

5. **Computational complexity**: Training large neural networks can be computationally expensive, requiring specialized hardware and significant time resources. This can limit their applicability in real-time or resource-constrained environments.

6**. Sensitivity to initial parameters**: The performance of neural networks can be sensitive to the initial choice of parameters, such as network architecture, learning rate, and activation functions. This can make it challenging to find the optimal configuration.

7. **Lack of interpretability**: Neural networks are often referred to as "black boxes" due to their complex internal structure and decision-making processes. This makes it difficult to understand why a particular output is generated, hindering interpretability and explainability.

8. **Potential for bias:** Neural networks can inherit biases present in the training data, leading to discriminatory or unfair outcomes. This raises ethical concerns and necessitates careful data curation and bias mitigation strategies.

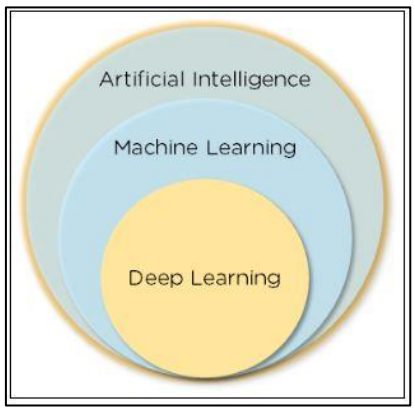
9. **Vulnerability to adversarial attacks**: Neural networks can be susceptible to adversarial attacks, where carefully crafted inputs can manipulate their decision-making process, leading to erroneous outputs. This poses security concerns in applications involving critical decision-making.

10**. Requirement for expert knowledge:** Developing and deploying neural networks effectively requires expertise in machine learning, deep learning algorithms, and optimization techniques. This can limit their accessibility to non-experts.

☢Deep learning:

Deep learning is a subfield of machine learning that deals with algorithms inspired by the structure and function of the human brain. These algorithms, called artificial neural networks (ANNs), are composed of interconnected nodes, or neurons, that process and transmit information.

- Deep learning algorithms are capable of learning from data and making predictions or decisions based on that data.
- Deep learning models are capable enough to focus on the accurate features themselves by requiring a little guidance from the programmer and are very helpful in solving out the problem of dimensionality.
- Deep learning algorithms are used, especially when we have a huge no of inputs and outputs.



♋ Why is Deep Learning Important?

Deep learning has revolutionized many fields, including computer vision, natural language processing, and robotics. It has led to the development of self-driving cars, voice assistants, and chatbots. Deep learning is also used in a variety of other applications, such as fraud detection, medical diagnosis, and drug discovery.

♍ How Does Deep Learning Work?

Deep learning algorithms are based on the idea of artificial neurons. Artificial neurons are mathematical functions that take input values, process them, and produce an output value. The output value of one neuron can become the input value of another neuron. This allows artificial neurons to be connected into networks, called artificial neural networks.

Artificial neural networks are trained using a process called supervised learning. In supervised learning, the neural network is given a set of training data that consists of input data and corresponding output data. The neural network is then trained to produce the correct output data for each input data.

What are the Different Types of Deep Learning Algorithms?

There are many different types of deep learning algorithms, but some of the most common include:

- **Feedforward neural networks**: In a feedforward neural network, the connections between neurons only go in one direction, from the input layer to the hidden layer(s) and then to the output layer.
- **Convolutional neural networks** (CNNs): CNNs are specifically designed for image recognition and analysis. They have a special type of layer called a convolutional layer, which is able to extract features from images without the need for hand-**engineered features.**
- **Recurrent neural networks** (RNNs): RNNs are able to process sequential data, such as text or time series data, by having connections that loop back to themselves. This allows RNNs to capture temporal dependencies in the data.

☸Architectures

- ***Deep Neural Networks***
  It is a neural network that incorporates the complexity of a certain level, which means several numbers of hidden layers are encompassed in between the input and output layers. They are highly proficient on model and process non-linear associations.
- ***Deep Belief Networks***
  A deep belief network is a class of Deep Neural Network that comprises of multi-layer belief networks.
  Steps to perform DBN:
  a. With the help of the Contrastive Divergence algorithm, a layer of features is learned from perceptible units.
  b. Next, the formerly trained features are treated as visible units, which perform learning of features.
  c. Lastly, when the learning of the final hidden layer is accomplished, then the whole DBN is trained.
- ***Recurrent Neural Networks***
  It permits parallel as well as sequential computation, and it is exactly similar to that of the human brain (large feedback network of connected neurons). Since they are capable enough to reminisce all of the imperative things related to the input they have received, so they are more precise.

✳Deep learning applications

- **Self-Driving Cars :** In self-driven cars, it is able to capture the images around it by processing a huge amount of data, and then it will decide which actions should be incorporated to take a left or right or should it stop. So, accordingly, it will decide what actions it should take, which will further reduce the accidents that happen every year.
- **Voice Controlled Assistance:** When we talk about voice control assistance, then Siri is the one thing that comes into our mind. So, you can tell Siri whatever you want it to do it for you, and it will search it for you and display it for you.
- **Automatic Image Caption Generation:** Whatever image that you upload, the algorithm will work in such a way that it will generate caption accordingly. If you say blue colored eye, it will display a blue-colored eye with a caption at the bottom of the image.
- **Automatic Machine Translation:** With the help of automatic machine translation, we are able to convert one language into another with the help of deep learning.

Advantages

- **It lessens the need for feature engineering.**
- **It eradicates all those costs that are needless.**
- **It easily identifies difficult defects.**
- **It results in the best-in-class performance on problems.**

Disadvantages

- **It requires an ample amount of data.**
- **It is quite expensive to train.**
- **It does not have strong theoretical groundwork.**

☣What is TensorFlow?

TensorFlow is a popular framework of machine learning and deep learning. It is a free and open-source library which is released on 9 November 2015 and developed by Google Brain Team.

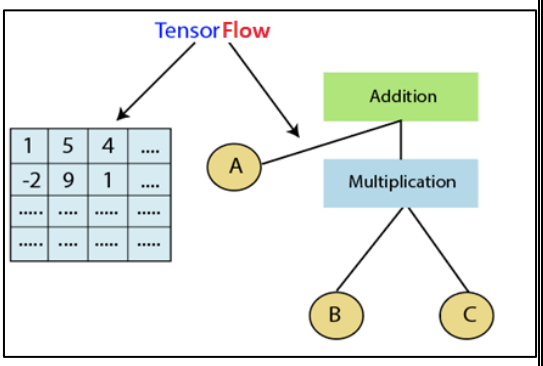
- It is entirely based on Python programming language and use for numerical computation and data flow, which makes machine learning faster and easier.
- TensorFlow can train and run the deep neural networks for image recognition, handwritten digit classification, recurrent neural network, word embedding, natural language processing, video detection, and many more. TensorFlow is run on multiple CPUs or GPUs and also mobile operating systems.

The word TensorFlow is made by two words, i.e., Tensor and Flow

Tensor is a multidimensional array

Flow is used to define the flow of data in operation.

TensorFlow is used to define the flow of data in operation on a multidimensional array or Tensor



✳History of TensorFlow

A couple of years ago, deep learning started to outperform all other machine learning algorithms when giving a massive amount of data. Google saw it could use these deep neural networks to improve its services:

- **Gmail**
- **Photo**
- **Google search engine**

They build a framework called Tensorflow to let researchers and developers work together on an AI model. Once developed and scaled, it allows lots of people to use it.

It was first made public in late 2015, while the first stable version appeared in 2017. It is open source under Apache Open Source license. You can use it, modify it and redistribute the modified version for a fee without paying anything to Google**.**

☣Key Features of TensorFlow:

1. **Flexible and Expressive:** TensorFlow offers a variety of APIs, including the eager execution API and the functional API, catering to different programming styles and preferences.

2. **Scalability and Performance:** TensorFlow can be scaled to run on various hardware platforms, including CPUs, GPUs, and TPUs, enabling efficient training and deployment of complex models.

3. **Distributed Computing:** TensorFlow supports distributed training across multiple machines or cloud instances, allowing for the handling of large datasets and accelerated model training.

4. **Integration with Other Tools:** TensorFlow integrates seamlessly with other popular machine learning tools and libraries, such as scikit-learn and matplotlib, providing a comprehensive ecosystem for data analysis and model development

Getting Started with TensorFlow:

- **Install TensorFlow:** TensorFlow can be installed using pip or Anaconda, depending on your preferred package manager.
- **Choose an API:** TensorFlow offers various APIs, including the eager execution API and the functional API. Select the API that suits your programming style and project requirements.
- **Learn the Basics:** Explore TensorFlow tutorials and documentation to familiarize yourself with the core concepts and syntax.
- **Practice with Examples**: Practice building and training simple machine learning models using TensorFlow tutorials and examples.
- **Join the Community**: Engage with the TensorFlow community through forums, discussion groups, and online resources to seek help and learn from experienced users.

Aspect	Machine Learning	Deep Learning	Artificial Neural Networks (ANNs)
Scope	Subset of AI focused on algorithms	Subset of machine learning using neural networks	A computational model inspired by the human brain's structure
Approach	Learns from data	Learns intricate patterns from data	Building blocks of deep learning using interconnected nodes
Depth of Processing	May involve feature engineering	Learns hierarchical representations	Consists of layers, connections, and activation functions
Complexity of Networks	May use simpler models	Involves complex, deep architectures	Basic structure of interconnected nodes in multiple layers
Representation Learning	May require manual feature extraction	Automatically learns features from data	Constructs representations from raw data
Task Suitability	Suitable for various tasks	Suitable for complex tasks	Forms the basis for deep learning architectures
Application Domains	Widespread across industries	Dominant in image, speech, and NLP tasks	Fundamental to various deep learning architectures
Hardware Resources	Moderate resource requirements	Demands substantial computational power	Basic building blocks adaptable to varying architectures
Model Interpretability	Easier to interpret models	Models can be complex and less interpretable	Basic structures, but complex networks may lack interpretability