Neural Networks

These computer systems have learning mechanisms similar to those of the human brain which help identify data patterns using a series of algorithms. It forms the foundation of deep learning and a subset of machine learning.

History Of Deep Learning

It all started in 1943 when the first computational model was created called threshold logic with advancements in 1958 by the perceptron which was a 2 layer network using simple addition and subtraction. By 1980, Neocognitron, a hierarchical, multilayered artificial neural network that has been used for pattern recognition problems was invented, but all these still took days of training time and were yet impractical for real life uses. In the mid-2000s we had 3-D object recognition automatically from cluttered scenes and how many-layered neural networks could be pre-trained one layer at a time and the term 'Deep Learning' gained popularity and error rates drop significantly. And finally in the early 2010's Google and Facebook started using AI for pattern recognition, optimized social media.

What is Deep Learning

"Deep Learning shows that with enough data, quick processors, sophisticated algorithms computers can start to perform tasks that were previously limited to human perception"

A Deep Learning model is trained by first defining its learning objective and then fine-tuning its parameters to maximize its output. This is not the same as traditional machine learning, which takes time and isn't necessarily efficient because it involves manually creating and picking features, properties of data that the system should look at.

Basic Neural Networks:

Perceptron is an algorithm that makes the neurons learn from the given information. A perceptron takes several binary inputs, x1, x2,..., and produces a single binary output. It is of two types Single Layer which contains one input and one output layer with multiple nodes each and Multi Layer, which has added internal layers. Each relationship has an associated weight, which impacts the preceding neuron's importance in the overall neural network and these weights are used to train models. The output is then passed to the activation function which is the backbone of deep learning which determines whether a neuron should be activated or not by introducing non linearity.

There are 4 main types of activation functions, each discussed below:

1. Threshold functions: It computes an output signal on the basis of weighted sum of input values from the preceding layer, if the input value is above a certain

threshold value then output is a certain value and if not then another value. Their output is discontinuous which makes it limited in the context of gradient based backpropagation.

2. Sigmoid Function: It produces values between 0 to 1 making it ideal for probabilities, logistic regression and moreover the smooth curve facilitates optimization of gradient techniques calculation of derivatives.
3. Rectified Linear Unit (ReLU): Its piecewise linear nature makes it computationally efficient and easier to train. If the input value is less than 0 then output is 0 and if the input is greater than zero, the output is x.
4. The hyperbolic tangent function: It is a sigmoid in shape but symmetric around the origin with values between -1 to +1.