Machine Learning

Course Code: AFI 124

Previous Class

1. Artificial Intelligence
   AI, AI Perspectives: acting and thinking humanly, acting and thinking rationally
2. History of AI
3. Foundations of AI: Philosophy, Economics, Psychology, Sociology, Linguistic, Neuroscience, Mathematics, Computer Science, Control Theory
4. Applications of AI
5. Handson Github

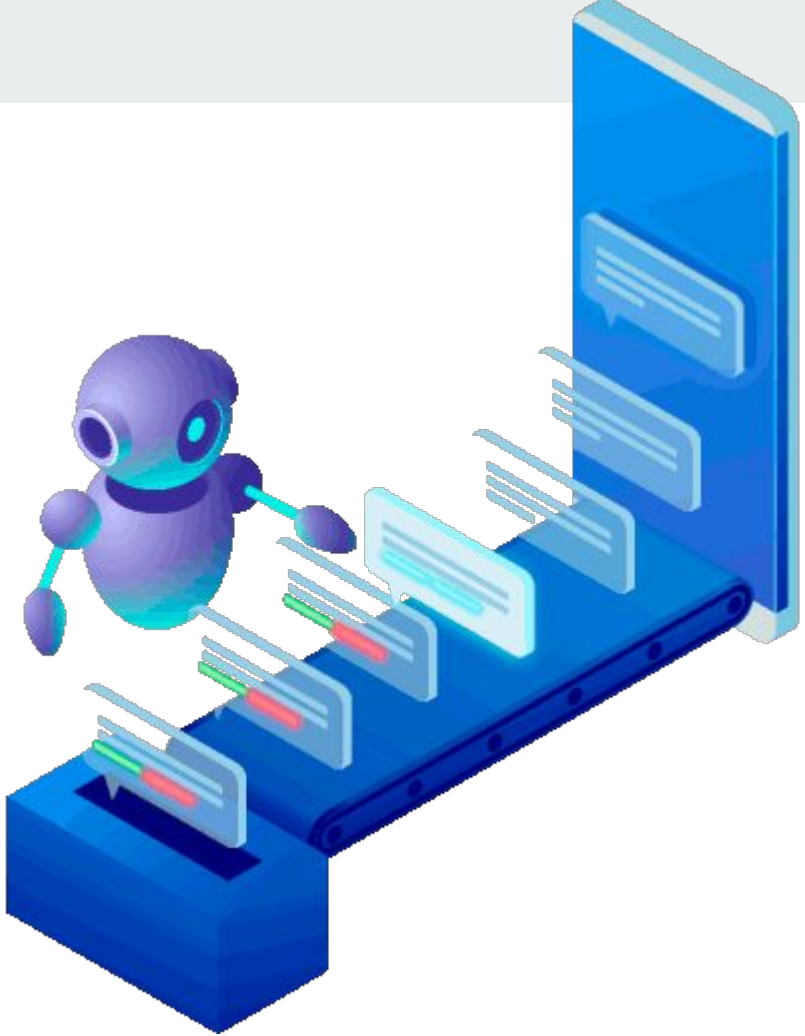
Contents

1. Introduction to Machine Learning, Concepts of Learning, Supervised, Unsupervised and
2. Reinforcement Learning
3. Statistical-based Learning: Naive Bayes Model
4. Learning with Neural Networks: Introduction, Biological Neural Networks Vs. Artificial Neural Networks (ANN), Mathematical Model of ANN, Activation Functions: Linear, Step Sigmoid, Types of ANN: Feed-forward, Recurrent, Single Layered, Multi-Layered, Application of Artificial Neural Networks, Learning by Training ANN, Supervised vs. Unsupervised Learning, Hebbian Learning, Perceptron Learning, Back - propagation Learning

1. Machine Learning

Machine learning is about extracting knowledge from the data. It can be defined as,

“Machine learning is a subfield of artificial intelligence, which enables machines to learn from past data or experiences without being explicitly programmed.”

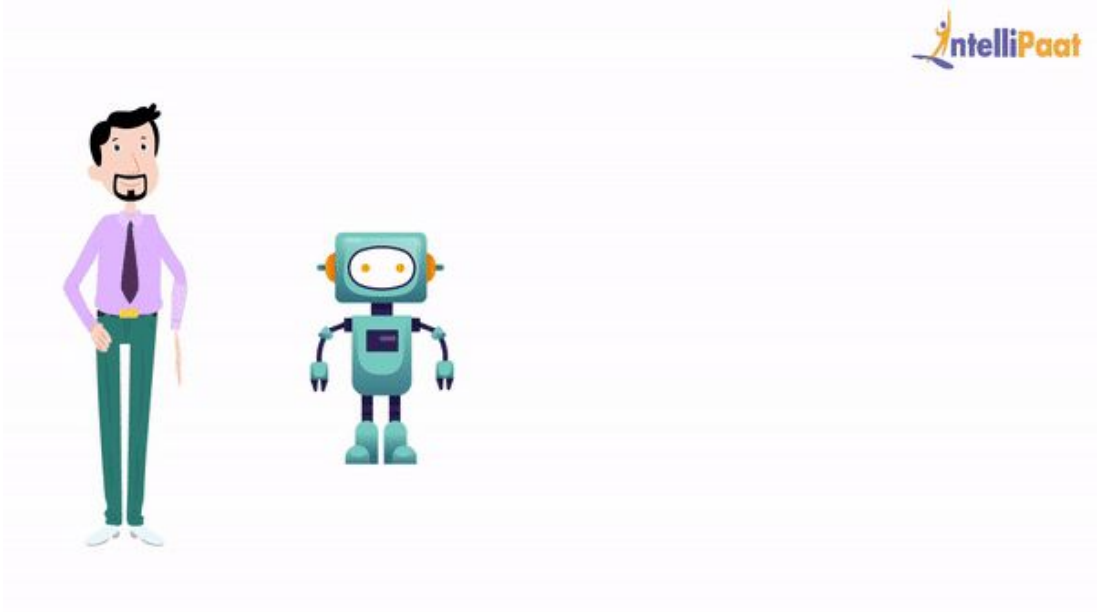


1.1. Concept of Machine Learning

- Supervised
- Unsupervised

1.1. Concept of Machine Learning: Supervised Learning

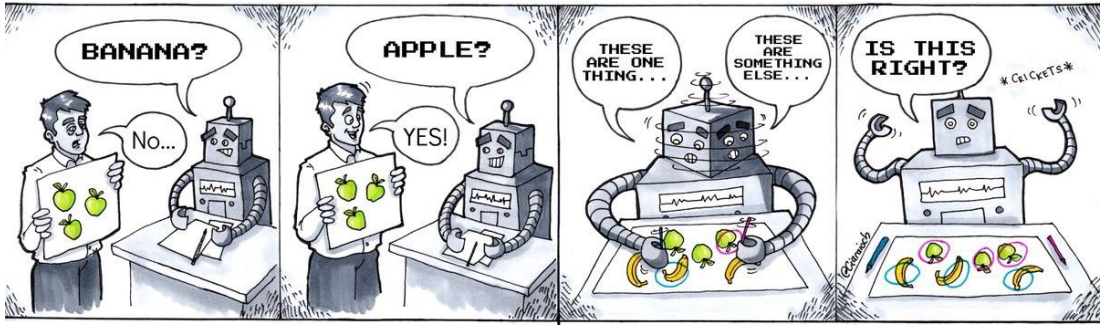
In **supervised learning** the agent observes input-output pairs and learns a function that maps from input to output.



1.1. Concept of Machine Learning: Unsupervised Learning

In **unsupervised learning** the agent learns patterns in the input without any explicit feedback. The most common unsupervised learning task is **clustering.**

Supervised and Unsupervised Learning



Supervised Learning

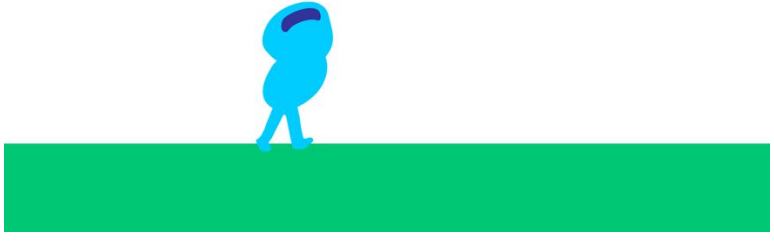
- Supervised machine learning relies on labelled input and output training data.
- Algorithms: Nearest Neighbor, Naive Bayes, Decision Trees, Linear Regression, Support Vector Machines (SVM), Neural Networks

Unsupervised Learning

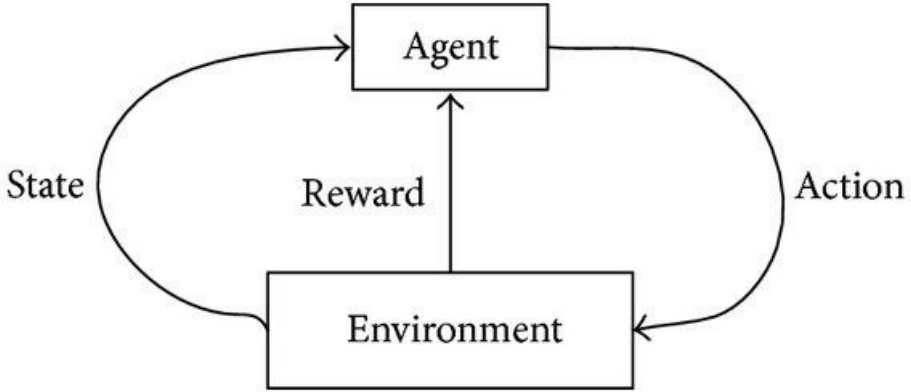
- Unsupervised learning processes unlabelled or raw data.
- Algorithms: K-means Clustering, Affinity clustering etc.

2. Reinforcement Learning

Method aims at using observations gathered from the interaction with the environment to take actions that would maximize the reward or minimize the risk. Reinforcement learning algorithm (called the agent) continuously learns from the environment in an iterative fashion. In the process, the agent learns from its experiences of the environment until it explores the full range of possible states.



2. Reinforcement Learning



In order to produce intelligent programs (also called agents), reinforcement learning goes through the following steps:

- Input state is observed by the agent.
- Decision making function is used to make the agent perform an action.
- After the action is performed, the agent receives reward or reinforcement from the environment.
- The state-action pair information about the reward is stored.

Use cases:

Some applications of the reinforcement learning algorithms are computer played board games (Chess, Go), robotic hands, and self-driving cars.

3. Statistical-based Learning: Naive Bayes Model

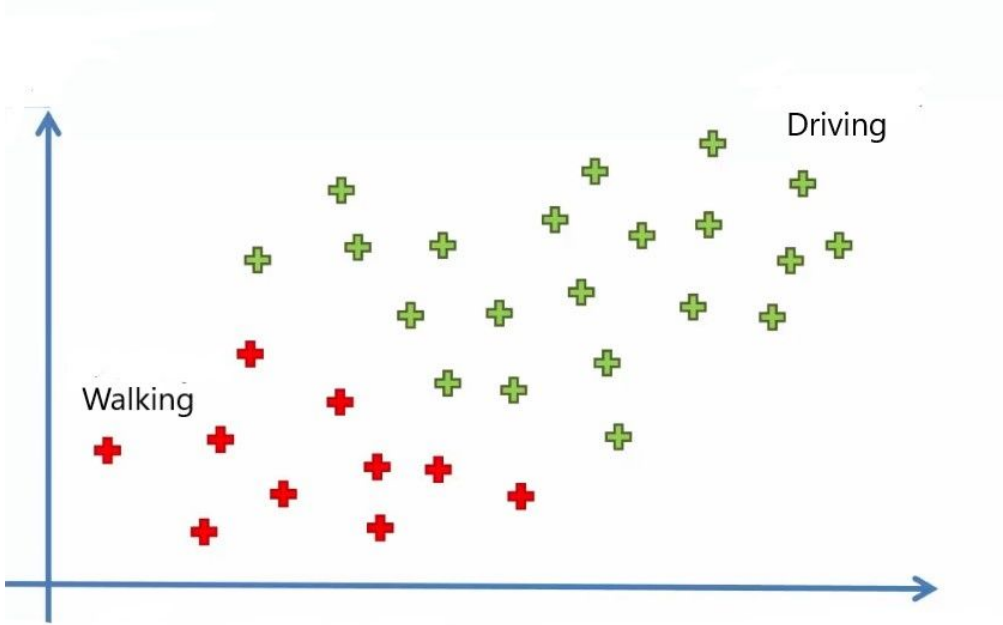
Naive Bayes Model is a classification technique based on **Bayes' Theorem** with an assumption of independence among predictors. In simple terms, a Naive Bayes classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature.

For example, a fruit may be considered to be an apple if it is red, round, and about 3 inches in diameter. Even if these features depend on each other or upon the existence of the other features, all of these properties independently contribute to the probability that this fruit is an apple and that is why it is known as ‘Naive’.

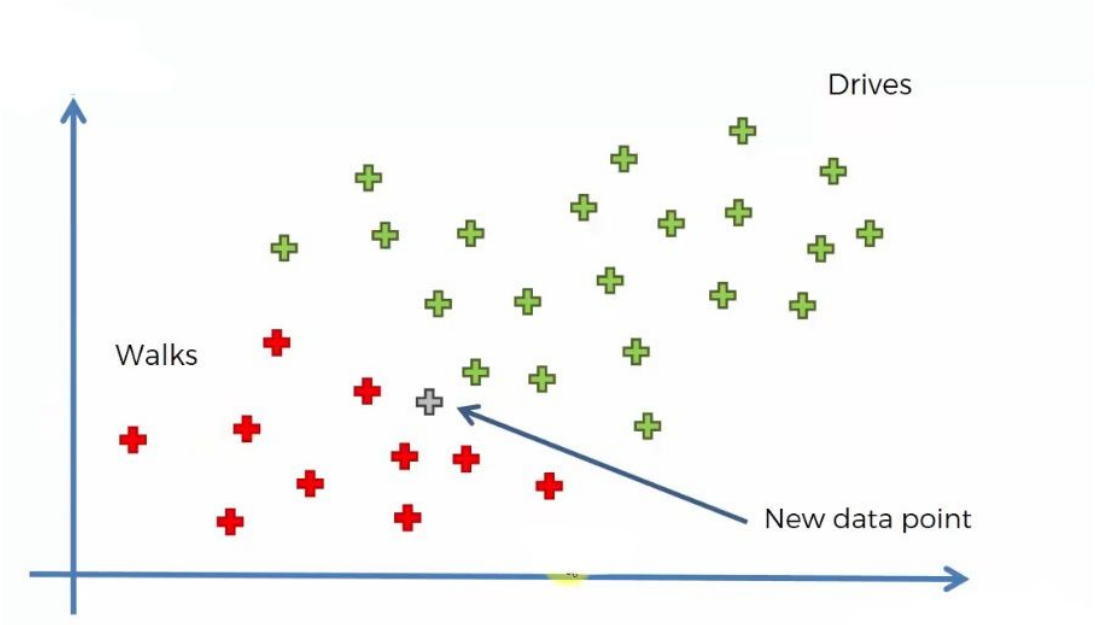
Bayes Theorem

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)}$$

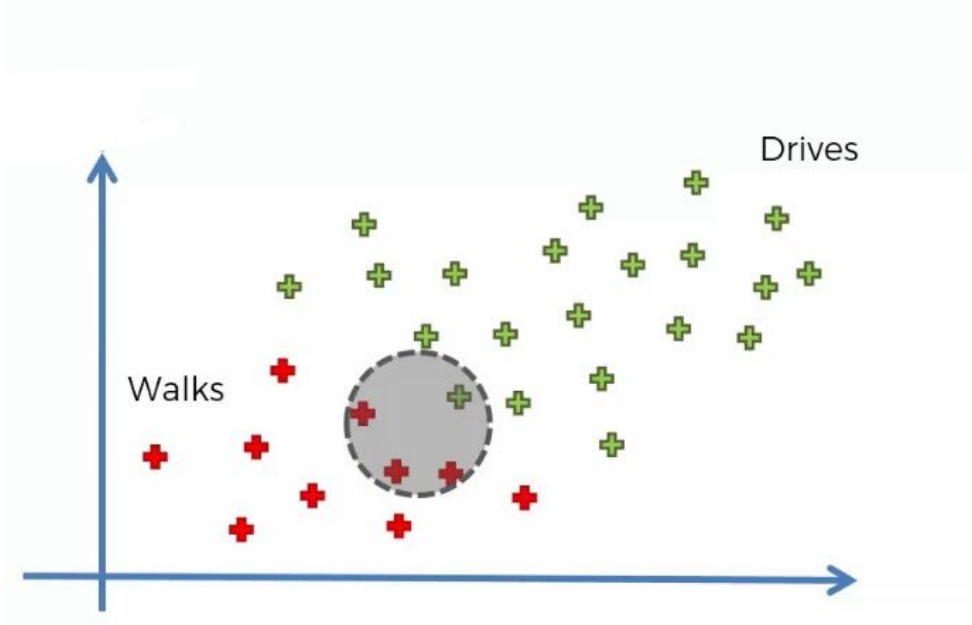
Naive Bayes Model Example



Naive Bayes Model Example



Naive Bayes Model Example

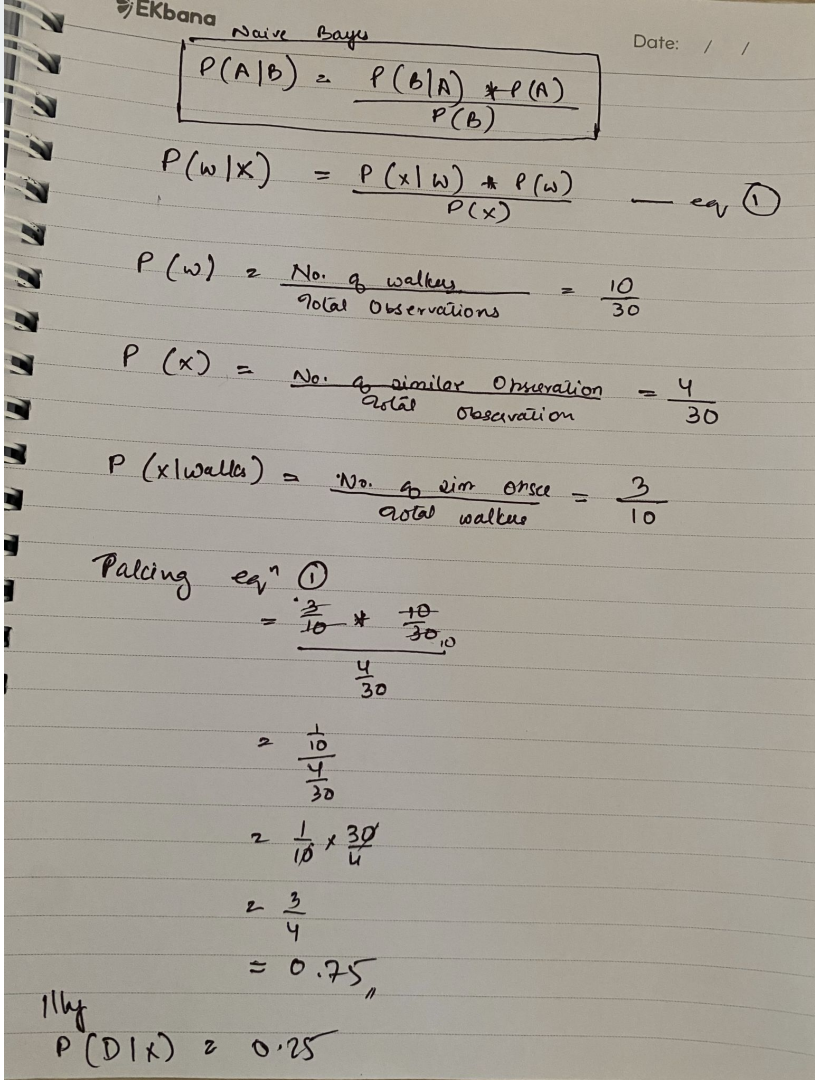


Step by Step

Naive Bayes classifier calculates the probability of an event in the following steps:

- Step 1: Calculate the prior probability for given class labels
- Step 2: Find Likelihood probability with each attribute for each class
- Step 3: Put these value in Bayes Formula and calculate posterior probability.
- Step 4: See which class has a higher probability, given the input belongs to the higher probability class.

[Notebook Link]



To learn more on Machine Learning

Course: https://www.udacity.com/course/intro-to-machine-learning--ud120

4. Learning with Neural Networks

To be Continued in Next Class….

CONTD....

Previous Class

1. Introduction to Machine Learning,
2. Concepts of Learning: Supervised, Unsupervised and
3. Reinforcement Learning
4. Statistical-based Learning: Naive Bayes Model

Contents

Learning with Neural Networks:

- Introduction: Deep Learning, Neural Network, Artificial Neural Network
- Biological Neural Networks Vs. Artificial Neural Networks (ANN),
- Example of Neural Network: House Price Prediction
- Activation Functions: Sigmoid, Tanh, ReLu
- Types of ANN: Feed-forward, Single Layered, Multi-Layered
- Application of Artificial Neural Networks,
- Learning Techniques in Neural Networks
- Perceptron Learning,
- Back - propagation Learning
- Transfer Learning

Deep Learning

Deep learning is a broad family of techniques for machine learning in which hypotheses take the form of complex algebraic circuits with tunable connection strengths. The word “deep” refers to the fact that the circuits are typically organized into many **layers**, which means that computation paths from inputs to outputs have many steps.

Neural Network

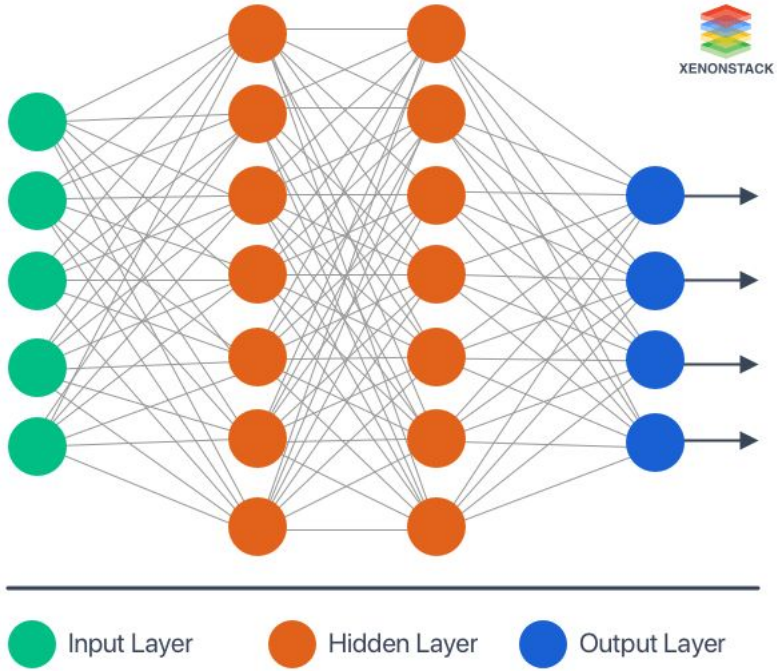
A neural network is **a series of algorithms that endeavors to recognize underlying relationships in a set of data through a process that mimics the way the human brain operates**. In this sense, neural networks refer to systems of neurons, either organic or artificial in nature.

Artificial neural network

Artificial neural networks, usually simply called neural networks, are computing systems inspired by the biological neural networks that constitute animal brains. An ANN is based on a collection of connected units or nodes called artificial neurons, which loosely model the neurons in a biological brain.

Artificial neural network Architecture

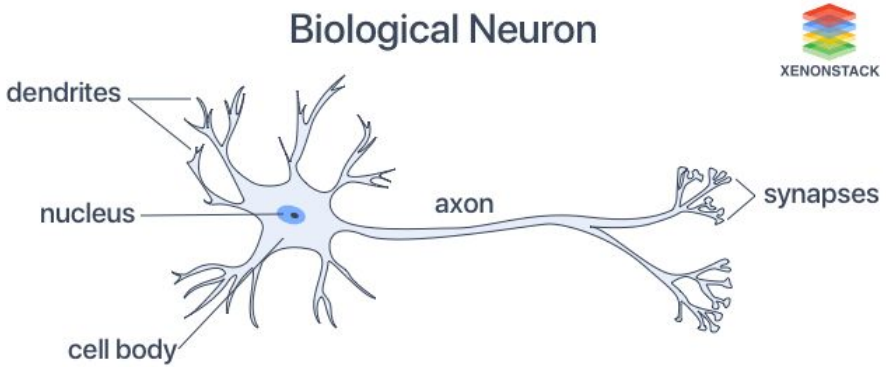
- **Input layer** - It contains those units (Artificia Neurons) which receive input from the outside world on which the network will learn, recognize about, or otherwise process.
- **Output layer** - It contains units that respond to the information about how it learn any task.
- **Hidden layer** - These units are in between input and output layers. The hidden layer's job is to transform the input into something that the output unit can use somehow.



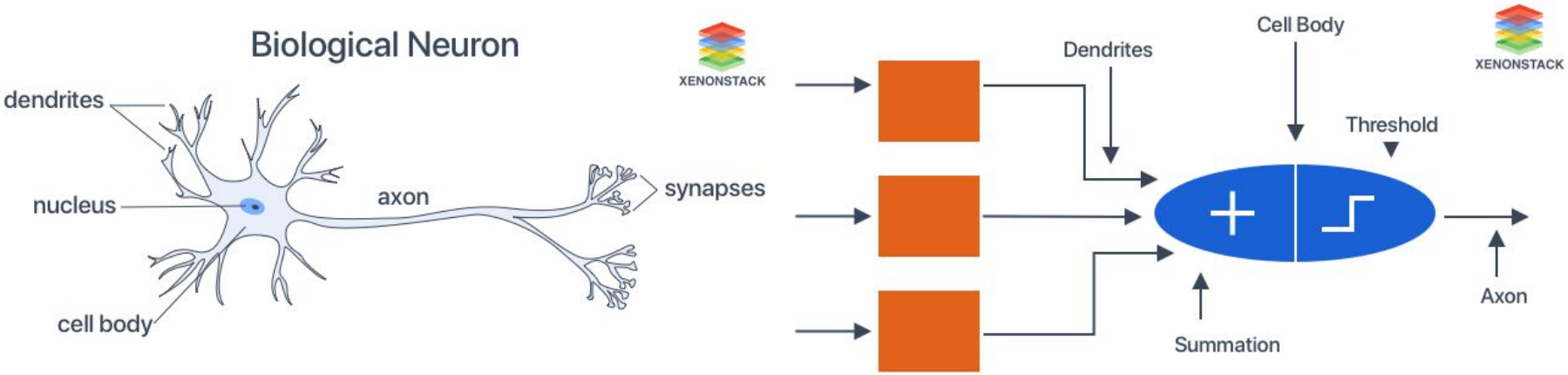
Biological Neuron

The typical nerve cell of the human brain comprises of four parts

- **Function of Dendrite** It receives signals from other neurons.
- **Soma (cell body)** It sums all the incoming signals to generate input.
- **Axon Structure** When the sum reaches a threshold value, the neuron fires, and the signal travels down the axon to the other neurons.
- **Synapses Working** The point of interconnection of one neuron with other neurons. The amount of signal transmitted depends upon the strength (synaptic weights) of the connections.



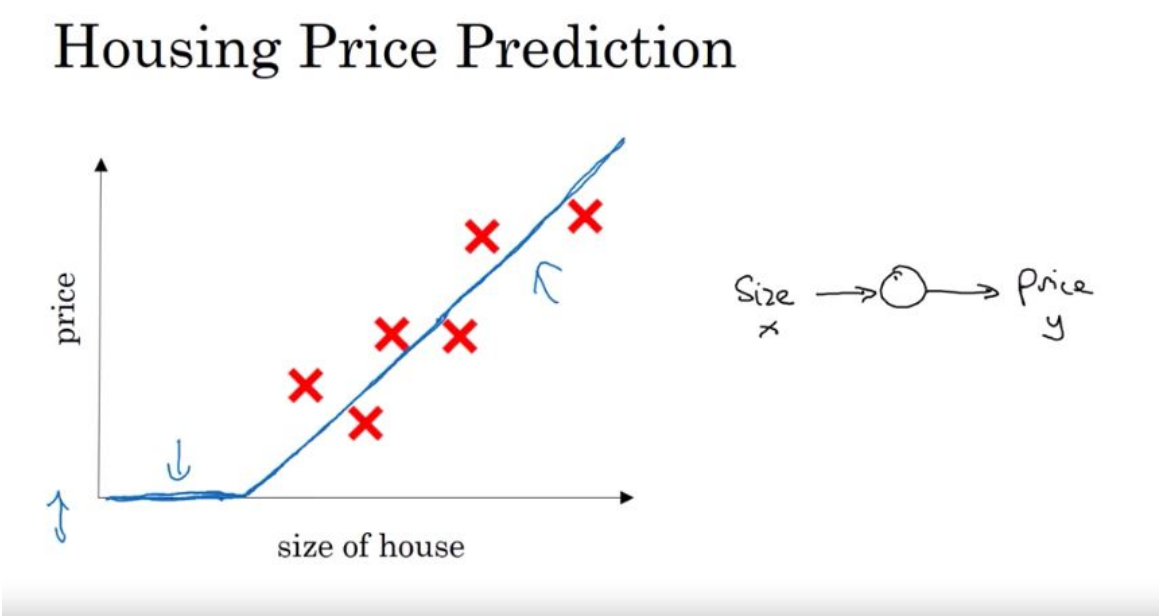
Biological Neural Networks Vs. Artificial Neural Networks (ANN)



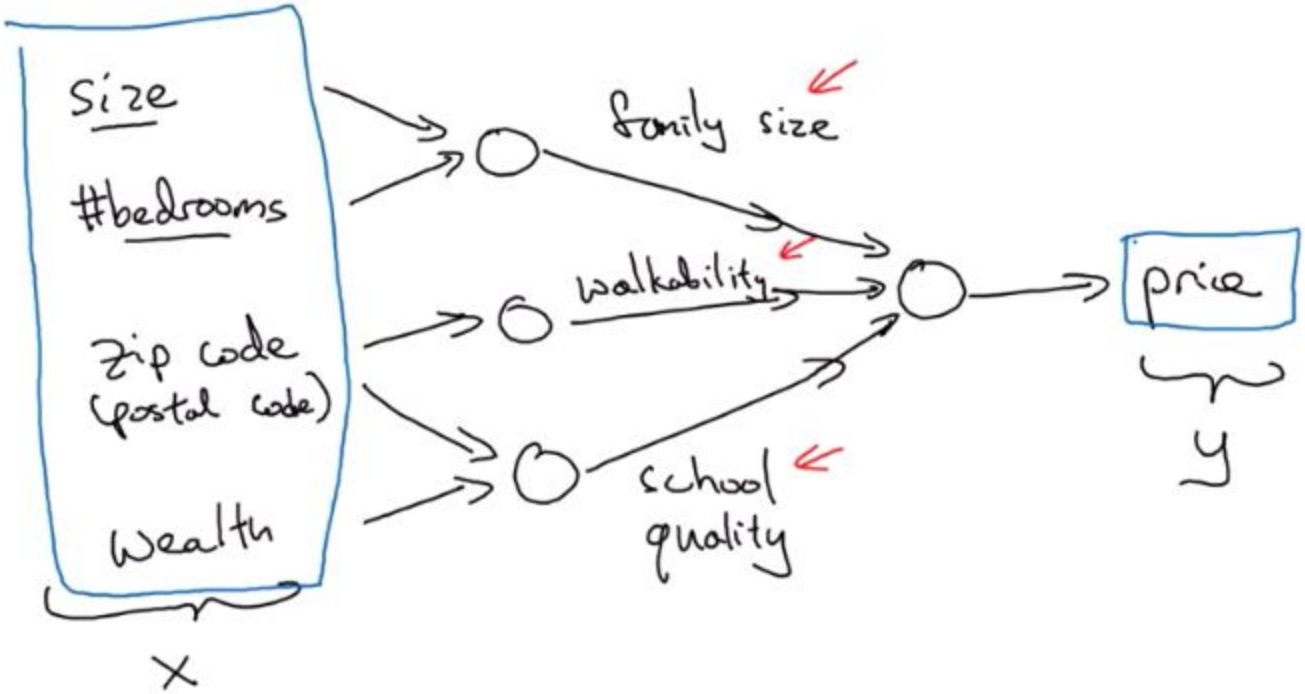
Biological Neural Networks Vs. Artificial Neural Networks (ANN)

Characteristics	Artificial Neural Network	Biological(Real) Neural Network
Speed	Faster in processing information. Response time is in nanoseconds.	Slower in processing information. The response time is in milliseconds.
Processing	Serial processing.	Massively parallel processing.
Size & Complexity	Less size & complexity. It does not perform complex pattern recognition tasks.	A highly complex and dense network of interconnected neurons containing neurons of the order of 1011 with 1015 of interconnections.<strong
Storage	Information storage is replaceable means replacing new data with an old one.	A highly complex and dense network of interconnected neurons containing neurons of the order of 1011 with 1015 of interconnections.
Fault tolerance	Fault intolerant. Corrupt information cannot retrieve in case of failure of the system.	Information storage is adaptable means new information is added by adjusting the interconnection strengths without destroying old information.
Control Mechanism	There is a control unit for controlling computing activities	No specific control mechanism external to the computing task.

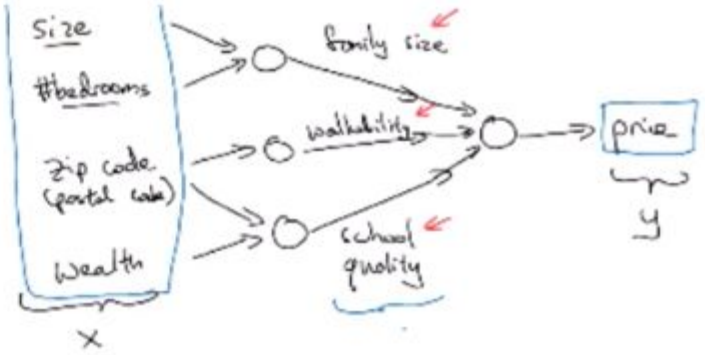
Example: House Price prediction

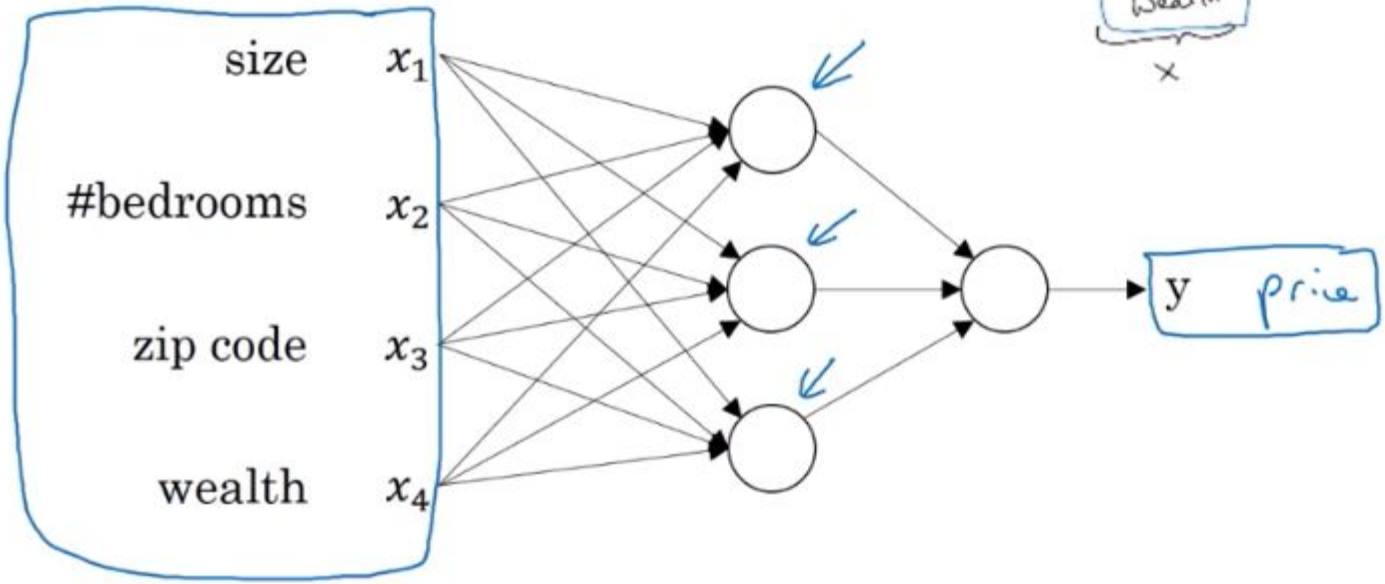


Housing Price Prediction



Housing Price Prediction





Activation Functions

Activation function defines the output of input.

Activation functions commonly used in deep learning systems:

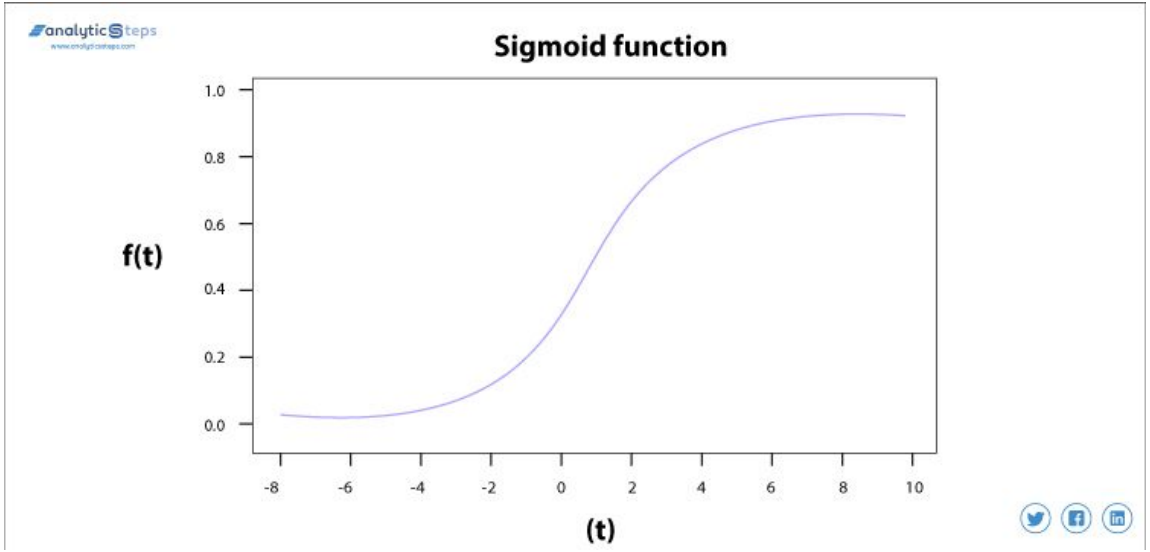
(a) the logistic or sigmoid function

(b) the ReLU function

(c) the tanh function.

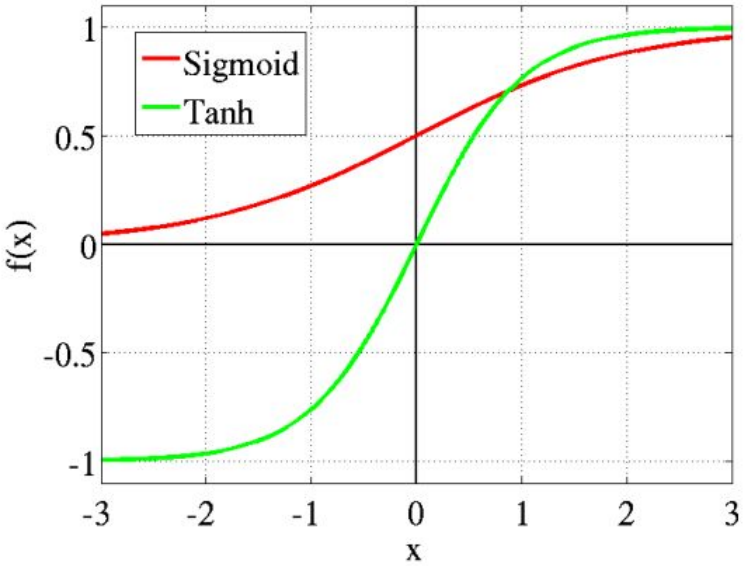
Sigmoid Functions

The sigmoid activation function is a probabilistic approach towards decision making and ranges in between 0 to 1.



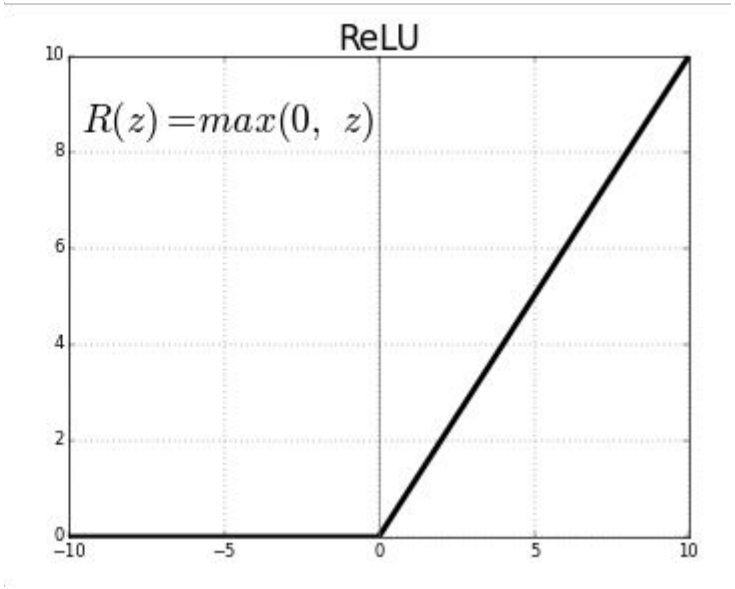
Tanh Function

Tanh is also like logistic sigmoid but better. The range of the tanh function is from (-1 to 1). tanh is also sigmoidal (s - shaped).



ReLU function

As you can see, the ReLU is half rectified (from bottom). f(z) is zero when z is less than zero and f(z) is equal to z when z is above or equal to zero.. The range of the ReLU function is from (0 to infinity)



Types of ANN

- Feed-forward,
- Single Layered,
- Multi-Layered,
- Recurrent

Types of ANN: Feedforward Neural Network

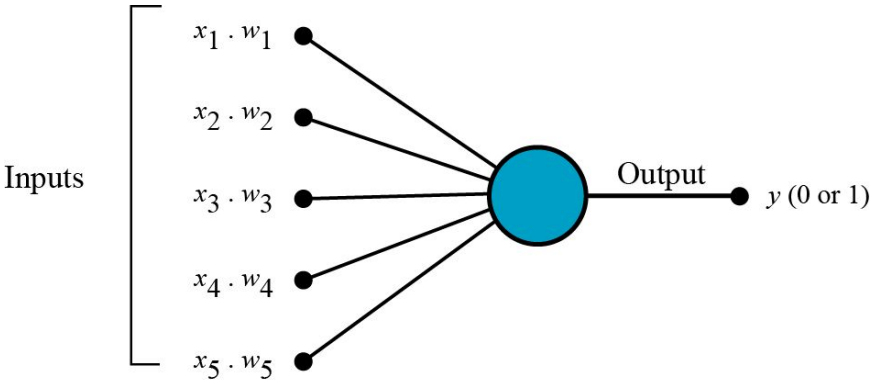
This neural network is one of the simplest forms of ANN, where the data or the input travels in one direction. The data passes through the input nodes and exit on the output nodes. This neural network may or may not have the hidden layers.

In simple words,The feed forward model is the simplest form of neural network as information is only processed in one direction. While the data may pass through multiple hidden nodes, it always moves in one direction and never backwards.

Types of ANN: Single Layer Feedforward Neural Network

In Single layer feed-forward network, the sum of the products of inputs and weights are calculated and fed to the output. The output is considered if it is above a certain value i.e threshold(usually 0) and the neuron fires with an activated output (usually 1) and if it does not fire, the deactivated value is emitted (usually -1).

A Feed Forward Neural Network is commonly seen in its simplest form as a single layer perceptron. In this model, a series of inputs enter the layer and are multiplied by the weights. Each value is then added together to get a sum of the weighted input values. If the sum of the values is above a specific threshold, usually set at zero, the value produced is often 1, whereas if the sum falls below the threshold, the output value is -1.



Types of ANN: Multi Layer Neural Network

A multi-layer neural network contains more than one layer of artificial neurons or nodes. They differ widely in design. It is important to note that while single-layer neural networks were useful early in the evolution of AI, the vast majority of networks used today have a multi-layer model. Basically, by adding more hidden layers / more neurons per layer you add more parameters to the model. Hence you **allow the model to fit more complex functions**.

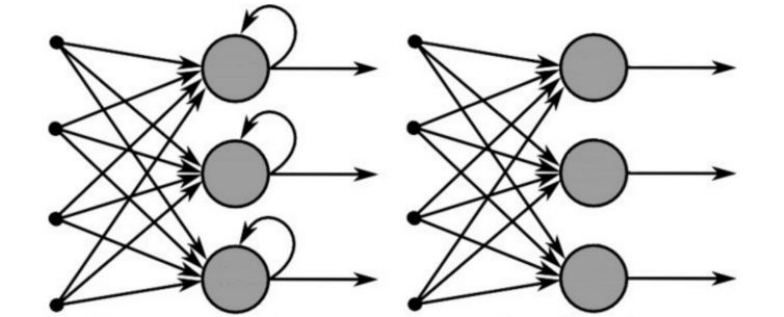
Examples of Multi layer Neural Networkare:

- Convolutional neural networks (CNNs)
- Recurrent neural networks,
- Deep networks and deep belief systems

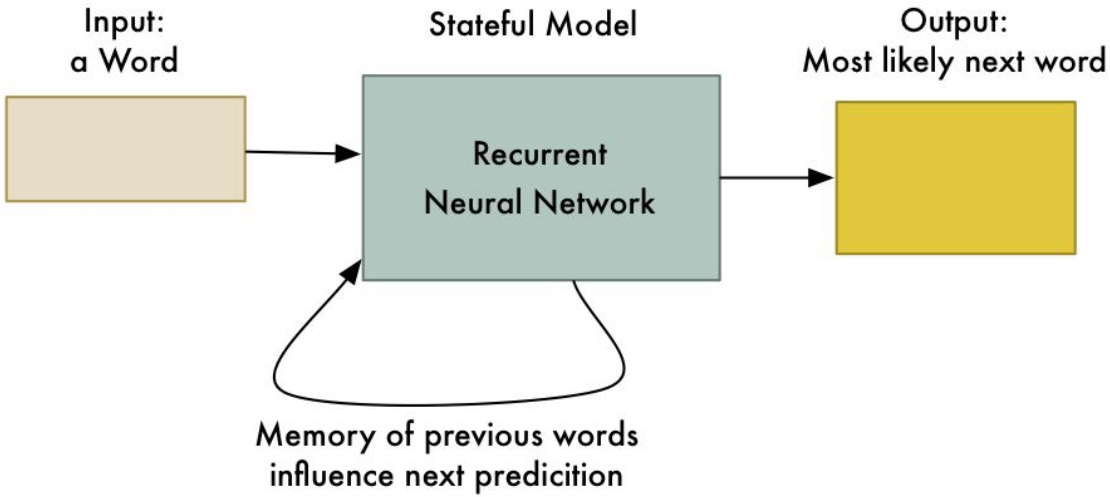
Types of ANN: Recurrent Neural Network

Recurrent neural networks (RNN) are **the state of the art algorithm for sequential data** and are used by Apple's Siri and and Google's voice search. It is the first algorithm that remembers its input, due to an internal memory, which makes it perfectly suited for machine learning problems that involve sequential data.

Recurrent neural networks recognize data's sequential characteristics and use patterns to predict the next likely scenario.



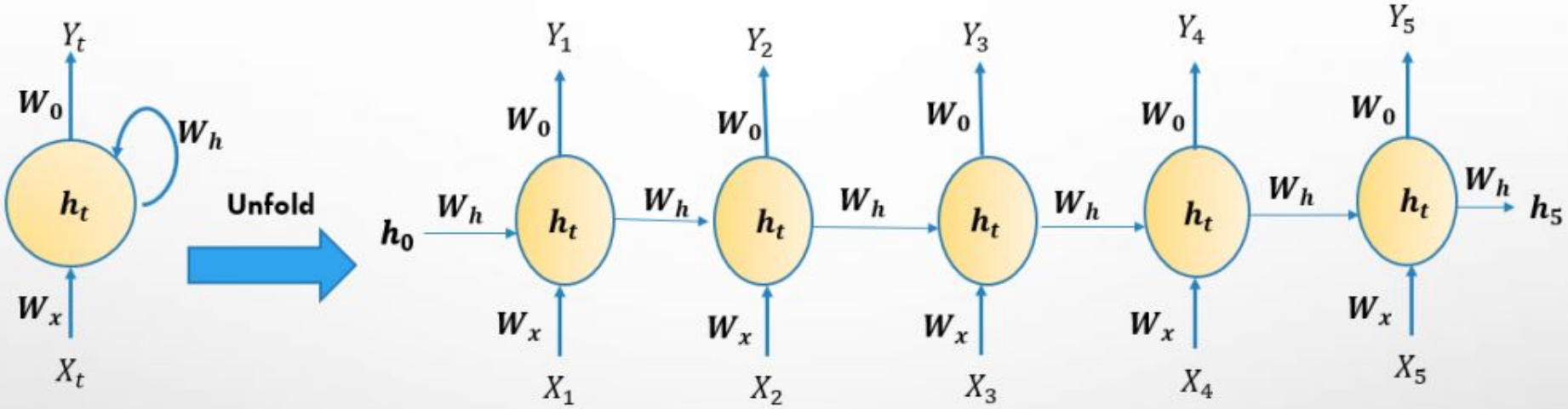
Example



Output so far:

Machine

UNFOLDING RNN



Recurrent Neural Network(RNN) – Long Short Term Memory:

Here, the first layer is formed similar to the feed forward neural network with the product of the sum of the weights and the features. The recurrent neural network process starts once this is computed, this means that from one time step to the next each neuron will remember some information it had in the previous time-step.

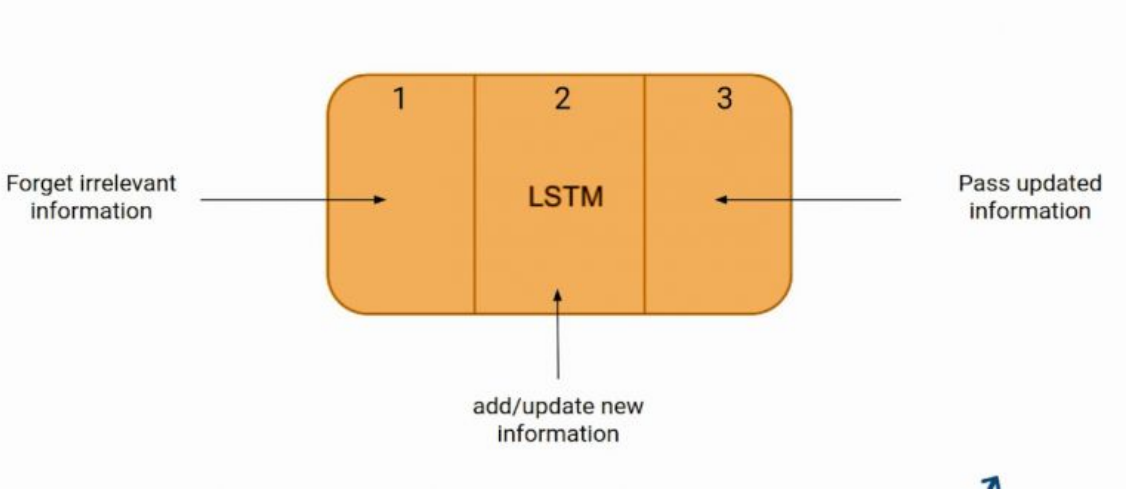
This makes each neuron act like a memory cell in performing computations. In this process, we need to let the neural network to work on the front propagation and remember what information it needs for later use. Here, if the prediction is wrong we use the learning rate or error correction to make small changes so that it will gradually work towards making the right prediction during the back propagation.

Recurrent Neural Network(RNN) – Long Short Term Memory:

The first part chooses whether the information coming from the previous timestamp is to be remembered or is irrelevant and can be forgotten.

In the second part, the cell tries to learn new information from the input to this cell.

At last, in the third part, the cell passes the updated information from the current timestamp to the next timestamp.

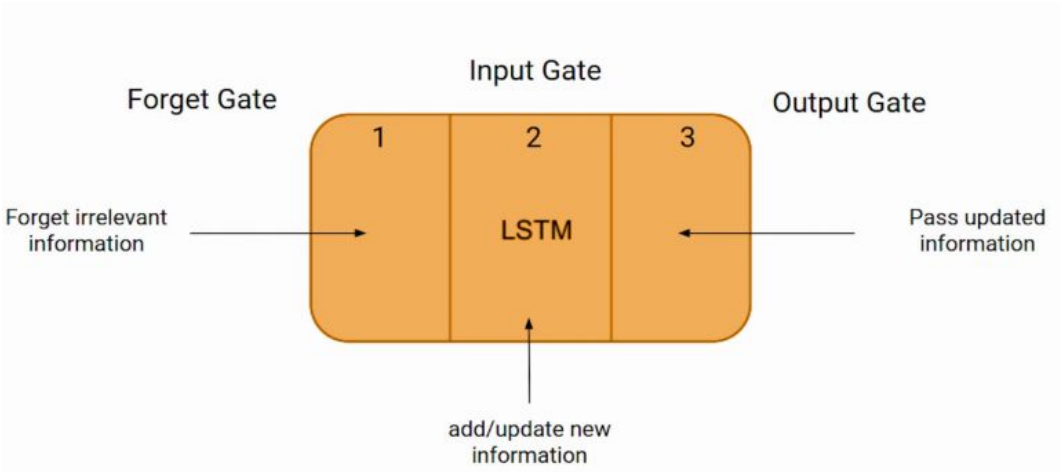


Recurrent Neural Network(RNN) – Long Short Term Memory:

These three parts of an LSTM cell are known as gates. The first part is called Forget gate, the second part is known as the Input gate and the last one is the Output gate.

Ex: Bob is a nice person. Dan, on the Other hand, is evil.
The first sentence is “Bob is a nice person” and the second sentence is “Dan, on the Other hand, is evil”. It is very clear, in the first sentence we are talking about Bob and as soon as we encounter the full stop(.) we started talking about Dan.

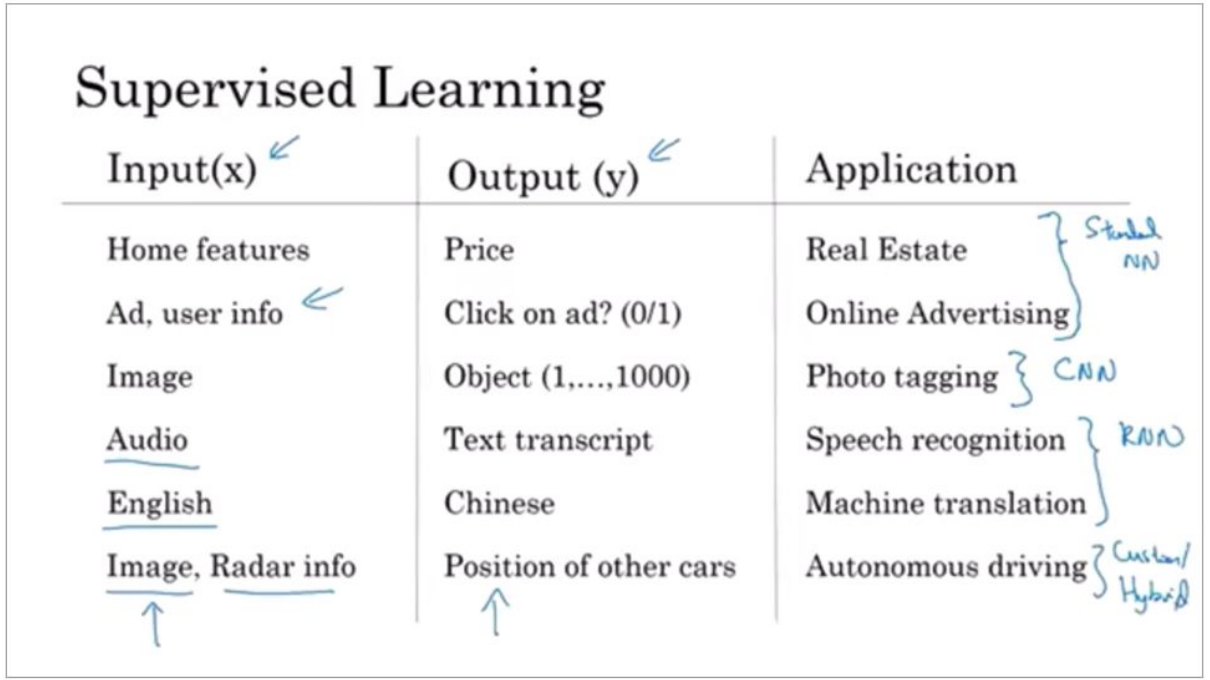
As we move from the first sentence to the second sentence, our network should realize that we are no more talking about Bob. Now our subject is Dan. Here, the Forget gate of the network allows it to forget about it. Let’s understand the roles played by these gates in LSTM architecture.



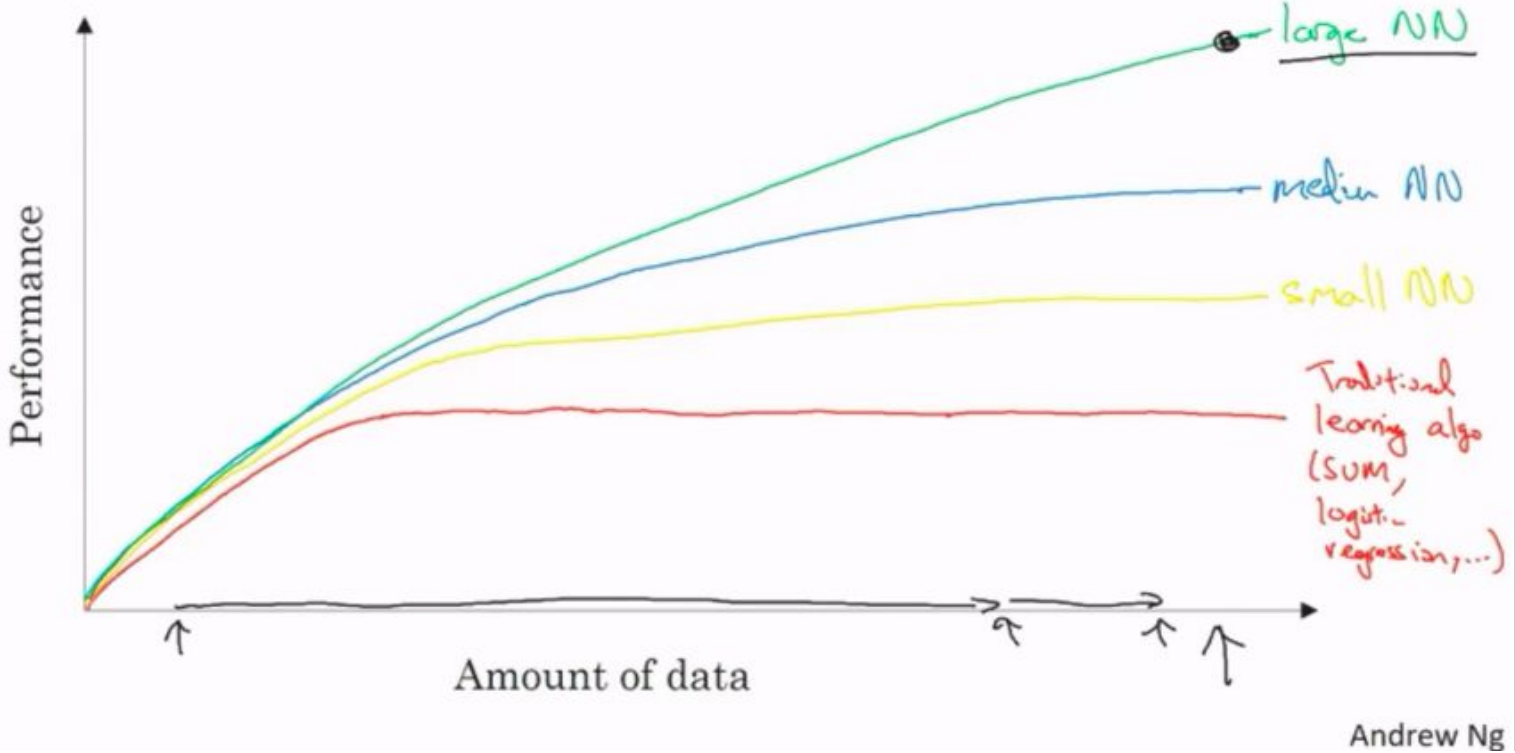
Application of Artificial Neural Networks

- Handwriting Recognition
- Image Compression
- Stock Exchange Prediction
- Text Classification
- Machine Translation
- Speech Recognition
- Facial Recognition

Algorithms used



Scale drives deep learning progress

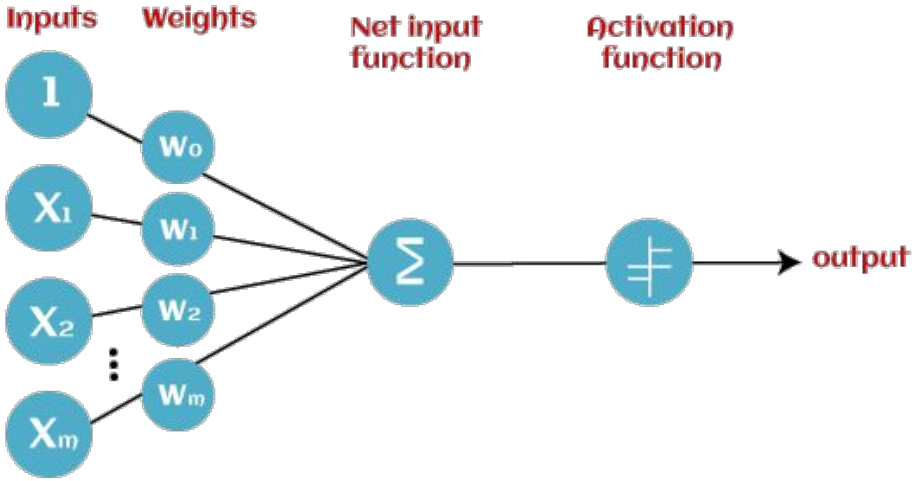


Learning Techniques in Neural Networks

- **Supervised Learning** In this learning, the training data is input to the network, and the desired output is known weights are adjusted until production yields desired value.
- **Unsupervised Learning** Use the input data to train the network whose output is known. The network classifies the input data and adjusts the weight by feature extraction in input data.
- **Reinforcement Learning** Here, the output value is unknown, but the network provides feedback on whether the output is right or wrong. It is Semi-Supervised Learning.
- **Offline Learning** The weight vector adjustment and threshold adjustment are made only after the training set is shown to the network. It is also called Batch Learning.
- **Online Learning** The adjustment of the weight and threshold is made after presenting each training sample to the network.

Perceptron Learning

In Machine Learning and Artificial Intelligence, Perceptron is the most commonly used term for all folks. It is the primary step to learn Machine Learning and Deep Learning technologies, which consists of a set of weights, input values or scores, and a threshold. ***Perceptron is a building block of an Artificial Neural Network***.



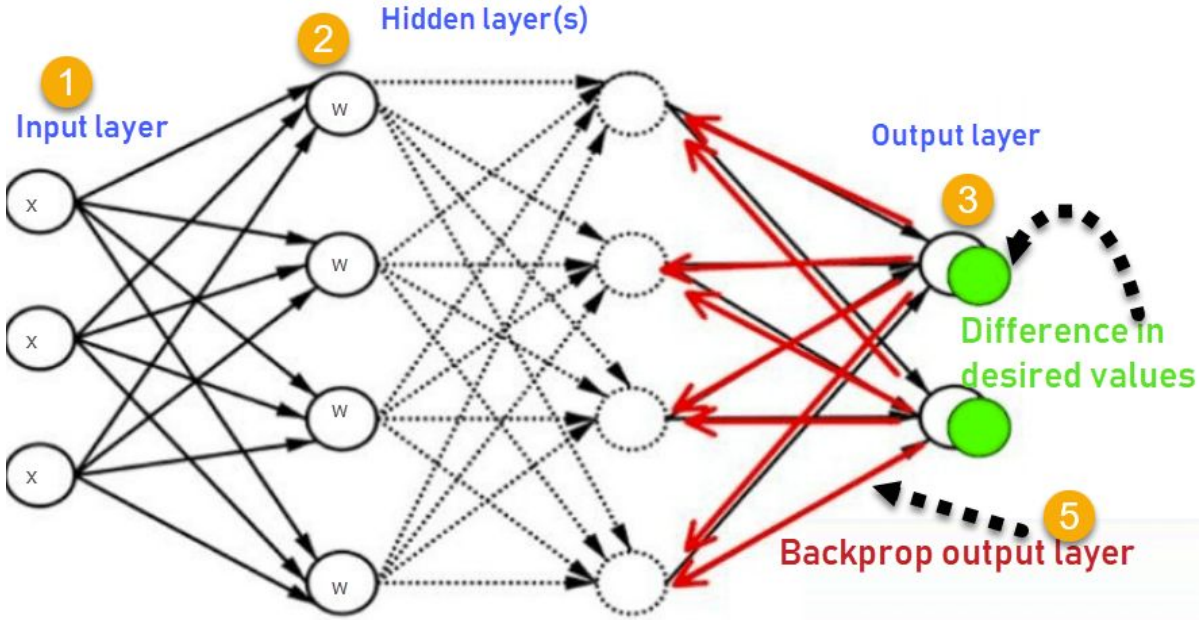
Types of Perceptron Models

Based on the layers, Perceptron models are divided into two types. These are as follows:

1. Single-layer Perceptron Model: Single layer perceptron model has one hidden layer.
2. Multi-layer Perceptron model: Like a single-layer perceptron model, a multi-layer perceptron model also has the same model structure but has a greater number of hidden layers. The multi-layer perceptron model is also known as the Backpropagation algorithm.

Back - propagation Learning

Backpropagation in neural network is a short form for “backward propagation of errors.” It is a standard method of training artificial neural networks. This method helps calculate the gradient of a loss function with respect to all the weights in the network.



Transfer Learning

Transfer learning is a machine learning method where a model developed for a task is reused as the starting point for a model on a second task.

Two common approaches are as follows:

1. Develop Model Approach
2. Pre-trained Model Approach

Refs

- https://www.xenonstack.com/blog/artificial-neural-network-applications
- https://vidyaesampally1998.medium.com/artificial-neural-network-v-s-biological-neural-network-a0862d12e9a8
- https://www.tutorialspoint.com/artificial_neural_network/artificial_neural_network_applications.htm
- http://neuralnetworksanddeeplearning.com/
- https://www.coursera.org/learn/neural-networks-deep-learning/home/week/1
- https://towardsdatascience.com/perceptron-learning-algorithm-d5db0deab975
- https://www.guru99.com/backpropogation-neural-network.html
- https://www.v7labs.com/blog/transfer-learning-guide#:~:text=In%20other%20words%2C%20transfer%20learning,when%20modeling%20the%20second%20task
- https://analyticsindiamag.com/6-types-of-artificial-neural-networks-currently-being-used-in-todays-technology/
- https://deepai.org/machine-learning-glossary-and-terms/feed-forward-neural-network .
- https://www.analyticsvidhya.com/blog/2021/03/introduction-to-long-short-term-memory-lstm/

Study Materials

Video Course: https://www.coursera.org/learn/neural-networks-deep-learning/home/week/1

Theory Course: http://neuralnetworksanddeeplearning.com/

Brief course: https://www.tutorialspoint.com/artificial_neural_network/artificial_neural_network_applications.htm