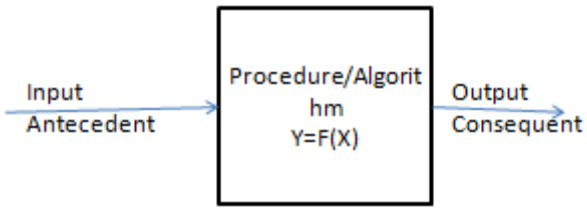
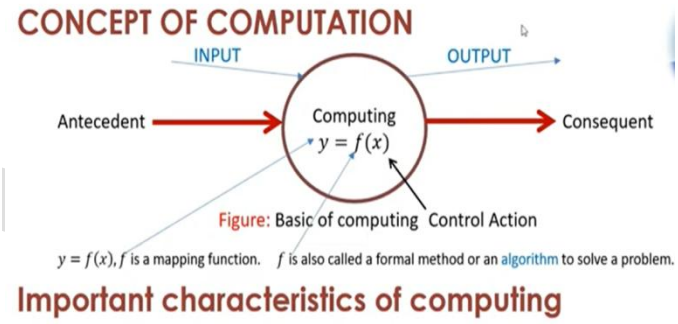
Hard computing

ante acurate values thesukuntadhi like 2.5 unte 3 ani 6.3unte 6 ani theskuntadhi. Precise results are guaranteed, control action is unambiguous, formally defined (that means, mathematical model or algorithm). Ex- Searching & sorting techniques, solving computational geometry, numerical problems (roots of polynomials, integration)

Soft Computing

- It is the study of science of reasoning, thinking, analyzing and detecting that correlates the real world problems to the biological inspired methods. It provides Optimal Solutions to the Problems.
  Soft computing ante accurate ga results evani vatiki manamu computing(calculations) apply cheyochu like float numbers
- Soft Computing is the fusion of methodologies designed to model and enable solutions to real world problems, which are not modeled or too difficult to model mathematically.
- The aim of Soft Computing is to exploit the tolerance for imprecision, uncertainty, approximate reasoning, and partial truth in order to achieve close resemblance with human like decision making.
- **Y=F(X)**, F is a mapping Function, Formal Method, Algorithm to solve the Function.





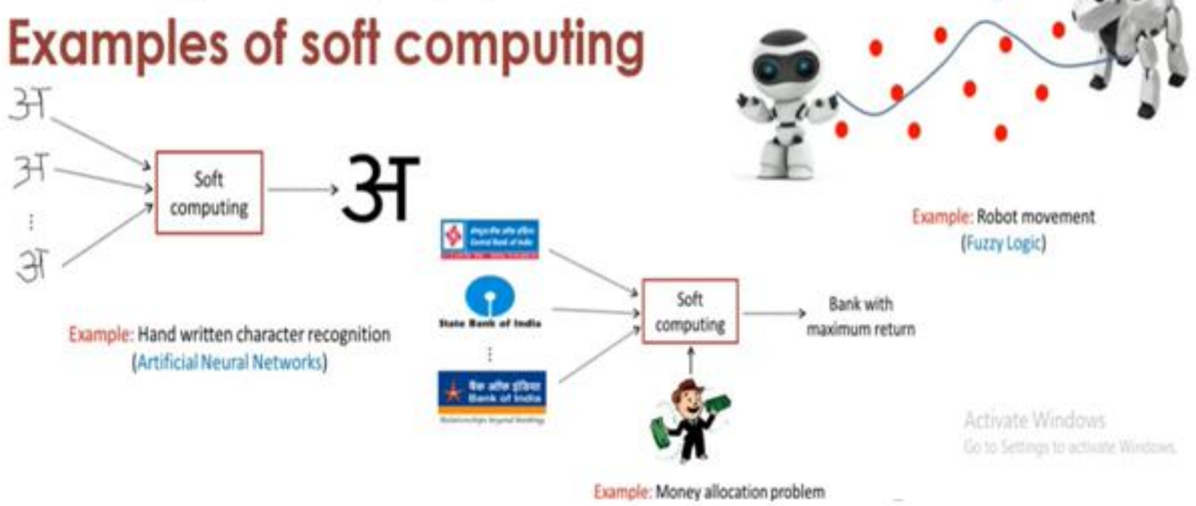
Hard Vs Soft Computing (Cont...)

Hard Computing	Soft Computing
Conventional computing requires a precisely stated analytical model.	Soft computing is tolerant of imprecision.
Often requires a lot of computation time.	Can solve some real world problems in reasonably less time.
Not suited for real world problems for which ideal model is not present.	Suitable for real world problems.
It requires full truth	Can work with partial truth
It is precise and accurate	Imprecise.
High cost for solution	Low cost for solution
Require programs to be written	Can evolve its own programs
Deterministic	Stochastic
Require exact input	Can deal with ambiguous and noisy data
Produce precise answer	Produce approximate answers

Hard computing	Soft computing
▪ It requires a precisely stated analytical model and often a lot of computation time.	▪ It is tolerant of imprecision, uncertainty, partial truth, and approximation.
▪ It is based on binary logic, crisp systems, numerical analysis and crisp software.	▪ It is based on fuzzy logic, neural nets and probabilistic reasoning.
▪ It has the characteristics of precision and categoricity.	▪ It has the characteristics of approximation and dispositionality.
▪ It is deterministic.	▪ It incorporates stochasticity.
▪ It requires exact input data.	▪ It can deal with ambiguous and noisy data.
▪ It is strictly sequential.	▪ It allows parallel computations.

<u>**Characteristics of soft computing**</u>

- It does not require any mathematical modeling of problem solving.
- It may not yield the precise solution.
- Algorithms are adaptive (i.e., it can adjust to the change of dynamic environment).
- Use some biological inspired methodologies such as genetics, evolution, Ant's behaviors, particles swarming, human nervous system, etc.).



<u>**Goals of Soft Computing**</u>

The main goal of Soft Computing is to develop intelligent machines to provide solutions to real world problems, which are not modelled, or too difficult to model mathematically. Its aim is to exploit the tolerance for :

- **Approximation** : here the model features are similar to the real ones, but not the same.
- **Uncertainty** : here we are not sure that the features of the model are the same as that of the entity (belief).

- **Imprecision and Partial** Truth in order to achieve close resemblance with human like decision making. Here the model features (quantities) are not the same as that of the real ones, but close to them.

Importance of Soft Computing

- Soft computing differs from hard (conventional) computing. Unlike hard computing, the soft computing is tolerant of imprecision, uncertainty, partial truth, and approximation
- The guiding principle of soft computing is to exploit these tolerance to achieve tractability, robustness and low solution cost. In effect, the role model for soft computing is the human mind.
- Soft computing is not a concoction, mixture, or combination, rather, Soft computing is a FUSION OF METHODOLOGIES THAT WORK SYNERGISTICALLY in one form or another, where each partner contributes a distinct methodology for addressing problems in its domain.
- In principal the constituent methodologies in Soft computing are complementary rather than competitive.
- Soft computing may be viewed as a foundation component for the emerging field of Conceptual Intelligence

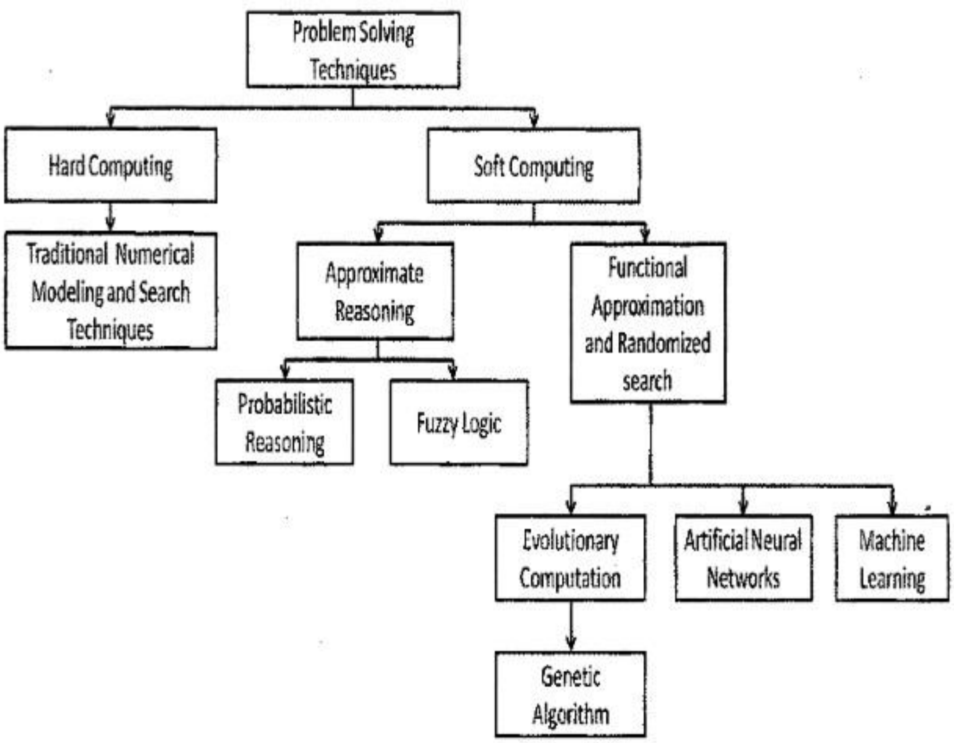
The four fields that constitute Soft Computing (SC) are :

- Fuzzy Computing (FC),
- Evolutionary Computing (EC),
- Neural computing (NC), and
- Probabilistic Computing (PC)

Properties of Soft Computing methods

These methods have in common: They

1. are non linear.
2. have ability to deal with non linearity.
3. follow more human like reasoning paths than classical methods.
4. utilize self learning.
5. utilize yet-to-be proven theorems.
6. are robust in the presence of noise or errors.



---

1. Artificial Neural network

- ANN is a parallel distributed information processing structure consisting of a number of nonlinear processing units called neurons.
- The neuron operates as a mathematical processor performing specific mathematical operations on its inputs to generate an output.
- It can be trained to recognize patterns and to identify incomplete patterns by resembling the human-brain processes of recognizing information, burying noise literally and retrieving information correctly.
- ANN are strongly interconnected systems of neurons which have simple behavior, but when connected they can solve complex problems. Changes may be made further to enhance its performance

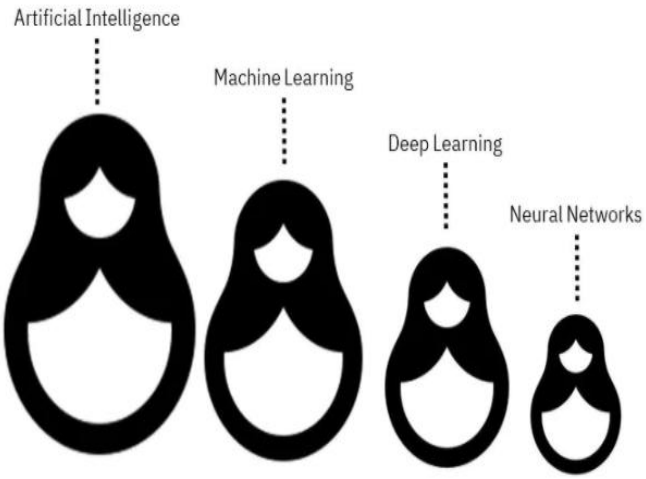
2. Genetic algorithms

- Evolutionary algorithms (EA) were invented to mimic some of the processes observed in natural evolution. Evolution occurs on chromosomes - organic devices for encoding the structure of living beings.
- Processes of natural selection then drive those chromosomes that encode successful structures to reproduce more frequent than those that encode failed structures. In other word, the chromosomes with the best evaluations tend to reproduce more often than those with bad evaluations.
- By using simple encodings and reproduction mechanisms, the algorithms can then display complicated behavior and turn out to solve some extremely difficult problems.
- Based on the principles of natural evolution, GAs are robust and adaptive methods to solve search and optimization problems.
- In addition, by simulating some features of biological evolution, GA can solve problems where traditional search and optimization methods are less effective.

3. Fuzzy Logic

- In the real world, information is often ambiguous or imprecise.
- When we state that it is warm today, the context is necessary to approximate the temperature.
- A warm day in January may be degrees Celsius, but a warm day in August may be 33 degrees. After a long spell of frigid days, we may call a milder but still chilly day relatively warm.
- Human reasoning filters and interprets information in order to arrive at conclusions or to dismiss it as inconclusive.
- Although machines cannot yet handle imprecise information in the same ways that humans do, computer programs with fuzzy logic are becoming quite useful when the sheer volume of tasks defines human analysis and action.
- An organized method for dealing with imprecise data is called fuzzy logic. The data sets engaged in fuzzy logic are considered as fuzzy sets.
- Traditional sets include or do not include an individual element; there is no other case than true or false. Fuzzy sets allow partial membership.
- Fuzzy Logic is basically a multi-valued logic that allows intermediate values to be defined between conventional evaluations like yes/no, true/false, black/white, etc.
- Notions like rather warm or pretty cold can be formulated mathematically and processed with the computer.
- In this way, an attempt is made to apply a more humanlike way of thinking in the programming of computers.
- Fuzzy logic is an extension of the classical propositional and predicate logic that rests on the principles of the binary truth functionality

---

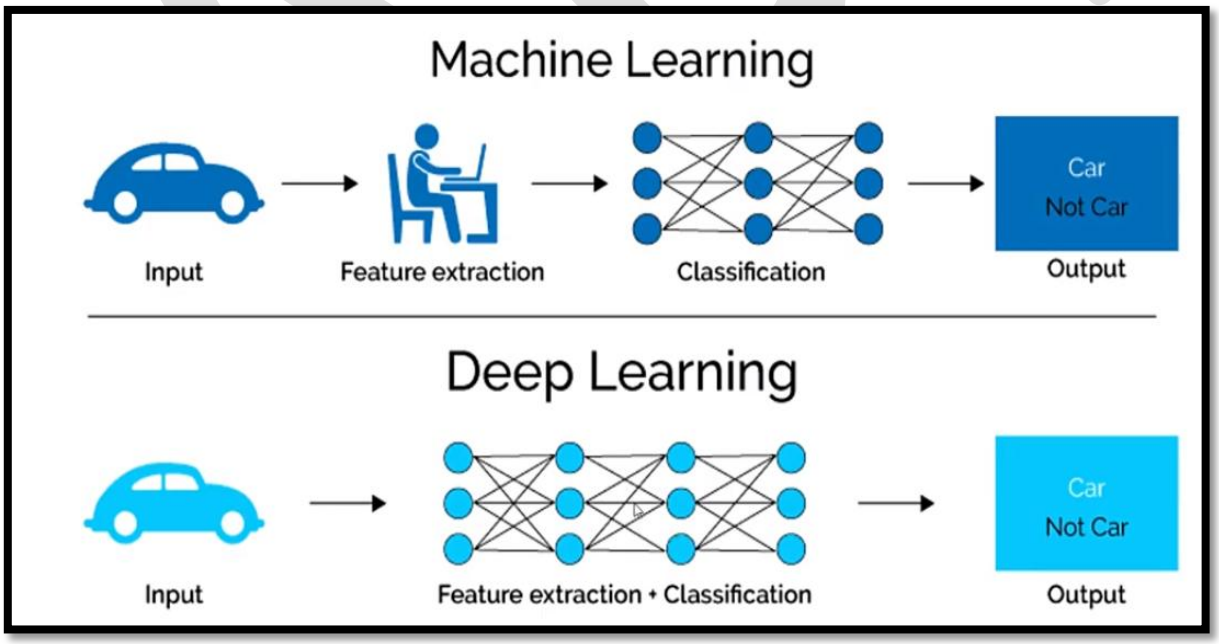


Artificial Intelligence - If the intelligence is demonstrated by the human it is called as natural intelligence ,else if it is demonstrated by the machines it is called as artificial intelligence i.e automation of intelligence behavior.

Machine Learning - A subset of artificial intelligence involved with the creation of algorithms which can modify itself without human intervention to produce desired output- by feeding itself through structured data.

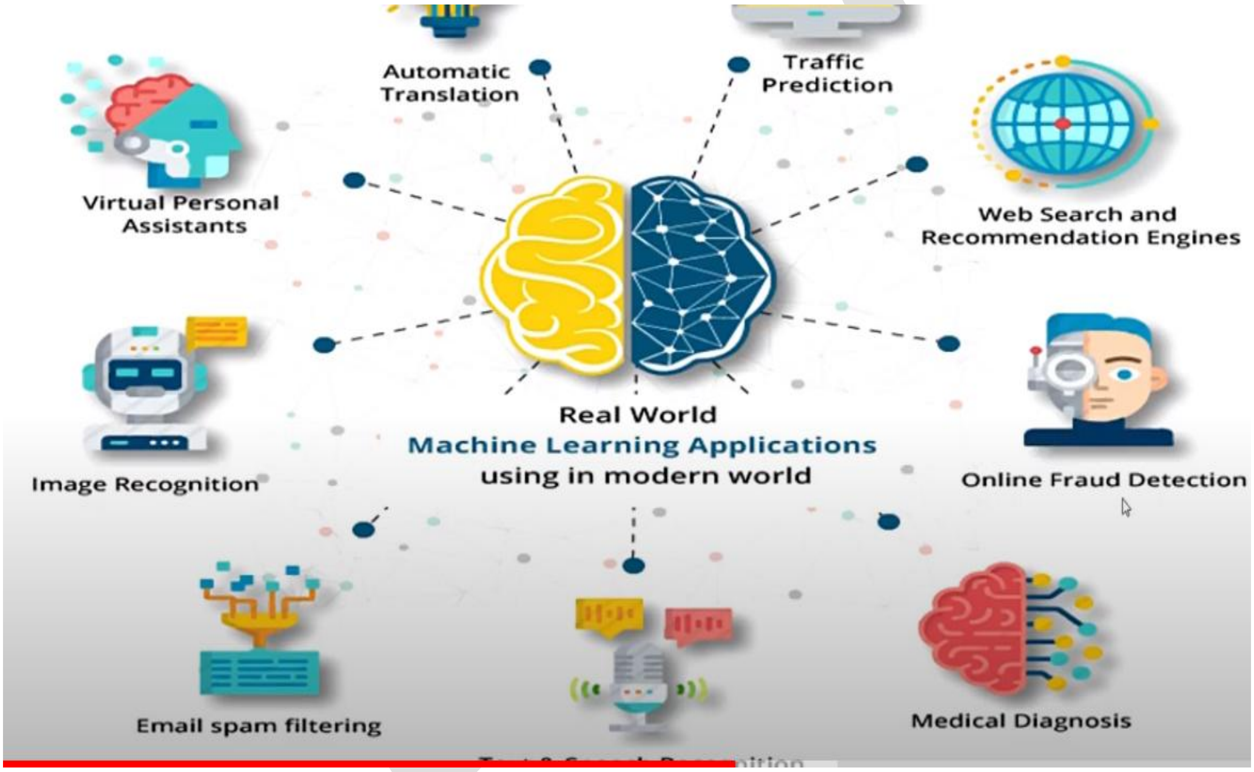
Deep learning - A subset of machine learning where algorithms are created and function similar to those in machine learning, but there are numerous layers of these algorithms- each providing a different interpretation to the data it feeds on. Such a network of algorithms are called artificial neural networks.

- DEEP LEARNING IS A CLASS OF MACHINE LEARNING ALGORITHMS THAT USES MULTIPLE LAYERS TO PROGRESSIVELY EXTRACT HIGHER LEVEL FEATURES FROM THE RAW INPUT.
- FOR EXAMPLE, IN IMAGE PROCESSING, LOWER LAYERS MAY IDENTIFY EDGES, WHILE HIGHER LAYERS MAY IDENTIFY THE CONCEPTS RELEVANT TO A HUMAN SUCH AS DIGITS OR LETTERS OR FACES.



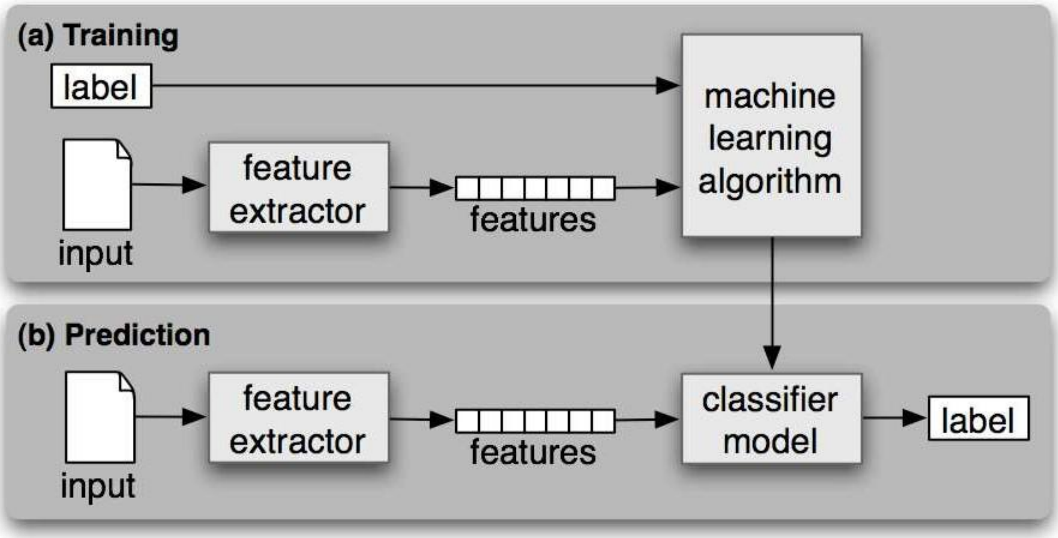
image, speech, text

Applications of machine learning and deep learning



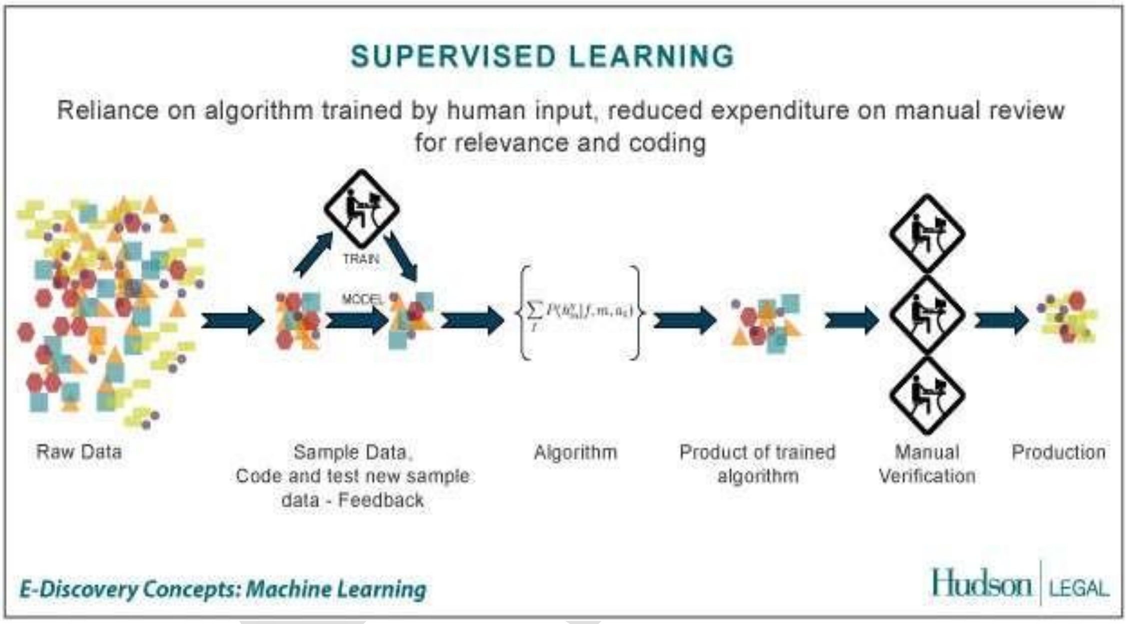
<u>ML TERMINOLOGIES</u>

- **Features**
  - ✓ The number of features or distinct traits that can be used to describe each item in a quantitative manner.
- **Samples**
  - ✓ A sample is an item to process (e.g. classify). It can be a document, a picture, a sound, a video, a row in database or CSV file, or whatever you can describe with a fixed set of quantitative traits.
- **Feature vector**
  - ✓ is an n-dimensional vector of numerical features that represent some object.
- **Feature extraction**
  - ✓ Preparation of feature vector
  - ✓ transforms the data in the high-dimensional space to a space of fewer dimensions.
- **Training/Evolution set**
  - ✓ Set of data to discover potentially predictive relationships.

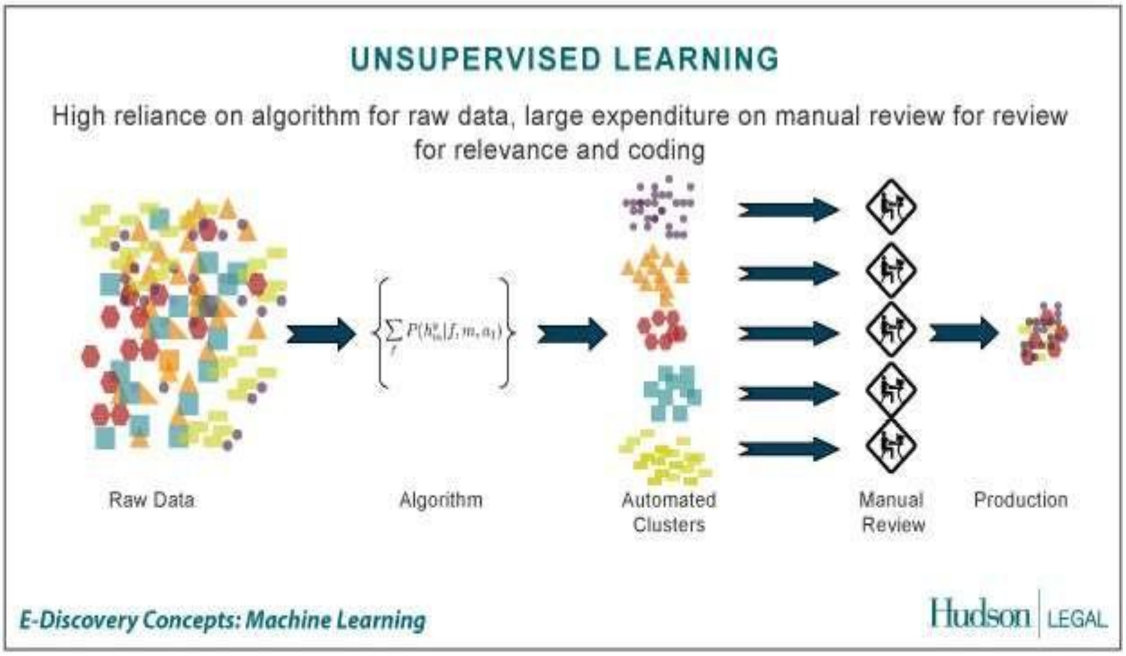


~~~~~~

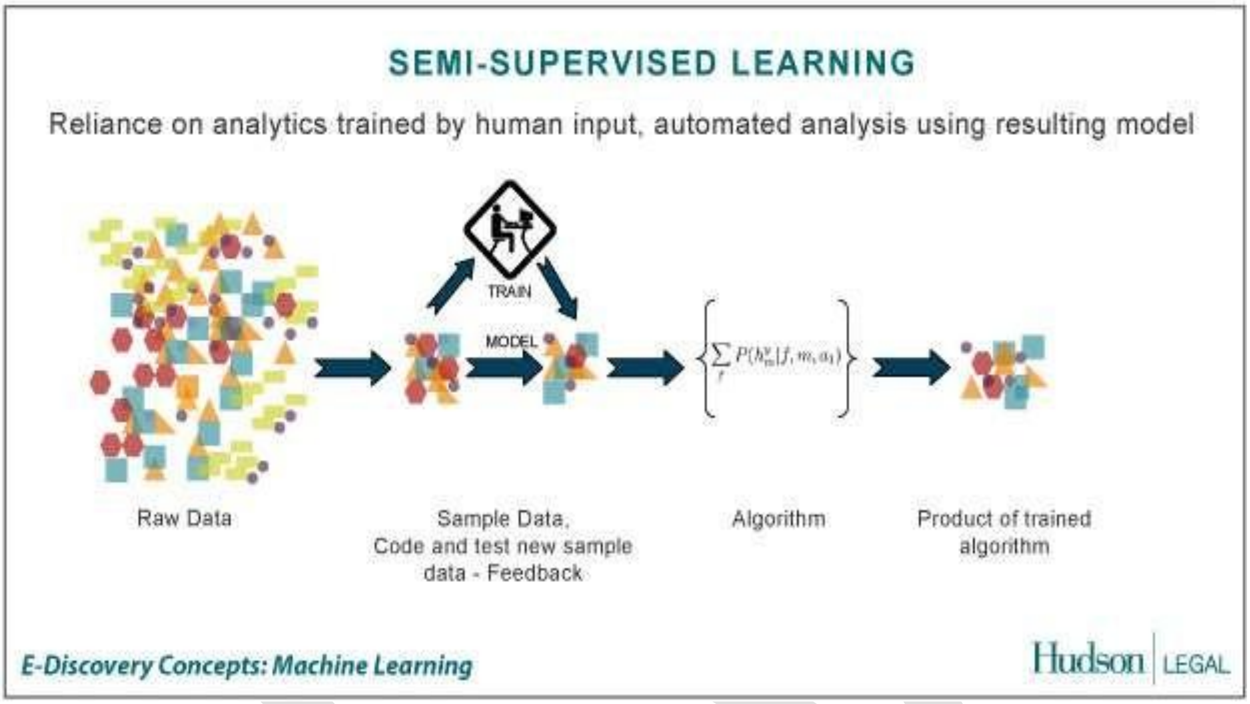
1. **Supervised Learning -**the correct classes of the training data are known



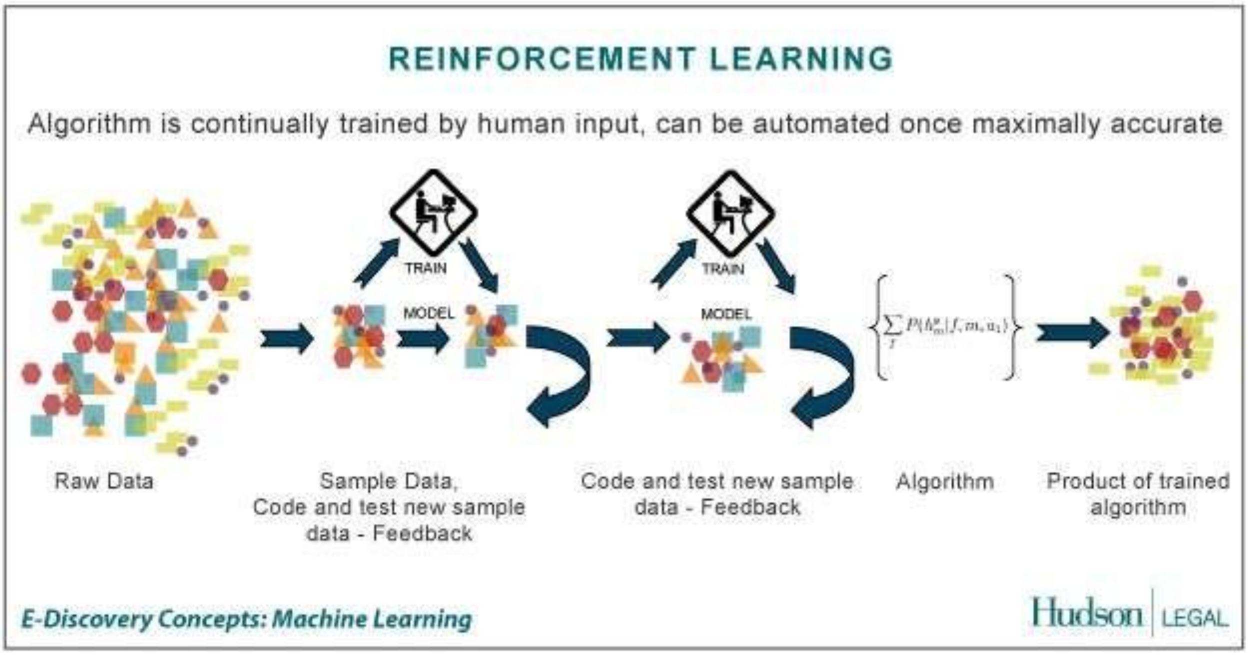
2. **Unsupervised Learning - the correct classes of the training data are not known**


~~~~~~

3. **Semi-Supervised Learning –** A Mix of Supervised and Unsupervised learning



4. **Reinforcement Learning –**
   - ✓ allows the machine or software agent to learn its behavior based on feedback from the environment.
   - ✓ This behavior can be learnt once and for all, or keep on adapting as time goes by.



~~~~~~

<u>**Machine Learning Techniques**</u>

- **classification**: predict class from observations
- **clustering**: group observations into "meaningful" groups
- **regression** (prediction): predict value from observations

~~~~~~

Subject	t	Feature 1	Feature 2	Target
Paul	1	1000	male	0
Paul	2	1100	male	0
Paul	3	1200	male	1
Paul	4	1300	male	1
Crista	4	20	female	0
Crista	5	100	female	0
Paulina	1	10000	female	0
Paulina	2	100000	female	1
Paulina	3	95000	female	1
Paulina	4	97000	female	1
Paulina	5	99000	female	1
Paulina	6	101000	female	1
George	1	50000	male	1
George	2	50000	male	1
George	3	50000	male	1
George	4	50000	male	1
George	5	50000	male	1
George	6	50000	male	1

Subject	t	Feature 1	Feature 2	Target
Paul	1	1000	male	0
Paulina	1	10000	female	0
George	1	50000	male	1
Paul	2	1100	male	0
Paulina	2	100000	female	1
George	2	50000	male	1
Paul	3	1200	male	1
Paulina	3	95000	female	1
George	3	50000	male	1
Paul	4	1300	male	1
Crista	4	20	female	0
Paulina	4	97000	female	1
George	4	50000	male	1
Crista	5	100	female	0
Paulina	5	99000	female	1
George	5	50000	male	1
Paulina	6	101000	female	1
George	6	50000	male	1

~~~~~~

CLASSIFICATION

- classify a document into a predefined category.
- documents can be text, images
- Popular one is Naive Bayes Classifier.
- Steps:
  - ✓ Step1 : Train the program (Building a Model) using a training set with a category for e.g. sports, cricket, news, Classifier will compute probability for each word, the probability that it makes a document belong to each of considered categories
  - ✓ Step2 : Test with a test data set against this Model

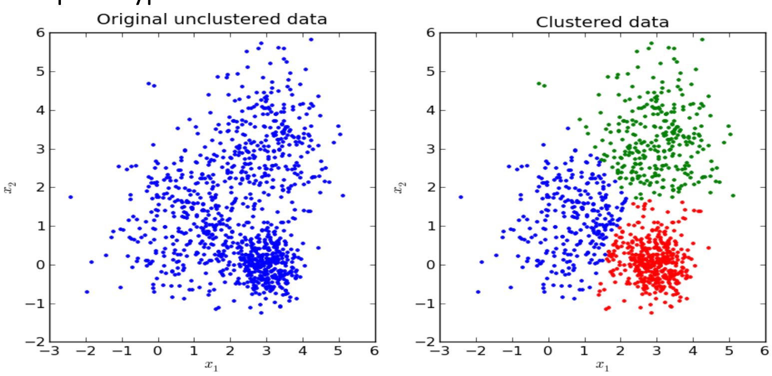
~~~~~~

CLUSTERING

- **clustering** is the task of grouping a set of objects in such a way that objects in the same group (called a **cluster**) are more similar to each other
- objects are not predefined
- For e.g. these keywords
  - ✓ "man's shoe"
  - ✓ "women's shoe"
  - ✓ "women's t-shirt"
  - ✓ "man's t-shirt"
  - ✓ can be cluster into 2 categories "shoe" and "t-shirt" or "man" and "women"
- Popular ones are **K-means clustering** and **Hierarchical clustering**

~~~~~~

K-means Clustering - partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster.

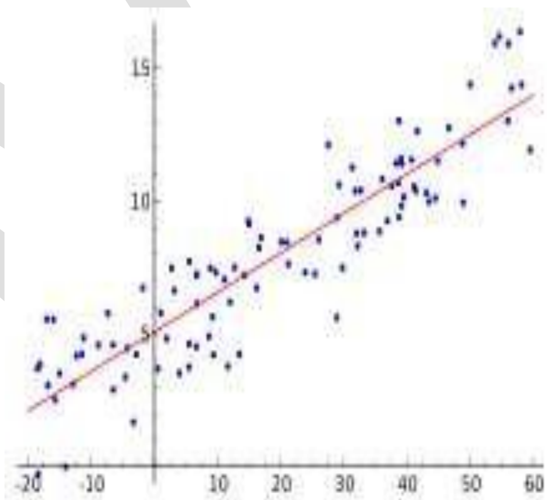

~~~~~~

<u>Hierarchical clustering</u>

- method of cluster analysis which seeks to build a hierarchy of clusters.
- There can be two strategies
  - **Agglomerative**:
    - ✓ This is a "bottom up" approach: each observation starts in its own cluster, and pairs of clusters are merged as one moves up the hierarchy.
    - ✓ Time complexity is O(n^3)
  - **Divisive**:
    - ✓ This is a "top down" approach: all observations start in one cluster, and splits are performed recursively as one moves down the hierarchy.
    - ✓ Time complexity is O(2^n)

<u>Regression</u>

- is a measure of the relation between the mean value of one variable (e.g. output) and corresponding values of other variables (e.g. time and cost).
- **regression analysis** is a statistical process for estimating the relationships among variables.
- Regression means to **predict** the output value using training data.
- Popular one is Logistic regression (binary regression)



<u>Regression vs Classification</u>

- Regression means to predict the output value using training data.
- regression to **predict** the house price from training data
- if it is a real number/continuous, then it is regression problem.

- Classification means to group the output into a class.
- classification to **predict** the type of tumor i.e. harmful or not harmful using training data
- if it is discrete/categorical variable, then it is classification problem

---

<u>Use-Cases of ML in real life</u>

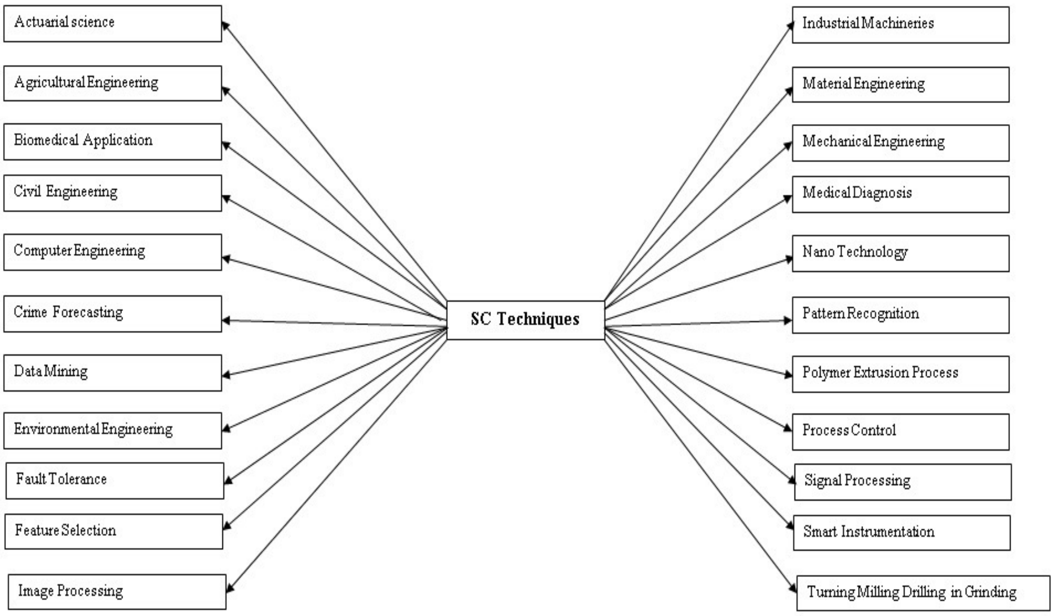
- Spam Email Detection
- Machine Translation (Language Translation)
- Image Search (Similarity)
- Clustering (KMeans) : Amazon Recommendations
- Classification : Google News
- Text Summarization - Google News
- Rating a Review/Comment: Yelp
- Fraud detection : Credit card Providers
- Decision Making : e.g. Bank/Insurance sector
- Sentiment Analysis
- Speech Understanding – iPhone with Siri
- Face Detection – Facebook's Photo tagging

<u>Popular Frameworks/Tools</u>

- Weka
- Carrot2
- Gate
- OpenNLP
- LingPipe
- Stanford NLP

- Mallet – Topic Modelling
- Gensim – Topic Modelling (Python)
- Apache Mahout
- MLib – Apache Spark
- scikit-learn - Python
- LIBSVM : Support Vector Machines

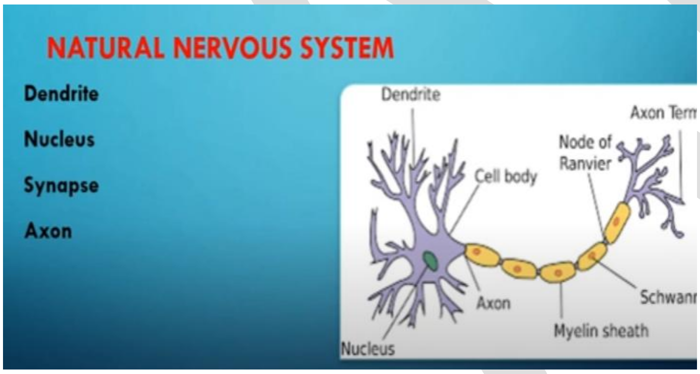
Application areas of soft computing



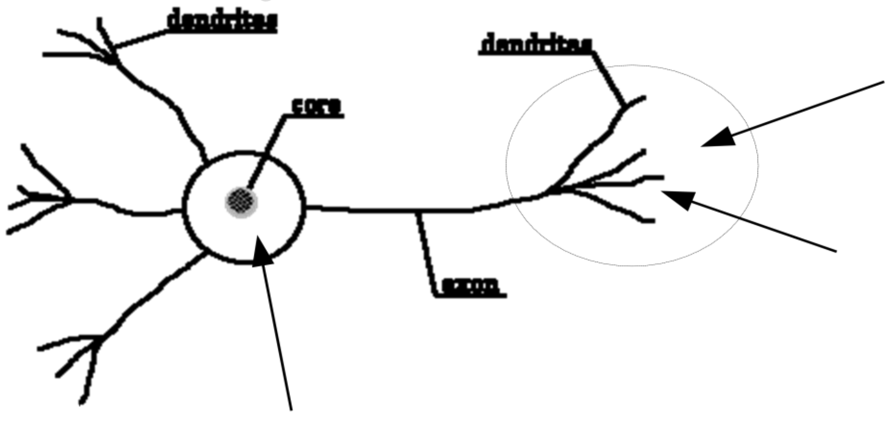
- Heavy industry - Robotic arms, Humanoid robots
- Home appliances - Washing machines, ACs, Refrigerators, cameras
- Automobiles - Travel Speed Estimation, Sleep Warning Systems, Driver-less cars
- Spacecrafts - Maneuvering of a Space Shuttle(FL), Optimization of Fuel- efficient Solutions for space craft

---

NATURAL NERVOUS SYSTEM

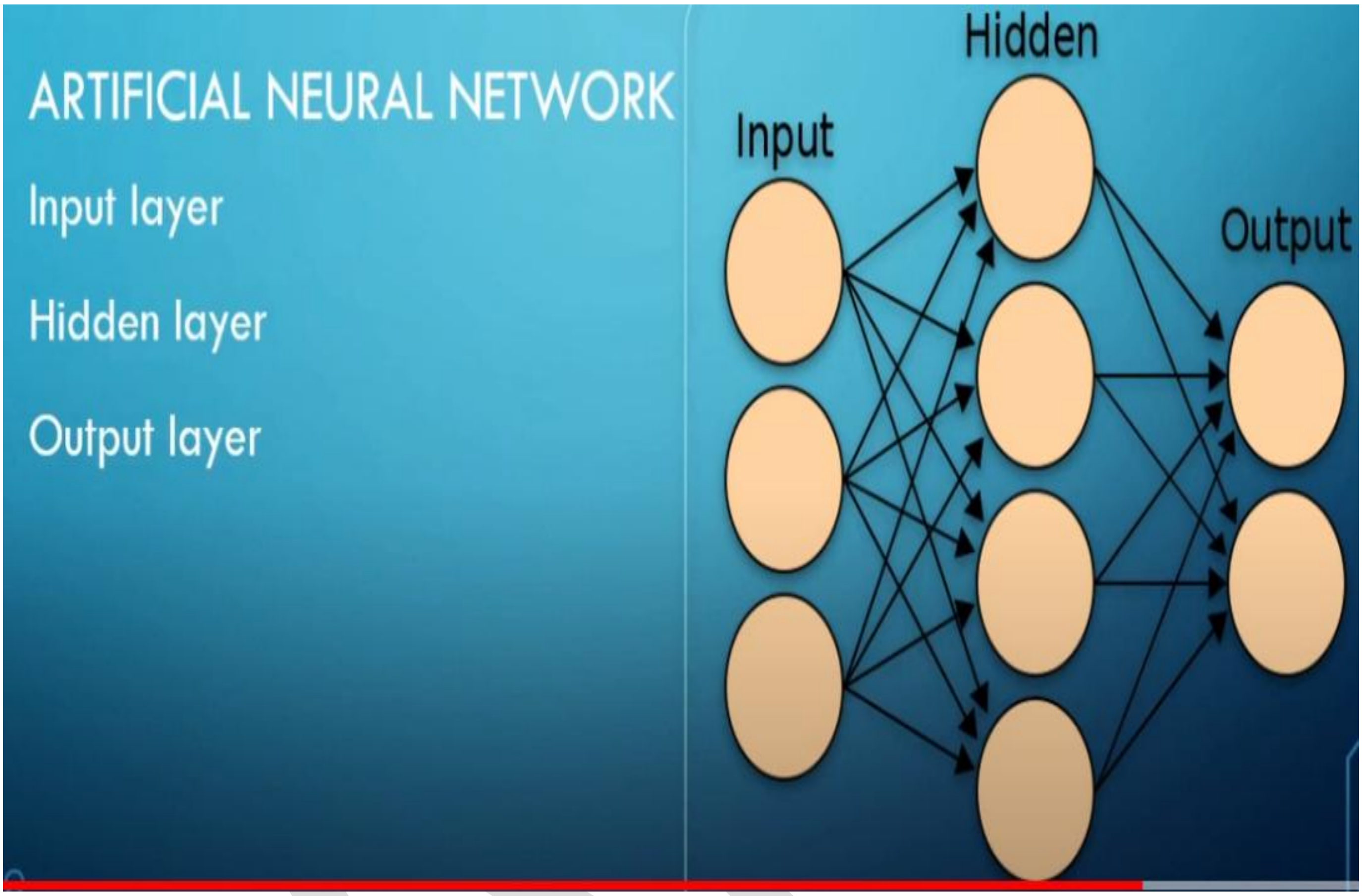


- any biological nervous system consists of a large number of interconnected processing units called neurons.
- Each neuron is approximately 10μm long and they can operate in parallel.
- Typically, a human brain consists of approximately 10^{11} neurons communicating with each other with the help of electrical impulses.



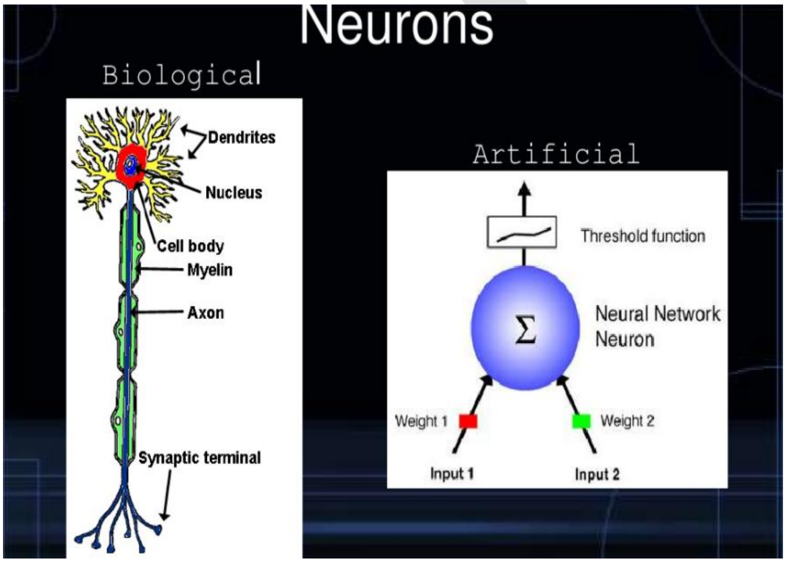
Dendrite : A bush of very thin fibre.
Axon : A long cylindrical fibre.
Soma : It is also called a cell body, and just like as a nucleus of cell.
Synapse : It is a junction where axon makes contact with the dendrites of neighboring dendrites

ARTIFICIAL NEURAL NETWORKS

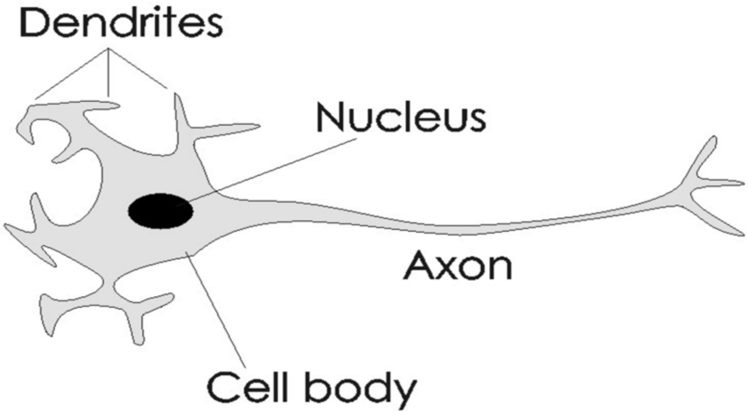


Artificial neural networks (ANNs) or simply we refer it as neural network (NNs), which are simplified models (i.e. imitations) of the biological nervous system, and obviously, therefore, have been motivated by the kind of computing performed by the human brain.
The behavior of a biolgical neural network can be captured by a simple model called Artificial neural network.

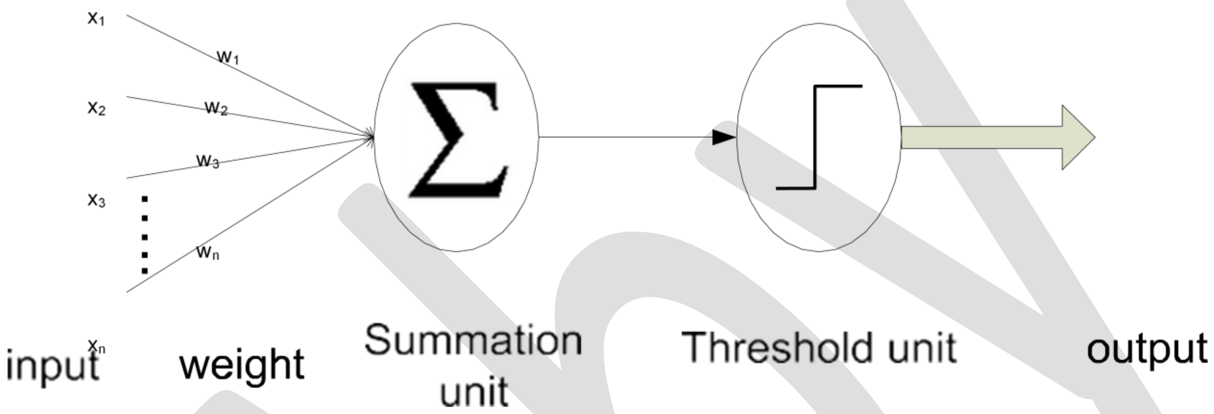
BIOLOGICAL NEURONS VS ARTIFICIAL NEURAL NETWORK



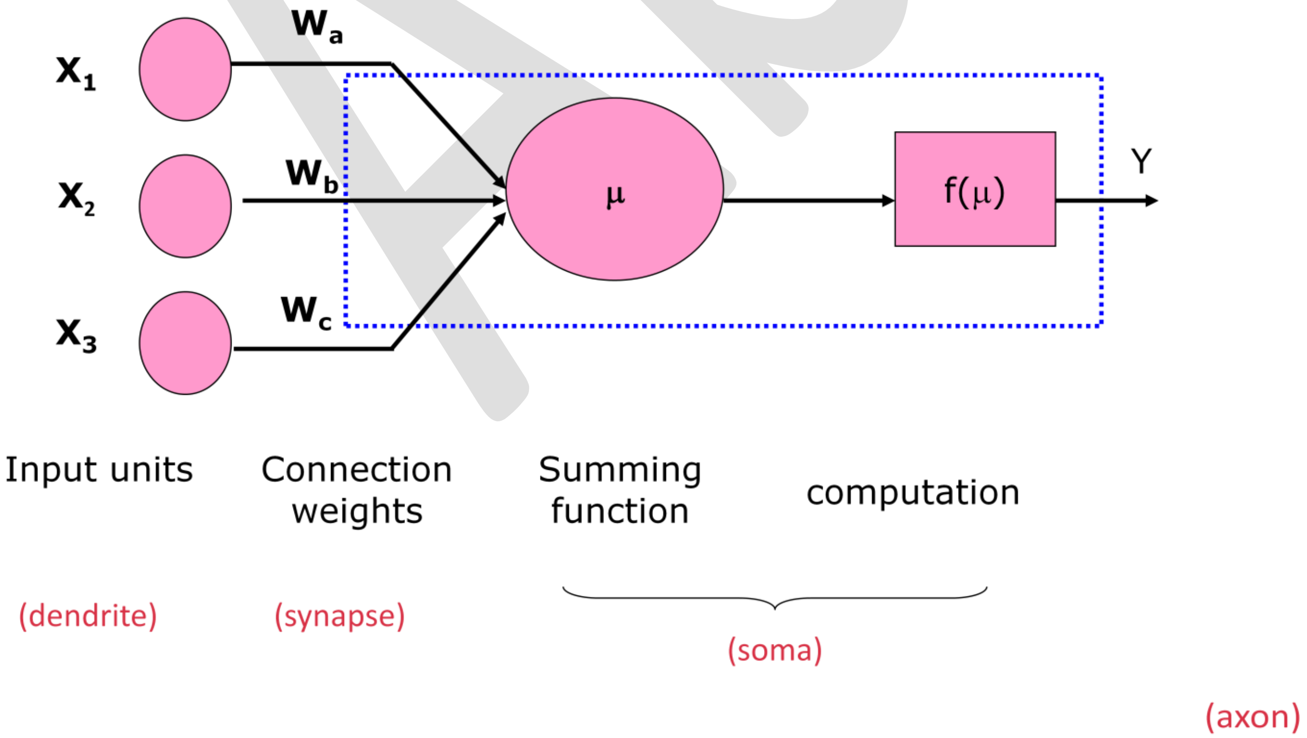
BIOLOGICAL NEURON



Artificial Neural Network



Here, x1, x2, · · · , xn are the n inputs to the artificial neuron. | w1, w2, · · · , wn are weights attached to the input links.



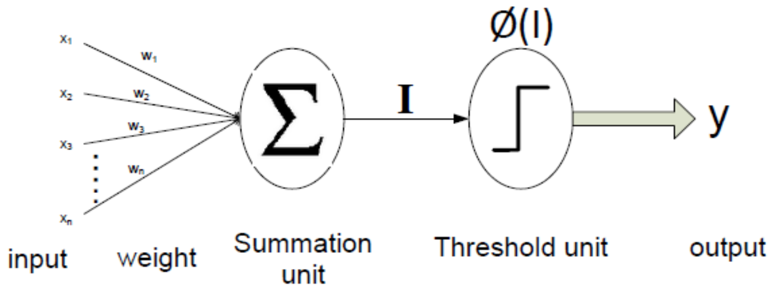
~~~~~~

- Note that, a biological neuron receives all inputs through the dendrites, sums them and produces an output if the sum is greater than a threshold value.
~~~~~~

- The input signals are passed on to the cell body through the synapse, which may accelerate or retard an arriving signal.
- It is this acceleration of the input signals that is modeled by the weights.
- An effective synapse, which transmits a stronger signal will have a correspondingly larger weights while a weak synapse will have smaller weights.
- Thus, weights here are multiplicative factors of the inputs to account for the strength of the synapse.
- Hence, the total input say I received by the soma of the artificial neuron is

$$I = w_1x_1 + w_2x_2 + \cdots + w_nx_n = \sum_{i=1}^{n} w_ix_i$$

- To generate the final output y, the sum is passed to a filter ϕ called transfer function, which releases the output.
- That is, $y = \phi(I)$



A very commonly known transfer function is the thresholding function. In this thresholding function, sum (i.e. I) is compared with a threshold value θ.If the value of I is greater than θ, then the output is 1 else it is 0 (this is just like a simple linear filter).

In other words,

$$y = \varphi\left(\sum_{i=1}^{n} w_i x_i - \theta\right)$$

$$\begin{matrix} 1 & , \text{if } I > \theta \\ 0 & , \text{if } I \leq \theta \end{matrix}$$

$$y = \phi\left(\sum_{i=1}^{n} w_ix_i - \theta\right)$$

where

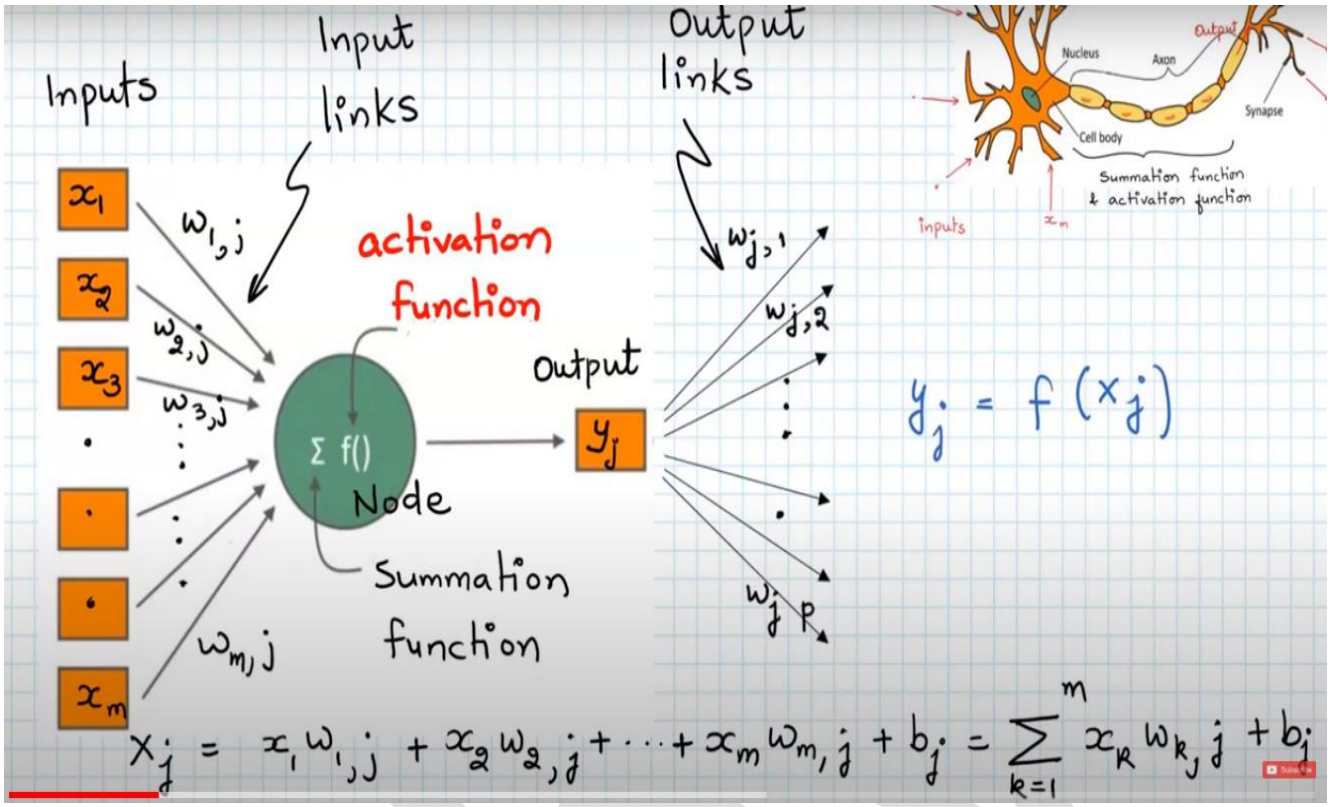
$$\phi(I) = \begin{cases} 1 & , \text{if } I > \theta \\ 0 & , \text{if } I \leq \theta \end{cases}$$

Such a Φ is called step function (also known as Heaviside function).

~~~~~~

ACTIVATION FUNCTION

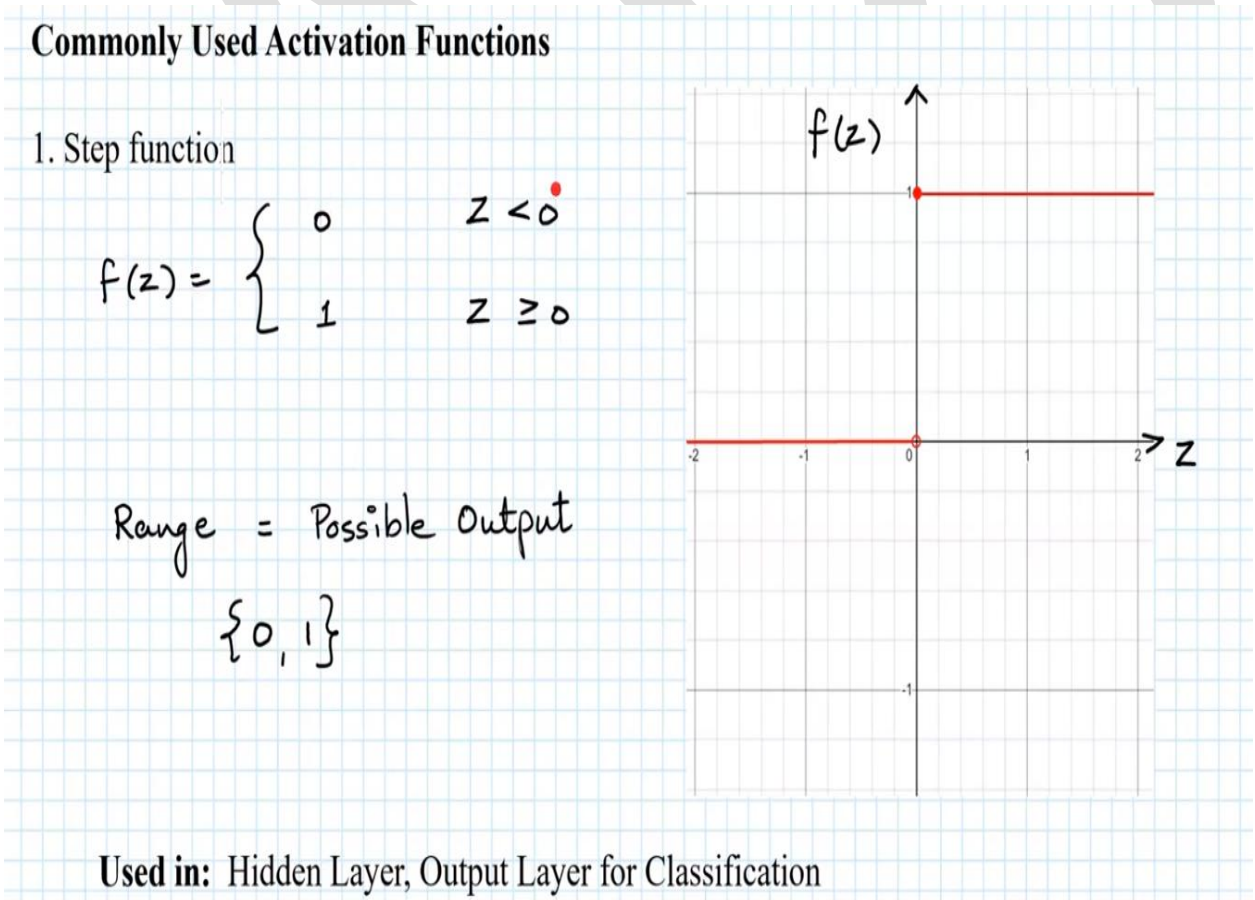
In artificial neural networks, the activation function of a node defines the output of that node given an input or set of inputs. A standard integrated circuit can be seen as a digital network of activation functions that can be "ON" or "OFF", depending on input.
~~~~~~



TYPES / EXAMPLE OF ACTIVATION FUNCTION

1. Identity - f(x) = x

2. Step Function



3. Binary step - f(x) = 1 if x >= q
f(x) = 0 otherwise

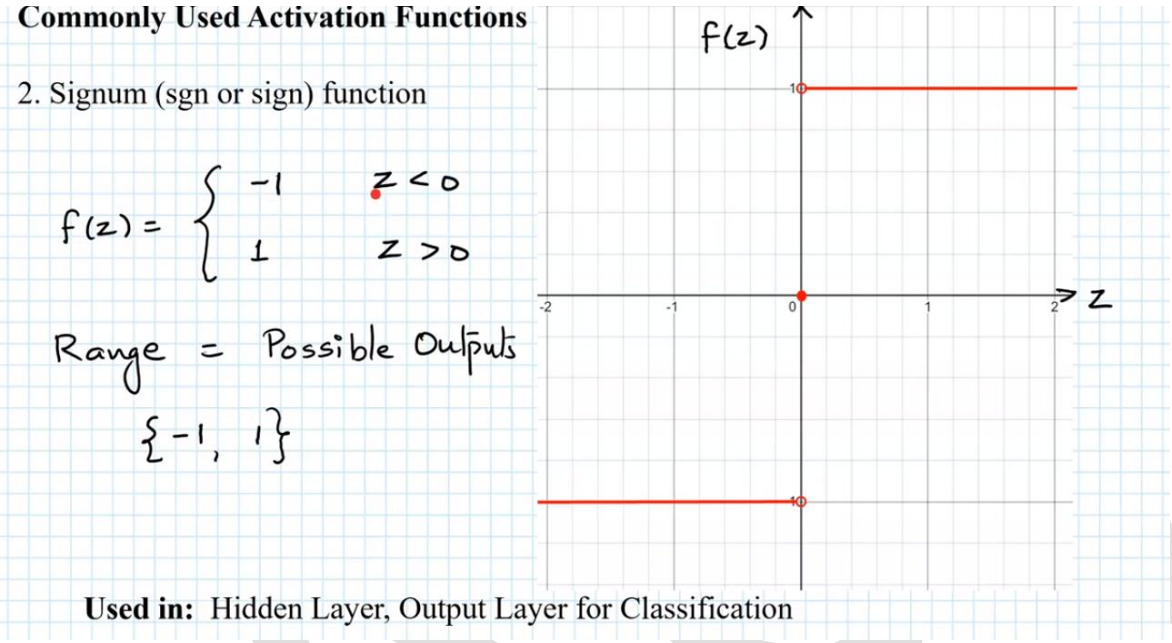
4. Binary sigmoid - f(x) = 1 / (1 + e-sx)

5. Bipolar sigmoid - f(x) = -1 + 2 / (1 + e-sx)

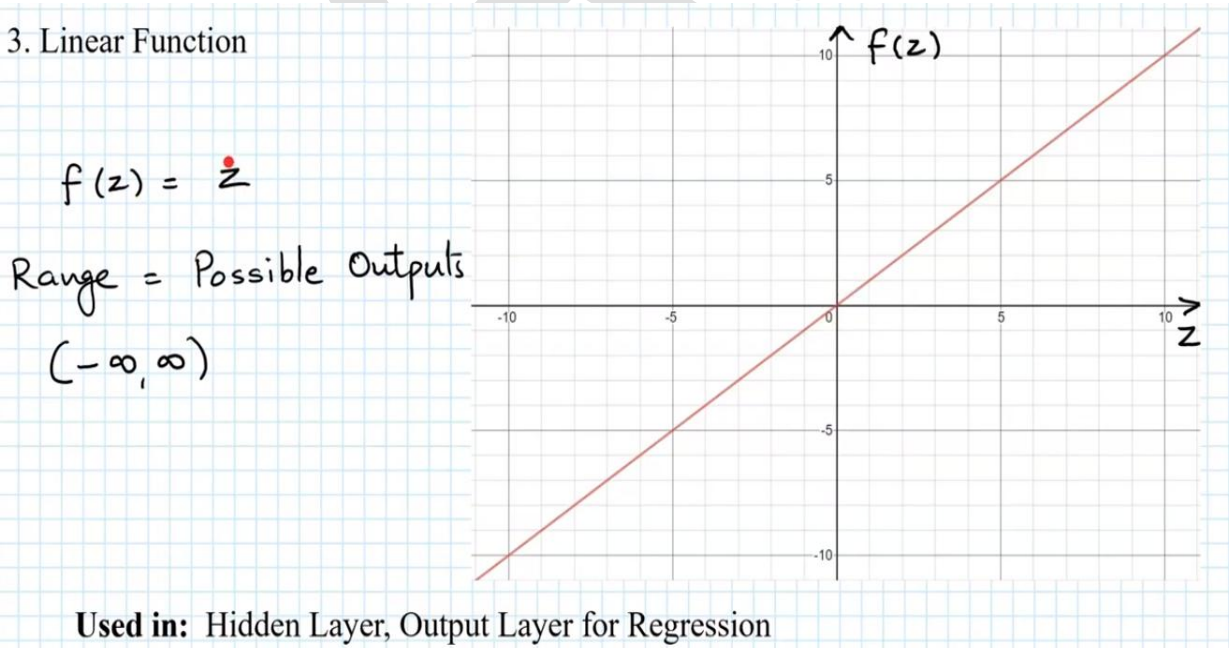
6. Hyperbolic tangent - f(x) = (ex – e-x) / (ex + e-x)

7. Hyperbolic function - tanh

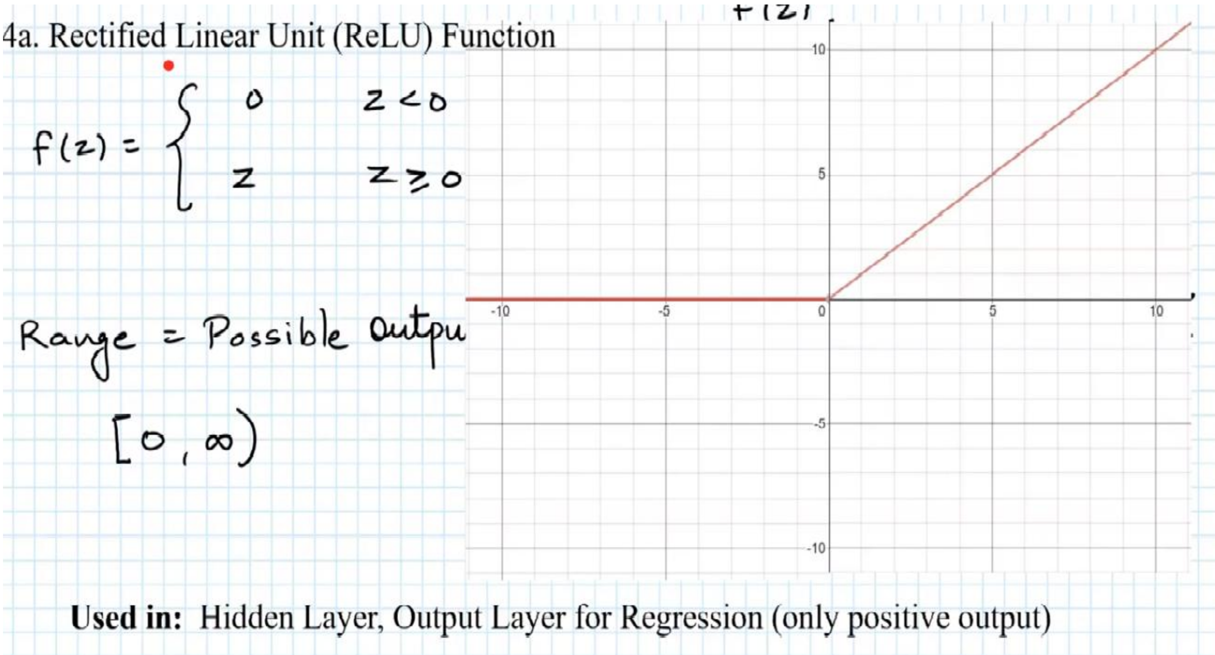
8. Signum (sgn or sign) function



9. Linear Activation Function



10. Relu Activation Function



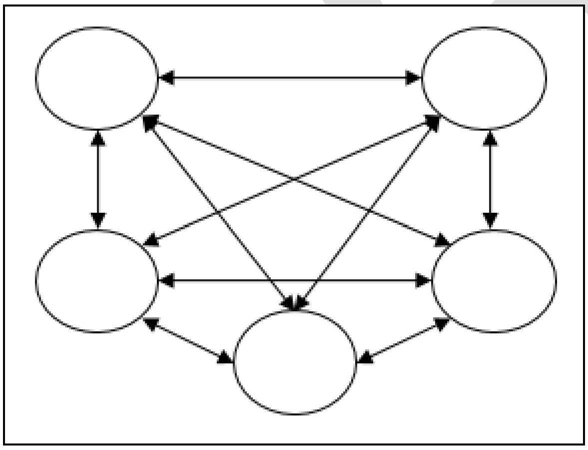
Advantages of ANN

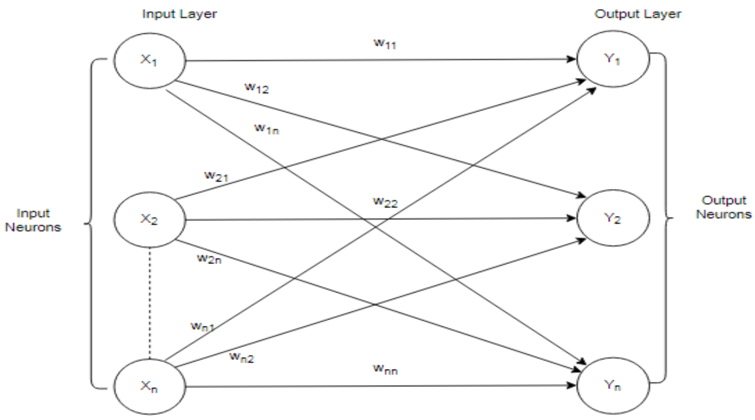
- ANNs exhibits mapping capabilities, that is, they can map input patterns to their associated output pattern.
- The ANNs learn by examples.
- The ANNs posses the capability to generalize
- The ANNs are robust system and fault tolerant. They can therefore, recall full patterns from incomplete, partial or noisy patterns.
- The ANNS can process information in parallel, at high speed and in a distributed manner

---

Neural Network Architecture

NN structure can be represented using Directed Graph. It is ordered tuple (V,E) vertices & Edges(v1,v2..Vn), (e1,e2,….En) Vertices of Graph represent neurons(I/O). Edges represent Synaptic links. The edges are labeled by the weights and attached to synaptic Links



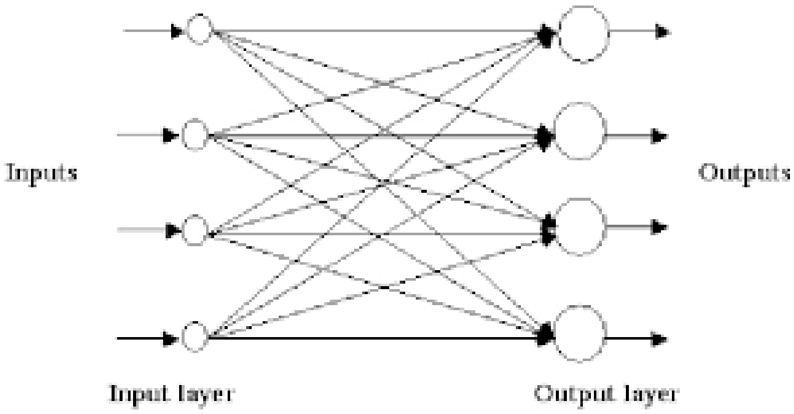


Types of NN

1. Single Layer Feed Forward Network

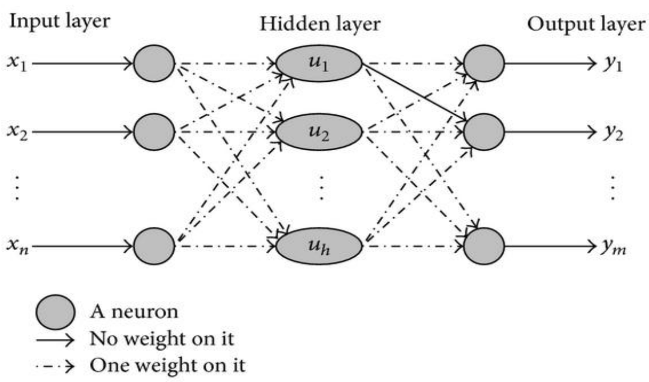
This network comprises 2 layers - 1.Input Layer 2.Out Put Layer

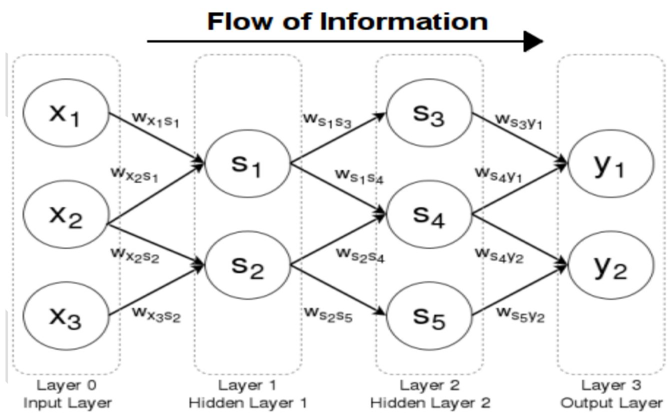
Input layer receives the Input signals and Out put layer receives the Out put Signals. The synaptic link connect every i/p neuron to o/p neuron. Hence it is Feed Forward. The i/p layer transmits the signals to o/p layer. The O/P layer performs computations , hence it is SLFFNN.



2. Multi layer Feed Forward Network

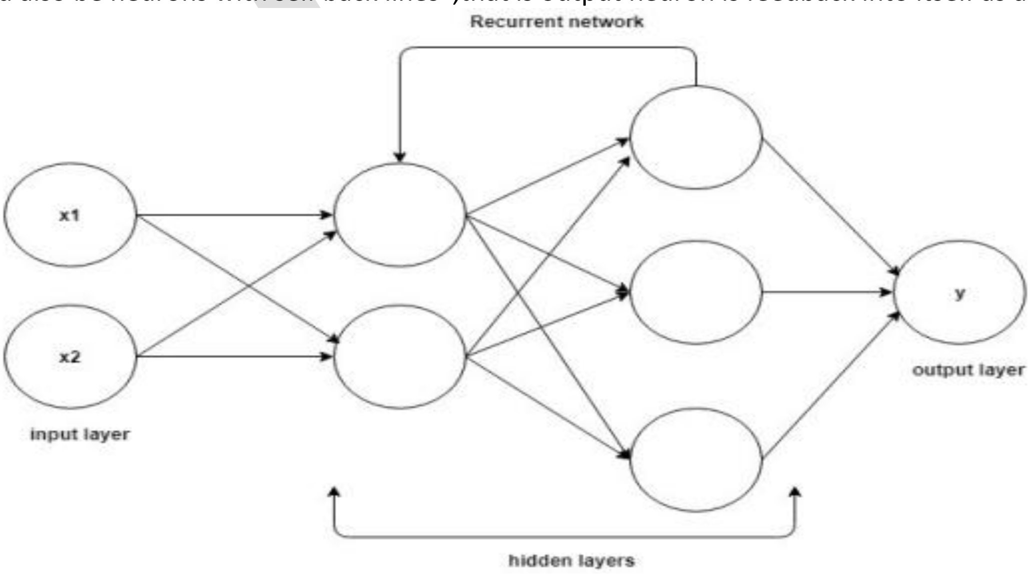
- MLFFNN (L-m1-m2-n) configuration.
- As name indicates N/W is made up of multi layers.
- I/P,O/P, Hidden layers(M1,M2)
- Computational units of hidden layer are called hidden neurons and hidden weights.
- I/P- hidden layer weights
- Hidden-O/P layer weights





3. Recurrent neural Network

- Differs from Single and Multi layer Architecture in the sence that there is a at least **one feed back loop** is there.
- There exist one layer with feedback connections
- Self feed back loops will be there
- O/P of neuron is I/P to the other in the form of feedback.
- There could also be neurons with self back lines ,that is output neuron is feedback into itself as a output.



Characteristics of NN

- NN exhibits mapping capabilities { mapping of I/P and O/P patterns}
- NN learns by example
- NN posses the capability to generalize, predict the new out comes from past trends
- NN are Robust and Fault Tolerant systems
- NN can process information in parallel at higher speed and distributed manner.

~~~~~~

LEARNING - Learning is a process of which free parameter of the neural networks are adopted through a process of change in environment in which neural network is present.

NN LEARNING METHODS -

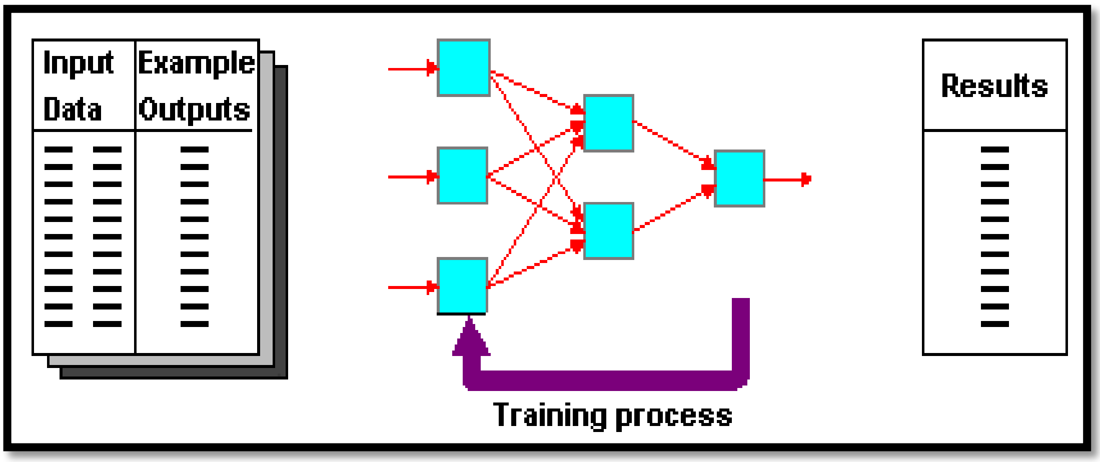
1. Supervised Learning
    a. Error Correction
        a1.Least mean Square
        a2.Back Propagation
    b. Stochastic
2. Un Supervised Learning
    a. Hebbian
    b. Competitive
3. Reinforced learning.

Supervised Learning –

- Every I/P pattern is used to train the N/W associated with O/P pattern.
- A teacher is assumed to be present during the learning process
- Comparison is made b/n N/W computed O/P and the expected O/P to determine the error.
- Error can be used to change N/W parameters which results in improvement of performance.

Example of Supervised Learning –

- **Supervised learning:**
  suppose you had a basket and it is fulled with some fresh fruits your task is to arrange the same type fruits at one place.
- suppose the fruits are **apple,banana,cherry,grape.**
- so you already know from your previous work that, the shape of each and every fruit so it is easy to arrange the same type of fruits at one place.
- here your previous work is called as **train data** in data mining.
- so you already learn the things from your train data, This is because of you have a **response variable** which says you that if some fruit have so and so features it is grape, like that for each and every fruit.
- This type of data you will get from the **train data.**
- This type of learning is called as **supervised learning**.
- This type solving problem come under **Classification.**
- So you already learn the things so you can do your job confidently.


~~~~~~

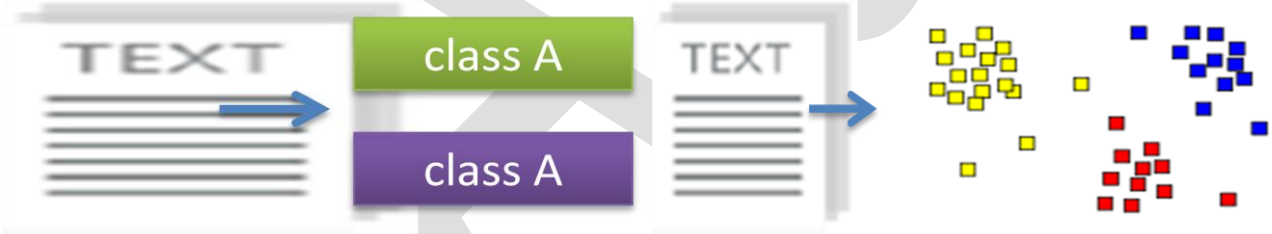
Unsupervised Learning –

Unsupervised learning

- suppose you had a basket and it is fulled with some fresh fruits your task is to arrange the same type fruits at one place.
- This time you don't know any thing about that fruits, you are first time seeing these fruits so how will you arrange the same type of fruits.
- What you will do first you take on fruit and you will select any physical character of that particular fruit. suppose you taken colours.
- Then the groups will be some thing like this.
- **RED COLOR GROUP: apples & cherry fruits.**
  GREEN COLOR AND SMALL SIZE: grapes.
- This type of learning is know **unsupervised learning.**

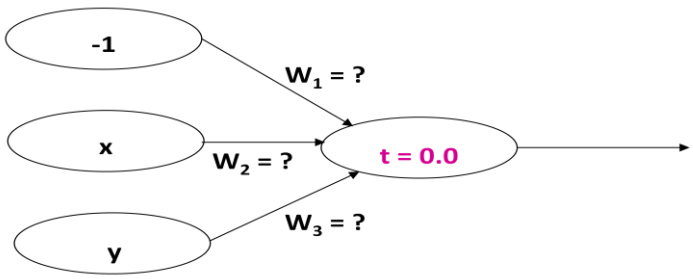
Types of Learning:

- Supervised: Learning with a labeled training set. Example: email classification with already labeled emails
- Unsupervised: Discover patterns in unlabeled data. Example: cluster similar documents based on text
- Reinforcement learning: learn to act based on feedback/reward. Example: learn to play Go, reward: win or lose



Classification Clustering

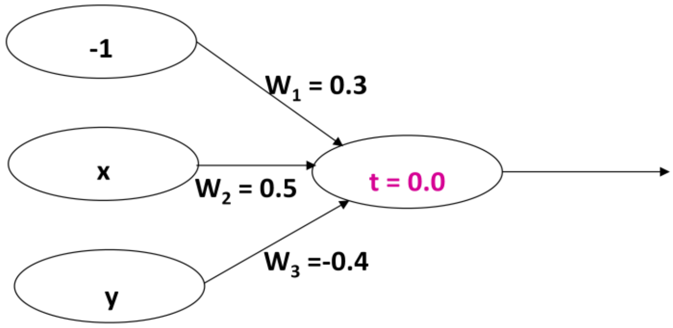
Training Perceptrons



For AND

A	B	Output
0	0	0
0	1	0
1	0	0
1	1	1

- What are the weight values?
- Initialize with random weight values



For AND

A	B	Output
0	0	0
0	1	0
1	0	0
1	1	1

I_1	I_2	I_3	Summation	Output
-1	0	0	(-1*0.3) + (0*0.5) + (0*-0.4) = -0.3	0
-1	0	1	(-1*0.3) + (0*0.5) + (1*-0.4) = -0.7	0
-1	1	0	(-1*0.3) + (1*0.5) + (0*-0.4) = 0.2	1
-1	1	1	(-1*0.3) + (1*0.5) + (1*-0.4) = -0.2	0

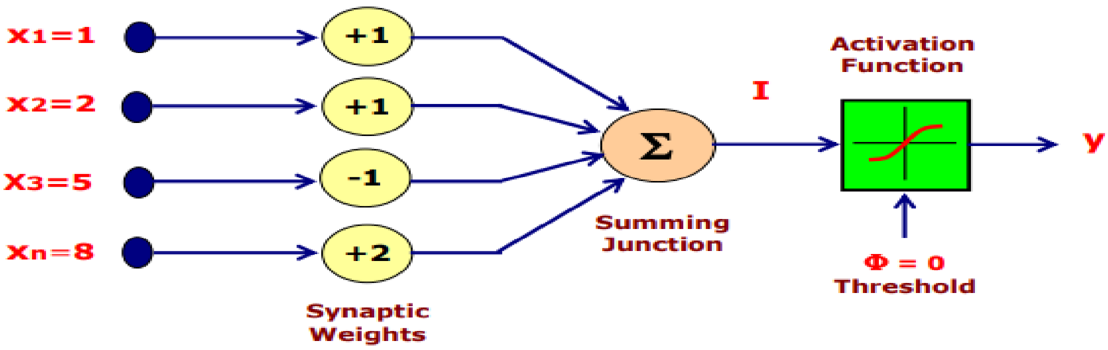


Fig Neuron Structure of Example

The output **I** of the network, prior to the activation function stage, is

$$I = X^T . W = \begin{bmatrix} 1 & 2 & 5 & 8 \end{bmatrix} \bullet \begin{pmatrix} +1 \\ +1 \\ -1 \\ +2 \end{pmatrix} = 14$$

$$= (1 \times 1) + (2 \times 1) + (5 \times -1) + (8 \times 2) = 14$$

With a binary activation function the outputs of the neuron is:

y (threshold) = 1;

~~~~~~

Stochastic Learning

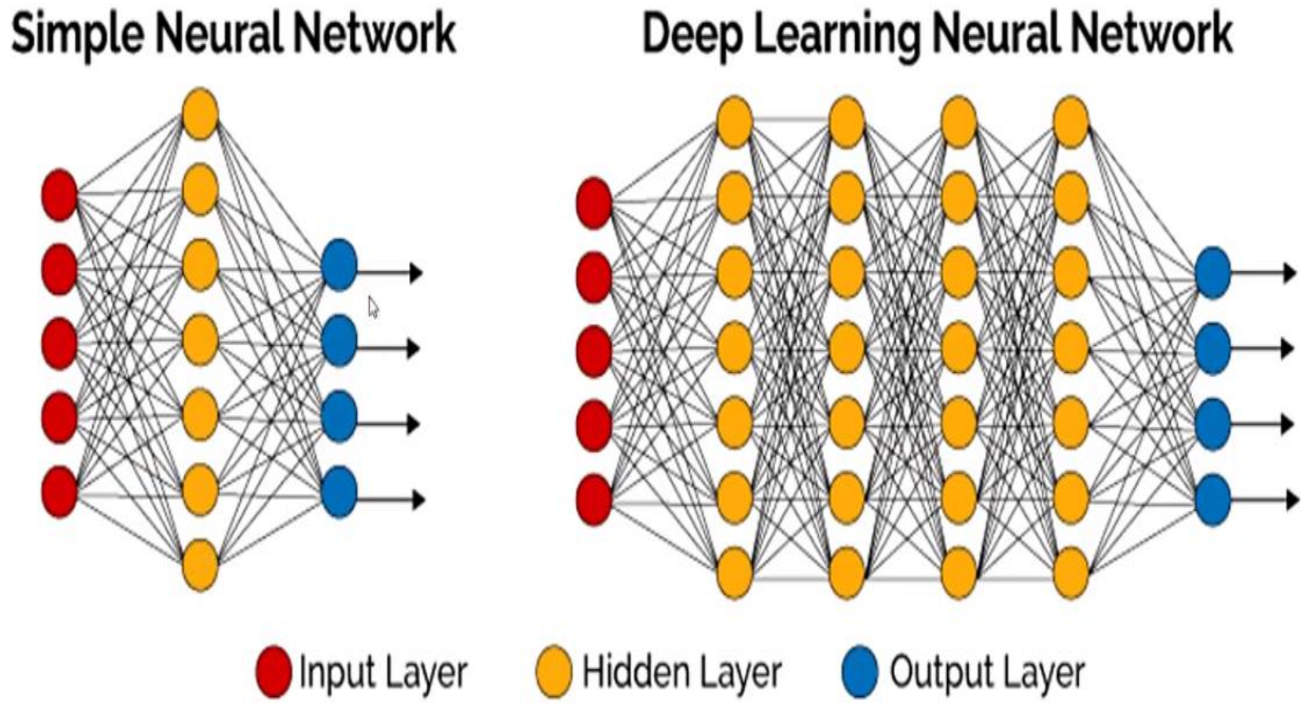
Weights are adjusted in probabilistic fashion. Learning mechanism is employed by Boltzmann and Cauchy machines.

~~~~~~

Hebbian Learning

- Proposed by Hebb and is based on correlative weight adjustment
- Oldest and inspired by biology
- In this I/P-O/P patterns pairs (Xi,Yi) are associated by weight matrix W, Known as Correlation Matrix
- It is compute as W= ∑Xi.Yi T.
- Yi T is the Transpose of the Yi

~~~~~~
~~~~~~


Simple Neural Network
Deep Learning Neural Network
Input Layer
Hidden Layer
Output Layer

~~~~~~
~~~~~~