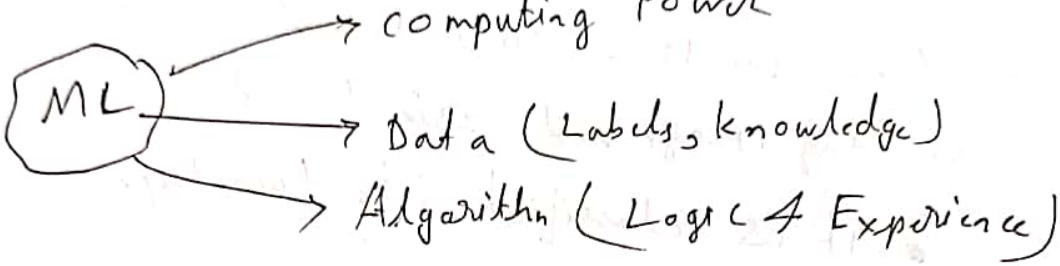
Machine Learning

Father Geoffrey Hinton.

- The art & science of giving computers the ability to learn & to make decisions from data without being explicitly programmed.

- A computer Program is said to learn from experience E with respect to some class of task T & performance measure P.



Goal: To build an automated data model for analytical reason

Objective: To build a system that learns from the data based on the applied algorithm

Reasons -

- Big data Explosion
- Hunger for new business & Revenue

Todays ML are mostly based on Supervised learning.

ML = mathematics + Probablity + Linear algebra + Statistic + decision theory + Algo + Neuro scien

Types of Learning :=>

(i) Supervised Learning : We trained machine using the labelled dataset. based on the training the machine predicts the output.

- Both input + output | Labelled data

Main goal is to map the input variable with the o/p variable.

(ex- we train with dog & cat images then give dog image and ask to identify it)

- it is classified into two type

(a) classification : classification algo are used to solve the classification problem in which o/p variable is categorized such as yes or no, Male or Female, Red or blue.

ex=> Spam detection, email filtering

popular algo are

- Random forest algorithm
- Decision tree
- Naive bayse
- Logistic Regression
- Suppart vector Machine

KNN number of neighbour

(b) Regression : Used to solve Regression problem in which there is linear

relationship between input and output variable.

ex → market trend, weather prediction etc

• has real value dollars / weights

Some popular algo —

- Simple linear Regressi.
- Multivariate "
- Decision Tree
- Lasso Regression

Adv • as it works on labelled data set

So we have exact idea about class of objects

• helpful to predict o/p on the basis of prior experience.

Dis • not able to solve complex problem

• requires lots of computational time.

• may predict wrong o/p

Applications

- image Segmentation
- Medical diagnosis
- Speech Recognition
- Fraud detection
- Spam detection

(ii) Unsupervised Learning :- (unlabelled data)

Main aim is to group or categaries the unsorted dataset according to the similarities, patterns and differences.

Machines are instructed to find the hidden pattern.

• only inputs

two types

(a) clustering - Grouping in the data. Such as - grouping customers by purchasing behaviour

(b) Association - discover rule that describe large portion of data

ex- People that buy X also tend to buy Y.

• Algo of clusterings are

- K means - k means number of iteration
- DBScan
- Mean shift
- PCA Principal component analysis (PCA)

Algo for associations are

- Apriori Algo
- FP growth
- Eclat

Adv :- • used for complex task

• useful for unlabel dataset

Dis :- • less accurate o/p

• working difficult

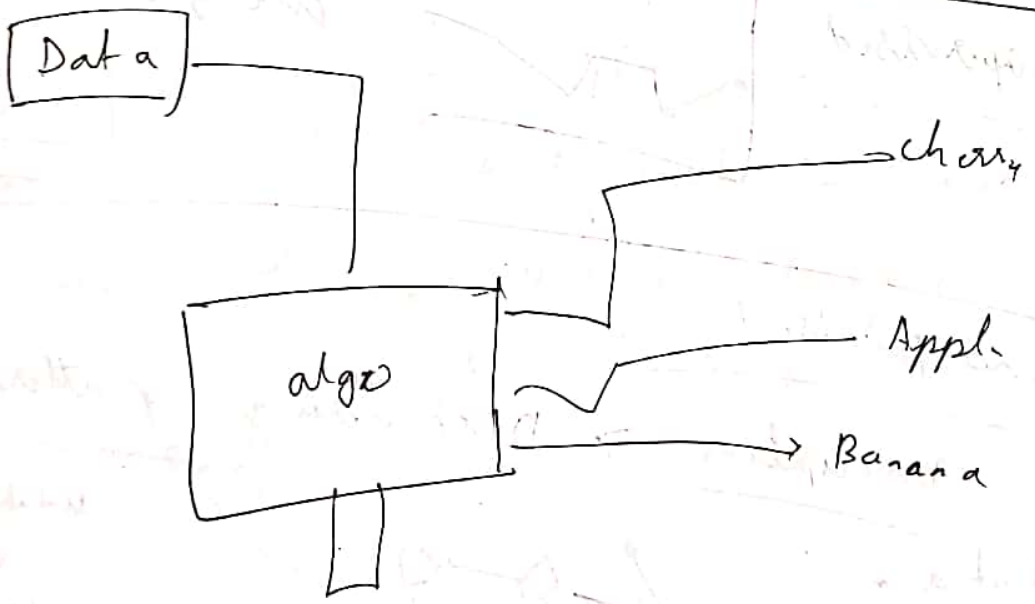
Applications are

- Recommendation System
- Network analysis
- Anomaly detection (fraud credit card)

(iii) Semi Supervised: lies between Supervised and Unsupervised

(iv) Reinforcement Learning: A class of problem where an agent operates in an environment and must learn to operate using feedback.

- The use of an environment means that there is no fixed training dataset rather a goal or set of goal that an agent is required to achieve through trial & error method.



Machine learning

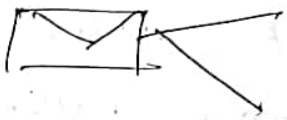
Supervised learning → Learning in our childhood

Unsupervised learning → Learning in our adulthood we learn on our own, similar patterns

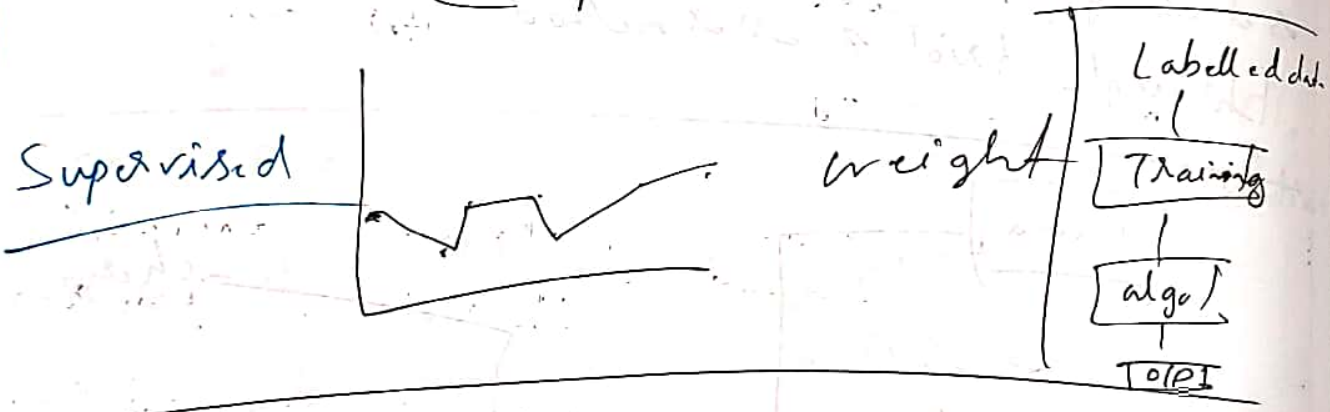
Reinforcement learning → if we are on an unknown island then we will learn where to eat etc by trial & error method

---

• Classification → classify the data in two or more type

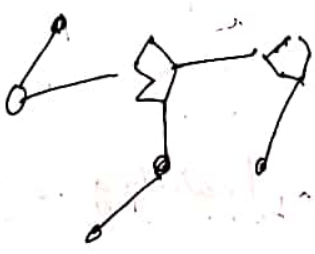


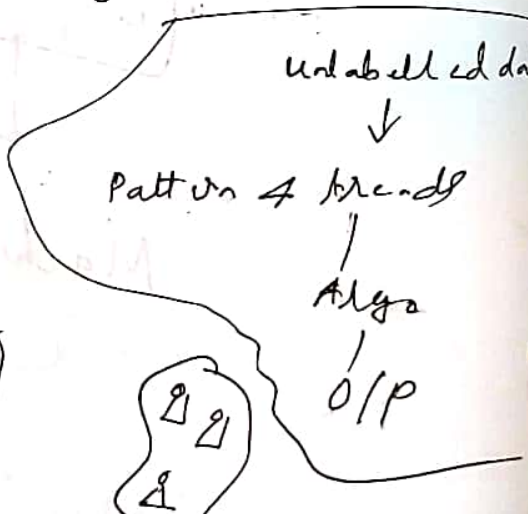
• Regression → used to predict continuous quantity. (continuous variable has infinite possibilities.



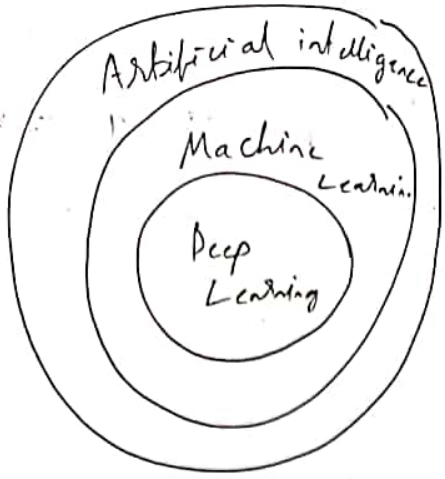
Unsupervised

Association → Discovering pattern and data.





clustering :=



(*) Deep Learning is a particular kind of ML that is inspired by the functionality of our brain cells called neurons which led to the concept of artificial neural network.

(*) ML algo easily works with smaller dataset but DL needs a large amount of data to achieve good performance

- DL need high level of Hardware.

Classification algo : is the person male or female?
is the mail is spam or not spam

Anomly detection : is there any fraud in CC?
is someone trying to hack

Clustering algo : What type of customer buy this product

Regression algo - market value of the house
Stock price prediction

Classification :⇒

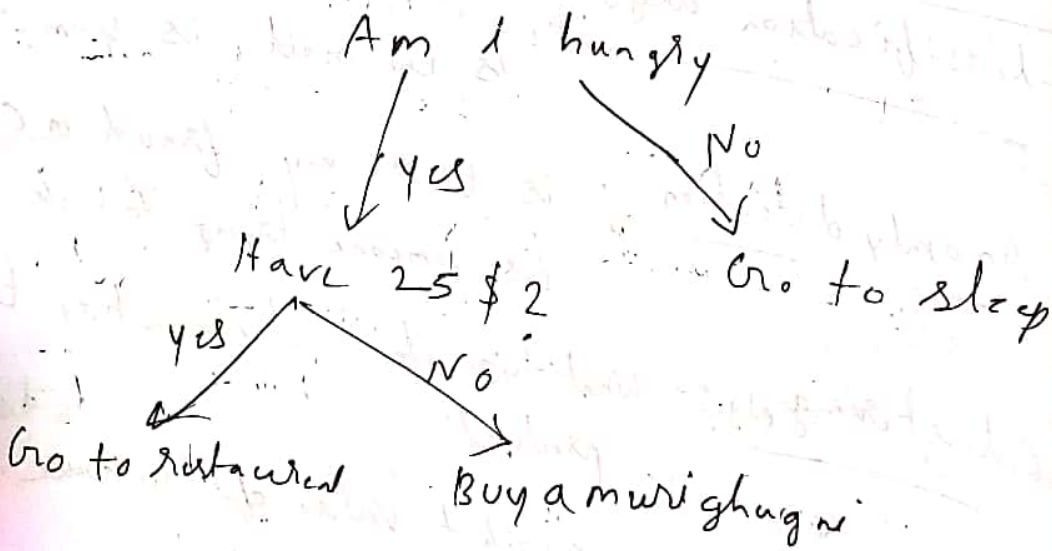
Classification is the process of dividing the datasets into different categories or groups by adding labels.

Types :-

(i) Decision tree (ii) Random forest
(iii) Naive Bayes (iv) KNN

Decision Tree :

- Graphical representation of all possible solution to a decision
- Decisions are based on some condition
- Decision made can be easily explained.



Random Forest:

- Build multiple decision trees and merge them together
- More accurate and stable prediction.
- Trained with bagging method (Combination of learning model increase overall result)
if we use different learning model & merge them together

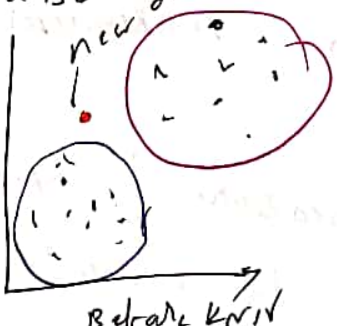
Naive Bayes:

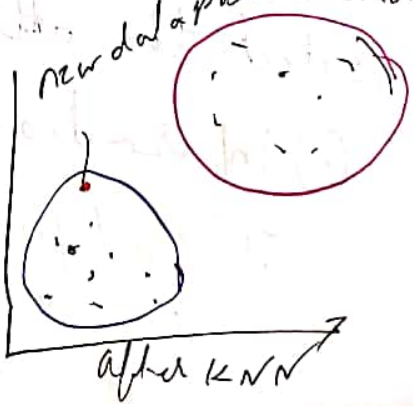
- classification technique based on Bayes theorem

Email spam filtering use it

K-nearest Neighbour:

- stores all the available data & classify a new data point based on the similarity
- it is called lazy learner algo as it does not learn from training data set instead when it gets new data, it performs action on the data set





Decision Tree terminologies

Root node : From where decision tree start

Leaf node : Final output node, tree can not be divided further

Splitting :⇒ process of dividing decision node into subnode

Branch / sub tree :⇒ A tree formed by splitting tree

pruning :⇒ process of removing the unwanted branches from the tree.

(X) In order to build a tree we use the CART (classification and Regression tree algorithm) algo

Attribute Selection Measure :⇒

Two popular technique are used to select the attribute (question asked for splitting)

(i) information gain : maximum information gain split first

(ii) Gini index : measure of impurity. attribute with low gini index is prefered.

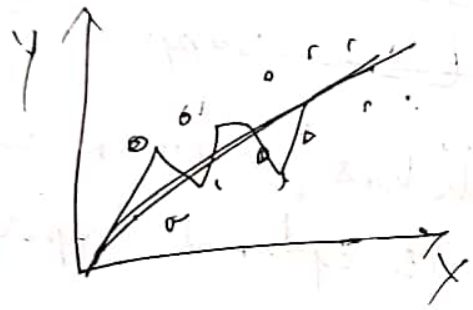
Regression :⇒ modelling technique which investigate the relationship between a dependent and independent variable

Uses → • Trends forcasting (trends in market)

• Forecasting an effect (how much sale in 1000 $ marketing)

Type ⇒ - Linear Regression ; Logistic Regression

Linear



$y = mx + c$

for every value of x there is y.

• Continuous variable

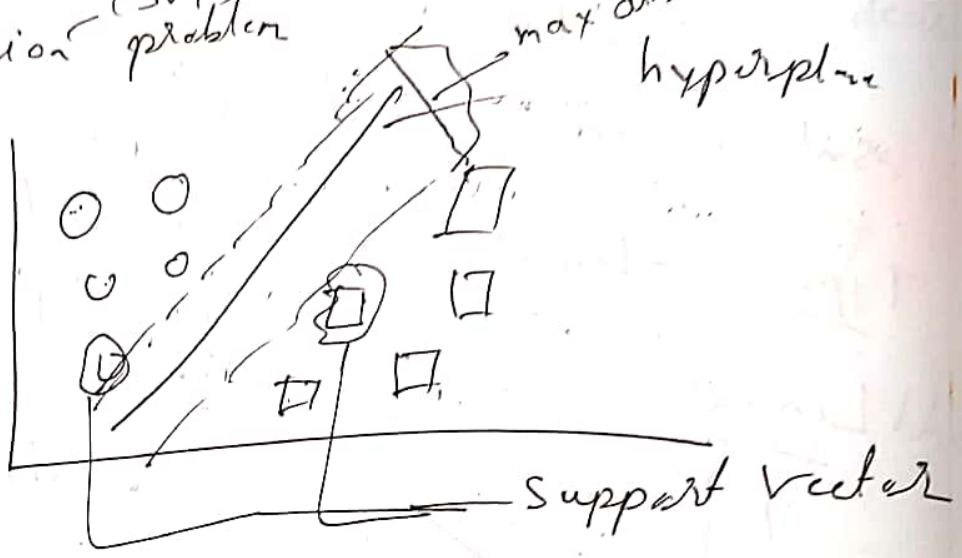
Logistic

• categorical variable

• probability of occurence of event

Support vector Machine (SVM) :

- SVM is a supervised classification method that separates data using hyperplanes
- It can be used for both classification & regression problem (SVR)



The hyperplane which has maximum distance from support vector is optimal hyperplane

(*) Support vector : The data points or vectors that are closest to the hyperplane and which effect the position of hyperplane. This vector support the hyperplane

SVM has two type.

(i) Linear SVM

(ii) Non Linear SVM

Clustering

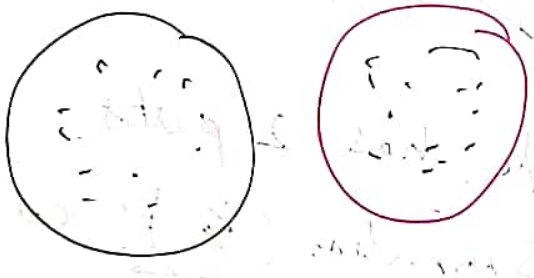
- It is the process of dividing the datasets into groups consisting of similar data points.
- It means grouping of objects based on the information found in the data

(*) The goal of clustering is to determine the intrinsic group in unlabelled data.

Uses :- in marketing, insurance companies Recommendation System.

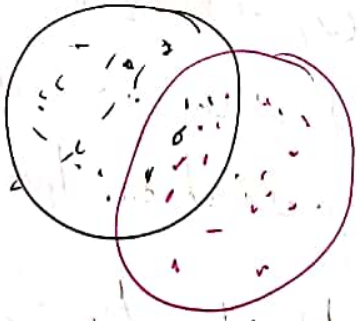
Types :⇒

(i) Exclusive clustering



K means clustering

(ii) overlapping clustering



(iii) Hierarchical clustering → parent child relationship

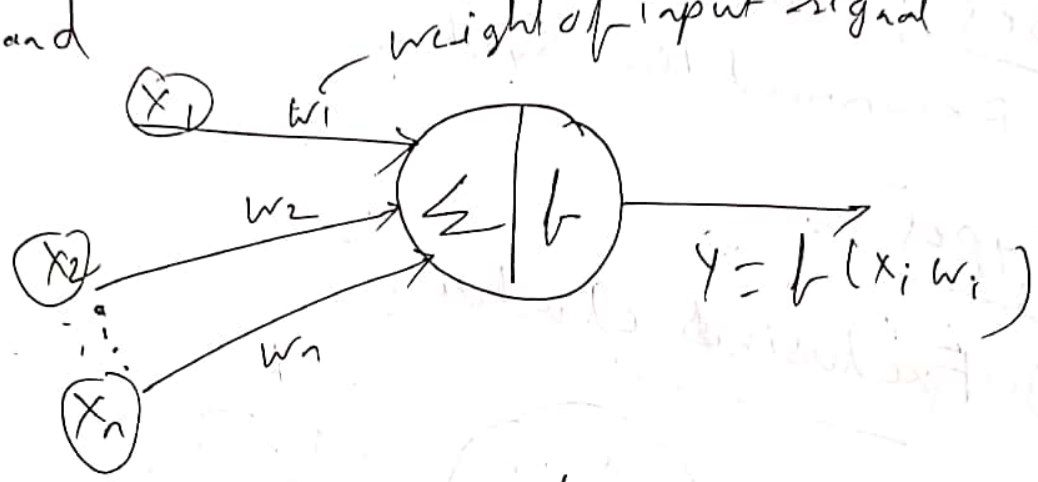
Artificial Neural Network :⇒

Our human brain has million of neurons. Neurons helps us to stimulate our action

(*) When natural functions are done artificially then it is called artificial neural network

Biological – Neurons
Formal – Nodes

ex– When we touch a hot vessel all neuron give signal to brain, we automatically take our hand



A node has 2 parts

(i) Summation (Σ): calculate the weighted sum of all the i/p

$= x_1 w_1 + x_2 w_2 + \cdots + x_n w_n$

Once weighted sum is calculated it is sent to activation function

(ii) Activation function :⇒ Generate the o/p on i/p given

⊛ Every node is connected with the help of connection link.

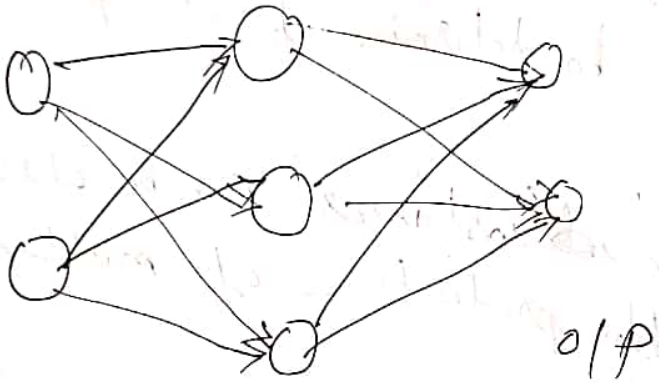
Representation of Artificial Neural network ⇒

ANN are divided into 3 parts

(i) input layer

(ii) Hidden layer

(iii) O/P layer



all the nodes should be connected with next layer

- I/P layer will receive input signal & then send to hidden layer.
- Hidden layer extract information of data of input layer & process it. Then send to O/P layer

Bayesian Learning

- most practical approach
- provide useful perspective to understand other learning algo.

Features :=

- instead of eliminating hypothesis, it will increase / decrease the estimated probability.
- Prior knowledge can be combined with observed data to determine final prob of a hypothesis
- New instances can be classified by combining the prediction of multiple hypothesis.

Issues :-

- prior knowledge needed
- computational cost high

Bayes. Theorem

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

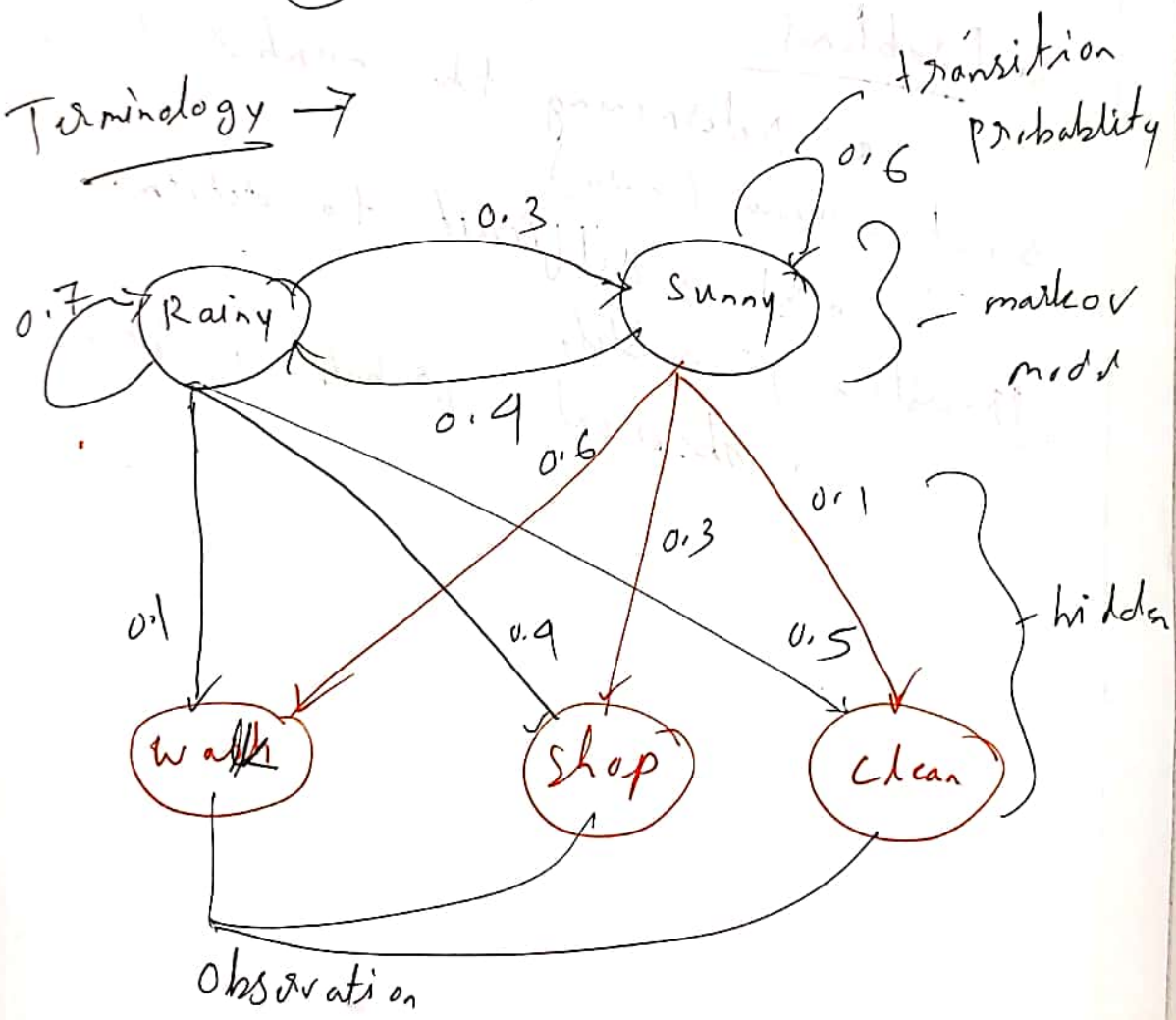
- $P(A|B)$: posterior prob
- $P(B|A)$: likelihood
- $P(B)$: marginality

Hidden Markov Model ⇒

HMM is a statistical Markov Model in which the system being modelled is similar to be a markov process with unobserve (hidden) states.

application —
(i) Reinforcement learning
(ii) Speech, gesture Recognization.
(iii) Bioinformatics

Terminology →



Property — • memory less
• Future & past are independent
prediction (depends on current state)

Goal is to make sequence of decision where a particular is influenced by earlier decision.

- HMM is set of states with transition probability, observation probability distribution
- HMM is an extension of Markov model
- HMM produce a sequence of observable symbol as output.

- Problems
  - Determining the number of states is not obvious (easy)
  - It is difficult to determine transition probability
  - observation probability is hidden.

(x) Descriptive model :⇒

describe real world event and the relationship between factors.

does not involve any target variable

ex= clustering, association

predictive model

predict the output.