PAGE :
DATE: / /

ML Final Prep:

Q-1 Diff. b/w ML & traditional programming

=>

Machine Learning →	Traditional Programming →
(1) It is a subset of artificial intelligence in which we find suitable models for our problems	(1) It is the normal coding we do to make a program which gives us output.
(2) In ML, we use various algorithms like classification, regression etc.	(2) In this we just code using language like Java, HTML, CSS, C++, C etc.
(3) It is a probabilistic approach	(3) It is a deterministic approach
(4) In this, the algorithm is learning itself continuously with new data.	(4) In this, there is no learning of the algorithm.
(5) Ex, ML is used for making predictions, assumptions & analysis of problem like Fraud Detection, etc	(5) Ex, It is used for making websites & software working on both front end & backend. etc
(6) Input + Output → Program	(6) Input + Program → Output

PAGE :
DATE: / /

Q2 Different Paradigms / type of ML ?

→ (1) Supervised Learning →

- Supervised Learning is a type of machine learning that uses labelled data to train machine learning models.

- In labelled data, the output is already known. The model just needs to map the inputs to the respective outputs.

- As an example of supervised learning is to train a system that identifies the image of an animal.

- Supervised learning algorithms take labelled inputs & map them to the known outputs, which means you already know the target variable.

- Supervised learning algorithms are generally used for solving classification & regression problems.
- Eg: Linear regression, Logistic regression, SVM, Nearest neighbors

(2) Unsupervised Learning →

- Unsupervised learning is a type of machine learning that uses unlabeled data to train machines.

- Unlabeled data doesn't have a fixed output variable. The model learns from the data, discovers the pattern & features in the data, & returns the output.

PAGE :
DATE: / /

- Example we put image & the model classifies the image as bee image or tech image.

- Unsupervised learning finds patterns & understands the trends in the data to discover the output. So the model tries to label the data based on the features of the input data.

- Unsupervised learning is used for solving clustering & association problems.

- Algorithms used in this are K means clustering, Hierarchical clustering etc.

(3) Reinforcement learning →

- Reinforcement learning trains a machine to take suitable actions & maximize its rewards in a particular situation.

- It uses an agent & an environment to produce actions & rewards.

- As an example could be to train a dog to catch the ball, & give it a reward.

- Reinforcement learning follows trial & error methods to get the desired result. After accomplishing a task, the agent receives an award.

PAGE :
DATE: / /

- Reinforcement learning algorithms are widely used in the gaming industry to build games. It is also used to train robots to do human tasks.

- Algorithms used in this are Q-learning, Deep Q learning etc.

Q3 Why machine learning is the future?

=> ① Easily identifies trends & patterns ->

machine learning can review large volume of data & discover specific trends & patterns that would not be apparent to humans. For instance, for an e-commerce website, it serves to understand the browsing behaviour & purchase history.

② No human intervention needed (automation) ->

With ML, you don't need to babysit your project every step of the way, since it means giving machines the ability to learn, it lets them make predictions & also improve the algorithms on their own.

③ Continuous Improvement ->

As ML algorithms gain experience, they keep improving in accuracy & efficiency. This lets them make better decisions. For instance while making a weather forecast model, as the amount of data

PAGE :
DATE: / /

Deep growing, the algorithm learns to make more accurate predictions faster.

(4) Handling multi-dimensional & multi-variety data:-

ML algorithms are good at handling data that are multidimensional & multi-variety, & they can do this in dynamic or uncertain environments

(5) Wide applications:-

You could be an e-tailor or a healthcare provider & make machine learning work for you. Where it does apply, it holds the capability to help deliver much more personal experience to customers while also targeting the right customers.

Q-4 Applications of ML?

⇒ Machine learning is the subset of AI which uses algorithms to find results like predictions. Machine learning is very useful in every aspect of life. It gives accurate results & helps saving time of human beings. It have various applications, some of which are →

(1) Social media → ML is very much trend in every field & social media is one of them. Posting pictures, making comments, adding friends, all are a part of social media. ML gives us recommendations of like mutual friends & of posts which we may like based on our activities on such platform. The biggest examples are facebook & insta

PAGE :
DATE: / /

(2) Speech Recognition → ML helps in recognizing voices & gives suitable results. Google & Alexa are using such features for searching or even for calling someone. For example, if you want to search for latest news we just have to say on the google & we will get the results.

(3) Spam filtering → This is done using ML. We get alerts from google when we get a spam mail & we also get a true caller notification with warning of spam.

(4) Online recommendations → ML gives us recommendations based on our search history, like if we have searched about stationary items, it will start giving us ads of that item. There is another example of recommendation i.e. movie recommendation by netflix.

(5) Medical diagnosis → ML helps in predicting diseases. The various parameters are given & on the basis of these parameters, we get our prediction of whether this person is suffering from that disease or not.

(6) Inventory management → ML helps in predicting sale of products. Like we give all our previous ~~data~~ data & ML helps in knowing which product is best & most purchased. Based on ~~their~~ this the owner can know which product he should stock & in what quantity.

PAGE :
DATE: / /

(7) Self driving cars → Companies are introducing self driving cars which are based on ML to imitate ML cars have sensors in them to sense all the requirements.

(8) Stock market prediction → It helps in the prediction of stocks like how much will the rise & fall of stocks based on the previous data.

Q What is regression?

⇒ Regression is a process of finding the correlations b/w dependent & independent variables. It helps in predicting the continuous variables such as prediction of market trends, prediction of house prices etc.

Ex simple linear regression, multiple linear regression, polynomial regression, support vector regression, decision tree regression, random forest regression

~~Q Different regression models~~

Q Different regression models?

⇒ Simple linear regression, multiple linear regression, polynomial ~~linear~~ regression, Support vector machine (SVM), ~~naive bayes~~, Decision Tree ~~classification~~, Random Forest regression.

PAGE :
DATE: / /

Q- What is simple linear regression?

=> It is a type of regression algorithms that models the relationship b/w a dependent variable & a single independent variable.

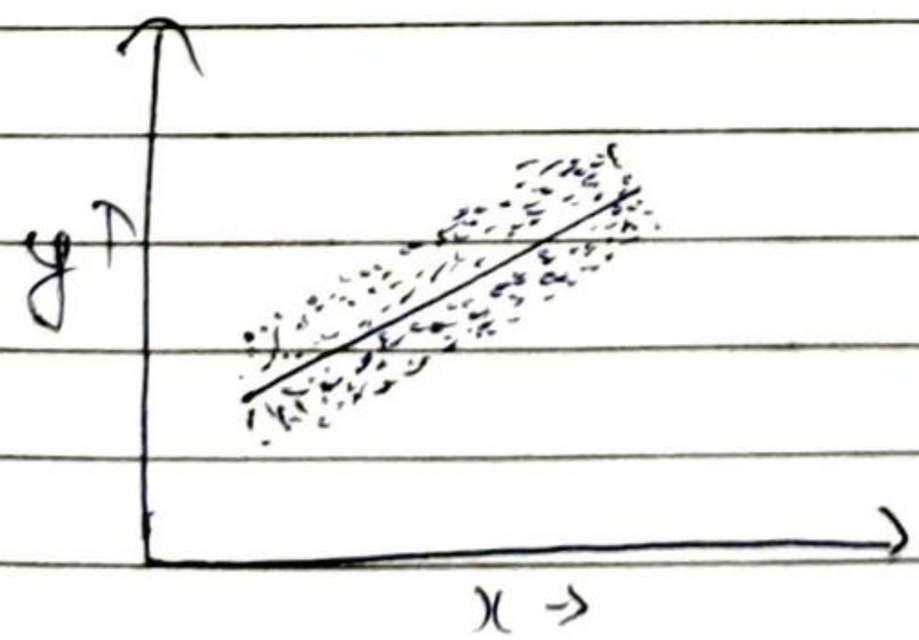
- The relationship shown by a simple Linear Regression model is linear or a sloped straight line.

- Simple linear regression algorithm has mainly two objectives:
(i) model the relationship between the two variables.
(ii) Forecasting new observations

- The simple linear equation is

$$\boxed{y = a_0 + a_1 x}$$

a_0 = intercept
a_1 = slope



PAGE :
DATE: / /

- Advantages →
(i) It performs exceptionally well for linearly separable data.
(ii) It is easier to implement & efficient to train

- Disadvantages →
(i) The assumption of ~~linearly~~ linearity b/w dependent & independent variables
(ii) It is often quite prone to noise & overfitting.

- Example of use case →
weather forecasting, relationship b/w income & expenditure.

Q numerical of simple linear regression

X	Y
1	3
2	4
3	5
4	7

⇒

	X	Y		XY	X^2
	1	3		3	1
	2	4		8	4
	3	5		15	9
	4	7		28	16
Σ	10	19		54	30

PAGE :
DATE: / /

We have to form equation

$$y = bx + a$$

$$\cancel{a = \frac{(\Sigma y)(\Sigma x^2) - (\Sigma x)(\Sigma xy)}{n(\Sigma x^2) - (\Sigma x)^2}}$$

(n = no. of terms)

$$b = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{n(\Sigma x^2) - (\Sigma x)^2}, \quad a = \frac{\Sigma y - b(\Sigma x)}{n}$$

$$a = \frac{(19)(30) - (10)(54)}{(4)(30) - (100)}$$

$$= \frac{570 - 540}{120 - 100}$$

$$= \frac{30}{20} = \frac{3}{2} = 1.5$$

$$b = \frac{(4)(54) - (10)(19)}{(4)(30) - (100)}$$

$$= \frac{216 - 190}{120 - 100}$$

$$= \frac{26}{20} = \frac{13}{10} = 1.3$$

$$\boxed{y = 1.3x + 1.5}$$

PAGE :
DATE: / /

now using the equation, calculate predicted value

p = predicted value

p	Error
2.8	0.2
4.1	0.1
5.4	0.4
6.7	0.3

Lower the error, better the model

Q What is multiple linear regression?

=> • It models the linear relationship b/w a single dependent continuous variable & more than one independent variable.

• It is an extension of simple linear regression as it takes more than one predictor variable to predict the response variable.

• The key point is each feature variable must model the linear relationship with the dependent variable.

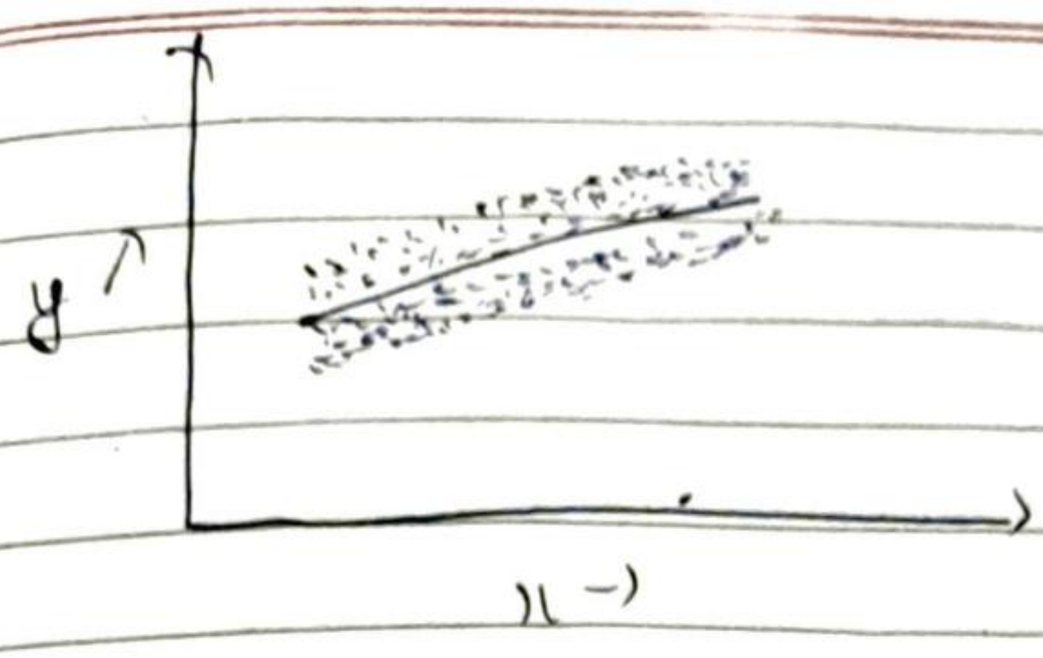
• The multiple linear regression equation →

$$y = a_0 + a_1 x_1 + a_2 x_2 + \dots + a_m x_m$$

$$\Rightarrow \boxed{y = a_0 + \sum_{i=1}^{m} a_i x_i}$$

a_i = regression coefficient
x_i = independent variable
y = dependent variable

PAGE :
DATE: / /



- Advantages →
  (i) Works well irrespective of the dataset size.
  (ii) Gives information about the relevance of feature.

- Disadvantages →
  (i) The assum in multiple linear regression.

- Example → Prediction of CO_2 emission based on engine size & number of cylinder in a car.

Q numerical Predict value of ?)

Subject	Y	X_1	X_2
1	-3.7	3	8
2	3.5	4	5
3	2.5	5	7
4	11.5	6	3
5	5.7	2	1
6	?	3	2

PAGE :
DATE: / /

=>

Q What is Polynomial Regression?

=> • It is a regression algorithm that models the relationship b/w a dependent & independent variable as nth degree Polynomial.

• It is a special case of multiple linear regression as we add some polynomial terms to the multiple linear regression equation to convert it to polynomial regression.

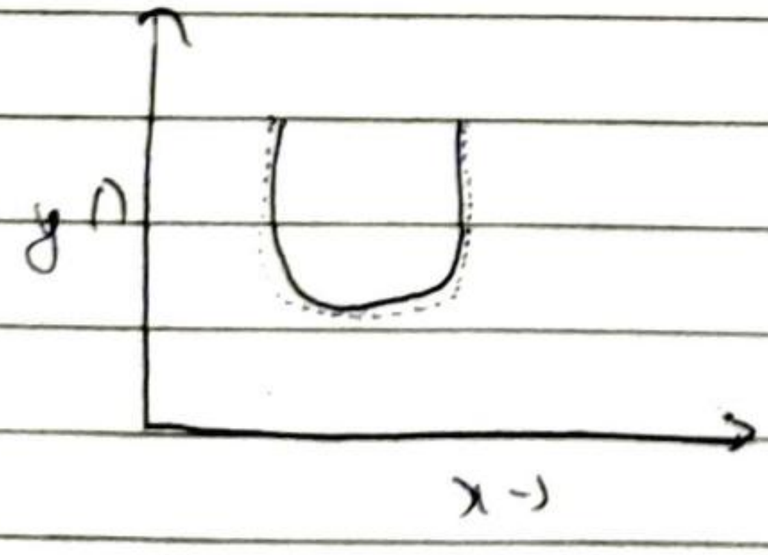
• The dataset used for training is of non-linear nature.

• The equation is →

$$y = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \dots + a_n x^n$$

=> $$\boxed{y = a_0 + \sum_{i=1}^{n} a_i x_i + F_p}$$

where F_p = polynomial function which holds degree of x

• 

PAGE :
DATE: / /

- Advantages →
(i) works very well on non linear problems
(ii) works on any size of dataset

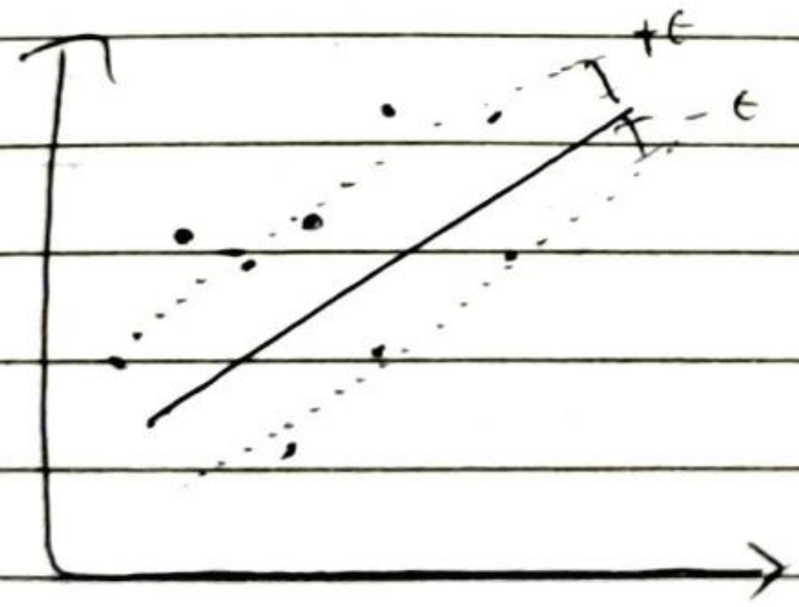
- Disadvantage →
(i) We need to choose the right polynomial degree for a good bias / variance trade off.

- Example → It is widely applied to predict the spread rate of COVID - 19 & other infectious diseases.

Q What is support vector regression?

=> • Support vector regression ~~it~~ ~~used~~ uses the same principle as support vector machine.

- SVR find the best line i.e. called the hyperplane that has the maximum number of points.

- 

- Example → Salary expected

- Advantages →
(i) It is easy to implement.

PAGE :
DATE: / /

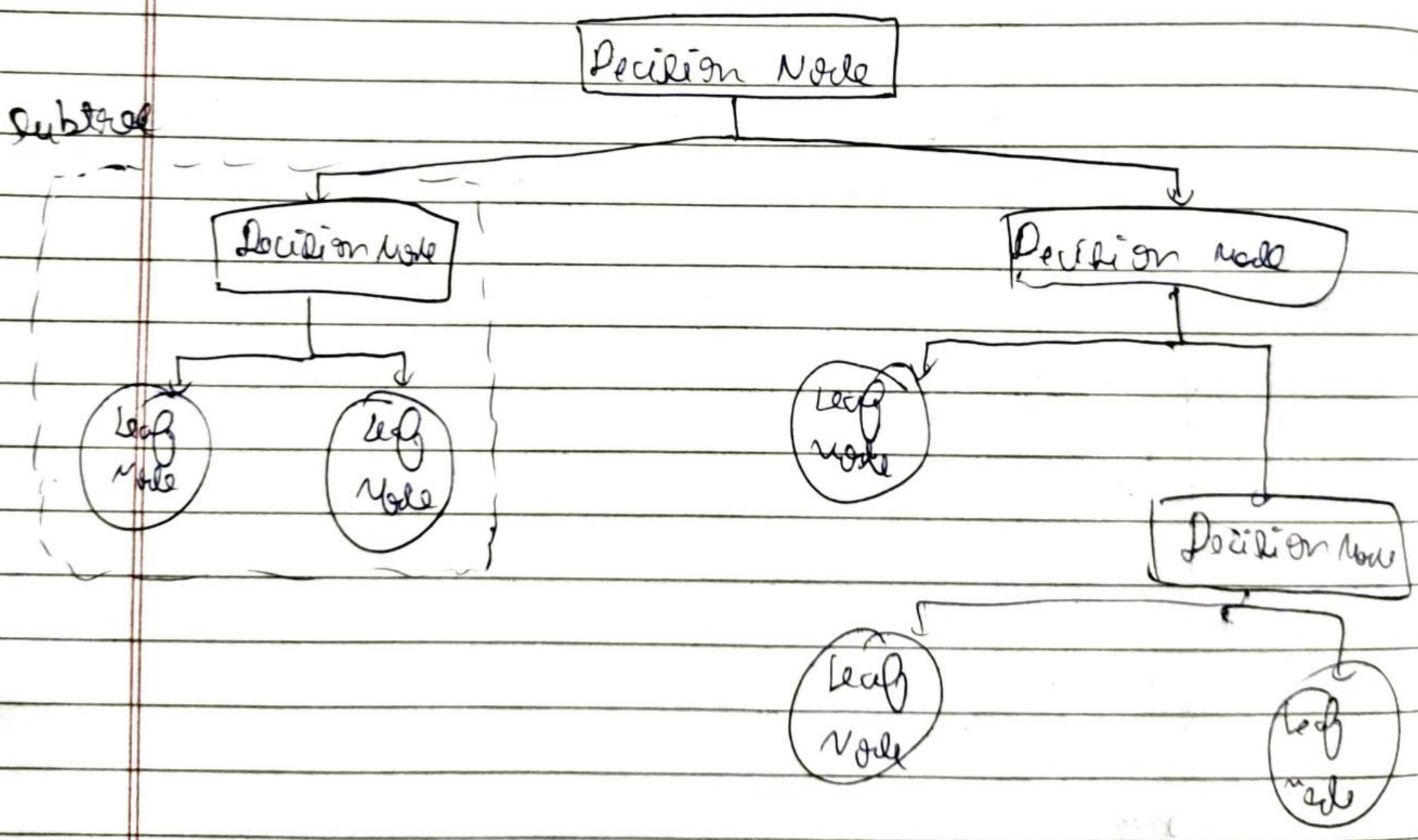
- Disadvantage →
  (i) They are not suitable for large datasets.

Q Decision Tree Regression?

⇒ - It is a tree structured regression model where the nodes represent the decision & leaf node represent the outcome.

- There are present two nodes that are decision node & leaf nodes.

- A decision tree simply asks a question, based on answer, it further splits the tree into subtree.



PAGE :
DATE: / /

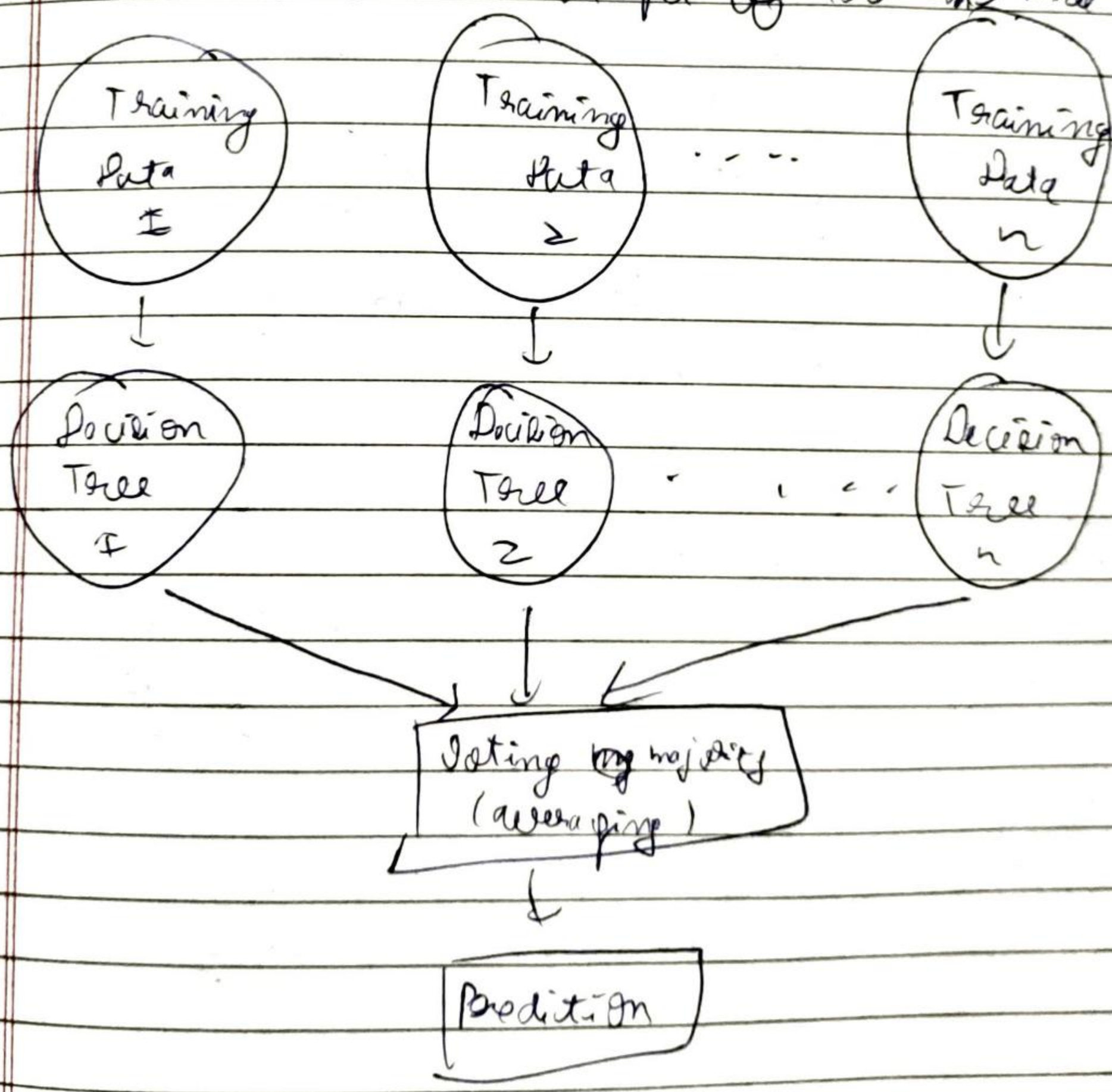
- Advantage →
(i) It works on both linear & non linear problems.

- Disadvantage →
i) Poor results on small datasets.

Q What is Random Forest Regression?

=) In Random Forest, it contains a number of decision trees on various subsets of the given dataset & takes the average voting majority to predict the final value.

- It is based on the concept of the ensemble learning



PAGE :
DATE: / /

- Advantages:
(i) It is powerful & accurate
(ii) It has good performance on many problems

- Disadvantages:
(i) We need to choose the number of decision trees.

Q What is Inductive Learning?

=> Inductive learning is where we are given examples of a function in the form of data (x) & the output of the function (f(x)). The goal of inductive learning is to learn the function for new data (x).
- Classification → when the function being learned is discrete
- Regression → when the function being learned is continuous
- Probability Estimation → when the output of the function is a probability.

- Example →
- Disease diagnosis →
  - The x are the properties of the patient
  - The f(x) is the disease they suffer from

- Face recognition →
  - The x are bitmaps of people faces
  - The f(x) is to assign a name to face.

PAGE :
DATE: / /

Q What is concept learning?

⇒. Concept learning can be formulated as problem of searching through a predefined space of potential hypotheses for that best fits the training examples (solution)

- It can be stated as acquiring the definition of a general category from given sample positive & negative training examples of the category

Q What is overfitting & underfitting?

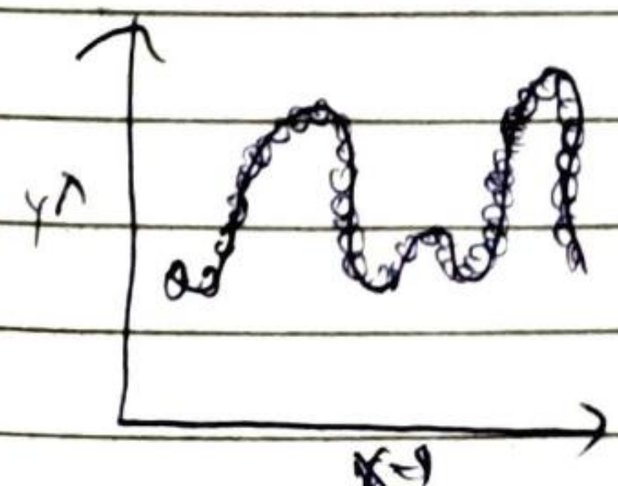
⇒ Overfitting →

- Overfitting occurs when our model tries to cover all the data points or more than required data points.

- Because of this the model starts catching noise & inaccurate values present in the dataset. All these factors reduce the efficiency & accuracy of the model.

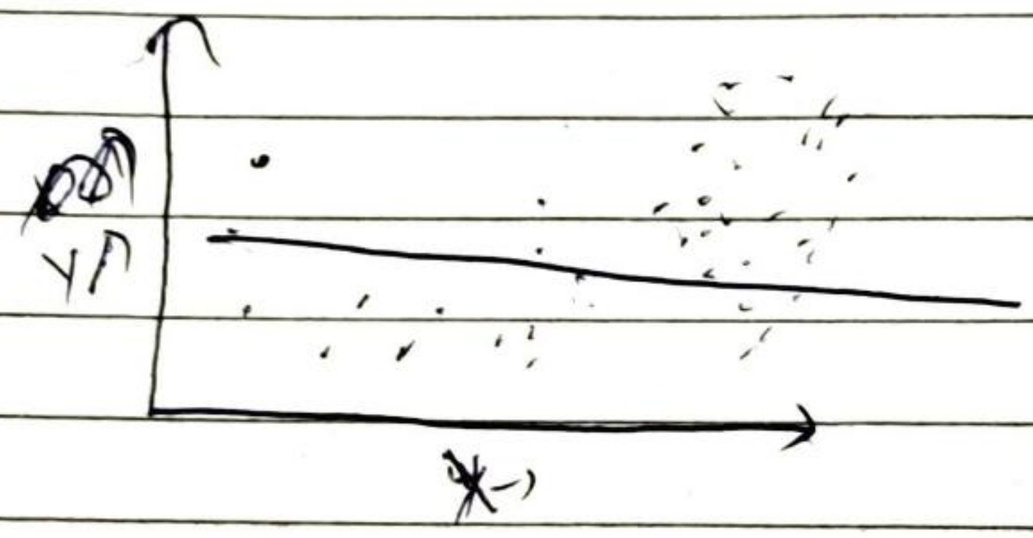
- The overfitted model has low bias & high variance.

- The more we train our model, the more chances of occurring the overfitted model.

- 

PAGE :
DATE: / /

Underfitting :-

- Underfitting occurs when our model is not able to capture the underlying trend of the data.
- To avoid overfitting, the feed of training data can be stopped at an early stage, due to which the model may not learn enough. As a result the model may not detect the underlying trend of the data.
- An underfitted model has high bias & low variance.
- Since the model is not able to learn enough, hence it reduces the accuracy & produces unreliable results.
- 

PAGE :
DATE: / /

Q Issues related to machine learning.

→ (1) Inadequate traing data →

The major issue that comes while using machine learning algorithms is the lack of quality & quantity of data. The data is inadequate, noisy & unclean.

(2) Poor quality of data →

Data quality can also be considered as a major common problem while processing machine learning algorithms.

(3) Overfitting & Underfitting →

Overfitting → Whenever a ML model is trained with a large amount of data, it starts capturing noise & inaccurate data into training dataset.

Underfitting → Whenever a ML model is trained with fewer amounts of data, it provides incomplete & inaccurate data & destroys the accuracy of model.

(4) Monitoring & maintenance →

Regular monitoring & maintaince is compulsory for generalized output.

PAGE :
DATE: / /

(5) Lack of skilled resources →

Although ML & AI are continuously growing in the market, still these industries are fresher in comparison to others.

(6) Slow implementation & results →

Although ML models are highly efficient in providing accurate results, but they are time consuming.

Q Various stages involved in designing a model.

(1) Collecting Data →

- It is of the upmost importance to collect reliable data so that your machine learning model can find the correct patterns.

- The quality of the data that has been fed to the machine will determine how accurate the model is.

- If the data is incorrect or outdated, there will be wrong predictions which are not relevant.

(2) Preparing the Data →

- Put together all the data & randomize it. This helps make sure that data is evenly distributed & ordering does not affect the model.

PAGE :
DATE: / /

- Clean the data to remove unwanted data, missing values, duplicate values etc.

- Visualize the data to understand the relationship b/w various variables.

- Split the data into training & testing set.

(3) Creating a model →

- It is important to create a model which is relevant to the task at hand.

- Find out to see if the model is suited for numerical or categorical data & choose accordingly.

(4) Training the model →

- In training the prepared data is passed to the machine learning model to find patterns & make predictions.

- It results in the model learning from the data so that it can accomplish the task.

- Over time with training the model gets better at predicting.

PAGE :
DATE: / /

(5) Evaluating the model →

- After training, model is checked for its performance.

- This is done by testing the performance of the model on previously unseen data.

(6) Parameter Tuning →

- Once the evaluation of the model is done, check if its accuracy can be improved in any way.

- This is done by tuning the parameters present in the model.

(7) making predictions →

- The model is now ready for interacting the new results.

Q What is validation set?

→ • The validation set is a set of data separate from the training set that is used to validate our model performance during training.

- This validation process gives information that helps tune the model's parameters.

PAGE :
DATE: / /

- It tells us whether the training is moving in the right direction or not.

~~What is classification?~~

Q What is the importance of Sikit-learn library in python ?

=> Sikit-learn (sklearn) is the most useful & robust library for machine learning in Python.

- It provides a selection of efficient tools for ML & statistical modeling including classification, regression, clustering & dimensionality reduction via a consistent interface in Python.

- This library is built upon NumPy, Scipy, & matplotlib

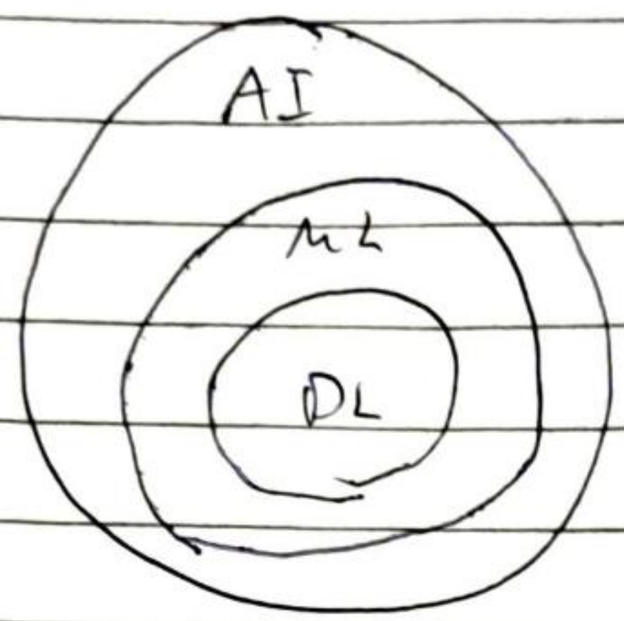
PAGE :
DATE: / /

Q How the performance of two ML models can be evaluated on the given data?

⇒

Q Diff. b/w AI, ML & DL

⇒



AI →	ML →	DL →
① It is study / process which makes machines to mimic human behaviour through of particular algorithm.	① It is the study that uses statistical methods enabling machines to improve with experience.	① It is the study that makes use of Neural Networks (similar to neurons present in human brain) to imitate functionality just like a human brain.
② AI is the broader family consisting of ML & DL as its components.	② ML is the subset of AI.	② DL is the subset of ML.
③ AI is a computer algorithm which	③ ML is an AI algo which allows system	③ DL is ML algo that uses deep neural network

PAGE :
DATE: / /

exhibits intelligence through decision making.	to learn from data.	to analyze data & provide output accordingly
(4) The aim is to basically increase chances of success not accuracy.	(4) The aim is to increase accuracy not caring much about the success ratio	(4) It attains higher rank in terms of accuracy when it is trained with large amount of data.
(5) The efficiency of AI is basically the efficiency provided by ML & DL	(5) Less efficient than DL as it can't work for longer dimensions or higher amount of data	(5) more powerful than ML as it can easily work for larger sets of data
(6) ML & DL are parts of AI or branches of AI	(6) ML consists of supervised, unsupervised & reinforcement learning etc	(6) DL consist of two> (i) CNN (ii) RNN
(7) Ex AI powered traditional, uber, autopilot in flights	(7) Ex virtual assistants like siri, alexa, Email spam filtering.	(7) Ex Image analysis, caption generation on live video.

Diff algo → Logistic regression, KNN, SVM, naive bayes, Decision Tree, Random Forest

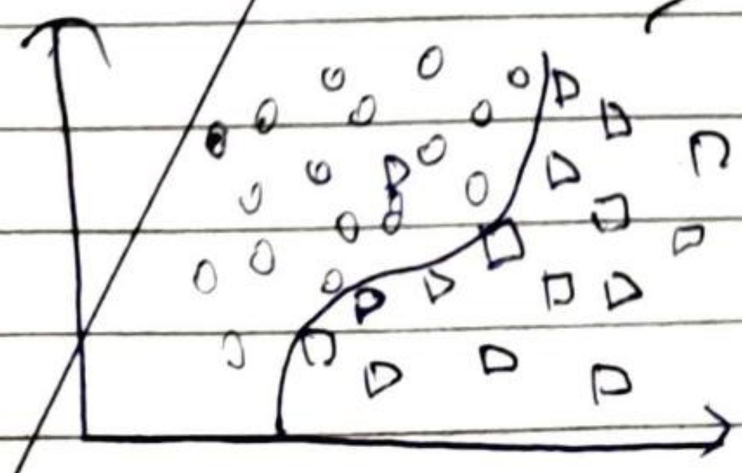
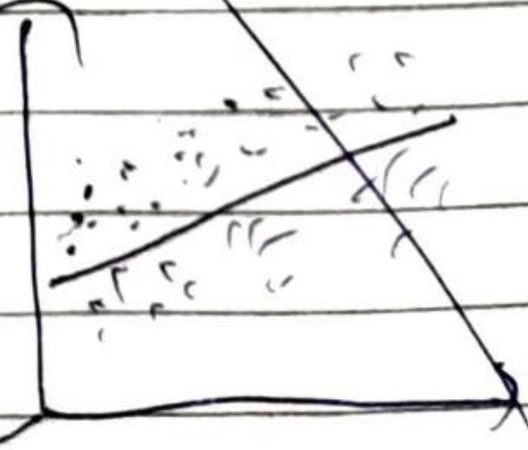
PAGE :
DATE: / /

Q What is classification?

⇒ The classification algorithm is a supervised learning technique that is used to identifying the category of new observations on the basis of training data.
In classification, a program learns from the given dataset & then classifies new observation into a number of classes or groups. such as yes/no, 0/1, spam/not spam, cat/dog etc.

Q Classification vs Regression

⇒

Regression algo	Classification algo
(1) In regression, the output variable must be of continuous nature or real value. such as price, salary, age etc.	(1) In classification, the output variable must be a discrete value. such as male/female, true/false, spam/not spam etc.
(2) Regression algo are used with continuous data.	(2) Classification algo are used with discrete data.
(3) In regression, we try to find the best fit line, which can predict the O/P more accurately	(3) In classification, we try to find the decision boundary, which can divide the dataset into different classes.
(4) 	(4) 

Wrong Order of graph

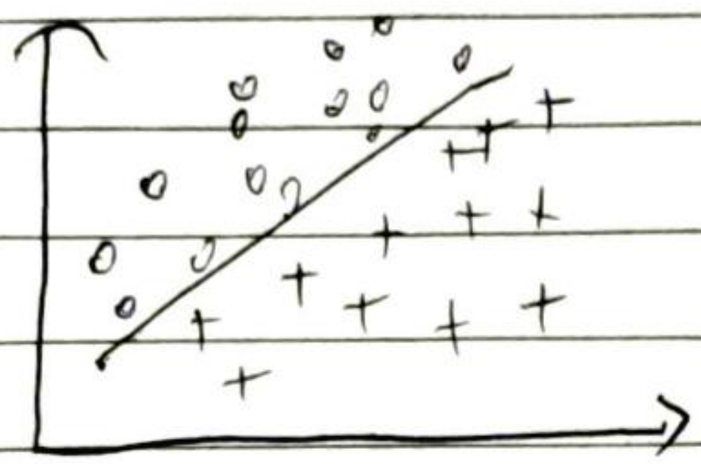
PAGE :
DATE: / /

Q Linear vs non linear classification.

=> Linear →

① Linear classification refers to categorizing a set of data points into a discrete class based on a linear combination of its variables

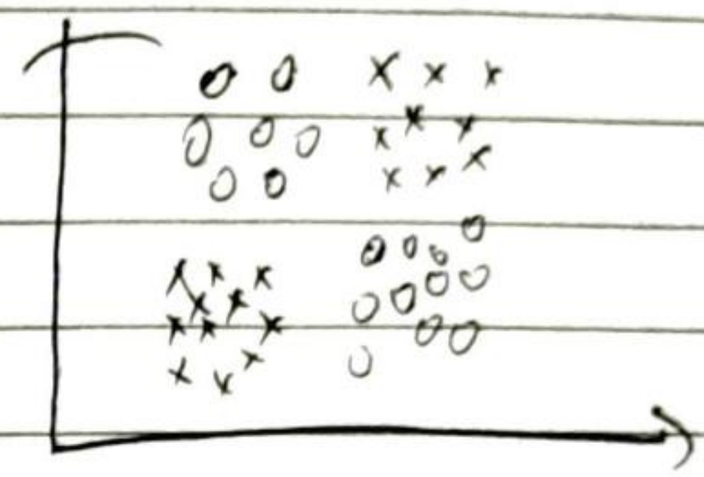
②



Non linear →

① Non linear classification refers to separating those instances that are not linearly separable.

②



Q Binary classification?

=> In binary classification, any of the samples from the dataset takes only one label out of two classes.
Ex mail spam filter.
2 classes only

Q multi class vs multi label classification?

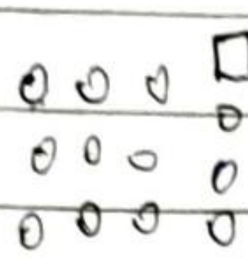
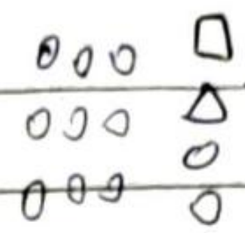
=> multi class →

① It is the problem of classification in ML where our task is to classify the more than two classes.

Multi label →

PAGE :
DATE: / /

=>

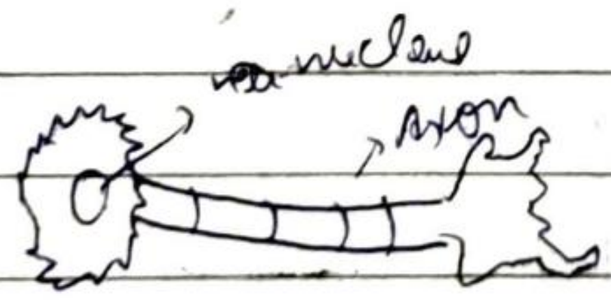
multi class	multi label
① Each input will have only one output class	① Each input will have multi output class
②  only one o/p class at a time	②  can have multiple o/p class at on.
③ Ex an animal can be cat or dog at a time.	③ Ex news article can be about sports, person, location at same time.

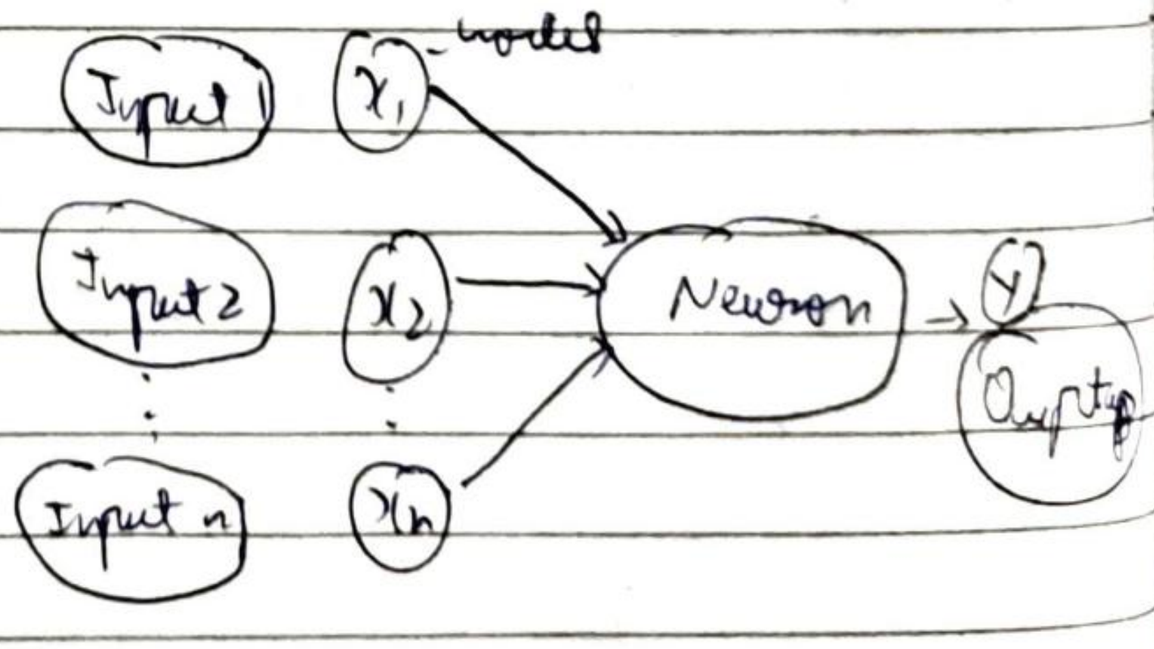
Q What are neural networks?

=>. The term artificial neural networks 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 networks

- These neurons are known as nodes.





PAGE :
DATE: / /

Q What is Perceptron?

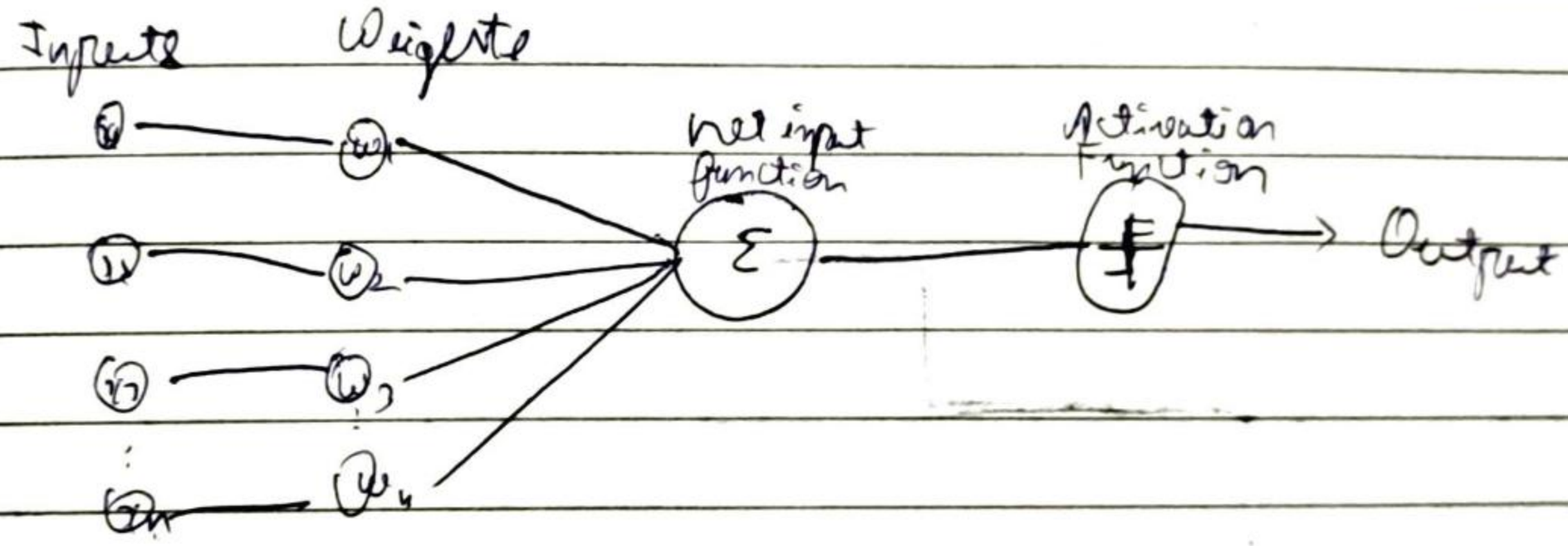
=> Perceptron is a algo for supervised learning of various binary classification tasks.

- Perceptron is also understood as an

Q What is Perceptron?

=> A perceptron is an algorithm for supervised learning of binary classifiers.

- The algorithm enables neurons to learn & process elements in the training set one at a time.



- Inputs nodes or Input layers ->

This is the primary component of perceptron which accepts the initial data into the system for further processing. Each node contains real numerical value.

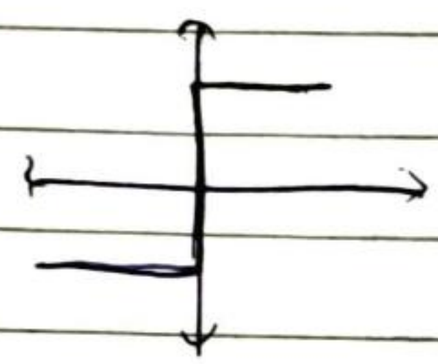
PAGE :
DATE: / /

- Weights :

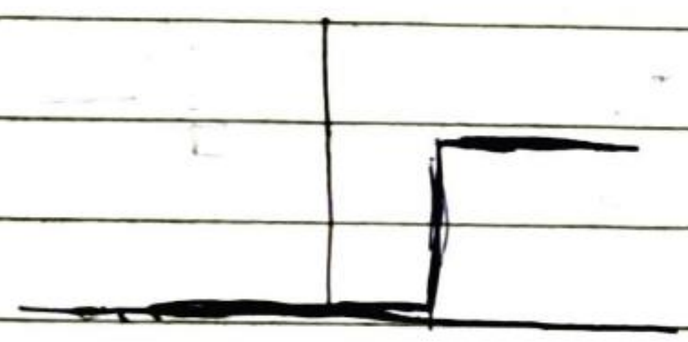
Weight parameter represents the strength of the connection between ~~unit~~ units.

- Activation Function →

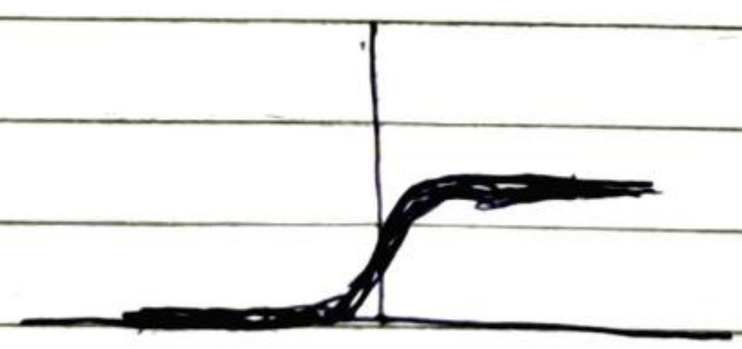
These are the final ~~compon~~ components that help to determine whether the neuron will fire or not.
These are of three type →
(i) Sign function



(ii) Step function



(iii) Sigmoid function



PAGE :
DATE: / /

Q Types of Perceptron models?

=> There are of two types → Single Layer Perceptron & multi Layer Perceptron.

① Single Layer Perceptron model →

- A single layer perceptron model consists feed forward network & also ~~in the~~ includes a threshold transfer function inside the model.

② multi-layered Perceptron model →

- Like a single layered perceptron model, a multi-layer perceptron model also has same structure but has greater number of hidden layers.

- It occurs in two stages →
  (i) Forward stage → Activation function start from the input layer in the forward stage & terminate on the output layer.
  (ii) Backward stage → In the backward stage, weight & bias values are modified as per the model's requirement.

- Advantages →
  (i) It can also solve non-linear problems
  (ii) It works well with both small & large input.

- Disadvantages →
  (i) Computations are difficult & time-consuming.
  (ii) It is difficult to predict affect of dependent variable

PAGE :
DATE: / /

Q What is logistic regression?

=> Logistic regression predicts the ~~outpt~~ output of a categorical dependent variable.

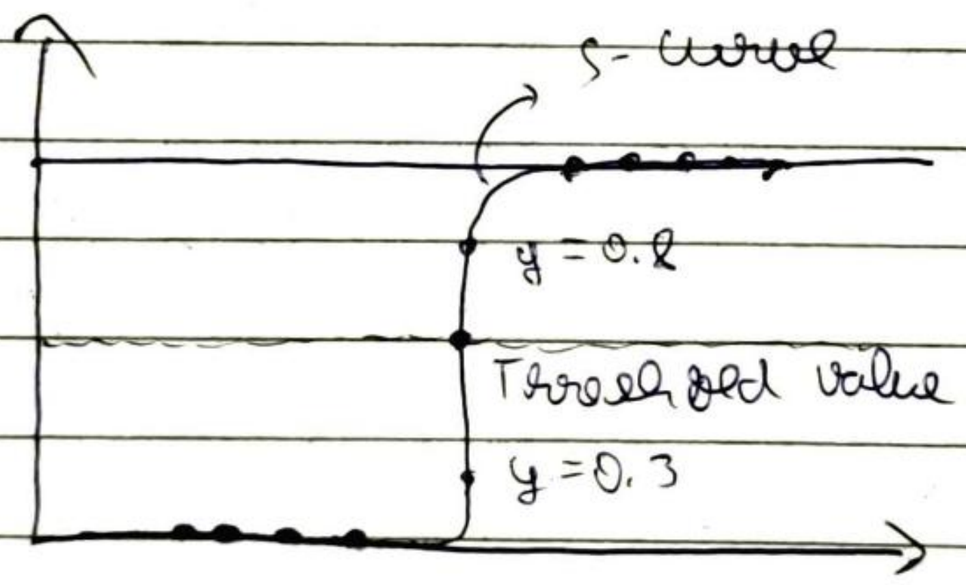
- Therefore the outcome must be a categorical or discrete value.

- ~~It can be either yes or no, 0 or 1, true or false~~

- Logistic regression gives the probabilistic values which lie between 0 & 1

- In logistic regression we fit a logistic function called sigmoid function which predicts from two maximum values (0 or 1).

- Due to the sigmoid function, the 'S' shaped graph is formed.



- Advantages -)

- It is designed for classification, & is most useful for understanding the influence of several independent variables on a single outcome variable

PAGE :
DATE: / /

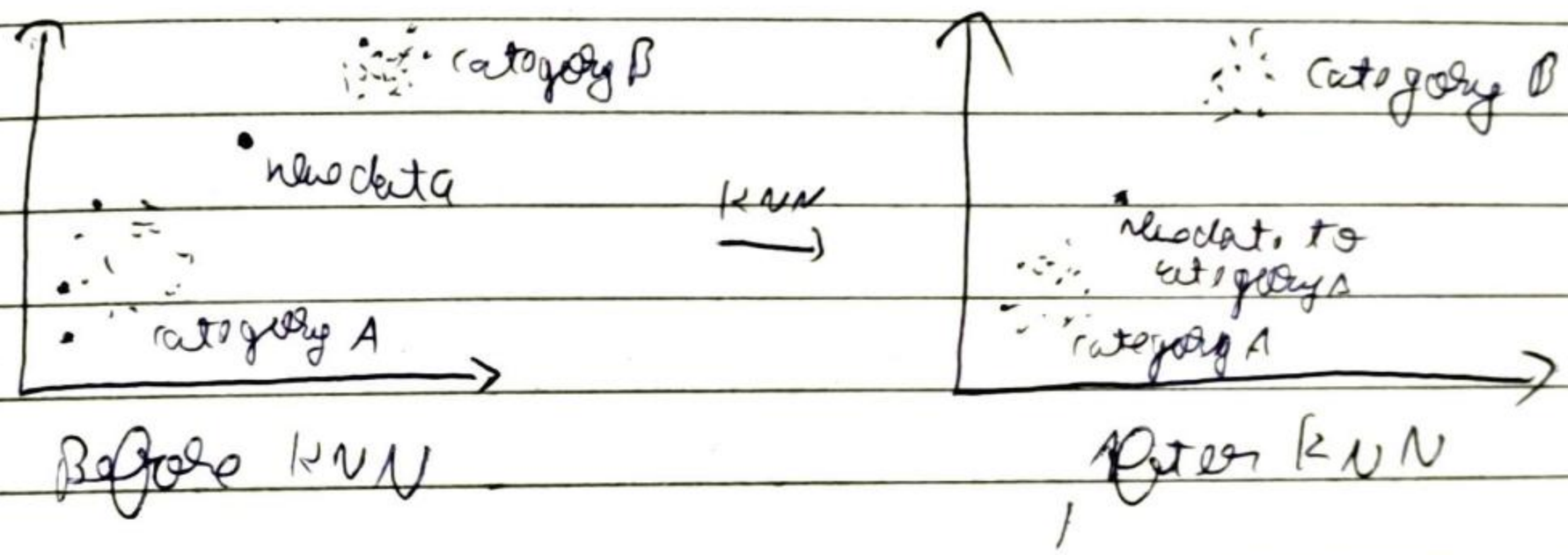
- Disadvantage →
- Works only when the predicted variable is binary.

Q What is K - nearest neighbours (KNN)?

→ KNN assumes the similarity b/w the new data & available data & put the new data into the category that is most similar to the available categories.

- KNN stores all the available data & classifies a new data point based on similarity.

- KNN uses Euclidean distance ~~to~~ between data points which is used to determine the distance b/w two points of K neighbours.



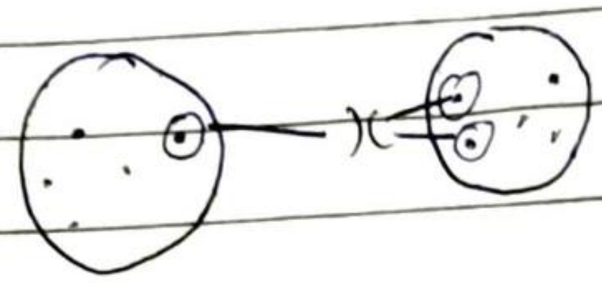
- Advantages →
  (i) It is simple to implement

- Disadvantage →
  (i) need to determine the value of K

DATE: / /

Q What is K in KNN?

⇒ In KNN, K is just the number of neighbours voting on the test example's class



If K=3, x will take voting from 3 nearest neighbours

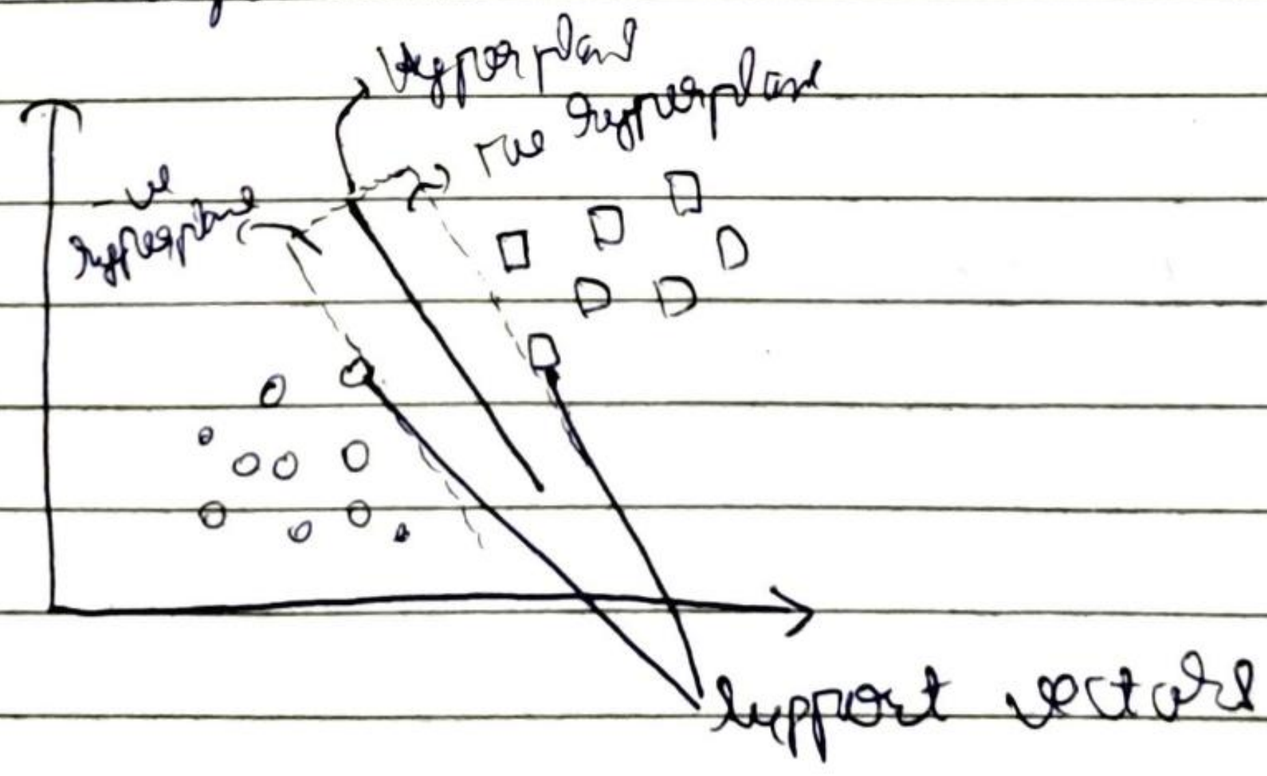
Q What is support vector machine?

⇒ • The goal of SVM algorithm is to create the best line or decision boundary that can separate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future.

• This best decision boundary is called a hyperplane.

• SVM chooses the extreme points that help in creating the hyperplane.

- These extreme points are called as support vectors, & hence algorithm is SVM.



PAGE :
DATE: / /

- Advantages →
(i) Effective in high dimensional spaces

- Disadvantages →
(i) It performs expensive calculations

Q What is naive Bayes?

⇒ - naive Bayes algo is a supervised learning algo, which is based on Bayes theorem.

- It is a probabilistic classifier, which means it predicts on the basis of the probability of an object.

- Bayes Theorem →

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

- Advantages →
(i) It is extremely fast

- Disadvantages →
(i) It is known to be a bad estimator

PAGE :
DATE: / /

Q Two machines are utilized to produce screwdrivers. First machine produces 35 screwdrivers/hr & second machine produces 45 screwdrivers/hr. Out of the total screwdrivers produced, 1.5% are defective & each machine contributes equally in defective parts. Using Bayes theorem, calculate the probability that the screwdriver produced by machine 1 is defective.

=> Production by machine 1 = 35 screwdrivers / hr
Production by machine 2 = 45 screwdrivers / hr

Defective screwdrivers = 1.5%

Contribution of machine 1 in defective parts = 50%

Contribution of machine 2 in defective parts = 50%

$$P(M_1) = \frac{35}{80} = 0.437$$

$$P(M_2) = \frac{45}{80} = 0.562$$

$$P(D) = 0.015 \quad (1.5 \div 100)$$

$$P(M_1 / D) = 0.5$$
$$P(M_2 / D) = 0.5$$

$$P(D / M_1) = ?$$

By Bayes thm

$$P(D/M_1) = \frac{P(M_1/D)\ P(D)}{P(M_1)}$$

$$= \frac{0.5 \times 0.015}{0.437}$$

$$= 0.017$$

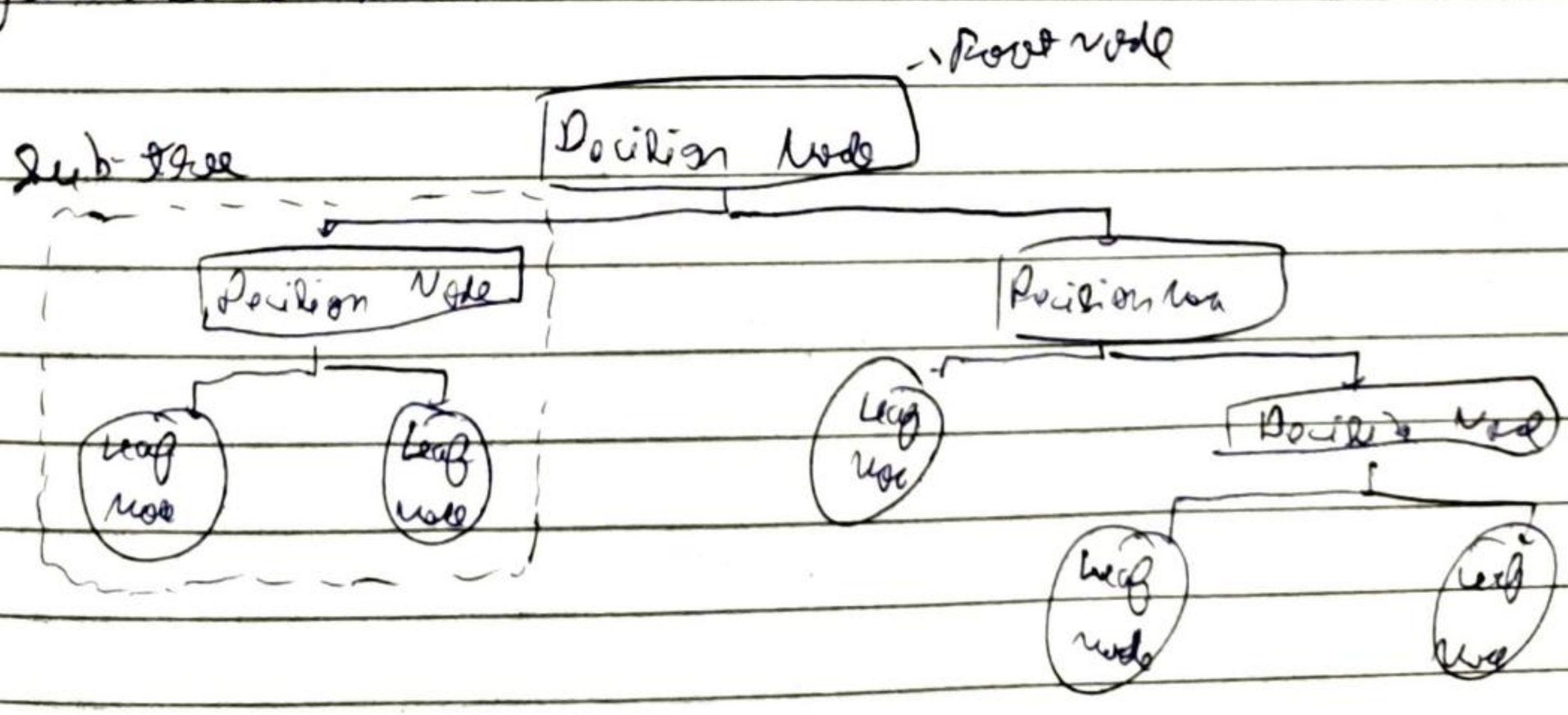
i.e. 1.7% Ans

Q What is decision tree classification?

⇒ • Decision tree is a tree structured classifier, where internal nodes represent the features of dataset, branches represent the decision rules & each leaf node represents outcome.

• In a decision tree, there are two nodes, which are decision node & leaf node. Decision nodes are used to make any decision & have multiple branches, whereas leaf nodes are the output of those decisions & do not contain any further branches.

• Sub-tree



PAGE :
DATE: / /

- Advantages :-
  (i) It is simple to understand & visualize.

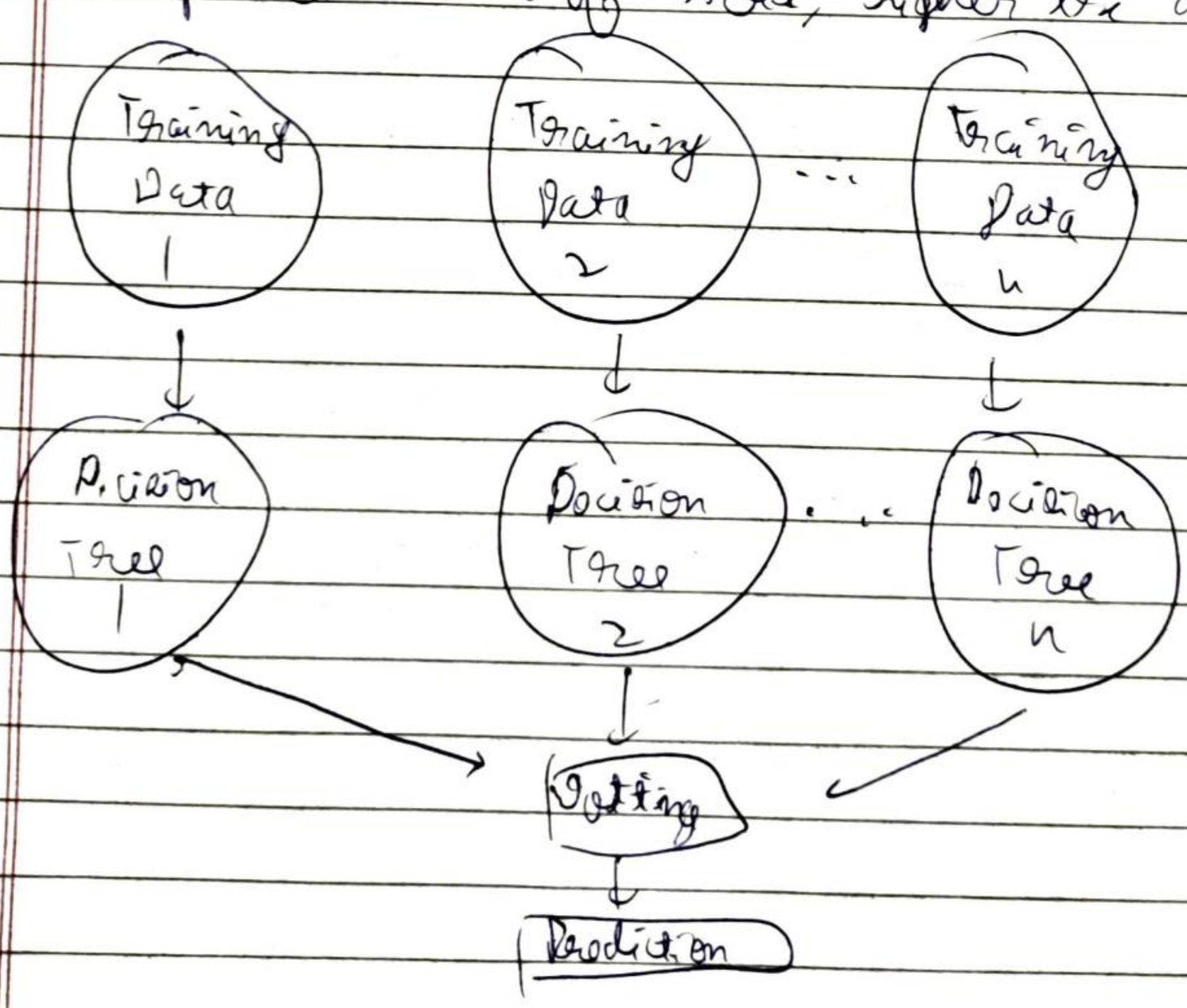
- Disadvantages :-
  (i) Sometimes trees can be complex.

Q What is Random Forest Classification?

=>. It is based on the concept of ensemble learning which ~~and~~ is a process of combining multiple classifiers to solve a complex problem & to improve the performance of the model.

- It contains no. of decision trees on various subsets of the given dataset & takes the average to improve the accuracy.

- The greater the no. of trees, higher the accuracy.



PAGE :
DATE: / /

Advantage →

(i) Reduction in over fitting

Disadvantage →

(i) It is slow, difficult to implement.

Q Diff b/w Regression & Classification ?

=) **Regression →**

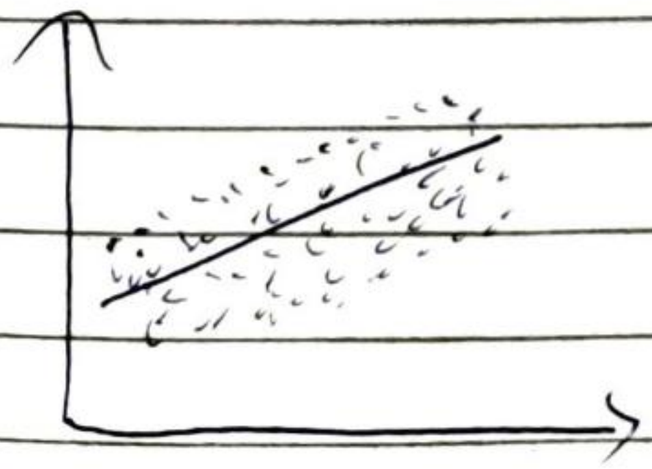
① In regression, the output variable must be of continuous nature or real value.
Such as price, salary, age

② Regression algo are used with continuous data

③ In regression, we try to find the best fit line, which can predict the output more accurately

④ It can be used to solve problems such as house price prediction etc

⑤



Classification →

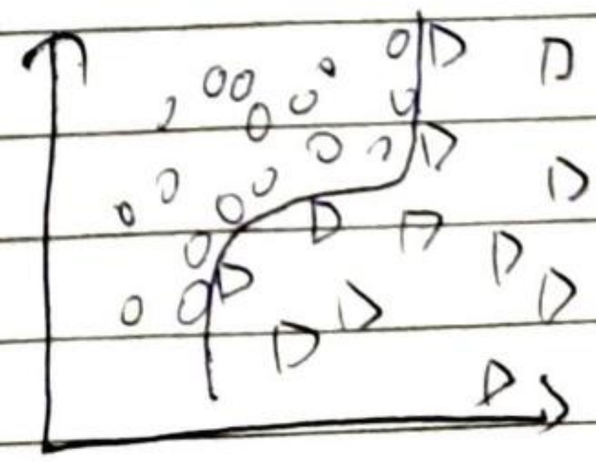
① In classification the output variable must be a discrete value. Such as male/female, yes/no, 0/1.

② Classification algo are used with discrete data

③ In classification we try to find the decision boundary, which can divide dataset into different classes

④ It can be used to find spam mails, cancer cells etc

⑤



PAGE :
DATE: / /

Q What is clustering?

A. It can be defined as way of grouping the data points into different clusters, consisting of similar data points. The objects with the possible similarities remain in a group that has less or no similarities with another group.

- It is an unsupervised learning method, hence no supervision is provided to the algorithm.

- It deals with unlabeled dataset.

Q Diff b/w Classification & Clustering?

A

Classification →	Clustering →
① The process of classifying the input data based on the corresponding class labels	① The grouping of the data based on their similarity without the help of class labels.
② It is used for supervised learning	② It is used for unsupervised learning.
③ It has labels so there is need of training & testing data	③ There is no need of training & testing data
④ It is more complex as compared to clustering	④ It is less complex as compared to classification.

PAGE :
DATE: / /

(S) Ex algos → Logistic regression, Naive Bayes, SVM etc

(C) Ex K-means clustering, K-mode etc.

Q What is Hierarchical Clustering?

=) In this technique, the dataset is divided into clusters to create a tree like structure, which is called a dendrogram.

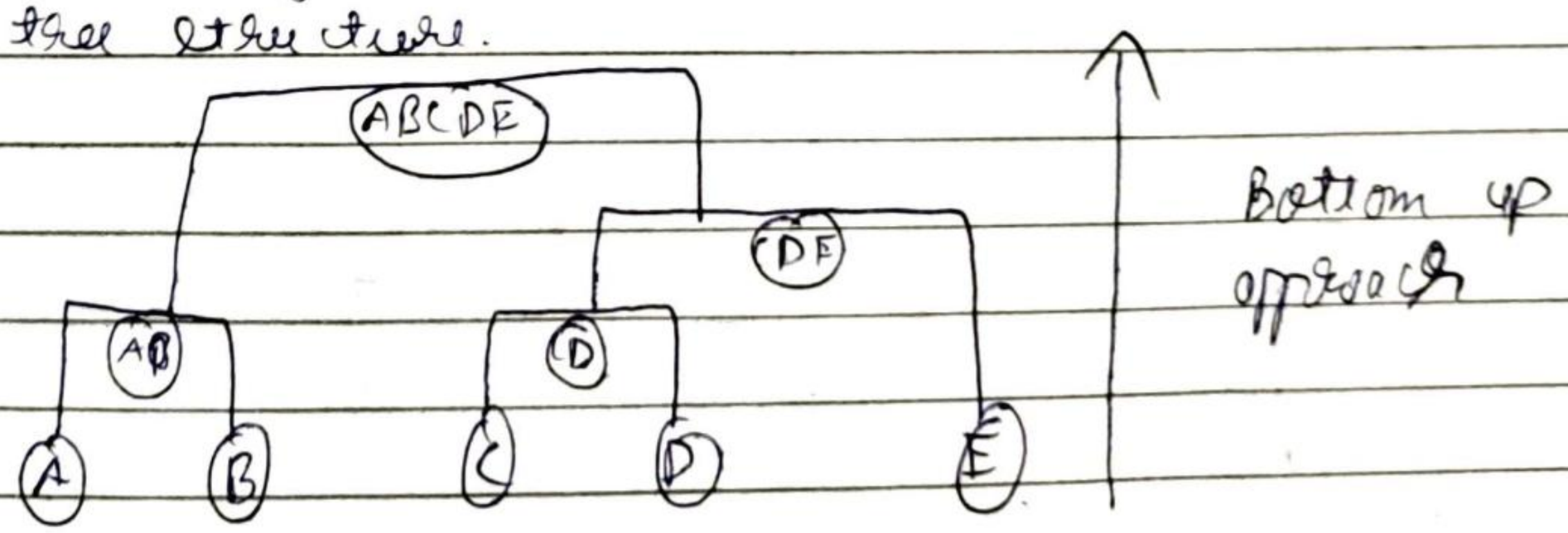
~~The Observations It is hierarchical~~

- It is another unsupervised learning algorithm that is used to group together the unlabeled data points having similar characteristic

- It has 2 categories → Agglomerative & Divisive

- Agglomerative Hierarchical ~~Cluster~~ algorithm →

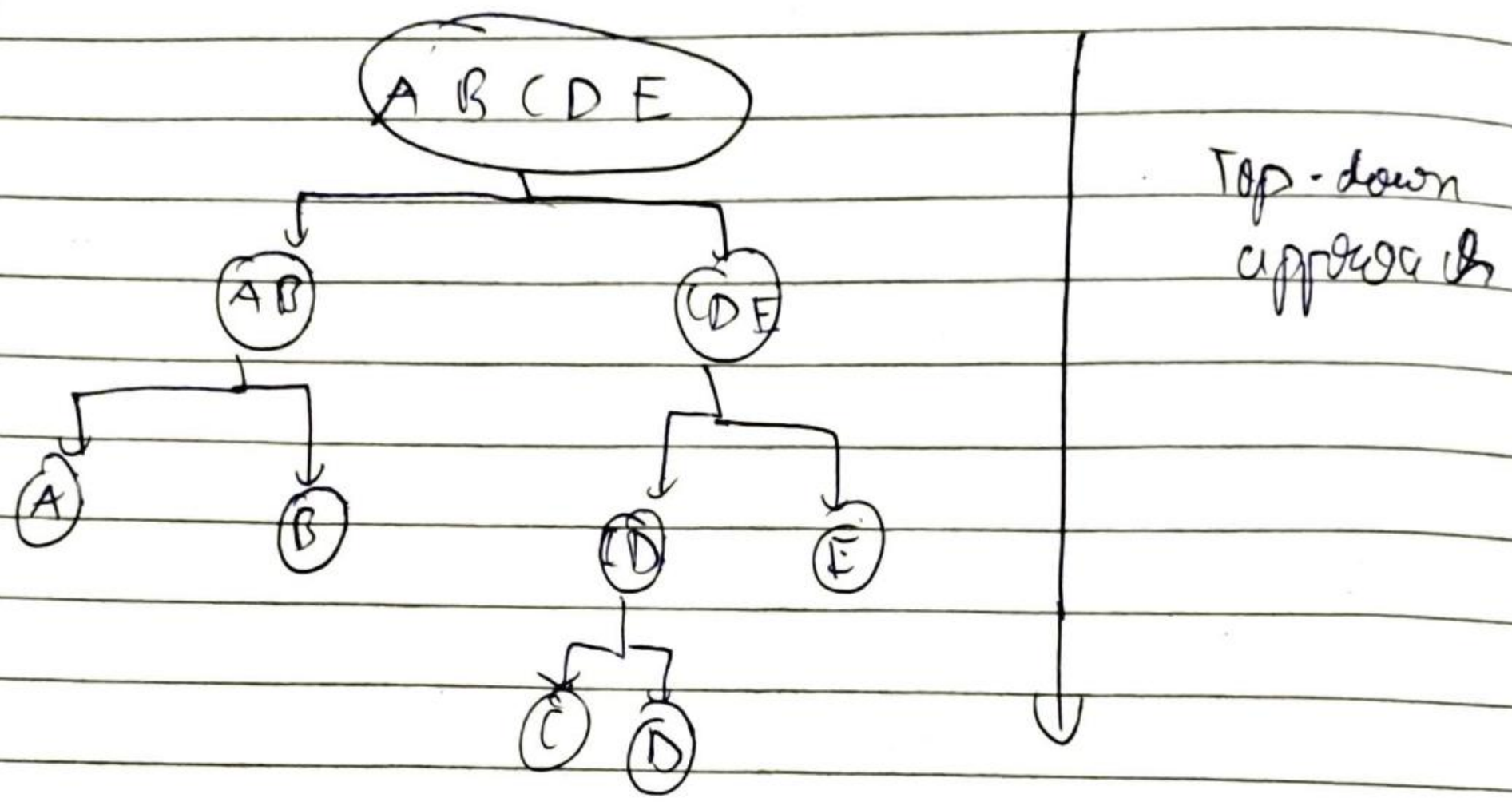
In agglomerative algo, each data point is treated as a single cluster & then recursively merge or agglomerate the pair of clusters. It follows a bottom up approach. The hierarchy of clusters is represented as a dendrogram or tree structure.



PAGE :
DATE: / /

- Divisive Hierarchical algo ->

In divisive hierarchical algo, all the data points are treated as one big cluster & the process of clustering involves dividing the one big cluster into various small clusters. It follows top down approach.



Q Working of Agglomerative Hierarchical Clustering

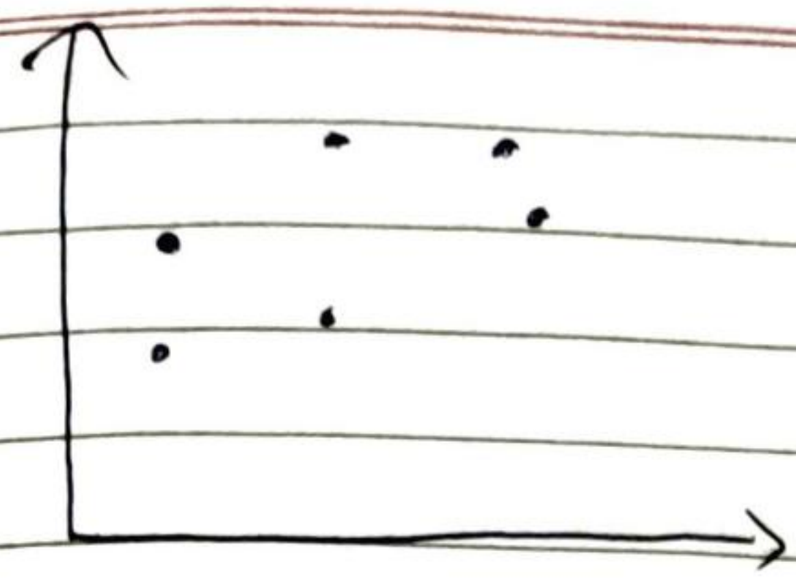
⇒ Firstly explain what agglomerative clustering is

Steps->

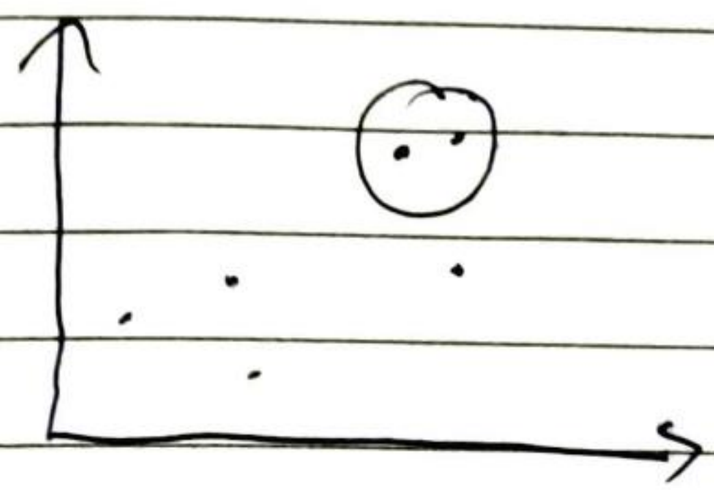
Step I->

Create each data point as a single cluster. Let's say there are N data points, so no. of clusters will be N.

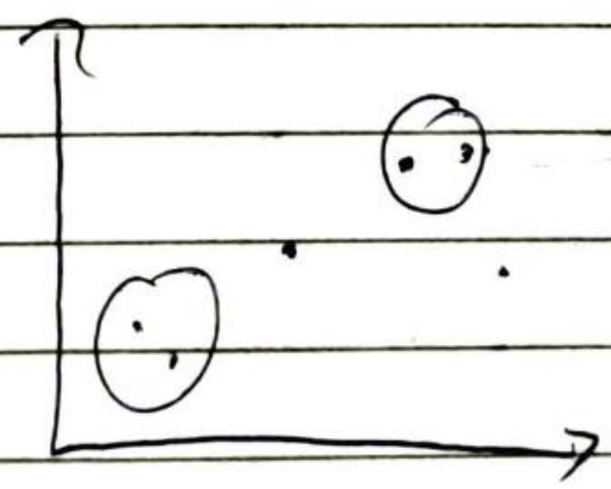
PAGE :
DATE: / /



Step 2 ->
Take two closest data points or clusters & merge them to form one cluster. Now there will be N-1 clusters.



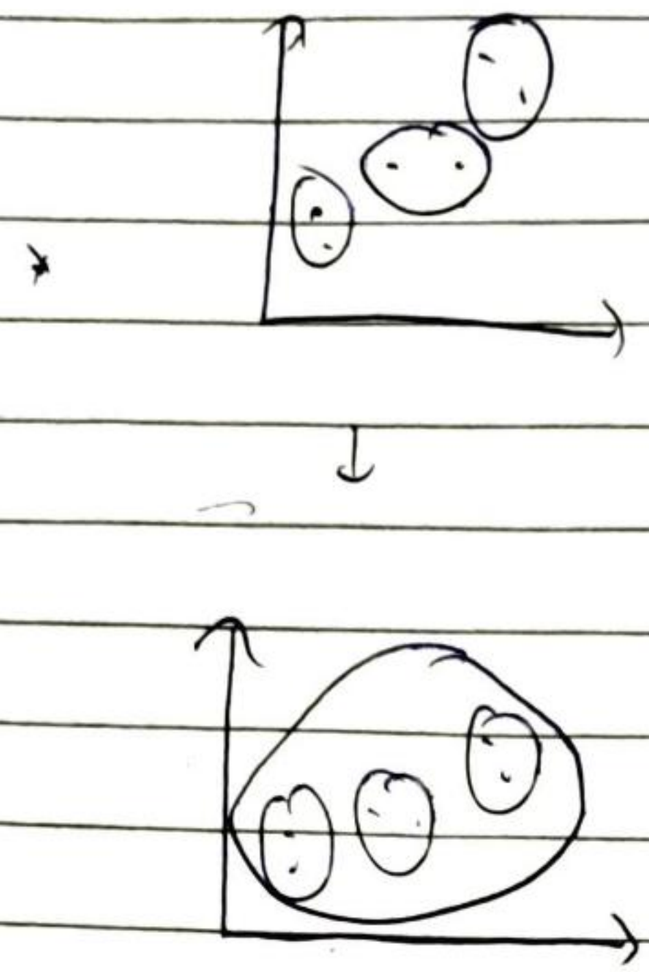
Step 3 ->
Again take the two closest clusters & merge them together to form one cluster. Now N-2 clusters are there.



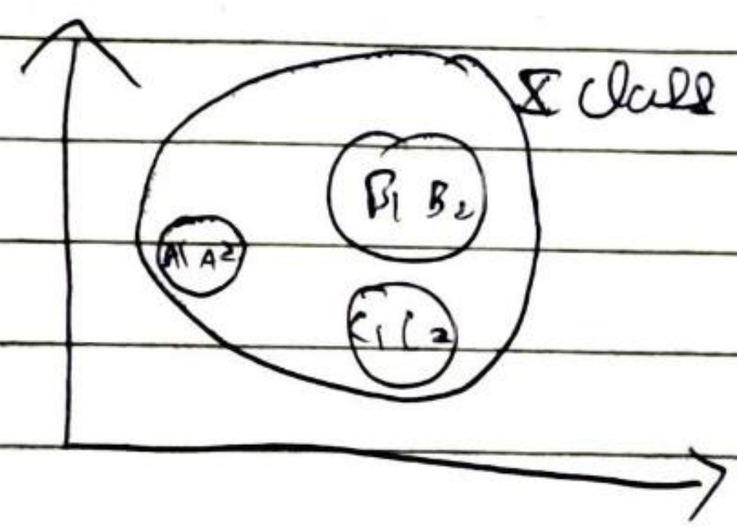
Step 4 ->
Repeat ~~the~~ step 3 until only one cluster is left.

PAGE :
DATE: / /

Steps ->
All the clusters are combined & made a single cluster



- Ex students of same class in different sections.



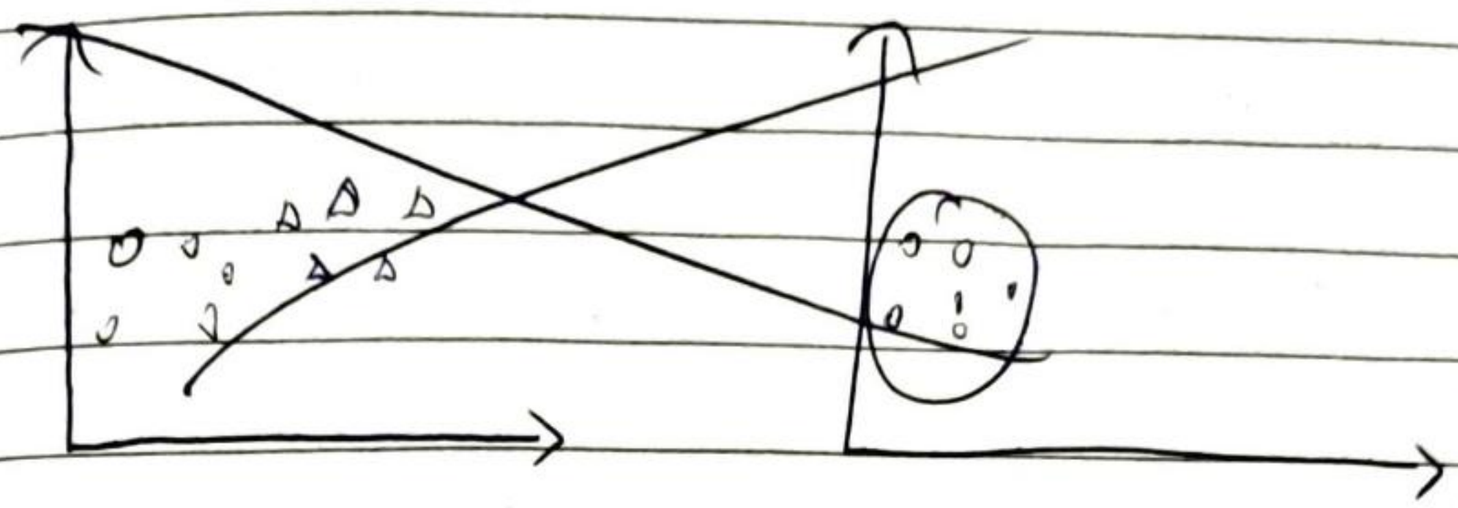
Q Why clustering is called unsupervised learning?

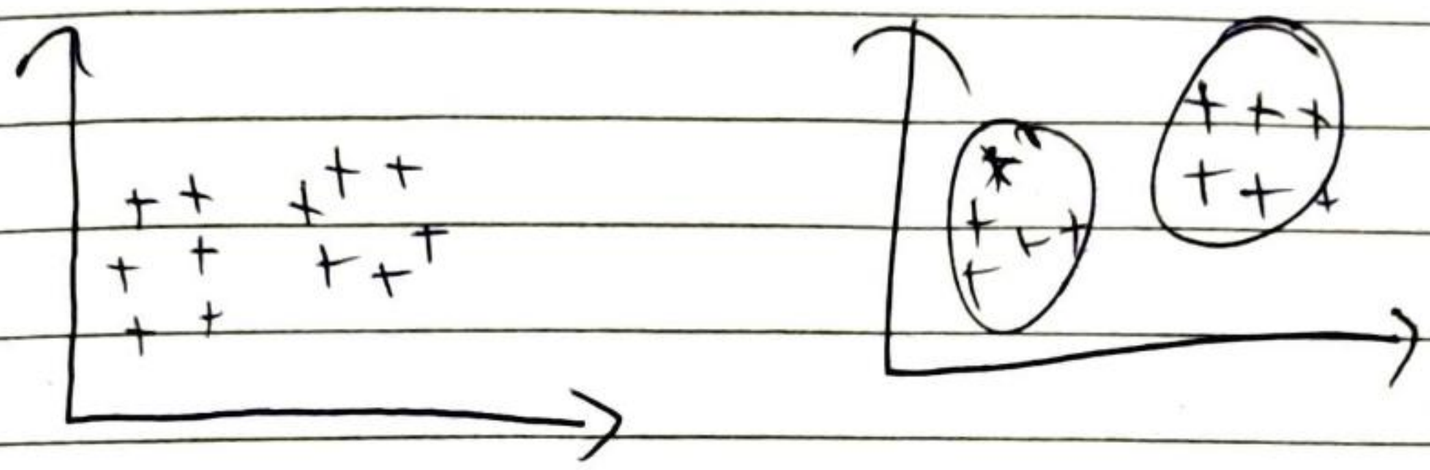
=> Clustering is called unsupervised learning because of the unlabelled datasets.

- The datasets provided in unsupervised learning are not organised & labelled.

PAGE :
DATE: / /

- Also the process of models is done under no supervision.
- As the dataset is unlabelled, it ~~have~~ real scattered datapoints, which are grouped together by forming clusters.





The similar datapoints are clubbed together in clusters.

PAGE :
DATE: / /

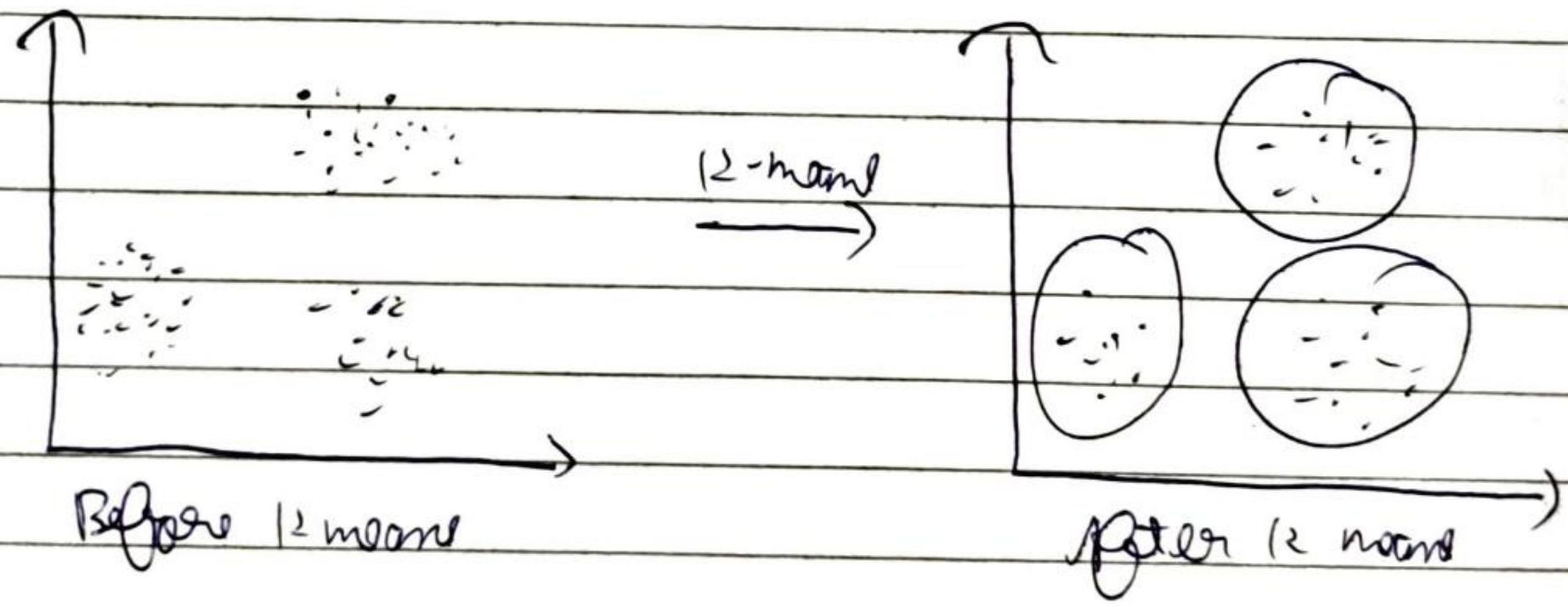
Q What is K means clustering?

→ K means clustering is an unsupervised learning algorithm which groups the unlabeled dataset into different clusters.

- Here K defines the no. of pre-defined clusters that need to be created in the process. As if K=2, there will be two clusters, K=3 there will be three clusters.

- It allows us to cluster the data into different groups & a convenient way to discover the categories of groups in the unlabeled dataset

- It is a centroid based algorithm where each cluster is associated with a centroid. The main aim of algo is to minimize the sum of distances b/w the data point & their corresponding clusters.

- 

- Advantages →
(i) It is easy to implement

- Disadvantages →
(i) Difficult to predict the number of clusters (K)

PAGE :
DATE: / /

Q Working of K means algo?

=> Step 1 ->
select the number k to decide number of clusters.

Step 2 ->
select random k points or centroids.

Step 3 ->
assign each data point to their closest centroid.

Step 4 ->
Calculate the variance & place new centroid of each cluster.

Step 5 ->
Repeat the steps 3 & 4.

Step 6 ->
The model is ready.

Q How to choose value of K in K means?

=> The performance of K means clustering algorithm depends upon the number of clusters it forms.

- The most optimal way to choose the K is Elbow method. This method uses concept of WCSS

PAGE :
DATE: / /

Q- What is Fuzzy Logic ?

⇒ • The 'Fuzzy' word means the things that are not clear or are vague.

- Sometimes we cannot decide in real life that given problem or statement is either true or false.

- At that time, this concept provides values between true & false & gives the flexibility to find the best solution to that problem.

- In the Boolean system, only two possibilities (0 & 1) exist, where 1 denotes the absolute truth value & 0 denotes the absolute false value.

- But in fuzzy system, there are multiple possibilities present b/w the 0 & 1, which are partially false & partially true.

- Image →

It is cold → Very much (0.9), Little (0.25), Very Less (0.1)] Fuzzy Logic

PAGE :
DATE: / /

Q What is Fuzzy set?

=> Fuzzy set is a set having degree of membership b/w 0 & 1, Fuzzy sets are represented with tilde character (~). For eg. number of cars following traffic signals at a particular time of all cars present will have membership value between [0,1].

- Partial membership exists when member of one fuzzy set can also be a part of other fuzzy sets in the same universe.

- The degree of membership or truth is not same as probability, fuzzy truth represents membership in vaguely defined sets.

- A fuzzy set $\tilde{A}$ in the universe of discourse U can be defined as a set of ordered pairs & it is given by

$$\tilde{A} = \{ (x, \underbrace{\mu_{\tilde{A}}(x)}_{\text{membership function}}) \mid x \in X \}$$

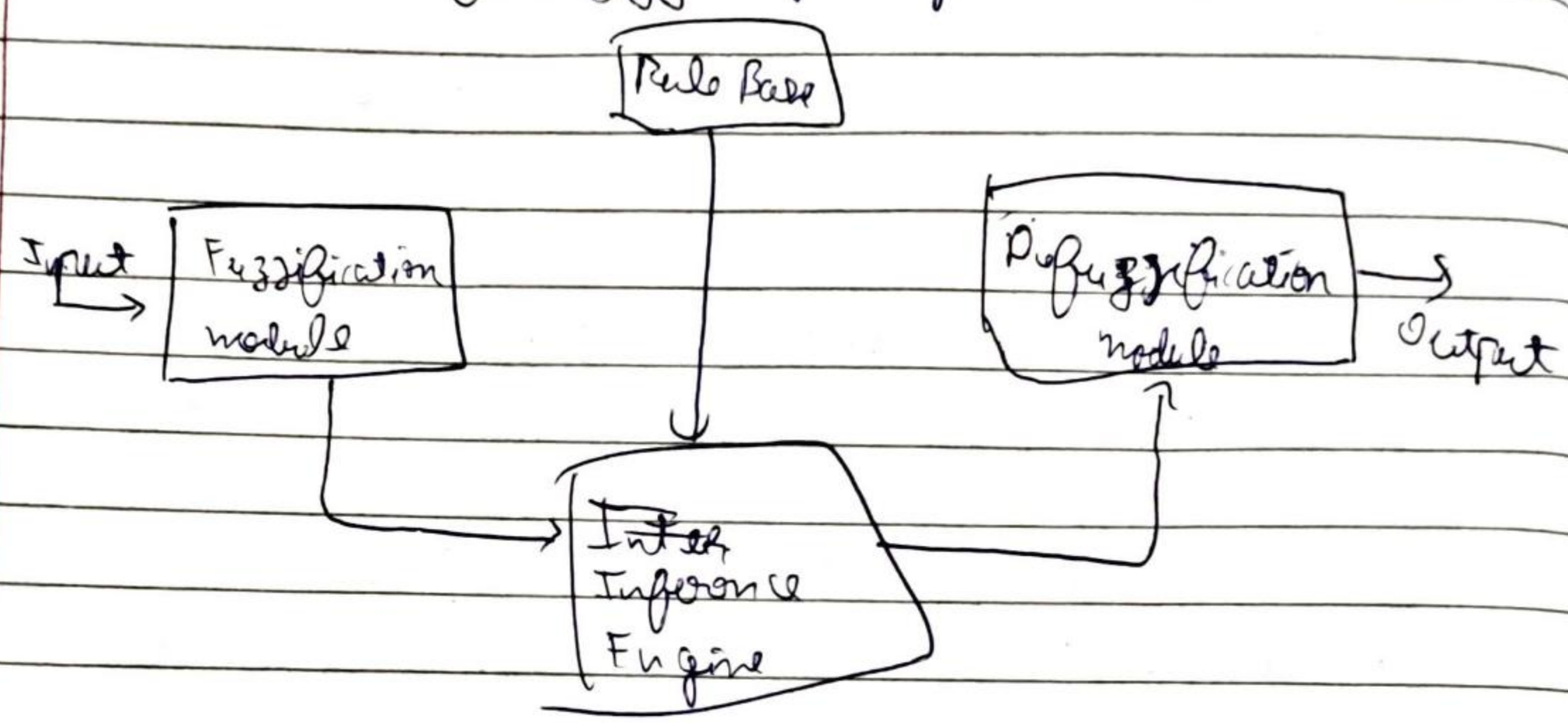
PAGE :
DATE: / /

Q What is Fuzzification?

=> It is used to convert inputs i.e. crisp numbers into fuzzy sets. Crisp inputs are basically the exact inputs measured by sensors & passed into the control system for processing, such as temperature, pressure, rpm's etc.

Q Architecture of Fuzzy Logic System

=>



① Rule Base →
It is a component used for storing the set of rules & the if then conditions given for decision making

② Fuzzification module →
It is a module for transforming the inputs i.e to convert crisp number into fuzzy sets.

③ Inference Engine →
All of the info is processed here. It allows users to find the matching degree b/w fuzzy inputs & rules.

PAGE :
DATE: / /

(c) Defuzzification module →
It is a module which takes fuzzy set ~~input~~ inputs & then transform them into a crisp value.

Q- Advantages & Disadvantages of Fuzzy Logic.

→ Advantages :-

- This concept works similarly as human reasoning.
- Any user can easily understand the structure of fuzzy logic.
- It does not need large memory.
- This concept is based on the set theory of mathematics.

Disadvantages :-

- The running time of system is slow.
- The possibilities are not always accurate
- There is ~~lot of~~ need of lot of testing data.

~~Q Applications of~~

PAGE :
DATE: / /

Q What are membership functions?

=> • The membership function is a function which represents the graph of fuzzy sets, & allows us to quantify the linguistic term.

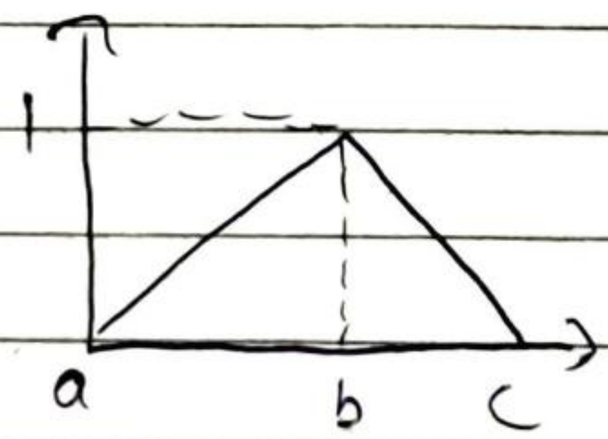
• It is a graph which is used for mapping each element of x to the value b/w 0 & 1

• For a fuzzy set A, membership function for X is defined as $\mu_{\tilde{A}}(x)$ where $x \in X$.

Different types of membership functions →

① Trimf :- Triangular membership function →

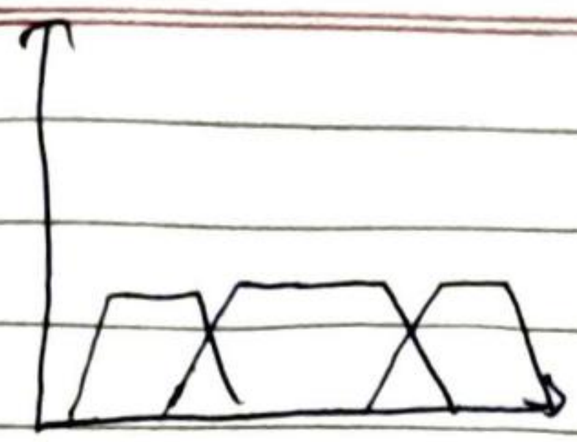
The triangle which fuzzifies the input can be defined by three parameters a, b, & c ~~where~~ where c defines the base & b defines the height of the Δ.



② ~~Tra~~ Trapmf :- Trapezoidal mf →

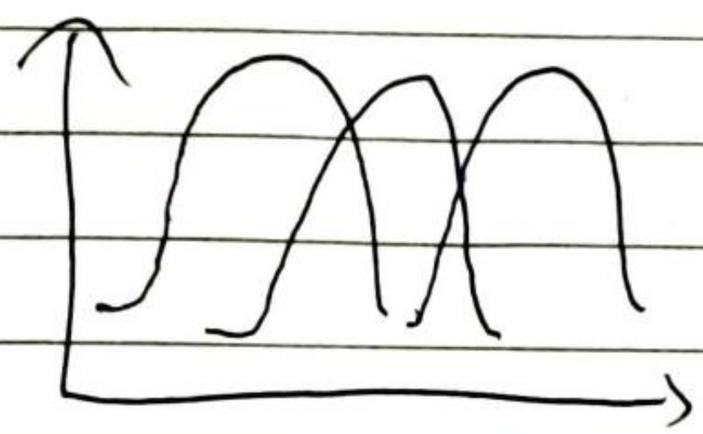
This fn computes fuzzy membership value using a trapezoidal membership function.

PAGE :
DATE: / /



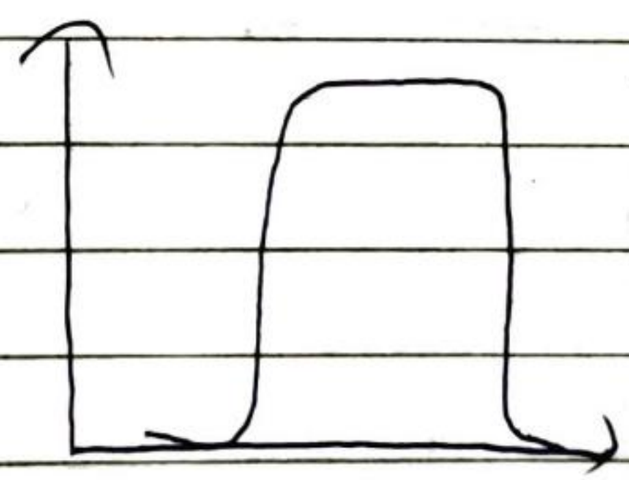
③ **gaussmf :- gaussian mf →**

The gaussian mf implements a membership fn based on a symmetric gaussian function.
The two functions are called gaussmf & gauss2mf



④ **gbellmf :- generalized bell mf →**

It is a symmetrical shape similar to a bell.



$X^* =$ crisp value

PAGE :
DATE: / /

Q What is Defuzzification?

=> It is the process of converting a fuzzified o/p into a single crisp value w.r.t fuzzy set. The defuzzified value in FLC represents the action to be taken in controlling the process.

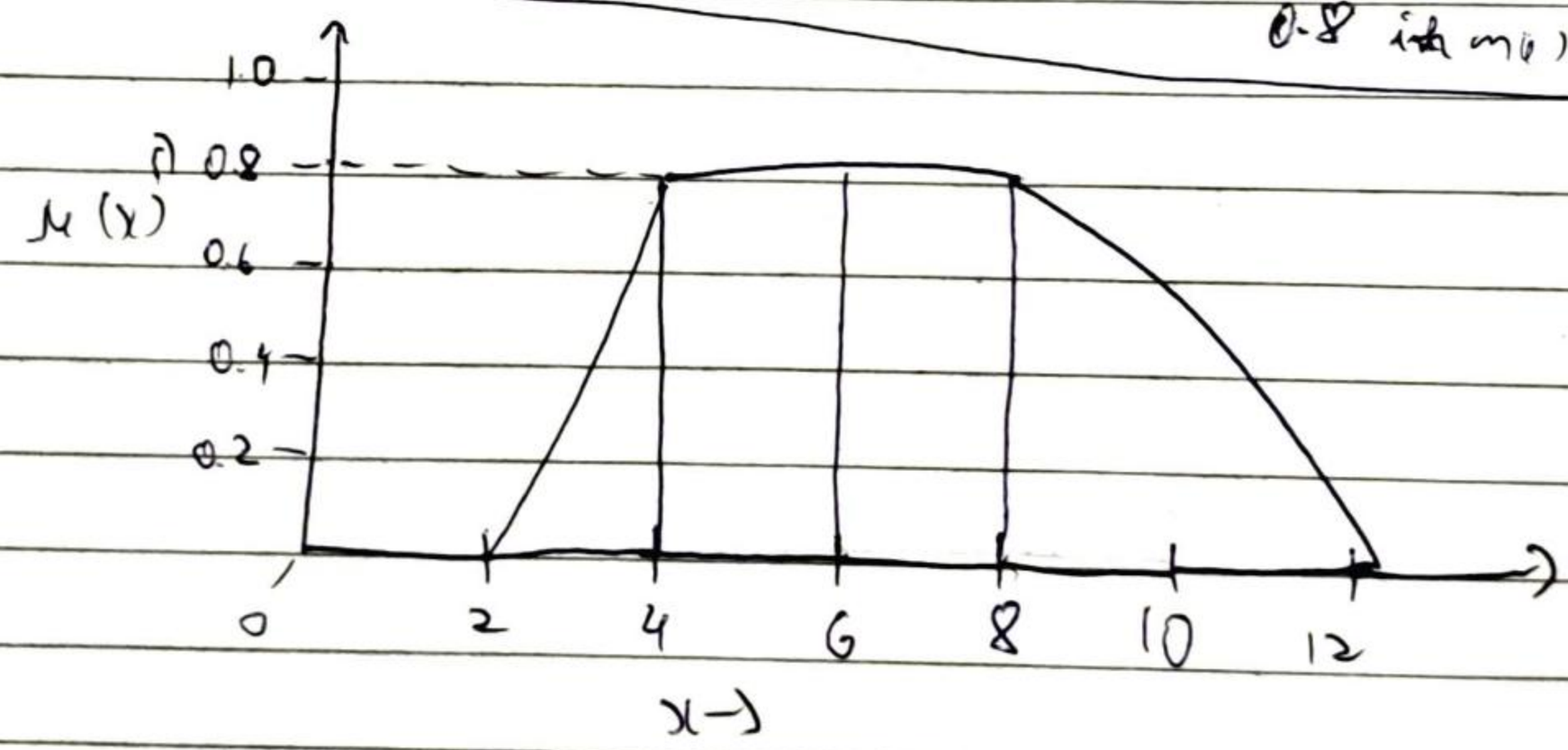
Q methods of Defuzzification.

=> (1) <u>maxima methods</u> →

It contains three sub methods
First of maxima, Last of maxima, mean of maxima

note
Height of fuzzy set = maxima (max^m $\mu(x)$ value)

0.8 is maxima



$$X^* = \frac{\sum x_i \in M \; x_i}{|M|}$$

$M = \{ x \mid \mu_A(x) =$ height of fuzzy set $\}$

PAGE :
DATE: / /

(I) First of maxima (FOM) →

Observe first value where membership value = height of graph

$$\boxed{x^* = 4}$$

(II) Last of maxima (LOM) →

Observe last value where membership value = height of graph

$$\boxed{x^* = 8}$$

(III) Mean of maxima (MOM) →

$$x^* = \frac{\sum x_i \in M \; x_i}{|M|}$$

$M = \{ x \mid \mu_A(x) = \text{height of fuzzy set} \}$

$x = 4, 6, 8$

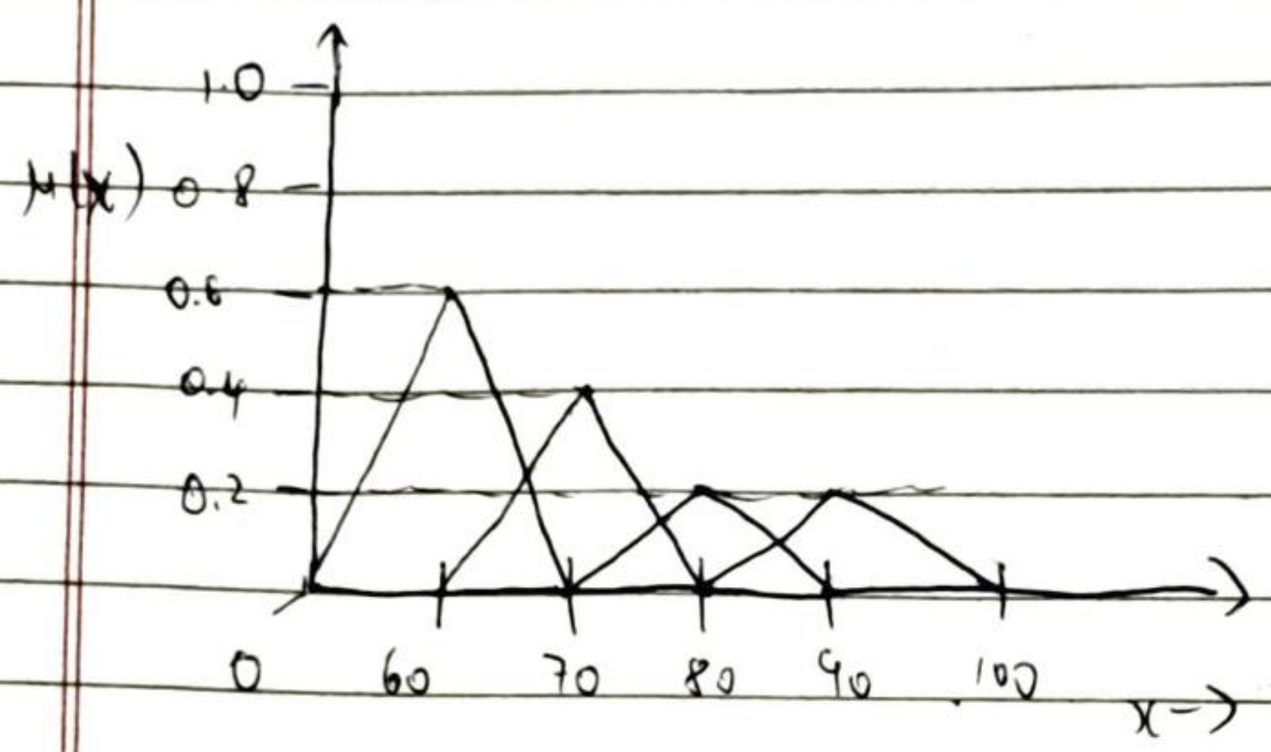
$$x^* = \frac{4+6+8}{3}$$

$$\boxed{x^* = 6}$$

PAGE :
DATE: / /

② <u>Weighted Average Method</u> :-

$$x^* = \frac{\Sigma \mu(x) \cdot x}{\Sigma \mu(x)}$$

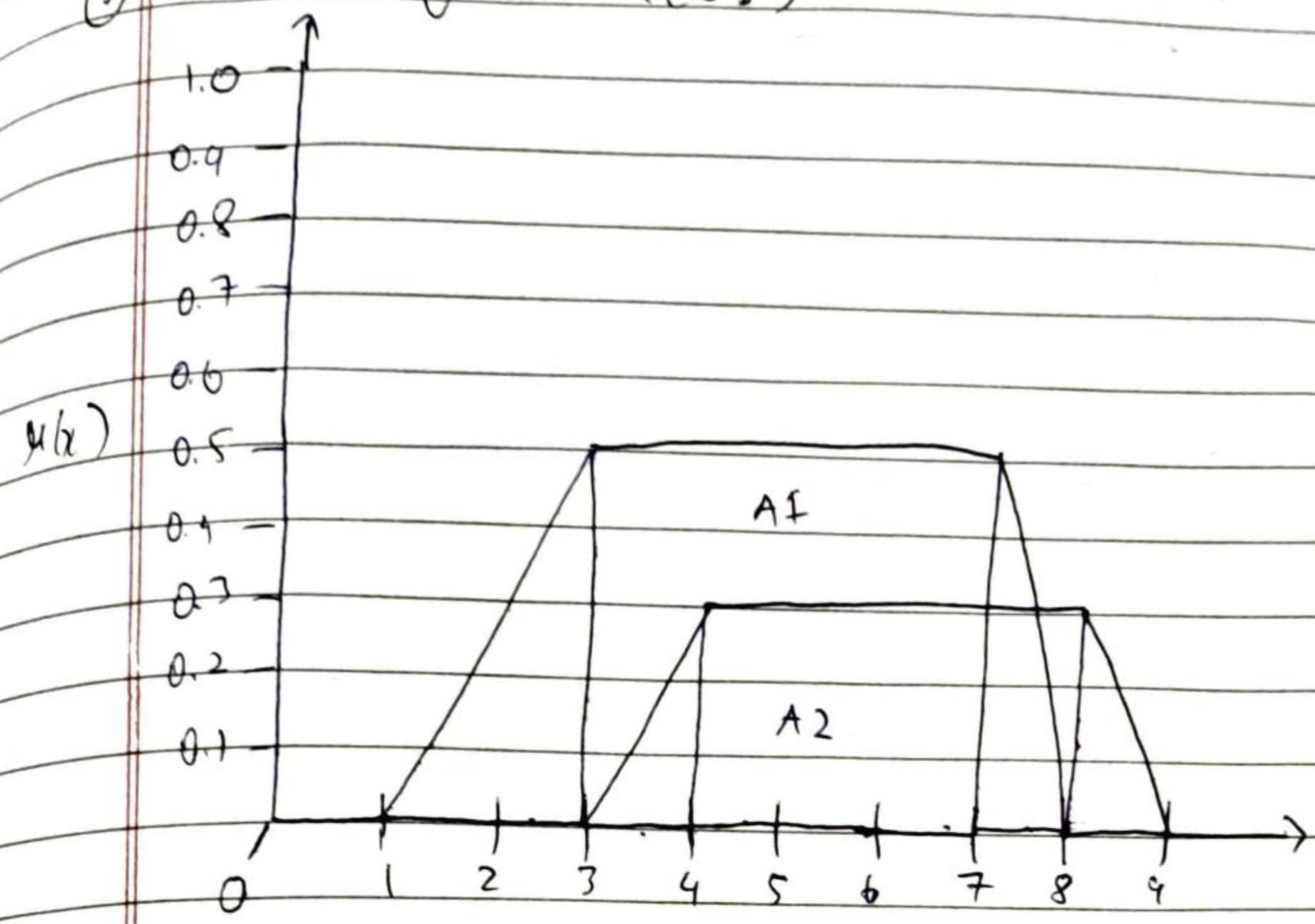


$$x^* = \frac{(60 \times 0.6 + 70 \times 0.4 + 80 \times 0.2 + 90 \times 0.2)}{0.6 + 0.4 + 0.2 + 0.2}$$

$$\boxed{x^* = 70}$$

PAGE :
DATE: / /

(7) Center of Sum (COS)



Fuzzy sets C_1 & C_2

$$X^* = \frac{\Sigma A \cdot x_c}{\Sigma A}$$

A = area of fuzzy set
x_c = center value

$$A1 = \left(\frac{1}{2} \times 2 \times 0.5\right) + \left(4 \times 0.5\right) + \left(\frac{1}{2} \times 1 \times 0.5\right)$$

$$= 0.5 + 2 + 0.25$$

$$= 2.75$$

$$A2 = \left(\frac{1}{2} \times 1 \times 0.3\right) + \left(4 \times 0.3\right) + \left(\frac{1}{2} \times 1 \times 0.3\right)$$

$$= 0.15 + 1.2 + 0.15$$

$$= 1.5$$

PAGE :
DATE: / /

$$X^* = \frac{A_1 X_{c_1} + A_2 X_{c_2}}{A_1 + A_2}$$

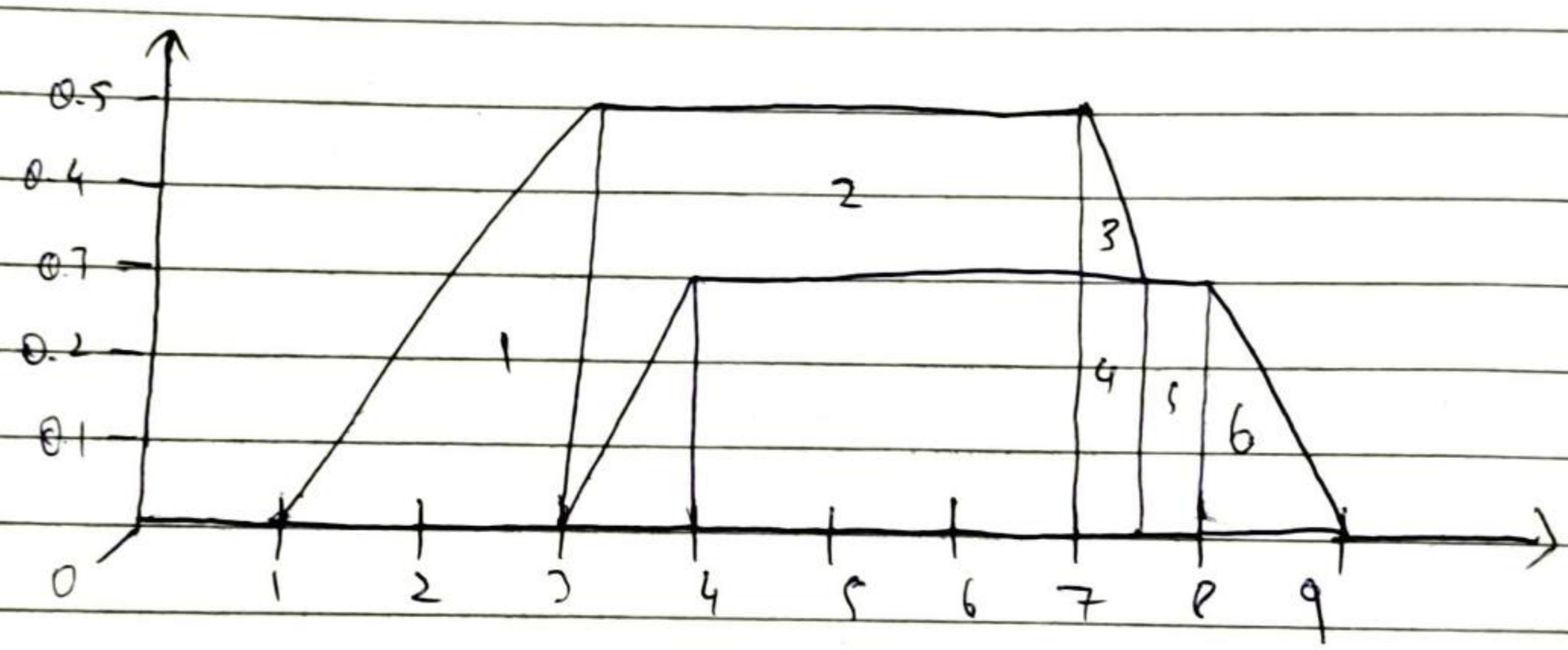
$X_{c_1} = 5$

$X_{c_2} = 6$

$$X^* = \frac{2.75 \times 5 + 1.50 \times 6}{2.75 + 1.5}$$

$$\boxed{X^* = 5.35}$$ Ans

④ <u>Center of gravity (COG) →</u>



Total area divided into 6 subarea

$$X^* = \frac{\sum_{i=1}^{n} x_i A_i}{\sum_{i=1}^{n} A_i}$$

x_i = centroid of i

A_i = area of i

PAGE :
DATE: / /

We calculate area & centroid

T.A. of subarea 1 $= \frac{1}{2} \times 2 \times 0.5 = 0.5$

TA of subarea 2 $= (7-3) \times 0.5 = 2$

TA of subarea 3 $= \frac{1}{2} \times (7.5-7) \times 2 = 0.05$

TA of subarea 4 $= 0.5 \times 0.3 = 0.15$

TA of subarea 5 $= 0.5 \times 0.3 = 0.15$

TA of subarea 6 $= \frac{1}{2} \times 1 \times 0.3 = 0.15$

Centroid of subarea 1 $= \frac{1+3+3}{3} = 2.33$

Centroid of subset 2 $= \frac{7+3}{2} = 5$

Centroid of subset 3 $= \left(\frac{7+7+7.5}{3}\right) = 7.16$

Centroid of subset 4 $= \frac{7+7.5}{2} = 7.25$

Centroid of subset 5 $= \frac{7.5+8}{2} = 7.75$

Centroid of subset 6 $= \frac{8+8+9}{3} = 8.33$

$$x^* = \frac{\sum_{i=1}^{n} x_i A_i}{\sum_{i=1}^{n} A_i}$$

$$= \frac{15.024}{3}$$

$$= 5.008$$

$$\boxed{x^* = 5.008}$$

Q How performance of two M.L models can be evaluated on the given data?

=> Confusion matrix can be constructed to evaluate the performance.

- A confusion matrix is a correlation between the predictions of a model & the actual class labels of the data points.

-

		Predicted	
		No	yes
Actual	No	TN	FP
	yes	FN	TP

PAGE :
DATE: / /

- **Accuracy**

It is correct predictions / total predictions

$$= \frac{TP + TN}{TN + FP + FN + TP \text{ (total)}}$$

- **Error Rate**

$$\text{Error rate} = \frac{FP + FN}{TN + FP + FN + TP}$$

- **Precision & Recall →**

$$\text{Precision} = \frac{TP}{\text{Predicted yes}}$$

$$\Rightarrow \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{\text{actual yes}}$$

$$= \frac{TP}{FN + TP}$$

- **F measure →**

$$F = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$