apsara
Date:

K=3 : choose 3 minm value

we got : 2, 3, 5 (These have min ~~value~~ dist.)

1 ← 2 → Pass

1 ← 3 → Pass

2 ← 5 → Pass

Since all Pass so
Ans → Pass.

22/02/23

Naive Bayes → supervised learning algo based on Bayes theorem. classification

text classif

Naive
↳ occurrence

quick prediction

Bayes theorem

predicts on basis of probability.

$$P\left(\frac{A}{B}\right) = \frac{P(B/A)\ P(A)}{P(B)}$$

Spam filteration

22/02/23 Naive Bayes

$$P\left(\frac{A}{B}\right) = \frac{P(B/A)\,P(A)}{P(B)}$$

$P(A/B)$ → Possib

$P(B/A)$ → likelihood

$P(A)$ → prior

$P(A)$ → hypothesis

$P(B)$ → marginal

$P(B)$ → evidence

Probability of event A given that ~~event~~ prob of event B has already occure

$P\left(\frac{B}{A}\right)$ → Probability of event given that hypothesis is true

$P(A)$ → Probability of hypothesis ~~before~~ before considering evidence

$$P\left(\frac{A}{B}\right) = \frac{P(A \cap B)}{P(B)}$$

$$P\left(\frac{B}{A}\right) = \frac{P(A \cap B)}{P(A)}$$

apsara

Date: ______

$$P(A \cap B) = P(B \cap A)$$

$$P(A/B)\ P(B) = P(A \cap B)$$

$$P(B/A)\ P(A) = P(B \cap A)$$

$$\cancel{P(A \cap B)}\ P(A/B)\ P(B) = P(B/A)\ P(A)$$

$$P(A/B) = \frac{P(B/A)\ P(A)}{P(B)}$$

$$\boxed{P(A/B) = \frac{P(B/A)\ P(A)}{P(B)}}$$

Q1 Prob. that from deck of card, a card drawn is a King, given that it was a face card.

Sol

$$P\left(\frac{K}{FC}\right) = \frac{P\left(\frac{FC}{K}\right) * P(K)}{P(FC)}$$

$$= \frac{\frac{4}{4} * \frac{4}{52}}{\frac{12}{52}}$$

$$= \frac{1}{3}$$

$$= 0.33$$

Q1 For the given dataset apply naive bayes algorithm and predict the outcomes for car {Red, Domestic, SUV}

	color	~~type~~ type	Origin	Stolen
1	Red	Sport	Domestic	Yes
2	Red	Sport	Domestic	No
3	Red	Sport	Domestic	Yes
4	Yellow	Sport	Domestic	No
5	Yellow	Sport	Imported	Yes
6	Yellow	SUV	Imported	No
7	Yellow	SUV	Imported	~~No~~ Yes
8	Yellow	SUV	Domestic	No
9	Red	SUV	Imported	No
10	Red	Sport	Imported	~~No~~ Yes

$\vec{Car} = \{R, D, SUV\}$

Red color, domestic origin, SUV stolen or not.

Sol

$$P(Red/Yes) = \frac{P(Yes/Red) * P(Red)}{P(yes)}$$

$$= \frac{\frac{3}{5} * \frac{5}{10}}{\frac{5}{10}}$$

$$= \frac{3}{5}$$

$$P(Domestic/Yes) = \frac{P(Yes/Domestic) * P(D)}{P(Y)}$$

$$= \frac{\frac{2}{5} * \frac{5}{10}}{\frac{5}{10}}$$

$$= \frac{2}{5}$$

apsara

Date

$$P\left(SUV/Yes\right) = \frac{P\left(Y/SUV\right) * P(SUV)}{P(Yes)}$$

$$= \frac{\frac{1}{4} * \frac{4}{10}}{\frac{5}{10}}$$

$$= \frac{1}{5}$$

$$P\left(X/Yes\right) = \frac{3}{5} * \frac{2}{5} * \frac{1}{5}$$

$$= \frac{6}{125}$$

$$= 0.048$$

$$P\left(Red/No\right) = \frac{P(No/Red) * P(Red)}{P(No)}$$

$$= \frac{\frac{2}{5} * \frac{5}{10}}{\frac{5}{10}}$$

$$= \frac{2}{5}$$

$$P\left(\text{Domestic}/\text{NO}\right) = \frac{P(N/D) * P(D)}{P(N)}$$

$$= \frac{\frac{3}{5} * \frac{5}{10}}{\frac{5}{10}}$$

$$= \frac{3}{5}$$

$$P\left(\text{SUV}/\text{NO}\right) = \frac{P(N/S) * P(S)}{P(N)}$$

$$= \frac{\frac{3}{4} * \frac{4}{10}}{\frac{5}{10}}$$

$$= \frac{3}{5}$$

$$P\left(X/\text{NO}\right) = \frac{2}{5} * \frac{3}{5} * \frac{3}{5}$$

$$= \frac{18}{125}$$

$$= 0.144$$

apsara

Date:

Q2 Consider the given dataset, Apply naive baye algorithm and predict if a fruit has following properties then which type of fruit it is.

fruit = { yellow, sweet, long }

fruit	Yellow	Sweet	long	Total
Mango	350	450	0	650
Banana	400	300	350	400
Others	50	100	50	150
Total	800	850	400	1200

Sol ~~Ass~~ Assume fruit is mango

$P(X/mango)$

$$P(Yellow/Mango) = \frac{P(M/y) \cdot P(y)}{P(M)}$$

$$= \frac{\frac{350}{800} * \frac{800}{1200}}{\frac{650}{1200}}$$

$$P(S/M) = \frac{P(M/S) \times P(S)}{P(M)}$$

$$= \frac{350}{650} = 0.53$$

$$= \frac{\frac{450}{850} * \cancel{\frac{850}{1200}}}{\frac{650}{1200}}$$

$P(S/M) = 0.69$

$P(l/M) = 0$

$P(X/M) = P(Y/M) \cdot P(S/M) \cdot P(l/M)$

$= 0$

Assume fruit is Banana

$P(Y/B) = 1$

P

$\frac{45}{65}$

23/02/23

1. Machine learning life cycle (8M)
2. Diff b/w types of ML (6M)
3. Diff b/w batch & online learning (2M)
4. Scope and limitation of ML (2M)
5. Relationship b/w AI, ML, DL & DS (6M)
6. Diff b/w Linear vs Logistic (2M)
7. Numerical (6M)
8. Perceptron convergence theorem (6M)
9. Naive Bias problem (6M)
10. KNN problem (2M)
11. Linear and non linear SVM (6M)
12. ~~Diff~~ Naive Bias difference (6M)
13. Bias variance tradeoff (2)

23/02/23

Type of Naive Bias model

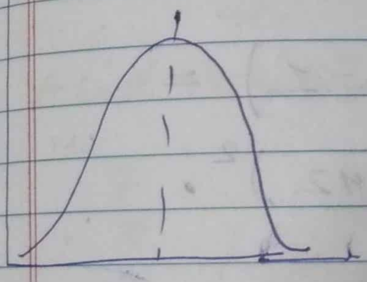
Gaussian

Variable is continous in nature

PDF =

$$\frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

$-\infty < e < \infty$



In Gaussian, Gaussian distribution is used,

Bernaulli

feature have binary value

$p(\text{success}) = p$

$P(\text{failure}) = q = 1-p$

Random variable x

$x = 1$ [success]

$x = 0$ [failure

In bernoulli, bernaulli distribution is used.

① $P(X=x) = p^{x}(1-p)^{1-x}$

② $P(x) = \begin{cases} p & \text{if } x=1 \\ q & \text{if } x=0 \end{cases}$

Multinomial

- Discrete count
- Multinomial distribution

$P(X_1 = x_1 \dots X_k = x_k)$

$$\frac{n!}{x_1! \cdots x_k!}\left[P_1^{x_1} \dots P_k^{x_k}\right]$$

Q: A random sample blood group of 6 Indians were taken. What is the prob. in the random sample one person has blood group O.

BG	O	A	B	AB
P	0.44	0.42	0.10	0.04

1 — O
2 — A
2 — B
1 — AB

Sol

$$P(X_1 = 1, X_2 = 2, X_3 = 2, X_4 = 1) =$$

$$\frac{6!}{1!\,2!\,2!\,1!}\left[(0.44)^1 \cdot (0.42)^2 \cdot (0.10)^2 \cdot (0.04)^1\right]$$

$$180 * [0.44 * 0.1764 * 0.01 * 0.04]$$

$$= 0.00558$$

(1) Explain working of SVM.

Linear SVM | Non linear SVM

(2) Bias variance tradeoff

(3) Perceptron

Q1 Explain diff steps of KNN algo.

Q. We have an objective testing with 2 attributes acid durability and strength to classify whether a special paper tissue is good or not

x_1 Acid Durability	x_2 length	Y
7	7	Bad
7	4	Bad
3	4	Good
1	4	Good

Now the factor produce a new paper tissue that pass laboratory test $x_1 = 3$, $x_2 = 7$, $k = 3$

Guess what type of classification is this tissue

Date: __.__.__

Q2 Estimate conditional probability of each attribute {color, legs, height, smelly} for specified class {M, H} using data given in table. Using the probability, estimate the probability value for new instance

{color = green, leg = 2, height = Tall, smelly = no

color	legs	height	smelly	species
white	3	Short	Yes	M
green	2	Tall	NO	M
green	3	Short	Yes	M
white	3	Short	Yes	M
green	2	Short	No	H
white	2	Tall	NO	H
white	2	Tall	NO	H
white	2	Short	Yes	H

color	M	H
white	$\frac{2}{5}$	$\frac{3}{5}$
green	$\frac{2}{3}$	$\frac{1}{3}$

$P(M) = \frac{1}{2}$

$P(H) = \frac{1}{2}$

legs	M	H
3	$\frac{3}{3}$	0
2	$\frac{1}{5}$	$\frac{4}{5}$

height	M	H
Short	$\frac{3}{5}$	$\frac{2}{5}$
Tall	$\frac{1}{3}$	$\frac{2}{3}$

smelly	M	H
Yes	$\frac{3}{4}$	$\frac{1}{4}$
NO	$\frac{1}{4}$	$\frac{3}{4}$

$P(W/M)$

$$P\left(G/M\right) = \frac{P(M/G) * P(G)}{P(M)}$$

$$= \frac{\frac{2}{3} * \frac{3}{8}}{\frac{4}{8}}$$

$$= \frac{1}{2}$$

$$P\left(L/M\right) = \frac{P(M/L) * P(L)}{P(M)}$$

$$= \frac{\frac{1}{5} * \frac{5}{8}}{\frac{4}{8}}$$

$$= \frac{1}{4}$$

$$P\left(T/M\right) = \frac{P(M/T) * P(T)}{P(M)}$$

$$= \frac{\frac{1}{3} * \frac{3}{8}}{\frac{4}{8}}$$

$$= \frac{1}{4}$$

$$P\left(\text{no s}/M\right) = \frac{P(M/\text{not s}) * P(\text{not s})}{P(M)}$$

$$= \frac{\frac{1}{4} * \frac{4}{8}}{\frac{4}{8}}$$

$$= \frac{1}{4}$$

$$P\left(X/M\right) = \frac{1}{2} * \frac{1}{4} * \frac{1}{4} * \frac{1}{4}$$

$$= \frac{1}{128}$$

$$P\left(G/M\right) = \frac{\frac{1}{3} * \frac{3}{8}}{\frac{4}{8}} = \frac{1}{4}$$

$$P\left(2l/M\right) = \frac{\frac{4}{5} * \frac{5}{8}}{\frac{4}{8}} = 1$$

$$P\left(T/M\right) = \frac{\frac{2}{3} * \frac{3}{8}}{\frac{4}{8}} = \frac{1}{2}$$

$$P(\text{not } s/H) = \frac{\frac{3}{4} * \frac{4}{8}}{\frac{4}{8}}$$

$$= \frac{3}{4}$$

$$P\left(\frac{x}{H}\right) = \frac{1}{4} * 1 * \frac{1}{2} * \frac{3}{4}$$

$$= \frac{3}{32}$$

	LV	HV
LB	ideal	overfitting
HB	Underfitting	

	LV	HV
LB	Ideal	Predic: Incon ACC
HB	Con InAc	Inc InAc

12/03/23

Classification model

Metrics

1. Classification model → Accuracy, Precision, Recall, AUC

2. Regression model → RMSE, R^2

3. Segmentation → IOU, Precision, Recall, fixed Accuracy. fixed Accuracy,

Divide data into parts.

Data Augmentation : Increasing data to improve accuracy,

Q1

	Predicted NO	Predicted NO
Actual NO	45	5
Actual Yes	5	95

TN	FN
FN	TP

Date:

consider confusion matrix given belove for a binary classifier predicting the presence of a disease.

The classifier made a total of 150 prediction. Out of these 150 cases, the classifier predicted Yes 100 times & No 50 times. Find out the accuracy, miss classification rate, True Positive Rate, False positive, False negative, True negative, precision, f1 score, prevalance.

CONFUSION MATRIX

Imp It is a table that is often used to describe performance of classification model.

Accuracy
How often a classifier is correct.

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

Missclassification
Overall of how often classifier is wrong.

Missclassification $= \frac{FN + FP}{TN + TP + FN + FP}$

True positive rate (Sensitivity / Recall)

when it is actually 'Yes', among how often does it predict 'Yes'.

$\frac{TP}{\text{Actual Yes}}$

$$\text{True Positive rate} = \frac{TP}{\text{Actual Yes}}$$

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

False positive rate

when it's actually 'NO', how often does it predict 'Yes'

$$\text{False positive} = \frac{FP}{\text{Actual No}}$$

$$= \frac{FP}{TN + FP}$$

True Negative rate

when it is actually no, how often does it predict no.

It is also known as specificity.

$$\text{TRUE Negative} = \frac{TN}{\text{Actual NO}}$$

Precision

what proportion of predicted positive is truly positive.

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

Prevelance

How often does the yes condition actually occur in our sample.

$$\text{Prevelance} = \frac{\text{Actual Yes}}{\text{Total Yes}}$$

~~F1 score~~

F1 score

It is function of precision and recall.

$$f_1 = \frac{2 * (\text{Precision} * \text{Recall})}{\text{Precision} + \text{Recall}}$$

Q

	Predicted No	Predicted Yes
Actual No	45 TN	5 FP
Actual Yes	5 FN	95 TP

① Accuracy $= \frac{TN + TP}{Total} = \frac{45+95}{150} = \frac{140}{150} =$

② Miss classification $= \frac{FN + FP}{Total} = \frac{10}{150}$

True positive $= \dfrac{TP}{\text{Actual Yes}}$

$= \dfrac{95}{100} = 0.95$

False Positive $= \dfrac{FP}{\text{Actual no}}$

$= \dfrac{5}{50}$

$=$

~~False -ve.~~

True -ve $= \dfrac{TN}{\text{Actual no}}$

$= \dfrac{45}{50}$

False +ve $= \dfrac{5}{100}$

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

$= \dfrac{95}{100}$

apsara

Date: ___

3/03/23

Baysian n/w

It is also known as decision n/w or relief n/w.

It is a probabilistic graphical model which represent a set of variable and their conditional dependencies using a directed acyclic graph.

It can be used for building model from datas and experts opinion.

Basically it consist of 2 part

1. Directed Acyclic graph

2. Table of conditional Probabilities.

It has mainly 2 components

1) Casual component

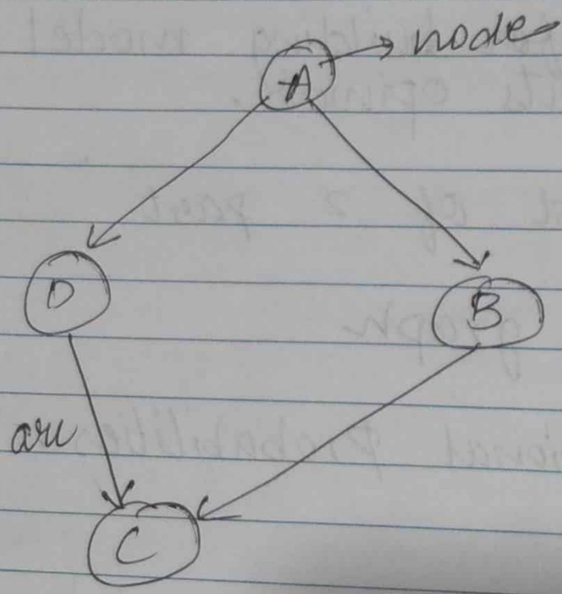
2) Actual no.

To propagate belief in ~~baseian~~ bayesian network initially DAG is converted into an undirected graph in which the arcs can be used to transmit probabilities in directed of evidence.

Use of Bayesian n/w

It is used for analyzing past, improving quality of decision.

Text analysis, fraud detection, cancer detection, image recognition.



DAG

Local Markov' property

It states that a node is conditionally independent of its non descendant given its parent.

$$P\left(D/A,B\right) = P\left(D/A\right)$$

b'coz D is independent of B.

$P(A/B)$

when A & B are dependent

$$\boxed{P(A/B) = \frac{P(A \text{ and } B)}{P(B)}}$$

when A & B are independent

$$\boxed{P\left(A/B\right) = P(A)}$$

$$P\left(A/B\right) = \frac{P(A)\cdot P(B)}{P(B)}$$

$$P\left(A/B\right) = P(A)$$

Joint probability

consider 3 variable a_1, a_2, a_3. By def^n the prob. of all diff possible combination of a_1, a_2, a_3 are called joint prob. distribution.

apsara

Date: ____

- If $[P(x_1, x_2 \ldots x_n)]$ is joint prob. from $a_1, a_2 \ldots a_n$ then joint prob. distribution is given by

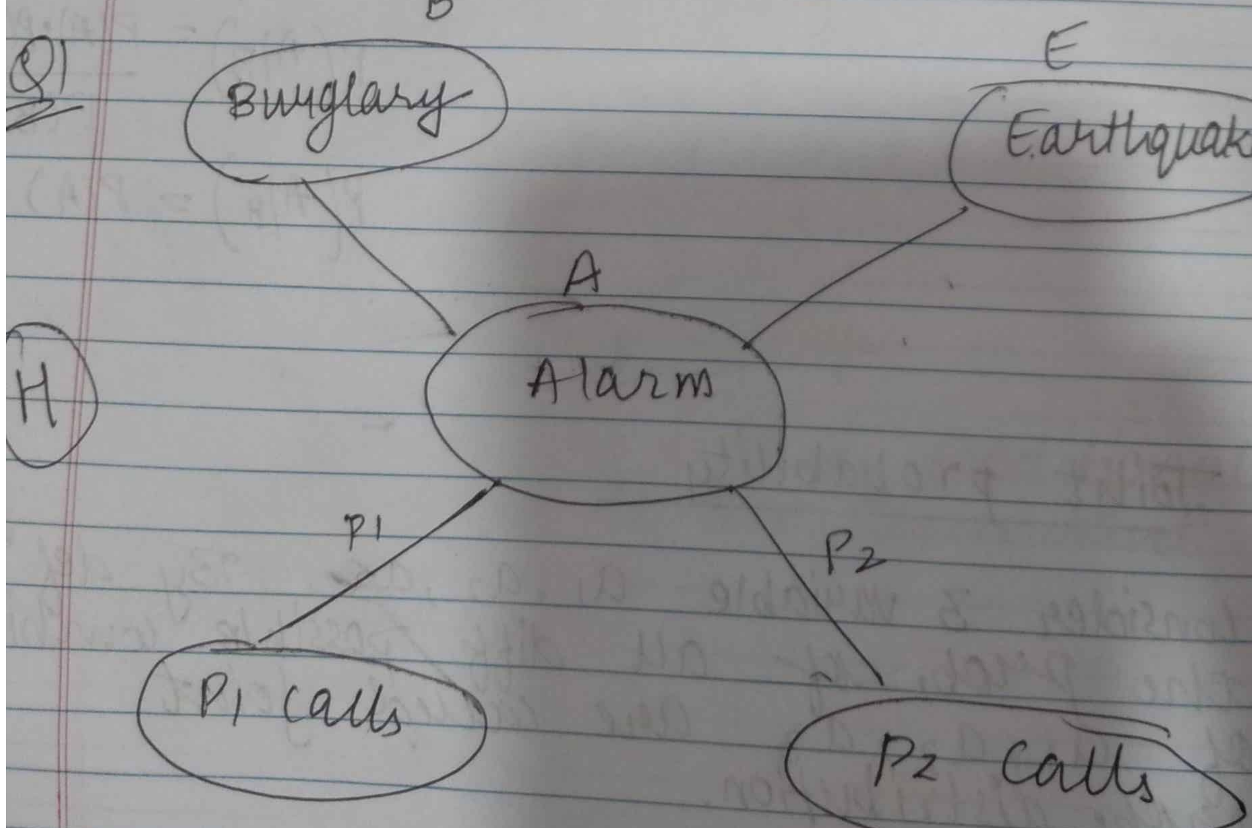
- $$P[a_1, a_2, \ldots a_n] = P[a_1 | a_2, a_3 \ldots a_n] * P[a_2, a_3 \ldots a_n]$$

$$= P[a_1 | a_2 \ldots a_n] * P[a_2 | a_3 \ldots a_n]$$

$$= P[X_i \mid X_{i-1}, \ldots, X_n]$$

$$= P[X_i \mid Parent(X_i)]$$

Q1



apsara

Date: ___

$P(B=T) = 0.001$

$P(B=F) = 0.999$

$P(E=T) = 0.002$

$P(E=F) = 0.998$

B	E	P(A=T)	P(A=F)
T	T	0.95	0.05
T	F	0.94	0.06
F	T	0.29	0.71
F	F	0.001	0.999

A	$P_1 = T$	$P_1 = F$
T	0.90	0.10
F	0.05	0.95

A	$P_2 = T$	$P_2 = F$
T	0.70	0.30
F	0.01	0.99

apsara

Date: ____

Find out probability

$P(P_1, P_2, A, \sim B, \sim E)$

Sol $P(P_1, P_2, A, \sim B, \sim E)$

$= P\left(P_1/A\right) \cdot P\left(P_2/A\right) \cdot \underline{P\left(A/\sim B, \sim E\right)}$

Since given in ques

$\cdot P(\sim B) \cdot P(\sim E)$

$= 0.90 * 0.70 * 0.001 * 0.999 *$
0.998

$=$

Naive
KNN

Perceptron
Bayesian n/w

14/03/22

Clustering

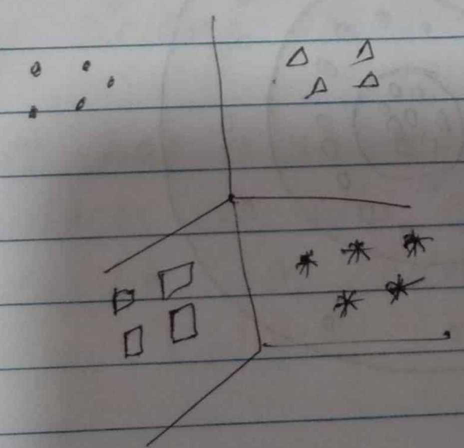
Clustering is a technique which prove the unlabelled dataset.

It is a way of grouping the data points into different clusters consisting of similar data points.

The objects with possible similarities remain in group that has less or no similarity with another group.

Ex: Amazon, Netflix uses it to show relcated ~~some~~ content.

① ~~Partioning~~ Partitioning



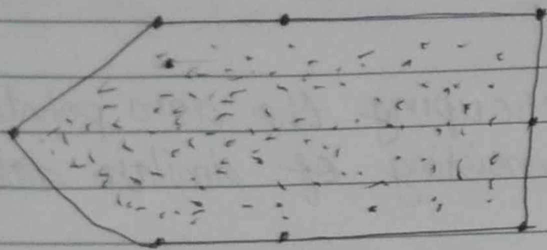
~~A~~lso called

centroid

apsara

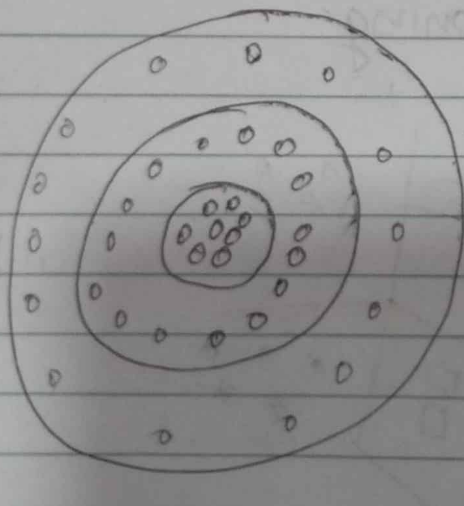
Date:

② Density based



It connects highly dense area into cluster and arbiterally shape distribition are formed as long as dense region can be connected.

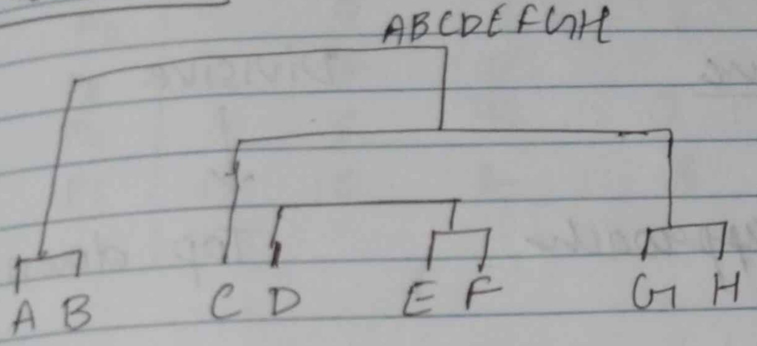
③ Distribution based



The data is divided based on probability of how a data set belongs to a particular distribution.

The grouping is done by assumming some distribution commonly known as gaussian distribution.

④ Hierarchial (Dendogram)



⑤ Fuzzy

It is a type of soft method in which data object may belong to more than one cluster.
Each data set has a set of membership coefficient which depends on degree of membership to be in cluster.

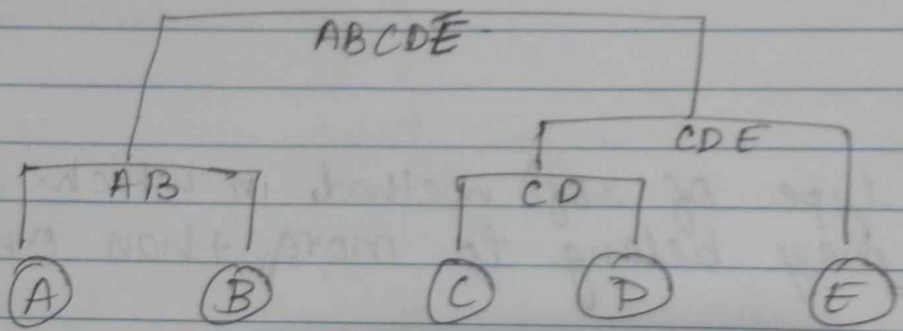
Hierarchical

1. Agglomarative
2. Divisive

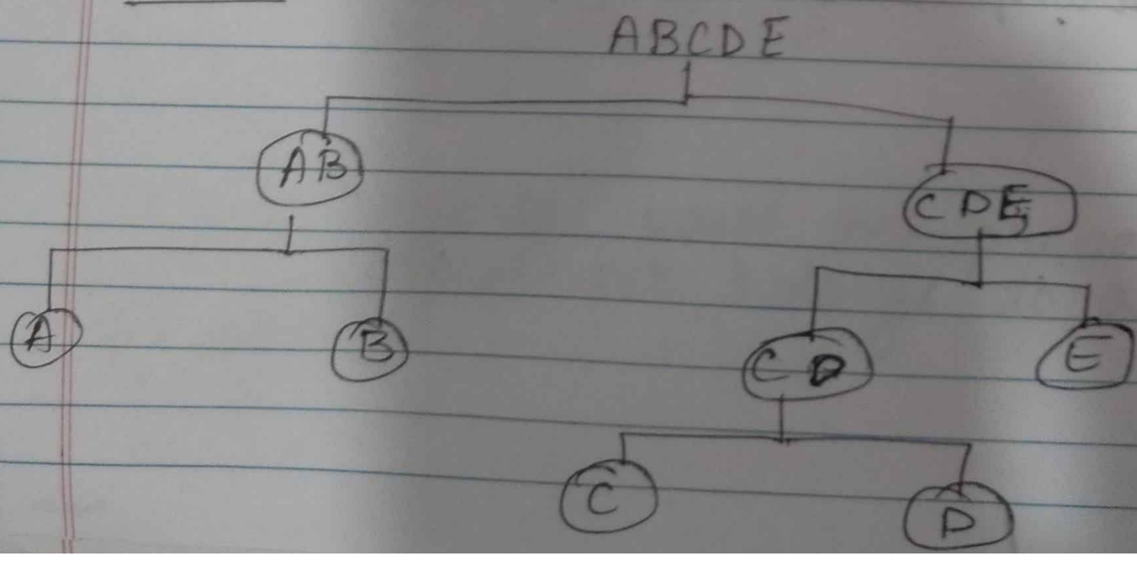
Agglomarative → Bottom up approach

Divisive → Top daonappro

Agglomarative



Divisive



single linkage → find minim dis.

apsara

Date:____

① Agglomarative single linkage

Q1

	P1	P2	P3	P4	P5
P1	0				
P2	9	0			
P3	3	7	0		
P4	6	5	9	0	
P5	11	10	2	8	0

Solve using Agglomarative single linkage.

Sol

Step 1 : Excluding 0 select the minm distance.

It is 2 ; (P5 - P3)
So P3 P5 will form cluster

	P1	P2	[P3 P5]	P4
P1	0			
P2	9	0		
[P3 P5]	$\boxed{3}$	$\boxed{7}$	0	
P4	6	5	$\boxed{8}$	0

$d(P_1, [P_3, P_5])$

$\min(d(P_1, P_3), P(P_1, P_5)]$

$\min(3, 11)$

3

$d[P_2, [P_3, P_5]]$

$= \min(7, 10) = 7$

$d(P_4, [P_3, P_5]) = \min(9, 8) = 8$

above table, min is 3,

It is P_1 & P_3 P_5

so combine P_1 P_3 P_5

		$P_1 P_3 P_5$	P_2	P_4
	$P_1 P_3 P_5$	0		
9, 7, 10	P_2	$\boxed{7}$	0	
6, 9, 8	P_4	$\boxed{6}$	5	0

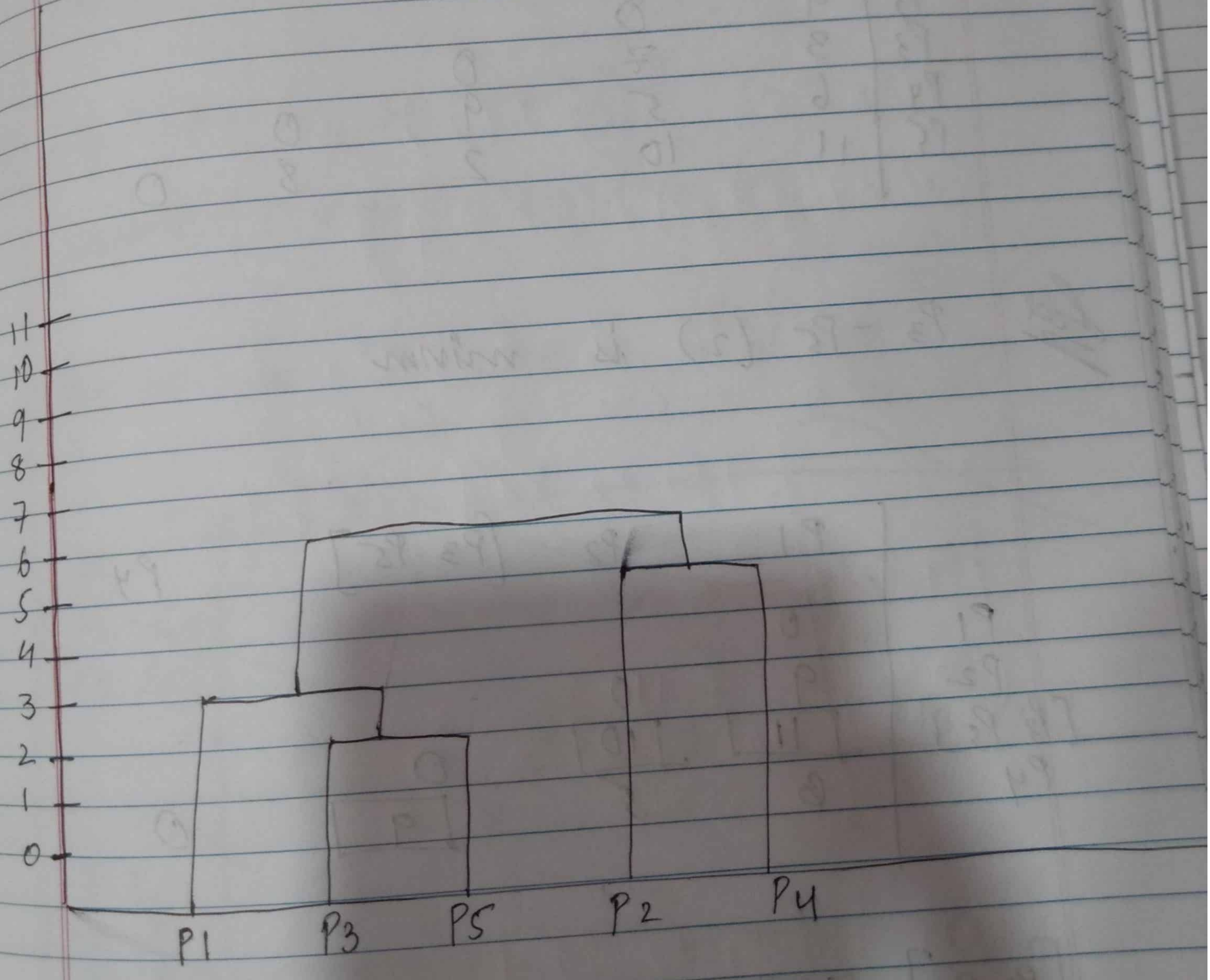
min is 5 ; combine P_2 P_4

$P_2 \rightarrow$ 9, 7, 10 $\rightarrow$ 7

$P_4 \rightarrow 6, 9, 8 \rightarrow 6$

Date:

	$P_1 P_3 P_5$	P_2 P_4
$P_1 P_3 P_5$	0	
$P_2 P_4$	6	0



16/03/23

Agglomarative Linkage :

Complete ~~[illegible]~~ Linkage

	P_1	P_2	P_3	P_4	P_5
P_1	0				
P_2	9	0			
P_3	3	7	0		
P_4	6	5	9	0	
P_5	11	10	2	8	0

Sol: $P_3 - P_5$ (2) is minm

	P_1	P_2	$[P_3\ P_5]$	P_4
P_1	0			
P_2	9	0		
$[P_3\ P_5]$	$\boxed{11}$	$\boxed{10}$	0	
P_4	6	5	$\boxed{9}$	0

$[P_3\ P_5]$ $[P_1]$

1. max $d(P_1\ (P_3\ P_5))$

$max(3, 11) = 11$

2. $max(d(P_2\ (P_3\ P_5)))$

$max(7, 10) = 10$

3. $d(P_4\ (P_3\ P_5))$

$max(9, 4) = 9$

min is 5 i.e $P_2\ P_4$

	~~P_1~~
~~P_1~~	
~~$[P_2\ P_4]$~~	
$[P_3\ P_5]$	

	P_1	$[P_2\ P_4]$	$[P_3\ P_5]$
P_1	0		
$[P_2\ P_4]$	$\boxed{9}$	0	
$[P_3\ P_5]$	$\boxed{11}$	$\boxed{10}$	0