21/3/25

Q.

Observation	Actual label	Predicted score	Predicted label
1	1	0.85	1
2	0	0.60	1 * (FP)
3	1	0.70	1
4	1	0.40	0 * (FN)
5	0	0.55	1 * (FP)
6	1	0.50	1
7	0	0.65	1 * (FP)
8	0	0.35	0
9	1	0.60	1
10	0	0.20	0

Threshold = 0.5

Find accuracy, precision, recall / sensitivity true +ve rate, specificity, false +ve rate (1 − specificity), F1-score, ROC plot, AUC.

Ans :- confusion matrix (For t = 0.5)

	Original +ve	Original −ve
Predicted +ve	4	3
Predicted −ve	1	2

1) Accuracy $= \frac{TP + TN}{\text{total}}$

$= \frac{4+2}{10}$

$= \frac{6}{10} = 0.6$

2) Precision $= \frac{TP}{TP + FP}$

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

$= \frac{4}{7} = 0.57$

3) true Recall / sensitivity

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

$= \frac{4}{4+1} = \frac{4}{5} = 0.8$

4) Specificity (-ve recall) $= \frac{TN}{TN+FP}$

$= \frac{2}{2+3}$

$= \frac{2}{5} = 0.4$

5) false +ve rate = 1 − specificity

$= 1 - 0.4 = 0.6$

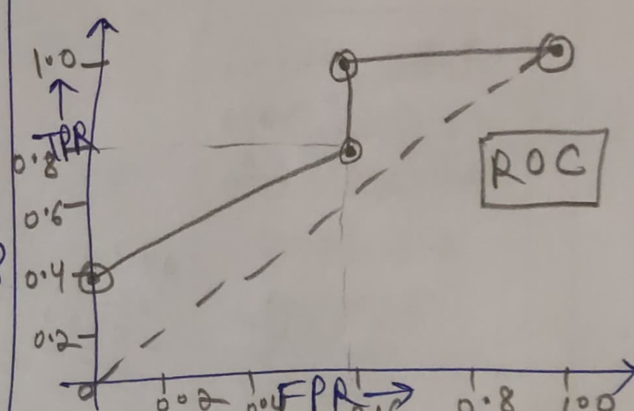
6) F1-score $= \frac{1}{\frac{\frac{1}{\text{prec}} + \frac{1}{\text{recall}}}{2}}$

F1 score $= \frac{1}{\frac{\frac{1}{0.57} + \frac{1}{0.8}}{2}}$

$F1 = \frac{1}{\frac{3}{2}}$

$F1 = \frac{1}{3} \times 2 = 0.66$

7) Given:- thresholds are − 0.7, 0.5, 0.4, 0.2



Obs.	Actual	Score	Predicted ($t = 0.7$)
1	1	0.85	1 (score > t)
2	0	0.60	0
3	1	0.70	1
4	1	0.40	0
5	0	0.55	0
6	1	0.50	0
7	0	0.65	0
8	0	0.35	0
9	1	0.60	0
10	0	0.20	0

For 't' = 0.7

T.P.	F.P.
~~2~~ (TP)	0 (FP)
3 (FN)	5 (TN)

TPR = ~~1 (true recall)~~ sensitivity

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

$= \frac{2}{2+3} = \frac{2}{5} = 0.4$

= ~~0.7~~

FPR = $1 - \frac{TN}{TN+FP}$ → specificity

$= 1 - \frac{5}{5+0} = \frac{8}{8} = 0$

Obs.	Actual score		Predicted (t=0.4)
1	0.85	1	1
2	0.60	0	1
3	0.70	1	1
4	0.40	1	1
5	0.55	0	1
6	0.50	1	1
7	0.65	0	1
8	0.35	0	0
9	0.60	1	1
10	0.20	0	0

For 't' = 0.4

TP (7)	FP (3)
FN (0)	TN (2)

TPR = $\frac{7}{7+0} = 1$

FPR = $1 - \frac{2}{2+3}$

$1 - \frac{2}{5}$

FPR = 0.6

~~Obs~~ Actual labels	Actual score	Predicted (t=0.2)
1	0.85	1
0	0.60	1
1	0.70	1
1	0.40	1
0	0.55	1
1	0.50	1
0	0.65	1
0	0.35	1
1	0.60	1
0	0.20	1

TP (5)	FP (5)
FN (0)	TN (0)

For 't' = 0.2

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

$FPR = 1 - \frac{0}{0+5} = 1 - 0 = 1$

8) $$\boxed{AUC = \sum_{i=1}^{4} \frac{(FPR_i - FPR_{i-1}) * (TPR_i + TPR_{i-1})}{2}}$$

AUC will be b/w 0.5 & 1

For $i = 1$, → i.e., 0.7 (t)

$$AUC = \frac{(FPR_1 - FPR_0) * (TPR_1 + TPR_0)}{2}$$

$AUC = 0$

For $i = 2$, → i.e., $t = 0.5$

$$AUC = \frac{(FPR_2 - FPR_1) * (TPR_2 + TPR_1)}{2}$$

$$AUC = \frac{(0.6 - 0) * (0.8 + 0.4)}{2}$$

$$AUC = \frac{0.6 \times 1.2}{2} = 0.36$$

For $i = 3$ → i.e., $t = 0.4$

$$AUC = \frac{(FPR_3 - FPR_2) * (TPR_3 + TPR_2)}{2}$$

$$AUC = \frac{(0.6 - 0.6) * (1 + 0.8)}{2} = 0$$

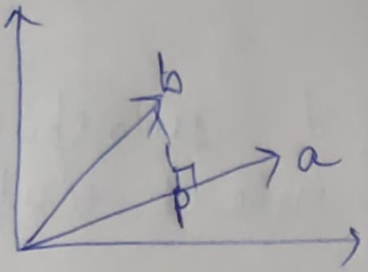
For $i = 4$; $t = 0.2$

$$AUC = \frac{(FPR_4 - FPR_3) * (TPR_4 + TPR_3)}{2}$$

$$AUC = \frac{(1 - 0.6) * (1 + 1)}{2} = \frac{0.8}{2} = 0.4$$

$\therefore$ $\boxed{AUC = 0.36 + 0.4 = 0.76}$

Q1



$$b = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

$$a = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Find projected point 'p' = [ ] ?

Ans :- $p = \dfrac{a a^T b}{a^T a}$

$$a^T = \begin{bmatrix} 1 & 2 \end{bmatrix}$$

$$p = \frac{\begin{bmatrix} 1 \\ 2 \end{bmatrix}_{2\times1} \begin{bmatrix} 1 & 2 \end{bmatrix}_{1\times2}}{\underbrace{\begin{bmatrix} 1 & 2 \end{bmatrix}_{1\times2} \begin{bmatrix} 1 \\ 2 \end{bmatrix}_{2\times1}}} \times \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

$$p = \frac{\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}_{\cancel{2}\times\cancel{2}} \begin{bmatrix} 3 \\ 4 \end{bmatrix}_{\cancel{2}\times1}}{\begin{bmatrix} 1 & 2 \end{bmatrix}_{1\times\cancel{2}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}_{\cancel{2}\times1}} = \frac{\begin{bmatrix} 11 \\ 22 \end{bmatrix}}{1+2\times2 = 5} = \begin{pmatrix} 11/5 \\ 22/5 \end{pmatrix} = \begin{pmatrix} 2.2 \\ 4.4 \end{pmatrix}$$

1×1 = Scalar

Q1 $X = \begin{bmatrix} 2 & 100 \\ 4 & 200 \\ 6 & 300 \end{bmatrix}_{3\times 2} \xrightarrow{\text{apply PCA}} Z = \begin{bmatrix} z_1? \\ z_2? \\ z_3? \end{bmatrix}_{3\times 1}$

(d=2) (d=1)

$\mathbb{R}^2 \longrightarrow \mathbb{R}^1$

Ans :- ① Standardise each column — (X_{std})

$z = \frac{x - \bar{x}}{\sigma}$ or $\frac{x - \mu}{\sigma}$

- μ_1 (mean of 1st column) $= \frac{2+4+6}{3} = 4$

  μ_2 (mean of 2nd column) $= \frac{100+200+300}{3} = 200$

- $\sigma^2 = \frac{1}{n-1} \sum (x_i - \mu)^2$

Feat Column-1 :- $\frac{1}{3-1} \left[(2-4)^2 + (\cancel{4-4})^2 + (6-4)^2 \right]$

$\sigma^2 :- \frac{1}{2} \times (4+4) = 4$

$\boxed{\sigma_1 = \sqrt{4} = 2}$

Feat Column-2 :- $\frac{1}{2} \left[(100-200)^2 + (\cancel{200-200})^2 + (300-200)^2 \right]$

$\sigma^2 = \frac{1}{2} \times (10,000 + 10,000) = 10,000$

$\boxed{\sigma_2 = 100}$

Feat Column-1 :- $z_{11} = (2-4)/2 = -1$

$z_{12} = (4-4)/2 = 0$

$z_{13} = (6-4)/2 = 1$

Feat Column-2 :- $z_{21} = (100-200)/100 = -1$

$z_{22} = (200-200)/100 = 0$

$z_{23} = (300-200)/100 = 1$

$$X_{std.} = \begin{bmatrix} -1 & -1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix} \Rightarrow \text{After step-1 (standardization)}$$

② $X^T X$ covariance matrix —

$$\frac{1}{n-1}\left(X^T_{std} \cdot X_{std}\right) \Rightarrow \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}_{2\times3} \begin{bmatrix} -1 & -1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix}_{3\times2}$$

$$\Rightarrow \begin{bmatrix} (-1\times-1+0\times0+1\times1) & (1+0+1) \\ (1+0+1) & (1+0+1) \end{bmatrix}_{2\times2}$$

$$(n \Rightarrow \text{samples}) \Rightarrow \frac{1}{n-1}\begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}_{2\times2}$$

$$\Rightarrow \frac{1}{2}\begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}_{2\times2}$$

$$\Rightarrow \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

③ Compute the eigen values & eigen vectors of the covariance matrix — $(\lambda_1, V_1), (\lambda_2, V_2)$ — (λ_d, V_d)

- determinant of $(A - \lambda I) = 0$

$$A - \lambda I \Rightarrow \underbrace{\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}}_{(A)} - \underbrace{\begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}}_{(\lambda I)}$$

$$\Rightarrow \begin{bmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{bmatrix} \rightarrow ①$$

$$\Rightarrow \det\begin{bmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{bmatrix} = [(1-\lambda)(1-\lambda)] - [(1\times1)] = (1-\lambda)^2 - 1$$

$$\Rightarrow \cancel{1} + \lambda^2 - 2\lambda - \cancel{1} = 0$$

$$\Rightarrow \lambda(\lambda - 2) = 0$$

$\Rightarrow \boxed{\lambda_1 = 0 \ , \lambda_2 = 2} \Rightarrow$ Eigen values

- Find eigen vectors

vector $\begin{bmatrix} x \\ y \end{bmatrix}$

$(A - \lambda I)\vec{V} = 0$

For $\lambda_2 = 2$

$(A - 2I) \Rightarrow \begin{bmatrix} 1-2 & 1 \\ 1 & 1-2 \end{bmatrix}$

$\Rightarrow \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$

$(A - 2I)V_2 \Rightarrow \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$-1x + y = 0$ or $(1 \times x) - y = 0$

$\boxed{x = y}$ $\boxed{x = y}$

square of lengths of components

$\sqrt{(1)^2 + (1)^2}$

$\boxed{V_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}}$

Normalization is done column wise

For $\lambda_1 = 0$

$(A - 0I) \Rightarrow \begin{bmatrix} 1-0 & 1 \\ 1 & 1-0 \end{bmatrix}$

$\Rightarrow \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

$(A - 0I)V_1 \Rightarrow \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$x + y = 0$ or $x + y = 0$

$\boxed{x = -y}$ $\boxed{x = -y}$

$\boxed{V_1 = \begin{bmatrix} -1 \\ +1 \end{bmatrix}}$

4. ~~divide~~ decide the needed % of variance explained & choose 'K' principal components such that (8) variance explained $\geq 95\%$

5. Transform 'X' into 'Z' (into new space of principal components)

6. Use this 'Z_{pca}' to train the ML model.

$Z_{pca} = X_{std} \times V_2$

bec. its $\lambda_2 > \lambda_1$ → Bec ↑λ ↑ variance

$Z_{pca} = \begin{bmatrix} -1 & -1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \Rightarrow \begin{bmatrix} -1 & -1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ \cancel{1/\sqrt{2}} \end{bmatrix}$

28/3/25

$$x_{pca} = \begin{bmatrix} \left(-1 \times \frac{1}{\sqrt{2}}\right) + \left(-1 \times \frac{1}{\sqrt{2}}\right) \\ (0 \times 1/\sqrt{2}) + (0 \times 1/\sqrt{2}) \\ (1 \times 1/\sqrt{2}) + (1 \times 1/\sqrt{2}) \end{bmatrix}$$

$$\boxed{x_{pca} = \begin{bmatrix} \frac{-2}{\sqrt{2}} \\ 0 \\ \frac{2}{\sqrt{2}} \end{bmatrix}_{3 \times 1}}$$

$\Rightarrow$ Transformed data points

Q.1

	x_1	x_2
S-1	1	1
S-2	1	2
S-3	2	2
S-4	8	8
S-5	8.	9
S-6	9	8

K = 2 (no. of clusters)

C_1 (cluster1) $= \{2, 3, 4, 6\}$

$C_2 = \{1, 5\}$

After first iteration —

$C_1 = \{4, 5, 6\}$

$C_2 = \{1, 2, 3\}$

$C_1 = \{2, 3, 4, 6\}$

① Centroid-1 $\begin{bmatrix} 5 \\ 5 \end{bmatrix}$ $\Rightarrow \dfrac{1+2+8+9}{4} = 5 \; (x_1)$ | $\dfrac{2+2+8+8}{4} = \dfrac{20}{4} = 5 \; (x_2)$

Centroid-2 $\begin{bmatrix} 4.5 \\ 5 \end{bmatrix}$ $\Rightarrow \dfrac{1+8}{2} = \dfrac{9}{2} = 4.5 \; (x_1)$ | $\dfrac{1+9}{2} = \dfrac{10}{2} = 5 \; (x_2)$

② Euclidean distance —

- $d_{\substack{S-1 \\ Cent1}} = \begin{bmatrix} 5-1 \\ 5-1 \end{bmatrix}$ (Cent1 − x_1 (S-1), Cent1 − x_2 (S-1)) $= \begin{bmatrix} (4)^2 \\ (4)^2 \end{bmatrix} = \sqrt{\begin{bmatrix} 16 + 16 \end{bmatrix}} = \sqrt{32} = 5.65$

$d_{\substack{S-1 \\ cent2}} = \begin{bmatrix} 4.5-1 \\ 5-1 \end{bmatrix} = \begin{bmatrix} (3.5)^2 \\ (4)^2 \end{bmatrix} = \begin{bmatrix} 12.25 \\ + \\ 16.00 \end{bmatrix} = \sqrt{18.25} = 4.27$

Since, $4.27 < 5.65$, so put sample-1 in cluster-2 (C_2).

- $d_{\substack{S-2 \\ cent2}} = \begin{bmatrix} 5-1 \\ 5-2 \end{bmatrix} = \begin{bmatrix} (4)^2 \\ (3)^2 \end{bmatrix} = \begin{bmatrix} 16 + 9 \end{bmatrix} = \sqrt{25} = 5$

$d_{\substack{S-2 \\ cent2}} = \begin{bmatrix} 4.5-1 \\ 5-2 \end{bmatrix} = \begin{bmatrix} (3.5)^2 \\ (3.0)^2 \end{bmatrix} = \begin{bmatrix} 12.25 \\ + \\ \cancel{6.25} \; 9.0 \end{bmatrix} = \sqrt{21.25} = 4.6$

$4.6 < 5$, so sample-2 goes in cluster-2.

- $d_{\substack{s-3\\ cent1}} = \begin{bmatrix} 5-2 \\ 5-2 \end{bmatrix} = \begin{bmatrix} (3)^2 \\ (3)^2 \end{bmatrix} = \begin{bmatrix} 9 \\ + \\ 9 \end{bmatrix} = \sqrt{18} = 4.24$

  $d_{\substack{s-3\\ cent2}} = \begin{bmatrix} 4.5-2 \\ 5.0-2 \end{bmatrix} = \begin{bmatrix} (2.5)^2 \\ \cancel{(2.5)^2} \\ (3)^2 \end{bmatrix} = \begin{bmatrix} 6.25 \\ + \\ \cancel{6.25} \\ 9.0 \end{bmatrix} = \sqrt{15.25} = 3.90$

  since, $3.53 < 4.24$, so sample-3 will go in cluster-2.

- $d_{\substack{s-4\\ cent1}} = \begin{bmatrix} 5-8 \\ 5-8 \end{bmatrix} = \begin{bmatrix} (-3)^2 \\ (-3)^2 \end{bmatrix} = \begin{bmatrix} 9 \\ + \\ 9 \end{bmatrix} = \sqrt{18} = 4.24$

  $d_{\substack{s-4\\ cent2}} = \begin{bmatrix} 4.5-8 \\ 5.0-8 \end{bmatrix} = \begin{bmatrix} (-3.5)^2 \\ (-3.0)^2 \end{bmatrix} = \begin{bmatrix} 12.25 \\ + \\ \cancel{12.25} \\ 9.0 \end{bmatrix} = \sqrt{21.25} = 4.60$

  since $4.24 < 4.60$, so sample-4 goes to cluster-1.

- $d_{\substack{s-5\\ cent1}} = \begin{bmatrix} 5-8 \\ 5-9 \end{bmatrix} = \begin{bmatrix} (-3)^2 \\ (-4)^2 \end{bmatrix} = \begin{bmatrix} 9 \\ + \\ 16 \end{bmatrix} = \sqrt{25} = 5$

  $d_{\substack{s-5\\ cent2}} = \begin{bmatrix} 4.5-8 \\ 5-9 \end{bmatrix} = \begin{bmatrix} (-3.5)^2 \\ (-4)^2 \end{bmatrix} = \begin{bmatrix} 12.25 \\ + \\ \cancel{20.25} \\ 16.0 \end{bmatrix} = \sqrt{28.25} = 5.31$

  since, $5 < 5.31$, so sample-5 will go in cluster-1.

- $d_{\substack{s-6\\ cent1}} = \begin{bmatrix} 5-9 \\ 5-8 \end{bmatrix} = \begin{bmatrix} (-4)^2 \\ (-3)^2 \end{bmatrix} = \begin{bmatrix} 16 \\ + \\ 9 \end{bmatrix} = \sqrt{25} = 5$

  $d_{\substack{s-6\\ cent2}} = \begin{bmatrix} 4.5-9 \\ 5-8 \end{bmatrix} = \begin{bmatrix} 20.25 \\ + \\ \cancel{12.25} \\ 9.0 \end{bmatrix} = \sqrt{29.25} = 5.4$

  Sample-6 goes in cluster-1 (bec; $5 < 5.4$).

- Compute new centroids — for next iteration till convergence.

* Example :-

	x_1	x_2
A	1	1
B	2	1
C	4	3
D	5	4
E	6	5

→ Every point is in $\mathbb{R}^2$ (2-D) space.

→ Single linkage (Euclidean distance)

Ans:- Step-1 :- compute the distance matrix (Pairwise distance – $^nC_2$)

Distance value (score) ↓ y-axis = Min. value in matrix = 1

	A	B	C	D	E
A	0	Min. 1	3.6	5	6.4
B	Min. 1	0	2.8	4.2	5.6
C	3.6	2.8	0	1.41	2.8
D	5	4.2	1.41	0	1.41
E	6.4	5.6	2.8	1.41	0

- $A = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, B = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \Rightarrow d_{AB} = \begin{bmatrix} (1-2)^2 \\ + \\ (1-1)^2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

  $\Rightarrow d_{AB} = \sqrt{1+0} = \sqrt{1} = 1$

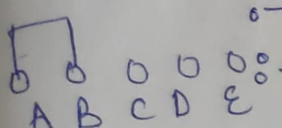
- $d_{AC} = \begin{bmatrix} (1-4)^2 \\ + \\ (1-3)^2 \end{bmatrix} = \begin{bmatrix} 9 \\ + \\ 4 \end{bmatrix} = \sqrt{13} = 3.6$

- $d_{AD} = \begin{bmatrix} (1-5)^2 \\ + \\ (1-4)^2 \end{bmatrix} = \begin{bmatrix} 16 \\ + \\ 9 \end{bmatrix} = \sqrt{25} = 5$

- $d_{AE} = \begin{bmatrix} 25 \\ + \\ 16 \end{bmatrix} = \sqrt{41} = 6.4$
- $d_{BA} = d_{AB}$, $d_{CA} = d_{AC}$, $d_{DA} = d_{AD}$, $d_{EA} = d_{AE}$.
- $d_{BC} = \begin{bmatrix} (2-4)^2 \\ + \\ (1-3)^2 \end{bmatrix} = \sqrt{8} = 2.8$
- $d_{BD} = \begin{bmatrix} (3)^2 \\ + \\ (3)^2 \end{bmatrix} = \sqrt{18} = 4.2$
- $d_{BE} = \begin{bmatrix} (4)^2 \\ + \\ (4)^2 \end{bmatrix} = \sqrt{32} = 5.6$
- $d_{CD} = \begin{bmatrix} (4-5)^2 \\ + \\ (3-4)^2 \end{bmatrix} = \sqrt{2} = 1.41$
- $d_{CE} = \begin{bmatrix} (4-6)^2 \\ + \\ (3-5)^2 \end{bmatrix} = \sqrt{8} = 2.8$
- $d_{DE} = \begin{bmatrix} (5-6)^2 \\ + \\ (4-5)^2 \end{bmatrix} = \sqrt{2} = 1.41$
- $d_{CB} = d_{BC}$, $d_{DB} = d_{BD}$, $d_{EB} = d_{BE}$
- $d_{DC} = d_{CD}$, $d_{EC} = d_{CE}$
- $d_{ED} = d_{DE}$

Step-2 :- Merge A & B since their pairwise distance is min., i.e., 1

:- New clusters are – (AB) C D E

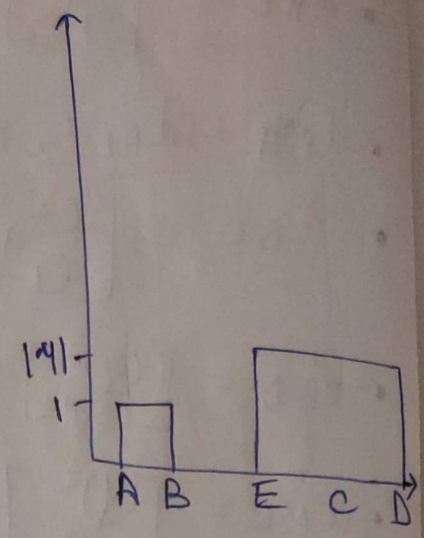


:- New distance matrix –

	(AB)	C	D	E
(AB)	0	2.82	4.2	5.6
C	2.82	0	1.41	2.8
D	4.2	1.41	0	1.41
E	5.6	2.8	1.41	0

Min. (New score)

$$d_{(AB)C} = \min(d_{CA}, d_{CB})$$
$$= \min(3.6, 2.8)$$
$$= 2.8$$



$$d_{(AB)D} = \min(d_{DA}, d_{DB})$$
$$= \min(5, 4.2)$$
$$= 4.2$$

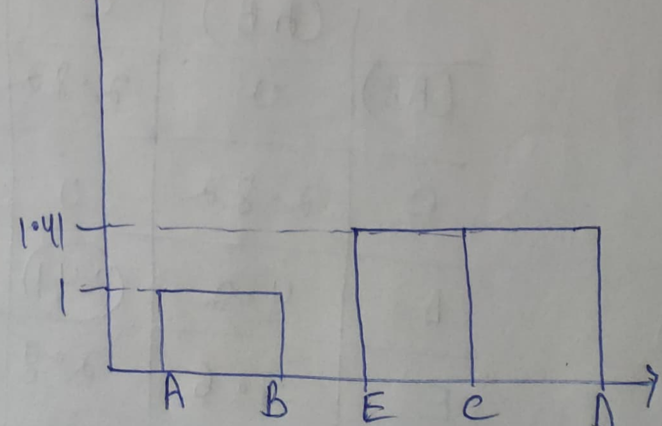
$$d_{(AB)E} = \min(6.4, 5.6) = 5.6$$

Step-3 :-

	AB	C	DE
AB	0	2.82	4.2
C	2.82	0	Min. 1.41
DE	4.2	1.41	0

$$d_{(AB)(DE)} = \min(d_{DA}, d_{EA}, d_{DB}, d_{EB})$$
$$= \min(5, 6.4, 4.2, 5.6)$$
$$= 4.2$$

$$d_{(DE)C} = \min(d_{CD}, d_{CE})$$
$$= \min(1.41, 2.8)$$
$$= 1.41$$



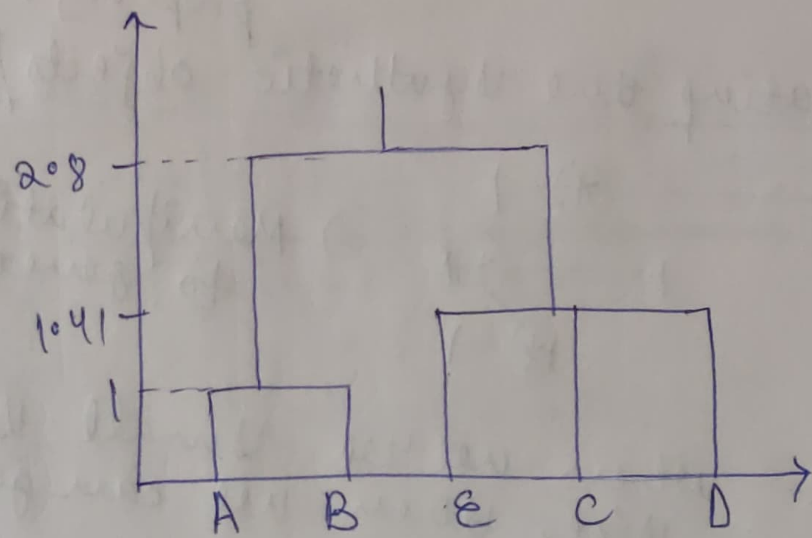
Step-4 :-

	AB	CDE
AB	0	2.8
CDE	2.8	0

$$d_{(AB)(CDE)} = \min\left(d_{CA}, d_{CB}, d_{DA}, d_{DB}, d_{EA}, d_{EB}\right)$$

$$d_{(AB)(CDE)} = \min\left(3.6, 2.8, 5, 4.2, 6.4, 5.6\right)$$

$$d_{(AB)(CDE)} = 2.8$$



e.g.:-

X	Y
1) Free win now	Spam
2) win a prize	Spam
3) Hello how are you	Not spam
4) Let's win it	Not spam
5) Free lunch today	Not spam

Test message :- "Free win" belongs to which class?

	Feature-1 (x_1) Word "free"	x_2 Word "win"	Label
Msg-1	Yes (true)	Yes (true)	Spam
2	No (free-not)	Yes	Spam
3	No	No	Not spam
4	No	Yes	Not spam
5	Yes	No	NA Spam

<u>Step-1</u> :- calculate prior probability —

$n=5$

$$P(y=1) = \frac{2}{5} = 0.4$$

$$P(y=0) = \frac{3}{5} = 0.6$$

<u>Step-2</u> :- compute the likelihood —

For 'spam' class :- "Free" appears 1 of 2 spam class

Free: yes, No → $1/2$ ⇐ (deno. ⇓ 2 ⇒ spam mo.)

Win: yes, yes → $2/2$ (spam, spam)

$$P(\text{Free} = \text{yes} \mid y=1) = \frac{1}{2} = 0.5$$

$$P(\text{Win} = \text{yes} \mid y=1) = \frac{2}{2} = 1$$

For 'not spam' class :-

$$P(\text{Free} = \text{yes} \mid y=0) = 1/3 = 0.33$$

$$P(\text{Win} = \text{yes} \mid y=0) = 1/3 = 0.33$$

<u>Step-3</u> :- Inference / Test phase (Bayes rule) —

Given test message :- Free = yes
win = yes

$$P(spam|X) = P(X|spam) \cdot P(spam)$$

↳ Baye's rule

$$P(X|spam) = P(X_1, X_2 | spam) = P(X_1|spam) \cdot P(X_2|spam)$$

Test msg. ⇒ $P(spam^{(y=1)}|X) = P(X_1|spam) \cdot P(X_2|spam) \cdot P(spam)$

$$= \frac{1}{\cancel{2}} \times \frac{\cancel{2}}{\cancel{2}} \times \frac{\cancel{2}}{5}$$

$$= \frac{1}{5} = 0.2$$

$P(Not\ spam|X) = P(X_1|Not\ spam) \cdot P(X_2|Not\ spam) \cdot P(Not\ spam)$

$$P(Not\ spam|X) = \frac{1}{3} \times \frac{1}{\cancel{3}} \times \frac{\cancel{3}}{5} = \frac{1}{15} = 0.066$$

Since, 0.2 > 0.066, so test message belongs to "spam" class.

Q. 1) Accuracy of test = 99% = 0.99
2) Tested +ve
3) Rare disease (1 in 10,000)

Ans:- Acc. to Bayes' theorem —

$$P(D=1 \mid T=1) = \frac{P(D) \times P(T \mid D)}{P(D) \cdot P(T \mid D) + P(D') \cdot P(T \mid D')}$$

or

$P(D \mid T)$

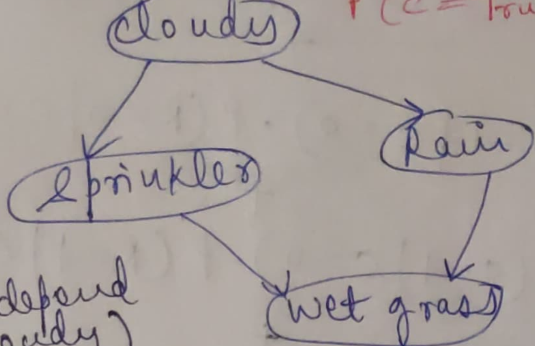
$$P(D|T) = \frac{0.99 \times 0.0001}{0.99 \times 0.0001 + 0.999 \times 0.01}$$

$$P(D|T) = \frac{0.000099}{0.010089} = 0.0098$$

e.g.:-

$P(C = False) = 0.5$

$P(C = True) = 0.5$



(wet grass doesn't depend on cloudy)

[X_i ⫫ ancestors; given parents(X_i)]
independent of

Note:- Time series are not taken into account in this.

:- Every variable rep. as nodes

$P(R|C) \Rightarrow$

Cloudy (C)	P(R=F)	P(R=T)
False(F)	0.8	0.2
True(T)	0.2	0.8

$(\theta, \alpha, \beta, \gamma)$

$\Rightarrow$ These tables are learned by the model (Baysian network)

$P(S|C) \Rightarrow$

Cloudy	P(S=F)	P(S=T)
F	0.5	0.5
T	0.9	0.1

$P(W|R,S) \Rightarrow$

Sprinkler	Rain	P(W=F)	P(W=T)
F	F	1.0	0.0
T	F	0.1	0.9
F	T	0.1	0.9
T	T	0.01	0.99

Dataset — Given

C	S	R	W
1	0	1	1
1	1	1	0
0	0	1	1

(a) $P(S=1)$?

Ans:- Joint prob. :- $\sum_{C=0}^{1} \sum_{R=0}^{1} \sum_{W=0}^{1} P(C, R, W, S=1)$

{Inference phase}

$= \sum_{CRW} P(C) \cdot P(S=1|C) \cdot P(R|C) \cdot P(W|S,R)$

$= P(C=0) \cdot P(S=1 \mid C=0) \cdot P(R=0 \mid C=0) \cdot P(W=0 \mid S=1, R=0)$

$+$

$P(C=0) \cdot P(S=1 \mid C=0) \cdot P(R=0 \mid C=0) \cdot P(W=1 \mid S=1, R=0)$

$+$

$P(C=0) \cdot P(S=1 \mid C=0) \cdot P(R=1 \mid C=0) \cdot P(W=0 \mid S=1, R=1)$

$+$

$P(C=0) \cdot P(S=1 \mid C=0) \cdot P(R=1 \mid C=0) \cdot P(W=1 \mid S=1, R=1)$

$+$

$P(C=1) \cdot P(S=1 \mid C=1) \cdot P(R=0 \mid C=1) \cdot P(W=0 \mid S=1, R=0)$

$+$

$P(C=1) \cdot P(S=1 \mid C=1) \cdot P(R=0 \mid C=1) \cdot P(W=1 \mid S=1, R=0)$

$+$

$P(C=1) \cdot P(S=1 \mid C=1) \cdot P(R=1 \mid C=1) \cdot P(W=0 \mid S=1, R=1)$

$+$

$P(C=1) \cdot P(S=1 \mid C=1) \cdot P(R=1 \mid C=1) \cdot P(W=1 \mid S=1, R=1)$

($2^{(3)} = 8$ combo)

C	R	W
0	0	0
0	0	1
0	1	0
0	1	1
1	0	0
1	0	1
1	1	0
1	1	1

$\Rightarrow (0.5 \times 0.5 \times 0.8 \times 0.1) + (0.5 \times 0.5 \times 0.8 \times 0.9) + (0.5 \times 0.5 \times 0.2 \times 0.01) + (0.5 \times 0.5 \times 0.2 \times 0.99) + (0.5 \times 0.1 \times 0.2 \times 0.1) + (0.5 \times 0.1 \times 0.2 \times 0.9) + (0.5 \times 0.1 \times 0.8 \times 0.01) + (0.5 \times 0.1 \times 0.8 \times 0.99)$

$\Rightarrow 0.02 + 0.18 + 0.0005 + 0.0495 + 0.001 + 0.009 + 0.0004 + 0.0396$

$\boxed{P(S=1) \Rightarrow 0.3}$