

Q4. Object X Y

| | | |
|-------|---|---|
| Med A | 1 | 1 |
| Med B | 2 | 1 |
| Med C | 4 | 3 |
| Med D | 5 | 4 |

Page No.

Date:

$K=2 \rightarrow$ Let's initialize Med A & Med C as two centroids

| Object | Distance to centroid 1 (Med A) (1,1) | Distance to centroid 2 (Med C) (4,3) | Cluster |
|-------------|--------------------------------------|--------------------------------------|---------|
| (1,1) Med A | 0 | 3.61 | 1 |
| (2,1) Med B | 1 | 2.83 | 1 |
| (4,3) Med C | 3.61 | 0 | 2 |
| (5,4) Med D | 5 | 1.41 | 2 |

Cluster 1 \Rightarrow (1,1) & (2,1)

$$\frac{1+2}{2} = 1.5 \quad \& \quad \frac{1+1}{2} = 1$$

New centroid 1 \Rightarrow (1.5, 1)

Cluster 2 \Rightarrow (4,3) & (5,4)

$$\frac{4+5}{2} = 4.5 \quad ; \quad \frac{3+4}{2} = 3.5$$

New centroid 2 \Rightarrow (4.5, 3.5)

| Object | Dist to centroid 1 (1.5, 1) | Dist to centroid 2 (4.5, 3.5) | Cluster |
|-------------|-----------------------------|-------------------------------|---------|
| (1,1) Med A | 0.5 | 4.30 | 1 |
| (2,1) B | 0.5 | 3.58 | 1 |
| (4,3) C | 3.20 | 0.71 | 2 |
| (5,4) D | 4.61 | 0.71 | 2 |

Same

Cluster 1 \Rightarrow Medicine A & Medicine B

Cluster 2 \Rightarrow Medicine C & Medicine D

Q5:-
apriori

Trans-ID

Items

A

1, 3, 4, 6

B

2, 3, 5, 7

C

1, 2, 3, 5, 8

D

2, 5, 9, 10

E

1, 4

$$\text{min support} = 30\% = \frac{30}{100}$$

$$\frac{30}{100} \times 100 = 30$$

Date:

$$\text{min confidence} = 70\% = \frac{70}{100}$$

$$\frac{70}{100} \times 100 = 70$$

1 item set $\leftarrow CI$ $CI \rightarrow$ Item Count

| Item | Count | Item | Count |
|------|-------|------|-------|
| 1 | 3 | 1 | 3 |
| 2 | 3 | 2 | 3 |
| 3 | 3 | 3 | 3 |
| 4 | 2 | 4 | 2 |
| 5 | 3 | 5 | 3 |
| 6 | 1 | | |
| 7 | 1 | | |
| 8 | 1 | | |
| 9 | 1 | | |
| 10 | 1 | | |

$\{1\}, \{2\}, \{3\}, \{4\}, \{5\}$

2 - item set

| Item | Count | Item | Count |
|------|-------|------|-------|
| 1, 2 | 1 | 1, 3 | 2 |
| 1, 3 | 2 | 1, 4 | 2 |
| 1, 4 | 2 | 2, 3 | 2 |
| 1, 5 | 1 | 2, 5 | 3 |
| 2, 3 | 2 | 3, 5 | 2 |
| 2, 4 | 0 | | |
| 2, 5 | 3 | | |
| 3, 4 | 1 | | |
| 3, 5 | 2 | | |
| 4, 5 | 0 | | |

$\{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 5\}, \{3, 5\}$

3 - item set

| Item | Item | Count |
|---------------|---------|-------|
| $\{1, 3, 4\}$ | 1, 3, 4 | 1 |
| $\{2, 3, 5\}$ | 2, 3, 5 | 2 |

Strong association rule

2-item data set

→ 1, 3

$$1 \rightarrow 3 = \frac{2}{3} = 66.67 \quad \times$$

$$3 \rightarrow 1 = \frac{2}{3} = 66.67 \quad \times$$

→ 1, 4

$$1 \rightarrow 4 = \frac{2}{3} = 66.67 \quad \times$$

$$4 \rightarrow 1 = \frac{2}{2} = 100 \quad \checkmark$$

$$\rightarrow 2, 3 = \frac{2}{3} = 66.67 \quad \times$$

$$3 \rightarrow 2 = \frac{2}{3} = 66.67 \quad \times$$

→ 2, 5

$$2 \rightarrow 5 = \frac{3}{3} = 100\% \quad \checkmark$$

$$5 \rightarrow 2 = \frac{3}{3} = 100\% \quad \checkmark$$

→ 3, 5

$$3 \rightarrow 5 = \frac{2}{3} = 66.67 \quad \times$$

$$5 \rightarrow 3 = \frac{2}{3} = 66.67 \quad \times$$

3-item set

(2, 3, 5)

$$2 \rightarrow 3, 5 = \frac{2}{3} = 66.67$$

$$3 \rightarrow 2, 5 = \frac{2}{3} = 66.67$$

$$5 \rightarrow 2, 3 = \frac{2}{3} = 66.67$$

$$3, 5 \rightarrow 2 = \frac{2}{2} = 100 \quad \checkmark$$

$$2, 5 \rightarrow 3 = \frac{2}{3} = 66.67$$

$$2, 3 \rightarrow 5 = \frac{2}{2} = 100 \quad \checkmark$$

Valid are → Strong association Rules

| Rule | Confidence |
|----------|------------|
| 4 → 1 | 100% |
| 2 → 5 | 100% |
| 5 → 2 | 100% |
| 3, 5 → 2 | 100% |
| 2, 3 → 5 | 100% |

Q6-1) ~~Linear~~ Linear regression

X Y X² XY

2 3 4 6

4 7 16 28

6 5 36 30

8 10 64 80

20 25 120 ~~144~~ 144

$$b = \frac{N(\sum XY) - (\sum X)(\sum Y)}{n(\sum X^2) - (\sum X)^2} = \frac{4(144) - (20)(25)}{4(120) - (20)^2} = 0.95$$

$$a = \frac{\sum Y - b(\sum X)}{N} = \frac{25 - (0.95)(20)}{4} = 1.5$$

$$Y = a + bX$$

$$Y = 1.5 + 0.95X$$