* Name-: Adke Prasad Sukdev.
* Div-; A. Roll No.-: 01.
* Class-: SYMCA Batch-: A1.
* Theory Assignment - II.
* subject -: Machine Learning.

Q.1. Using k-means clustering cluster following data into two clusters.
(2,3), (3,4), (5,6), (8,7), (7,3), (4,6). show each of the steps.

→ solution-:

Consider,
$c_1 = (2,3)$ as a first centroid.
$c_2 = (4,6)$ as a second centroid.

	$c_1(2,3)$	$c_2(4,6)$	Class
(2,3)	0	3.61	C1
(3,4)	1.41	2.23	C1
(5,6)	4.24	1	C2
(8,7)	7.21	4.12	C2
(7,3)	~~2.05~~	4.24	C2
(4,6)	3.66	0	C2

∴ cluster 1 = [(2,3), (3,4)]
cluster 2 = [(5,6), (8,7), (7,3), (4,6)]

Now, new centroids will be,

$$c_1' = \left(\frac{2+3}{2}, \frac{3+4}{2}\right) = (2.5, 3.5)$$

$$\boxed{c_1' = (2.5, 3.5)}$$

$$c_2' = \left(\frac{5+8+7+4}{4}, \frac{6+7+3+6}{4}\right)$$

$$\boxed{C_2' = (6, 5.5)}$$

Now,

	$C_1(2.5, 3.5)$	$C_2(6, 5.5)$	class
(2,3)	0.7071	4.7169	C_1
(3,4)	0.7071	3.3541	C_1
(5,6)	3.5355	1.11	C_2
(8,7)	6.5	2.5	C_2
(7,3)	4.5	2.69	C_2
(4,6)	2.9	2.06	C_2

∴ cluster 1 = [(2,3), (3,4)]

cluster 2 = [(5,6), (8,7), (7,3), (4,6)]

∴ As there is no changes in previous & new clusters

⇒ C_1 [(2,3), (3,4)] & C_2 [(5,6), (8,7), (7,3), (4,6)] are the final clusters

Q.2. Find all association rules using apriori algorithm in the following database with minimum support = 2 and minimum confidence = 65%.

Transactions	Data Items.
T1	Pen, Pencil, Notebook.
T2	Pencil, File.
T3	Pen, Pencil, Notebook, File.
T4	Pen, Notebook.
T5	Pencil, Scale, File.
T6	Pencil, Scale.
T7	Pen, Pencil, Scale.

→ solution:-

Now, We have to calculate support & confidence for one data item. (k=1).

∴

Item	support
Pen	4
Pencil	6
Notebook	3
File	3
Scale.	3

Now, for k=2.

Item	support
Pen, Pencil	3
Pen, Notebook	3
Pen, File	1
Pen, Scale	1
Pencil, Notebook	2

Pencil, File	3
Pencil, Scale	3
File, Scale	2

Now, for minimum support = 2,

Item	min_support = 2
Pen, Pencil	3
Pen, Notebook	3
Pencil, Notebook	2
Pencil, File	3
Pencil, Scale	3
File, Scale	2

Now, Generate 3-itemsets

3-Itemsets	Freqncy
{Pen, Pencil, Notebook}	2
{Pencil, Notebook, file}	1
{Pencil, File, Scale}	1

itemset with minimum_support ≥ 2
i.e {Pen, Pencil, Notebook}

Generating Association Rules.
Rules from {Pen, Pencil} {support = 3)
- Pen ⇒ Pencil: confidence = (3/4)*100 = 75% (Pass)
- Pencil ⇒ Pen: confidence = (3/6)*100 = 50% (fail)

Rules from {Pen, Notebook} (support = 3)

- Pen ⇒ Notebook : confidence = (3/4) * 100 = 75%. (Pass)
- Notebook ⇒ Pen : confidence = (3/3) * 100 = 100%. (Pass)

Rules from {Pencil, Notebook} (support = 2)

- Pencil ⇒ Notebook : confidence = (2/6) * 100 = 33.33%. (fail)
- Notebook ⇒ Pencil : confidence = (2/3) * 100 = 66.67%. (Pass)

Rules from {Pencil, File} (support = 2)

- Pencil ⇒ file : confidence = (2/6) * 100 = 33.33%. (fail)
- File ⇒ Pencil : confidence = (2/3) * 100 = 66.67%. (Pass)

Rules from {Pencil, Scale} (support = 3)

- Pencil ⇒ scale : confidence = (3/6) * 100 = 50%. (fail)
- Scale ⇒ Pencil : confidence = (3/3) * 100 = 100%. (pass)

Rules from {File, scale} (support = 2)

- File ⇒ scale : confidence = (2/3) * 100 = 66.67%. (Pass)
- Scale ⇒ File : confidence = (2/3) * 100 = 66.67%. (Pass)

Rules from {Pen, Pencil, Notebook} (support = 2)

- Pen, Pencil ⇒ Notebook : confidence = (2/3) * 100 = 66.67%. (Pass)
- Pen, Notebook ⇒ Pencil : confidence = (2/3) * 100 = 66.67%. (Pass)
- Pencil, Notebook ⇒ Pen : confidence = (2/2) * 100 = 100%. Pass.

So, Finally, (Minimum Confidence ≥ 65%)

1. Pen ⇒ Pencil (75%)
2. Pen ⇒ Notebook (75%)
3. Notebook ⇒ Pen (100%)

4. Notebook ⇒ Pencil (66.67%)
5. File ⇒ Pencil (66.67%)
6. Scale ⇒ Pencil (100%)
7. File ⇒ scale (66.67%)
8. Scale ⇒ File (66.67%)
9. Pen, Pencil ⇒ Notebook (66.67%)
10. Pen, Notebook ⇒ Pencil (66.67%)
11. Pencil, Notebook ⇒ Pen (100%)

These are association rules with min_support = 2 and minimum confidence = 65%.