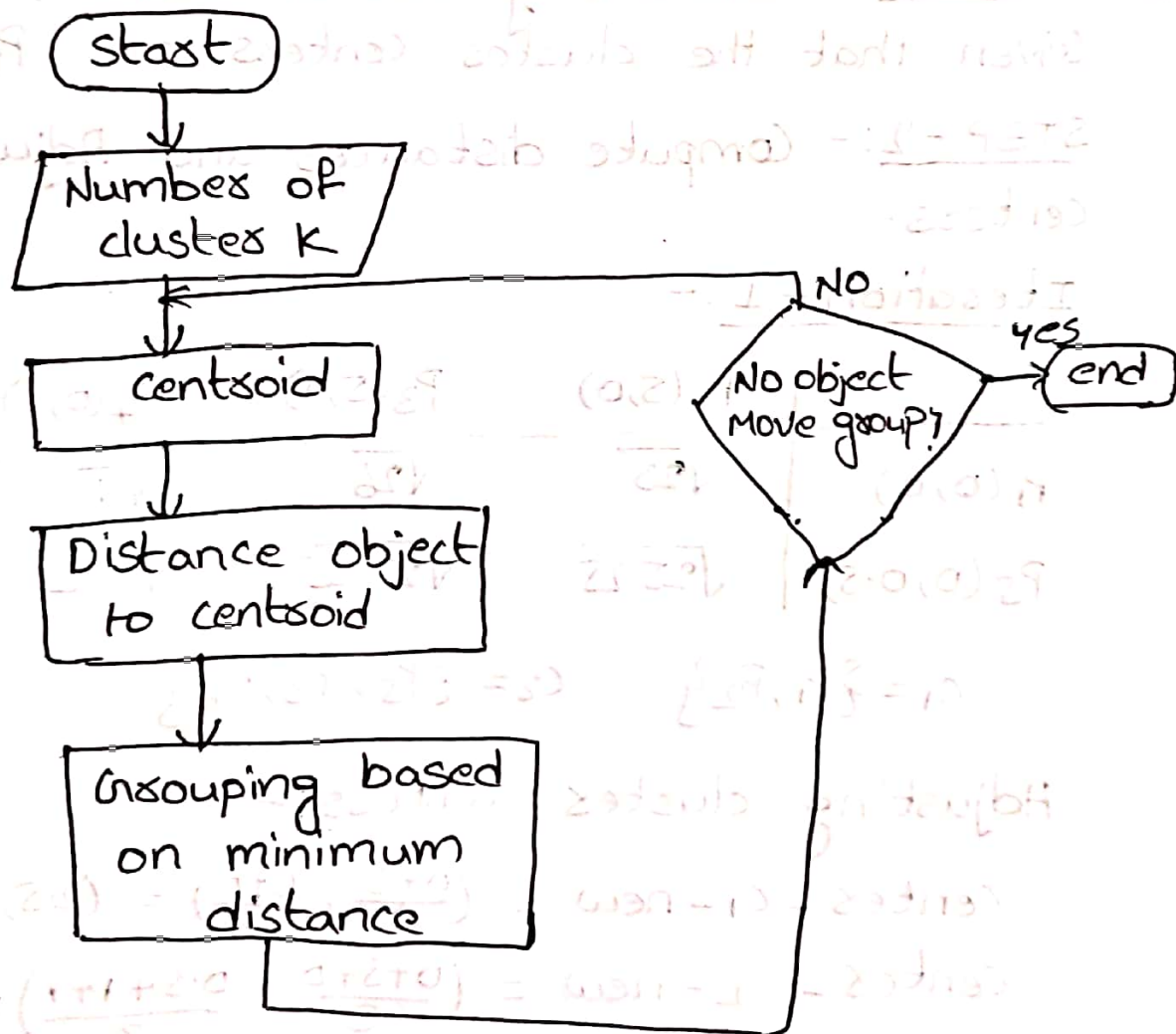


KMeans clustering:-

KMeans algorithm will do the three steps below until convergence.

1. Determine the centroid co-ordinate
2. Determine the distance of each object to the centroid.
3. Group the object based on minimum distance.

Flow chart:



Problem on KMeans clustering:-

1) Consider a set of five 2-dimensional points

$$P_1 = (0, 0)$$

$$P_2 = (5, 0)$$

$$P_3 = (5, 1)$$

$$P_4 = (0, 1)$$

$$P_5 = (0, 0.5)$$

Apply the K-means algorithm to cluster these points into two clusters. The initial cluster centers are P_1 and P_5 .

A) STEP-1:- Initializing cluster centers.
Given that the cluster centers are P_1 and P_5

STEP-2:- Compute distances and Adjust cluster centers.

Iteration-1:-

	$P_2(5, 0)$	$P_3(5, 1)$	$P_4(0, 1)$
$P_1(0, 0)$	$\sqrt{25}$	$\sqrt{26}$	$\sqrt{1}$
$P_5(0, 0.5)$	$\sqrt{25.25}$	$\sqrt{25.25}$	$\sqrt{0.25}$

$$C_1 = \{P_1, P_2\} \quad C_2 = \{P_5, P_3, P_4\}$$

Adjusting cluster centers:-

$$\text{center } C_1 - \text{new} = \left(\frac{0+5}{2}, \frac{0+0}{2} \right) = (2.5, 0)$$

$$\text{center } C_2 - \text{new} = \left(\frac{0+5+0}{3}, \frac{0.5+1+1}{3} \right) = \left(\frac{5}{3}, \frac{2.5}{3} \right)$$

Iteration - 2:-

	P_1 (0,0)	P_2 (5,0)	P_3 (5,1)	P_4 (0,1)	P_5 (0,0.5)
C_1 -new (2.5, 0)	$\sqrt{6.25}$	$\sqrt{6.25}$	$\sqrt{7.25}$	$\sqrt{7.25}$	$\sqrt{6.5}$
C_2 -new ($\frac{5}{2}, \frac{2.5}{3}$)	$\sqrt{3.47}$	$\sqrt{11.8}$	$\sqrt{11.18}$	$\sqrt{2.8}$	$\sqrt{2.88}$

$$C_1 = \{P_2, P_3\} \quad C_2 = \{P_1, P_4, P_5\}$$

Adjusting cluster Centres:-

$$\text{Centres} - C_1 - \text{new} = \left(\frac{5+5}{2}, \frac{0+1}{2} \right) = (5, 0.5)$$

$$\text{Centres} - C_2 - \text{new} = \left(\frac{0+0+0}{3}, \frac{0+1+0.5}{3} \right) = (0, 0.5)$$

Iteration - 3:-

	P_1 (0,0)	P_2 (5,0)	P_3 (5,1)	P_4 (0,1)	P_5 (0,0.5)
C_1 -new (5, 0.5)	$\sqrt{25.25}$	$\sqrt{0.25}$	$\sqrt{0.25}$	$\sqrt{25.25}$	$\sqrt{25}$
C_2 -new (0, 0.5)	$\sqrt{0.25}$	$\sqrt{25.25}$	$\sqrt{25.25}$	$\sqrt{0.25}$	$\sqrt{0}$

$$C_1 = \{P_2, P_3\} \quad C_2 = \{P_1, P_4, P_5\}$$

- Comparing the grouping of last iteration reveals that the objects does not move group anymore. Thus, the computation of the k-means clustering has reached its stability and no more iteration is needed.

- We get the final grouping as the results
 $C_1 = \{P_2, P_3\} \quad ; \quad C_2 = \{P_1, P_4, P_5\}$