INDIAN INSTITUTE OF INFORMATION TECHNOLOGY KALYANI

Kalyani, Nadia-741235, West Bengal
(An institute of National Importance)

Course Title: Machine Learning Lab   Paper Code: CSC-612   Spring 2025

Lab Test

Date: 04/03/2025   Due date: 04/03/2025   1 PM   10 Marks

Implement the K-Means clustering algorithm and show the results of each of the following steps. Consider K = 2.

1. Randomly assign means m1 and m2.
2. Initialize C1 and C2 by the mean values m1 and m2.
3. Calculate D_{1i} (Euclidian distance from C_1 to i^{th} value) and D_{2i} (Euclidian distance from C_2 to i^{th} value) in a table.
4. Among D_{n1} and D_{n2} check the least value and assign the nth datapoint to the corresponding cluster.
5. Calculate the mean (M_i) of datapoints for each cluster generated in step 4.
6. Set $C_i = M_i$ for each cluster.
7. Repeat step 4, 5, 6 until there is no change in the mean values in two consecutive iterations.

Consider the following example for reference: [2, 4, 10, 12, 3, 20, 30, 11, 25] and k = 2

Datapoint	D1	D2	Cluster
2	2	9	C1
4	0	7	C1
10	6	1	C2
12	8	1	C2
3	1	8	C1
20	16	9	C2
30	26	19	C2
11	7	0	C2
25	21	14	C2

Datapoint	D1	D2	Cluster
2	1	16	C1
4	1	14	C1
3	0	15	C1
10	7	8	C1
12	9	6	C2
20	17	2	C2
30	27	12	C2
11	8	7	C2
25	22	7	C2

Datapoint	D1	D2	Cluster
2	2.75	17.6	C1
4	0.75	15.6	C1
3	1.75	16.6	C1
10	5.25	9.6	C1
12	7.25	7.6	C1
20	15.25	0.4	C2
30	25.25	10.4	C2
11	6.25	8.6	C1
25	20.25	5.4	C2

8. Repeat the same procedure for the three dimensional list: [{2,3,4}, {10,11,12}, {3,15,20}, {30,9,11}, {25,22,20}, {15,9,25}, {12,8,25}, {9,18,21}, {6,9,12}, {15,14,13}] and k = 2
9. Show the points in 3 dimensional space.
10. Show the groups using colour coding (Red for cluster 1 and Blue for cluster 2) in 3 dimensional space after running the algorithm.