# NPTEL Week-10 Live Session

Powered by:

on Machine Learning and Deep Learning - Fundamentals and Applications (noc24_ee146)

A course offered by: Prof. Manas Kamal Bhuyan, IIT Guwahati

## NPTEL Quiz Solution: Week-9: Clustering


By

**Arka Roy**

NPTEL PMRF TA

Prime Minister's Research Fellow

Department of Electrical Engineering, IIT Patna

Web: https://sites.google.com/view/arka-roy/home


PMRF

Prime Minister's Research Fellows

Ministry of Education

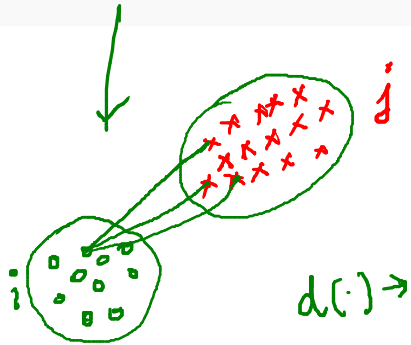
Government of India

1) Which of the following statements is true about HAC with complete linkage?

- It tends to produce elongated clusters.
- It produces spherical clusters.
- It tends to merge clusters with the minimum distance between centroids.
- It considers the maximum distance between points in different clusters.

Single linkage, Complete linkage ⇒ Distance metric used in the Hierarchical Agglomerative clustering method.

In case of Complete linkage: ⇒ "the maximum distance between any pair of points, one from each cluster".




$d(\cdot)$ → Euclidean distance.

Complete linkage = $\max\{d_{ij}\}$

2) Lets consider we have cluster points P1(1,3) , P2(2,2) , P3(5,8) , P4(8,5) , P5(3,9) , P6(10,7) , P7(3,3) , P8(9,4) , P9(3,7). First, we take our K value as 3 and we assume that our Initial cluster centers are P7(3,3), P9(3,7), P8(9,4) as C1, C2, C3. Then find out the new centroids after 1 iterations for the above data points. **1 point**

- (2, 2.7), (3, 6), (9, 5.3)
- (2, 2.7), (3.7, 8), (9, 5.3)
- (2, 6), (3.7, 8), (4, 9.86)
- (2, 9), (3, 2), (9, 7)

$$d(P_i, C_j) \Big|_{\substack{i=1 \\ j=1}}^{\substack{i=9 \\ j=3}} \downarrow \sqrt{(x_2-x_1)^2 + (y_2-y_1)^2}$$

| Points | $C_1$ (3,3) | $C_2$ (3,7) | $C_3$ (9,4) | Decision of Cent |
|---|---|---|---|---|
| | 2 | 4.5 | 8.1 | $C_1$ |
| $P_1$ (1,3) | 1.4 | 5.1 | 7.3 | $C_1$ |
| $P_2$ (2,2) | 5.3 | 2.2 | 5.7 | $C_2$ |
| $P_3$ (5,8) | 5.4 | 5.4 | 5.1 | $C_3$ |
| $P_4$ (8,5) | 6 | 2 | 7.9 | $C_2$ |
| $P_5$ (3,9) | 8.1 | 7 | 3.2 | $C_3$ |
| $P_6$ (10,7) | 0 | 4 | 6.1 | $C_1$ |
| $P_7$ (3,3) | 6.1 | 6.7 | 0 | $C_3$ |
| $P_8$ (9,4) | — | 0 | — | $C_2$ |
| $P_9$ (3,7) | | | | |

$C_1 :- \{P_1, P_2, P_7\}$

$C_2 :- \{P_3, P_5, P_9\}$

$C_3 :- \{P_4, P_6, P_8\}$

2) Lets consider we have cluster points P1(1,3) , P2(2,2) , P3(5,8) , P4(8,5) , P5(3,9) , P6(10,7) , P7(3,3) , P8(9,4) , P9(3,7). First, we take our K value as 3 and we assume that our Initial cluster centers are P7(3,3), P9(3,7), P8(9,4) as C1, C2, C3. Then find out the new centroids after 1 iterations for the above data points. **1 point**

- (2, 2.7), (3, 6), (9, 5.3)
- (2, 2.7), (3.7, 8), (9, 5.3)
- (2, 6), (3.7, 8), (4, 9.86)
- (2, 9), (3, 2), (9, 7)

$C_1 :- \{P_1, P_2, P_7\}$

$C_2 :- \{P_3, P_5, P_9\}$

$C_3 :- \{P_4, P_6, P_8\}$

After iteration - 1

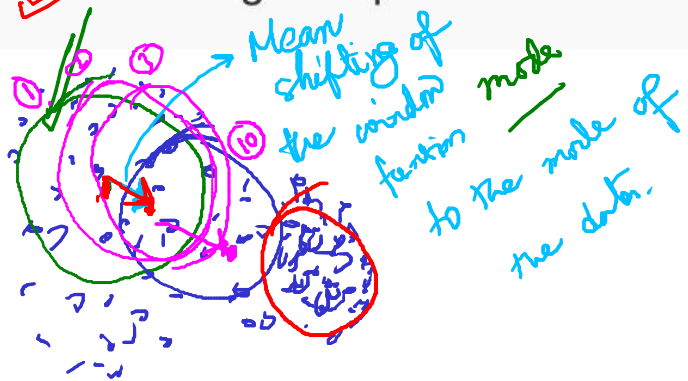
Centroid for cluster 1 $= \left(\frac{1+2+3}{3}, \frac{3+2+3}{3}\right) = (2, 2.7)$

cluster 2 $= \left(\frac{5+3+3}{3}, \frac{8+9+7}{3}\right) = (3.7, 8)$

cluster 3 $= \left(\frac{8+10+9}{3}, \frac{5+7+4}{3}\right) = (9, 5.3)$

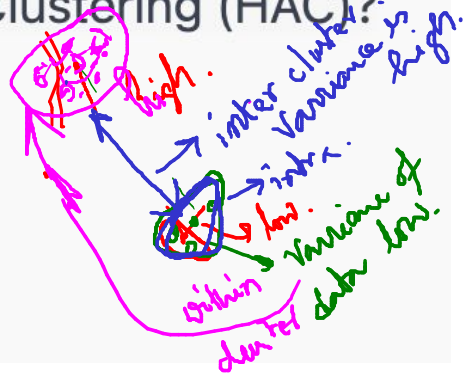
3) What is the primary idea behind Mean-Shift clustering?

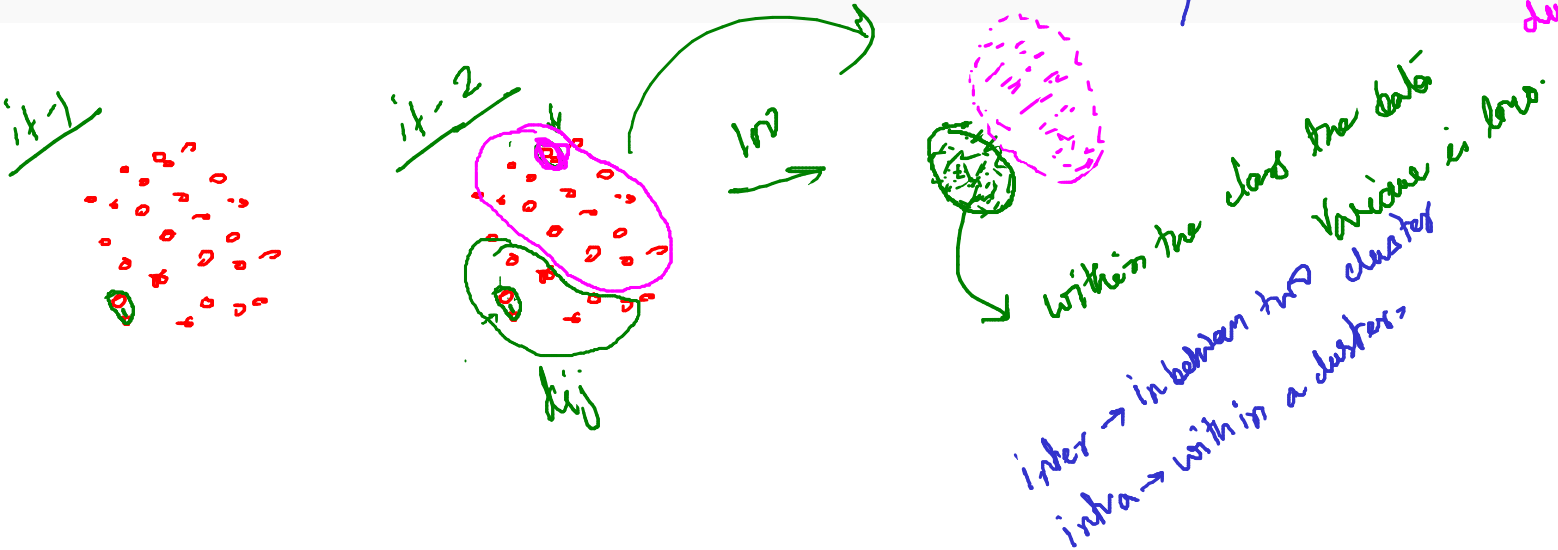
- Iteratively shifting data points towards the mean of the entire dataset.
- Assigning data points to clusters based on predefined centroids.
- Minimizing the sum of squared distances within each cluster.
- Shifting data points towards the mode (highest density area) of the data distribution.



4) What is the main objective of Hierarchical Agglomerative Clustering (HAC)?

- Minimize within-cluster variance
- Maximize between-cluster variance
- Minimize the number of clusters
- Maximize the number of data points in each cluster



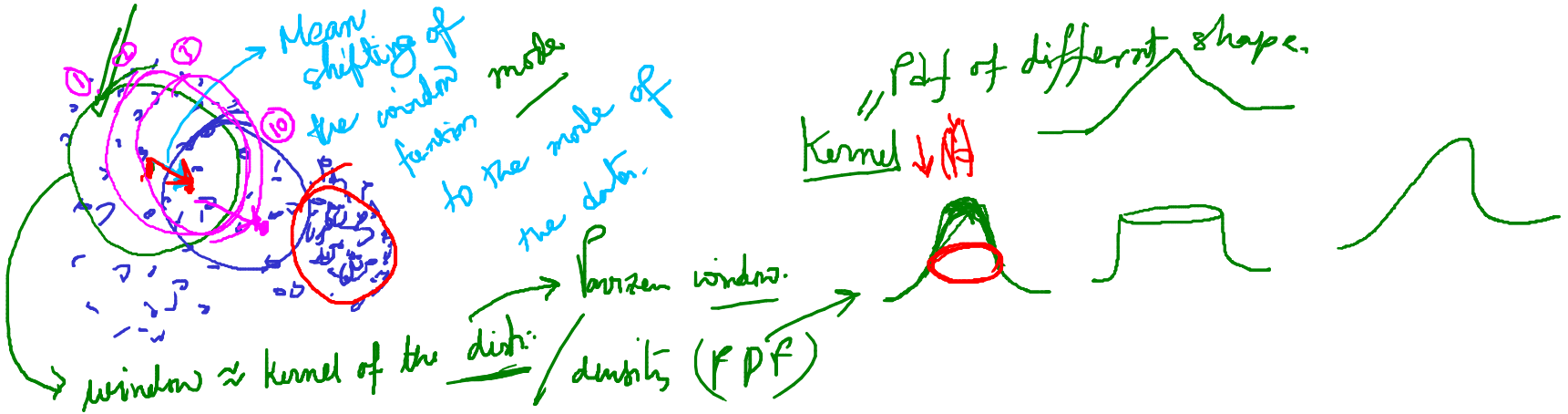
no uncertainity case → no fuzziness → Standard K means.

5) When m=1 in Fuzzy K-means, the algorithm behaves similarly to:

- Hierarchical Clustering.
- Mean-Shift Clustering.
- Standard K-means.
- DBSCAN.

6) In Mean Shift clustering, what is the "kernel" used for?

- It defines the number of clusters.
- It specifies the dimensionality of the data.
- It represents the shape of the density function.
- It determines the number of iterations.



7) What is the key difference between Fuzzy K-Means and the traditional K-Means clustering algorithm?

- Fuzzy K-Means allows each data point to belong to multiple clusters with varying degrees of membership.
- Fuzzy K-Means does not require the number of clusters to be specified.
- Fuzzy K-Means assigns each data point to exactly one cluster.
- Fuzzy K-Means is a supervised learning algorithm, while K-Means is unsupervised.

8) If we utilize the Hierarchical Agglomerative Clustering method on given below data with the aim of identifying clusters based on the smallest values obtained from Euclidean distance calculations, the dendrogram will be -

$\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$

| point | x coordinate | y coordinate |
|---|---|---|
| p1 | 0.4005 | 0.5306 |
| p2 | 0.2148 | 0.3854 |
| p3 | 0.3457 | 0.3156 |
| p4 | 0.2652 | 0.1875 |
| p5 | 0.0789 | 0.4139 |
| p6 | 0.4548 | 0.3022 |

**Table 1:** X-Y coordinates of six points.

| | p1 | p2 | p3 | p4 | p5 | p6 |
|---|---|---|---|---|---|---|
| p1 | 0.0000 | 0.2357 | 0.2218 | 0.3688 | 0.3421 | 0.2347 |
| p2 | 0.2357 | 0.0000 | 0.1483 | 0.2042 | 0.1388 | 0.2540 |
| p3 | 0.2218 | 0.1483 | 0.0000 | 0.1513 | 0.2843 | 0.1100 |
| p4 | 0.3688 | 0.2042 | 0.1513 | 0.0000 | 0.2932 | 0.2216 |
| p5 | 0.3421 | 0.1388 | 0.2843 | 0.2932 | 0.0000 | 0.3921 |
| p6 | 0.2347 | 0.2540 | 0.1100 | 0.2216 | 0.3921 | 0.0000 |

**Table 2:** Distance Matrix for Six Points

At iteration 1:- First cluster of data pair is = $P_3 P_6$.

As distance between $P_3 P_6$ is lowest in table 2

$P_3 P_6 \rightarrow C_1$

distance matrix: Table 3

| | $P_1$ | $P_2$ | $P_4$ | $P_5$ | $C_1 (P_3 P_6)$ |
|---|---|---|---|---|---|
| $P_1$ | 0 | 0.2357 | 0.3688 | 0.3421 | 0.2218 |
| $P_2$ | 0.2357 | 0 | 0.2042 | 0.1388 | 0.1483 |
| $P_4$ | 0.3688 | 0.2042 | 0 | 0.2932 | 0.1513 |
| $P_5$ | 0.3421 | 0.1388 | 0.2932 | 0 | 0.2843 |
| $C_1$ | 0.2218 | 0.1483 | 0.1513 | 0.2843 | 0 |

iteration 2 :-

$d(C, P) \approx d(C_1, P_1) = \min\{d(P_3 P_1), d(P_6 P_1)\} = 0.2218$

$d(C_1, P_2) = \min\{d(P_3, P_2), d(P_6, P_2)\}$

$= 0.1483$

$C_1 = \{P_3, P_6\}$.

$d(C_1, P_1)$

$= \min\{d(P_3 P_1), d\{P_6, P_1\}\}$.

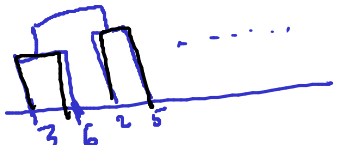
check $\min\{T3\}$

$= 0.1388$

$= d(P_2 P_5)$

$d(C_1, P_4) = \min\{d(P_3, P_4), d(P_6, P_4)\}$

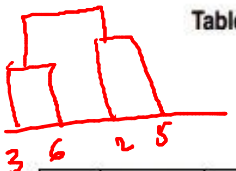
$= 0.1513$



8) If we utilize the Hierarchical Agglomerative Clustering method on given below data with the aim of identifying clusters based on the smallest values obtained from Euclidean distance calculations, the dendrogram will be -

| point | x coordinate | y coordinate |
|---|---|---|
| p1 | 0.4005 | 0.5306 |
| p2 | 0.2148 | 0.3854 |
| p3 | 0.3457 | 0.3156 |
| p4 | 0.2652 | 0.1875 |
| p5 | 0.0789 | 0.4139 |
| p6 | 0.4548 | 0.3022 |

**Table :** X-Y coordinates of six points.




| | p1 | p2 | p3 | p4 | p5 | p6 |
|---|---|---|---|---|---|---|
| p1 | 0.0000 | 0.2357 | 0.2218 | 0.3688 | 0.3421 | 0.2347 |
| p2 | 0.2357 | 0.0000 | 0.1483 | 0.2042 | 0.1388 | 0.2540 |
| p3 | 0.2218 | 0.1483 | 0.0000 | 0.1513 | 0.2843 | 0.1100 |
| p4 | 0.3688 | 0.2042 | 0.1513 | 0.0000 | 0.2932 | 0.2216 |
| p5 | 0.3421 | 0.1388 | 0.2843 | 0.2932 | 0.0000 | 0.3921 |
| p6 | 0.2347 | 0.2540 | 0.1100 | 0.2216 | 0.3921 | 0.0000 |

**Table :** Distance Matrix for Six Points

iteration 3

| | $P_1$ | $P_4$ | $P_2-P_5$ | $P_3-P_6$ |
|---|---|---|---|---|
| $P_1$ | 0 | 0.3688 | 0.2357 | 0.2218 |
| $P_4$ | 0.3688 | 0 | 0.2042 | 0.1513 |
| $P_2-P_5$ | 0.2357 | 0.2042 | 0 | 0.1483 |
| $P_3-P_6$ | 0.2218 | 0.1513 | 0.1483 | 0 |

$C_2-C_1$ is creating another cluster.

$$d(C_2,P_1) = \min\{d(P_2,P_1), d(P_5,P_1)\}$$

$$d(C_2,P_4) = \min\{d(P_2,P_4), d(P_5,P_4)\}$$

$$d(C_2,C_1) = \min\{d(P_2-P_5, P_3-P_6)\} = \min\{d(P_2,P_3), d(P_2,P_6), d(P_5,P_3), d(P_5,P_6)\}$$

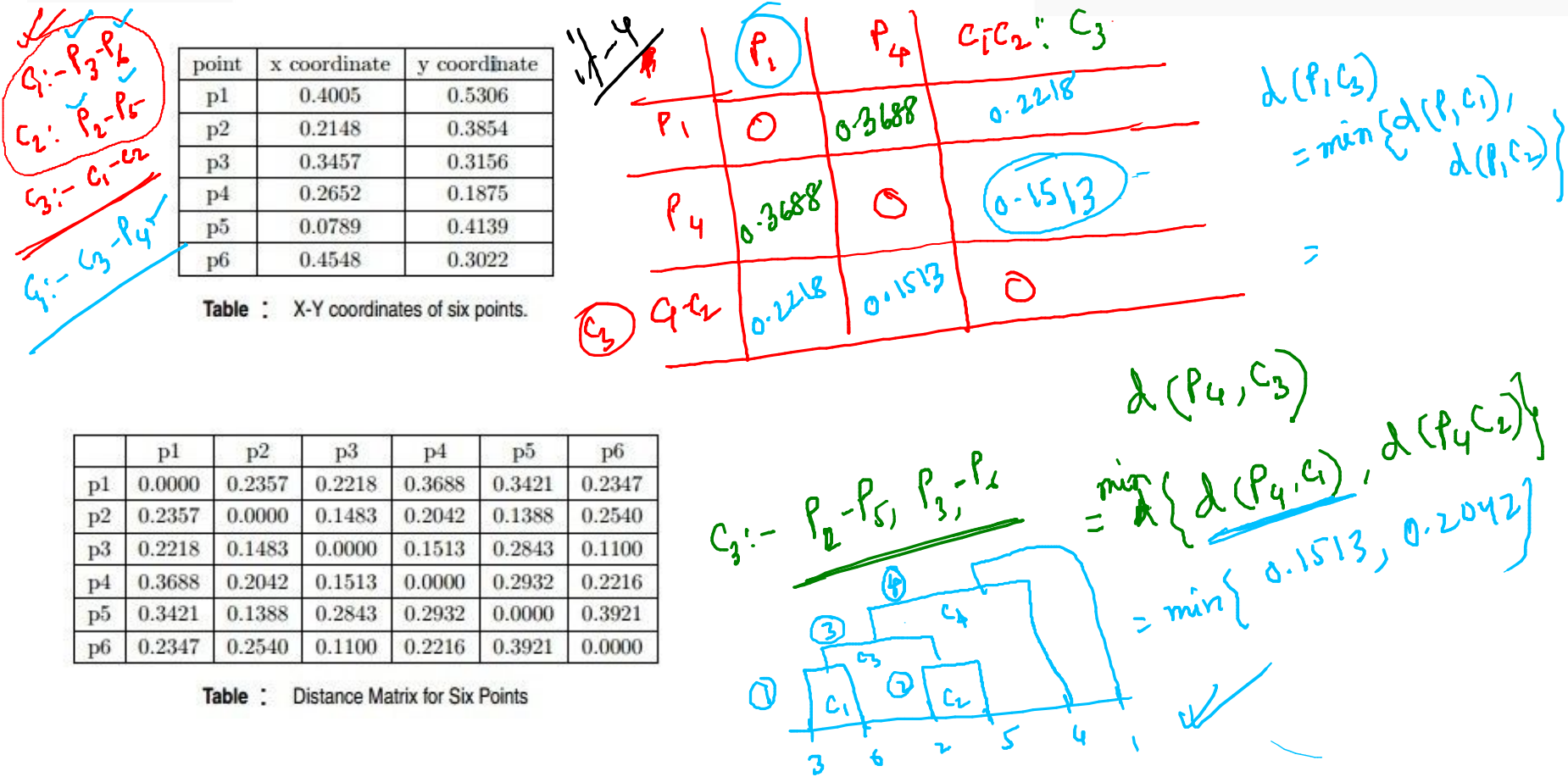
8) If we utilize the Hierarchical Agglomerative Clustering method on given below data with the aim of identifying clusters based on the smallest values obtained from Euclidean distance calculations, the dendrogram will be -

| point | x coordinate | y coordinate |
|---|---|---|
| p1 | 0.4005 | 0.5306 |
| p2 | 0.2148 | 0.3854 |
| p3 | 0.3457 | 0.3156 |
| p4 | 0.2652 | 0.1875 |
| p5 | 0.0789 | 0.4139 |
| p6 | 0.4548 | 0.3022 |

**Table :** X-Y coordinates of six points.

| | p1 | p2 | p3 | p4 | p5 | p6 |
|---|---|---|---|---|---|---|
| p1 | 0.0000 | 0.2357 | 0.2218 | 0.3688 | 0.3421 | 0.2347 |
| p2 | 0.2357 | 0.0000 | 0.1483 | 0.2042 | 0.1388 | 0.2540 |
| p3 | 0.2218 | 0.1483 | 0.0000 | 0.1513 | 0.2843 | 0.1100 |
| p4 | 0.3688 | 0.2042 | 0.1513 | 0.0000 | 0.2932 | 0.2216 |
| p5 | 0.3421 | 0.1388 | 0.2843 | 0.2932 | 0.0000 | 0.3921 |
| p6 | 0.2347 | 0.2540 | 0.1100 | 0.2216 | 0.3921 | 0.0000 |

**Table :** Distance Matrix for Six Points



9) How does hierarchical agglomerative clustering work?

- It randomly assigns data points to different clusters, and then iteratively adjusts the assignments until a stopping criterion is met.
- It applies a clustering algorithm to different subsets of the data, and then combines the results using an ensemble method.
- It starts with each data point as its own cluster, and then iteratively merges the closest pairs of clusters until a stopping criterion is met.
- It uses a decision tree to recursively split the data into smaller clusters until a stopping criterion is met.

10) Which parameter in Fuzzy K-Means controls the fuzziness of the clustering?

- Number of clusters (K).
- Membership exponent (m).
- Number of iterations.
- Initial cluster centroids.