MACHINE LEARNING

What is the most appropriate no. of clusters for the data points represented by the following dendrogram:

Ans- b) 4

2. In which of the following cases will K-Means clustering fail to give good results?

Ans- d) 1,2 and 4

3. The most important part of is selecting the variables on which clustering is based.

Ans- b) selecting a clustering procedure

4. The most commonly used measure of similarity is the or its square

Ans- d) Manhattan distance

5. is a clustering procedure where all objects start out in one giant cluster. Clusters are formed by dividing this cluster into smaller and smaller clusters.

Ans-

6. Which of the following is required by K-means clustering?

Ans- d) All answers are correct

7. The goal of clustering is to-

Ans- a) Divide the data points into groups

8. Clustering is a-

Ans- b) Unsupervised learning

9. Which of the following clustering algorithms suffers from the problem of convergence at local optima?

Ans- a) K- Means clustering

10. Which version of the clustering algorithm is most sensitive to outliers?

Ans- a) K-means clustering algorithm

11. Which of the following is a bad characteristic of a dataset for clustering analysis

Ans- d) All of the above

12. For clustering, we do not require-

Ans- a) Labeled data

13. How is cluster analysis calculated?

Ans- Cluster analysis can be calculated by measuring the distance from each data point then link it to cluster closes to it.

14. How is cluster quality measured?

Ans- For cluster quality measurement we have to use the average silhouette coefficient value of all objects in data sets.

15. What is cluster analysis and its types?

Ans- Cluster analysis is a data analysis techniques that explores the naturally occurring groups within a data set.

Its types- HARD and SOFT clusters