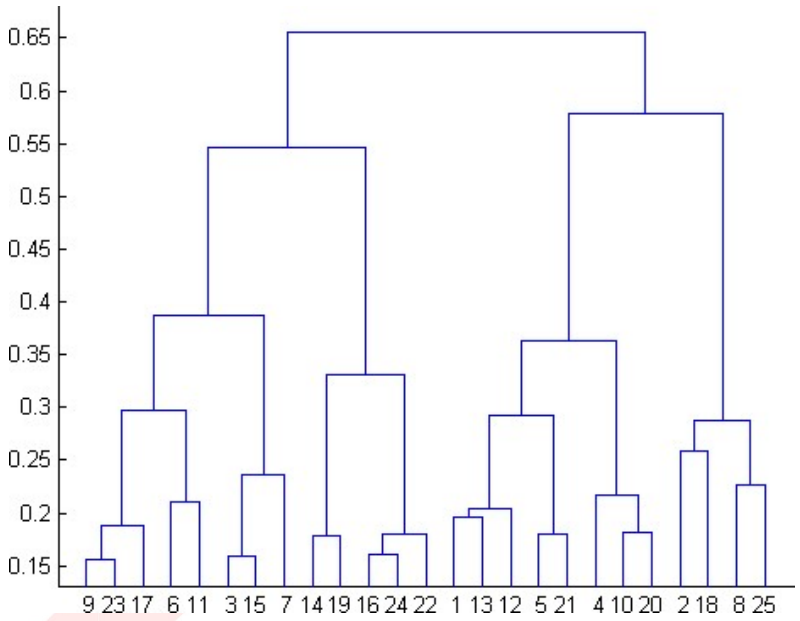
MACHINE LEARNING

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



Ans: 4

2. In which of the following cases will K-Means clustering fail to give good results?
   1. Data points with outliers
   2. Data points with different densities
   3. Data points with round shapes
   4. Data points with non-convex shapes

   Ans: 1, 2 and 4

3. The most important part of____is selecting the variables on which clustering is based.
   Ans: formulating the clustering problem

4. The most commonly used measure of similarity is the____or its square.
   Ans: Euclidean 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: Divisive clustering

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

   Ans: All answers are correct

MACHINE LEARNING

7. The goal of clustering is to-
   Ans: Divide the data points into groups

8. Clustering is a-
   Ans: Unsupervised learning

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

10. Which version of the clustering algorithm is most sensitive to outliers?
    Ans: K-means clustering algorithm

11. Which of the following is a bad characteristic of a dataset for clustering analysis-
    Ans: All of the above

12. For clustering, we do not require-
    Ans: Labeled data

13. How is cluster analysis calculated?

    Ans: Finding the data in same in one cluster and different from other cluster. It is find data points using shortest distance method.

14. How is cluster quality measured?
    Ans: As small the pure cluster, the score is high.

15. What is cluster analysis and its types?
    Ans: Cluster analysis used for unsupervised ML, which is the grouping method on which how closely data points associated.