

MACHINE LEARNING

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

Ans:- a) 2

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

Ans:- a) 1 and 2

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

Ans:- a) interpreting and profiling clusters

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

Ans:- a) 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:- b) Divisive clustering

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

Ans:- a) Defined distance metric

7. The goal of clustering is to

Ans:- d) All of the above

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:- c) Diverse clustering

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

Ans:- d) None

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

Ans:- a) Data points with outliers

12. For clustering, we do not require

Ans:- b) Unlabeled data