MACHINE LEARNING

Answer Keys

1. b) 4

   Explanation: The decision of the no. of clusters that can best depict different groups can be chosen by observing the dendrogram. The best choice of the no. of clusters is the no. of vertical lines in the dendrogram cut by a horizontal line that can transverse the maximum distance vertically without intersecting a cluster.

In the above example, the best choice of no. of clusters will be 4 as the red horizontal line in the dendrogram below covers maximum vertical distance AB.

2. d) 1, 2 and 4
3. d) formulating the clustering problem
4. a) Euclidean Distance
5. b) Divisive clustering
6. d) All answers are correct
7. a) Divide the data points into groups
8. b) Unsupervised learning
9. d) All of the above
10. a) K-means clustering algorithm
11. d) All of the above
12. a) Labeled data
13. The hierarchical cluster analysis follows three basic steps: 1) calculate the distances, 2) link the clusters, and 3) choose a solution by selecting the right number of clusters.
14. To measure a cluster's fitness within a clustering, we can compute the average silhouette coefficient value of all objects in the cluster. To measure the quality of a clustering, we can use the average silhouette coefficient value of all objects in the data set.
15. Cluster analysis is the task of grouping a set of data points in such a way that they can be characterized by their relevance to one another. These types are Centroid Clustering, Density Clustering Distribution Clustering, and Connectivity Clustering.