MACHINE LEARNING 1. ASSIGNMENT

Q1. What is the most appropriate no. of clusters for the data points represented by the following dendrogram:

ANS :- 4

Q2. In which of the following cases will K-Means clustering fail to give good results?

ANS :- 2 AND 4

Q3. The most important part of is selecting the variables on which clustering is based

AND :- interpreting and profiling clusters

Q4. The most commonly used measure of similarity is the or its square

ANS :- Euclidean distance

Q5. 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

Q6. Which of the following is required by K-means clustering?

ANS :- All answers are correct

Q7. The goal of clustering is to

ANS :- Predict the output values of input data points

Q8. Clustering is a

ANS :- Unsupervised learning

Q9. Which of the following clustering algorithms suffers from the problem of convergence at local optima ?

ANS :- K- Means clustering

Q10. Which version of the clustering algorithm is most sensitive to outliers?

ANS :- K-means clustering algorithm

Q11. Which of the following is a bad characteristic of a dataset for clustering analysis

ANS :- All of the above

Q12. . For clustering, we do not require

ANS :- Labeled data

Q13. How is cluster analysis calculated?

ANS :- Measuring the distance between each data point and its centroid, squaring this distance, and summing these squares across one cluster.

Q14. How is cluster quality measured?

ANS :- we can use the average silhouette coefficient value of all objects in the data set.

Q15. What is cluster analysis and its types?

ANS :- It is a type of clustering model closely related to statistics based on the modals of distribution