MACHINE LEARNING ASSIGNMENT -1

ANSWERS:

1. (b) 4
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. (a) K-Means clustering
10. (a) K-means clustering algorithm
11. (d) All of the above
12. (a) Labelled data
13. Cluster Analysis is calculated in 3 main steps:
    - Initialising the clustering model with defined clusters and fitting the dataset.
    - Checking the model/cluster quality using metrics such as silhouette_score.
    - Picking a clustering technique upon scaling the metric data and choosing the number of clusters with appropriate distances.
14. Cluster quality can be measured using any of the clustering accuracy metrics such as silhouette_score.
15. Cluster analysis is an Unsupervised Machine learning technique which tries to distributes the data into most appropriate number of smaller group/cluster based on similarities with other data points. It's not required any label of analysis.
    There are four types of Cluster Analysis:
    - Hierarchical Cluster Analysis: In this method, a cluster is made and then added to another cluster (the most similar or closest one) to form one single cluster.
    - Centroid Based Clustering: In this method K-means methods of clustering is used in this method, where k are the cluster and objects are assignment to the nearest cluster centres.'
    - Distribution Based Clustering: In this type of clustering model closely related to statistics based on the modals of distribution. It belongs to same distribution are put into a single cluster.
    - Density Based Clustering: In this type of clustering, cluster are defined by the areas of density that are higher than the remaining of the data set and in which most popular methos is DBSCAN.