Assignment-Machine Learning

Q1.(b)

Q2. (d)

Q3. (d)

Q4 .(a)

Q5.(d)

Q6.(d)

Q7.(a)

Q8.(b)

Q9.(a)

Q 10.(a)

Q11.(d)

Q12.(a)

Q13.) As follows:

Step 1 : Decide the number of clusters k
Step 2: Make an initial selection of K centroids
Step 3:Calculate the Euclidean distance between every data point and K centroid of the different cluster and Assign respective data points to that cluster which have minimum distance

Q14.) As follows:

We can measure the quality of clustering by using Dissimilarity/Similarity metric in most situations.

Dissimilarity metric:
The similarity between the clusters can be expressed in terms of distance functions represented by d(I, j), which is a Euclidian distance.

Cluster completeness:
Cluster completeness is high if the data objects are having similar characteristics and are of same category.

Ragbag:

When data objects of few categories cannot be merged with the other objects then the quality of those clusters can be measured by ragbag.

Q 15. Cluster analysis is a data analysis technique that explores the naturally occurring groups within a dataset known as clusters.
Cluster analysis is an unsupervised learning as it does not require data points to group into any predefined groups.

Types of clustering Analysis:

There are six types of clustering algorithm in Machine learning:

1.) Centroid based
2.) density based
3.) distribution based
4.) hierarchal based
5.) constraint based
6.) fuzzy clustering