Assignment – 1

Machine Learning

- Q1 – b) 4

- Q2 -- d) 1,2 and 4

- Q3 - d) formulating the clustering problem

- Q4 - a) Euclidean distance

- Q5 - b) Divisive clustering

- Q6 - d) All answers are correct

- Q7 - a) Divide the data points into groups

- Q8 - b) Unsupervised learning

- Q9 - a) K- Means clustering

- Q10 - a) K-means clustering algorithm

- Q11 - d) All of the above

- Q12 - b) Unlabeled data

- Q13 - How is cluster analysis calculated?

Ans – Cluster analysis is nothing but grouping of data having similar characteristics, it forms different groups but characteristics of data points among the group is different.

Steps of cluster analysis :-
1) Calculate the distance between the different datapoints or attributes
2) Grouping of clusters with minimal distance ie having similar characteristics
3) Find a solution by selecting the best number of clusters

- Q14 - How is cluster quality measured?

Ans –
1. The quality of a cluster depends on the similarity in the intra class attributes
2. And also on the dissimilarity between the inter class
3. It depend on the method which we use for similarity measurement and its implementation

4. It is also measured by its ability to find some or all of the hidden patterns

- Q 15 - What is cluster analysis and its types?

Ans :- Cluster analysis is nothing but grouping of unlabelled variables/attributes in given dataset. It is the basic step in data mining. It is a very common technique for statistical data analysis.

- Types of cluster Analysis :-
    1) Hierarchical Cluster Analysis
    2) Centroid-based Clustering
    3) Distribution based clustering
    4) Density-based Clustering