

## MACHINE LEARNING 2

1-D

2-D

3-A

4-A

5-B

6-B

7-A

8-D

9-A

10-D

11-D

12-K-Means clustering is an unsupervised learning algorithm which aims to partition  $n$  observations into  $k$  clusters in which each observation belongs to the cluster with the nearest centroid. The algorithm aims to minimize the squared Euclidean distances between the observation and the centroid of cluster to which it belongs. But sometime K-Means algorithm does not give best results. It is sensitive to outliers. An outlier is a point which is different from the rest of data points.

13-Advantages of K-means cluster.

High Performance -K-Means algorithm has linear time complexity and it can be used with large datasets conveniently. With unlabeled big data K-Means offers many insights and benefits as an unsupervised clustering algorithm.

- Easy to Use -K-Means is also easy to use.
- Unlabeled Data -This one is a general unsupervised machine learning algorithm that also applies to K-Means.If your data has no labels (class values or targets) or even column headers, K-Means will still successfully cluster your data.

- Result Interpretation -K-Means returns clusters which can be easily interpreted and even visualized. This simplicity makes it highly useful in some cases when you need a quick overview of the data segments.

14-K-Means is a non-deterministic algorithm. This means that a compiler cannot solve the problem in polynomial time and doesn't clearly know the next step. This is because some problems have a great degree of randomness to them. These algorithms usually have 2 steps — 1)Guessing step 2)Assignment step. On similar lines is the K-means algorithm. The K-Means algorithm divides the data space into K clusters such that the total variance of all data points with respect to the cluster mean is minimized. This means that running the algorithm several times on the same data, could give different results.