

# Machine Learning Assignment-2

---

1. A
2. D
3. A
4. A
5. B
6. B
7. A
8. D
9. A
10. C
11. D
12. Yes, K is sensitive to outliers. The k-means algorithm updates the cluster centers by taking the average of all data points that are closer to each cluster center. As a result, this will push your cluster center closer to the outlier.

E.g. Data set point are 1 2 3 7 8 80

80 is the outlier

K=2

C1=1 c2 =7

After iteration

C1 = 2 C2 = 31.67

As 80 data point which is outlier comes to cluster 2

Cluster 2 centroid changes to accommodate 80

Therefore K means sensitive to outliers.

13.

K-Means is a data partitioning algorithm which is among the most immediate choices as a clustering algorithm. Some reasons for the popularity of k-Means are:

1. Fast to Execute.
2. Online and Mini-Batch Implementations are also available thus requiring less memory.
3. Easy interpretation. The centroid of a cluster often gives a fair idea of the data present in the cluster
4. Results of k-Means can be used as starting points for other algorithms. It is often a practice to Use the centroids of k-Means as starting points for Gaussian Mixture Models.

14.

One of the significant drawbacks of K-Means is its non-deterministic nature. K-Means starts with a random set of data points as initial centroids. This random selection influences the quality of the resulting clusters. Besides, each run of the algorithm for the same dataset may yield a different output.