Machine learning

Objective Answers:

1. a
2. d
3. a
4. a
5. b
6. b
7. a
8. d
9. a
10. d
11. d

Subjective Answers:

12. Is K sensitive to outliers?

Yes, it is very sensitive to Outliers as centroids can be dragged by outliers, or outliers might get their own cluster instead of being ignored thus will make the model perform poorly.

13. **Why is K means better?**

K means is better due to following reasons:

1)It is relatively simple to implement and identify unknown groups of data from complex data sets. The results are presented in an easy and simple manner.

2)It is very easy to tune as number of clusters is only significant parameter choice.

3)It is suitable for large number of dataset and it's computed much faster than the smaller dataset. It can also produce higher clusters.

4) It doesn't take more time in classifying similar characteristics in data like hierarchical algorithms.

5) Compared to using other clustering methods, a k-means clustering technique is fast and efficient in terms of its computational cost .

14.Is K means a deterministic algorithm?

The basic k-means clustering is based on a non-deterministic algorithm. This means that running the algorithm several times on the same data, could give yield different outputs. The non-deterministic nature of K-Means is due to its random selection of data points as initial centroids. The key idea of the algorithm is to select data points which belong to dense regions and which are adequately separated in feature space as the initial centroids.