**MACHINE LEARNING**

1. Movie Recommendation systems are an example of

1. 2 Only

2. Sentiment Analysis is an example of:

d) 1, 2 and 4

3. Can decision trees be used for performing clustering?

a) True

4. Which of the following is the most appropriate strategy for data cleaning before performing clustering analysis, given less than desirable number of data points:

a) 1 only

5. What is the minimum no. of variables/ features required to perform clustering?

b) 1

6. For two runs of K-Mean clustering is it expected to get same clustering results?

b) No

7. Is it possible that Assignment of observations to clusters does not change between successive iterations in K-Means?

a) Yes

8. Which of the following can act as possible termination conditions in K-Means?

d) All of the above

9. Which of the following algorithms is most sensitive to outliers?

a) K-means clustering algorithm

10. How can Clustering (Unsupervised Learning) be used to improve the accuracy of Linear Regression model (Supervised Learning):

d) All of the above

11. What could be the possible reason(s) for producing two different dendrograms using agglomerative clustering algorithms for the same dataset?

d) All of the above

12. Is K sensitive to outliers?

Ans: - Yes K is sensitive to outliers, Because K-Means algorithm is about finding mean of clusters, the k-means algorithm updates the cluster centers by taking the average of all the data points that are closer to each cluster center. When all the points are packed nicely together, the average makes sense. However, when you have outliers, this can affect the average calculation of the whole cluster. As a result, this will push your cluster center closer to the outlier.

13. Why is K means better?

Ans: - K means it is the simplest algorithm which uses unsupervised learning method to solve non known Clustering issues. It's really work with large data sets, it is simple to implement

it also guarantees the convergence and can also warm start the position of centroids Easily adapts to new examples and can generalize to cluster of different shape and size such as elliptical clusters.

14. Is K means a deterministic algorithm?

Ans: - No, the k is not deterministic algorithm rather it is non - deterministic. it is due to it random selection of data points as initial centroids. The key idea of algorithm is to select data point which belongs to dense region and which adequately separated in feature space as the initial centroids.