1. Movie Recommendation systems are an example of:

a) 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?

K-Means clustering algorithm is most sensitive to outliers as it uses the mean of cluster data points to find the cluster center.

13. Why is K means better?

Guarantees convergence. Can warm-start the positions of centroids. Easily adapts to new examples. Generalizes to clusters of different shapes and sizes, such as elliptical clusters.

14. Is K means a deterministic algorithm

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.

$$\arg\min_{\mathbf{S}} \sum_{i=1}^{k} \sum_{\mathbf{x} \in S_i} \left\| \mathbf{x} - \boldsymbol{\mu}_i \right\|^2$$

Fig (1) Function to be minimized. S(i) are clusters.

However, the approach that compiler takes does not involve Multivariate Calculus as it seems. Rather, the approach taken is iterative. Now, like any deterministic algorithm it has 2 phases. Guessing phase: Randomly initializing k means in the data space(Mu(k)s). Now, all the data points X(i)s (1,m) are assigned to clusters in accordance to which cluster mean they are closer to. Mathematically, this step tries to minimize the within cluster variance. Hence, every point is now assigned a cluster. Next is the assignment step. All the cluster means (Mu(k)s)are now assigned to

the mean of the data points in the cluster. This step is repeated a couple of times. Refer to the image below.