**MACHINE LEARNING- WorkSheet2**

1. Movie Recommendation systems are an example of:
2. 2 Only
3. Sentiment Analysis is an example of:

d) 1, 2 and 4

1. Can decision trees be used for performing clustering?
2. True
3. Which of the following is the most appropriate strategy for data cleaning before performing clustering analysis, given less than desirable number of data points:
4. 1 only
5. What is the minimum no. of variables/ features required to perform clustering?
6. 1
7. For two runs of K-Mean clustering is it expected to get same clustering results?
8. No
9. Is it possible that Assignment of observations to clusters does not change between successive iterations in K-Means?
10. Yes
11. Which of the following can act as possible termination conditions in K-Means?
12. All of the above
13. Which of the following algorithms is most sensitive to outliers?
14. K-means clustering algorithm
15. How can Clustering (Unsupervised Learning) be used to improve the accuracy of Linear Regression model (Supervised Learning):
16. All of the above
17. 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

1. Is K sensitive to outliers?

Ans : K-Means is very sensitive to outliers as any extreme outlier values can significantly change the mean of the data set.

1. Why is K means better?

Ans: 1. It works well with large data set. 2. It is simple and easy to implement 3.It gives a good idea of the data structure

1. 14. Is K means a deterministic algorithm?

Ans: No K-Means is a not a deterministic algorithm because it selects random data points to be considered as centroids. This means we will get different results from the same dataset depending upon change in the centroids.