ASSIGNMENT – 2

Machine Learning

1. Movie Recommendation systems are an example of:
   Ans: Option a – clustering
2. Sentiment Analysis is an example of:
   Ans: Option b – regression, classification and reinforcement
3. Can decision trees be used for performing clustering?
   Ans: Option 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:
   Ans: Option a - Capping and flouring of variables
5. What is the minimum no. of variables/ features required to perform clustering?
   Ans: Option b – 1
6. For two runs of K-Mean clustering is it expected to get same clustering results?
   Ans: Option b - No
7. Is it possible that Assignment of observations to clusters does not change between successive iterations in K-Means?
   Ans: Option a - Yes
8. Which of the following can act as possible termination conditions in K-Means?
   Ans: Option d – all the above
9. Which of the following algorithms is most sensitive to outliers?
   Ans: Option a – K-means clustering algorithm
10. How can Clustering (Unsupervised Learning) be used to improve the accuracy of Linear Regression model (Supervised Learning):
    Ans: Option d – all the above
11. What could be the possible reason(s) for producing two different dendrograms using agglomerative clustering algorithms for the same dataset?
    Ans: Option d – All the above
12. Is K sensitive to outliers?
    Ans: Yes, as it uses the mean of cluster data points to find the cluster center.
13. Why is K means better?
    Ans: K-means has been around since the 1970s and fares better than other clustering algorithms like density-based, expectation-maximization. It is one of the most robust methods, especially for image segmentation and image annotation projects.
    K-means is very simple and easy to implement, gives good results.
14. Is K means a deterministic algorithm?
    Ans: No, k-means clustering is based on a non-deterministic algorithm.