

MACHINE LEARNING

1. Movie Recommendation systems are an example of:

Ans :- b) 1 and 2

2. Sentiment Analysis is an example of:

Ans :- d) 1, 2 and 4

3. Can decision trees be used for performing clustering?

Ans :- 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 :- i) Capping and flooring of variables

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

Ans :- b) 1

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

Ans :- b) No

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

Ans :- a) Yes

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

Ans :- d) All of the above

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

Ans :- a) K-means clustering algorithm

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

Ans :- 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?

Ans :- d) All of the above

12. Is K sensitive to outliers?

Ans :- K-Means Clustering is an unsupervised learning algorithm which aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest centroid.

The algorithm aims to minimize the squared Euclidean distances between the observation and the centroid of cluster to which it belongs.

But sometime K-Means algorithm does not give best results.

It is sensitive to outliers. An outlier is a point which is different from the rest of data points.

13. Why is K means better?

Ans :- Relatively simple to implement

Scales to large data sets

Guarantees convergence

Generalizes to clusters of different shapes and sizes, such as elliptical clusters.

14. Is K means a deterministic algorithm?

Ans :- Cluster analysis aims to group a set of objects/events in such a way that objects/events in the same group (i.e. a cluster) are more similar (in some sense or another) to each other than to those in other groups

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 different results.

However, to ensure consistent results, FCS Express performs k -means clustering using a deterministic method.