MACHINE LEARNING

Q1 to Q11 have only one correct answer. Choose the correct option to answer your question.

1. Movie Recommendation systems are an example of:
   i) Classification
   ii) Clustering
   iii) Regression
   Options:
   a) 2 Only
   b) 1 and 2
   c) 1 and 3
   d) 2 and 3

   Correct Option: a

2. Sentiment Analysis is an example of:
   i) Regression
   ii) Classification
   iii) Clustering
   iv) Reinforcement
   Options:
   a) 1 Only
   b) 1 and 2
   c) 1 and 3
   d) 1, 2 and 4

   Correct Option: d

3. Can decision trees be used for performing clustering?
   a) True
   b) False

   Correct Option: a

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:
   i) Capping and flooring of variables
   ii) Removal of outliers
   Options:
   a) 1 only
   b) 2 only
   c) 1 and 2
   d) None of the above

   Correct Option: a

5. What is the minimum no. of variables/ features required to perform clustering?
   a) 0
   b) 1
   c) 2
   d) 3

MACHINE LEARNING

Correct Option: b

6. For two runs of K-Mean clustering is it expected to get same clustering results?
   a) Yes
   b) No

   Correct Option: b

7. Is it possible that Assignment of observations to clusters does not change between successive iterations in K-Means?
   a) Yes
   b) No
   c) Can't say
   d) None of these

   Correct Option: a

8. Which of the following can act as possible termination conditions in K-Means?
   i) For a fixed number of iterations.
   ii) Assignment of observations to clusters does not change between iterations. Except for cases with a bad local minimum.
   iii) Centroids do not change between successive iterations.
   iv) Terminate when RSS falls below a threshold.
   Options:
   a) 1, 3 and 4
   b) 1, 2 and 3
   c) 1, 2 and 4
   d) All of the above

   Correct Option: d

9.Which of the following algorithms is most sensitive tooutliers?

   b) K-medians clustering algorithm
   c) K-modes clustering algorithm
   d) K-medoids clustering algorithm

   Correct Option: a

MACHINE LEARNING

10. HowcanClustering(UnsupervisedLearning)beusedtoimprovetheaccuracyofLinearRegression

i) Creating different models for different cluster groups.
ii) Creating an input feature for cluster ids as an ordinal variable.
iii) Creating an input feature for cluster centroids as a continuous variable.
iv) Creating an input feature for cluster size as a continuous variable.

Options:

a) 1 only
b) 2 only
c) 3 and 4
d) All of the above

Correct Option: d

11. Whatcouldbethepossiblereason(s)forproducingtwodifferentdendrogramsusingagglomerative

a) Proximity function used
b) of data points used
c) of variables used
d) All of the above

Correct Option: d

Q12 to Q14 aresubjective answers type questions, Answers them in their own words briefly

12.Is K sensitive to outliers?

Yes, k is sensitive to outliners

13.WhyisKmeansbetter?

K-means becomes a great solution for unsupervised learning which is used when we have unlabeled data, pre-clustering, reducing the space into disjoint smaller Sub-spaces where other clustering algorithms can be applied.

14.Is K means a deterministicalgorithm?

No, K means is a Non deterministic algorithm