



Score: 93%

No. of questions: 8

Correct answer: 7

Incorrect answer: 1

Show incorrect attempt only ☐**Question 1**

1 Mark

Let's say you are working on a dataset which have 950 observations and 25 features. You want to use K-Means clustering algorithm to segment observations. What would be the maximum possible number of clusters in this case?

A 950



B 25

C 1000

D 20

Correct Answer: A. 950

If we assign each observation as separate clusters, the maximum number of clusters that we can have is equal to the number of observations in the dataset. And hence in this case, the answer will be 950.

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



A Yes



B No

Correct Answer: A. Yes

When the K-Means algorithm has reached the local or global minima, it will not alter the assignment of data points to clusters for two successive iterations.

Question 3

3 marks

Can we get different results for different runs of K-Means clustering?

A Yes



B No

Correct Answer: A. Yes

K-Means clustering algorithm instead converges on local minima which might also correspond to the global minima in some cases but not always. Therefore, it's advised to run the K-Means algorithm multiple times before drawing inferences about the clusters. However, note that it's possible to receive the same clustering results from K-means by setting the same seed value for each run. But that is done by simply making the algorithm choose the set of the same random number for each run.

Question 4

1 Mark

What is the minimum number of variables/ features required to perform clustering?



Prev

TextChapter
(machine-

Next



Final Test

(machine-learning/final-

B 1



Module Test



H



C 2

D 3

Correct Answer: B. 1

At least a single variable is required to perform clustering analysis. Clustering analysis with a single variable can be visualized with the help of a histogram.

Question 5

2 Marks

In which of the following cases will K-Means clustering will fail to give good results?

- a. Data points with outliers.
- b. Data points with different densities.
- c. Data points with round shapes.
- d. Data points with non-convex shapes.

A a and b

B b and c

C b and d

D a, b, and d



E a, b, c, and d

Correct Answer: D. a, b, and d



K-Means clustering algorithm fails to give good results when the data contains outliers, the density spread of data points across the data space is different and the data points follow non-convex shapes.
(machine-learning/screen/6251)




(machine-learning/final-

test)

Question 6

3 marks

Feature scaling is an important step before applying K-Mean algorithm. What is the reason behind this?

- A In distance calculation it will give the same weights for all features 
- B You always get the same clusters even if you use or don't use feature scaling
- C Feature scaling is an important step in Manhattan distance but not in the case of Euclidean distance
- D None of the above


Correct Answer: A. In distance calculation it will give the same weights for all features

Feature scaling ensures that all the features get the same weight in the clustering analysis. Consider a scenario of clustering people based on their weight (KG) with range 55-110 and height (inches) with range 5.6 to 6.4. In this case, the clusters produced without scaling can be very misleading as the range of weight is much higher than that of height. Therefore, it's necessary to bring them to the same scale so that they have equal weight on the clustering result.

Question 7

1 Mark

Which of the following method is used for finding optimal of cluster in K-Mean algorithm?

- A Elbow method
- B Manhattan method
- C Euclidean method
- D All of the above 



Prev

Text Chapter

(machine-

learning/screen/6251)
Correct Answer: A. Elbow method

Next



Final Test

(machine-learning/final-

test)



Out of the given options, elbow method is used for finding the optimal number of clusters. The elbow method looks at the percentage of variance explained as a function of the number of clusters: One should choose a number of clusters so that adding another cluster doesn't give much better modeling of the data.

Question 8

2 Marks

Which of the following corresponds to the properties of clusters?

- A Points within the same cluster should be as different from each other as possible
- B Points within the same cluster should be as similar to each other as possible
- C Points in different clusters should be as different from each other as possible
- D Points in different clusters should be as similar to each other as possible
- E Both B and C

Correct Answer: E. Both B and C

According to the properties of clusters, points within the same cluster should be similar to each other and points from different clusters should be different from each other.

