

Google Discussions



Exam Professional Machine Learning Engineer All Questions

View all questions & answers for the Professional Machine Learning Engineer exam

Go to Exam

EXAM PROFESSIONAL MACHINE LEARNING ENGINEER TOPIC 1 QUESTION 63 DISCUSSIO..

Actual exam question from Google's Professional Machine Learning Engineer

Question #: 63

Topic #: 1

[\[All Professional Machine Learning Engineer Questions\]](#)

You work for a large retailer and have been asked to segment your customers by their purchasing habits. The purchase history of all customers has been uploaded to BigQuery. You suspect that there may be several distinct customer segments, however you are unsure of how many, and you don't yet understand the commonalities in their behavior. You want to find the most efficient solution. What should you do?

- A. Create a k-means clustering model using BigQuery ML. Allow BigQuery to automatically optimize the number of clusters.
- B. Create a new dataset in Dataprep that references your BigQuery table. Use Dataprep to identify similarities within each column.
- C. Use the Data Labeling Service to label each customer record in BigQuery. Train a model on your labeled data using AutoML Tables. Review the evaluation metrics to understand whether there is an underlying pattern in the data.
- D. Get a list of the customer segments from your company's Marketing team. Use the Data Labeling Service to label each customer record in BigQuery according to the list. Analyze the distribution of labels in your dataset using Data Studio.

Show Suggested Answer

by Vedjha at Dec. 7, 2022, 11:30 p.m.

Comments

Type your comment...

Submit

PhilipKoku 5 months ago

Selected Answer: A

A) K-means is ideal for unsupervised clustering

upvoted 1 times

MultiCloudIronMan 7 months, 1 week ago

Selected Answer: A

K-means algorithm is used for grouping/clustering data in unsupervised learning experiments.

upvoted 3 times

M25 1 year, 6 months ago

Selected Answer: A

Went with A

upvoted 4 times

CloudKida 1 year, 6 months ago

Selected Answer: A

when to use k-means : Your data may contain natural groupings or clusters of data. You may want to identify these groupings descriptively in order to make data-driven decisions. For example, a retailer may want to identify natural groupings of customers who have similar purchasing habits or locations. This process is known as customer segmentation.
<https://cloud.google.com/bigquery/docs/kmeans-tutorial>

upvoted 3 times

tavva_prudhvi 1 year, 7 months ago

A

This is the most efficient solution for segmenting customers based on their purchasing habits, as it utilizes BigQuery's built-in machine learning capabilities to identify distinct clusters of customers based on their purchasing behavior. By allowing BigQuery to automatically optimize the number of clusters, you can ensure that the model identifies the most appropriate number of segments based on the data, without having to manually select the number of clusters.

upvoted 2 times

ares81 1 year, 10 months ago

Selected Answer: A

I correct myself. It's A:

According to the documentation, if you omit the num_clusters option, BigQuery ML will choose a reasonable default based on the total number of rows in the training data.

upvoted 2 times

hiromi 1 year, 10 months ago

Selected Answer: A

A

<https://cloud.google.com/bigquery-ml/docs/kmeans-tutorial>

<https://towardsdatascience.com/how-to-use-k-means-clustering-in-bigquery-ml-to-understand-and-describe-your-data-better-c972c6f5733b>

upvoted 3 times

wish0035 1 year, 10 months ago

Selected Answer: A

ans: A, pretty sure.

C, D => discarded, very time consuming.

B => yes, you can identify similarities within each column, but when i read "you don't yet understand the commonalities in their behavior" i understand that this job would be difficult, because there could be many columns to analyze, and i don't think that this would be efficient.

A => BigQuery ML is compatible with kmeans clustering, it's easy and efficient to create, and i would automatically detect the number of clusters.

Also from the BigQuery ML docs: "K-means clustering for data segmentation; for example, identifying customer segments."
(Source: https://cloud.google.com/bigquery-ml/docs/introduction#supported_models_in)

upvoted 4 times

LearnSodas 1 year, 10 months ago

Selected Answer: A



K-means is a good unsupervised learning algorithm to segment a population based on similarity.

K-means is a good unsupervised learning algorithm to segment a population based on similarity

We can use K-means directly in BQ, so I think it's "the most efficient way"

Labeling is not a good option since we don't really know what makes a customer similar to another, and why Dataprep if we can use directly BQ?

   upvoted 3 times

  **ares81** 1 year, 10 months ago

It seems B, to me.



   upvoted 1 times

  **neochaotic** 1 year, 11 months ago



Selected Answer: B


It's B! Dataprep provides Data profiling functionalities

   upvoted 1 times

  **japoji** 1 year, 11 months ago

The question is about commonalities of clients by characteristics, not about characteristics by client. I mean with B you are looking for segments of the characteristics which define a client. But you need segments of clients defined by characteristics.

   upvoted 1 times

  **Vedjha** 1 year, 11 months ago

Will go for 'A' as it is easy to build a model in BQML where data is already present and optimization would be automatic in case of K-mean algorithm

   upvoted 4 times

Start Learning for free



Social Media

[Facebook](#) , [Twitter](#)

[YouTube](#) , [Reddit](#)

[Pinterest](#)

EXAMTOPICS

We are the biggest and most updated IT certification exam material website.

Using our own resources, we strive to strengthen the IT professionals community for free.



© 2024 ExamTopics

ExamTopics doesn't offer Real Microsoft Exam Questions. ExamTopics doesn't offer Real Amazon Exam Questions. ExamTopics Materials do not contain actual questions and answers from Cisco's Certification Exams.

CFA Institute does not endorse, promote or warrant the accuracy or quality of ExamTopics. CFA® and Chartered Financial Analyst® are registered trademarks owned by CFA Institute.