## EXAM PROFESSIONAL MACHINE LEARNING ENGINEER TOPIC 1 QUESTION 222 DISCUSSI...

Actual exam question from Google's Professional Machine Learning Engineer
Question #: 222
Topic #: 1
[All Professional Machine Learning Engineer Questions]

You developed a custom model by using Vertex AI to predict your application's user churn rate. You are using Vertex AI Model Monitoring for skew detection. The training data stored in BigQuery contains two sets of features - demographic and behavioral. You later discover that two separate models trained on each set perform better than the original model. You need to configure a new model monitoring pipeline that splits traffic among the two models. You want to use the same prediction-sampling-rate and monitoring-frequency for each model. You also want to minimize management effort. What should you do?

A. Keep the training dataset as is. Deploy the models to two separate endpoints, and submit two Vertex AI Model Monitoring jobs with appropriately selected feature-thresholds parameters.

B. Keep the training dataset as is. Deploy both models to the same endpoint and submit a Vertex AI Model Monitoring job with a monitoring-config-from-file parameter that accounts for the model IDs and feature selections.

C. Separate the training dataset into two tables based on demographic and behavioral features. Deploy the models to two separate endpoints, and submit two Vertex AI Model Monitoring jobs.

D. Separate the training dataset into two tables based on demographic and behavioral features. Deploy both models to the same endpoint, and submit a Vertex AI Model Monitoring job with a monitoring-config-from-file parameter that accounts for the model IDs and training datasets.

Show Suggested Answer

by pikachu007 at *Jan. 13, 2024, 7:15 a.m.*

## Comments

Type your comment...

Submit

**guilhermebutzke** Highly Voted 8 months, 1 week ago

Selected Answer: B

My answer: B

If you're using Vertex AI Model Monitoring for skew detection and your data is stored in BigQuery, it's not strictly necessary to separate the data into two tables. Vertex AI Model Monitoring can indeed analyze each feature individually to detect skew. So, isn't necessary to separate data.

Then, the `monitoring-config-from-file` parameter lets you specify unique configurations for each model, including ID and training data information. This ensures targeted monitoring and analysis and a unique monitoring job.

upvoted 6 times

**Dirtie_Sinkie** Most Recent 3 weeks, 4 days ago

Selected Answer: D

My vote is D, have to separate the training dataset

upvoted 1 times

**bfdf9c8** 2 months, 2 weeks ago

Selected Answer: D

The question mentions skew, yo need to configure the model monitoring with this in mind, so the better option is to separate in two diferent tables to user skew detection

upvoted 1 times

**fitri001** 6 months, 1 week ago

Selected Answer: B

Reduced Management Effort: You only need to deploy and monitor a single endpoint, minimizing complexity compared to managing two separate endpoints and monitoring jobs (Option A and C).
Efficient Data Usage: Maintaining the original training dataset simplifies data management and avoids the need to split it into separate tables (Option C and D).
Granular Monitoring: The monitoring-config-from-file parameter allows you to specify configurations for each model within the same monitoring job. You can define the model ID and the features to monitor for potential skew or drift for each model independently.

upvoted 3 times

**fitri001** 6 months, 1 week ago

A. Separate Endpoints and Monitoring Jobs: This approach requires managing two endpoints and monitoring jobs, increasing complexity.
C. Separate Training Data and Separate Endpoints: While it separates training data, it requires managing separate endpoints and monitoring jobs, similar to option A. Additionally, splitting the data might be unnecessary for monitoring purposes in this scenario.
D. Separate Training Data (Optional) and Single Endpoint: Splitting the data (optional) adds complexity, and while you can use a single endpoint, defining configurations for each model within the monitoring job is more efficient using the monitoring-config-from-file parameter (option B).

upvoted 2 times

**pinimichele01** 6 months, 2 weeks ago

Selected Answer: B

I don't understand why it is necessary to separate dataset when there is Vertex AI Monitoring

upvoted 1 times

**SausageMuffins** 4 months, 2 weeks ago

For training-skew detection, you require the training dataset. Hence, by splitting the original dataset into the two features, it would make management easier later on.

Correct me if I'm wrong, but you would have to update the monitoring job when you retrain the model to keep the monitoring job updated as well. Hence splitting it makes sense. Agreed that same endpoint would be easier to manage as opposed to two.

As a result, my answer is D.

upvoted 1 times

**shuvs** 6 months, 3 weeks ago

**Selected Answer: D**

Not B, as training on separate datasets is recommended.

upvoted 1 times

**pinimichele01** 6 months, 2 weeks ago

why? i don't understand sorry

upvoted 1 times

**Yan_X** 7 months, 2 weeks ago

**Selected Answer: D**

D

Separate data to 2 tables to make sure both models are trained with most relevant data.

upvoted 1 times

**b1a8fae** 9 months, 1 week ago

**Selected Answer: D**

D.
You need to split the training dataset for each respective model. Furthermore, you only need to control for 2 differences between models in monitoring-config-from-file: model ID, and training set. Feature selection should be the same in both models.

upvoted 1 times

**vaibavi** 8 months, 2 weeks ago

Why not B?

upvoted 2 times

**shadz10** 9 months, 1 week ago

**Selected Answer: D**

D - makes more sense two models to be trained seperately and more accuarately also submits a Vertex AI Model Monitoring job with a monitoring-config-from parameter which would enable the skew detecttion to work for each model

upvoted 2 times

**pikachu007** 9 months, 1 week ago

**Selected Answer: B**

A. Separate Endpoints: This approach involves more management overhead and potentially complicates monitoring configurations.
C. Separate Datasets: Splitting the dataset into two tables is unnecessary for model monitoring and could introduce data management complexities.
D. Separate Datasets, Same Endpoint: While feasible, this option lacks the flexibility of granular feature control provided by monitoring-config-from-file.

upvoted 4 times

## Social Media

Facebook , Twitter
YouTube , Reddit
Pinterest

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.