



- Expert Verified, Online, Free.

MENU



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 185 DISCUSSI...

Actual exam question from Google's Professional Machine Learning Engineer

Question #: 185

Topic #: 1

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

You have developed a BigQuery ML model that predicts customer churn, and deployed the model to Vertex AI Endpoints. You want to automate the retraining of your model by using minimal additional code when model feature values change. You also want to minimize the number of times that your model is retrained to reduce training costs. What should you do?



- A. 1. Enable request-response logging on Vertex AI Endpoints
2. Schedule a TensorFlow Data Validation job to monitor prediction drift
3. Execute model retraining if there is significant distance between the distributions
- B. 1. Enable request-response logging on Vertex AI Endpoints
2. Schedule a TensorFlow Data Validation job to monitor training/serving skew
3. Execute model retraining if there is significant distance between the distributions
- C. 1. Create a Vertex AI Model Monitoring job configured to monitor prediction drift
2. Configure alert monitoring to publish a message to a Pub/Sub queue when a monitoring alert is detected
3. Use a Cloud Function to monitor the Pub/Sub queue, and trigger retraining in BigQuery
- D. 1. Create a Vertex AI Model Monitoring job configured to monitor training/serving skew
2. Configure alert monitoring to publish a message to a Pub/Sub queue when a monitoring alert is detected
3. Use a Cloud Function to monitor the Pub/Sub queue, and trigger retraining in BigQuery

Show Suggested Answer

Comments

Type your comment...

Submit

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

Selected Answer: D

My answer: D

Given the emphasis on "model feature values change" in the question, the most suitable option would be D.

Although option C involves monitoring prediction drift, which may indirectly capture changes in feature values, option D directly addresses the need to monitor training/serving skew. By detecting discrepancies between the training and serving data distributions, option D is more aligned with the requirement to automate retraining when model feature values change. Therefore, option D is the most suitable choice in this context.

   upvoted 6 times

  **bobjr** Most Recent  4 months, 2 weeks ago

Selected Answer: C

Skew should be detected at the beginning of the productionalisation of the model -> skew test the training data Vs the real data -> a skew indicates you trained in a dataset that is not alined with your data that you have in input

Drift is used when the model works well at the beginning, but the world change and the data input changes -> drift is more long term

here it is a drift issue

   upvoted 3 times

  **Prakzz** 3 months, 3 weeks ago

Agreed

   upvoted 1 times

  **Shno** 5 months, 3 weeks ago

if the model training is done through bigquery ML, we don't have access to the training data after export, so I don't understand how training/serving skew can be applied. Can someone who is voting in favour of D clarify?

   upvoted 1 times

  **gscharly** 6 months, 1 week ago

Selected Answer: D

I go with D

   upvoted 1 times

  **pinimichele01** 6 months, 2 weeks ago

Selected Answer: D

It's D

   upvoted 1 times

  **pinimichele01** 6 months ago




see guilhermebutzke

   upvoted 1 times

  **CHARLIE2108** 8 months, 1 week ago

Selected Answer: D

changed my mind it's D

   upvoted 3 times

  **CHARLIE2108** 8 months, 2 weeks ago

Selected Answer: C

I go with C but D is pretty similar.

C -> Prediction drift (When the overall distribution of predictions changes significantly between training and serving data).

D -> Training/serving skew (When the distribution of specific features between training and serving data differs significantly).

   upvoted 2 times

👍 🔄 🚩 upvoted 2 times

🗨️ 👤 **CHARLIE2108** 8 months, 1 week ago

It's D

👍 🔄 🚩 upvoted 1 times

🗨️ 👤 **ddogg** 8 months, 3 weeks ago

Selected Answer: C

Option C:

This option directly addresses your requirements:

Vertex AI Model Monitoring: It allows efficient monitoring of prediction drift through metrics like Mean Squared Error or AUC-ROC.

Pub/Sub alerts: Alert triggers notification upon significant drift, minimizing unnecessary retraining.

Cloud Function: It reacts to Pub/Sub messages and triggers retraining in BigQuery using minimal additional code.

👍 🔄 🚩 upvoted 2 times

🗨️ 👤 **b1a8fae** 9 months, 1 week ago

Selected Answer: C

After reconsidering, I think it is C:

- No need to use TF to enable model monitoring as stated here: <https://cloud.google.com/vertex-ai/docs/model-monitoring/using-model-monitoring>

(even if it uses it under the hood: <https://cloud.google.com/vertex-ai/docs/model-monitoring/overview#calculating-skew-and-drift>)

- The problem speaks about alerting of model feature changes, which happens over time, and uses a baseline of the historical production data -> prediction skew. (if the problem specified that it changes compared to training data, then it would be training-skew) (https://cloud.google.com/vertex-ai/docs/model-monitoring/monitor-explainable-ai#feature_attribution_training-serving_skew_and_prediction_drift)

👍 🔄 🚩 upvoted 2 times

🗨️ 👤 **b1a8fae** 9 months, 1 week ago

Selected Answer: D

I would avoid using TensorFlow validation to minimize code written. That leaves us with options C and D. Now, since it is the values of the features that we want to flag and not the value of the predictions, this sounds more like training-serving skew situation than prediction drift. Hence, I would go for D.

👍 🔄 🚩 upvoted 4 times

🗨️ 👤 **BlehMaks** 9 months, 1 week ago

Selected Answer: D

i've changed my mind) it's D

<https://www.evidentlyai.com/blog/machine-learning-monitoring-data-and-concept-drift>

👍 🔄 🚩 upvoted 1 times

🗨️ 👤 **BlehMaks** 9 months, 1 week ago

Selected Answer: D

we might need to retrain if the feature data distribution in the production and training are significantly different(training-serving skew). Prediction drift occurs when feature data distribution in production changes significantly over time. Should we retrain our model every time when we meet prediction drift? I dont think so, better to analyze why this drift happens.

<https://cloud.google.com/vertex-ai/docs/model-monitoring/overview#considerations>

👍 🔄 🚩 upvoted 1 times

🗨️ 👤 **36bdc1e** 9 months, 1 week ago

C

The best option for automating the retraining of your model by using minimal additional code when model feature values change, and minimizing the number of times that your model is retrained to reduce training costs, is to create a Vertex AI Model Monitoring job configured to monitor prediction drift, configure alert monitoring to publish a message to a Pub/Sub queue when a monitoring alert is detected, and use a Cloud Function to monitor the Pub/Sub queue, and trigger retraining in BigQuery. This option allows you to leverage the power and simplicity of Vertex AI, Pub/Sub, and Cloud Functions to monitor your model performance and retrain your model when needed. Vertex AI is a unified platform for building and deploying machine learning solutions on Google Cloud.

👍 🔄 🚩 upvoted 2 times



🗨️ 👤 **pikachu007** 9 months, 2 weeks ago

Selected Answer: C

A and B: TensorFlow Data Validation jobs require more setup and maintenance, and they might not integrate as seamlessly with Vertex AI Endpoints for automated retraining.

D: Monitoring training-serving skew focuses on differences between training and deployment environments, which might not directly address feature value changes.

👍 🔄 🚩 upvoted 1 times

  **vale_76_na_xxx** 9 months, 2 weeks ago

I go with : C.

1. Create a Vertex AI Model Monitoring job configured to monitor prediction drift -> if the model is already in production we have to consider Prediction drift
2. Configure alert monitoring to publish a message to a Pub/Sub queue when a monitoring alert is detected -> set Pub/Sub notification channels.
3. Use a Cloud Function to monitor the Pub/Sub queue, and trigger retraining in BigQuery -> to import new data in BQ

   upvoted 2 times

Start Learning for free



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.



© 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.