



- 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 138 DISCUSSI...

Actual exam question from Google's Professional Machine Learning Engineer

Question #: 138

Topic #: 1

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

You work for a company that manages a ticketing platform for a large chain of cinemas. Customers use a mobile app to search for movies they're interested in and purchase tickets in the app. Ticket purchase requests are sent to Pub/Sub and are processed with a Dataflow streaming pipeline configured to conduct the following steps:

1. Check for availability of the movie tickets at the selected cinema.
2. Assign the ticket price and accept payment.
3. Reserve the tickets at the selected cinema.
4. Send successful purchases to your database.

Each step in this process has low latency requirements (less than 50 milliseconds). You have developed a logistic regression model with BigQuery ML that predicts whether offering a promo code for free popcorn increases the chance of a ticket purchase, and this prediction should be added to the ticket purchase process. You want to identify the simplest way to deploy this model to production while adding minimal latency. What should you do?

- A. Run batch inference with BigQuery ML every five minutes on each new set of tickets issued.
- B. Export your model in TensorFlow format, and add a `tfx_bsl.public.beam.RunInference` step to the Dataflow pipeline.
- C. Export your model in TensorFlow format, deploy it on Vertex AI, and query the prediction endpoint from your streaming pipeline.
- D. Convert your model with TensorFlow Lite (TFLite), and add it to the mobile app so that the promo code and the incoming request arrive together in Pub/Sub.

Share Question & Answer

by  mil_spyro at Dec. 13, 2022, 6:43 p.m.

Comments

Type your comment...

[Submit](#)

  **behzadsw** [Highly Voted](#)  1 year, 10 months ago

[Selected Answer: D](#)

D as you want to do the prediction before the purchase

   upvoted 9 times

  **hiromi** [Highly Voted](#)  1 year, 10 months ago

[Selected Answer: D](#)

D

- <https://www.tensorflow.org/lite/guide>

   upvoted 5 times

  **bobjr** [Most Recent](#)  5 months ago

[Selected Answer: C](#)

D makes no senses -> if the prediction is made on the phone, why send it to the server ?

C is the best choice because it splits the responsibility, and use best practices and scalable tools

   upvoted 2 times

  **omermahgoub** 6 months, 3 weeks ago

[Selected Answer: B](#)

B. Export your model in TensorFlow format, and add a `tfx_bsl.public.beam.RunInference` step to the Dataflow pipeline.

Here's why this approach offers minimal latency:

In-Pipeline Prediction: The model is integrated directly into the Dataflow pipeline, enabling real-time predictions for each ticket purchase request without external calls.

Dataflow Integration: `tfx_bsl.public.beam.RunInference` is a Beam utility specifically designed for integrating TensorFlow models into Dataflow pipelines, ensuring efficient execution.

   upvoted 1 times

  **Yan_X** 7 months, 3 weeks ago

[Selected Answer: B](#)

B

For D - How can we assume the model does be feasible to convert to Mobile app?

   upvoted 1 times

  **Krish6488** 12 months ago

[Selected Answer: D](#)

Question looks ambiguous! However considering some keywords like low latency and more importantly ML usage for maximising the ticket purchase requests using the promo code means that model embedded to the device looks more appropriate, however there are a lot of downsides to it like model management and upgrades but that does not seem to be the consideration here anyway. Just looking at low latency and ML to maximise the ticket sales, I will go with D as thats much simpler to implement

   upvoted 2 times

  **andresvelasco** 1 year, 1 month ago

the whole question does not make too much sense to me.

first of all, it seems that the Dataflow streaming job would "accept payment" meaning it communicates with payment gateways and back to the user, which does not sound right to do in Dataflow.

the model "predicts whether offering a promo code for free popcorn increases the chance of a ticket purchase" is necessarily executed before processing payment, so D seems the best.

Awkward

   upvoted 2 times

  **M25** 1 year, 5 months ago

👤 **M23** 1 year, 5 months ago

Selected Answer: D

Went with D

👍 🔄 🚩 upvoted 1 times

👤 **TNT87** 1 year, 6 months ago

Answer D

👍 🔄 🚩 upvoted 4 times

👤 **TNT87** 1 year, 7 months ago

Selected Answer: B

is the simplest way to deploy the logistic regression model to production with minimal latency. Exporting the model in TensorFlow format and adding a `tfx_bsl.public.beam.RunInference` step to the existing Dataflow pipeline enables the model to be integrated directly into the ticket purchase process.

👍 🔄 🚩 upvoted 1 times

👤 **tavva_prudhvi** 1 year, 3 months ago

would also not be suitable because adding a `tfx_bsl.public.beam.RunInference` step to the Dataflow pipeline would still require the model to be executed within the same pipeline, potentially introducing additional latency and computational overhead.

👍 🔄 🚩 upvoted 1 times

👤 **TNT87** 1 year, 8 months ago

Selected Answer: C

Option C is the best solution. Since the entire process has low latency requirements, running batch inference every five minutes is not a suitable option. Option B requires a TensorFlow model format, which may not be available since the model is created using BigQuery ML. Option D is not recommended because it requires deploying the model to the mobile app, which may not be feasible or desired. Deploying the model on Vertex AI and querying the prediction endpoint from the streaming pipeline adds minimal latency and is the simplest solution.

👍 🔄 🚩 upvoted 1 times

👤 **TNT87** 1 year, 7 months ago

Answer B

👍 🔄 🚩 upvoted 1 times

👤 **TNT87** 1 year, 8 months ago

Aiiii between B and C

👍 🔄 🚩 upvoted 1 times

👤 **Scipione_** 1 year, 8 months ago

Selected Answer: D

I perfectly agree with behzadsw.

You send a Pub/Sub request when you already want to buy, you must add the coupon before this process.

👍 🔄 🚩 upvoted 3 times

👤 **John_Pongthorn** 1 year, 8 months ago

Selected Answer: B

B (is it possible) along with What I get from this question.

1. this prediction should be added to the ticket purchase process (it mean that it have to be include in rocessed with a Dataflow streaming pipeline
2. Each step in this process has low latency requirements (less than 50 milliseconds) , which signifies that whatever you will process in dataflow, there are no latency requirements issues

👍 🔄 🚩 upvoted 3 times

👤 **John_Pongthorn** 1 year, 8 months ago

https://www.tensorflow.org/tfx/tfx_bsl/api_docs/python/tfx_bsl/public/beam/RunInference

👍 🔄 🚩 upvoted 1 times

👤 **TNT87** 1 year, 10 months ago

Answer D

<https://www.tensorflow.org/lite/guide>

👍 🔄 🚩 upvoted 1 times

👤 **TNT87** 1 year, 8 months ago

Nope answer is B



👍 🔄 🚩 upvoted 1 times

👤 **mil_spyro** 1 year, 10 months ago

Selected Answer: C

By deploying your model on Vertex AI, you can quickly and easily add the prediction step to your streaming pipeline, without the need to add additional infrastructure or manage model deployment and management.

   upvoted 1 times

  **mymy9418** 1 year, 10 months ago


but maybe D is faster?

   upvoted 1 times

  **mil_spyro** 1 year, 10 months ago

Hey I think you're right, predictions on the device itself, which avoids the need to send requests over the network. Should be D

   upvoted 4 times

  **behzadsw** 1 year, 10 months ago

You also want to do the prediction before the purchase right? so D is correct

   upvoted 3 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.