EXAM PROFESSIONAL MACHINE LEARNING ENGINEER TOPIC 1 QUESTION 211 DISCUSSI...

Actual exam question from Google's Professional Machine Learning Engineer
Question #: 211
Topic #: 1
[All Professional Machine Learning Engineer Questions]

You need to develop a custom TensorFlow model that will be used for online predictions. The training data is stored in BigQuery You need to apply instance-level data transformations to the data for model training and serving. You want to use the same preprocessing routine during model training and serving. How should you configure the preprocessing routine?

A. Create a BigQuery script to preprocess the data, and write the result to another BigQuery table.

B. Create a pipeline in Vertex AI Pipelines to read the data from BigQuery and preprocess it using a custom preprocessing component.

C. Create a preprocessing function that reads and transforms the data from BigQuery. Create a Vertex AI custom prediction routine that calls the preprocessing function at serving time.

D. Create an Apache Beam pipeline to read the data from BigQuery and preprocess it by using TensorFlow Transform and Dataflow.

Show Suggested Answer

by pikachu007 at *Jan. 13, 2024, 5:52 a.m.*

Comments

Type your comment...

Submit

guilhermebutzke Highly Voted 8 months, 3 weeks ago

Selected Answer: D

My answer: D

According to this documentation, it is very clear that using BigQuery is not a good approach for online prediction at the instance level. That's because we won't use the same code for both training and prediction serving. In the same documentation, the final table on the page recommends using Dataflow with TensorFlow Transform for instance-level data transformation.

https://www.tensorflow.org/tfx/guide/tft_bestpractices

upvoted 7 times

pinimichele01 Most Recent 6 months, 3 weeks ago

Selected Answer: D

https://www.tensorflow.org/tfx/guide/tft_bestpractices#preprocessing_options_summary

upvoted 1 times

Yan_X 8 months, 4 weeks ago

Selected Answer: D

D - Apache Beam + tf.transform or Dataflow.
https://notebook.community/GoogleCloudPlatform/training-data-analyst/courses/machine_learning/deepdive/04_advanced_preprocessing/a_dataflow

upvoted 2 times

BlehMaks 9 months, 3 weeks ago

Selected Answer: A

the simplest way

upvoted 1 times

shadz10 9 months, 3 weeks ago

Selected Answer: D

D- Vertex AI isn't designed for instance-level data transformations

upvoted 1 times

shadz10 9 months, 3 weeks ago

This document also provides an overview of TensorFlow Transform (tf.Transform), a library for TensorFlow that lets you define both instance-level and full-pass data transformation through data preprocessing pipelines. These pipelines are executed with Apache Beam, and they create artifacts that let you apply the same transformations during prediction as when the model is served.
https://www.tensorflow.org/tfx/guide/tft_bestpractices

upvoted 1 times

shadz10 9 months, 3 weeks ago

D- Vertex AI isn't designed for instance-level data transformations

upvoted 3 times

pikachu007 9 months, 3 weeks ago

Selected Answer: C

Addressing limitations of other options:

A. Data validation: While essential, it doesn't guarantee consistency if the preprocessing logic itself differs between pipeline and endpoint.
C. Sharing code with end users: This shifts the preprocessing burden to end users, potentially leading to inconsistencies and errors, and isn't feasible for real-time inference.
D. Batching real-time requests: This introduces latency and might not align with real-time requirements, as users expect immediate responses.

upvoted 1 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.