EXAMTOPICS

- 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 228 DISCUSSI...**

Actual exam question from Google's Professional Machine Learning Engineer

Question #: 228

Topic #: 1

[All Professional Machine Learning Engineer Questions]

You are investigating the root cause of a misclassification error made by one of your models. You used Vertex AI Pipelines to train and deploy the model. The pipeline reads data from BigQuery. creates a copy of the data in Cloud Storage in TFRecord format, trains the model in Vertex AI Training on that copy, and deploys the model to a Vertex AI endpoint. You have identified the specific version of that model that misclassified, and you need to recover the data this model was trained on. How should you find that copy of the data?

A. Use Vertex AI Feature Store. Modify the pipeline to use the feature store, and ensure that all training data is stored in it. Search the feature store for the data used for the training.

B. Use the lineage feature of Vertex AI Metadata to find the model artifact. Determine the version of the model and identify the step that creates the data copy and search in the metadata for its location.

C. Use the logging features in the Vertex AI endpoint to determine the timestamp of the model's deployment. Find the pipeline run at that timestamp. Identify the step that creates the data copy, and search in the logs for its location.

D. Find the job ID in Vertex AI Training corresponding to the training for the model. Search in the logs of that job for the data used for the training.

Show Suggested Answer

by 👤 pikachu007 at *Jan. 13, 2024, 7:42 a.m.*

Comments

Comments

Type your comment...

Submit

⊟ 👤 **fitri001** 6 months, 3 weeks ago

Selected Answer: B

Vertex AI Metadata Lineage: This feature tracks the relationships between pipeline components and the artifacts they produce. By identifying the model version's lineage, you can pinpoint the specific pipeline run that generated it.
Data Copy Step: Within the pipeline run, locate the step responsible for creating the data copy in TFRecord format for training.
Metadata Search: Vertex AI Metadata likely stores information about the data copy's location in Cloud Storage, allowing you to access it.

👍 ↩ 🏳 upvoted 3 times

⊟ 👤 **fitri001** 6 months, 3 weeks ago

A. Feature Store: Feature Store is designed for managing feature engineering and serving preprocessed features, not necessarily raw training data. While it could be a good practice for future pipelines, it wouldn't help recover historical data.
C. Endpoint Logs: Endpoint logs primarily focus on model deployment details and might not provide information about the specific training data used for a particular version.
D. Training Job Logs: Training job logs might contain references to the data used, but they might not be as detailed or structured as Vertex AI Metadata lineage, making it harder to pinpoint the exact data copy location.

👍 ↩ 🏳 upvoted 1 times

⊟ 👤 **pinimichele01** 7 months ago

Selected Answer: B

agree with shadz10 and pikachu007

👍 ↩ 🏳 upvoted 1 times

⊟ 👤 **shadz10** 9 months, 3 weeks ago

Selected Answer: B

https://cloud.google.com/vertex-ai/docs/ml-metadata/introduction

👍 ↩ 🏳 upvoted 2 times

⊟ 👤 **pikachu007** 9 months, 3 weeks ago

Selected Answer: B

A. Feature Store: While useful for managing features, it might not store complete training datasets, and modifying the pipeline would not help recover historical data.
C. Endpoint Logs and Pipeline Run: This approach involves more manual searching and might be less precise for identifying the exact data copy.
D. Training Job Logs: Training job logs might not reliably contain complete data paths or might be purged after a certain period.

👍 ↩ 🏳 upvoted 3 times

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.