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 230 DISCUSSI...**

Actual exam question from Google's Professional Machine Learning Engineer

Question #: 230

Topic #: 1

[All Professional Machine Learning Engineer Questions]

You are training models in Vertex AI by using data that spans across multiple Google Cloud projects. You need to find, track, and compare the performance of the different versions of your models. Which Google Cloud services should you include in your ML workflow?

- A. Dataplex, Vertex AI Feature Store, and Vertex AI TensorBoard
- B. Vertex AI Pipelines, Vertex AI Feature Store, and Vertex AI Experiments
- C. Dataplex, Vertex AI Experiments, and Vertex AI ML Metadata
- D. Vertex AI Pipelines, Vertex AI Experiments, and Vertex AI Metadata

Show Suggested Answer

by 👤 b1a8fae at *Jan. 17, 2024, 2:57 p.m.*

Comments

Type your comment...

Submit

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

Why not the others?

A. Dataplex & Vertex AI Feature Store: While Dataplex can manage data across projects, it's not directly tied to model versioning and comparison. Feature Store focuses on feature engineering, not model version management.

B. Vertex AI Feature Store & Vertex AI TensorBoard: Similar to option A, Feature Store isn't directly involved in model version tracking, and TensorBoard is primarily for visualizing training data and metrics, not model version comparison across projects.

C. Dataplex & Vertex AI ML Metadata: Dataplex, as mentioned earlier, doesn't directly address model version comparison. While ML Metadata tracks lineage, it might not have the experiment management features of Vertex AI Experiments.

👍 ↩ 🚩 upvoted 3 times

👤 **fitri001** 6 months, 3 weeks ago

Vertex AI Pipelines (Optional): While optional, pipelines can automate your training workflow, including data access from BigQuery tables in different projects. It helps orchestrate the training process across projects.

Vertex AI Experiments: This service is crucial for tracking and comparing the performance of different model versions. It allows you to:

Run multiple training experiments with different configurations.

Track experiment metrics like accuracy, precision, recall, etc.

Compare the performance of different model versions trained in various projects.

Vertex AI Metadata: This service provides a centralized view of your ML workflow, including model lineage and versioning. It's particularly helpful in your scenario because:

It tracks the origin and relationships between models, including the specific data used for training, regardless of the project.

You can see how different model versions (potentially trained across projects) relate to each other and the data they were trained on.

👍 ↩ 🚩 upvoted 2 times

👤 **Aastha_Vashist** 7 months, 2 weeks ago

went with D

👍 ↩ 🚩 upvoted 2 times

👤 **Yan_X** 8 months ago

I would go with option D.

No Vertex AI pipeline no orchestration. So rule out A and C.

Vertex AI Metadata is for 'spans across multiple Google Cloud projects' data used by the model.

👍 ↩ 🚩 upvoted 2 times

👤 **Carlose2108** 8 months, 1 week ago

Why not Option D?

👍 ↩ 🚩 upvoted 1 times

👤 **guilhermebutzke** 8 months, 2 weeks ago

My Answer: B

Vertex AI Pipelines: to create, deploy, and manage ML pipelines, which are essential for orchestrating your ML workflow, especially when dealing with data spanning multiple projects.

Vertex AI Feature Store: It's crucial for managing feature data across different projects.

Vertex AI Experiments: track and compare the performance of different versions of your models, enabling you to experiment

Why not the other:

Dataplex: not specifically tailored for managing ML workflows or model training.

Vertex AI ML metadata: not sufficient on its own to cover all aspects of managing the ML workflow across multiple projects.

Vertex AI TensorBoard: not specifically designed for managing the end-to-end ML workflow or tracking model versions across multiple projects.

👍 ↩ 🚩 upvoted 2 times

👤 **tavva_prudhvi** 6 months ago

I feel, Vertex AI Feature Store is valuable for managing and serving features for ML models, but it doesn't address the need for tracking experiments and managing metadata, right?

👍 ↩ 🚩 upvoted 1 times

👤 **SKDE** 9 months, 1 week ago

Dataplex works well with the data across projects and even on-prem, but doesn't work well with the ML related data like tracking and performance. So options A and C are considered wrong.

Metadata is to store metadata. So it is not required while we consider to compare the model performance. So option D is

Metadata is to store metadata. So it is not required while we consider to compare the model performance. So option D is wrong.

On the other hand Feature store brings meaningful data for comparing the Models performance based on feature data. So Option B is correct

👍 ↩ 🚩 upvoted 1 times

⊟ 👤 **b1a8fae** 9 months, 3 weeks ago

Selected Answer: C

I go with C.
Dataplex to centralize different Google projects.
Vertex AI experiments + ML Metadata to track experiment lineage, parameter usage etc and compare models.

👍 ↩ 🚩 upvoted 3 times

⊟ 👤 **daidai75** 9 months, 2 weeks ago

How about Option D? the Pipeline can also do the cross project data processing.

👍 ↩ 🚩 upvoted 5 times