EXAM PROFESSIONAL MACHINE LEARNING ENGINEER TOPIC 1 QUESTION 192 DISCUSSI...

Actual exam question from Google's Professional Machine Learning Engineer
Question #: 192
Topic #: 1
[All Professional Machine Learning Engineer Questions]

Your team has a model deployed to a Vertex AI endpoint. You have created a Vertex AI pipeline that automates the model training process and is triggered by a Cloud Function. You need to prioritize keeping the model up-to-date, but also minimize retraining costs. How should you configure retraining?

A. Configure Pub/Sub to call the Cloud Function when a sufficient amount of new data becomes available

B. Configure a Cloud Scheduler job that calls the Cloud Function at a predetermined frequency that fits your team's budget

C. Enable model monitoring on the Vertex AI endpoint. Configure Pub/Sub to call the Cloud Function when anomalies are detected

D. Enable model monitoring on the Vertex AI endpoint. Configure Pub/Sub to call the Cloud Function when feature drift is detected

Show Suggested Answer

by winston9 at *Jan. 11, 2024, 1:36 p.m.*

Comments

fitri001 6 months, 2 weeks ago

Selected Answer: D

Data-driven Retraining: Monitoring for feature drift identifies significant changes in the underlying data distribution used to train the model. Retraining based on drift detection ensures the model stays relevant to evolving data patterns, prioritizing model accuracy.
Reduced Cost: Triggering retraining only when drift is detected avoids unnecessary training runs, minimizing costs associated with Vertex AI training jobs.

upvoted 2 times

fitri001 6 months, 2 weeks ago

A. New Data Availability: While new data is important, it might not always necessitate retraining, especially if the new data aligns with existing patterns.
B. Predetermined Frequency: Fixed scheduling can lead to either under-training (data evolves faster than the schedule) or over-training (drift happens slower than the schedule), potentially wasting resources.
C. Anomaly Detection: Anomalies might not directly indicate feature drift, and retraining based solely on anomalies could introduce noise into the model.

upvoted 1 times

ddogg 9 months, 1 week ago

Selected Answer: D

D) Makes the most sense and scales

upvoted 2 times

b1a8fae 9 months, 3 weeks ago

Selected Answer: D

Keep the model up to date -> monitoring drift (distribution of production data doesnt change wildly). Only rerun training when necessary.

upvoted 1 times

pikachu007 9 months, 3 weeks ago

Selected Answer: D

It proactively triggers retraining when feature drift is detected, ensuring the model adapts to changing data patterns and maintains accuracy.

upvoted 1 times

winston9 9 months, 4 weeks ago

Selected Answer: D

feature drifting detecting to trigger retraining

upvoted 1 times