EXAM PROFESSIONAL MACHINE LEARNING ENGINEER TOPIC 1 QUESTION 245 DISCUSSI...

Actual exam question from Google's Professional Machine Learning Engineer
Question #: 245
Topic #: 1
[All Professional Machine Learning Engineer Questions]

You are developing an ML model that predicts the cost of used automobiles based on data such as location, condition, model type, color, and engine/battery efficiency. The data is updated every night. Car dealerships will use the model to determine appropriate car prices. You created a Vertex AI pipeline that reads the data splits the data into training/evaluation/test sets performs feature engineering trains the model by using the training dataset and validates the model by using the evaluation dataset. You need to configure a retraining workflow that minimizes cost. What should you do?

A. Compare the training and evaluation losses of the current run. If the losses are similar, deploy the model to a Vertex AI endpoint. Configure a cron job to redeploy the pipeline every night.

B. Compare the training and evaluation losses of the current run. If the losses are similar, deploy the model to a Vertex AI endpoint with training/serving skew threshold model monitoring. When the model monitoring threshold is triggered redeploy the pipeline.

C. Compare the results to the evaluation results from a previous run. If the performance improved deploy the model to a Vertex AI endpoint. Configure a cron job to redeploy the pipeline every night.

D. Compare the results to the evaluation results from a previous run. If the performance improved deploy the model to a Vertex AI endpoint with training/serving skew threshold model monitoring. When the model monitoring threshold is triggered redeploy the pipeline.

Show Suggested Answer

by pikachu007 at *Jan. 13, 2024, 9:13 a.m.*

Comments

Type your comment...

Submit

pikachu007 Highly Voted 9 months, 3 weeks ago

Selected Answer: B

Option A: Redeploying the pipeline every night without checking for degradation wastes resources if model performance is stable.
Option C: Comparing results to a previous run doesn't guarantee model degradation detection in the current run.
Option D: Comparing to a previous run and using model monitoring is redundant; model monitoring alone is sufficient.

upvoted 5 times

fitri001 Most Recent 6 months, 3 weeks ago

Selected Answer: D

Since the goal is to minimize cost while maintaining accuracy, Option D provides a more targeted approach for retraining based on the likelihood of the model being outdated due to data changes. Option B might trigger retraining more frequently even if the performance difference doesn't necessarily stem from a significant shift in the data distribution.

upvoted 4 times

fitri001 6 months, 3 weeks ago

Option D: Utilizes training/serving skew monitoring. This specifically focuses on identifying discrepancies between the training data and the real-world data the deployed model encounters. This is a strong indicator of when the model might be outdated due to changes in the data distribution.
Option B: Utilizes training/serving loss monitoring. Training loss tells you how well the model performs on the training data, while serving loss tells you how well it performs on real-world data. While high serving loss can indicate a problem, it might not necessarily be due to training/serving skew. Other factors like data quality issues or concept drift (gradual changes in the underlying data patterns) could also lead to high serving loss.

upvoted 2 times

omermahgoub 6 months, 3 weeks ago

Selected Answer: D

D. Compare the results to the evaluation results from a previous run. If the performance improved, deploy the model to a Vertex AI endpoint with training/serving skew threshold model monitoring. When the model monitoring threshold is triggered, redeploy the pipeline.

upvoted 3 times

pinimichele01 6 months, 3 weeks ago

i agree, see guilhermebutzke

upvoted 1 times

guilhermebutzke 8 months, 3 weeks ago

Selected Answer: D

My answer D:
A and C: Not Correct: Schedule a retrain every night is not necessary since the model is performing well.
B. Not Correct: This approach focuses on internal consistency within the current training run, train versus loss evaluation. Comparing similar training and validation losses doesn't guarantee better performance than previous models. This is an approach to identity overfitting, for example, or model quality.
D. Correct: This approach focuses on identifying performance changes over time. Comparing to previous runs helps assess if the new model performs better than the old one on the evaluation set. we will check if this new version is better or not than the old one

https://www.youtube.com/watch?v=1ykDWsnL2LE&ab_channel=GoogleCloudTech

upvoted 4 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.