Testpassport Q&A

La meilleure qualité le meilleur service

http://www.testpassport.fr

Service de mise à jour gratuit pendant un an

Exam : **Professional Machine Learning Engineer**

Title : Google Professional Machine Learning Engineer

Version : DEMO

1.You work for a company that manages a ticketing platform for a large chain of cinemas. Customers use a mobile app to search for movies they're interested in and purchase tickets in the app.
Ticket purchase requests are sent to Pub/Sub and are processed with a Dataflow streaming pipeline configured to conduct the following steps:
1. Check for availability of the movie tickets at the selected cinema.
2. Assign the ticket price and accept payment.
3. Reserve the tickets at the selected cinema.
4. Send successful purchases to your database.
Each step in this process has low latency requirements (less than 50 milliseconds). You have developed a logistic regression model with BigQuery ML that predicts whether offering a promo code for free popcorn increases the chance of a ticket purchase, and this prediction should be added to the ticket purchase process. You want to identify the simplest way to deploy this model to production while adding minimal latency.
What should you do?
A. Run batch inference with BigQuery ML every five minutes on each new set of tickets issued.
B. Export your model in TensorFlow format, and add a tfx_bsl.public.beam.RunInference step to the Dataflow pipeline.
C. Export your model in TensorFlow format, deploy it on Vertex AI, and query the prediction endpoint from your streaming pipeline.
D. Convert your model with TensorFlow Lite (TFLite), and add it to the mobile app so that the promo code and the incoming request arrive together in Pub/Sub.
Answer: A

2.You have recently created a proof-of-concept (POC) deep learning model. You are satisfied with the overall architecture, but you need to determine the value for a couple of hyperparameters. You want to perform hyperparameter tuning on Vertex AI to determine both the appropriate embedding dimension for a categorical feature used by your model and the optimal learning rate.
You configure the following settings:
For the embedding dimension, you set the type to INTEGER with a minValue of 16 and maxValue of 64.
For the learning rate, you set the type to DOUBLE with a minValue of 10e-05 and maxValue of 10e-02.
You are using the default Bayesian optimization tuning algorithm, and you want to maximize model accuracy. Training time is not a concern.
How should you set the hyperparameter scaling for each hyperparameter and the maxParallelTrials?
A. Use UNIT_LINEAR_SCALE for the embedding dimension, UNIT_LOG_SCALE for the learning rate, and a large number of parallel trials.
B. Use UNIT_LINEAR_SCALE for the embedding dimension, UNIT_LOG_SCALE for the learning rate, and a small number of parallel trials.
C. Use UNIT_LOG_SCALE for the embedding dimension, UNIT_LINEAR_SCALE for the learning rate, and a large number of parallel trials.
D. Use UNIT_LOG_SCALE for the embedding dimension, UNIT_LINEAR_SCALE for the learning rate, and a small number of parallel trials.
Answer: B

3.You are building an ML model to predict trends in the stock market based on a wide range of factors.

While exploring the data, you notice that some features have a large range. You want to ensure that the features with the largest magnitude don't overfit the model.
What should you do?
A. Standardize the data by transforming it with a logarithmic function.
B. Apply a principal component analysis (PCA) to minimize the effect of any particular feature.
C. Use a binning strategy to replace the magnitude of each feature with the appropriate bin number.
D. Normalize the data by scaling it to have values between 0 and 1.
Answer: D

4.You work for a magazine distributor and need to build a model that predicts which customers will renew their subscriptions for the upcoming year. Using your company's historical data as your training set, you created a TensorFlow model and deployed it to AI Platform. You need to determine which customer attribute has the most predictive power for each prediction served by the model.
What should you do?
A. Use AI Platform notebooks to perform a Lasso regression analysis on your model, which will eliminate features that do not provide a strong signal.
B. Stream prediction results to BigQuery. Use BigQuery's CORR (X1, X2) function to calculate the Pearson correlation coefficient between each feature and the target variable.
C. Use the AI Explanations feature on AI Platform. Submit each prediction request with the 'explain' keyword to retrieve feature attributions using the sampled Shapley method.
D. Use the What-If tool in Google Cloud to determine how your model will perform when individual features are excluded. Rank the feature importance in order of those that caused the most significant performance drop when removed from the model.
Answer: D

5.You recently built the first version of an image segmentation model for a self-driving car. After deploying the model, you observe a decrease in the area under the curve (AUC) metric. When analyzing the video recordings, you also discover that the model fails in highly congested traffic but works as expected when there is less traffic.
What is the most likely reason for this result?
A. The model is overfitting in areas with less traffic and underfitting in areas with more traffic.
B. AUC is not the correct metric to evaluate this classification model.
C. Too much data representing congested areas was used for model training.
D. Gradients become small and vanish while backpropagating from the output to input nodes.
Answer: D