PROFESSIONAL-MACHINE-LEARNING-ENGINEER Q&As

Professional Machine Learning Engineer

Pass Google PROFESSIONAL-MACHINE-LEARNING-ENGINEER Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

https://www.certbus.com/professional-machine-learning-engineer.html

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by Google Official Exam Center

QUESTION 1

You work for an online publisher that delivers news articles to over 50 million readers. You have built an AI model that recommends content for the company\\'s weekly newsletter. A recommendation is considered successful if the article is opened within two days of the newsletter\\'s published date and the user remains on the page for at least one minute.

All the information needed to compute the success metric is available in BigQuery and is updated hourly. The model is trained on eight weeks of data, on average its performance degrades below the acceptable baseline after five weeks, and training time is 12 hours. You want to ensure that the model\\'s performance is above the acceptable baseline while minimizing cost. How should you monitor the model to determine when retraining is necessary?

A. Use Vertex AI Model Monitoring to detect skew of the input features with a sample rate of 100% and a monitoring frequency of two days.

B. Schedule a cron job in Cloud Tasks to retrain the model every week before the newsletter is created.

C. Schedule a weekly query in BigQuery to compute the success metric.

D. Schedule a daily Dataflow job in Cloud Composer to compute the success metric.

Correct Answer: C

https://cloud.google.com/blog/topics/developers-practitioners/continuous-model-evaluation-bigquery-ml-stored-procedures-and-cloud-scheduler

QUESTION 2

You are training an LSTM-based model on AI Platform to summarize text using the following job submission script:

gcloud ai-platform jobs submit training $JOB_NAME \

--package-path $TRAINER_PACKAGE_PATH \

--module-name $MAIN_TRAINER_MODULE \

--job-dir $JOB_DIR \

--region $REGION \

--scale-tier basic \

-- \

--epochs 20 \

--batch_size=32 \

--learning_rate=0.001 \

You want to ensure that training time is minimized without significantly compromising the accuracy of your model. What should you do?

A. Modify the `epochs\\' parameter.

B. Modify the `scale-tier\\' parameter.

C. Modify the `batch size\\' parameter.

D. Modify the `learning rate\\' parameter.

Correct Answer: B

Changing the scale tier does not impact performance?nly speeds up training time. Epochs, Batch size, and learning rate all are hyperparameters that might impact model accuracy.

QUESTION 3

You work for an online travel agency that also sells advertising placements on its website to other companies. You have been asked to predict the most relevant web banner that a user should see next. Security is important to your company. The model latency requirements are 300ms@p99, the inventory is thousands of web banners, and your exploratory analysis has shown that navigation context is a good predictor. You want to Implement the simplest solution. How should you configure the prediction pipeline?

A. Embed the client on the website, and then deploy the model on AI Platform Prediction.

B. Embed the client on the website, deploy the gateway on App Engine, deploy the database on Firestore for writing and for reading the user\\'s navigation context, and then deploy the model on AI Platform Prediction.

C. Embed the client on the website, deploy the gateway on App Engine, deploy the database on Cloud Bigtable for writing and for reading the user\\'s navigation context, and then deploy the model on AI Platform Prediction.

D. Embed the client on the website, deploy the gateway on App Engine, deploy the database on Memorystore for writing and for reading the user\\'s navigation context, and then deploy the model on Google Kubernetes Engine.

Correct Answer: C

QUESTION 4

You work for a startup that has multiple data science workloads. Your compute infrastructure is currently on-premises, and the data science workloads are native to PySpark. Your team plans to migrate their data science workloads to Google Cloud. You need to build a proof of concept to migrate one data science job to Google Cloud. You want to propose a migration process that requires minimal cost and effort. What should you do first?

A. Create a n2-standard-4 VM instance and install Java, Scala, and Apache Spark dependencies on it.

B. Create a Google Kubernetes Engine cluster with a basic node pool configuration, install Java, Scala, and Apache Spark dependencies on it.

C. Create a Standard (1 master, 3 workers) Dataproc cluster, and run a Vertex AI Workbench notebook instance on it.

D. Create a Vertex AI Workbench notebook with instance type n2-standard-4.

Correct Answer: C

QUESTION 5

You are implementing a batch inference ML pipeline in Google Cloud. The model was developed using TensorFlow and is stored in SavedModel format in Cloud Storage. You need to apply the model to a historical dataset containing 10 TB of data that is stored in a BigQuery table. How should you perform the inference?

A. Export the historical data to Cloud Storage in Avro format. Configure a Vertex AI batch prediction job to generate predictions for the exported data

B. Import the TensorFlow model by using the CREATE MODEL statement in BigQuery ML. Apply the historical data to the TensorFlow model

C. Export the historical data to Cloud Storage in CSV format. Configure a Vertex AI batch prediction job to generate predictions for the exported data

D. Configure a Vertex AI batch prediction job to apply the model to the historical data in BigQuery

Correct Answer: B

QUESTION 6

You have trained a model on a dataset that required computationally expensive preprocessing operations. You need to execute the same preprocessing at prediction time. You deployed the model on AI Platform for high-throughput online prediction. Which architecture should you use?

A. Validate the accuracy of the model that you trained on preprocessed data. Create a new model that uses the raw data and is available in real time. Deploy the new model onto AI Platform for online prediction.

B. Send incoming prediction requests to a Pub/Sub topic. Transform the incoming data using a Dataflow job. Submit a prediction request to AI Platform using the transformed data. Write the predictions to an outbound Pub/Sub queue.

C. Stream incoming prediction request data into Cloud Spanner. Create a view to abstract your preprocessing logic. Query the view every second for new records. Submit a prediction request to AI Platform using the transformed data. Write the predictions to an outbound Pub/Sub queue.

D. Send incoming prediction requests to a Pub/Sub topic. Set up a Cloud Function that is triggered when messages are published to the Pub/Sub topic. Implement your preprocessing logic in the Cloud Function. Submit a prediction request to AI Platform using the transformed data. Write the predictions to an outbound Pub/Sub queue.

Correct Answer: B

https://cloud.google.com/architecture/data-preprocessing-for-ml-with-tf-transform-pt1#where_to_do_preprocessing

QUESTION 7

You work for a semiconductor manufacturing company. You need to create a real-time application that automates the quality control process. High-definition images of each semiconductor are taken at the end of the assembly line in real time. The photos are uploaded to a Cloud Storage bucket along with tabular data that includes each semiconductor\\'s batch number, serial number, dimensions, and weight. You need to configure model training and serving while maximizing model accuracy. What should you do?

A. Use Vertex AI Data Labeling Service to label the images, and tram an AutoML image classification model. Deploy the model, and configure Pub/Sub to publish a message when an image is categorized into the failing class.

B. Use Vertex AI Data Labeling Service to label the images, and train an AutoML image classification model. Schedule a daily batch prediction job that publishes a Pub/Sub message when the job completes.

C. Convert the images into an embedding representation. Import this data into BigQuery, and train a BigQuery ML K-means clustering model with two clusters. Deploy the model and configure Pub/Sub to publish a message when a semiconductor\\'s data is categorized into the failing cluster.

D. Import the tabular data into BigQuery, use Vertex AI Data Labeling Service to label the data and train an AutoML tabular classification model. Deploy the model, and configure Pub/Sub to publish a message when a semiconductor\\'s data is categorized into the failing class.

Correct Answer: A

QUESTION 8

You are a data scientist at an industrial equipment manufacturing company. You are developing a regression model to estimate the power consumption in the company\\'s manufacturing plants based on sensor data collected from all of the plants. The sensors collect tens of millions of records every day. You need to schedule daily training runs for your model that use all the data collected up to the current date. You want your model to scale smoothly and require minimal development work. What should you do?

A. Develop a custom TensorFlow regression model, and optimize it using Vertex AI Training.

B. Develop a regression model using BigQuery ML.

C. Develop a custom scikit-learn regression model, and optimize it using Vertex AI Training.

D. Develop a custom PyTorch regression model, and optimize it using Vertex AI Training.

Correct Answer: B

Minimal development effort => BigQueryML

QUESTION 9

You work for a large retailer, and you need to build a model to predict customer chum. The company has a dataset of historical customer data, including customer demographics purchase history, and website activity. You need to create the model in BigQuery ML and thoroughly evaluate its performance. What should you do?

A. Create a linear regression model in BigQuery ML, and register the model in Vertex AI Model Registry. Evaluate the model performance in Vertex AI .

B. Create a logistic regression model in BigQuery ML and register the model in Vertex AI Model Registry. Evaluate the model performance in Vertex AI .

C. Create a linear regression model in BigQuery ML. Use the ML.EVALUATE function to evaluate the model performance.

D. Create a logistic regression model in BigQuery ML. Use the ML.CONFUSION_MATRIX function to evaluate the model performance.

Correct Answer: B

QUESTION 10

You need to execute a batch prediction on 100 million records in a BigQuery table with a custom TensorFlow DNN regressor model, and then store the predicted results in a BigQuery table. You want to minimize the effort required to build this inference pipeline. What should you do?

A. Import the TensorFlow model with BigQuery ML, and run the ml.predict function.

B. Use the TensorFlow BigQuery reader to load the data, and use the BigQuery API to write the results to BigQuery.

C. Create a Dataflow pipeline to convert the data in BigQuery to TFRecords. Run a batch inference on Vertex AI Prediction, and write the results to BigQuery.

D. Load the TensorFlow SavedModel in a Dataflow pipeline. Use the BigQuery I/O connector with a custom function to perform the inference within the pipeline, and write the results to BigQuery.

Correct Answer: A

https://cloud.google.com/bigquery-ml/docs/making-predictions-with-imported-tensorflow-models#api
https://towardsdatascience.com/how-to-do-batch-predictions-of-tensorflow-models-directly-in-bigquery-ffa843ebdba6

QUESTION 11

You work for a large technology company that wants to modernize their contact center. You have been asked to develop a solution to classify incoming calls by product so that requests can be more quickly routed to the correct support team. You have already transcribed the calls using the Speech-to-Text API. You want to minimize data preprocessing and development time. How should you build the model?

A. Use the AI Platform Training built-in algorithms to create a custom model.

B. Use AutoMlL Natural Language to extract custom entities for classification.

C. Use the Cloud Natural Language API to extract custom entities for classification.

D. Build a custom model to identify the product keywords from the transcribed calls, and then run the keywords through a classification algorithm.

Correct Answer: B

QUESTION 12

You have been asked to productionize a proof-of-concept ML model built using Keras. The model was trained in a Jupyter notebook on a data scientist\\'s local machine. The notebook contains a cell that performs data validation and a cell that performs model analysis. You need to orchestrate the steps contained in the notebook and automate the execution of these steps for weekly retraining. You expect much more training data in the future. You want your solution to take advantage of managed services while minimizing cost. What should you do?

A. Move the Jupyter notebook to a Notebooks instance on the largest N2 machine type, and schedule the execution of the steps in the Notebooks instance using Cloud Scheduler.

B. Write the code as a TensorFlow Extended (TFX) pipeline orchestrated with Vertex AI Pipelines. Use standard TFX components for data validation and model analysis, and use Vertex AI Pipelines for model retraining.

C. Rewrite the steps in the Jupyter notebook as an Apache Spark job, and schedule the execution of the job on ephemeral Dataproc clusters using Cloud Scheduler.

D. Extract the steps contained in the Jupyter notebook as Python scripts, wrap each script in an Apache Airflow BashOperator, and run the resulting directed acyclic graph (DAG) in Cloud Composer.

Correct Answer: B

QUESTION 13

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.

Correct Answer: C

https://cloud.google.com/ai-platform/prediction/docs/ai-explanations/overview

QUESTION 14

You are creating an ML pipeline for data processing, model training, and model deployment that uses different Google Cloud services. You have developed code for each individual task, and you expect a high frequency of new files. You now need to create an orchestration layer on top of these tasks. You only want this orchestration pipeline to run if new files are present in your dataset in a Cloud Storage bucket. You also want to minimize the compute node costs. What should you do?

A. Create a pipeline in Vertex AI Pipelines. Configure the first step to compare the contents of the bucket to the last time the pipeline was run. Use the scheduler API to run the pipeline periodically.

B. Create a Cloud Function that uses a Cloud Storage trigger and deploys a Cloud Composer directed acyclic graph (DAG).

C. Create a pipeline in Vertex AI Pipelines. Create a Cloud Function that uses a Cloud Storage trigger and deploys the pipeline.

D. Deploy a Cloud Composer directed acyclic graph (DAG) with a GCSObjectUpdateSensor class that detects when a

new file is added to the Cloud Storage bucket.

Correct Answer: C

QUESTION 15

You are developing a process for training and running your custom model in production. You need to be able to show lineage for your model and predictions. What should you do?

A. 1. Create a Vertex AI managed dataset.

2.

Use a Vertex AI training pipeline to train your model.

3.

Generate batch predictions in Vertex AI.

B. 1. Use a Vertex AI Pipelines custom training job component to tram your model.

2. Generate predictions by using a Vertex AI Pipelines model batch predict component.

C. 1. Upload your dataset to BigQuery.

2.

Use a Vertex AI custom training job to train your model.

3.

Generate predictions by using Vertex Al SDK custom prediction routines.

D. 1. Use Vertex AI Experiments to train your model.

2.

Register your model in Vertex AI Model Registry.

3.

Generate batch predictions in Vertex AI.

Correct Answer: D