Custom View Settings

Question #45 *Topic 1*

You are training a TensorFlow model on a structured dataset with 100 billion records stored in several CSV files. You need to improve the input/output execution performance. What should you do?

A. Load the data into BigQuery, and read the data from BigQuery.

B. Load the data into Cloud Bigtable, and read the data from Bigtable.

C. Convert the CSV files into shards of TFRecords, and store the data in Cloud Storage.

D. Convert the CSV files into shards of TFRecords, and store the data in the Hadoop Distributed File System (HDFS).

Question #46 *Topic 1*

As the lead ML Engineer for your company, you are responsible for building ML models to digitize scanned customer forms. You have developed a TensorFlow model that converts the scanned images into text and stores them in Cloud Storage. You need to use your ML model on the aggregated data collected at the end of each day with minimal manual intervention. What should you do?

A. Use the batch prediction functionality of AI Platform.

B. Create a serving pipeline in Compute Engine for prediction.

C. Use Cloud Functions for prediction each time a new data point is ingested.

D. Deploy the model on AI Platform and create a version of it for online inference.

Question #47 *Topic 1*

You recently joined an enterprise-scale company that has thousands of datasets. You know that there are accurate descriptions for each table in BigQuery, and you are searching for the proper BigQuery table to use for a model you are building on AI Platform. How should you find the data that you need?

A. Use Data Catalog to search the BigQuery datasets by using keywords in the table description.

B. Tag each of your model and version resources on AI Platform with the name of the BigQuery table that was used for training.

C. Maintain a lookup table in BigQuery that maps the table descriptions to the table ID. Query the lookup table to find the correct table ID for the data that you need.

D. Execute a query in BigQuery to retrieve all the existing table names in your project using the INFORMATION_SCHEMA metadata tables that are native to BigQuery. Use the result o find the table that you need.

Question #48 *Topic 1*

You started working on a classification problem with time series data and achieved an area under the receiver operating characteristic curve (AUC ROC) value of
99% for training data after just a few experiments. You haven't explored using any sophisticated algorithms or spent any time on hyperparameter tuning. What should your next step be to identify and fix the problem?

A. Address the model overfitting by using a less complex algorithm.

B. Address data leakage by applying nested cross-validation during model training.

C. Address data leakage by removing features highly correlated with the target value.

D. Address the model overfitting by tuning the hyperparameters to reduce the AUC ROC value.

Previous Questions Next Questions