Sample Question 1

You work for a retail company. You have a managed tabular dataset in Vertex AI that contains sales data from three different stores. The dataset includes several features, such as store name and sale timestamp. You want to use the data to train a model that makes sales predictions for a new store that will open soon. You need to split the data between the training, validation, and test sets. What approach should you use to split the data?

A: Use Vertex AI manual split, using the store name feature to assign one store for each set
B: Use Vertex AI default data split
C: Use Vertex AI chronological split, and specify the sales timestamp feature as the time variable
D: Use Vertex AI random split, assigning 70% of the rows to the training set, 10% to the validation set, and 20% to the test set

Sample Question 1

You work for a retail company. You have a managed tabular dataset in Vertex AI that contains sales data from three different stores. The dataset includes several features, such as store name and sale timestamp. You want to use the data to train a model that makes sales predictions for a new store that will open soon. You need to split the data between the training, validation, and test sets. What approach should you use to split the data?

A: Use Vertex AI manual split, using the store name feature to assign one store for each set
B: Use Vertex AI default data split
C: Use Vertex AI chronological split, and specify the sales timestamp feature as the time variable
D: Use Vertex AI random split, assigning 70% of the rows to the training set, 10% to the validation set, and 20% to the test set

Sample Question 2

You are developing a model to identify traffic signs in images extracted from videos taken from the dashboard of a vehicle. You have a dataset of 100,000 images that were cropped to show one out of ten different traffic signs. The images have been labeled accordingly for model training, and are stored in a Cloud Storage bucket. You need to be able to tune the model during each training run. How should you train the model?

A: Train a model for object detection by using Vertex AI AutoML.
B: Train a model for image classification by using Vertex AI AutoML.
C: Develop the model training code for object detection, and train a model by using Vertex AI custom training.
D: Develop the model training code for image classification, and train a model by using Vertex AI custom training.

Sample Question 2

You are developing a model to identify traffic signs in images extracted from videos taken from the dashboard of a vehicle. You have a dataset of 100,000 images that were cropped to show one out of ten different traffic signs. The images have been labeled accordingly for model training, and are stored in a Cloud Storage bucket. You need to be able to tune the model during each training run. How should you train the model?

A: Train a model for object detection by using Vertex AI AutoML.
B: Train a model for image classification by using Vertex AI AutoML.
C: Develop the model training code for object detection, and train a model by using Vertex AI custom training.
D: Develop the model training code for image classification, and train a model by using Vertex AI custom training.

Sample Question 130

You are developing a model to detect fraudulent credit card transactions. You need to prioritize detection, because missing even one fraudulent transaction could severely impact the credit card holder. You used AutoML to tram a model on users' profile information and credit card transaction data After training the initial model, you notice that the model is failing to detect many fraudulent transactions. How should you adjust the training parameters in AutoML to improve model performance? (Choose two.)

A: Increase the score threshold
B: Decrease the score threshold.
C: Add more positive examples to the training set
D: Add more negative examples to the training set

Sample Question 130

You are developing a model to detect fraudulent credit card transactions. You need to prioritize detection, because missing even one fraudulent transaction could severely impact the credit card holder. You used AutoML to tram a model on users' profile information and credit card transaction data After training the initial model, you notice that the model is failing to detect many fraudulent transactions. How should you adjust the training parameters in AutoML to improve model performance? (Choose two.)

A: Increase the score threshold
B: Decrease the score threshold.
C: Add more positive examples to the training set
D: Add more negative examples to the training set

Sample Question 3

You are working on a binary classification ML algorithm that detects whether an image of a classified scanned document contains a company's logo. In the dataset, 96% of examples don't have the logo, so the dataset is very skewed. Which metric would give you the most confidence in your model?

A: Precision
B: Recall
C: RMSE
D: F1 Score

Sample Question 3

You are working on a binary classification ML algorithm that detects whether an image of a classified scanned document contains a company's logo. In the dataset, 96% of examples don't have the logo, so the dataset is very skewed. Which metric would give you the most confidence in your model?

A: Precision
B: Recall
C: RMSE
D: F1 Score

Sample Question 127

You are building a predictive maintenance model to preemptively detect part defects in bridges. You plan to use high definition images of the bridges as model inputs. You need to explain the output of the model to the relevant stakeholders so they can take appropriate action. How should you build the model?

A: Use scikit-learn to build a tree-based model, and use SHAP values to explain the model output.
B: Use scikit-learn to build a tree-based model, and use partial dependence plots (PDP) to explain the model output.
C: Use TensorFlow to create a deep learning-based model, and use Integrated Gradients to explain the model output.
D: Use TensorFlow to create a deep learning-based model, and use the sampled Shapley method to explain the model output.

Sample Question 127

You are building a predictive maintenance model to preemptively detect part defects in bridges. You plan to use high definition images of the bridges as model inputs. You need to explain the output of the model to the relevant stakeholders so they can take appropriate action. How should you build the model?

A: Use scikit-learn to build a tree-based model, and use SHAP values to explain the model output.
B: Use scikit-learn to build a tree-based model, and use partial dependence plots (PDP) to explain the model output.
C: Use TensorFlow to create a deep learning-based model, and use Integrated Gradients to explain the model output.
D: Use TensorFlow to create a deep learning-based model, and use the sampled Shapley method to explain the model output.

Sample Question 36

You recently deployed an ML model. Three months after deployment, you notice that your model is underperforming on certain subgroups, thus potentially leading to biased results. You suspect that the inequitable performance is due to class imbalances in the training data, but you cannot collect more data. What should you do? (Choose two.)

A: Remove training examples of high-performing subgroups, and retrain the model.
B: Add an additional objective to penalize the model more for errors made on the minority class, and retrain the model
C: Remove the features that have the highest correlations with the majority class.
D: Upsample or reweight your existing training data, and retrain the model

Sample Question 36

You recently deployed an ML model. Three months after deployment, you notice that your model is underperforming on certain subgroups, thus potentially leading to biased results. You suspect that the inequitable performance is due to class imbalances in the training data, but you cannot collect more data. What should you do? (Choose two.)

A: Remove training examples of high-performing subgroups, and retrain the model.

B: Add an additional objective to penalize the model more for errors made on the minority class, and retrain the model

C: Remove the features that have the highest correlations with the majority class.

D: Upsample or reweight your existing training data, and retrain the model

Sample Question 100

You are an ML engineer at a manufacturing company. You are creating a classification model for a predictive maintenance use case. You need to predict whether a crucial machine will fail in the next three days so that the repair crew has enough time to fix the machine before it breaks. Regular maintenance of the machine is relatively inexpensive, but a failure would be very costly. You have trained several binary classifiers to predict whether the machine will fail, where a prediction of 1 means that the ML model predicts a failure.

You are now evaluating each model on an evaluation dataset. You want to choose a model that prioritizes detection while ensuring that more than 50% of the maintenance jobs triggered by your model address an imminent machine failure. Which model should you choose?

A: The model with the highest area under the receiver operating characteristic curve (AUC ROC) and precision greater than 0.5
B: The model with the lowest root mean squared error (RMSE) and recall greater than 0.5.
C: The model with the highest recall where precision is greater than 0.5.
D: The model with the highest precision where recall is greater than 0.5.

Sample Question 100

You are an ML engineer at a manufacturing company. You are creating a classification model for a predictive maintenance use case. You need to predict whether a crucial machine will fail in the next three days so that the repair crew has enough time to fix the machine before it breaks. Regular maintenance of the machine is relatively inexpensive, but a failure would be very costly. You have trained several binary classifiers to predict whether the machine will fail, where a prediction of 1 means that the ML model predicts a failure.

You are now evaluating each model on an evaluation dataset. You want to choose a model that prioritizes detection while ensuring that more than 50% of the maintenance jobs triggered by your model address an imminent machine failure. Which model should you choose?

A: The model with the highest area under the receiver operating characteristic curve (AUC ROC) and precision greater than 0.5
B: The model with the lowest root mean squared error (RMSE) and recall greater than 0.5.
C: The model with the highest recall where precision is greater than 0.5.
D: The model with the highest precision where recall is greater than 0.5.

Sample Question 59

You are developing a custom TensorFlow classification model based on tabular data. Your raw data is stored in BigQuery. contains hundreds of millions of rows, and includes both categorical and numerical features. You need to use a MaxMin scaler on some numerical features, and apply a one-hot encoding to some categorical features such as SKU names. Your model will be trained over multiple epochs. You want to minimize the effort and cost of your solution. What should you do?

A: 1. Write a SQL query to create a separate lookup table to scale the numerical features.
2. Deploy a TensorFlow-based model from Hugging Face to BigQuery to encode the text features.
3. Feed the resulting BigQuery view into Vertex AI Training.

B: 1. Use BigQuery to scale the numerical features.
2. Feed the features into Vertex AI Training.
3. Allow TensorFlow to perform the one-hot text encoding.

C: 1. Use TFX components with Dataflow to encode the text features and scale the numerical features.
2. Export results to Cloud Storage as TFRecords.
3. Feed the data into Vertex AI Training.

D: 1. Write a SQL query to create a separate lookup table to scale the numerical features.
2. Perform the one-hot text encoding in BigQuery.
3. Feed the resulting BigQuery view into Vertex AI Training.

Sample Question 59

You are developing a custom TensorFlow classification model based on tabular data. Your raw data is stored in BigQuery. contains hundreds of millions of rows, and includes both categorical and numerical features. You need to use a MaxMin scaler on some numerical features, and apply a one-hot encoding to some categorical features such as SKU names. Your model will be trained over multiple epochs. You want to minimize the effort and cost of your solution. What should you do?

A: 1. Write a SQL query to create a separate lookup table to scale the numerical features.
2. Deploy a TensorFlow-based model from Hugging Face to BigQuery to encode the text features.
3. Feed the resulting BigQuery view into Vertex AI Training.

B: 1. Use BigQuery to scale the numerical features.
2. Feed the features into Vertex AI Training.
3. Allow TensorFlow to perform the one-hot text encoding.

C: 1. Use TFX components with Dataflow to encode the text features and scale the numerical features.
2. Export results to Cloud Storage as TFRecords.
3. Feed the data into Vertex AI Training.

D: 1. Write a SQL query to create a separate lookup table to scale the numerical features.
2. Perform the one-hot text encoding in BigQuery.
3. Feed the resulting BigQuery view into Vertex AI Training.

Sample Question 55

You work at a large organization that recently decided to move their ML and data workloads to Google Cloud. The data engineering team has exported the structured data to a Cloud Storage bucket in Avro format. You need to propose a workflow that performs analytics, creates features, and hosts the features that your ML models use for online prediction. How should you configure the pipeline?

A: Ingest the Avro files into Cloud Spanner to perform analytics. Use a Dataflow pipeline to create the features, and store them in Vertex AI Feature Store for online prediction.
B: Ingest the Avro files into BigQuery to perform analytics. Use a Dataflow pipeline to create the features, and store them in Vertex AI Feature Store for online prediction.
C: Ingest the Avro files into Cloud Spanner to perform analytics. Use a Dataflow pipeline to create the features, and store them in BigQuery for online prediction.
D: Ingest the Avro files into BigQuery to perform analytics. Use BigQuery SQL to create features and store them in a separate BigQuery table for online prediction.

Sample Question 55

You work at a large organization that recently decided to move their ML and data workloads to Google Cloud. The data engineering team has exported the structured data to a Cloud Storage bucket in Avro format. You need to propose a workflow that performs analytics, creates features, and hosts the features that your ML models use for online prediction. How should you configure the pipeline?

A: Ingest the Avro files into Cloud Spanner to perform analytics. Use a Dataflow pipeline to create the features, and store them in Vertex AI Feature Store for online prediction.

B: Ingest the Avro files into BigQuery to perform analytics. Use a Dataflow pipeline to create the features, and store them in Vertex AI Feature Store for online prediction.

C: Ingest the Avro files into Cloud Spanner to perform analytics. Use a Dataflow pipeline to create the features, and store them in BigQuery for online prediction.

D: Ingest the Avro files into BigQuery to perform analytics. Use BigQuery SQL to create features and store them in a separate BigQuery table for online prediction.

Sample Question 67

You are an ML engineer at a retail company. You have built a model that predicts a coupon to offer an ecommerce customer at checkout based on the items in their cart. When a customer goes to checkout, your serving pipeline, which is hosted on Google Cloud, joins the customer's existing cart with a row in a BigQuery table that contains the customers' historic purchase behavior and uses that as the model's input. The web team is reporting that your model is returning predictions too slowly to load the coupon offer with the rest of the web page. How should you speed up your model's predictions?

A: Attach an NVIDIA P100 GPU to your deployed model’s instance.
B: Use a low latency database for the customers’ historic purchase behavior.
C: Deploy your model to more instances behind a load balancer to distribute traffic.
D: Create a materialized view in BigQuery with the necessary data for predictions.

Sample Question 67

You are an ML engineer at a retail company. You have built a model that predicts a coupon to offer an ecommerce customer at checkout based on the items in their cart. When a customer goes to checkout, your serving pipeline, which is hosted on Google Cloud, joins the customer's existing cart with a row in a BigQuery table that contains the customers' historic purchase behavior and uses that as the model's input. The web team is reporting that your model is returning predictions too slowly to load the coupon offer with the rest of the web page. How should you speed up your model's predictions?

A: Attach an NVIDIA P100 GPU to your deployed model's instance.
B: Use a low latency database for the customers' historic purchase behavior.
C: Deploy your model to more instances behind a load balancer to distribute traffic.
D: Create a materialized view in BigQuery with the necessary data for predictions.

Sample Question 84

You need to develop a custom TensorFlow model that will be used for online predictions. The training data is stored in BigQuery You need to apply instance-level data transformations to the data for model training and serving. You want to use the same preprocessing routine during model training and serving. How should you configure the preprocessing routine?

A: Create a BigQuery script to preprocess the data, and write the result to another BigQuery table.
B: Create a pipeline in Vertex AI Pipelines to read the data from BigQuery and preprocess it using a custom preprocessing component.
C: Create a preprocessing function that reads and transforms the data from BigQuery. Create a Vertex AI custom prediction routine that calls the preprocessing function at serving time.
D: Create an Apache Beam pipeline to read the data from BigQuery and preprocess it by using TensorFlow Transform and Dataflow.

Sample Question 84

You need to develop a custom TensorFlow model that will be used for online predictions. The training data is stored in BigQuery You need to apply instance-level data transformations to the data for model training and serving. You want to use the same preprocessing routine during model training and serving. How should you configure the preprocessing routine?

A: Create a BigQuery script to preprocess the data, and write the result to another BigQuery table.
B: Create a pipeline in Vertex AI Pipelines to read the data from BigQuery and preprocess it using a custom preprocessing component.
C: Create a preprocessing function that reads and transforms the data from BigQuery. Create a Vertex AI custom prediction routine that calls the preprocessing function at serving time.
D: Create an Apache Beam pipeline to read the data from BigQuery and preprocess it by using TensorFlow Transform and Dataflow.

Sample Question 109

You recently deployed a model to a Vertex AI endpoint and set up online serving in Vertex AI Feature Store. You have configured a daily batch ingestion job to update your featurestore. During the batch ingestion jobs, you discover that CPU utilization is high in your featurestore's online serving nodes and that feature retrieval latency is high. You need to improve online serving performance during the daily batch ingestion. What should you do?

A: Schedule an increase in the number of online serving nodes in your featurestore prior to the batch ingestion jobs
B: Enable autoscaling of the online serving nodes in your featurestore
C: Enable autoscaling for the prediction nodes of your DeployedModel in the Vertex AI endpoint
D: Increase the worker_count in the ImportFeatureValues request of your batch ingestion job

Sample Question 109

You recently deployed a model to a Vertex AI endpoint and set up online serving in Vertex AI Feature Store. You have configured a daily batch ingestion job to update your featurestore. During the batch ingestion jobs, you discover that CPU utilization is high in your featurestore's online serving nodes and that feature retrieval latency is high. You need to improve online serving performance during the daily batch ingestion. What should you do?

A: Schedule an increase in the number of online serving nodes in your featurestore prior to the batch ingestion jobs
B: Enable autoscaling of the online serving nodes in your featurestore
C: Enable autoscaling for the prediction nodes of your DeployedModel in the Vertex AI endpoint
D: Increase the worker_count in the ImportFeatureValues request of your batch ingestion job

Sample Question 136

You are working with a dataset that contains customer transactions. You need to build an ML model to predict customer purchase behavior. You plan to develop the model in BigQuery ML, and export it to Cloud Storage for online prediction. You notice that the input data contains a few categorical features, including product category and payment method. You want to deploy the model as quickly as possible. What should you do?

A: Use the TRANSFORM clause with the ML.ONE_HOT_ENCODER function on the categorical features at model creation and select the categorical and non-categorical features.
B: Use the ML.ONE_HOT_ENCODER function on the categorical features and select the encoded categorical features and non-categorical features as inputs to create your model.
C: Use the CREATE MODEL statement and select the categorical and non-categorical features.
D: Use the ML.MULTI_HOT_ENCODER function on the categorical features, and select the encoded categorical features and non-categorical features as inputs to create your model.

Sample Question 136

You are working with a dataset that contains customer transactions. You need to build an ML model to predict customer purchase behavior. You plan to develop the model in BigQuery ML, and export it to Cloud Storage for online prediction. You notice that the input data contains a few categorical features, including product category and payment method. You want to deploy the model as quickly as possible. What should you do?

A: Use the TRANSFORM clause with the ML.ONE_HOT_ENCODER function on the categorical features at model creation and select the categorical and non-categorical features.

B: Use the ML.ONE_HOT_ENCODER function on the categorical features and select the encoded categorical features and non-categorical features as inputs to create your model.

C: Use the CREATE MODEL statement and select the categorical and non-categorical features.

D: Use the ML.MULTI_HOT_ENCODER function on the categorical features, and select the encoded categorical features and non-categorical features as inputs to create your model.

Sample Question 113

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.

Sample Question 113

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.

Sample Question 123

You are using Keras and TensorFlow to develop a fraud detection model. Records of customer transactions are stored in a large table in BigQuery. You need to preprocess these records in a cost-effective and efficient way before you use them to train the model. The trained model will be used to perform batch inference in BigQuery. How should you implement the preprocessing workflow?

A: Implement a preprocessing pipeline by using Apache Spark, and run the pipeline on Dataproc. Save the preprocessed data as CSV files in a Cloud Storage bucket.
B: Load the data into a pandas DataFrame. Implement the preprocessing steps using pandas transformations, and train the model directly on the DataFrame.
C: Perform preprocessing in BigQuery by using SQL. Use the BigQueryClient in TensorFlow to read the data directly from BigQuery.
D: Implement a preprocessing pipeline by using Apache Beam, and run the pipeline on Dataflow. Save the preprocessed data as CSV files in a Cloud Storage bucket.

Sample Question 123

You are using Keras and TensorFlow to develop a fraud detection model. Records of customer transactions are stored in a large table in BigQuery. You need to preprocess these records in a cost-effective and efficient way before you use them to train the model. The trained model will be used to perform batch inference in BigQuery. How should you implement the preprocessing workflow?

A: Implement a preprocessing pipeline by using Apache Spark, and run the pipeline on Dataproc. Save the preprocessed data as CSV files in a Cloud Storage bucket.
B: Load the data into a pandas DataFrame. Implement the preprocessing steps using pandas transformations, and train the model directly on the DataFrame.
C: Perform preprocessing in BigQuery by using SQL. Use the BigQueryClient in TensorFlow to read the data directly from BigQuery.
D: Implement a preprocessing pipeline by using Apache Beam, and run the pipeline on Dataflow. Save the preprocessed data as CSV files in a Cloud Storage bucket.

Sample Question 129

You work for a food product company. Your company's historical sales data is stored in BigQuery.You need to use Vertex AI's custom training service to train multiple TensorFlow models that read the data from BigQuery and predict future sales. You plan to implement a data preprocessing algorithm that performs mm-max scaling and bucketing on a large number of features before you start experimenting with the models. You want to minimize preprocessing time, cost, and development effort. How should you configure this workflow?

A: Write the transformations into Spark that uses the spark-bigquery-connector, and use Dataproc to preprocess the data.
B: Write SQL queries to transform the data in-place in BigQuery.
C: Add the transformations as a preprocessing layer in the TensorFlow models.
D: Create a Dataflow pipeline that uses the BigQueryIO connector to ingest the data, process it, and write it back to BigQuery.

Sample Question 129

You work for a food product company. Your company's historical sales data is stored in BigQuery.You need to use Vertex AI's custom training service to train multiple TensorFlow models that read the data from BigQuery and predict future sales. You plan to implement a data preprocessing algorithm that performs mm-max scaling and bucketing on a large number of features before you start experimenting with the models. You want to minimize preprocessing time, cost, and development effort. How should you configure this workflow?

A: Write the transformations into Spark that uses the spark-bigquery-connector, and use Dataproc to preprocess the data.

B: Write SQL queries to transform the data in-place in BigQuery.

C: Add the transformations as a preprocessing layer in the TensorFlow models.

D: Create a Dataflow pipeline that uses the BigQueryIO connector to ingest the data, process it, and write it back to BigQuery.

Sample Question 131

You want to train an AutoML model to predict house prices by using a small public dataset stored in BigQuery. You need to prepare the data and want to use the simplest, most efficient approach. What should you do?

A: Write a query that preprocesses the data by using BigQuery and creates a new table. Create a Vertex AI managed dataset with the new table as the data source.
B: Use Dataflow to preprocess the data. Write the output in TFRecord format to a Cloud Storage bucket.
C: Write a query that preprocesses the data by using BigQuery. Export the query results as CSV files, and use those files to create a Vertex AI managed dataset.
D: Use a Vertex AI Workbench notebook instance to preprocess the data by using the pandas library. Export the data as CSV files, and use those files to create a Vertex AI managed dataset.

Sample Question 131

You want to train an AutoML model to predict house prices by using a small public dataset stored in BigQuery. You need to prepare the data and want to use the simplest, most efficient approach. What should you do?

A: Write a query that preprocesses the data by using BigQuery and creates a new table. Create a Vertex AI managed dataset with the new table as the data source.

B: Use Dataflow to preprocess the data. Write the output in TFRecord format to a Cloud Storage bucket.

C: Write a query that preprocesses the data by using BigQuery. Export the query results as CSV files, and use those files to create a Vertex AI managed dataset.

D: Use a Vertex AI Workbench notebook instance to preprocess the data by using the pandas library. Export the data as CSV files, and use those files to create a Vertex AI managed dataset.

Sample Question 60

You work for a delivery company. You need to design a system that stores and manages features such as parcels delivered and truck locations over time. The system must retrieve the features with low latency and feed those features into a model for online prediction. The data science team will retrieve historical data at a specific point in time for model training. You want to store the features with minimal effort. What should you do?

A: Store features in Bigtable as key/value data.
B: Store features in Vertex AI Feature Store.
C: Store features as a Vertex AI dataset, and use those features to train the models hosted in Vertex AI endpoints.
D: Store features in BigQuery timestamp partitioned tables, and use the BigQuery Storage Read API to serve the features.

Sample Question 60

You work for a delivery company. You need to design a system that stores and manages features such as parcels delivered and truck locations over time. The system must retrieve the features with low latency and feed those features into a model for online prediction. The data science team will retrieve historical data at a specific point in time for model training. You want to store the features with minimal effort. What should you do?

A: Store features in Bigtable as key/value data.
B: Store features in Vertex AI Feature Store.
C: Store features as a Vertex AI dataset, and use those features to train the models hosted in Vertex AI endpoints.
D: Store features in BigQuery timestamp partitioned tables, and use the BigQuery Storage Read API to serve the features.

Sample Question 18

You are developing a model to help your company create more targeted online advertising campaigns. You need to create a dataset that you will use to train the model. You want to avoid creating or reinforcing unfair bias in the model. What should you do? (Choose two.)

A: Include a comprehensive set of demographic features
B: Include only the demographic groups that most frequently interact with advertisements
C: Collect a random sample of production traffic to build the training dataset
D: Collect a stratified sample of production traffic to build the training dataset

Sample Question 18

You are developing a model to help your company create more targeted online advertising campaigns. You need to create a dataset that you will use to train the model. You want to avoid creating or reinforcing unfair bias in the model. What should you do? (Choose two.)

A: Include a comprehensive set of demographic features
B: Include only the demographic groups that most frequently interact with advertisements
C: Collect a random sample of production traffic to build the training dataset
D: Collect a stratified sample of production traffic to build the training dataset

Sample Question 61

You are developing a training pipeline for a new XGBoost classification model based on tabular data. The data is stored in a BigQuery table. You need to complete the following steps:

1. Randomly split the data into training and evaluation datasets in a 65/35 ratio
2. Conduct feature engineering
3. Obtain metrics for the evaluation dataset
4. Compare models trained in different pipeline executions

How should you execute these steps?

A: 1. Using Vertex AI Pipelines, add a component to divide the data into training and evaluation sets, and add another component for feature engineering.
2. Enable autologging of metrics in the training component.
3. Compare pipeline runs in Vertex AI Experiments.

B: 1. Using Vertex AI Pipelines, add a component to divide the data into training and evaluation sets, and add another component for feature engineering.
2. Enable autologging of metrics in the training component.
3. Compare models using the artifacts' lineage in Vertex ML Metadata.

C: 1. In BigQuery ML, use the CREATE MODEL statement with BOOSTED_TREE_CLASSIFIER as the model type and use BigQuery to handle the data splits.
2. Use a SQL view to apply feature engineering and train the model using the data in that view.
3. Compare the evaluation metrics of the models by using a SQL query with the ML.TRAINING_INFO statement.

D: 1. In BigQuery ML, use the CREATE MODEL statement with BOOSTED_TREE_CLASSIFIER as the model type and use

Sample Question 61

You are developing a training pipeline for a new XGBoost classification model based on tabular data. The data is stored in a BigQuery table. You need to complete the following steps:

1. Randomly split the data into training and evaluation datasets in a 65/35 ratio
2. Conduct feature engineering
3. Obtain metrics for the evaluation dataset
4. Compare models trained in different pipeline executions

How should you execute these steps?

A: 1. Using Vertex AI Pipelines, add a component to divide the data into training and evaluation sets, and add another component for feature engineering.
2. Enable autologging of metrics in the training component.
3. Compare pipeline runs in Vertex AI Experiments.

B: 1. Using Vertex AI Pipelines, add a component to divide the data into training and evaluation sets, and add another component for feature engineering.
2. Enable autologging of metrics in the training component.
3. Compare models using the artifacts' lineage in Vertex ML Metadata.

C: 1. In BigQuery ML, use the CREATE MODEL statement with BOOSTED_TREE_CLASSIFIER as the model type and use BigQuery to handle the data splits.
2. Use a SQL view to apply feature engineering and train the model using the data in that view.
3. Compare the evaluation metrics of the models by using a SQL query with the ML.TRAINING_INFO statement.

D: 1. In BigQuery ML, use the CREATE MODEL statement with BOOSTED_TREE_CLASSIFIER as the model type and use

Sample Question 53

You work for a multinational organization that has recently begun operations in Spain. Teams within your organization will need to work with various Spanish documents, such as business, legal, and financial documents. You want to use machine learning to help your organization get accurate translations quickly and with the least effort. Your organization does not require domain-specific terms or jargon. What should you do?

A: Create a Vertex AI Workbench notebook instance. In the notebook, extract sentences from the documents, and train a custom AutoML text model.
B: Use Google Translate to translate 1,000 phrases from Spanish to English. Using these translated pairs, train a custom AutoML Translation model.
C: Use the Document Translation feature of the Cloud Translation API to translate the documents.
D: Create a Vertex AI Workbench notebook instance. In the notebook, convert the Spanish documents into plain text, and create a custom TensorFlow seq2seq translation model.

Sample Question 53

You work for a multinational organization that has recently begun operations in Spain. Teams within your organization will need to work with various Spanish documents, such as business, legal, and financial documents. You want to use machine learning to help your organization get accurate translations quickly and with the least effort. Your organization does not require domain-specific terms or jargon. What should you do?

A: Create a Vertex AI Workbench notebook instance. In the notebook, extract sentences from the documents, and train a custom AutoML text model.
B: Use Google Translate to translate 1,000 phrases from Spanish to English. Using these translated pairs, train a custom AutoML Translation model.
C: Use the Document Translation feature of the Cloud Translation API to translate the documents.
D: Create a Vertex AI Workbench notebook instance. In the notebook, convert the Spanish documents into plain text, and create a custom TensorFlow seq2seq translation model.

Sample Question 50

You work at a gaming startup that has several terabytes of structured data in Cloud Storage. This data includes gameplay time data, user metadata, and game metadata. You want to build a model that recommends new games to users that requires the least amount of coding. What should you do?

A: Load the data in BigQuery. Use BigQuery ML to train an Autoencoder model.
B: Load the data in BigQuery. Use BigQuery ML to train a matrix factorization model.
C: Read data to a Vertex AI Workbench notebook. Use TensorFlow to train a two-tower model.
D: Read data to a Vertex AI Workbench notebook. Use TensorFlow to train a matrix factorization model.

Sample Question 50

You work at a gaming startup that has several terabytes of structured data in Cloud Storage. This data includes gameplay time data, user metadata, and game metadata. You want to build a model that recommends new games to users that requires the least amount of coding. What should you do?

A: Load the data in BigQuery. Use BigQuery ML to train an Autoencoder model.
B: Load the data in BigQuery. Use BigQuery ML to train a matrix factorization model.
C: Read data to a Vertex AI Workbench notebook. Use TensorFlow to train a two-tower model.
D: Read data to a Vertex AI Workbench notebook. Use TensorFlow to train a matrix factorization model.

Sample Question 46

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.

Sample Question 46

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.

Sample Question 45

You need to build classification workflows over several structured datasets currently stored in BigQuery. Because you will be performing the classification several times, you want to complete the following steps without writing code: exploratory data analysis, feature selection, model building, training, and hyperparameter tuning and serving. What should you do?

A: Train a TensorFlow model on Vertex AI.
B: Train a classification Vertex AutoML model.
C: Run a logistic regression job on BigQuery ML.
D: Use scikit-learn in Vertex AI Workbench user-managed notebooks with pandas library.

Sample Question 45

You need to build classification workflows over several structured datasets currently stored in BigQuery. Because you will be performing the classification several times, you want to complete the following steps without writing code: exploratory data analysis, feature selection, model building, training, and hyperparameter tuning and serving. What should you do?

A: Train a TensorFlow model on Vertex AI.
B: Train a classification Vertex AutoML model.
C: Run a logistic regression job on BigQuery ML.
D: Use scikit-learn in Vertex AI Workbench user-managed notebooks with pandas library.

Sample Question 54

You work at an organization that maintains a cloud-based communication platform that integrates conventional chat, voice, and video conferencing into one platform. The audio recordings are stored in Cloud Storage. All recordings have an 8 kHz sample rate and are more than one minute long. You need to implement a new feature in the platform that will automatically transcribe voice call recordings into a text for future applications, such as call summarization and sentiment analysis. How should you implement the voice call transcription feature following Google-recommended best practices?

A: Use the original audio sampling rate, and transcribe the audio by using the Speech-to-Text API with synchronous recognition.
B: Use the original audio sampling rate, and transcribe the audio by using the Speech-to-Text API with asynchronous recognition.
C: Upsample the audio recordings to 16 kHz, and transcribe the audio by using the Speech-to-Text API with synchronous recognition.
D: Upsample the audio recordings to 16 kHz, and transcribe the audio by using the Speech-to-Text API with asynchronous recognition.

Sample Question 54

You work at an organization that maintains a cloud-based communication platform that integrates conventional chat, voice, and video conferencing into one platform. The audio recordings are stored in Cloud Storage. All recordings have an 8 kHz sample rate and are more than one minute long. You need to implement a new feature in the platform that will automatically transcribe voice call recordings into a text for future applications, such as call summarization and sentiment analysis. How should you implement the voice call transcription feature following Google-recommended best practices?

A: Use the original audio sampling rate, and transcribe the audio by using the Speech-to-Text API with synchronous recognition.

B: Use the original audio sampling rate, and transcribe the audio by using the Speech-to-Text API with asynchronous recognition.

C: Upsample the audio recordings to 16 kHz, and transcribe the audio by using the Speech-to-Text API with synchronous recognition.

D: Upsample the audio recordings to 16 kHz, and transcribe the audio by using the Speech-to-Text API with asynchronous recognition.

Sample Question 69

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

Sample Question 69

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

Sample Question 85

You work for a hotel and have a dataset that contains customers' written comments scanned from paper-based customer feedback forms, which are stored as PDF files. Every form has the same layout. You need to quickly predict an overall satisfaction score from the customer comments on each form. How should you accomplish this task?

A: Use the Vision API to parse the text from each PDF file. Use the Natural Language API analyzeSentiment feature to infer overall satisfaction scores.
B: Use the Vision API to parse the text from each PDF file. Use the Natural Language API analyzeEntitySentiment feature to infer overall satisfaction scores.
C: Uptrain a Document AI custom extractor to parse the text in the comments section of each PDF file. Use the Natural Language API analyzeSentiment feature to infer overall satisfaction scores.

D: Uptrain a Document AI custom extractor to parse the text in the comments section of each PDF file. Use the Natural Language API analyzeEntitySentiment feature to infer overall satisfaction scores.

Sample Question 85

You work for a hotel and have a dataset that contains customers' written comments scanned from paper-based customer feedback forms, which are stored as PDF files. Every form has the same layout. You need to quickly predict an overall satisfaction score from the customer comments on each form. How should you accomplish this task?

A: Use the Vision API to parse the text from each PDF file. Use the Natural Language API analyzeSentiment feature to infer overall satisfaction scores.

B: Use the Vision API to parse the text from each PDF file. Use the Natural Language API analyzeEntitySentiment feature to infer overall satisfaction scores.

C: Uptrain a Document AI custom extractor to parse the text in the comments section of each PDF file. Use the Natural Language API analyzeSentiment feature to infer overall satisfaction scores.

D: Uptrain a Document AI custom extractor to parse the text in the comments section of each PDF file. Use the Natural Language API analyzeEntitySentiment feature to infer overall satisfaction scores.

Sample Question 74

You work for a retail company. You have been tasked with building a model to determine the probability of churn for each customer. You need the predictions to be interpretable so the results can be used to develop marketing campaigns that target at-risk customers. What should you do?

A: Build a random forest regression model in a Vertex AI Workbench notebook instance. Configure the model to generate feature importances after the model is trained.
B: Build an AutoML tabular regression model. Configure the model to generate explanations when it makes predictions.
C: Build a custom TensorFlow neural network by using Vertex AI custom training. Configure the model to generate explanations when it makes predictions.
D: Build a random forest classification model in a Vertex AI Workbench notebook instance. Configure the model to generate feature importances after the model is trained.

Sample Question 74

You work for a retail company. You have been tasked with building a model to determine the probability of churn for each customer. You need the predictions to be interpretable so the results can be used to develop marketing campaigns that target at-risk customers. What should you do?

A: Build a random forest regression model in a Vertex AI Workbench notebook instance. Configure the model to generate feature importances after the model is trained.

B: Build an AutoML tabular regression model. Configure the model to generate explanations when it makes predictions.

C: Build a custom TensorFlow neural network by using Vertex AI custom training. Configure the model to generate explanations when it makes predictions.

D: Build a random forest classification model in a Vertex AI Workbench notebook instance. Configure the model to generate feature importances after the model is trained.

Sample Question 81

You work at a leading healthcare firm developing state-of-the-art algorithms for various use cases. You have unstructured textual data with custom labels. You need to extract and classify various medical phrases with these labels. What should you do?

A: Use the Healthcare Natural Language API to extract medical entities
B: Use a BERT-based model to fine-tune a medical entity extraction model
C: Use AutoML Entity Extraction to train a medical entity extraction model
D: Use TensorFlow to build a custom medical entity extraction model

Sample Question 81

You work at a leading healthcare firm developing state-of-the-art algorithms for various use cases. You have unstructured textual data with custom labels. You need to extract and classify various medical phrases with these labels. What should you do?

A: Use the Healthcare Natural Language API to extract medical entities
B: Use a BERT-based model to fine-tune a medical entity extraction model
C: Use AutoML Entity Extraction to train a medical entity extraction model
D: Use TensorFlow to build a custom medical entity extraction model

Sample Question 43

You work for a company that sells corporate electronic products to thousands of businesses worldwide. Your company stores historical customer data in BigQuery. You need to build a model that predicts customer lifetime value over the next three years. You want to use the simplest approach to build the model and you want to have access to visualization tools. What should you do?

A: Create a Vertex AI Workbench notebook to perform exploratory data analysis. Use IPython magics to create a new BigQuery table with input features. Use the BigQuery console to run the CREATE MODEL statement. Validate the results by using the ML.EVALUATE and ML.PREDICT statements.
B: Run the CREATE MODEL statement from the BigQuery console to create an AutoML model. Validate the results by using the ML.EVALUATE and ML.PREDICT statements.
C: Create a Vertex AI Workbench notebook to perform exploratory data analysis and create input features. Save the features as a CSV file in Cloud Storage. Import the CSV file as a new BigQuery table. Use the BigQuery console to run the CREATE MODEL statement. Validate the results by using the ML.EVALUATE and ML.PREDICT statements.
D: Create a Vertex AI Workbench notebook to perform exploratory data analysis. Use IPython magics to create a new BigQuery table with input features, create the model, and validate the results by using the CREATE MODEL, ML.EVALUATE, and ML.PREDICT statements.

Sample Question 43

You work for a company that sells corporate electronic products to thousands of businesses worldwide. Your company stores historical customer data in BigQuery. You need to build a model that predicts customer lifetime value over the next three years. You want to use the simplest approach to build the model and you want to have access to visualization tools. What should you do?

A: Create a Vertex AI Workbench notebook to perform exploratory data analysis. Use IPython magics to create a new BigQuery table with input features. Use the BigQuery console to run the CREATE MODEL statement. Validate the results by using the ML.EVALUATE and ML.PREDICT statements.
B: Run the CREATE MODEL statement from the BigQuery console to create an AutoML model. Validate the results by using the ML.EVALUATE and ML.PREDICT statements.
C: Create a Vertex AI Workbench notebook to perform exploratory data analysis and create input features. Save the features as a CSV file in Cloud Storage. Import the CSV file as a new BigQuery table. Use the BigQuery console to run the CREATE MODEL statement. Validate the results by using the ML.EVALUATE and ML.PREDICT statements.
D: Create a Vertex AI Workbench notebook to perform exploratory data analysis. Use IPython magics to create a new BigQuery table with input features, create the model, and validate the results by using the CREATE MODEL, ML.EVALUATE, and ML.PREDICT statements.

Sample Question 87

You work for a company that captures live video footage of checkout areas in their retail stores. You need to use the live video footage to build a model to detect the number of customers waiting for service in near real time. You want to implement a solution quickly and with minimal effort. How should you build the model?

A: Use the Vertex AI Vision Occupancy Analytics model.
B: Use the Vertex AI Vision Person/vehicle detector model.
C: Train an AutoML object detection model on an annotated dataset by using Vertex AutoML.
D: Train a Seq2Seq+ object detection model on an annotated dataset by using Vertex AutoML.

Sample Question 87

You work for a company that captures live video footage of checkout areas in their retail stores. You need to use the live video footage to build a model to detect the number of customers waiting for service in near real time. You want to implement a solution quickly and with minimal effort. How should you build the model?

A: Use the Vertex AI Vision Occupancy Analytics model.
B: Use the Vertex AI Vision Person/vehicle detector model.
C: Train an AutoML object detection model on an annotated dataset by using Vertex AutoML.
D: Train a Seq2Seq+ object detection model on an annotated dataset by using Vertex AutoML.

Sample Question 88

Your company stores a large number of audio files of phone calls made to your customer call center in an on-premises database. Each audio file is in wav format and is approximately 5 minutes long. You need to analyze these audio files for customer sentiment. You plan to use the Speech-to-Text API You want to use the most efficient approach. What should you do?

A: 1. Upload the audio files to Cloud Storage
2. Call the speech:longrunningrecognize API endpoint to generate transcriptions
3. Call the predict method of an AutoML sentiment analysis model to analyze the transcriptions.

B: 1. Upload the audio files to Cloud Storage.
2. Call the speech:longrunningrecognize API endpoint to generate transcriptions
3. Create a Cloud Function that calls the Natural Language API by using the analyzeSentiment method

C: 1. Iterate over your local files in Python
2. Use the Speech-to-Text Python library to create a speech.RecognitionAudio object, and set the content to the audio file data
3. Call the speech:recognize API endpoint to generate transcriptions
4. Call the predict method of an AutoML sentiment analysis model to analyze the transcriptions.

D: 1. Iterate over your local files in Python
2. Use the Speech-to-Text Python Library to create a speech.RecognitionAudio object and set the content to the audio file data
3. Call the speech:longrunningrecognize API endpoint to generate transcriptions.
4. Call the Natural Language API by using the analyzeSentiment method

Sample Question 88

Your company stores a large number of audio files of phone calls made to your customer call center in an on-premises database. Each audio file is in wav format and is approximately 5 minutes long. You need to analyze these audio files for customer sentiment. You plan to use the Speech-to-Text API You want to use the most efficient approach. What should you do?

A: 1. Upload the audio files to Cloud Storage
2. Call the speech:longrunningrecognize API endpoint to generate transcriptions
3. Call the predict method of an AutoML sentiment analysis model to analyze the transcriptions.

B: 1. Upload the audio files to Cloud Storage.
2. Call the speech:longrunningrecognize API endpoint to generate transcriptions
3. Create a Cloud Function that calls the Natural Language API by using the analyzeSentiment method

C: 1. Iterate over your local files in Python
2. Use the Speech-to-Text Python library to create a speech.RecognitionAudio object, and set the content to the audio file data
3. Call the speech:recognize API endpoint to generate transcriptions
4. Call the predict method of an AutoML sentiment analysis model to analyze the transcriptions.

D: 1. Iterate over your local files in Python
2. Use the Speech-to-Text Python Library to create a speech.RecognitionAudio object and set the content to the audio file data
3. Call the speech:longrunningrecognize API endpoint to generate transcriptions.
4. Call the Natural Language API by using the analyzeSentiment method

Sample Question 97

You work for a pet food company that manages an online forum. Customers upload photos of their pets on the forum to share with others. About 20 photos are uploaded daily. You want to automatically and in near real time detect whether each uploaded photo has an animal. You want to prioritize time and minimize cost of your application development and deployment. What should you do?

A: Send user-submitted images to the Cloud Vision API. Use object localization to identify all objects in the image and compare the results against a list of animals.
B: Download an object detection model from TensorFlow Hub. Deploy the model to a Vertex AI endpoint. Send new user-submitted images to the model endpoint to classify whether each photo has an animal.
C: Manually label previously submitted images with bounding boxes around any animals. Build an AutoML object detection model by using Vertex AI. Deploy the model to a Vertex AI endpoint Send new user-submitted images to your model endpoint to detect whether each photo has an animal.
D: Manually label previously submitted images as having animals or not. Create an image dataset on Vertex AI. Train a classification model by using Vertex AutoML to distinguish the two classes. Deploy the model to a Vertex AI endpoint. Send new user-submitted images to your model endpoint to classify whether each photo has an animal.

Sample Question 97

You work for a pet food company that manages an online forum. Customers upload photos of their pets on the forum to share with others. About 20 photos are uploaded daily. You want to automatically and in near real time detect whether each uploaded photo has an animal. You want to prioritize time and minimize cost of your application development and deployment. What should you do?

A: Send user-submitted images to the Cloud Vision API. Use object localization to identify all objects in the image and compare the results against a list of animals.

B: Download an object detection model from TensorFlow Hub. Deploy the model to a Vertex AI endpoint. Send new user-submitted images to the model endpoint to classify whether each photo has an animal.

C: Manually label previously submitted images with bounding boxes around any animals. Build an AutoML object detection model by using Vertex AI. Deploy the model to a Vertex AI endpoint Send new user-submitted images to your model endpoint to detect whether each photo has an animal.

D: Manually label previously submitted images as having animals or not. Create an image dataset on Vertex AI. Train a classification model by using Vertex AutoML to distinguish the two classes. Deploy the model to a Vertex AI endpoint. Send new user-submitted images to your model endpoint to classify whether each photo has an animal.

Sample Question 99

You need to develop an image classification model by using a large dataset that contains labeled images in a Cloud Storage bucket. What should you do?

A: Use Vertex AI Pipelines with the Kubeflow Pipelines SDK to create a pipeline that reads the images from Cloud Storage and trains the model.
B: Use Vertex AI Pipelines with TensorFlow Extended (TFX) to create a pipeline that reads the images from Cloud Storage and trains the model.
C: Import the labeled images as a managed dataset in Vertex AI and use AutoML to train the model.
D: Convert the image dataset to a tabular format using Dataflow Load the data into BigQuery and use BigQuery ML to train the model.

Sample Question 99

You need to develop an image classification model by using a large dataset that contains labeled images in a Cloud Storage bucket. What should you do?

A: Use Vertex AI Pipelines with the Kubeflow Pipelines SDK to create a pipeline that reads the images from Cloud Storage and trains the model.
B: Use Vertex AI Pipelines with TensorFlow Extended (TFX) to create a pipeline that reads the images from Cloud Storage and trains the model.
C: Import the labeled images as a managed dataset in Vertex AI and use AutoML to train the model.
D: Convert the image dataset to a tabular format using Dataflow Load the data into BigQuery and use BigQuery ML to train the model.

Sample Question 102

You work for an auto insurance company. You are preparing a proof-of-concept ML application that uses images of damaged vehicles to infer damaged parts. Your team has assembled a set of annotated images from damage claim documents in the company's database. The annotations associated with each image consist of a bounding box for each identified damaged part and the part name. You have been given a sufficient budget to train models on Google Cloud. You need to quickly create an initial model. What should you do?

A: Download a pre-trained object detection model from TensorFlow Hub. Fine-tune the model in Vertex AI Workbench by using the annotated image data.
B: Train an object detection model in AutoML by using the annotated image data.
C: Create a pipeline in Vertex AI Pipelines and configure the AutoMLTrainingJobRunOp component to train a custom object detection model by using the annotated image data.
D: Train an object detection model in Vertex AI custom training by using the annotated image data.

Sample Question 102

You work for an auto insurance company. You are preparing a proof-of-concept ML application that uses images of damaged vehicles to infer damaged parts. Your team has assembled a set of annotated images from damage claim documents in the company's database. The annotations associated with each image consist of a bounding box for each identified damaged part and the part name. You have been given a sufficient budget to train models on Google Cloud. You need to quickly create an initial model. What should you do?

A: Download a pre-trained object detection model from TensorFlow Hub. Fine-tune the model in Vertex AI Workbench by using the annotated image data.

B: Train an object detection model in AutoML by using the annotated image data.

C: Create a pipeline in Vertex AI Pipelines and configure the AutoMLTrainingJobRunOp component to train a custom object detection model by using the annotated image data.

D: Train an object detection model in Vertex AI custom training by using the annotated image data.

Sample Question 107

You work for a hospital that wants to optimize how it schedules operations. You need to create a model that uses the relationship between the number of surgeries scheduled and beds used. You want to predict how many beds will be needed for patients each day in advance based on the scheduled surgeries. You have one year of data for the hospital organized in 365 rows.

The data includes the following variables for each day:
• Number of scheduled surgeries
• Number of beds occupied
• Date

You want to maximize the speed of model development and testing. What should you do?

A: Create a BigQuery table. Use BigQuery ML to build a regression model, with number of beds as the target variable, and number of scheduled surgeries and date features (such as day of week) as the predictors.
B: Create a BigQuery table. Use BigQuery ML to build an ARIMA model, with number of beds as the target variable, and date as the time variable.
C: Create a Vertex AI tabular dataset. Train an AutoML regression model, with number of beds as the target variable, and number of scheduled minor surgeries and date features (such as day of the week) as the predictors.
D: Create a Vertex AI tabular dataset. Train an AutoML regression model, with number of beds as the target variable, and number of scheduled minor surgeries and date features (such as day of the week) as the predictors.

Sample Question 107

You work for a hospital that wants to optimize how it schedules operations. You need to create a model that uses the relationship between the number of surgeries scheduled and beds used. You want to predict how many beds will be needed for patients each day in advance based on the scheduled surgeries. You have one year of data for the hospital organized in 365 rows.

The data includes the following variables for each day:
• Number of scheduled surgeries
• Number of beds occupied
• Date

You want to maximize the speed of model development and testing. What should you do?

A: Create a BigQuery table. Use BigQuery ML to build a regression model, with number of beds as the target variable, and number of scheduled surgeries and date features (such as day of week) as the predictors.
B: Create a BigQuery table. Use BigQuery ML to build an ARIMA model, with number of beds as the target variable, and date as the time variable.
C: Create a Vertex AI tabular dataset. Train an AutoML regression model, with number of beds as the target variable, and number of scheduled minor surgeries and date features (such as day of the week) as the predictors.
D: Create a Vertex AI tabular dataset. Train an AutoML regression model, with number of beds as the target variable, and number of scheduled minor surgeries and date features (such as day of the week) as the predictors.

Sample Question 111

You created a model that uses BigQuery ML to perform linear regression. You need to retrain the model on the cumulative data collected every week. You want to minimize the development effort and the scheduling cost. What should you do

A: Use BigQuery's scheduling service to run the model retraining query periodically.
B: Create a pipeline in Vertex AI Pipelines that executes the retraining query, and use the Cloud Scheduler API to run the query weekly.
C: Use Cloud Scheduler to trigger a Cloud Function every week that runs the query for retraining the model.
D: Use the BigQuery API Connector and Cloud Scheduler to trigger Workflows every week that retrains the model.

Sample Question 111

You created a model that uses BigQuery ML to perform linear regression. You need to retrain the model on the cumulative data collected every week. You want to minimize the development effort and the scheduling cost. What should you do

A: Use BigQuery's scheduling service to run the model retraining query periodically.
B: Create a pipeline in Vertex AI Pipelines that executes the retraining query, and use the Cloud Scheduler API to run the query weekly.
C: Use Cloud Scheduler to trigger a Cloud Function every week that runs the query for retraining the model.
D: Use the BigQuery API Connector and Cloud Scheduler to trigger Workflows every week that retrains the model.

Sample Question 126

You work for an online retailer. Your company has a few thousand short lifecycle products. Your company has five years of sales data stored in BigQuery. You have been asked to build a model that will make monthly sales predictions for each product. You want to use a solution that can be implemented quickly with minimal effort. What should you do?

A: Use Prophet on Vertex AI Training to build a custom model.
B: Use Vertex AI Forecast to build a NN-based model.
C: Use BigQuery ML to build a statistical ARIMA_PLUS model.
D: Use TensorFlow on Vertex AI Training to build a custom model.

Sample Question 126

You work for an online retailer. Your company has a few thousand short lifecycle products. Your company has five years of sales data stored in BigQuery. You have been asked to build a model that will make monthly sales predictions for each product. You want to use a solution that can be implemented quickly with minimal effort. What should you do?

A: Use Prophet on Vertex AI Training to build a custom model.
B: Use Vertex AI Forecast to build a NN-based model.
C: Use BigQuery ML to build a statistical ARIMA_PLUS model.
D: Use TensorFlow on Vertex AI Training to build a custom model.

Sample Question 73

You work for a company that is developing an application to help users with meal planning. You want to use machine learning to scan a corpus of recipes and extract each ingredient (e.g., carrot, rice, pasta) and each kitchen cookware (e.g., bowl, pot, spoon) mentioned. Each recipe is saved in an unstructured text file. What should you do?

A: Create a text dataset on Vertex AI for entity extraction Create two entities called “ingredient” and “cookware”, and label at least 200 examples of each entity. Train an AutoML entity extraction model to extract occurrences of these entity types. Evaluate performance on a holdout dataset.

B: Create a multi-label text classification dataset on Vertex AI. Create a test dataset, and label each recipe that corresponds to its ingredients and cookware. Train a multi-class classification model. Evaluate the model’s performance on a holdout dataset.

C: Use the Entity Analysis method of the Natural Language API to extract the ingredients and cookware from each recipe. Evaluate the model's performance on a prelabeled dataset.

D: Create a text dataset on Vertex AI for entity extraction. Create as many entities as there are different ingredients and cookware. Train an AutoML entity extraction model to extract those entities. Evaluate the model’s performance on a holdout dataset.

Sample Question 73

You work for a company that is developing an application to help users with meal planning. You want to use machine learning to scan a corpus of recipes and extract each ingredient (e.g., carrot, rice, pasta) and each kitchen cookware (e.g., bowl, pot, spoon) mentioned. Each recipe is saved in an unstructured text file. What should you do?

A: Create a text dataset on Vertex AI for entity extraction Create two entities called “ingredient” and “cookware”, and label at least 200 examples of each entity. Train an AutoML entity extraction model to extract occurrences of these entity types. Evaluate performance on a holdout dataset.

B: Create a multi-label text classification dataset on Vertex AI. Create a test dataset, and label each recipe that corresponds to its ingredients and cookware. Train a multi-class classification model. Evaluate the model’s performance on a holdout dataset.

C: Use the Entity Analysis method of the Natural Language API to extract the ingredients and cookware from each recipe. Evaluate the model's performance on a prelabeled dataset.

D: Create a text dataset on Vertex AI for entity extraction. Create as many entities as there are different ingredients and cookware. Train an AutoML entity extraction model to extract those entities. Evaluate the model’s performance on a holdout dataset.

Sample Question 42

You work at an ecommerce startup. You need to create a customer churn prediction model. Your company's recent sales records are stored in a BigQuery table. You want to understand how your initial model is making predictions. You also want to iterate on the model as quickly as possible while minimizing cost. How should you build your first model?

A: Export the data to a Cloud Storage bucket. Load the data into a pandas DataFrame on Vertex AI Workbench and train a logistic regression model with scikit-learn.
B: Create a tf.data.Dataset by using the TensorFlow BigQueryClient. Implement a deep neural network in TensorFlow.
C: Prepare the data in BigQuery and associate the data with a Vertex AI dataset. Create an AutoMLTabularTrainingJob to tram a classification model.
D: Export the data to a Cloud Storage bucket. Create a tf.data.Dataset to read the data from Cloud Storage. Implement a deep neural network in TensorFlow.

Sample Question 42

You work at an ecommerce startup. You need to create a customer churn prediction model. Your company's recent sales records are stored in a BigQuery table. You want to understand how your initial model is making predictions. You also want to iterate on the model as quickly as possible while minimizing cost. How should you build your first model?

A: Export the data to a Cloud Storage bucket. Load the data into a pandas DataFrame on Vertex AI Workbench and train a logistic regression model with scikit-learn.
B: Create a tf.data.Dataset by using the TensorFlow BigQueryClient. Implement a deep neural network in TensorFlow.
C: Prepare the data in BigQuery and associate the data with a Vertex AI dataset. Create an AutoMLTabularTrainingJob to tram a classification model.
D: Export the data to a Cloud Storage bucket. Create a tf.data.Dataset to read the data from Cloud Storage. Implement a deep neural network in TensorFlow.

Sample Question 68

You work for a social media company. You want to create a no-code image classification model for an iOS mobile application to identify fashion accessories. You have a labeled dataset in Cloud Storage. You need to configure a training workflow that minimizes cost and serves predictions with the lowest possible latency. What should you do?

A: Train the model by using AutoML, and register the model in Vertex AI Model Registry. Configure your mobile application to send batch requests during prediction.
B: Train the model by using AutoML Edge, and export it as a Core ML model. Configure your mobile application to use the .mlmodel file directly.
C: Train the model by using AutoML Edge, and export the model as a TFLite model. Configure your mobile application to use the .tflite file directly.
D: Train the model by using AutoML, and expose the model as a Vertex AI endpoint. Configure your mobile application to invoke the endpoint during prediction.

Sample Question 68

You work for a social media company. You want to create a no-code image classification model for an iOS mobile application to identify fashion accessories. You have a labeled dataset in Cloud Storage. You need to configure a training workflow that minimizes cost and serves predictions with the lowest possible latency. What should you do?

A: Train the model by using AutoML, and register the model in Vertex AI Model Registry. Configure your mobile application to send batch requests during prediction.

B: Train the model by using AutoML Edge, and export it as a Core ML model. Configure your mobile application to use the .mlmodel file directly.

C: Train the model by using AutoML Edge, and export the model as a TFLite model. Configure your mobile application to use the .tflite file directly.

D: Train the model by using AutoML, and expose the model as a Vertex AI endpoint. Configure your mobile application to invoke the endpoint during prediction.

Sample Question 24

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

Sample Question 24

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

Sample Question 23

You work for an international manufacturing organization that ships scientific products all over the world. Instruction manuals for these products need to be translated to 15 different languages. Your organization's leadership team wants to start using machine learning to reduce the cost of manual human translations and increase translation speed. You need to implement a scalable solution that maximizes accuracy and minimizes operational overhead. You also want to include a process to evaluate and fix incorrect translations. What should you do?

A: Create a workflow using Cloud Function triggers. Configure a Cloud Function that is triggered when documents are uploaded to an input Cloud Storage bucket. Configure another Cloud Function that translates the documents using the Cloud Translation API, and saves the translations to an output Cloud Storage bucket. Use human reviewers to evaluate the incorrect translations.

B: Create a Vertex AI pipeline that processes the documents launches, an AutoML Translation training job, evaluates the translations and deploys the model to a Vertex AI endpoint with autoscaling and model monitoring. When there is a predetermined skew between training and live data, re-trigger the pipeline with the latest data.

C: Use AutoML Translation to train a model. Configure a Translation Hub project, and use the trained model to translate the documents. Use human reviewers to evaluate the incorrect translations.

D: Use Vertex AI custom training jobs to fine-tune a state-of-the-art open source pretrained model with your data. Deploy the model to a Vertex AI endpoint with autoscaling and model monitoring. When there is a predetermined skew between the training and live data, configure a trigger to run another training job with the latest data.

Sample Question 23

You work for an international manufacturing organization that ships scientific products all over the world. Instruction manuals for these products need to be translated to 15 different languages. Your organization's leadership team wants to start using machine learning to reduce the cost of manual human translations and increase translation speed. You need to implement a scalable solution that maximizes accuracy and minimizes operational overhead. You also want to include a process to evaluate and fix incorrect translations. What should you do?

A: Create a workflow using Cloud Function triggers. Configure a Cloud Function that is triggered when documents are uploaded to an input Cloud Storage bucket. Configure another Cloud Function that translates the documents using the Cloud Translation API, and saves the translations to an output Cloud Storage bucket. Use human reviewers to evaluate the incorrect translations.

B: Create a Vertex AI pipeline that processes the documents launches, an AutoML Translation training job, evaluates the translations and deploys the model to a Vertex AI endpoint with autoscaling and model monitoring. When there is a predetermined skew between training and live data, re-trigger the pipeline with the latest data.

C: Use AutoML Translation to train a model. Configure a Translation Hub project, and use the trained model to translate the documents. Use human reviewers to evaluate the incorrect translations.

D: Use Vertex AI custom training jobs to fine-tune a state-of-the-art open source pretrained model with your data. Deploy the model to a Vertex AI endpoint with autoscaling and model monitoring. When there is a predetermined skew between the training and live data, configure a trigger to run another training job with the latest data.

Sample Question 22

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.

Sample Question 22

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.

Sample Question 11

You are collaborating on a model prototype with your team. You need to create a Vertex AI Workbench environment for the members of your team and also limit access to other employees in your project. What should you do?

A: 1. Create a new service account and grant it the Notebook Viewer role
2. Grant the Service Account User role to each team member on the service account
3. Grant the Vertex AI User role to each team member
4. Provision a Vertex AI Workbench user-managed notebook instance that uses the new service account

B: 1. Grant the Vertex AI User role to the default Compute Engine service account
2. Grant the Service Account User role to each team member on the default Compute Engine service account
3. Provision a Vertex AI Workbench user-managed notebook instance that uses the default Compute Engine service account.

C: 1. Create a new service account and grant it the Vertex AI User role
2. Grant the Service Account User role to each team member on the service account
3. Grant the Notebook Viewer role to each team member.
4. Provision a Vertex AI Workbench user-managed notebook instance that uses the new service account

D: 1. Grant the Vertex AI User role to the primary team member
2. Grant the Notebook Viewer role to the other team members
3. Provision a Vertex AI Workbench user-managed notebook instance that uses the primary user's account

Sample Question 11

You are collaborating on a model prototype with your team. You need to create a Vertex AI Workbench environment for the members of your team and also limit access to other employees in your project. What should you do?

A: 1. Create a new service account and grant it the Notebook Viewer role
2. Grant the Service Account User role to each team member on the service account
3. Grant the Vertex AI User role to each team member
4. Provision a Vertex AI Workbench user-managed notebook instance that uses the new service account

B: 1. Grant the Vertex AI User role to the default Compute Engine service account
2. Grant the Service Account User role to each team member on the default Compute Engine service account
3. Provision a Vertex AI Workbench user-managed notebook instance that uses the default Compute Engine service account.

C: 1. Create a new service account and grant it the Vertex AI User role
2. Grant the Service Account User role to each team member on the service account
3. Grant the Notebook Viewer role to each team member.
4. Provision a Vertex AI Workbench user-managed notebook instance that uses the new service account

D: 1. Grant the Vertex AI User role to the primary team member
2. Grant the Notebook Viewer role to the other team members
3. Provision a Vertex AI Workbench user-managed notebook instance that uses the primary user's account

Sample Question 41

You are training models in Vertex AI by using data that spans across multiple Google Cloud projects. You need to find, track, and compare the performance of the different versions of your models. Which Google Cloud services should you include in your ML workflow?

A: Dataplex, Vertex AI Feature Store, and Vertex AI TensorBoard
B: Vertex AI Pipelines, Vertex AI Feature Store, and Vertex AI Experiments
C: Dataplex, Vertex AI Experiments, and Vertex AI ML Metadata
D: Vertex AI Pipelines, Vertex AI Experiments, and Vertex AI Metadata

Sample Question 41

You are training models in Vertex AI by using data that spans across multiple Google Cloud projects. You need to find, track, and compare the performance of the different versions of your models. Which Google Cloud services should you include in your ML workflow?

A: Dataplex, Vertex AI Feature Store, and Vertex AI TensorBoard
B: Vertex AI Pipelines, Vertex AI Feature Store, and Vertex AI Experiments
C: Dataplex, Vertex AI Experiments, and Vertex AI ML Metadata
D: Vertex AI Pipelines, Vertex AI Experiments, and Vertex AI Metadata

Sample Question 75

You are developing a model to predict whether a failure will occur in a critical machine part. You have a dataset consisting of a multivariate time series and labels indicating whether the machine part failed. You recently started experimenting with a few different preprocessing and modeling approaches in a Vertex AI Workbench notebook. You want to log data and track artifacts from each run. How should you set up your experiments?

A: 1. Use the Vertex AI SDK to create an experiment and set up Vertex ML Metadata.
2. Use the log_time_series_metrics function to track the preprocessed data, and use the log_merrics function to log loss values.

B: 1. Use the Vertex AI SDK to create an experiment and set up Vertex ML Metadata.
2. Use the log_time_series_metrics function to track the preprocessed data, and use the log_metrics function to log loss values.

C: 1. Create a Vertex AI TensorBoard instance and use the Vertex AI SDK to create an experiment and associate the TensorBoard instance.
2. Use the assign_input_artifact method to track the preprocessed data and use the log_time_series_metrics function to log loss values.

D: 1. Create a Vertex AI TensorBoard instance, and use the Vertex AI SDK to create an experiment and associate the TensorBoard instance.
2. Use the log_time_series_metrics function to track the preprocessed data, and use the log_metrics function to log loss values

Sample Question 75

You are developing a model to predict whether a failure will occur in a critical machine part. You have a dataset consisting of a multivariate time series and labels indicating whether the machine part failed. You recently started experimenting with a few different preprocessing and modeling approaches in a Vertex AI Workbench notebook. You want to log data and track artifacts from each run. How should you set up your experiments?

A: 1. Use the Vertex AI SDK to create an experiment and set up Vertex ML Metadata.
2. Use the log_time_series_metrics function to track the preprocessed data, and use the log_merrics function to log loss values.

B: 1. Use the Vertex AI SDK to create an experiment and set up Vertex ML Metadata.
2. Use the log_time_series_metrics function to track the preprocessed data, and use the log_metrics function to log loss values.

C: 1. Create a Vertex AI TensorBoard instance and use the Vertex AI SDK to create an experiment and associate the TensorBoard instance.
2. Use the assign_input_artifact method to track the preprocessed data and use the log_time_series_metrics function to log loss values.

D: 1. Create a Vertex AI TensorBoard instance, and use the Vertex AI SDK to create an experiment and associate the TensorBoard instance.
2. Use the log_time_series_metrics function to track the preprocessed data, and use the log_metrics function to log loss values

Sample Question 86

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 AI 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.

Sample Question 86

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 AI 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.

Sample Question 92

You are working on a prototype of a text classification model in a managed Vertex AI Workbench notebook. You want to quickly experiment with tokenizing text by using a Natural Language Toolkit (NLTK) library. How should you add the library to your Jupyter kernel?

A: Install the NLTK library from a terminal by using the pip install nltk command.
B: Write a custom Dataflow job that uses NLTK to tokenize your text and saves the output to Cloud Storage.
C: Create a new Vertex AI Workbench notebook with a custom image that includes the NLTK library.
D: Install the NLTK library from a Jupyter cell by using the !pip install nltk --user command.

Sample Question 92

You are working on a prototype of a text classification model in a managed Vertex AI Workbench notebook. You want to quickly experiment with tokenizing text by using a Natural Language Toolkit (NLTK) library. How should you add the library to your Jupyter kernel?

A: Install the NLTK library from a terminal by using the pip install nltk command.
B: Write a custom Dataflow job that uses NLTK to tokenize your text and saves the output to Cloud Storage.
C: Create a new Vertex AI Workbench notebook with a custom image that includes the NLTK library.
D: Install the NLTK library from a Jupyter cell by using the !pip install nltk --user command.

Sample Question 98

You work with a team of researchers to develop state-of-the-art algorithms for financial analysis. Your team develops and debugs complex models in TensorFlow. You want to maintain the ease of debugging while also reducing the model training time. How should you set up your training environment?

A: Configure a v3-8 TPU VM. SSH into the VM to train and debug the model.
B: Configure a v3-8 TPU node. Use Cloud Shell to SSH into the Host VM to train and debug the model.
C: Configure a n1 -standard-4 VM with 4 NVIDIA P100 GPUs. SSH into the VM and use ParameterServerStraregv to train the model.
D: Configure a n1-standard-4 VM with 4 NVIDIA P100 GPUs. SSH into the VM and use MultiWorkerMirroredStrategy to train the model.

Sample Question 98

You work with a team of researchers to develop state-of-the-art algorithms for financial analysis. Your team develops and debugs complex models in TensorFlow. You want to maintain the ease of debugging while also reducing the model training time. How should you set up your training environment?

A: Configure a v3-8 TPU VM. SSH into the VM to train and debug the model.
B: Configure a v3-8 TPU node. Use Cloud Shell to SSH into the Host VM to train and debug the model.
C: Configure a n1 -standard-4 VM with 4 NVIDIA P100 GPUs. SSH into the VM and use ParameterServerStraregv to train the model.
D: Configure a n1-standard-4 VM with 4 NVIDIA P100 GPUs. SSH into the VM and use MultiWorkerMirroredStrategy to train the model.

Sample Question 110

You want to migrate a scikit-learn classifier model to TensorFlow. You plan to train the TensorFlow classifier model using the same training set that was used to train the scikit-learn model, and then compare the performances using a common test set. You want to use the Vertex AI Python SDK to manually log the evaluation metrics of each model and compare them based on their F1 scores and confusion matrices. How should you log the metrics?

A: Use the aiplatform.log_classification_metrics function to log the F1 score, and use the aiplatform.log_metrics function to log the confusion matrix.
B: Use the aiplatform.log_classification_metrics function to log the F1 score and the confusion matrix.
C: Use the aiplatform.log_metrics function to log the F1 score and the confusion matrix.
D: Use the aiplatform.log_metrics function to log the F1 score: and use the aiplatform.log_classification_metrics function to log the confusion matrix.

Sample Question 110

You want to migrate a scikit-learn classifier model to TensorFlow. You plan to train the TensorFlow classifier model using the same training set that was used to train the scikit-learn model, and then compare the performances using a common test set. You want to use the Vertex AI Python SDK to manually log the evaluation metrics of each model and compare them based on their F1 scores and confusion matrices. How should you log the metrics?

A: Use the aiplatform.log_classification_metrics function to log the F1 score, and use the aiplatform.log_metrics function to log the confusion matrix.
B: Use the aiplatform.log_classification_metrics function to log the F1 score and the confusion matrix.
C: Use the aiplatform.log_metrics function to log the F1 score and the confusion matrix.
D: Use the aiplatform.log_metrics function to log the F1 score: and use the aiplatform.log_classification_metrics function to log the confusion matrix.

Sample Question 116

You work on a team that builds state-of-the-art deep learning models by using the TensorFlow framework. Your team runs multiple ML experiments each week, which makes it difficult to track the experiment runs. You want a simple approach to effectively track, visualize, and debug ML experiment runs on Google Cloud while minimizing any overhead code. How should you proceed?

A: Set up Vertex AI Experiments to track metrics and parameters. Configure Vertex AI TensorBoard for visualization.
B: Set up a Cloud Function to write and save metrics files to a Cloud Storage bucket. Configure a Google Cloud VM to host TensorBoard locally for visualization.
C: Set up a Vertex AI Workbench notebook instance. Use the instance to save metrics data in a Cloud Storage bucket and to host TensorBoard locally for visualization.
D: Set up a Cloud Function to write and save metrics files to a BigQuery table. Configure a Google Cloud VM to host TensorBoard locally for visualization

Sample Question 116

You work on a team that builds state-of-the-art deep learning models by using the TensorFlow framework. Your team runs multiple ML experiments each week, which makes it difficult to track the experiment runs. You want a simple approach to effectively track, visualize, and debug ML experiment runs on Google Cloud while minimizing any overhead code. How should you proceed?

A: Set up Vertex AI Experiments to track metrics and parameters. Configure Vertex AI TensorBoard for visualization.

B: Set up a Cloud Function to write and save metrics files to a Cloud Storage bucket. Configure a Google Cloud VM to host TensorBoard locally for visualization.

C: Set up a Vertex AI Workbench notebook instance. Use the instance to save metrics data in a Cloud Storage bucket and to host TensorBoard locally for visualization.

D: Set up a Cloud Function to write and save metrics files to a BigQuery table. Configure a Google Cloud VM to host TensorBoard locally for visualization

Sample Question 118

You have created a Vertex AI pipeline that automates custom model training. You want to add a pipeline component that enables your team to most easily collaborate when running different executions and comparing metrics both visually and programmatically. What should you do?

A: Add a component to the Vertex AI pipeline that logs metrics to a BigQuery table. Query the table to compare different executions of the pipeline. Connect BigQuery to Looker Studio to visualize metrics.
B: Add a component to the Vertex AI pipeline that logs metrics to a BigQuery table. Load the table into a pandas DataFrame to compare different executions of the pipeline. Use Matplotlib to visualize metrics.
C: Add a component to the Vertex AI pipeline that logs metrics to Vertex ML Metadata. Use Vertex AI Experiments to compare different executions of the pipeline. Use Vertex AI TensorBoard to visualize metrics.
D: Add a component to the Vertex AI pipeline that logs metrics to Vertex ML Metadata. Load the Vertex ML Metadata into a pandas DataFrame to compare different executions of the pipeline. Use Matplotlib to visualize metrics.

Sample Question 118

You have created a Vertex AI pipeline that automates custom model training. You want to add a pipeline component that enables your team to most easily collaborate when running different executions and comparing metrics both visually and programmatically. What should you do?

A: Add a component to the Vertex AI pipeline that logs metrics to a BigQuery table. Query the table to compare different executions of the pipeline. Connect BigQuery to Looker Studio to visualize metrics.

B: Add a component to the Vertex AI pipeline that logs metrics to a BigQuery table. Load the table into a pandas DataFrame to compare different executions of the pipeline. Use Matplotlib to visualize metrics.

C: Add a component to the Vertex AI pipeline that logs metrics to Vertex ML Metadata. Use Vertex AI Experiments to compare different executions of the pipeline. Use Vertex AI TensorBoard to visualize metrics.

D: Add a component to the Vertex AI pipeline that logs metrics to Vertex ML Metadata. Load the Vertex ML Metadata into a pandas DataFrame to compare different executions of the pipeline. Use Matplotlib to visualize metrics.

Sample Question 119

You are investigating the root cause of a misclassification error made by one of your models. You used Vertex AI Pipelines to train and deploy the model. The pipeline reads data from BigQuery. creates a copy of the data in Cloud Storage in TFRecord format, trains the model in Vertex AI Training on that copy, and deploys the model to a Vertex AI endpoint. You have identified the specific version of that model that misclassified, and you need to recover the data this model was trained on. How should you find that copy of the data?

A: Use Vertex AI Feature Store. Modify the pipeline to use the feature store, and ensure that all training data is stored in it. Search the feature store for the data used for the training.
B: Use the lineage feature of Vertex AI Metadata to find the model artifact. Determine the version of the model and identify the step that creates the data copy and search in the metadata for its location.
C: Use the logging features in the Vertex AI endpoint to determine the timestamp of the model's deployment. Find the pipeline run at that timestamp. Identify the step that creates the data copy, and search in the logs for its location.
D: Find the job ID in Vertex AI Training corresponding to the training for the model. Search in the logs of that job for the data used for the training.

Sample Question 119

You are investigating the root cause of a misclassification error made by one of your models. You used Vertex AI Pipelines to train and deploy the model. The pipeline reads data from BigQuery. creates a copy of the data in Cloud Storage in TFRecord format, trains the model in Vertex AI Training on that copy, and deploys the model to a Vertex AI endpoint. You have identified the specific version of that model that misclassified, and you need to recover the data this model was trained on. How should you find that copy of the data?

A: Use Vertex AI Feature Store. Modify the pipeline to use the feature store, and ensure that all training data is stored in it. Search the feature store for the data used for the training.

B: Use the lineage feature of Vertex AI Metadata to find the model artifact. Determine the version of the model and identify the step that creates the data copy and search in the metadata for its location.

C: Use the logging features in the Vertex AI endpoint to determine the timestamp of the model's deployment. Find the pipeline run at that timestamp. Identify the step that creates the data copy, and search in the logs for its location.

D: Find the job ID in Vertex AI Training corresponding to the training for the model. Search in the logs of that job for the data used for the training.

Sample Question 122

You need to use TensorFlow to train an image classification model. Your dataset is located in a Cloud Storage directory and contains millions of labeled images. Before training the model, you need to prepare the data. You want the data preprocessing and model training workflow to be as efficient, scalable, and low maintenance as possible. What should you do?

A: 1. Create a Dataflow job that creates sharded TFRecord files in a Cloud Storage directory.
2. Reference tf.data.TFRecordDataset in the training script.
3. Train the model by using Vertex AI Training with a V100 GPU.

B: 1. Create a Dataflow job that moves the images into multiple Cloud Storage directories, where each directory is named according to the corresponding label
2. Reference tfds.folder_dataset:ImageFolder in the training script.
3. Train the model by using Vertex AI Training with a V100 GPU.

C: 1. Create a Jupyter notebook that uses an nt-standard-64 V100 GPU Vertex AI Workbench instance.
2. Write a Python script that creates sharded TFRecord files in a directory inside the instance.
3. Reference tf.data.TFRecordDataset in the training script.
4. Train the model by using the Workbench instance.

D: 1. Create a Jupyter notebook that uses an n1-standard-64, V100 GPU Vertex AI Workbench instance.
2. Write a Python script that copies the images into multiple Cloud Storage directories, where each. directory is named according to the corresponding label.
3. Reference tfds.foladr_dataset.ImageFolder in the training script.
4. Train the model by using the Workbench instance.

Sample Question 122

You need to use TensorFlow to train an image classification model. Your dataset is located in a Cloud Storage directory and contains millions of labeled images. Before training the model, you need to prepare the data. You want the data preprocessing and model training workflow to be as efficient, scalable, and low maintenance as possible. What should you do?

A: 1. Create a Dataflow job that creates sharded TFRecord files in a Cloud Storage directory.
2. Reference tf.data.TFRecordDataset in the training script.
3. Train the model by using Vertex AI Training with a V100 GPU.

B: 1. Create a Dataflow job that moves the images into multiple Cloud Storage directories, where each directory is named according to the corresponding label
2. Reference tfds.folder_dataset:ImageFolder in the training script.
3. Train the model by using Vertex AI Training with a V100 GPU.

C: 1. Create a Jupyter notebook that uses an nt-standard-64 V100 GPU Vertex AI Workbench instance.
2. Write a Python script that creates sharded TFRecord files in a directory inside the instance.
3. Reference tf.data.TFRecordDataset in the training script.
4. Train the model by using the Workbench instance.

D: 1. Create a Jupyter notebook that uses an n1-standard-64, V100 GPU Vertex AI Workbench instance.
2. Write a Python script that copies the images into multiple Cloud Storage directories, where each. directory is named according to the corresponding label.
3. Reference tfds.foladr_dataset.ImageFolder in the training script.
4. Train the model by using the Workbench instance.

Sample Question 124

You are pre-training a large language model on Google Cloud. This model includes custom TensorFlow operations in the training loop. Model training will use a large batch size, and you expect training to take several weeks. You need to configure a training architecture that minimizes both training time and compute costs. What should you do?

A: Implement 8 workers of a2-megagpu-16g machines by using tf.distribute.MultiWorkerMirroredStrategy.
B: Implement a TPU Pod slice with -accelerator-type=v4-l28 by using tf.distribute.TPUStrategy.
C: Implement 16 workers of c2d-highcpu-32 machines by using tf.distribute.MirroredStrategy.
D: Implement 16 workers of a2-highgpu-8g machines by using tf.distribute.MultiWorkerMirroredStrategy.

Sample Question 124

You are pre-training a large language model on Google Cloud. This model includes custom TensorFlow operations in the training loop. Model training will use a large batch size, and you expect training to take several weeks. You need to configure a training architecture that minimizes both training time and compute costs. What should you do?

A: Implement 8 workers of a2-megagpu-16g machines by using tf.distribute.MultiWorkerMirroredStrategy.
B: Implement a TPU Pod slice with -accelerator-type=v4-l28 by using tf.distribute.TPUStrategy.
C: Implement 16 workers of c2d-highcpu-32 machines by using tf.distribute.MirroredStrategy.
D: Implement 16 workers of a2-highgpu-8g machines by using tf.distribute.MultiWorkerMirroredStrategy.

Sample Question 1

You are training a custom language model for your company using a large dataset. You plan to use the Reduction Server strategy on Vertex AI.
You need to configure the worker pools of the distributed training job. What should you do?

A: Configure the machines of the first two worker pools to have GPUs, and to use a container image where your training code runs. Configure
the third worker pool to have GPUs, and use the reductionserver container image.
B: Configure the machines of the first two worker pools to have GPUs and to use a container image where your training code runs. Configure the third worker pool to use the reductionserver container image without accelerators, and choose a machine type that prioritizes bandwidth.
C: Configure the machines of the first two worker pools to have TPUs and to use a container image where your training code runs. Configure the third worker pool without accelerators, and use the reductionserver container image without accelerators, and choose a machine type that
prioritizes bandwidth.
D: Configure the machines of the first two pools to have TPUs, and to use a container image where your training code runs. Configure the third
pool to have TPUs, and use the reductionserver container image.

Sample Question 1

You are training a custom language model for your company using a large dataset. You plan to use the Reduction Server strategy on Vertex AI.
You need to configure the worker pools of the distributed training job. What should you do?

A: Configure the machines of the first two worker pools to have GPUs, and to use a container image where your training code runs. Configure
the third worker pool to have GPUs, and use the reductionserver container image.

B: Configure the machines of the first two worker pools to have GPUs and to use a container image where your training code runs. Configure the third worker pool to use the reductionserver container image without accelerators, and choose a machine type that prioritizes bandwidth.

C: Configure the machines of the first two worker pools to have TPUs and to use a container image where your training code runs. Configure the third worker pool without accelerators, and use the reductionserver container image without accelerators, and choose a machine type that
prioritizes bandwidth.

D: Configure the machines of the first two pools to have TPUs, and to use a container image where your training code runs. Configure the third
pool to have TPUs, and use the reductionserver container image.

Sample Question 132

You are training an ML model on a large dataset. You are using a TPU to accelerate the training process. You notice that the training process is taking longer than expected. You discover that the TPU is not reaching its full capacity. What should you do?

A: Increase the learning rate
B: Increase the number of epochs
C: Decrease the learning rate
D: Increase the batch size

Sample Question 132

You are training an ML model on a large dataset. You are using a TPU to accelerate the training process. You notice that the training process is taking longer than expected. You discover that the TPU is not reaching its full capacity. What should you do?

A: Increase the learning rate
B: Increase the number of epochs
C: Decrease the learning rate
D: Increase the batch size

Sample Question 133

You created an ML pipeline with multiple input parameters. You want to investigate the tradeoffs between different parameter combinations. The parameter options are
• Input dataset
• Max tree depth of the boosted tree regressor
• Optimizer learning rate

You need to compare the pipeline performance of the different parameter combinations measured in F1 score, time to train, and model complexity. You want your approach to be reproducible, and track all pipeline runs on the same platform. What should you do?

A: 1. Use BigQueryML to create a boosted tree regressor, and use the hyperparameter tuning capability.
2. Configure the hyperparameter syntax to select different input datasets: max tree depths, and optimizer learning rates. Choose the grid search option.
B: 1. Create a Vertex AI pipeline with a custom model training job as part of the pipeline. Configure the pipeline's parameters to include those you are investigating.
2. In the custom training step, use the Bayesian optimization method with F1 score as the target to maximize.
C: 1. Create a Vertex AI Workbench notebook for each of the different input datasets.
2. In each notebook, run different local training jobs with different combinations of the max tree depth and optimizer learning rate parameters.
3. After each notebook finishes, append the results to a BigQuery table.
D: 1. Create an experiment in Vertex AI Experiments.
2. Create a Vertex AI pipeline with a custom model training job as part of the pipeline. Configure the pipeline's parameters to include those you are investigating.

Sample Question 133

You created an ML pipeline with multiple input parameters. You want to investigate the tradeoffs between different parameter combinations. The parameter options are
• Input dataset
• Max tree depth of the boosted tree regressor
• Optimizer learning rate

You need to compare the pipeline performance of the different parameter combinations measured in F1 score, time to train, and model complexity. You want your approach to be reproducible, and track all pipeline runs on the same platform. What should you do?

A: 1. Use BigQueryML to create a boosted tree regressor, and use the hyperparameter tuning capability.
2. Configure the hyperparameter syntax to select different input datasets: max tree depths, and optimizer learning rates. Choose the grid search option.
B: 1. Create a Vertex AI pipeline with a custom model training job as part of the pipeline. Configure the pipeline's parameters to include those you are investigating.
2. In the custom training step, use the Bayesian optimization method with F1 score as the target to maximize.
C: 1. Create a Vertex AI Workbench notebook for each of the different input datasets.
2. In each notebook, run different local training jobs with different combinations of the max tree depth and optimizer learning rate parameters.
3. After each notebook finishes, append the results to a BigQuery table.
D: 1. Create an experiment in Vertex AI Experiments.
2. Create a Vertex AI pipeline with a custom model training job as part of the pipeline. Configure the pipeline's parameters to include those you are investigating.

Sample Question 72

You are developing an ML model in a Vertex AI Workbench notebook. You want to track artifacts and compare models during experimentation using different approaches. You need to rapidly and easily transition successful experiments to production as you iterate on your model implementation. What should you do?

A: 1. Initialize the Vertex SDK with the name of your experiment. Log parameters and metrics for each experiment, and attach dataset and model artifacts as inputs and outputs to each execution.
2. After a successful experiment create a Vertex AI pipeline.

B: 1. Initialize the Vertex SDK with the name of your experiment. Log parameters and metrics for each experiment, save your dataset to a Cloud Storage bucket, and upload the models to Vertex AI Model Registry.
2. After a successful experiment, create a Vertex AI pipeline.

C: 1. Create a Vertex AI pipeline with parameters you want to track as arguments to your PipelineJob. Use the Metrics, Model, and Dataset artifact types from the Kubeflow Pipelines DSL as the inputs and outputs of the components in your pipeline.
2. Associate the pipeline with your experiment when you submit the job.

D: 1. Create a Vertex AI pipeline. Use the Dataset and Model artifact types from the Kubeflow Pipelines DSL as the inputs and outputs of the components in your pipeline.
2. In your training component, use the Vertex AI SDK to create an experiment run. Configure the log_params and log_metrics functions to track parameters and metrics of your experiment.

Sample Question 72

You are developing an ML model in a Vertex AI Workbench notebook. You want to track artifacts and compare models during experimentation using different approaches. You need to rapidly and easily transition successful experiments to production as you iterate on your model implementation. What should you do?

A: 1. Initialize the Vertex SDK with the name of your experiment. Log parameters and metrics for each experiment, and attach dataset and model artifacts as inputs and outputs to each execution.
2. After a successful experiment create a Vertex AI pipeline.

B: 1. Initialize the Vertex SDK with the name of your experiment. Log parameters and metrics for each experiment, save your dataset to a Cloud Storage bucket, and upload the models to Vertex AI Model Registry.
2. After a successful experiment, create a Vertex AI pipeline.

C: 1. Create a Vertex AI pipeline with parameters you want to track as arguments to your PipelineJob. Use the Metrics, Model, and Dataset artifact types from the Kubeflow Pipelines DSL as the inputs and outputs of the components in your pipeline.
2. Associate the pipeline with your experiment when you submit the job.

D: 1. Create a Vertex AI pipeline. Use the Dataset and Model artifact types from the Kubeflow Pipelines DSL as the inputs and outputs of the components in your pipeline.
2. In your training component, use the Vertex AI SDK to create an experiment run. Configure the log_params and log_metrics functions to track parameters and metrics of your experiment.

Sample Question 26

Your team is training a large number of ML models that use different algorithms, parameters, and datasets. Some models are trained in Vertex AI Pipelines, and some are trained on Vertex AI Workbench notebook instances. Your team wants to compare the performance of the models across both services. You want to minimize the effort required to store the parameters and metrics. What should you do?

A: Implement an additional step for all the models running in pipelines and notebooks to export parameters and metrics to BigQuery.
B: Create a Vertex AI experiment. Submit all the pipelines as experiment runs. For models trained on notebooks log parameters and metrics by using the Vertex AI SDK.
C: Implement all models in Vertex AI Pipelines Create a Vertex AI experiment, and associate all pipeline runs with that experiment.
D: Store all model parameters and metrics as model metadata by using the Vertex AI Metadata API.

Sample Question 26

Your team is training a large number of ML models that use different algorithms, parameters, and datasets. Some models are trained in Vertex AI Pipelines, and some are trained on Vertex AI Workbench notebook instances. Your team wants to compare the performance of the models across both services. You want to minimize the effort required to store the parameters and metrics. What should you do?

A: Implement an additional step for all the models running in pipelines and notebooks to export parameters and metrics to BigQuery.

B: Create a Vertex AI experiment. Submit all the pipelines as experiment runs. For models trained on notebooks log parameters and metrics by using the Vertex AI SDK.

C: Implement all models in Vertex AI Pipelines Create a Vertex AI experiment, and associate all pipeline runs with that experiment.

D: Store all model parameters and metrics as model metadata by using the Vertex AI Metadata API.

Sample Question 16

You are training models in Vertex AI by using data that spans across multiple Google Cloud projects. You need to find, track, and compare the performance of the different versions of your models. Which Google Cloud services should you include in your ML workflow?

A: Dataplex, Vertex AI Feature Store, and Vertex AI TensorBoard
B: Vertex AI Pipelines, Vertex AI Feature Store, and Vertex AI Experiments
C: Dataplex, Vertex AI Experiments, and Vertex AI ML Metadata
D: Vertex AI Pipelines, Vertex AI Experiments, and Vertex AI Metadata

Sample Question 16

You are training models in Vertex AI by using data that spans across multiple Google Cloud projects. You need to find, track, and compare the performance of the different versions of your models. Which Google Cloud services should you include in your ML workflow?

A: Dataplex, Vertex AI Feature Store, and Vertex AI TensorBoard
B: Vertex AI Pipelines, Vertex AI Feature Store, and Vertex AI Experiments
C: Dataplex, Vertex AI Experiments, and Vertex AI ML Metadata
D: Vertex AI Pipelines, Vertex AI Experiments, and Vertex AI Metadata

Sample Question 52

You have a custom job that runs on Vertex AI on a weekly basis. The job is implemented using a proprietary ML workflow that produces the datasets, models, and custom artifacts, and sends them to a Cloud Storage bucket. Many different versions of the datasets and models were created. Due to compliance requirements, your company needs to track which model was used for making a particular prediction, and needs access to the artifacts for each model. How should you configure your workflows to meet these requirements?

A: Use the Vertex AI Metadata API inside the custom job to create context, execution, and artifacts for each model, and use events to link them together.
B: Create a Vertex AI experiment, and enable autologging inside the custom job.
C: Configure a TensorFlow Extended (TFX) ML Metadata database, and use the ML Metadata API.
D: Register each model in Vertex AI Model Registry, and use model labels to store the related dataset and model information.

Sample Question 52

You have a custom job that runs on Vertex AI on a weekly basis. The job is implemented using a proprietary ML workflow that produces the datasets, models, and custom artifacts, and sends them to a Cloud Storage bucket. Many different versions of the datasets and models were created. Due to compliance requirements, your company needs to track which model was used for making a particular prediction, and needs access to the artifacts for each model. How should you configure your workflows to meet these requirements?

A: Use the Vertex AI Metadata API inside the custom job to create context, execution, and artifacts for each model, and use events to link them together.

B: Create a Vertex AI experiment, and enable autologging inside the custom job.

C: Configure a TensorFlow Extended (TFX) ML Metadata database, and use the ML Metadata API.

D: Register each model in Vertex AI Model Registry, and use model labels to store the related dataset and model information.

Sample Question 48

You manage a team of data scientists who use a cloud-based backend system to submit training jobs. This system has become very difficult to administer, and you want to use a managed service instead. The data scientists you work with use many different frameworks, including Keras, PyTorch, theano, scikit-learn, and custom libraries. What should you do?

A: Use the Vertex AI Training to submit training jobs using any framework.
B: Configure Kubeflow to run on Google Kubernetes Engine and submit training jobs through TFJob.
C: Create a library of VM images on Compute Engine, and publish these images on a centralized repository.
D: Set up Slurm workload manager to receive jobs that can be scheduled to run on your cloud infrastructure.

Sample Question 48

You manage a team of data scientists who use a cloud-based backend system to submit training jobs. This system has become very difficult to administer, and you want to use a managed service instead. The data scientists you work with use many different frameworks, including Keras, PyTorch, theano, scikit-learn, and custom libraries. What should you do?

A: Use the Vertex AI Training to submit training jobs using any framework.
B: Configure Kubeflow to run on Google Kubernetes Engine and submit training jobs through TFJob.
C: Create a library of VM images on Compute Engine, and publish these images on a centralized repository.
D: Set up Slurm workload manager to receive jobs that can be scheduled to run on your cloud infrastructure.

Sample Question 37

You built a custom ML model using scikit-learn. Training time is taking longer than expected. You decide to migrate your model to Vertex AI Training, and you want to improve the model's training time. What should you try out first?

A: Train your model in a distributed mode using multiple Compute Engine VMs.
B: Train your model using Vertex AI Training with CPUs.
C: Migrate your model to TensorFlow, and train it using Vertex AI Training.
D: Train your model using Vertex AI Training with GPUs.

Sample Question 37

You built a custom ML model using scikit-learn. Training time is taking longer than expected. You decide to migrate your model to Vertex AI Training, and you want to improve the model's training time. What should you try out first?

A: Train your model in a distributed mode using multiple Compute Engine VMs.
B: Train your model using Vertex AI Training with CPUs.
C: Migrate your model to TensorFlow, and train it using Vertex AI Training.
D: Train your model using Vertex AI Training with GPUs.

Sample Question 56

You need to train an XGBoost model on a small dataset. Your training code requires custom dependencies. You want to minimize the startup time of your training job. How should you set up your Vertex AI custom training job?

A: Store the data in a Cloud Storage bucket, and create a custom container with your training application. In your training application, read the data from Cloud Storage and train the model.
B: Use the XGBoost prebuilt custom container. Create a Python source distribution that includes the data and installs the dependencies at runtime. In your training application, load the data into a pandas DataFrame and train the model.
C: Create a custom container that includes the data. In your training application, load the data into a pandas DataFrame and train the model.
D: Store the data in a Cloud Storage bucket, and use the XGBoost prebuilt custom container to run your training application. Create a Python source distribution that installs the dependencies at runtime. In your training application, read the data from Cloud Storage and train the model.

Sample Question 56

You need to train an XGBoost model on a small dataset. Your training code requires custom dependencies. You want to minimize the startup time of your training job. How should you set up your Vertex AI custom training job?

A: Store the data in a Cloud Storage bucket, and create a custom container with your training application. In your training application, read the data from Cloud Storage and train the model.

B: Use the XGBoost prebuilt custom container. Create a Python source distribution that includes the data and installs the dependencies at runtime. In your training application, load the data into a pandas DataFrame and train the model.

C: Create a custom container that includes the data. In your training application, load the data into a pandas DataFrame and train the model.

D: Store the data in a Cloud Storage bucket, and use the XGBoost prebuilt custom container to run your training application. Create a Python source distribution that installs the dependencies at runtime. In your training application, read the data from Cloud Storage and train the model.

Sample Question 51

You have recently developed a custom model for image classification by using a neural network. You need to automatically identify the values for learning rate, number of layers, and kernel size. To do this, you plan to run multiple jobs in parallel to identify the parameters that optimize performance. You want to minimize custom code development and infrastructure management. What should you do?

A: Train an AutoML image classification model.
B: Create a custom training job that uses the Vertex AI Vizier SDK for parameter optimization.
C: Create a Vertex AI hyperparameter tuning job.
D: Create a Vertex AI pipeline that runs different model training jobs in parallel.

Sample Question 51

You have recently developed a custom model for image classification by using a neural network. You need to automatically identify the values for learning rate, number of layers, and kernel size. To do this, you plan to run multiple jobs in parallel to identify the parameters that optimize performance. You want to minimize custom code development and infrastructure management. What should you do?

A: Train an AutoML image classification model.
B: Create a custom training job that uses the Vertex AI Vizier SDK for parameter optimization.
C: Create a Vertex AI hyperparameter tuning job.
D: Create a Vertex AI pipeline that runs different model training jobs in parallel.

Sample Question 58

You work for a rapidly growing social media company. Your team builds TensorFlow recommender models in an on-premises CPU cluster. The data contains billions of historical user events and 100,000 categorical features. You notice that as the data increases, the model training time increases. You plan to move the models to Google Cloud. You want to use the most scalable approach that also minimizes training time. What should you do?

A: Deploy the training jobs by using TPU VMs with TPUv3 Pod slices, and use the TPUEmbeading API
B: Deploy the training jobs in an autoscaling Google Kubernetes Engine cluster with CPUs
C: Deploy a matrix factorization model training job by using BigQuery ML
D: Deploy the training jobs by using Compute Engine instances with A100 GPUs, and use the tf.nn.embedding_lookup API

Sample Question 58

You work for a rapidly growing social media company. Your team builds TensorFlow recommender models in an on-premises CPU cluster. The data contains billions of historical user events and 100,000 categorical features. You notice that as the data increases, the model training time increases. You plan to move the models to Google Cloud. You want to use the most scalable approach that also minimizes training time. What should you do?

A: Deploy the training jobs by using TPU VMs with TPUv3 Pod slices, and use the TPUEmbeading API
B: Deploy the training jobs in an autoscaling Google Kubernetes Engine cluster with CPUs
C: Deploy a matrix factorization model training job by using BigQuery ML
D: Deploy the training jobs by using Compute Engine instances with A100 GPUs, and use the tf.nn.embedding_lookup API

Sample Question 30

You are developing a recommendation engine for an online clothing store. The historical customer transaction data is stored in BigQuery and Cloud Storage. You need to perform exploratory data analysis (EDA), preprocessing and model training. You plan to rerun these EDA, preprocessing, and training steps as you experiment with different types of algorithms. You want to minimize the cost and development effort of running these steps as you experiment. How should you configure the environment?

A: Create a Vertex AI Workbench user-managed notebook using the default VM instance, and use the %%bigquerv magic commands in Jupyter to query the tables.
B: Create a Vertex AI Workbench managed notebook to browse and query the tables directly from the JupyterLab interface.
C: Create a Vertex AI Workbench user-managed notebook on a Dataproc Hub, and use the %%bigquery magic commands in Jupyter to query the tables.
D: Create a Vertex AI Workbench managed notebook on a Dataproc cluster, and use the spark-bigquery-connector to access

Sample Question 30

You are developing a recommendation engine for an online clothing store. The historical customer transaction data is stored in BigQuery and Cloud Storage. You need to perform exploratory data analysis (EDA), preprocessing and model training. You plan to rerun these EDA, preprocessing, and training steps as you experiment with different types of algorithms. You want to minimize the cost and development effort of running these steps as you experiment. How should you configure the environment?

A: Create a Vertex AI Workbench user-managed notebook using the default VM instance, and use the %%bigquerv magic commands in Jupyter to query the tables.

B: Create a Vertex AI Workbench managed notebook to browse and query the tables directly from the JupyterLab interface.

C: Create a Vertex AI Workbench user-managed notebook on a Dataproc Hub, and use the %%bigquery magic commands in Jupyter to query the tables.

D: Create a Vertex AI Workbench managed notebook on a Dataproc cluster, and use the spark-bigquery-connector to access

Sample Question 35

You have trained a DNN regressor with TensorFlow to predict housing prices using a set of predictive features. Your default precision is tf.float64, and you use a standard TensorFlow estimator. Your model performs well, but just before deploying it to production, you discover that your current serving latency is 10ms @ 90 percentile and you currently serve on CPUs. Your production requirements expect a model latency of 8ms @ 90 percentile. You're willing to accept a small decrease in performance in order to reach the latency requirement.
Therefore your plan is to improve latency while evaluating how much the model's prediction decreases. What should you first try to quickly lower the serving latency?

A: Switch from CPU to GPU serving.
B: Apply quantization to your SavedModel by reducing the floating point precision to tf.float16.
C: Increase the dropout rate to 0.8 and retrain your model.
D: Increase the dropout rate to 0.8 in _PREDICT mode by adjusting the TensorFlow Serving parameters.

Sample Question 35

You have trained a DNN regressor with TensorFlow to predict housing prices using a set of predictive features. Your default precision is tf.float64, and you use a standard TensorFlow estimator. Your model performs well, but just before deploying it to production, you discover that your current serving latency is 10ms @ 90 percentile and you currently serve on CPUs. Your production requirements expect a model latency of 8ms @ 90 percentile. You're willing to accept a small decrease in performance in order to reach the latency requirement.
Therefore your plan is to improve latency while evaluating how much the model's prediction decreases. What should you first try to quickly lower the serving latency?

A: Switch from CPU to GPU serving.
B: Apply quantization to your SavedModel by reducing the floating point precision to tf.float16.
C: Increase the dropout rate to 0.8 and retrain your model.
D: Increase the dropout rate to 0.8 in _PREDICT mode by adjusting the TensorFlow Serving parameters.

Sample Question 29

You developed a Transformer model in TensorFlow to translate text. Your training data includes millions of documents in a Cloud Storage bucket. You plan to use distributed training to reduce training time. You need to configure the training job while minimizing the effort required to modify code and to manage the cluster's configuration. What should you do?

A: Create a Vertex AI custom training job with GPU accelerators for the second worker pool. Use tf.distribute.MultiWorkerMirroredStrategy for distribution.
B: Create a Vertex AI custom distributed training job with Reduction Server. Use N1 high-memory machine type instances for the first and second pools, and use N1 high-CPU machine type instances for the third worker pool.
C: Create a training job that uses Cloud TPU VMs. Use tf.distribute.TPUStrategy for distribution.
D: Create a Vertex AI custom training job with a single worker pool of A2 GPU machine type instances. Use tf.distribute.MirroredStrategy for distribution.

Sample Question 29

You developed a Transformer model in TensorFlow to translate text. Your training data includes millions of documents in a Cloud Storage bucket. You plan to use distributed training to reduce training time. You need to configure the training job while minimizing the effort required to modify code and to manage the cluster's configuration. What should you do?

A: Create a Vertex AI custom training job with GPU accelerators for the second worker pool. Use tf.distribute.MultiWorkerMirroredStrategy for distribution.

B: Create a Vertex AI custom distributed training job with Reduction Server. Use N1 high-memory machine type instances for the first and second pools, and use N1 high-CPU machine type instances for the third worker pool.

C: Create a training job that uses Cloud TPU VMs. Use tf.distribute.TPUStrategy for distribution.

D: Create a Vertex AI custom training job with a single worker pool of A2 GPU machine type instances. Use tf.distribute.MirroredStrategy for distribution.

Sample Question 44

You are developing an ML model to identify your company's products in images. You have access to over one million images in a Cloud Storage bucket. You plan to experiment with different TensorFlow models by using Vertex AI Training. You need to read images at scale during training while minimizing data I/O bottlenecks. What should you do?

A: Load the images directly into the Vertex AI compute nodes by using Cloud Storage FUSE. Read the images by using the tf.data.Dataset.from_tensor_slices function
B: Create a Vertex AI managed dataset from your image data. Access the AIP_TRAINING_DATA_URI environment variable to read the images by using the tf.data.Dataset.list_files function.
C: Convert the images to TFRecords and store them in a Cloud Storage bucket. Read the TFRecords by using the tf.data.TFRecordDataset function.
D: Store the URLs of the images in a CSV file. Read the file by using the tf.data.experimental.CsvDataset function.

Sample Question 44

You are developing an ML model to identify your company's products in images. You have access to over one million images in a Cloud Storage bucket. You plan to experiment with different TensorFlow models by using Vertex AI Training. You need to read images at scale during training while minimizing data I/O bottlenecks. What should you do?

A: Load the images directly into the Vertex AI compute nodes by using Cloud Storage FUSE. Read the images by using the tf.data.Dataset.from_tensor_slices function

B: Create a Vertex AI managed dataset from your image data. Access the AIP_TRAINING_DATA_URI environment variable to read the images by using the tf.data.Dataset.list_files function.

C: Convert the images to TFRecords and store them in a Cloud Storage bucket. Read the TFRecords by using the tf.data.TFRecordDataset function.

D: Store the URLs of the images in a CSV file. Read the file by using the tf.data.experimental.CsvDataset function.

Sample Question 34

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.

Sample Question 34

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.

Sample Question 39

You work for a bank. You have been asked to develop an ML model that will support loan application decisions. You need to determine which Vertex AI services to include in the workflow. You want to track the model's training parameters and the metrics per training epoch. You plan to compare the performance of each version of the model to determine the best model based on your chosen metrics. Which Vertex AI services should you use?

A: Vertex ML Metadata, Vertex AI Feature Store, and Vertex AI Vizier
B: Vertex AI Pipelines, Vertex AI Experiments, and Vertex AI Vizier
C: Vertex ML Metadata, Vertex AI Experiments, and Vertex AI TensorBoard
D: Vertex AI Pipelines, Vertex AI Feature Store, and Vertex AI TensorBoard

Sample Question 39

You work for a bank. You have been asked to develop an ML model that will support loan application decisions. You need to determine which Vertex AI services to include in the workflow. You want to track the model's training parameters and the metrics per training epoch. You plan to compare the performance of each version of the model to determine the best model based on your chosen metrics. Which Vertex AI services should you use?

A: Vertex ML Metadata, Vertex AI Feature Store, and Vertex AI Vizier
B: Vertex AI Pipelines, Vertex AI Experiments, and Vertex AI Vizier
C: Vertex ML Metadata, Vertex AI Experiments, and Vertex AI TensorBoard
D: Vertex AI Pipelines, Vertex AI Feature Store, and Vertex AI TensorBoard

Sample Question 101

You recently created a new Google Cloud project. After testing that you can submit a Vertex AI Pipeline job from the Cloud Shell, you want to use a Vertex AI Workbench user-managed notebook instance to run your code from that instance. You created the instance and ran the code but this time the job fails with an insufficient permissions error. What should you do?

A: Ensure that the Workbench instance that you created is in the same region of the Vertex AI Pipelines resources you will use.
B: Ensure that the Vertex AI Workbench instance is on the same subnetwork of the Vertex AI Pipeline resources that you will use.
C: Ensure that the Vertex AI Workbench instance is assigned the Identity and Access Management (IAM) Vertex AI User role.
D: Ensure that the Vertex AI Workbench instance is assigned the Identity and Access Management (IAM) Notebooks Runner role.

Sample Question 101

You recently created a new Google Cloud project. After testing that you can submit a Vertex AI Pipeline job from the Cloud Shell, you want to use a Vertex AI Workbench user-managed notebook instance to run your code from that instance. You created the instance and ran the code but this time the job fails with an insufficient permissions error. What should you do?

A: Ensure that the Workbench instance that you created is in the same region of the Vertex AI Pipelines resources you will use.
B: Ensure that the Vertex AI Workbench instance is on the same subnetwork of the Vertex AI Pipeline resources that you will use.
C: Ensure that the Vertex AI Workbench instance is assigned the Identity and Access Management (IAM) Vertex AI User role.
D: Ensure that the Vertex AI Workbench instance is assigned the Identity and Access Management (IAM) Notebooks Runner role.

Sample Question 121

You are developing an ML pipeline using Vertex AI Pipelines. You want your pipeline to upload a new version of the XGBoost model to Vertex AI Model Registry and deploy it to Vertex AI Endpoints for online inference. You want to use the simplest approach. What should you do?

A: Use the Vertex AI REST API within a custom component based on a vertex-ai/prediction/xgboost-cpu image
B: Use the Vertex AI ModelEvaluationOp component to evaluate the model
C: Use the Vertex AI SDK for Python within a custom component based on a python:3.10 image
D: Chain the Vertex AI ModelUploadOp and ModelDeployOp components together

Sample Question 121

You are developing an ML pipeline using Vertex AI Pipelines. You want your pipeline to upload a new version of the XGBoost model to Vertex AI Model Registry and deploy it to Vertex AI Endpoints for online inference. You want to use the simplest approach. What should you do?

A: Use the Vertex AI REST API within a custom component based on a vertex-ai/prediction/xgboost-cpu image
B: Use the Vertex AI ModelEvaluationOp component to evaluate the model
C: Use the Vertex AI SDK for Python within a custom component based on a python:3.10 image
D: Chain the Vertex AI ModelUploadOp and ModelDeployOp components together

Sample Question 120

You are building a MLOps platform to automate your company's ML experiments and model retraining. You need to organize the artifacts for dozens of pipelines. How should you store the pipelines' artifacts?

A: Store parameters in Cloud SQL, and store the models' source code and binaries in GitHub.
B: Store parameters in Cloud SQL, store the models' source code in GitHub, and store the models' binaries in Cloud Storage.
C: Store parameters in Vertex ML Metadata, store the models' source code in GitHub, and store the models' binaries in Cloud Storage.
D: Store parameters in Vertex ML Metadata and store the models' source code and binaries in GitHub.

Sample Question 120

You are building a MLOps platform to automate your company's ML experiments and model retraining. You need to organize the artifacts for dozens of pipelines. How should you store the pipelines' artifacts?

A: Store parameters in Cloud SQL, and store the models' source code and binaries in GitHub.
B: Store parameters in Cloud SQL, store the models' source code in GitHub, and store the models' binaries in Cloud Storage.
C: Store parameters in Vertex ML Metadata, store the models' source code in GitHub, and store the models' binaries in Cloud Storage.
D: Store parameters in Vertex ML Metadata and store the models' source code and binaries in GitHub.

Sample Question 47

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.

Sample Question 47

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.

Sample Question 106

You recently deployed a pipeline in Vertex AI Pipelines that trains and pushes a model to a Vertex AI endpoint to serve real-time traffic. You need to continue experimenting and iterating on your pipeline to improve model performance. You plan to use Cloud Build for CI/CD You want to quickly and easily deploy new pipelines into production, and you want to minimize the chance that the new pipeline implementations will break in production. What should you do?

A: Set up a CI/CD pipeline that builds and tests your source code. If the tests are successful, use the Google. Cloud console to upload the built container to Artifact Registry and upload the compiled pipeline to Vertex AI Pipelines.
B: Set up a CI/CD pipeline that builds your source code and then deploys built artifacts into a pre-production environment. Run unit tests in the pre-production environment. If the tests are successful deploy the pipeline to production.
C: Set up a CI/CD pipeline that builds and tests your source code and then deploys built artifacts into a pre-production environment. After a successful pipeline run in the pre-production environment, deploy the pipeline to production.
D: Set up a CI/CD pipeline that builds and tests your source code and then deploys built artifacts into a pre-production environment. After a successful pipeline run in the pre-production environment, rebuild the source code and deploy the artifacts to production.

Sample Question 106

You recently deployed a pipeline in Vertex AI Pipelines that trains and pushes a model to a Vertex AI endpoint to serve real-time traffic. You need to continue experimenting and iterating on your pipeline to improve model performance. You plan to use Cloud Build for CI/CD You want to quickly and easily deploy new pipelines into production, and you want to minimize the chance that the new pipeline implementations will break in production. What should you do?

A: Set up a CI/CD pipeline that builds and tests your source code. If the tests are successful, use the Google. Cloud console to upload the built container to Artifact Registry and upload the compiled pipeline to Vertex AI Pipelines.

B: Set up a CI/CD pipeline that builds your source code and then deploys built artifacts into a pre-production environment. Run unit tests in the pre-production environment. If the tests are successful deploy the pipeline to production.

C: Set up a CI/CD pipeline that builds and tests your source code and then deploys built artifacts into a pre-production environment. After a successful pipeline run in the pre-production environment, deploy the pipeline to production.

D: Set up a CI/CD pipeline that builds and tests your source code and then deploys built artifacts into a pre-production environment. After a successful pipeline run in the pre-production environment, rebuild the source code and deploy the artifacts to production.

Sample Question 93

You have been tasked with deploying prototype code to production. The feature engineering code is in PySpark and runs on Dataproc Serverless. The model training is executed by using a Vertex AI custom training job. The two steps are not connected, and the model training must currently be run manually after the feature engineering step finishes. You need to create a scalable and maintainable production process that runs end-to-end and tracks the connections between steps. What should you do?

A: Create a Vertex AI Workbench notebook. Use the notebook to submit the Dataproc Serverless feature engineering job. Use the same notebook to submit the custom model training job. Run the notebook cells sequentially to tie the steps together end-to-end.

B: Create a Vertex AI Workbench notebook. Initiate an Apache Spark context in the notebook and run the PySpark feature engineering code. Use the same notebook to run the custom model training job in TensorFlow. Run the notebook cells sequentially to tie the steps together end-to-end.

C: Use the Kubeflow pipelines SDK to write code that specifies two components:
- The first is a Dataproc Serverless component that launches the feature engineering job
- The second is a custom component wrapped in the create_custom_training_job_from_component utility that launches the custom model training job
Create a Vertex AI Pipelines job to link and run both components

D: Use the Kubeflow pipelines SDK to write code that specifies two components
- The first component initiates an Apache Spark context that runs the PySpark feature engineering code
- The second component runs the TensorFlow custom model training code
Create a Vertex AI Pipelines job to link and run both components.

Sample Question 93

You have been tasked with deploying prototype code to production. The feature engineering code is in PySpark and runs on Dataproc Serverless. The model training is executed by using a Vertex AI custom training job. The two steps are not connected, and the model training must currently be run manually after the feature engineering step finishes. You need to create a scalable and maintainable production process that runs end-to-end and tracks the connections between steps. What should you do?

A: Create a Vertex AI Workbench notebook. Use the notebook to submit the Dataproc Serverless feature engineering job. Use the same notebook to submit the custom model training job. Run the notebook cells sequentially to tie the steps together end-to-end.

B: Create a Vertex AI Workbench notebook. Initiate an Apache Spark context in the notebook and run the PySpark feature engineering code. Use the same notebook to run the custom model training job in TensorFlow. Run the notebook cells sequentially to tie the steps together end-to-end.

C: Use the Kubeflow pipelines SDK to write code that specifies two components:
- The first is a Dataproc Serverless component that launches the feature engineering job
- The second is a custom component wrapped in the create_custom_training_job_from_component utility that launches the custom model training job
Create a Vertex AI Pipelines job to link and run both components

D: Use the Kubeflow pipelines SDK to write code that specifies two components
- The first component initiates an Apache Spark context that runs the PySpark feature engineering code
- The second component runs the TensorFlow custom model training code
Create a Vertex AI Pipelines job to link and run both components.

Sample Question 8

You are building a custom image classification model and plan to use Vertex AI Pipelines to implement the end-to-end training. Your dataset consists of images that need to be preprocessed before they can be used to train the model. The preprocessing steps include resizing the images, converting them to grayscale, and extracting features. You have already implemented some Python functions for the preprocessing tasks. Which components should you use in your pipeline?

A: DataprocSparkBatchOp and CustomTrainingJobOp
B: DataflowPythonJobOp, WaitGcpResourcesOp, and CustomTrainingJobOp
C: dsl.ParallelFor, dsl.component, and CustomTrainingJobOp
D: ImageDatasetImportDataOp, dsl.component, and AutoMLImageTrainingJobRunOp

Sample Question 8

You are building a custom image classification model and plan to use Vertex AI Pipelines to implement the end-to-end training. Your dataset consists of images that need to be preprocessed before they can be used to train the model. The preprocessing steps include resizing the images, converting them to grayscale, and extracting features. You have already implemented some Python functions for the preprocessing tasks. Which components should you use in your pipeline?

A: DataprocSparkBatchOp and CustomTrainingJobOp
B: DataflowPythonJobOp, WaitGcpResourcesOp, and CustomTrainingJobOp
C: dsl.ParallelFor, dsl.component, and CustomTrainingJobOp
D: ImageDatasetImportDataOp, dsl.component, and AutoMLImageTrainingJobRunOp

Sample Question 117

You are building a TensorFlow text-to-image generative model by using a dataset that contains billions of images with their respective captions. You want to create a low maintenance, automated workflow that reads the data from a Cloud Storage bucket collects statistics, splits the dataset into training/validation/test datasets performs data transformations trains the model using the training/validation datasets, and validates the model by using the test dataset. What should you do?

A: Use the Apache Airflow SDK to create multiple operators that use Dataflow and Vertex AI services. Deploy the workflow on Cloud Composer.
B: Use the MLFlow SDK and deploy it on a Google Kubernetes Engine cluster. Create multiple components that use Dataflow and Vertex AI services.
C: Use the Kubeflow Pipelines (KFP) SDK to create multiple components that use Dataflow and Vertex AI services. Deploy the workflow on Vertex AI Pipelines.
D: Use the TensorFlow Extended (TFX) SDK to create multiple components that use Dataflow and Vertex AI services. Deploy the workflow on Vertex AI Pipelines.

Sample Question 117

You are building a TensorFlow text-to-image generative model by using a dataset that contains billions of images with their respective captions. You want to create a low maintenance, automated workflow that reads the data from a Cloud Storage bucket collects statistics, splits the dataset into training/validation/test datasets performs data transformations trains the model using the training/validation datasets, and validates the model by using the test dataset. What should you do?

A: Use the Apache Airflow SDK to create multiple operators that use Dataflow and Vertex AI services. Deploy the workflow on Cloud Composer.
B: Use the MLFlow SDK and deploy it on a Google Kubernetes Engine cluster. Create multiple components that use Dataflow and Vertex AI services.
C: Use the Kubeflow Pipelines (KFP) SDK to create multiple components that use Dataflow and Vertex AI services. Deploy the workflow on Vertex AI Pipelines.
D: Use the TensorFlow Extended (TFX) SDK to create multiple components that use Dataflow and Vertex AI services. Deploy the workflow on Vertex AI Pipelines.

Sample Question 95

You are creating a model training pipeline to predict sentiment scores from text-based product reviews. You want to have control over how the model parameters are tuned, and you will deploy the model to an endpoint after it has been trained. You will use Vertex AI Pipelines to run the pipeline. You need to decide which Google Cloud pipeline components to use. What components should you choose?

A: TabularDatasetCreateOp, CustomTrainingJobOp, and EndpointCreateOp
B: TextDatasetCreateOp, AutoMLTextTrainingOp, and EndpointCreateOp
C: TabularDatasetCreateOp. AutoMLTextTrainingOp, and ModelDeployOp
D: TextDatasetCreateOp, CustomTrainingJobOp, and ModelDeployOp

Sample Question 95

You are creating a model training pipeline to predict sentiment scores from text-based product reviews. You want to have control over how the model parameters are tuned, and you will deploy the model to an endpoint after it has been trained. You will use Vertex AI Pipelines to run the pipeline. You need to decide which Google Cloud pipeline components to use. What components should you choose?

A: TabularDatasetCreateOp, CustomTrainingJobOp, and EndpointCreateOp
B: TextDatasetCreateOp, AutoMLTextTrainingOp, and EndpointCreateOp
C: TabularDatasetCreateOp. AutoMLTextTrainingOp, and ModelDeployOp
D: TextDatasetCreateOp, CustomTrainingJobOp, and ModelDeployOp

Sample Question 9

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.

Sample Question 9

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.

Sample Question 33

You are the Director of Data Science at a large company, and your Data Science team has recently begun using the Kubeflow Pipelines SDK to orchestrate their training pipelines. Your team is struggling to integrate their custom Python code into the Kubeflow Pipelines SDK. How should you instruct them to proceed in order to quickly integrate their code with the Kubeflow Pipelines SDK?

A: Use the func_to_container_op function to create custom components from the Python code.
B: Use the predefined components available in the Kubeflow Pipelines SDK to access Dataproc, and run the custom code there.
C: Package the custom Python code into Docker containers, and use the load_component_from_file function to import the containers into the pipeline.
D: Deploy the custom Python code to Cloud Functions, and use Kubeflow Pipelines to trigger the Cloud Function

Sample Question 33

You are the Director of Data Science at a large company, and your Data Science team has recently begun using the Kubeflow Pipelines SDK to orchestrate their training pipelines. Your team is struggling to integrate their custom Python code into the Kubeflow Pipelines SDK. How should you instruct them to proceed in order to quickly integrate their code with the Kubeflow Pipelines SDK?

A: Use the func_to_container_op function to create custom components from the Python code.

B: Use the predefined components available in the Kubeflow Pipelines SDK to access Dataproc, and run the custom code there.

C: Package the custom Python code into Docker containers, and use the load_component_from_file function to import the containers into the pipeline.

D: Deploy the custom Python code to Cloud Functions, and use Kubeflow Pipelines to trigger the Cloud Function

Sample Question 0

208. You recently developed a wide and deep model in TensorFlow. You generated training datasets using a SQL script that preprocessed raw data in
BigQuery by performing instance-level transformations of the data. You need to create a training pipeline to retrain the model on a weekly basis.
The trained model will be used to generate daily recommendations. You want to minimize model development and training time. How should you
develop the training pipeline?

A: Use the Kubeflow Pipelines SDK to implement the pipeline. Use the BigQueryJobOp component to run the preprocessing script and the
CustomTrainingJobOp component to launch a Vertex AI training job.
B: Use the Kubeflow Pipelines SDK to implement the pipeline. Use the DataflowPythonJobOp component to preprocess the data and the CustomTrainingJobOp component to launch a Vertex AI training job.
C: Use the TensorFlow Extended SDK to implement the pipeline Use the ExampleGen component with the BigQuery executor to ingest the data the Transform component to preprocess the data, and the Trainer component to launch a Vertex AI training job.
D: Use the TensorFlow Extended SDK to implement the pipeline Implement the preprocessing steps as part of the input_fn of the model. Use
the ExampleGen component with the BigQuery executor to ingest the data and the Trainer component to launch a Vertex AI training job.

Sample Question 0

208. You recently developed a wide and deep model in TensorFlow. You generated training datasets using a SQL script that preprocessed raw data in
BigQuery by performing instance-level transformations of the data. You need to create a training pipeline to retrain the model on a weekly basis.
The trained model will be used to generate daily recommendations. You want to minimize model development and training time. How should you
develop the training pipeline?

A: Use the Kubeflow Pipelines SDK to implement the pipeline. Use the BigQueryJobOp component to run the preprocessing script and the
CustomTrainingJobOp component to launch a Vertex AI training job.
B: Use the Kubeflow Pipelines SDK to implement the pipeline. Use the DataflowPythonJobOp component to preprocess the data and the CustomTrainingJobOp component to launch a Vertex AI training job.
C: Use the TensorFlow Extended SDK to implement the pipeline Use the ExampleGen component with the BigQuery executor to ingest the data the Transform component to preprocess the data, and the Trainer component to launch a Vertex AI training job.
D: Use the TensorFlow Extended SDK to implement the pipeline Implement the preprocessing steps as part of the input_fn of the model. Use
the ExampleGen component with the BigQuery executor to ingest the data and the Trainer component to launch a Vertex AI training job.

Sample Question 19

You work as an analyst at a large banking firm. You are developing a robust scalable ML pipeline to train several regression and classification models. Your primary focus for the pipeline is model interpretability. You want to productionize the pipeline as quickly as possible. What should you do?

A: Use Tabular Workflow for Wide & Deep through Vertex AI Pipelines to jointly train wide linear models and deep neural networks
B: Use Google Kubernetes Engine to build a custom training pipeline for XGBoost-based models
C: Use Tabular Workflow for TabNet through Vertex AI Pipelines to train attention-based models
D: Use Cloud Composer to build the training pipelines for custom deep learning-based models

Sample Question 19

You work as an analyst at a large banking firm. You are developing a robust scalable ML pipeline to train several regression and classification models. Your primary focus for the pipeline is model interpretability. You want to productionize the pipeline as quickly as possible. What should you do?

A: Use Tabular Workflow for Wide & Deep through Vertex AI Pipelines to jointly train wide linear models and deep neural networks
B: Use Google Kubernetes Engine to build a custom training pipeline for XGBoost-based models
C: Use Tabular Workflow for TabNet through Vertex AI Pipelines to train attention-based models
D: Use Cloud Composer to build the training pipelines for custom deep learning-based models

Sample Question 20

Your company manages an ecommerce website. You developed an ML model that recommends additional products to users in near real time based on items currently in the user's cart. The workflow will include the following processes:

1. The website will send a Pub/Sub message with the relevant data and then receive a message with the prediction from Pub/Sub 2. Predictions will be stored in BigQuery
3. The model will be stored in a Cloud Storage bucket and will be updated frequently

You want to minimize prediction latency and the effort required to update the model. How should you reconfigure the architecture?

A: Write a Cloud Function that loads the model into memory for prediction. Configure the function to be triggered when messages are sent to Pub/Sub.

B: Create a pipeline in Vertex AI Pipelines that performs preprocessing, prediction, and postprocessing. Configure the pipeline to be triggered by a Cloud Function when messages are sent to Pub/Sub.

C: Expose the model as a Vertex AI endpoint. Write a custom DoFn in a Dataflow job that calls the endpoint for prediction.

D: Use the RunInference API with WatchFilePattern in a Dataflow job that wraps around the model and serves predictions.

Sample Question 20

Your company manages an ecommerce website. You developed an ML model that recommends additional products to users in near real time based on items currently in the user's cart. The workflow will include the following processes:

1. The website will send a Pub/Sub message with the relevant data and then receive a message with the prediction from Pub/Sub 2. Predictions will be stored in BigQuery
3. The model will be stored in a Cloud Storage bucket and will be updated frequently

You want to minimize prediction latency and the effort required to update the model. How should you reconfigure the architecture?

A: Write a Cloud Function that loads the model into memory for prediction. Configure the function to be triggered when messages are sent to Pub/Sub.

B: Create a pipeline in Vertex AI Pipelines that performs preprocessing, prediction, and postprocessing. Configure the pipeline to be triggered by a Cloud Function when messages are sent to Pub/Sub.

C: Expose the model as a Vertex AI endpoint. Write a custom DoFn in a Dataflow job that calls the endpoint for prediction.

D: Use the RunInference API with WatchFilePattern in a Dataflow job that wraps around the model and serves predictions.

Sample Question 25

You developed a Vertex AI ML pipeline that consists of preprocessing and training steps and each set of steps runs on a separate custom Docker
image. Your organization uses GitHub and GitHub Actions as CI/CD to run unit and integration tests. You need to automate the model retraining
workflow so that it can be initiated both manually and when a new version of the code is merged in the main branch. You want to minimize the
steps required to build the workflow while also allowing for maximum flexibility. How should you configure the CI/CD workflow?

A: Trigger a Cloud Build workflow to run tests, build custom Docker images, push the images to Artifact Registry, and launch the pipeline in
Vertex AI Pipelines.
B: Trigger GitHub Actions to run the tests, launch a job on Cloud Run to build custom Docker images, push the images to Artifact Registry, and
launch the pipeline in Vertex AI Pipelines.
C: Trigger GitHub Actions to run the tests, build custom Docker images, push the images to Artifact Registry, and launch the pipeline in Vertex
AI Pipelines.
D: Trigger GitHub Actions to run the tests, launch a Cloud Build workflow to build custom Docker images, push the images to Artifact Registry,
and launch the pipeline in Vertex AI Pipelines.

Sample Question 25

You developed a Vertex AI ML pipeline that consists of preprocessing and training steps and each set of steps runs on a separate custom Docker
image. Your organization uses GitHub and GitHub Actions as CI/CD to run unit and integration tests. You need to automate the model retraining
workflow so that it can be initiated both manually and when a new version of the code is merged in the main branch. You want to minimize the
steps required to build the workflow while also allowing for maximum flexibility. How should you configure the CI/CD workflow?

A: Trigger a Cloud Build workflow to run tests, build custom Docker images, push the images to Artifact Registry, and launch the pipeline in
Vertex AI Pipelines.

B: Trigger GitHub Actions to run the tests, launch a job on Cloud Run to build custom Docker images, push the images to Artifact Registry, and
launch the pipeline in Vertex AI Pipelines.

C: Trigger GitHub Actions to run the tests, build custom Docker images, push the images to Artifact Registry, and launch the pipeline in Vertex
AI Pipelines.

D: Trigger GitHub Actions to run the tests, launch a Cloud Build workflow to build custom Docker images, push the images to Artifact Registry,
and launch the pipeline in Vertex AI Pipelines.

Sample Question 128

You are analyzing customer data for a healthcare organization that is stored in Cloud Storage. The data contains personally identifiable information (PII). You need to perform data exploration and preprocessing while ensuring the security and privacy of sensitive fields. What should you do?

A: Use the Cloud Data Loss Prevention (DLP) API to de-identify the PII before performing data exploration and preprocessing.
B: Use customer-managed encryption keys (CMEK) to encrypt the PII data at rest, and decrypt the PII data during data exploration and preprocessing.
C: Use a VM inside a VPC Service Controls security perimeter to perform data exploration and preprocessing.
D: Use Google-managed encryption keys to encrypt the PII data at rest, and decrypt the PII data during data exploration and preprocessing.

Sample Question 128

You are analyzing customer data for a healthcare organization that is stored in Cloud Storage. The data contains personally identifiable information (PII). You need to perform data exploration and preprocessing while ensuring the security and privacy of sensitive fields. What should you do?

A: Use the Cloud Data Loss Prevention (DLP) API to de-identify the PII before performing data exploration and preprocessing.

B: Use customer-managed encryption keys (CMEK) to encrypt the PII data at rest, and decrypt the PII data during data exploration and preprocessing.

C: Use a VM inside a VPC Service Controls security perimeter to perform data exploration and preprocessing.

D: Use Google-managed encryption keys to encrypt the PII data at rest, and decrypt the PII data during data exploration and preprocessing.

Sample Question 32

You have built a model that is trained on data stored in Parquet files. You access the data through a Hive table hosted on Google Cloud. You preprocessed these data with PySpark and exported it as a CSV file into Cloud Storage. After preprocessing, you execute additional steps to train and evaluate your model. You want to parametrize this model training in Kubeflow Pipelines. What should you do?

A: Remove the data transformation step from your pipeline.
B: Containerize the PySpark transformation step, and add it to your pipeline.
C: Add a ContainerOp to your pipeline that spins a Dataproc cluster, runs a transformation, and then saves the transformed data in Cloud Storage.
D: Deploy Apache Spark at a separate node pool in a Google Kubernetes Engine cluster. Add a ContainerOp to your pipeline that invokes a corresponding transformation job for this Spark instance.

Sample Question 32

You have built a model that is trained on data stored in Parquet files. You access the data through a Hive table hosted on Google Cloud. You preprocessed these data with PySpark and exported it as a CSV file into Cloud Storage. After preprocessing, you execute additional steps to train and evaluate your model. You want to parametrize this model training in Kubeflow Pipelines. What should you do?

A: Remove the data transformation step from your pipeline.
B: Containerize the PySpark transformation step, and add it to your pipeline.
C: Add a ContainerOp to your pipeline that spins a Dataproc cluster, runs a transformation, and then saves the transformed data in Cloud Storage.
D: Deploy Apache Spark at a separate node pool in a Google Kubernetes Engine cluster. Add a ContainerOp to your pipeline that invokes a corresponding transformation job for this Spark instance.

Sample Question 6

You developed a Vertex AI pipeline that trains a classification model on data stored in a large BigQuery table. The pipeline has four steps, where each step is created by a Python function that uses the KubeFlow v2 API. You perform many model iterations by adjusting the code and parameters of the training step. You observe high costs associated with the development, particularly the data export and preprocessing steps. You need to reduce model development costs. What should you do?

A: Change the components' YAML filenames to export.yaml, preprocess,yaml, f "train-{dt}.yaml", f"calibrate-{dt).vaml".
B: Add the {"kubeflow.v1.caching": True} parameter to the set of params provided to your PipelineJob.
C: Move the first step of your pipeline to a separate step, and provide a cached path to Cloud Storage as an input to the main pipeline.
D: Change the name of the pipeline to f"my-awesome-pipeline-{dt}".

Sample Question 6

You developed a Vertex AI pipeline that trains a classification model on data stored in a large BigQuery table. The pipeline has four steps, where each step is created by a Python function that uses the KubeFlow v2 API. You perform many model iterations by adjusting the code and parameters of the training step. You observe high costs associated with the development, particularly the data export and preprocessing steps. You need to reduce model development costs. What should you do?

A: Change the components' YAML filenames to export.yaml, preprocess,yaml, f "train-{dt}.yaml", f"calibrate-{dt).vaml".
B: Add the {"kubeflow.v1.caching": True} parameter to the set of params provided to your PipelineJob.
C: Move the first step of your pipeline to a separate step, and provide a cached path to Cloud Storage as an input to the main pipeline.
D: Change the name of the pipeline to f"my-awesome-pipeline-{dt}".

Sample Question 21

You have recently used TensorFlow to train a classification model on tabular data. You have created a Dataflow pipeline that can transform several terabytes of data into training or prediction datasets consisting of TFRecords. You now need to productionize the model, and you want the predictions to be automatically uploaded to a BigQuery table on a weekly schedule. What should you do?

A: Import the model into Vertex AI and deploy it to a Vertex AI endpoint. On Vertex AI Pipelines, create a pipeline that uses the DataflowPythonJobOp and the ModelBacthPredictOp components.
B: Import the model into Vertex AI and deploy it to a Vertex AI endpoint. Create a Dataflow pipeline that reuses the data processing logic sends requests to the endpoint, and then uploads predictions to a BigQuery table.
C: Import the model into Vertex AI. On Vertex AI Pipelines, create a pipeline that uses the DataflowPvthonJobOp and the ModelBatchPredictOp components.
D: Import the model into BigQuery. Implement the data processing logic in a SQL query. On Vertex AI Pipelines create a pipeline that uses the BigquervQueryJobOp and the BigqueryPredictModelJobOp components.

Sample Question 21

You have recently used TensorFlow to train a classification model on tabular data. You have created a Dataflow pipeline that can transform several terabytes of data into training or prediction datasets consisting of TFRecords. You now need to productionize the model, and you want the predictions to be automatically uploaded to a BigQuery table on a weekly schedule. What should you do?

A: Import the model into Vertex AI and deploy it to a Vertex AI endpoint. On Vertex AI Pipelines, create a pipeline that uses the DataflowPythonJobOp and the ModelBacthPredictOp components.

B: Import the model into Vertex AI and deploy it to a Vertex AI endpoint. Create a Dataflow pipeline that reuses the data processing logic sends requests to the endpoint, and then uploads predictions to a BigQuery table.

C: Import the model into Vertex AI. On Vertex AI Pipelines, create a pipeline that uses the DataflowPythonJobOp and the ModelBatchPredictOp components.

D: Import the model into BigQuery. Implement the data processing logic in a SQL query. On Vertex AI Pipelines create a pipeline that uses the BigquervQueryJobOp and the BigqueryPredictModelJobOp components.

Sample Question 83

Your team frequently creates new ML models and runs experiments. Your team pushes code to a single repository hosted on Cloud Source Repositories. You want to create a continuous integration pipeline that automatically retrains the models whenever there is any modification of the code. What should be your first step to set up the CI pipeline?

A: Configure a Cloud Build trigger with the event set as "Pull Request"
B: Configure a Cloud Build trigger with the event set as "Push to a branch"
C: Configure a Cloud Function that builds the repository each time there is a code change
D: Configure a Cloud Function that builds the repository each time a new branch is created

Sample Question 83

Your team frequently creates new ML models and runs experiments. Your team pushes code to a single repository hosted on Cloud Source Repositories. You want to create a continuous integration pipeline that automatically retrains the models whenever there is any modification of the code. What should be your first step to set up the CI pipeline?

A: Configure a Cloud Build trigger with the event set as "Pull Request"
B: Configure a Cloud Build trigger with the event set as "Push to a branch"
C: Configure a Cloud Function that builds the repository each time there is a code change
D: Configure a Cloud Function that builds the repository each time a new branch is created

Sample Question 31

You have created a Vertex AI pipeline that includes two steps. The first step preprocesses 10 TB data completes in about 1 hour, and saves the result in a Cloud Storage bucket. The second step uses the processed data to train a model. You need to update the model's code to allow you to test different algorithms. You want to reduce pipeline execution time and cost while also minimizing pipeline changes. What should you do?

A: Add a pipeline parameter and an additional pipeline step. Depending on the parameter value, the pipeline step conducts or skips data preprocessing, and starts model training.
B: Create another pipeline without the preprocessing step, and hardcode the preprocessed Cloud Storage file location for model training.
C: Configure a machine with more CPU and RAM from the compute-optimized machine family for the data preprocessing step.
D: Enable caching for the pipeline job, and disable caching for the model training step.

Sample Question 31

You have created a Vertex AI pipeline that includes two steps. The first step preprocesses 10 TB data completes in about 1 hour, and saves the result in a Cloud Storage bucket. The second step uses the processed data to train a model. You need to update the model's code to allow you to test different algorithms. You want to reduce pipeline execution time and cost while also minimizing pipeline changes. What should you do?

A: Add a pipeline parameter and an additional pipeline step. Depending on the parameter value, the pipeline step conducts or skips data preprocessing, and starts model training.
B: Create another pipeline without the preprocessing step, and hardcode the preprocessed Cloud Storage file location for model training.
C: Configure a machine with more CPU and RAM from the compute-optimized machine family for the data preprocessing step.
D: Enable caching for the pipeline job, and disable caching for the model training step.

Sample Question 125

You work at a bank. You need to develop a credit risk model to support loan application decisions. You decide to implement the model by using a neural network in TensorFlow. Due to regulatory requirements, you need to be able to explain the model's predictions based on its features. When the model is deployed, you also want to monitor the model's performance over time. You decided to use Vertex AI for both model development and deployment. What should you do?

A: Use Vertex Explainable AI with the sampled Shapley method, and enable Vertex AI Model Monitoring to check for feature distribution drift.
B: Use Vertex Explainable AI with the sampled Shapley method, and enable Vertex AI Model Monitoring to check for feature distribution skew.
C: Use Vertex Explainable AI with the XRAI method, and enable Vertex AI Model Monitoring to check for feature distribution drift.
D: Use Vertex Explainable AI with the XRAI method, and enable Vertex AI Model Monitoring to check for feature distribution skew.

Sample Question 125

You work at a bank. You need to develop a credit risk model to support loan application decisions. You decide to implement the model by using a neural network in TensorFlow. Due to regulatory requirements, you need to be able to explain the model's predictions based on its features. When the model is deployed, you also want to monitor the model's performance over time. You decided to use Vertex AI for both model development and deployment. What should you do?

A: Use Vertex Explainable AI with the sampled Shapley method, and enable Vertex AI Model Monitoring to check for feature distribution drift.

B: Use Vertex Explainable AI with the sampled Shapley method, and enable Vertex AI Model Monitoring to check for feature distribution skew.

C: Use Vertex Explainable AI with the XRAI method, and enable Vertex AI Model Monitoring to check for feature distribution drift.

D: Use Vertex Explainable AI with the XRAI method, and enable Vertex AI Model Monitoring to check for feature distribution skew.

Sample Question 13

You recently trained a XGBoost model that you plan to deploy to production for online inference. Before sending a predict request to your model's binary, you need to perform a simple data preprocessing step. This step exposes a REST API that accepts requests in your internal VPC Service Controls and returns predictions. You want to configure this preprocessing step while minimizing cost and effort. What should you do?

A: Store a pickled model in Cloud Storage. Build a Flask-based app, package the app in a custom container image, and deploy the model to Vertex AI Endpoints.
B: Build a Flask-based app, package the app and a pickled model in a custom container image, and deploy the model to Vertex AI Endpoints.
C: Build a custom predictor class based on XGBoost Predictor from the Vertex AI SDK, package it and a pickled model in a custom container image based on a Vertex built-in image, and deploy the model to Vertex AI Endpoints.
D: Build a custom predictor class based on XGBoost Predictor from the Vertex AI SDK, and package the handler in a custom container image based on a Vertex built-in container image. Store a pickled model in Cloud Storage, and deploy the model to Vertex AI Endpoints.

Sample Question 13

You recently trained a XGBoost model that you plan to deploy to production for online inference. Before sending a predict request to your model's binary, you need to perform a simple data preprocessing step. This step exposes a REST API that accepts requests in your internal VPC Service Controls and returns predictions. You want to configure this preprocessing step while minimizing cost and effort. What should you do?

A: Store a pickled model in Cloud Storage. Build a Flask-based app, package the app in a custom container image, and deploy the model to Vertex AI Endpoints.
B: Build a Flask-based app, package the app and a pickled model in a custom container image, and deploy the model to Vertex AI Endpoints.
C: Build a custom predictor class based on XGBoost Predictor from the Vertex AI SDK, package it and a pickled model in a custom container image based on a Vertex built-in image, and deploy the model to Vertex AI Endpoints.
D: Build a custom predictor class based on XGBoost Predictor from the Vertex AI SDK, and package the handler in a custom container image based on a Vertex built-in container image. Store a pickled model in Cloud Storage, and deploy the model to Vertex AI Endpoints.

Sample Question 49

You work for a bank with strict data governance requirements. You recently implemented a custom model to detect fraudulent transactions. You want your training code to download internal data by using an API endpoint hosted in your project's network. You need the data to be accessed in the most secure way, while mitigating the risk of data exfiltration. What should you do?

A: Enable VPC Service Controls for peerings, and add Vertex AI to a service perimeter.
B: Create a Cloud Run endpoint as a proxy to the data. Use Identity and Access Management (IAM) authentication to secure access to the endpoint from the training job.
C: Configure VPC Peering with Vertex AI, and specify the network of the training job.
D: Download the data to a Cloud Storage bucket before calling the training job.

Sample Question 49

You work for a bank with strict data governance requirements. You recently implemented a custom model to detect fraudulent transactions. You want your training code to download internal data by using an API endpoint hosted in your project's network. You need the data to be accessed in the most secure way, while mitigating the risk of data exfiltration. What should you do?

A: Enable VPC Service Controls for peerings, and add Vertex AI to a service perimeter.

B: Create a Cloud Run endpoint as a proxy to the data. Use Identity and Access Management (IAM) authentication to secure access to the endpoint from the training job.

C: Configure VPC Peering with Vertex AI, and specify the network of the training job.

D: Download the data to a Cloud Storage bucket before calling the training job.

Sample Question 76

You have developed an application that uses a chain of multiple scikit-learn models to predict the optimal price for your company's products. Members of your team use the individual models in other solution workflows. You want to deploy this workflow while ensuring version control for each individual model and the overall workflow. Your application needs to be able to scale down to zero. You want to minimize the compute resource utilization and the manual effort required to manage this solution. What should you do?

A: Expose each individual model as an endpoint in Vertex AI Endpoints. Create a custom container endpoint to orchestrate the workflow.
B: Create a custom container endpoint for the workflow that loads each model's individual files Track the versions of each individual model in BigQuery.
C: Expose each individual model as an endpoint in Vertex AI Endpoints. Use Cloud Run to orchestrate the workflow.
D: Load each model's individual files into Cloud Run. Use Cloud Run to orchestrate the workflow. Track the versions of each individual model in BigQuery.

Sample Question 76

You have developed an application that uses a chain of multiple scikit-learn models to predict the optimal price for your company's products. Members of your team use the individual models in other solution workflows. You want to deploy this workflow while ensuring version control for each individual model and the overall workflow. Your application needs to be able to scale down to zero. You want to minimize the compute resource utilization and the manual effort required to manage this solution. What should you do?

A: Expose each individual model as an endpoint in Vertex AI Endpoints. Create a custom container endpoint to orchestrate the workflow.
B: Create a custom container endpoint for the workflow that loads each model's individual files Track the versions of each individual model in BigQuery.
C: Expose each individual model as an endpoint in Vertex AI Endpoints. Use Cloud Run to orchestrate the workflow.
D: Load each model's individual files into Cloud Run. Use Cloud Run to orchestrate the workflow. Track the versions of each individual model in BigQuery.

Sample Question 115

You recently trained an XGBoost model on tabular data. You plan to expose the model for internal use as an HTTP microservice. After deployment, you expect a small number of incoming requests. You want to productionize the model with the least amount of effort and latency. What should you do?

A: Deploy the model to BigQuery ML by using CREATE MODEL with the BOOSTED_TREE_REGRESSOR statement, and invoke the BigQuery API from the microservice.
B: Build a Flask-based app. Package the app in a custom container on Vertex AI, and deploy it to Vertex AI Endpoints.
C: Build a Flask-based app. Package the app in a Docker image, and deploy it to Google Kubernetes Engine in Autopilot mode.
D: Use a prebuilt XGBoost Vertex container to create a model, and deploy it to Vertex AI Endpoints.

Sample Question 115

You recently trained an XGBoost model on tabular data. You plan to expose the model for internal use as an HTTP microservice. After deployment, you expect a small number of incoming requests. You want to productionize the model with the least amount of effort and latency. What should you do?

A: Deploy the model to BigQuery ML by using CREATE MODEL with the BOOSTED_TREE_REGRESSOR statement, and invoke the BigQuery API from the microservice.
B: Build a Flask-based app. Package the app in a custom container on Vertex AI, and deploy it to Vertex AI Endpoints.
C: Build a Flask-based app. Package the app in a Docker image, and deploy it to Google Kubernetes Engine in Autopilot mode.
D: Use a prebuilt XGBoost Vertex container to create a model, and deploy it to Vertex AI Endpoints.

Sample Question 114

You have trained an XGBoost model that you plan to deploy on Vertex AI for online prediction. You are now uploading your model to Vertex AI Model Registry, and you need to configure the explanation method that will serve online prediction requests to be returned with minimal latency. You also want to be alerted when feature attributions of the model meaningfully change over time. What should you do?

A: 1. Specify sampled Shapley as the explanation method with a path count of 5.
2. Deploy the model to Vertex AI Endpoints.
3. Create a Model Monitoring job that uses prediction drift as the monitoring objective.

B: 1. Specify Integrated Gradients as the explanation method with a path count of 5.
2. Deploy the model to Vertex AI Endpoints.
3. Create a Model Monitoring job that uses prediction drift as the monitoring objective.

C: 1. Specify sampled Shapley as the explanation method with a path count of 50.
2. Deploy the model to Vertex AI Endpoints.
3. Create a Model Monitoring job that uses training-serving skew as the monitoring objective.

D: 1. Specify Integrated Gradients as the explanation method with a path count of 50.
2. Deploy the model to Vertex AI Endpoints.
3. Create a Model Monitoring job that uses training-serving skew as the monitoring objective.

Sample Question 114

You have trained an XGBoost model that you plan to deploy on Vertex AI for online prediction. You are now uploading your model to Vertex AI Model Registry, and you need to configure the explanation method that will serve online prediction requests to be returned with minimal latency. You also want to be alerted when feature attributions of the model meaningfully change over time. What should you do?

A: 1. Specify sampled Shapley as the explanation method with a path count of 5.
2. Deploy the model to Vertex AI Endpoints.
3. Create a Model Monitoring job that uses prediction drift as the monitoring objective.

B: 1. Specify Integrated Gradients as the explanation method with a path count of 5.
2. Deploy the model to Vertex AI Endpoints.
3. Create a Model Monitoring job that uses prediction drift as the monitoring objective.

C: 1. Specify sampled Shapley as the explanation method with a path count of 50.
2. Deploy the model to Vertex AI Endpoints.
3. Create a Model Monitoring job that uses training-serving skew as the monitoring objective.

D: 1. Specify Integrated Gradients as the explanation method with a path count of 50.
2. Deploy the model to Vertex AI Endpoints.
3. Create a Model Monitoring job that uses training-serving skew as the monitoring objective.

Sample Question 7

You work for an online grocery store. You recently developed a custom ML model that recommends a recipe when a user arrives at the website. You chose the machine type on the Vertex AI endpoint to optimize costs by using the queries per second (QPS) that the model can serve, and you deployed it on a single machine with 8 vCPUs and no accelerators. A holiday season is approaching and you anticipate four times more traffic during this time than the typical daily traffic. You need to ensure that the model can scale efficiently to the increased demand. What should you do?

A: A. 1. Maintain the same machine type on the endpoint.
2. Set up a monitoring job and an alert for CPU usage.
3. If you receive an alert, add a compute node to the endpoint.
B: 1. Change the machine type on the endpoint to have 32 vCPUs.
2. Set up a monitoring job and an alert for CPU usage.
3. If you receive an alert, scale the vCPUs further as needed.
C: 1. Maintain the same machine type on the endpoint Configure the endpoint to enable autoscaling based on vCPU usage.
2. Set up a monitoring job and an alert for CPU usage.
3. If you receive an alert, investigate the cause.
D: 1. Change the machine type on the endpoint to have a GPU. Configure the endpoint to enable autoscaling based on the GPU usage.
2. Set up a monitoring job and an alert for GPU usage.
3. If you receive an alert, investigate the cause.

Sample Question 7

You work for an online grocery store. You recently developed a custom ML model that recommends a recipe when a user arrives at the website. You chose the machine type on the Vertex AI endpoint to optimize costs by using the queries per second (QPS) that the model can serve, and you deployed it on a single machine with 8 vCPUs and no accelerators. A holiday season is approaching and you anticipate four times more traffic during this time than the typical daily traffic. You need to ensure that the model can scale efficiently to the increased demand. What should you do?

A: A. 1. Maintain the same machine type on the endpoint.
2. Set up a monitoring job and an alert for CPU usage.
3. If you receive an alert, add a compute node to the endpoint.
B: 1. Change the machine type on the endpoint to have 32 vCPUs.
2. Set up a monitoring job and an alert for CPU usage.
3. If you receive an alert, scale the vCPUs further as needed.
C: 1. Maintain the same machine type on the endpoint Configure the endpoint to enable autoscaling based on vCPU usage.
2. Set up a monitoring job and an alert for CPU usage.
3. If you receive an alert, investigate the cause.
D: 1. Change the machine type on the endpoint to have a GPU. Configure the endpoint to enable autoscaling based on the GPU usage.
2. Set up a monitoring job and an alert for GPU usage.
3. If you receive an alert, investigate the cause.

Sample Question 4

You work for a telecommunications company. You're building a model to predict which customers may fail to pay their next phone bill. The purpose of this model is to proactively offer at-risk customers assistance such as service discounts and bill deadline extensions. The data is stored in BigQuery and the predictive features that are available for model training include:

- Customer ID
- Age
- Salary
- Sex
- Average bill value
- Number of phone calls in last month
- Average duration of phone calls in last month

You need to investigate and mitigate potential bias against disadvantaged groups, while preserving model accuracy. What should you do?

A: Determine whether there is a meaningful correlation between the sensitive features and the other features. Train a BigQuery ML boosted trees classification model and exclude the sensitive features and any meaningfully correlated features.

B: Train a BigQuery ML boosted trees classification model with all features. Use the ML.GLOBAL_EXPLAIN method to calculate the global attribution values for each feature of the model. If the feature importance value for any of the sensitive features exceeds a threshold, discard the model and train without this feature.

C: Train a BigQuery ML boosted trees classification model with all features. Use the ML.EXPLAIN_PREDICT method to

Sample Question 4

You work for a telecommunications company. You're building a model to predict which customers may fail to pay their next phone bill. The purpose of this model is to proactively offer at-risk customers assistance such as service discounts and bill deadline extensions. The data is stored in BigQuery and the predictive features that are available for model training include:

- Customer ID
- Age
- Salary
- Sex
- Average bill value
- Number of phone calls in last month
- Average duration of phone calls in last month

You need to investigate and mitigate potential bias against disadvantaged groups, while preserving model accuracy. What should you do?

A: Determine whether there is a meaningful correlation between the sensitive features and the other features. Train a BigQuery ML boosted trees classification model and exclude the sensitive features and any meaningfully correlated features.

B: Train a BigQuery ML boosted trees classification model with all features. Use the ML.GLOBAL_EXPLAIN method to calculate the global attribution values for each feature of the model. If the feature importance value for any of the sensitive features exceeds a threshold, discard the model and train without this feature.

C: Train a BigQuery ML boosted trees classification model with all features. Use the ML.EXPLAIN_PREDICT method to

Sample Question 112

You work for an organization that operates a streaming music service. You have a custom production model that is serving a “next song” recommendation based on a user's recent listening history. Your model is deployed on a Vertex AI endpoint. You recently retrained the same model by using fresh data. The model received positive test results offline. You now want to test the new model in production while minimizing complexity. What should you do?

A: Create a new Vertex AI endpoint for the new model and deploy the new model to that new endpoint. Build a service to randomly send 5% of production traffic to the new endpoint. Monitor end-user metrics such as listening time. If end-user metrics improve between models over time, gradually increase the percentage of production traffic sent to the new endpoint.

B: Capture incoming prediction requests in BigQuery. Create an experiment in Vertex AI Experiments. Run batch predictions for both models using the captured data. Use the user’s selected song to compare the models performance side by side. If the new model’s performance metrics are better than the previous model, deploy the new model to production.

C: Deploy the new model to the existing Vertex AI endpoint. Use traffic splitting to send 5% of production traffic to the new model. Monitor end-user metrics, such as listening time. If end-user metrics improve between models over time, gradually increase the percentage of production traffic sent to the new model.

D: Configure a model monitoring job for the existing Vertex AI endpoint. Configure the monitoring job to detect prediction drift and set a threshold for alerts. Update the model on the endpoint from the previous model to the new model. If you receive an alert of prediction drift, revert to the previous model.

Sample Question 112

You work for an organization that operates a streaming music service. You have a custom production model that is serving a “next song” recommendation based on a user's recent listening history. Your model is deployed on a Vertex AI endpoint. You recently retrained the same model by using fresh data. The model received positive test results offline. You now want to test the new model in production while minimizing complexity. What should you do?

A: Create a new Vertex AI endpoint for the new model and deploy the new model to that new endpoint. Build a service to randomly send 5% of production traffic to the new endpoint. Monitor end-user metrics such as listening time. If end-user metrics improve between models over time, gradually increase the percentage of production traffic sent to the new endpoint.

B: Capture incoming prediction requests in BigQuery. Create an experiment in Vertex AI Experiments. Run batch predictions for both models using the captured data. Use the user’s selected song to compare the models performance side by side. If the new model’s performance metrics are better than the previous model, deploy the new model to production.

C: Deploy the new model to the existing Vertex AI endpoint. Use traffic splitting to send 5% of production traffic to the new model. Monitor end-user metrics, such as listening time. If end-user metrics improve between models over time, gradually increase the percentage of production traffic sent to the new model.

D: Configure a model monitoring job for the existing Vertex AI endpoint. Configure the monitoring job to detect prediction drift and set a threshold for alerts. Update the model on the endpoint from the previous model to the new model. If you receive an alert of prediction drift, revert to the previous model.

Sample Question 5

You trained a model packaged it with a custom Docker container for serving, and deployed it to Vertex AI Model Registry. When you submit a batch prediction job, it fails with this error: "Error model server never became ready. Please validate that your model file or container configuration are valid. " There are no additional errors in the logs. What should you do?

A: Add a logging configuration to your application to emit logs to Cloud Logging
B: Change the HTTP port in your model's configuration to the default value of 8080
C: Change the healthRoute value in your model's configuration to /healthcheck
D: Pull the Docker image locally, and use the docker run command to launch it locally. Use the docker logs command to explore the error logs

Sample Question 5

You trained a model packaged it with a custom Docker container for serving, and deployed it to Vertex AI Model Registry. When you submit a batch prediction job, it fails with this error: "Error model server never became ready. Please validate that your model file or container configuration are valid. " There are no additional errors in the logs. What should you do?

A: Add a logging configuration to your application to emit logs to Cloud Logging
B: Change the HTTP port in your model's configuration to the default value of 8080
C: Change the healthRoute value in your model's configuration to /healthcheck
D: Pull the Docker image locally, and use the docker run command to launch it locally. Use the docker logs command to explore the error logs

Sample Question 2

You are an ML engineer at a retail company. You have built a model that predicts a coupon to offer an ecommerce customer at checkout based on the items in their cart. When a customer goes to checkout, your serving pipeline, which is hosted on Google Cloud, joins the customer's existing cart with a row in a BigQuery table that contains the customers' historic purchase behavior and uses that as the model's input. The web team is reporting that your model is returning predictions too slowly to load the coupon offer with the rest of the web page. How should you speed up your model's predictions?

A: Attach an NVIDIA P100 GPU to your deployed model's instance.
B: Use a low latency database for the customers' historic purchase behavior.
C: Deploy your model to more instances behind a load balancer to distribute traffic.
D: Create a materialized view in BigQuery with the necessary data for predictions.

Sample Question 2

You are an ML engineer at a retail company. You have built a model that predicts a coupon to offer an ecommerce customer at checkout based on the items in their cart. When a customer goes to checkout, your serving pipeline, which is hosted on Google Cloud, joins the customer's existing cart with a row in a BigQuery table that contains the customers' historic purchase behavior and uses that as the model's input. The web team is reporting that your model is returning predictions too slowly to load the coupon offer with the rest of the web page. How should you speed up your model's predictions?

A: Attach an NVIDIA P100 GPU to your deployed model's instance.
B: Use a low latency database for the customers' historic purchase behavior.
C: Deploy your model to more instances behind a load balancer to distribute traffic.
D: Create a materialized view in BigQuery with the necessary data for predictions.

Sample Question 38

You work for a small company that has deployed an ML model with autoscaling on Vertex AI to serve online predictions in a production environment. The current model receives about 20 prediction requests per hour with an average response time of one second. You have retrained the same model on a new batch of data, and now you are canary testing it, sending ~10% of production traffic to the new model. During this canary test, you notice that prediction requests for your new model are taking between 30 and 180 seconds to complete. What should you do?

A: Submit a request to raise your project quota to ensure that multiple prediction services can run concurrently.
B: Turn off auto-scaling for the online prediction service of your new model. Use manual scaling with one node always available.
C: Remove your new model from the production environment. Compare the new model and existing model codes to identify the cause of the performance bottleneck.
D: Remove your new model from the production environment. For a short trial period, send all incoming prediction requests to BigQuery. Request batch predictions from your new model, and then use the Data Labeling Service to validate your model's performance before promoting it to production.

Sample Question 38

You work for a small company that has deployed an ML model with autoscaling on Vertex AI to serve online predictions in a production environment. The current model receives about 20 prediction requests per hour with an average response time of one second. You have retrained the same model on a new batch of data, and now you are canary testing it, sending ~10% of production traffic to the new model. During this canary test, you notice that prediction requests for your new model are taking between 30 and 180 seconds to complete. What should you do?

A: Submit a request to raise your project quota to ensure that multiple prediction services can run concurrently.
B: Turn off auto-scaling for the online prediction service of your new model. Use manual scaling with one node always available.
C: Remove your new model from the production environment. Compare the new model and existing model codes to identify the cause of the performance bottleneck.
D: Remove your new model from the production environment. For a short trial period, send all incoming prediction requests to BigQuery. Request batch predictions from your new model, and then use the Data Labeling Service to validate your model's performance before promoting it to production.

Sample Question 134

You need to deploy a scikit-leam classification model to production. The model must be able to serve requests 24/7, and you expect millions of requests per second to the production application from 8 am to 7 pm. You need to minimize the cost of deployment. What should you do?

A: Deploy an online Vertex AI prediction endpoint. Set the max replica count to 1
B: Deploy an online Vertex AI prediction endpoint. Set the max replica count to 100
C: Deploy an online Vertex AI prediction endpoint with one GPU per replica. Set the max replica count to 1
D: Deploy an online Vertex AI prediction endpoint with one GPU per replica. Set the max replica count to 100

Sample Question 134

You need to deploy a scikit-leam classification model to production. The model must be able to serve requests 24/7, and you expect millions of requests per second to the production application from 8 am to 7 pm. You need to minimize the cost of deployment. What should you do?

A: Deploy an online Vertex AI prediction endpoint. Set the max replica count to 1
B: Deploy an online Vertex AI prediction endpoint. Set the max replica count to 100
C: Deploy an online Vertex AI prediction endpoint with one GPU per replica. Set the max replica count to 1
D: Deploy an online Vertex AI prediction endpoint with one GPU per replica. Set the max replica count to 100

Sample Question 10

You have recently trained a scikit-learn model that you plan to deploy on Vertex AI. This model will support both online and batch prediction. You need to preprocess input data for model inference. You want to package the model for deployment while minimizing additional code. What should you do?

A: 1. Upload your model to the Vertex AI Model Registry by using a prebuilt scikit-ieam prediction container.
2. Deploy your model to Vertex AI Endpoints, and create a Vertex AI batch prediction job that uses the instanceConfig.instanceType setting to transform your input data.

B: 1. Wrap your model in a custom prediction routine (CPR). and build a container image from the CPR local model. 2. Upload your scikit learn model container to Vertex AI Model Registry.
3. Deploy your model to Vertex AI Endpoints, and create a Vertex AI batch prediction job

C: 1. Create a custom container for your scikit learn model.
2. Define a custom serving function for your model.
3. Upload your model and custom container to Vertex AI Model Registry.
4. Deploy your model to Vertex AI Endpoints, and create a Vertex AI batch prediction job

D: 1. Create a custom container for your scikit learn model.
2. Upload your model and custom container to Vertex AI Model Registry.
3. Deploy your model to Vertex AI Endpoints, and create a Vertex AI batch prediction job that uses the instanceConfig.instanceType setting to transform your input data.

Sample Question 10

You have recently trained a scikit-learn model that you plan to deploy on Vertex AI. This model will support both online and batch prediction. You need to preprocess input data for model inference. You want to package the model for deployment while minimizing additional code. What should you do?

A: 1. Upload your model to the Vertex AI Model Registry by using a prebuilt scikit-ieam prediction container.
2. Deploy your model to Vertex AI Endpoints, and create a Vertex AI batch prediction job that uses the instanceConfig.instanceType setting to transform your input data.

B: 1. Wrap your model in a custom prediction routine (CPR). and build a container image from the CPR local model.
2. Upload your scikit learn model container to Vertex AI Model Registry.
3. Deploy your model to Vertex AI Endpoints, and create a Vertex AI batch prediction job

C: 1. Create a custom container for your scikit learn model.
2. Define a custom serving function for your model.
3. Upload your model and custom container to Vertex AI Model Registry.
4. Deploy your model to Vertex AI Endpoints, and create a Vertex AI batch prediction job

D: 1. Create a custom container for your scikit learn model.
2. Upload your model and custom container to Vertex AI Model Registry.
3. Deploy your model to Vertex AI Endpoints, and create a Vertex AI batch prediction job that uses the instanceConfig.instanceType setting to transform your input data.

Sample Question 62

Your work for a textile manufacturing company. Your company has hundreds of machines, and each machine has many sensors. Your team used the sensory data to build hundreds of ML models that detect machine anomalies. Models are retrained daily, and you need to deploy these models in a cost-effective way. The models must operate 24/7 without downtime and make sub millisecond predictions. What should you do?

A: Deploy a Dataflow batch pipeline and a Vertex AI Prediction endpoint.
B: Deploy a Dataflow batch pipeline with the RunInference API, and use model refresh.
C: Deploy a Dataflow streaming pipeline and a Vertex AI Prediction endpoint with autoscaling.
D: Deploy a Dataflow streaming pipeline with the RunInference API, and use automatic model refresh.

Sample Question 62

Your work for a textile manufacturing company. Your company has hundreds of machines, and each machine has many sensors. Your team used the sensory data to build hundreds of ML models that detect machine anomalies. Models are retrained daily, and you need to deploy these models in a cost-effective way. The models must operate 24/7 without downtime and make sub millisecond predictions. What should you do?

A: Deploy a Dataflow batch pipeline and a Vertex AI Prediction endpoint.
B: Deploy a Dataflow batch pipeline with the RunInference API, and use model refresh.
C: Deploy a Dataflow streaming pipeline and a Vertex AI Prediction endpoint with autoscaling.
D: Deploy a Dataflow streaming pipeline with the RunInference API, and use automatic model refresh.

Sample Question 105

You built a deep learning-based image classification model by using on-premises data. You want to use Vertex AI to deploy the model to production. Due to security concerns, you cannot move your data to the cloud. You are aware that the input data distribution might change over time. You need to detect model performance changes in production. What should you do?

A: Use Vertex Explainable AI for model explainability. Configure feature-based explanations.
B: Use Vertex Explainable AI for model explainability. Configure example-based explanations.
C: Create a Vertex AI Model Monitoring job. Enable training-serving skew detection for your model.
D: Create a Vertex AI Model Monitoring job. Enable feature attribution skew and drift detection for your model.

Sample Question 105

You built a deep learning-based image classification model by using on-premises data. You want to use Vertex AI to deploy the model to production. Due to security concerns, you cannot move your data to the cloud. You are aware that the input data distribution might change over time. You need to detect model performance changes in production. What should you do?

A: Use Vertex Explainable AI for model explainability. Configure feature-based explanations.
B: Use Vertex Explainable AI for model explainability. Configure example-based explanations.
C: Create a Vertex AI Model Monitoring job. Enable training-serving skew detection for your model.
D: Create a Vertex AI Model Monitoring job. Enable feature attribution skew and drift detection for your model.

Sample Question 63

You are using Vertex AI and TensorFlow to develop a custom image classification model. You need the model's decisions and the rationale to be understandable to your company's stakeholders. You also want to explore the results to identify any issues or potential biases. What should you do?

A: 1. Use TensorFlow to generate and visualize features and statistics.
2. Analyze the results together with the standard model evaluation metrics.

B: 1. Use TensorFlow Profiler to visualize the model execution.
2. Analyze the relationship between incorrect predictions and execution bottlenecks.

C: 1. Use Vertex Explainable AI to generate example-based explanations.
2. Visualize the results of sample inputs from the entire dataset together with the standard model evaluation metrics.

D: 1. Use Vertex Explainable AI to generate feature attributions. Aggregate feature attributions over the entire dataset.
2. Analyze the aggregation result together with the standard model evaluation metrics.

Sample Question 63

You are using Vertex AI and TensorFlow to develop a custom image classification model. You need the model's decisions and the rationale to be understandable to your company's stakeholders. You also want to explore the results to identify any issues or potential biases. What should you do?

A: 1. Use TensorFlow to generate and visualize features and statistics.
2. Analyze the results together with the standard model evaluation metrics.

B: 1. Use TensorFlow Profiler to visualize the model execution.
2. Analyze the relationship between incorrect predictions and execution bottlenecks.

C: 1. Use Vertex Explainable AI to generate example-based explanations.
2. Visualize the results of sample inputs from the entire dataset together with the standard model evaluation metrics.

D: 1. Use Vertex Explainable AI to generate feature attributions. Aggregate feature attributions over the entire dataset.
2. Analyze the aggregation result together with the standard model evaluation metrics.

Sample Question 64

You have trained a model by using data that was preprocessed in a batch Dataflow pipeline. Your use case requires real-time inference. You want to ensure that the data preprocessing logic is applied consistently between training and serving. What should you do?

A: Perform data validation to ensure that the input data to the pipeline is the same format as the input data to the endpoint.
B: Refactor the transformation code in the batch data pipeline so that it can be used outside of the pipeline. Use the same code in the endpoint.
C: Refactor the transformation code in the batch data pipeline so that it can be used outside of the pipeline. Share this code with the end users of the endpoint.
D: Batch the real-time requests by using a time window and then use the Dataflow pipeline to preprocess the batched requests. Send the preprocessed requests to the endpoint.

Sample Question 64

You have trained a model by using data that was preprocessed in a batch Dataflow pipeline. Your use case requires real-time inference. You want to ensure that the data preprocessing logic is applied consistently between training and serving. What should you do?

A: Perform data validation to ensure that the input data to the pipeline is the same format as the input data to the endpoint.

B: Refactor the transformation code in the batch data pipeline so that it can be used outside of the pipeline. Use the same code in the endpoint.

C: Refactor the transformation code in the batch data pipeline so that it can be used outside of the pipeline. Share this code with the end users of the endpoint.

D: Batch the real-time requests by using a time window and then use the Dataflow pipeline to preprocess the batched requests. Send the preprocessed requests to the endpoint.

Sample Question 65

You have developed a BigQuery ML model that predicts customer chum, and deployed the model to Vertex AI Endpoints. You want to automate the retraining of your model by using minimal additional code when model feature values change. You also want to minimize the number of times that your model is retrained to reduce training costs. What should you do?

A: 1 Enable request-response logging on Vertex AI Endpoints
2. Schedule a TensorFlow Data Validation job to monitor prediction drift
3. Execute model retraining if there is significant distance between the distributions
B: 1. Enable request-response logging on Vertex AI Endpoints
2. Schedule a TensorFlow Data Validation job to monitor training/serving skew
3. Execute model retraining if there is significant distance between the distributions
C: 1. Create a Vertex AI Model Monitoring job configured to monitor prediction drift
2. Configure alert monitoring to publish a message to a Pub/Sub queue when a monitoring alert is detected
3. Use a Cloud Function to monitor the Pub/Sub queue, and trigger retraining in BigQuery
D: 1. Create a Vertex AI Model Monitoring job configured to monitor training/serving skew
2. Configure alert monitoring to publish a message to a Pub/Sub queue when a monitoring alert is detected
3. Use a Cloud Function to monitor the Pub/Sub queue, and trigger retraining in BigQuery

Sample Question 65

You have developed a BigQuery ML model that predicts customer chum, and deployed the model to Vertex AI Endpoints. You want to automate the retraining of your model by using minimal additional code when model feature values change. You also want to minimize the number of times that your model is retrained to reduce training costs. What should you do?

A: 1 Enable request-response logging on Vertex AI Endpoints
2. Schedule a TensorFlow Data Validation job to monitor prediction drift
3. Execute model retraining if there is significant distance between the distributions
B: 1. Enable request-response logging on Vertex AI Endpoints
2. Schedule a TensorFlow Data Validation job to monitor training/serving skew
3. Execute model retraining if there is significant distance between the distributions
C: 1. Create a Vertex AI Model Monitoring job configured to monitor prediction drift
2. Configure alert monitoring to publish a message to a Pub/Sub queue when a monitoring alert is detected
3. Use a Cloud Function to monitor the Pub/Sub queue, and trigger retraining in BigQuery
D: 1. Create a Vertex AI Model Monitoring job configured to monitor training/serving skew
2. Configure alert monitoring to publish a message to a Pub/Sub queue when a monitoring alert is detected
3. Use a Cloud Function to monitor the Pub/Sub queue, and trigger retraining in BigQuery

Sample Question 27

You received a training-serving skew alert from a Vertex AI Model Monitoring job running in production. You retrained the model with more recent
training data, and deployed it back to the Vertex AI endpoint, but you are still receiving the same alert. What should you do?

A: Update the model monitoring job to use a lower sampling rate.
B: Update the model monitoring job to use the more recent training data that was used to retrain the model.
C: Temporarily disable the alert. Enable the alert again after a sufficient amount of new production traffic has passed through the Vertex AI
endpoint.
D: Temporarily disable the alert until the model can be retrained again on newer training data. Retrain the model again after a sufficient
amount of new production traffic has passed through the Vertex AI endpoint.

Sample Question 27

You received a training-serving skew alert from a Vertex AI Model Monitoring job running in production. You retrained the model with more recent
training data, and deployed it back to the Vertex AI endpoint, but you are still receiving the same alert. What should you do?

A: Update the model monitoring job to use a lower sampling rate.
B: Update the model monitoring job to use the more recent training data that was used to retrain the model.
C: Temporarily disable the alert. Enable the alert again after a sufficient amount of new production traffic has passed through the Vertex AI
endpoint.
D: Temporarily disable the alert until the model can be retrained again on newer training data. Retrain the model again after a sufficient
amount of new production traffic has passed through the Vertex AI endpoint.

Sample Question 135

You developed a custom model by using Vertex AI to forecast the sales of your company's products based on historical transactional data. You anticipate changes in the feature distributions and the correlations between the features in the near future. You also expect to receive a large volume of prediction requests. You plan to use Vertex AI Model Monitoring for drift detection and you want to minimize the cost. What should you do?

A: Use the features for monitoring. Set a monitoring-frequency value that is higher than the default.
B: Use the features for monitoring. Set a prediction-sampling-rate value that is closer to 1 than 0.
C: Use the features and the feature attributions for monitoring. Set a monitoring-frequency value that is lower than the default.
D: Use the features and the feature attributions for monitoring. Set a prediction-sampling-rate value that is closer to 0 than 1.

Sample Question 135

You developed a custom model by using Vertex AI to forecast the sales of your company's products based on historical transactional data. You anticipate changes in the feature distributions and the correlations between the features in the near future. You also expect to receive a large volume of prediction requests. You plan to use Vertex AI Model Monitoring for drift detection and you want to minimize the cost. What should you do?

A: Use the features for monitoring. Set a monitoring-frequency value that is higher than the default.
B: Use the features for monitoring. Set a prediction-sampling-rate value that is closer to 1 than 0.
C: Use the features and the feature attributions for monitoring. Set a monitoring-frequency value that is lower than the default.
D: Use the features and the feature attributions for monitoring. Set a prediction-sampling-rate value that is closer to 0 than 1.

Sample Question 70

You work for a pharmaceutical company based in Canada. Your team developed a BigQuery ML model to predict the number of flu infections for the next month in Canada. Weather data is published weekly, and flu infection statistics are published monthly. You need to configure a model retraining policy that minimizes cost. What should you do?

A: Download the weather and flu data each week. Configure Cloud Scheduler to execute a Vertex AI pipeline to retrain the model weekly.
B: Download the weather and flu data each month. Configure Cloud Scheduler to execute a Vertex AI pipeline to retrain the model monthly.
C: Download the weather and flu data each week. Configure Cloud Scheduler to execute a Vertex AI pipeline to retrain the model every month.
D: Download the weather data each week, and download the flu data each month. Deploy the model to a Vertex AI endpoint with feature drift monitoring, and retrain the model if a monitoring alert is detected.

Sample Question 70

You work for a pharmaceutical company based in Canada. Your team developed a BigQuery ML model to predict the number of flu infections for the next month in Canada. Weather data is published weekly, and flu infection statistics are published monthly. You need to configure a model retraining policy that minimizes cost. What should you do?

A: Download the weather and flu data each week. Configure Cloud Scheduler to execute a Vertex AI pipeline to retrain the model weekly.
B: Download the weather and flu data each month. Configure Cloud Scheduler to execute a Vertex AI pipeline to retrain the model monthly.
C: Download the weather and flu data each week. Configure Cloud Scheduler to execute a Vertex AI pipeline to retrain the model every month.
D: Download the weather data each week, and download the flu data each month. Deploy the model to a Vertex AI endpoint with feature drift monitoring, and retrain the model if a monitoring alert is detected.

Sample Question 71

You are developing a model to help your company create more targeted online advertising campaigns. You need to create a dataset that you will use to train the model. You want to avoid creating or reinforcing unfair bias in the model. What should you do? (Choose two.)

A: Include a comprehensive set of demographic features
B: Include only the demographic groups that most frequently interact with advertisements
C: Collect a random sample of production traffic to build the training dataset
D: Collect a stratified sample of production traffic to build the training dataset

Sample Question 71

You are developing a model to help your company create more targeted online advertising campaigns. You need to create a dataset that you will use to train the model. You want to avoid creating or reinforcing unfair bias in the model. What should you do? (Choose two.)

A: Include a comprehensive set of demographic features
B: Include only the demographic groups that most frequently interact with advertisements
C: Collect a random sample of production traffic to build the training dataset
D: Collect a stratified sample of production traffic to build the training dataset

Sample Question 57

You work for a large bank that serves customers through an application hosted in Google Cloud that is running in the US and Singapore. You have developed a PyTorch model to classify transactions as potentially fraudulent or not. The model is a three-layer perceptron that uses both numerical and categorical features as input, and hashing happens within the model.

You deployed the model to the us-central1 region on nl-highcpu-16 machines, and predictions are served in real time. The model's current median response latency is 40 ms. You want to reduce latency, especially in Singapore, where some customers are experiencing the longest delays. What should you do?

A: Attach an NVIDIA T4 GPU to the machines being used for online inference.
B: Change the machines being used for online inference to nl-highcpu-32.
C: Deploy the model to Vertex AI private endpoints in the us-central1 and asia-southeast1 regions, and allow the application to choose the appropriate endpoint.
D: Create another Vertex AI endpoint in the asia-southeast1 region, and allow the application to choose the appropriate endpoint.

Sample Question 57

You work for a large bank that serves customers through an application hosted in Google Cloud that is running in the US and Singapore. You have developed a PyTorch model to classify transactions as potentially fraudulent or not. The model is a three-layer perceptron that uses both numerical and categorical features as input, and hashing happens within the model.

You deployed the model to the us-central1 region on nl-highcpu-16 machines, and predictions are served in real time. The model's current median response latency is 40 ms. You want to reduce latency, especially in Singapore, where some customers are experiencing the longest delays. What should you do?

A: Attach an NVIDIA T4 GPU to the machines being used for online inference.
B: Change the machines being used for online inference to nl-highcpu-32.
C: Deploy the model to Vertex AI private endpoints in the us-central1 and asia-southeast1 regions, and allow the application to choose the appropriate endpoint.
D: Create another Vertex AI endpoint in the asia-southeast1 region, and allow the application to choose the appropriate endpoint.

Sample Question 77

You work at a mobile gaming startup that creates online multiplayer games. Recently, your company observed an increase in players cheating in the games, leading to a loss of revenue and a poor user experience You built a binary classification model to determine whether a player cheated after a completed game session, and then send a message to other downstream systems to ban the player that cheated. Your model has performed well during testing, and you now need to deploy the model to production. You want your serving solution to provide immediate classifications after a completed game session to avoid further loss of revenue. What should you do?

A: Import the model into Vertex AI Model Registry. Use the Vertex Batch Prediction service to run batch inference jobs.
B: Save the model files in a Cloud Storage bucket. Create a Cloud Function to read the model files and make online inference requests on the Cloud Function.
C: Save the model files in a VM. Load the model files each time there is a prediction request, and run an inference job on the VM
D: Import the model into Vertex AI Model Registry. Create a Vertex AI endpoint that hosts the model, and make online inference requests

Sample Question 77

You work at a mobile gaming startup that creates online multiplayer games. Recently, your company observed an increase in players cheating in the games, leading to a loss of revenue and a poor user experience You built a binary classification model to determine whether a player cheated after a completed game session, and then send a message to other downstream systems to ban the player that cheated. Your model has performed well during testing, and you now need to deploy the model to production. You want your serving solution to provide immediate classifications after a completed game session to avoid further loss of revenue. What should you do?

A: Import the model into Vertex AI Model Registry. Use the Vertex Batch Prediction service to run batch inference jobs.
B: Save the model files in a Cloud Storage bucket. Create a Cloud Function to read the model files and make online inference requests on the Cloud Function.
C: Save the model files in a VM. Load the model files each time there is a prediction request, and run an inference job on the VM
D: Import the model into Vertex AI Model Registry. Create a Vertex AI endpoint that hosts the model, and make online inference requests

Sample Question 79

You work for a manufacturing company. You need to train a custom image classification model to detect product defects at the end of an assembly line. Although your model is performing well, some images in your holdout set are consistently mislabeled with high confidence. You want to use Vertex AI to understand your model's results. What should you do?

A: Configure feature-based explanations by using Integrated Gradients. Set visualization type to PIXELS, and set clip_percent_upperbound to 95.
B: Create an index by using Vertex AI Matching Engine. Query the index with your mislabeled images.
C: Configure feature-based explanations by using XRAI. Set visualization type to OUTLINES, and set polarity to positive.
D: Configure example-based explanations. Specify the embedding output layer to be used for the latent space representation.

Sample Question 79

You work for a manufacturing company. You need to train a custom image classification model to detect product defects at the end of an assembly line. Although your model is performing well, some images in your holdout set are consistently mislabeled with high confidence. You want to use Vertex AI to understand your model's results. What should you do?

A: Configure feature-based explanations by using Integrated Gradients. Set visualization type to PIXELS, and set clip_percent_upperbound to 95.

B: Create an index by using Vertex AI Matching Engine. Query the index with your mislabeled images.

C: Configure feature-based explanations by using XRAI. Set visualization type to OUTLINES, and set polarity to positive.

D: Configure example-based explanations. Specify the embedding output layer to be used for the latent space representation.

Sample Question 80

You developed a custom model by using Vertex AI to predict your application's user churn rate. You are using Vertex AI Model Monitoring for skew detection. The training data stored in BigQuery contains two sets of features - demographic and behavioral. You later discover that two separate models trained on each set perform better than the original model. You need to configure a new model monitoring pipeline that splits traffic among the two models. You want to use the same prediction-sampling-rate and monitoring-frequency for each model. You also want to minimize management effort. What should you do?

A: Keep the training dataset as is. Deploy the models to two separate endpoints, and submit two Vertex AI Model Monitoring jobs with appropriately selected feature-thresholds parameters.
B: Keep the training dataset as is. Deploy both models to the same endpoint and submit a Vertex AI Model Monitoring job with a monitoring-config-from-file parameter that accounts for the model IDs and feature selections.
C: Separate the training dataset into two tables based on demographic and behavioral features. Deploy the models to two separate endpoints, and submit two Vertex AI Model Monitoring jobs.
D: Separate the training dataset into two tables based on demographic and behavioral features. Deploy both models to the same endpoint, and submit a Vertex AI Model Monitoring job with a monitoring-config-from-file parameter that accounts for the model IDs and training datasets.

Sample Question 80

You developed a custom model by using Vertex AI to predict your application's user churn rate. You are using Vertex AI Model Monitoring for skew detection. The training data stored in BigQuery contains two sets of features - demographic and behavioral. You later discover that two separate models trained on each set perform better than the original model. You need to configure a new model monitoring pipeline that splits traffic among the two models. You want to use the same prediction-sampling-rate and monitoring-frequency for each model. You also want to minimize management effort. What should you do?

A: Keep the training dataset as is. Deploy the models to two separate endpoints, and submit two Vertex AI Model Monitoring jobs with appropriately selected feature-thresholds parameters.
B: Keep the training dataset as is. Deploy both models to the same endpoint and submit a Vertex AI Model Monitoring job with a monitoring-config-from-file parameter that accounts for the model IDs and feature selections.
C: Separate the training dataset into two tables based on demographic and behavioral features. Deploy the models to two separate endpoints, and submit two Vertex AI Model Monitoring jobs.
D: Separate the training dataset into two tables based on demographic and behavioral features. Deploy both models to the same endpoint, and submit a Vertex AI Model Monitoring job with a monitoring-config-from-file parameter that accounts for the model IDs and training datasets.

Sample Question 108

You work for a retail company that is using a regression model built with BigQuery ML to predict product sales. This model is being used to serve online predictions. Recently you developed a new version of the model that uses a different architecture (custom model). Initial analysis revealed that both models are performing as expected. You want to deploy the new version of the model to production and monitor the performance over the next two months. You need to minimize the impact to the existing and future model users. How should you deploy the model?

A: Import the new model to the same Vertex AI Model Registry as a different version of the existing model. Deploy the new model to the same Vertex AI endpoint as the existing model, and use traffic splitting to route 95% of production traffic to the BigQuery ML model and 5% of production traffic to the new model.
B: Import the new model to the same Vertex AI Model Registry as the existing model. Deploy the models to one Vertex AI endpoint. Route 95% of production traffic to the BigQuery ML model and 5% of production traffic to the new model.

C: Import the new model to the same Vertex AI Model Registry as the existing model. Deploy each model to a separate Vertex AI endpoint.
D: Deploy the new model to a separate Vertex AI endpoint. Create a Cloud Run service that routes the prediction requests to the corresponding endpoints based on the input feature values.

Sample Question 108

You work for a retail company that is using a regression model built with BigQuery ML to predict product sales. This model is being used to serve online predictions. Recently you developed a new version of the model that uses a different architecture (custom model). Initial analysis revealed that both models are performing as expected. You want to deploy the new version of the model to production and monitor the performance over the next two months. You need to minimize the impact to the existing and future model users. How should you deploy the model?

A: Import the new model to the same Vertex AI Model Registry as a different version of the existing model. Deploy the new model to the same Vertex AI endpoint as the existing model, and use traffic splitting to route 95% of production traffic to the BigQuery ML model and 5% of production traffic to the new model.

B: Import the new model to the same Vertex AI Model Registry as the existing model. Deploy the models to one Vertex AI endpoint. Route 95% of production traffic to the BigQuery ML model and 5% of production traffic to the new model.

C: Import the new model to the same Vertex AI Model Registry as the existing model. Deploy each model to a separate Vertex AI endpoint.

D: Deploy the new model to a separate Vertex AI endpoint. Create a Cloud Run service that routes the prediction requests to the corresponding endpoints based on the input feature values.

Sample Question 82

You have built a custom model that performs several memory-intensive preprocessing tasks before it makes a prediction. You deployed the model to a Vertex AI endpoint, and validated that results were received in a reasonable amount of time. After routing user traffic to the endpoint, you discover that the endpoint does not autoscale as expected when receiving multiple requests. What should you do?

A: Use a machine type with more memory
B: Decrease the number of workers per machine
C: Increase the CPU utilization target in the autoscaling configurations.
D: Decrease the CPU utilization target in the autoscaling configurations

Sample Question 82

You have built a custom model that performs several memory-intensive preprocessing tasks before it makes a prediction. You deployed the model to a Vertex AI endpoint, and validated that results were received in a reasonable amount of time. After routing user traffic to the endpoint, you discover that the endpoint does not autoscale as expected when receiving multiple requests. What should you do?

A: Use a machine type with more memory
B: Decrease the number of workers per machine
C: Increase the CPU utilization target in the autoscaling configurations.
D: Decrease the CPU utilization target in the autoscaling configurations

Sample Question 40

You have deployed a scikit-team model to a Vertex AI endpoint using a custom model server. You enabled autoscaling: however, the deployed model fails to scale beyond one replica, which led to dropped requests. You notice that CPU utilization remains low even during periods of high load. What should you do?

A: Attach a GPU to the prediction nodes
B: Increase the number of workers in your model server
C: Schedule scaling of the nodes to match expected demand
D: Increase the minReplicaCount in your DeployedModel configuration

Sample Question 40

You have deployed a scikit-team model to a Vertex AI endpoint using a custom model server. You enabled autoscaling: however, the deployed model fails to scale beyond one replica, which led to dropped requests. You notice that CPU utilization remains low even during periods of high load. What should you do?

A: Attach a GPU to the prediction nodes
B: Increase the number of workers in your model server
C: Schedule scaling of the nodes to match expected demand
D: Increase the minReplicaCount in your DeployedModel configuration

Sample Question 89

You recently deployed a scikit-learn model to a Vertex AI endpoint. You are now testing the model on live production traffic. While monitoring the endpoint, you discover twice as many requests per hour than expected throughout the day. You want the endpoint to efficiently scale when the demand increases in the future to prevent users from experiencing high latency. What should you do?

A: Deploy two models to the same endpoint, and distribute requests among them evenly
B: Configure an appropriate minReplicaCount value based on expected baseline traffic
C: Set the target utilization percentage in the autoscailngMetricSpecs configuration to a higher value
D: Change the model's machine type to one that utilizes GPUs

Sample Question 89

You recently deployed a scikit-learn model to a Vertex AI endpoint. You are now testing the model on live production traffic. While monitoring the endpoint, you discover twice as many requests per hour than expected throughout the day. You want the endpoint to efficiently scale when the demand increases in the future to prevent users from experiencing high latency. What should you do?

A: Deploy two models to the same endpoint, and distribute requests among them evenly
B: Configure an appropriate minReplicaCount value based on expected baseline traffic
C: Set the target utilization percentage in the autoscailngMetricSpecs configuration to a higher value
D: Change the model's machine type to one that utilizes GPUs

Sample Question 90

You work for a bank. You have created a custom model to predict whether a loan application should be flagged for human review. The input features are stored in a BigQuery table. The model is performing well, and you plan to deploy it to production. Due to compliance requirements the model must provide explanations for each prediction. You want to add this functionality to your model code with minimal effort and provide explanations that are as accurate as possible. What should you do?

A: Create an AutoML tabular model by using the BigQuery data with integrated Vertex Explainable AI
B: Create a BigQuery ML deep neural network model and use the ML.EXPLAIN_PREDICT method with the num_integral_steps parameter.
C: Upload the custom model to Vertex AI Model Registry and configure feature-based attribution by using sampled Shapley with input baselines.
D: Update the custom serving container to include sampled Shapley-based explanations in the prediction outputs.

Sample Question 90

You work for a bank. You have created a custom model to predict whether a loan application should be flagged for human review. The input features are stored in a BigQuery table. The model is performing well, and you plan to deploy it to production. Due to compliance requirements the model must provide explanations for each prediction. You want to add this functionality to your model code with minimal effort and provide explanations that are as accurate as possible. What should you do?

A: Create an AutoML tabular model by using the BigQuery data with integrated Vertex Explainable AI
B: Create a BigQuery ML deep neural network model and use the ML.EXPLAIN_PREDICT method with the num_integral_steps parameter.
C: Upload the custom model to Vertex AI Model Registry and configure feature-based attribution by using sampled Shapley with input baselines.
D: Update the custom serving container to include sampled Shapley-based explanations in the prediction outputs.

Sample Question 91

You recently used BigQuery ML to train an AutoML regression model. You shared results with your team and received positive feedback. You need to deploy your model for online prediction as quickly as possible. What should you do?

A: Retrain the model by using BigQuery ML, and specify Vertex AI as the model registry. Deploy the model from Vertex AI Model Registry to a Vertex AI endpoint
B: Retrain the model by using Vertex Al Deploy the model from Vertex AI Model. Registry to a Vertex AI endpoint.
C: Alter the model by using BigQuery ML, and specify Vertex AI as the model registry. Deploy the model from Vertex AI Model Registry to a Vertex AI endpoint
D: Export the model from BigQuery ML to Cloud Storage. Import the model into Vertex AI Model Registry. Deploy the model to a Vertex AI endpoint.

Sample Question 91

You recently used BigQuery ML to train an AutoML regression model. You shared results with your team and received positive feedback. You need to deploy your model for online prediction as quickly as possible. What should you do?

A: Retrain the model by using BigQuery ML, and specify Vertex AI as the model registry. Deploy the model from Vertex AI Model Registry to a Vertex AI endpoint

B: Retrain the model by using Vertex AI Deploy the model from Vertex AI Model. Registry to a Vertex AI endpoint.

C: Alter the model by using BigQuery ML, and specify Vertex AI as the model registry. Deploy the model from Vertex AI Model Registry to a Vertex AI endpoint

D: Export the model from BigQuery ML to Cloud Storage. Import the model into Vertex AI Model Registry. Deploy the model to a Vertex AI endpoint.

Sample Question 94

You work at a bank. You have a custom tabular ML model that was provided by the bank's vendor. The training data is not available due to its sensitivity. The model is packaged as a Vertex AI Model serving container, which accepts a string as input for each prediction instance. In each string, the feature values are separated by commas. You want to deploy this model to production for online predictions and monitor the feature distribution over time with minimal effort. What should you do?

A: 1. Upload the model to Vertex AI Model Registry, and deploy the model to a Vertex AI endpoint
2. Create a Vertex AI Model Monitoring job with feature drift detection as the monitoring objective, and provide an instance schema

B: 1. Upload the model to Vertex AI Model Registry, and deploy the model to a Vertex AI endpoint
2. Create a Vertex AI Model Monitoring job with feature skew detection as the monitoring objective, and provide an instance schema

C: 1. Refactor the serving container to accept key-value pairs as input format
2. Upload the model to Vertex AI Model Registry, and deploy the model to a Vertex AI endpoint
3. Create a Vertex AI Model Monitoring job with feature drift detection as the monitoring objective.

D: 1. Refactor the serving container to accept key-value pairs as input format
2. Upload the model to Vertex AI Model Registry, and deploy the model to a Vertex AI endpoint
3. Create a Vertex AI Model Monitoring job with feature skew detection as the monitoring objective

Sample Question 94

You work at a bank. You have a custom tabular ML model that was provided by the bank's vendor. The training data is not available due to its sensitivity. The model is packaged as a Vertex AI Model serving container, which accepts a string as input for each prediction instance. In each string, the feature values are separated by commas. You want to deploy this model to production for online predictions and monitor the feature distribution over time with minimal effort. What should you do?

A: 1. Upload the model to Vertex AI Model Registry, and deploy the model to a Vertex AI endpoint
2. Create a Vertex AI Model Monitoring job with feature drift detection as the monitoring objective, and provide an instance schema

B: 1. Upload the model to Vertex AI Model Registry, and deploy the model to a Vertex AI endpoint
2. Create a Vertex AI Model Monitoring job with feature skew detection as the monitoring objective, and provide an instance schema

C: 1. Refactor the serving container to accept key-value pairs as input format
2. Upload the model to Vertex AI Model Registry, and deploy the model to a Vertex AI endpoint
3. Create a Vertex AI Model Monitoring job with feature drift detection as the monitoring objective.

D: 1. Refactor the serving container to accept key-value pairs as input format
2. Upload the model to Vertex AI Model Registry, and deploy the model to a Vertex AI endpoint
3. Create a Vertex AI Model Monitoring job with feature skew detection as the monitoring objective

Sample Question 96

You are training and deploying updated versions of a regression model with tabular data by using Vertex AI Pipelines, Vertex AI Training, Vertex AI Experiments, and Vertex AI Endpoints. The model is deployed in a Vertex AI endpoint, and your users call the model by using the Vertex AI endpoint. You want to receive an email when the feature data distribution changes significantly, so you can retrigger the training pipeline and deploy an updated version of your model. What should you do?

A: Use Vertex AI Model Monitoring. Enable prediction drift monitoring on the endpoint, and specify a notification email.
B: In Cloud Logging, create a logs-based alert using the logs in the Vertex AI endpoint. Configure Cloud Logging to send an email when the alert is triggered.
C: In Cloud Monitoring create a logs-based metric and a threshold alert for the metric. Configure Cloud Monitoring to send an email when the alert is triggered.
D: Export the container logs of the endpoint to BigQuery. Create a Cloud Function to run a SQL query over the exported logs and send an email. Use Cloud Scheduler to trigger the Cloud Function.

Sample Question 96

You are training and deploying updated versions of a regression model with tabular data by using Vertex AI Pipelines, Vertex AI Training, Vertex AI Experiments, and Vertex AI Endpoints. The model is deployed in a Vertex AI endpoint, and your users call the model by using the Vertex AI endpoint. You want to receive an email when the feature data distribution changes significantly, so you can retrigger the training pipeline and deploy an updated version of your model. What should you do?

A: Use Vertex AI Model Monitoring. Enable prediction drift monitoring on the endpoint, and specify a notification email.

B: In Cloud Logging, create a logs-based alert using the logs in the Vertex AI endpoint. Configure Cloud Logging to send an email when the alert is triggered.

C: In Cloud Monitoring create a logs-based metric and a threshold alert for the metric. Configure Cloud Monitoring to send an email when the alert is triggered.

D: Export the container logs of the endpoint to BigQuery. Create a Cloud Function to run a SQL query over the exported logs and send an email. Use Cloud Scheduler to trigger the Cloud Function.

Sample Question 15

You recently used XGBoost to train a model in Python that will be used for online serving. Your model prediction service will be called by a backend service implemented in Golang running on a Google Kubernetes Engine (GKE) cluster. Your model requires pre and postprocessing steps. You need to implement the processing steps so that they run at serving time. You want to minimize code changes and infrastructure maintenance, and deploy your model into production as quickly as possible. What should you do?

A: Use FastAPI to implement an HTTP server. Create a Docker image that runs your HTTP server, and deploy it on your organization's GKE cluster.
B: Use FastAPI to implement an HTTP server. Create a Docker image that runs your HTTP server, Upload the image to Vertex AI Model Registry and deploy it to a Vertex AI endpoint.
C: Use the Predictor interface to implement a custom prediction routine. Build the custom container, upload the container to Vertex AI Model Registry and deploy it to a Vertex AI endpoint.
D: Use the XGBoost prebuilt serving container when importing the trained model into Vertex AI. Deploy the model to a Vertex AI endpoint. Work with the backend engineers to implement the pre- and postprocessing steps in the Golang backend service.

Sample Question 15

You recently used XGBoost to train a model in Python that will be used for online serving. Your model prediction service will be called by a backend service implemented in Golang running on a Google Kubernetes Engine (GKE) cluster. Your model requires pre and postprocessing steps. You need to implement the processing steps so that they run at serving time. You want to minimize code changes and infrastructure maintenance, and deploy your model into production as quickly as possible. What should you do?

A: Use FastAPI to implement an HTTP server. Create a Docker image that runs your HTTP server, and deploy it on your organization's GKE cluster.
B: Use FastAPI to implement an HTTP server. Create a Docker image that runs your HTTP server, Upload the image to Vertex AI Model Registry and deploy it to a Vertex AI endpoint.
C: Use the Predictor interface to implement a custom prediction routine. Build the custom container, upload the container to Vertex AI Model Registry and deploy it to a Vertex AI endpoint.
D: Use the XGBoost prebuilt serving container when importing the trained model into Vertex AI. Deploy the model to a Vertex AI endpoint. Work with the backend engineers to implement the pre- and postprocessing steps in the Golang backend service.

Sample Question 28

You have developed a BigQuery ML model that predicts customer chum, and deployed the model to Vertex AI Endpoints. You want to automate the retraining of your model by using minimal additional code when model feature values change. You also want to minimize the number of times that your model is retrained to reduce training costs. What should you do?

A: 1 Enable request-response logging on Vertex AI Endpoints
2. Schedule a TensorFlow Data Validation job to monitor prediction drift
3. Execute model retraining if there is significant distance between the distributions
B: 1. Enable request-response logging on Vertex AI Endpoints
2. Schedule a TensorFlow Data Validation job to monitor training/serving skew
3. Execute model retraining if there is significant distance between the distributions
C: 1. Create a Vertex AI Model Monitoring job configured to monitor prediction drift
2. Configure alert monitoring to publish a message to a Pub/Sub queue when a monitoring alert is detected
3. Use a Cloud Function to monitor the Pub/Sub queue, and trigger retraining in BigQuery
D: 1. Create a Vertex AI Model Monitoring job configured to monitor training/serving skew
2. Configure alert monitoring to publish a message to a Pub/Sub queue when a monitoring alert is detected
3. Use a Cloud Function to monitor the Pub/Sub queue, and trigger retraining in BigQuery

Sample Question 28

You have developed a BigQuery ML model that predicts customer chum, and deployed the model to Vertex AI Endpoints. You want to automate the retraining of your model by using minimal additional code when model feature values change. You also want to minimize the number of times that your model is retrained to reduce training costs. What should you do?

A: 1 Enable request-response logging on Vertex AI Endpoints
2. Schedule a TensorFlow Data Validation job to monitor prediction drift
3. Execute model retraining if there is significant distance between the distributions

B: 1. Enable request-response logging on Vertex AI Endpoints
2. Schedule a TensorFlow Data Validation job to monitor training/serving skew
3. Execute model retraining if there is significant distance between the distributions

C: 1. Create a Vertex AI Model Monitoring job configured to monitor prediction drift
2. Configure alert monitoring to publish a message to a Pub/Sub queue when a monitoring alert is detected
3. Use a Cloud Function to monitor the Pub/Sub queue, and trigger retraining in BigQuery

D: 1. Create a Vertex AI Model Monitoring job configured to monitor training/serving skew
2. Configure alert monitoring to publish a message to a Pub/Sub queue when a monitoring alert is detected
3. Use a Cloud Function to monitor the Pub/Sub queue, and trigger retraining in BigQuery

Sample Question 12

You are using Vertex AI and TensorFlow to develop a custom image classification model. You need the model's decisions and the rationale to be understandable to your company's stakeholders. You also want to explore the results to identify any issues or potential biases. What should you do?

A: 1. Use TensorFlow to generate and visualize features and statistics.
2. Analyze the results together with the standard model evaluation metrics.

B: 1. Use TensorFlow Profiler to visualize the model execution.
2. Analyze the relationship between incorrect predictions and execution bottlenecks.

C: 1. Use Vertex Explainable AI to generate example-based explanations.
2. Visualize the results of sample inputs from the entire dataset together with the standard model evaluation metrics.

D: 1. Use Vertex Explainable AI to generate feature attributions. Aggregate feature attributions over the entire dataset.
2. Analyze the aggregation result together with the standard model evaluation metrics.

Sample Question 12

You are using Vertex AI and TensorFlow to develop a custom image classification model. You need the model's decisions and the rationale to be understandable to your company's stakeholders. You also want to explore the results to identify any issues or potential biases. What should you do?

A: 1. Use TensorFlow to generate and visualize features and statistics.
2. Analyze the results together with the standard model evaluation metrics.

B: 1. Use TensorFlow Profiler to visualize the model execution.
2. Analyze the relationship between incorrect predictions and execution bottlenecks.

C: 1. Use Vertex Explainable AI to generate example-based explanations.
2. Visualize the results of sample inputs from the entire dataset together with the standard model evaluation metrics.

D: 1. Use Vertex Explainable AI to generate feature attributions. Aggregate feature attributions over the entire dataset.
2. Analyze the aggregation result together with the standard model evaluation metrics.

Sample Question 103

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

Sample Question 103

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

Sample Question 104

You work for a retail company. You have created a Vertex AI forecast model that produces monthly item sales predictions. You want to quickly create a report that will help to explain how the model calculates the predictions. You have one month of recent actual sales data that was not included in the training dataset. How should you generate data for your report?

A: Create a batch prediction job by using the actual sales data. Compare the predictions to the actuals in the report.

B: Create a batch prediction job by using the actual sales data, and configure the job settings to generate feature attributions. Compare the results in the report.

C: Generate counterfactual examples by using the actual sales data. Create a batch prediction job using the actual sales data and the counterfactual examples. Compare the results in the report.

D: Train another model by using the same training dataset as the original, and exclude some columns. Using the actual sales data create one batch prediction job by using the new model and another one with the original model. Compare the two sets of predictions in the report.

Sample Question 104

You work for a retail company. You have created a Vertex AI forecast model that produces monthly item sales predictions. You want to quickly create a report that will help to explain how the model calculates the predictions. You have one month of recent actual sales data that was not included in the training dataset. How should you generate data for your report?

A: Create a batch prediction job by using the actual sales data. Compare the predictions to the actuals in the report.

B: Create a batch prediction job by using the actual sales data, and configure the job settings to generate feature attributions. Compare the results in the report.

C: Generate counterfactual examples by using the actual sales data. Create a batch prediction job using the actual sales data and the counterfactual examples. Compare the results in the report.

D: Train another model by using the same training dataset as the original, and exclude some columns. Using the actual sales data create one batch prediction job by using the new model and another one with the original model. Compare the two sets of predictions in the report.

Sample Question 17

You recently deployed a model to a Vertex AI endpoint. Your data drifts frequently, so you have enabled request-response logging and created a Vertex AI Model Monitoring job. You have observed that your model is receiving higher traffic than expected. You need to reduce the model monitoring cost while continuing to quickly detect drift. What should you do?

A: Replace the monitoring job with a DataFlow pipeline that uses TensorFlow Data Validation (TFDV)
B: Replace the monitoring job with a custom SQL script to calculate statistics on the features and predictions in BigQuery
C: Decrease the sample_rate parameter in the RandomSampleConfig of the monitoring job
D: Increase the monitor_interval parameter in the ScheduleConfig of the monitoring job

Sample Question 17

You recently deployed a model to a Vertex AI endpoint. Your data drifts frequently, so you have enabled request-response logging and created a Vertex AI Model Monitoring job. You have observed that your model is receiving higher traffic than expected. You need to reduce the model monitoring cost while continuing to quickly detect drift. What should you do?

A: Replace the monitoring job with a DataFlow pipeline that uses TensorFlow Data Validation (TFDV)
B: Replace the monitoring job with a custom SQL script to calculate statistics on the features and predictions in BigQuery
C: Decrease the sample_rate parameter in the RandomSampleConfig of the monitoring job
D: Increase the monitor_interval parameter in the ScheduleConfig of the monitoring job

Sample Question 14

You are deploying a new version of a model to a production Vertex AI endpoint that is serving traffic. You plan to direct all user traffic to the new model. You need to deploy the model with minimal disruption to your application. What should you do?

A: 1. Create a new endpoint
2. Create a new model. Set it as the default version. Upload the model to Vertex AI Model Registry
3. Deploy the new model to the new endpoint
4. Update Cloud DNS to point to the new endpoint

B: 1. Create a new endpoint
2. Create a new model. Set the parentModel parameter to the model ID of the currently deployed model and set it as the default version. Upload the model to Vertex AI Model Registry
3. Deploy the new model to the new endpoint, and set the new model to 100% of the traffic.

C: 1. Create a new model. Set the parentModel parameter to the model ID of the currently deployed model. Upload the model to Vertex AI Model Registry.
2. Deploy the new model to the existing endpoint, and set the new model to 100% of the traffic

D: 1. Create a new model. Set it as the default version. Upload the model to Vertex AI Model Registry
2. Deploy the new model to the existing endpoint

Sample Question 14

You are deploying a new version of a model to a production Vertex AI endpoint that is serving traffic. You plan to direct all user traffic to the new model. You need to deploy the model with minimal disruption to your application. What should you do?

A: 1. Create a new endpoint
2. Create a new model. Set it as the default version. Upload the model to Vertex AI Model Registry
3. Deploy the new model to the new endpoint
4. Update Cloud DNS to point to the new endpoint

B: 1. Create a new endpoint
2. Create a new model. Set the parentModel parameter to the model ID of the currently deployed model and set it as the default version. Upload the model to Vertex AI Model Registry
3. Deploy the new model to the new endpoint, and set the new model to 100% of the traffic.

C: 1. Create a new model. Set the parentModel parameter to the model ID of the currently deployed model. Upload the model to Vertex AI Model Registry.
2. Deploy the new model to the existing endpoint, and set the new model to 100% of the traffic

D: 1. Create a new model. Set it as the default version. Upload the model to Vertex AI Model Registry
2. Deploy the new model to the existing endpoint