

Google

Exam Questions Professional-Machine-Learning-Engineer

Google Professional Machine Learning Engineer



NEW QUESTION 1

You want to rebuild your ML pipeline for structured data on Google Cloud. You are using PySpark to conduct data transformations at scale, but your pipelines are taking over 12 hours to run. To speed up development and pipeline run time, you want to use a serverless tool and SQL syntax. You have already moved your raw data into Cloud Storage. How should you build the pipeline on Google Cloud while meeting the speed and processing requirements?

- A. Use Data Fusion's GUI to build the transformation pipelines, and then write the data into BigQuery
- B. Convert your PySpark into SparkSQL queries to transform the data and then run your pipeline on Dataproc to write the data into BigQuery.
- C. Ingest your data into Cloud SQL convert your PySpark commands into SQL queries to transform the data, and then use federated queries from BigQuery for machine learning
- D. Ingest your data into BigQuery using BigQuery Load, convert your PySpark commands into BigQuery SQL queries to transform the data, and then write the transformations to a new table

Answer: B

NEW QUESTION 2

Your organization wants to make its internal shuttle service route more efficient. The shuttles currently stop at all pick-up points across the city every 30 minutes between 7 am and 10 am. The development team has already built an application on Google Kubernetes Engine that requires users to confirm their presence and shuttle station one day in advance. What approach should you take?

- A. 1. Build a tree-based regression model that predicts how many passengers will be picked up at each shuttle station.* 2. Dispatch an appropriately sized shuttle and provide the map with the required stops based on the prediction.
- B. 1. Build a tree-based classification model that predicts whether the shuttle should pick up passengers at each shuttle station.* 2. Dispatch an available shuttle and provide the map with the required stops based on the prediction
- C. 1. Define the optimal route as the shortest route that passes by all shuttle stations with confirmed attendance at the given time under capacity constraints.* 2 Dispatch an appropriately sized shuttle and indicate the required stops on the map
- D. 1. Build a reinforcement learning model with tree-based classification models that predict the presence of passengers at shuttle stops as agents and a reward function around a distance-based metric* 2. Dispatch an appropriately sized shuttle and provide the map with the required stops based on the simulated outcome.

Answer: D

NEW QUESTION 3

You recently designed and built a custom neural network that uses critical dependencies specific to your organization's framework. You need to train the model using a managed training service on Google Cloud. However, the ML framework and related dependencies are not supported by AI Platform Training. Also, both your model and your data are too large to fit in memory on a single machine. Your ML framework of choice uses the scheduler, workers, and servers distribution structure. What should you do?

- A. Use a built-in model available on AI Platform Training
- B. Build your custom container to run jobs on AI Platform Training
- C. Build your custom containers to run distributed training jobs on AI Platform Training
- D. Reconfigure your code to a ML framework with dependencies that are supported by AI Platform Training

Answer: C

NEW QUESTION 4

You need to train a computer vision model that predicts the type of government ID present in a given image using a GPU-powered virtual machine on Compute Engine. You use the following parameters:

- Optimizer: SGD
- Image shape = 224x224
- Batch size = 64
- Epochs = 10
- Verbose = 2

During training you encounter the following error: ResourceExhaustedError: out of Memory (oom) when allocating tensor. What should you do?

- A. Change the optimizer
- B. Reduce the batch size
- C. Change the learning rate
- D. Reduce the image shape

Answer: A

NEW QUESTION 5

You work for a large hotel chain and have been asked to assist the marketing team in gathering predictions for a targeted marketing strategy. You need to make predictions about user lifetime value (LTV) over the next 30 days so that marketing can be adjusted accordingly. The customer dataset is in BigQuery, and you are preparing the tabular data for training with AutoML Tables. This data has a time signal that is spread across multiple columns. How should you ensure that AutoML fits the best model to your data?

- A. Manually combine all columns that contain a time signal into an array Allow AutoML to interpret this array appropriately Choose an automatic data split across the training, validation, and testing sets
- B. Submit the data for training without performing any manual transformations Allow AutoML to handle the appropriate transformations Choose an automatic data split across the training, validation, and testing sets
- C. Submit the data for training without performing any manual transformations, and indicate an appropriate column as the Time column Allow AutoML to split your data based on the time signal provided, and reserve the more recent data for the validation and testing sets
- D. Submit the data for training without performing any manual transformations Use the columns that have a time signal to manually split your data Ensure that the data in your validation set is from 30 days after the data in your training set and that the data in your testing set is from 30 days after your validation set

Answer: D

NEW QUESTION 6

You need to design a customized deep neural network in Keras that will predict customer purchases based on their purchase history. You want to explore model performance using multiple model architectures, store training data, and be able to compare the evaluation metrics in the same dashboard. What should you do?

- A. Create multiple models using AutoML Tables
- B. Automate multiple training runs using Cloud Composer
- C. Run multiple training jobs on AI Platform with similar job names
- D. Create an experiment in Kubeflow Pipelines to organize multiple runs

Answer: C

NEW QUESTION 7

You work for a toy manufacturer that has been experiencing a large increase in demand. You need to build an ML model to reduce the amount of time spent by quality control inspectors checking for product defects. Faster defect detection is a priority. The factory does not have reliable Wi-Fi. Your company wants to implement the new ML model as soon as possible. Which model should you use?

- A. AutoML Vision model
- B. AutoML Vision Edge mobile-versatile-1 model
- C. AutoML Vision Edge mobile-low-latency-1 model
- D. AutoML Vision Edge mobile-high-accuracy-1 model

Answer: A

NEW QUESTION 8

Your team has been tasked with creating an ML solution in Google Cloud to classify support requests for one of your platforms. You analyzed the requirements and decided to use TensorFlow to build the classifier so that you have full control of the model's code, serving, and deployment. You will use Kubeflow pipelines for the ML platform. To save time, you want to build on existing resources and use managed services instead of building a completely new model. How should you build the classifier?

- A. Use the Natural Language API to classify support requests
- B. Use AutoML Natural Language to build the support requests classifier
- C. Use an established text classification model on AI Platform to perform transfer learning
- D. Use an established text classification model on AI Platform as-is to classify support requests

Answer: D

NEW QUESTION 9

You are developing a Kubeflow pipeline on Google Kubernetes Engine. The first step in the pipeline is to issue a query against BigQuery. You plan to use the results of that query as the input to the next step in your pipeline. You want to achieve this in the easiest way possible. What should you do?

- A. Use the BigQuery console to execute your query and then save the query results into a new BigQuery table.
- B. Write a Python script that uses the BigQuery API to execute queries against BigQuery. Execute this script as the first step in your Kubeflow pipeline.
- C. Use the Kubeflow Pipelines domain-specific language to create a custom component that uses the Python BigQuery client library to execute queries.
- D. Locate the Kubeflow Pipelines repository on GitHub. Find the BigQuery Query Component, copy that component's URL, and use it to load the component into your pipeline.
- E. Use the component to execute queries against BigQuery.

Answer: A

NEW QUESTION 10

You developed an ML model with AI Platform, and you want to move it to production. You serve a few thousand queries per second and are experiencing latency issues. Incoming requests are served by a load balancer that distributes them across multiple Kubeflow CPU-only pods running on Google Kubernetes Engine (GKE). Your goal is to improve the serving latency without changing the underlying infrastructure. What should you do?

- A. Significantly increase the max_batch_size TensorFlow Serving parameter
- B. Switch to the tensorflow-model-server-universal version of TensorFlow Serving
- C. Significantly increase the max_enqueued_batches TensorFlow Serving parameter
- D. Recompile TensorFlow Serving using the source to support CPU-specific optimizations. Instruct GKE to choose an appropriate baseline minimum CPU platform for serving nodes.

Answer: A

NEW QUESTION 10

Your team is building an application for a global bank that will be used by millions of customers. You built a forecasting model that predicts customers' account balances 3 days in the future. Your team will use the results in a new feature that will notify users when their account balance is likely to drop below \$25. How should you serve your predictions?

- A. 1. Create a Pub/Sub topic for each user* 2. Deploy a Cloud Function that sends a notification when your model predicts that a user's account balance will drop below the \$25 threshold.
- B. 1. Create a Pub/Sub topic for each user* 2. Deploy an application on the App Engine standard environment that sends a notification when your model predicts that a user's account balance will drop below the \$25 threshold.
- C. 1. Build a notification system on Firebase* 2. Register each user with a user ID on the Firebase Cloud Messaging server, which sends a notification when the average of all account balance predictions drops below the \$25 threshold.
- D. 1. Build a notification system on Firebase* 2. Register each user with a user ID on the Firebase Cloud Messaging server, which sends a notification when your model predicts that a user's account balance will drop below the \$25 threshold.

Answer: B

NEW QUESTION 12

You are building a linear regression model on BigQuery ML to predict a customer's likelihood of purchasing your company's products. Your model uses a city name variable as a key predictive component. In order to train and serve the model, your data must be organized in columns. You want to prepare your data using the least amount of coding while maintaining the predictable variables. What should you do?

- A. Create a new view with BigQuery that does not include a column with city information
- B. Use Dataprep to transform the state column using a one-hot encoding method, and make each city a column with binary values.
- C. Use Cloud Data Fusion to assign each city to a region labeled as 1, 2, 3, 4, or 5r and then use that number to represent the city in the model.
- D. Use TensorFlow to create a categorical variable with a vocabulary list Create the vocabulary file, and upload it as part of your model to BigQuery ML.

Answer: C

NEW QUESTION 15

You are an ML engineer at a bank that has a mobile application. Management has asked you to build an ML-based biometric authentication for the app that verifies a customer's identity based on their fingerprint. Fingerprints are considered highly sensitive personal information and cannot be downloaded and stored into the bank databases. Which learning strategy should you recommend to train and deploy this ML model?

- A. Differential privacy
- B. Federated learning
- C. MD5 to encrypt data
- D. Data Loss Prevention API

Answer: B

NEW QUESTION 18

You are an ML engineer at a global car manufacturer. You need to build an ML model to predict car sales in different cities around the world. Which features or feature crosses should you use to train city-specific relationships between car type and number of sales?

- A. Three individual features binned latitude, binned longitude, and one-hot encoded car type
- B. One feature obtained as an element-wise product between latitude, longitude, and car type
- C. One feature obtained as an element-wise product between binned latitude, binned longitude, and one-hot encoded car type
- D. Two feature crosses as a element-wise product the first between binned latitude and one-hot encoded car type, and the second between binned longitude and one-hot encoded car type

Answer: A

NEW QUESTION 20

You are developing ML models with AI Platform for image segmentation on CT scans. You frequently update your model architectures based on the newest available research papers, and have to rerun training on the same dataset to benchmark their performance. You want to minimize computation costs and manual intervention while having version control for your code. What should you do?

- A. Use Cloud Functions to identify changes to your code in Cloud Storage and trigger a retraining job
- B. Use the gcloud command-line tool to submit training jobs on AI Platform when you update your code
- C. Use Cloud Build linked with Cloud Source Repositories to trigger retraining when new code is pushed to the repository
- D. Create an automated workflow in Cloud Composer that runs daily and looks for changes in code in Cloud Storage using a sensor.

Answer: A

NEW QUESTION 21

You are designing an ML recommendation model for shoppers on your company's ecommerce website. You will use Recommendations AI to build, test, and deploy your system. How should you develop recommendations that increase revenue while following best practices?

- A. Use the "Other Products You May Like" recommendation type to increase the click-through rate
- B. Use the "Frequently Bought Together" recommendation type to increase the shopping cart size for each order.
- C. Import your user events and then your product catalog to make sure you have the highest quality event stream
- D. Because it will take time to collect and record product data, use placeholder values for the product catalog to test the viability of the model.

Answer: C

NEW QUESTION 22

You are developing models to classify customer support emails. You created models with TensorFlow Estimators using small datasets on your on-premises system, but you now need to train the models using large datasets to ensure high performance. You will port your models to Google Cloud and want to minimize code refactoring and infrastructure overhead for easier migration from on-prem to cloud. What should you do?

- A. Use AI Platform for distributed training
- B. Create a cluster on Dataproc for training
- C. Create a Managed Instance Group with autoscaling
- D. Use Kubeflow Pipelines to train on a Google Kubernetes Engine cluster.

Answer: D

NEW QUESTION 24

You work for a social media company. You need to detect whether posted images contain cars. Each training example is a member of exactly one class. You have trained an object detection neural network and deployed the model version to AI Platform Prediction for evaluation. Before deployment, you created an evaluation job and attached it to the AI Platform Prediction model version. You notice that the precision is lower than your business requirements allow. How should you adjust the model's final layer softmax threshold to increase precision?

- A. Increase the recall
- B. Decrease the recall.
- C. Increase the number of false positives
- D. Decrease the number of false negatives

Answer: D

NEW QUESTION 29

You were asked to investigate failures of a production line component based on sensor readings. After receiving the dataset, you discover that less than 1% of the readings are positive examples representing failure incidents. You have tried to train several classification models, but none of them converge. How should you resolve the class imbalance problem?

- A. Use the class distribution to generate 10% positive examples
- B. Use a convolutional neural network with max pooling and softmax activation
- C. Downsample the data with upweighting to create a sample with 10% positive examples
- D. Remove negative examples until the numbers of positive and negative examples are equal

Answer: D

NEW QUESTION 34

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

```
gcloud ai-platform jobs submit training $JOB_NAME \
  --package-path $TRAINER_PACKAGE_PATH \
  --module-name $MAIN_TRAINER_MODULE \
  --job-dir $JOB_DIR \
  --region $REGION \
  --scale-tier basic \
  -- \
  --epochs 20 \
  --batch_size=32 \
  --learning_rate=0.001 \
```

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

- A. Modify the 'epochs' parameter
- B. Modify the 'scale-tier' parameter
- C. Modify the batch size' parameter
- D. Modify the 'learning rate' parameter

Answer: A

NEW QUESTION 35

You are building a model to predict daily temperatures. You split the data randomly and then transformed the training and test datasets. Temperature data for model training is uploaded hourly. During testing, your model performed with 97% accuracy; however, after deploying to production, the model's accuracy dropped to 66%. How can you make your production model more accurate?

- A. Normalize the data for the training, and test datasets as two separate steps.
- B. Split the training and test data based on time rather than a random split to avoid leakage
- C. Add more data to your test set to ensure that you have a fair distribution and sample for testing
- D. Apply data transformations before splitting, and cross-validate to make sure that the transformations are applied to both the training and test sets.

Answer: C

NEW QUESTION 38

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

- A. Use the AI Platform Training built-in algorithms to create a custom model
- B. Use AutoML Natural Language to extract custom entities for classification
- C. Use the Cloud Natural Language API to extract custom entities for classification
- D. Build a custom model to identify the product keywords from the transcribed calls, and then run the keywords through a classification algorithm

Answer: A

NEW QUESTION 43

Your team is working on an NLP research project to predict political affiliation of authors based on articles they have written. You have a large training dataset that is structured like this:


```
AuthorA:Political Party A
  TextA1: [SentenceA11, SentenceA12, SentenceA13, ...]
  TextA2: [SentenceA21, SentenceA22, SentenceA23, ...]
  ...

AuthorB:Political Party B
  TextB1: [SentenceB11, SentenceB12, SentenceB13, ...]
  TextB2: [SentenceB21, SentenceB22, SentenceB23, ...]
  ...

AuthorC:Political Party B
  TextC1: [SentenceC11, SentenceC12, SentenceC13, ...]
  TextC2: [SentenceC21, SentenceC22, SentenceC23, ...]
  ...

AuthorD:Political Party A
  TextD1: [SentenceD11, SentenceD12, SentenceD13, ...]
  TextD2: [SentenceD21, SentenceD22, SentenceD23, ...]
  ...

...
```

A)

Distribute texts randomly across the train-test-eval subsets:

```
Train set: [TextA1, TextB2, ...]
Test set: [TextA2, TextC1, TextD2, ...]
Eval set: [TextB1, TextC2, TextD1, ...]
```

B)

Distribute authors randomly across the train-test-eval subsets: (*)

```
Train set: [TextA1, TextA2, TextD1, TextD2, ...]
Test set: [TextB1, TextB2, ...]
Eval set: [TextC1, TextC2, ...]
```

C)

Distribute sentences randomly across the train-test-eval subsets:

```
Train set: [SentenceA11, SentenceA21, Sentence B11, SentenceB21, SentenceC11, SentenceD21, ...]
Test set: [SentenceA12, SentenceA22, Sentence B12, SentenceC22, SentenceC12, SentenceD22, ...]
Eval set: [SentenceA13, SentenceA23, Sentence B13, SentenceC23, SentenceC13, SentenceD31, ...]
```

D)

Distribute paragraphs of texts (i.e., chunks of consecutive sentences) across the train-test-eval subsets:

```
Train set: [SentenceA11, SentenceA12, Sentence D11, SentenceD12, ...]
Test set: [SentenceA13, SentenceB13, Sentence B21, SentenceD23, SentenceC12, SentenceD13, ...]
Eval set: [SentenceA11, SentenceA22, Sentence B13, SentenceD22, SentenceC23, SentenceD11, ...]
```

- A. Option A
- B. Option B
- C. Option C
- D. Option D

Answer: D

NEW QUESTION 44

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. Configure AutoML Tables to perform the classification task
- B. Run a BigQuery ML task to perform logistic regression for the classification
- C. Use AI Platform Notebooks to run the classification model with pandas library
- D. Use AI Platform to run the classification model job configured for hyperparameter tuning

Answer: C

NEW QUESTION 45

Your data science team needs to rapidly experiment with various features, model architectures, and hyperparameters. They need to track the accuracy metrics for

various experiments and use an API to query the metrics over time. What should they use to track and report their experiments while minimizing manual effort?

- A. Use Kubeflow Pipelines to execute the experiments Export the metrics file, and query the results using the Kubeflow Pipelines API.
- B. Use AI Platform Training to execute the experiments Write the accuracy metrics to BigQuery, and query the results using the BigQueryAPI.
- C. Use AI Platform Training to execute the experiments Write the accuracy metrics to Cloud Monitoring, and query the results using the Monitoring API.
- D. Use AI Platform Notebooks to execute the experiment
- E. Collect the results in a shared Google Sheetsfile, and query the results using the Google Sheets API

Answer: A

NEW QUESTION 50

You work for a public transportation company and need to build a model to estimate delay times for multiple transportation routes. Predictions are served directly to users in an app in real time. Because different seasons and population increases impact the data relevance, you will retrain the model every month. You want to follow Google-recommended best practices. How should you configure the end-to-end architecture of the predictive model?

- A. Configure Kubeflow Pipelines to schedule your multi-step workflow from training to deploying your model.
- B. Use a model trained and deployed on BigQuery ML and trigger retraining with the scheduled query feature in BigQuery
- C. Write a Cloud Functions script that launches a training and deploying job on Ai Platform that is triggered by Cloud Scheduler
- D. Use Cloud Composer to programmatically schedule a Dataflow job that executes the workflow from training to deploying your model

Answer: B

NEW QUESTION 51

You are an ML engineer at a global shoe store. You manage the ML models for the company's website. You are asked to build a model that will recommend new products to the user based on their purchase behavior and similarity with other users. What should you do?

- A. Build a classification model
- B. Build a knowledge-based filtering model
- C. Build a collaborative-based filtering model
- D. Build a regression model using the features as predictors

Answer: C

NEW QUESTION 52

You work on a growing team of more than 50 data scientists who all use AI Platform. You are designing a strategy to organize your jobs, models, and versions in a clean and scalable way. Which strategy should you choose?

- A. Set up restrictive IAM permissions on the AI Platform notebooks so that only a single user or group can access a given instance.
- B. Separate each data scientist's work into a different project to ensure that the jobs, models, and versions created by each data scientist are accessible only to that user.
- C. Use labels to organize resources into descriptive categorie
- D. Apply a label to each created resource so that users can filter the results by label when viewing or monitoring the resources
- E. Set up a BigQuery sink for Cloud Logging logs that is appropriately filtered to capture information about AI Platform resource usage In BigQuery create a SQL view that maps users to the resources they are using.

Answer: B

NEW QUESTION 54

You have written unit tests for a Kubeflow Pipeline that require custom libraries. You want to automate the execution of unit tests with each new push to your development branch in Cloud Source Repositories. What should you do?

- A. Write a script that sequentially performs the push to your development branch and executes the unit tests on Cloud Run
- B. Using Cloud Build, set an automated trigger to execute the unit tests when changes are pushed to your development branch.
- C. Set up a Cloud Logging sink to a Pub/Sub topic that captures interactions with Cloud Source Repositories Configure a Pub/Sub trigger for Cloud Run, and execute the unit tests on Cloud Run.
- D. Set up a Cloud Logging sink to a Pub/Sub topic that captures interactions with Cloud Source Repositorie
- E. Execute the unit tests using a Cloud Function that is triggered when messages are sent to the Pub/Sub topic

Answer: B

NEW QUESTION 59

.....

Thank You for Trying Our Product

We offer two products:

1st - We have Practice Tests Software with Actual Exam Questions

2nd - Questions and Answers in PDF Format

Professional-Machine-Learning-Engineer Practice Exam Features:

- * Professional-Machine-Learning-Engineer Questions and Answers Updated Frequently
- * Professional-Machine-Learning-Engineer Practice Questions Verified by Expert Senior Certified Staff
- * Professional-Machine-Learning-Engineer Most Realistic Questions that Guarantee you a Pass on Your FirstTry
- * Professional-Machine-Learning-Engineer Practice Test Questions in Multiple Choice Formats and Updatesfor 1 Year

100% Actual & Verified — Instant Download, Please Click
[Order The Professional-Machine-Learning-Engineer Practice Test Here](#)