

# Exam Questions Professional-Machine-Learning-Engineer

Google Professional Machine Learning Engineer

<https://www.2passeasy.com/dumps/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 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 4**

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 5**

You have deployed multiple versions of an image classification model on AI Platform. You want to monitor the performance of the model versions overtime. How should you perform this comparison?

- A. Compare the loss performance for each model on a held-out dataset.
- B. Compare the loss performance for each model on the validation data
- C. Compare the receiver operating characteristic (ROC) curve for each model using the What-If Tool
- D. Compare the mean average precision across the models using the Continuous Evaluation feature

**Answer: B**

**NEW QUESTION 6**

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 7

You are building a real-time prediction engine that streams files which may contain Personally Identifiable Information (PII) to Google Cloud. You want to use the Cloud Data Loss Prevention (DLP) API to scan the files. How should you ensure that the PII is not accessible by unauthorized individuals?

- A. Stream all files to Google CloudT and then write the data to BigQuery Periodically conduct a bulk scan of the table using the DLP API.
- B. Stream all files to Google Cloud, and write batches of the data to BigQuery While the data is being written to BigQuery conduct a bulk scan of the data using the DLP API.
- C. Create two buckets of data Sensitive and Non-sensitive Write all data to the Non-sensitive bucket Periodically conduct a bulk scan of that bucket using the DLP API, and move the sensitive data to the Sensitive bucket
- D. Create three buckets of data: Quarantine, Sensitive, and Non-sensitive Write all data to the Quarantine bucket.
- E. Periodically conduct a bulk scan of that bucket using the DLP API, and move the data to either the Sensitive or Non-Sensitive bucket

**Answer: A**

#### NEW QUESTION 8

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 9

You are responsible for building a unified analytics environment across a variety of on-premises data marts. Your company is experiencing data quality and security challenges when integrating data across the servers, caused by the use of a wide range of disconnected tools and temporary solutions. You need a fully managed, cloud-native data integration service that will lower the total cost of work and reduce repetitive work. Some members on your team prefer a codeless interface for building Extract, Transform, Load (ETL) process. Which service should you use?

- A. Dataflow
- B. Dataprep
- C. Apache Flink
- D. Cloud Data Fusion

**Answer: D**

#### NEW QUESTION 10

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 10

Your team trained and tested a DNN regression model with good results. Six months after deployment, the model is performing poorly due to a change in the distribution of the input data. How should you address the input differences in production?

- A. Create alerts to monitor for skew, and retrain the model.
- B. Perform feature selection on the model, and retrain the model with fewer features
- C. Retrain the model, and select an L2 regularization parameter with a hyperparameter tuning service
- D. Perform feature selection on the model, and retrain the model on a monthly basis with fewer features

**Answer: C**

#### NEW QUESTION 14

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 18

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 20

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-team, and custom libraries. What should you do?

- A. Use the AI Platform custom containers feature to receive training jobs using any framework
- B. Configure Kubeflow to run on Google Kubernetes Engine and receive 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.

**Answer: D**

#### NEW QUESTION 25

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

- A. • Validate the accuracy of the model that you trained on preprocessed data• Create a new model that uses the raw data and is available in real time• Deploy the new model onto AI Platform for online prediction
- B. • Send incoming prediction requests to a Pub/Sub topic• Transform the incoming data using a Dataflow job• Submit a prediction request to AI Platform using the transformed data• Write the predictions to an outbound Pub/Sub queue
- C. • Stream incoming prediction request data into Cloud Spanner• Create a view to abstract your preprocessing logic• Query the view every second for new records• Submit a prediction request to AI Platform using the transformed data• Write the predictions to an outbound Pub/Sub queue.
- D. • Send incoming prediction requests to a Pub/Sub topic• Set up a Cloud Function that is triggered when messages are published to the Pub/Sub topic. Implement your preprocessing logic in the Cloud Function• Submit a prediction request to AI Platform using the transformed data• Write the predictions to an outbound Pub/Sub queue

**Answer: D**

#### NEW QUESTION 28

Your team needs to build a model that predicts whether images contain a driver's license, passport, or credit card. The data engineering team already built the pipeline and generated a dataset composed of 10,000 images with driver's licenses, 1,000 images with passports, and 1,000 images with credit cards. You now have to train a model with the following label map: ['driverslicense', 'passport', 'credit\_card']. Which loss function should you use?

- A. Categorical hinge
- B. Binary cross-entropy
- C. Categorical cross-entropy
- D. Sparse categorical cross-entropy

**Answer: B**

#### NEW QUESTION 30

.....

## THANKS FOR TRYING THE DEMO OF OUR PRODUCT

Visit Our Site to Purchase the Full Set of Actual Professional-Machine-Learning-Engineer Exam Questions With Answers.

We Also Provide Practice Exam Software That Simulates Real Exam Environment And Has Many Self-Assessment Features. Order the Professional-Machine-Learning-Engineer Product From:

<https://www.2passeasy.com/dumps/Professional-Machine-Learning-Engineer/>

## Money Back Guarantee

### **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