DBCLOUD
EXAMS

# QUESTIONS & ANSWERS

verified



**Technology Provider:** Google
**Exam:** Professional Machine Learning Engineer
**Number of questions in the database:** 349
**Exam Version:** NOV.2026

**Topic 1 - Single Topic**

Question #1*Topic 1*

You are building an ML model to detect anomalies in real-time sensor data. You will use Pub/Sub to handle incoming requests. You want to store the results for analytics and visualization. How should you configure the pipeline?

- A. 1 = Dataflow, 2 = AI Platform, 3 = BigQuery
- B. 1 = DataProc, 2 = AutoML, 3 = Cloud Bigtable
- C. 1 = BigQuery, 2 = AutoML, 3 = Cloud Functions
- D. 1 = BigQuery, 2 = AI Platform, 3 = Cloud Storage

**Correct Answer:** *C*
Reference:
https://cloud.google.com/solutions/building-anomaly-detection-dataflow-bigqueryml-dlp

***Respuesta por mayor votación de la comunidad***

A (100%)

Question #2*Topic 1*

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

https://dbcloud.academy/

a distance-based metric. 2. Dispatch an appropriately sized shuttle and provide the map with the required stops based on the simulated outcome.

**Correct Answer:** *A*

***Respuesta por mayor votación de la comunidad***

C (94%)

6%

Question #3*Topic 1*

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.

**Correct Answer:** *B*
Reference:
https://towardsdatascience.com/convolution-neural-networks-a-beginners-guide-implementing-a-mnist-hand-written-digit-8aa60330d022

***Respuesta por mayor votación de la comunidad***

C (93%)

7%

Question #4*Topic 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.

https://dbcloud.academy/

- 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.

**Correct Answer:** *B*

***Respuesta por mayor votación de la comunidad***

D (73%)

B (19%)

8%

Question #5*Topic 1*

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 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 TF Job.
- 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.

**Correct Answer:** *D*

***Respuesta por mayor votación de la comunidad***

A (95%)

5%

Question #6*Topic 1*

You work for an online retail company that is creating a visual search engine. You have set up an end-to-end ML pipeline on Google Cloud to classify whether an image contains your company's product. Expecting the release of new products in the near future, you configured a retraining functionality in the pipeline so that new data can be fed into your ML models. You also want to use

https://dbcloud.academy/

AI Platform's continuous evaluation service to ensure that the models have high accuracy on your test dataset. What should you do?

- A. Keep the original test dataset unchanged even if newer products are incorporated into retraining.
- B. Extend your test dataset with images of the newer products when they are introduced to retraining.
- C. Replace your test dataset with images of the newer products when they are introduced to retraining.
- D. Update your test dataset with images of the newer products when your evaluation metrics drop below a pre-decided threshold.

**Correct Answer:** *C*

***Respuesta por mayor votación de la comunidad***

B (82%)

D (18%)

Question #7*Topic 1*

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.

**Correct Answer:** *B*
BigQuery ML supports supervised learningλ€ with the logistic regression model type.
Reference:
https://cloud.google.com/bigquery-ml/docs/logistic-regression-prediction

***Respuesta por mayor votación de la comunidad***

A (100%)

Question #8*Topic 1*

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

https://dbcloud.academy/