

PROFESSIONAL-MACHINE- LEARNING-ENGINEER^{Q&As}

Professional Machine Learning Engineer

**Pass Google PROFESSIONAL-MACHINE-LEARNING-
ENGINEER Exam with 100% Guarantee**

Free Download Real Questions & Answers **PDF** and **VCE** file from:

<https://www.leads4pass.com/professional-machine-learning-engineer.html>

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by Google
Official Exam Center

- ⚙️ **Instant Download** After Purchase
- ⚙️ **100% Money Back** Guarantee
- ⚙️ **365 Days** Free Update
- ⚙️ **800,000+** Satisfied Customers



QUESTION 1

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.

Correct Answer: D

QUESTION 2

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 20 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 sets from 30 days after your validation set.

Correct Answer: D

QUESTION 3

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 categories. Apply a label to each created resource so that users can filter the results by label when viewing or monitoring the resources.

D. 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

Correct Answer: C

https://cloud.google.com/ai-platform/prediction/docs/resource-labels#overview_of_labels

QUESTION 4

You are an ML engineer at a regulated insurance company. You are asked to develop an insurance approval model that accepts or rejects insurance applications from potential customers. What factors should you consider before building the model?

A. Redaction, reproducibility, and explainability

B. Traceability, reproducibility, and explainability

C. Federated learning, reproducibility, and explainability D. Differential privacy, federated learning, and explainability

Correct Answer: B

QUESTION 5

You work for a magazine distributor and need to build a model that predicts which customers will renew their subscriptions for the upcoming year. Using your company's historical data as your training set, you created a TensorFlow model and deployed it to AI Platform. You need to determine which customer attribute has the most predictive power for each prediction served by the model. What should you do?

A. Use AI Platform notebooks to perform a Lasso regression analysis on your model, which will eliminate features that do not provide a strong signal.

B. Stream prediction results to BigQuery. Use BigQuery's CORR(X1, X2) function to calculate the Pearson correlation coefficient between each feature and the target variable.

C. Use the AI Explanations feature on AI Platform. Submit each prediction request with the `explain` keyword to retrieve feature attributions using the sampled Shapley method.

D. Use the What-If tool in Google Cloud to determine how your model will perform when individual features are excluded. Rank the feature importance in order of those that caused the most significant performance drop when removed from the model.

Correct Answer: C

<https://cloud.google.com/ai-platform/prediction/docs/ai-explanations/overview>

QUESTION 6

You work for a large retailer and have been asked to segment your customers by their purchasing habits. The purchase history of all customers has been uploaded to BigQuery. You suspect that there may be several distinct customer segments, however you are unsure of how many, and you don't yet understand the commonalities in their behavior. You want to find the most efficient solution. What should you do?

- A. Create a k-means clustering model using BigQuery ML. Allow BigQuery to automatically optimize the number of clusters.
- B. Create a new dataset in Dataprep that references your BigQuery table. Use Dataprep to identify similarities within each column.
- C. Use the Data Labeling Service to label each customer record in BigQuery. Train a model on your labeled data using AutoML Tables. Review the evaluation metrics to understand whether there is an underlying pattern in the data.
- D. Get a list of the customer segments from your company's Marketing team. Use the Data Labeling Service to label each customer record in BigQuery according to the list. Analyze the distribution of labels in your dataset using Data Studio.

Correct Answer: A

<https://cloud.google.com/bigquery/docs/kmeans-tutorial>

QUESTION 7

You recently built the first version of an image segmentation model for a self-driving car. After deploying the model, you observe a decrease in the area under the curve (AUC) metric. When analyzing the video recordings, you also discover that the model fails in highly congested traffic but works as expected when there is less traffic. What is the most likely reason for this result?

- A. The model is overfitting in areas with less traffic and underfitting in areas with more traffic.
- B. AUC is not the correct metric to evaluate this classification model.
- C. Too much data representing congested areas was used for model training.
- D. Gradients become small and vanish while backpropagating from the output to input nodes.

Correct Answer: A

The most likely reason for this result is the model is overfitting in areas with less traffic and underfitting in areas with more traffic. Probably because the model was trained on a dataset that did not have enough examples of congested traffic. As a result, the model is not able to generalise well. When the model is validated on congested traffic, it makes mistakes because it has not seen this type of data before.

QUESTION 8

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

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

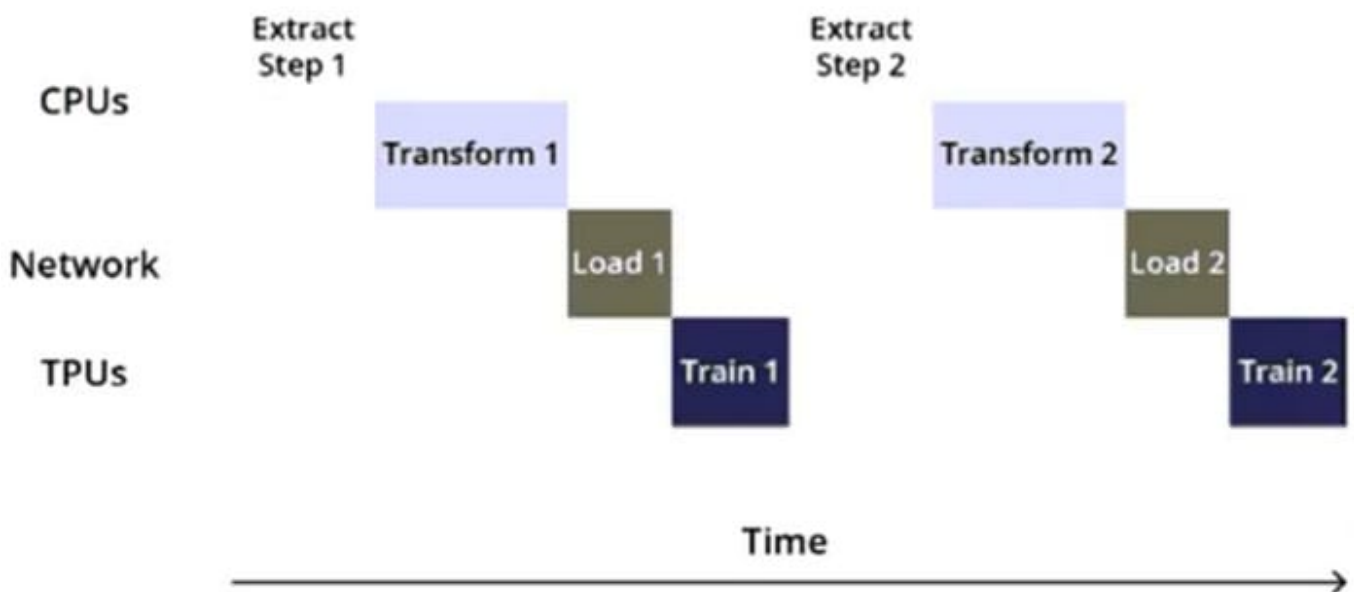
- B. Address data leakage by applying nested cross-validation during model training.
- C. Address data leakage by removing features highly correlated with the target value.
- D. Address the model overfitting by tuning the hyperparameters to reduce the AUC ROC value.

Correct Answer: B

<https://towardsdatascience.com/time-series-nested-cross-validation-76adba623eb9>

QUESTION 9

You are training an object detection model using a Cloud TPU v2. Training time is taking longer than expected. Based on this simplified trace obtained with a Cloud TPU profile, what action should you take to decrease training time in a cost-efficient way?



- A. Move from Cloud TPU v2 to Cloud TPU v3 and increase batch size.
- B. Move from Cloud TPU v2 to 8 NVIDIA V100 GPUs and increase batch size.
- C. Rewrite your input function to resize and reshape the input images.
- D. Rewrite your input function using parallel reads, parallel processing, and prefetch.

Correct Answer: D

Based on the profile, it appears that the Compute time is relatively low compared to the HostToDevice and DeviceToHost time. This suggests that the data transfer between the host (CPU) and the TPU device is a bottleneck. Therefore, the best action to decrease training time in a cost-efficient way would be to reduce the amount of data transferred between the host and the device.

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.

Correct Answer: D

Firebase is designed for exactly this sort of scenario. Also, it would not be possible to create millions of pubsub topics due to GCP quotas <https://cloud.google.com/pubsub/quotas#quotas> <https://firebase.google.com/docs/cloud-messaging>

QUESTION 11

You have deployed multiple versions of an image classification model on AI Platform. You want to monitor the performance of the model versions over time. 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.

Correct Answer: D

Compare the mean average precision across the models using the Continuous Evaluation feature

<https://cloud.google.com/vertex-ai/docs/evaluation/introduction>

Vertex AI provides model evaluation metrics, such as precision and recall, to help you determine the performance of your models...

Vertex AI supports evaluation of the following model types:

AuPRC: The area under the precision-recall (PR) curve, also referred to as average precision. This value ranges from zero to one, where a higher value indicates a higher-quality model.

QUESTION 12

During batch training of a neural network, you notice that there is an oscillation in the loss. How should you adjust your model to ensure that it converges?

- A. Decrease the size of the training batch.
- B. Decrease the learning rate hyperparameter.
- C. Increase the learning rate hyperparameter.
- D. Increase the size of the training batch.

Correct Answer: B

<https://ai.stackexchange.com/questions/14079/what-could-an-oscillating-training-loss-curve-represent#:~:text=Try%20lowering%20the%20learning%20rate,step%20and%20overshoot%20it%20again.>

QUESTION 13

You are an ML engineer at a travel company. You have been researching customers' travel behavior for many years, and you have deployed models that predict customers' vacation patterns. You have observed that customers' vacation destinations vary based on seasonality and holidays; however, these seasonal variations are similar across years. You want to quickly and easily store and compare the model versions and performance statistics across years. What should you do?

- A. Store the performance statistics in Cloud SQL. Query that database to compare the performance statistics across the model versions.
- B. Create versions of your models for each season per year in Vertex AI. Compare the performance statistics across the models in the Evaluate tab of the Vertex AI UI.
- C. Store the performance statistics of each pipeline run in Kubeflow under an experiment for each season per year. Compare the results across the experiments in the Kubeflow UI.
- D. Store the performance statistics of each version of your models using seasons and years as events in Vertex ML Metadata. Compare the results across the slices.

Correct Answer: D

<https://cloud.google.com/vertex-ai/docs/ml-metadata/analyzing#filtering>

QUESTION 14

You work for a company that is developing a new video streaming platform. You have been asked to create a recommendation system that will suggest the next video for a user to watch. After a review by an AI Ethics team, you are approved to start development. Each video asset in your company's catalog has useful metadata (e.g., content type, release date, country), but you do not have any historical user event data. How should you build the recommendation system for the first version of the product?

- A. Launch the product without machine learning. Present videos to users alphabetically, and start collecting user event

data so you can develop a recommender model in the future.

- B. Launch the product without machine learning. Use simple heuristics based on content metadata to recommend similar videos to users, and start collecting user event data so you can develop a recommender model in the future.
- C. Launch the product with machine learning. Use a publicly available dataset such as MovieLens to train a model using the Recommendations AI, and then apply this trained model to your data.
- D. Launch the product with machine learning. Generate embeddings for each video by training an autoencoder on the content metadata using TensorFlow. Cluster content based on the similarity of these embeddings, and then recommend videos from the same cluster.

Correct Answer: B

<https://developers.google.com/machine-learning/guides/rules-of-ml>

QUESTION 15

While monitoring your model training's GPU utilization, you discover that you have a native synchronous implementation. The training data is split into multiple files. You want to reduce the execution time of your input pipeline. What should you do?

- A. Increase the CPU load
- B. Add caching to the pipeline
- C. Increase the network bandwidth
- D. Add parallel interleave to the pipeline

Correct Answer: D

https://www.tensorflow.org/guide/data_performance

[Latest PROFESSIONAL-MACHINE-LEARNING-ENGINEER Dumps](#)

[PROFESSIONAL-MACHINE-LEARNING-ENGINEER Exam Questions](#)

[PROFESSIONAL-MACHINE-LEARNING-ENGINEER Braindumps](#)