

Google Google Professional-Machine-Learning-Engineer PDF

Google Google Professional-Machine-Learning-Engineer PDF Questions Available Here at:

<https://www.certification-questions.com/google-dumps/professional-machine-learning-engineer.html>

Enrolling now you will get access to **283 questions** in a unique set of Google Professional-Machine-Learning-Engineer

Question 1

As the lead ML Engineer for your company, you are responsible for building ML models to digitize scanned customer forms. You have developed a TensorFlow model that converts the scanned images into text and stores them in Cloud Storage. You need to use your ML model on the aggregated data collected at the end of each day with minimal manual intervention. What should you do?

Options:

- A. Use the batch prediction functionality of AI Platform
- B. Create a serving pipeline in Compute Engine for prediction
- C. Use Cloud Functions for prediction each time a new data point is ingested
- D. Deploy the model on AI Platform and create a version of it for online inference.

Answer: A

Explanation:

Batch prediction is the process of using an ML model to make predictions on a large set of data points. Batch prediction is suitable for scenarios where the predictions are not time-sensitive and can be done in batches, such as digitizing scanned customer forms at the end of each day. Batch prediction can also handle large volumes of data and scale up or down the resources as needed. AI Platform provides a batch prediction service that allows users to submit a job with their TensorFlow model and input data stored in Cloud Storage, and receive the output predictions in Cloud Storage as well. This service requires minimal manual intervention and can be automated with Cloud Scheduler or Cloud Functions. Therefore, using the batch prediction functionality of AI Platform is the best option for this use case.

Reference:

Batch prediction overview

Using batch prediction

Question 2

You work for a global footwear retailer and need to predict when an item will be out of stock based on historical inventory data

a. Customer behavior is highly dynamic since footwear demand is influenced by many different factors. You want to serve models that are trained on all available data, but track your performance on specific subsets of data before pushing to production. What is the most streamlined and reliable way to perform this validation?

Options:

- A. Use the TFX ModelValidator tools to specify performance metrics for production readiness
- B. Use k-fold cross-validation as a validation strategy to ensure that your model is ready for production.
- C. Use the last relevant week of data as a validation set to ensure that your model is performing accurately on current data
- D. Use the entire dataset and treat the area under the receiver operating characteristics curve (AUC ROC) as the main metric.

Answer: A

Explanation:

TFX ModelValidator is a tool that allows you to compare new models against a baseline model and evaluate their performance on different metrics and data slices¹. You can use this tool to validate your models before deploying them to production and ensure that they meet your expectations and requirements.

k-fold cross-validation is a technique that splits the data into k subsets and trains the model on k-1 subsets while testing it on the remaining subset. This is repeated k times and the average performance is reported². This technique is useful for estimating the generalization error of a model, but it does not account for the dynamic nature of customer behavior or the potential changes in data distribution over time.

Using the last relevant week of data as a validation set is a simple way to check the model's performance on recent data, but it may not be representative of the entire data or capture the long-term trends and patterns. It also does not allow you to compare the model with a baseline or evaluate it on different data slices.

Using the entire dataset and treating the AUC ROC as the main metric is not a good practice because it does not leave any data for validation or testing. It also assumes that the AUC ROC is the only metric that matters, which may not be true for your business problem. You may want to consider other metrics such as precision, recall, or revenue.

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?

Options:

- 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.

Answer: C

Explanation:

Labels are key-value pairs that can be attached to any AI Platform resource, such as jobs, models, versions, or endpoints¹. Labels can help you organize your resources into descriptive categories, such as project, team, environment, or purpose. You can use labels to filter the results when you list or monitor your resources, or to group them for billing or quota purposes². Using labels is a simple and scalable way to manage your AI Platform resources without creating unnecessary complexity or overhead. Therefore, using labels to organize resources is the best strategy for this use case.

Reference:

Using labels

Filtering and grouping by labels

Question 4

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?

Options:

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

D. Decrease the learning rate hyperparameter

Answer: D

Explanation:

Oscillation in the loss during batch training of a neural network means that the model is overshooting the optimal point of the loss function and bouncing back and forth. This can prevent the model from converging to the minimum loss value. One of the main reasons for this phenomenon is that the learning rate hyperparameter, which controls the size of the steps that the model takes along the gradient, is too high. Therefore, decreasing the learning rate hyperparameter can help the model take smaller and more precise steps and avoid oscillation. This is a common technique to improve the stability and performance of neural network training¹².

Reference:

Interpreting Loss Curves

Is learning rate the only reason for training loss oscillation after few epochs?

Question 5

You are building a linear model with over 100 input features, all with values between -1 and 1. You suspect that many features are non-informative. You want to remove the non-informative features from your model while keeping the informative ones in their original form. Which technique should you use?

Options:

- A. Use Principal Component Analysis to eliminate the least informative features.
- B. Use L1 regularization to reduce the coefficients of uninformative features to 0.
- C. After building your model, use Shapley values to determine which features are the most informative.
- D. Use an iterative dropout technique to identify which features do not degrade the model when removed.

Answer: B

Explanation:

L1 regularization, also known as Lasso regularization, adds the sum of the absolute values of the model's coefficients to the loss function¹. It encourages sparsity in the model by shrinking some coefficients to precisely zero². This way, L1 regularization can perform feature selection and remove the non-informative features from the model while keeping the informative ones in their original form. Therefore, using L1 regularization is the best technique for this use case.

Reference:

Regularization in Machine Learning - GeeksforGeeks

Regularization in Machine Learning (with Code Examples) - Dataquest

L1 And L2 Regularization Explained & Practical How To Examples

L1 and L2 as Regularization for a Linear Model

Question 6

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?

Options:

- 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: C

Explanation:

Transfer learning is a technique that leverages the knowledge and weights of a pre-trained model and adapts them to a new task or domain¹. Transfer learning can save time and resources by avoiding training a model from scratch, and can also improve the performance and generalization of the model by using a larger and more diverse dataset². AI Platform provides several established text classification models that can be used for transfer learning, such as BERT, ALBERT, or XLNet³. These models are based on state-of-the-art natural language processing techniques and can handle various text classification tasks, such as sentiment analysis, topic classification, or spam detection⁴. By using one of these models on AI Platform, you can customize the model's code, serving, and deployment, and use Kubeflow pipelines for the ML platform. Therefore, using an established text classification model on AI Platform to perform transfer learning is the best option for this use case.

Reference:

Transfer Learning - Machine Learning's Next Frontier

A Comprehensive Hands-on Guide to Transfer Learning with Real-World Applications in Deep Learning

Text classification models

Text Classification with Pre-trained Models in TensorFlow

Question 7

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:

You followed the standard 80%-10%-10% data distribution across the training, testing, and evaluation subsets. How should you distribute the training examples across the train-test-eval subsets while maintaining the 80-10-10 proportion?

AuthorA: Political Party A

A) TextA1: [SentenceA11, SentenceA12, SentenceA13, ...]

TextA2: [SentenceA21, SentenceA22, SentenceA23, ...]

B) ...

AuthorB: Political Party B

C) TextB1: [SentenceB11, SentenceB12, SentenceB13, ...]

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

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, ...]

Eval set: [TextC1, TextC2, ...]

Options:

A. Option A

AuthorD: Political Party A

B. Option B

C. Option C: [SentenceD11, SentenceD12, SentenceD13, ...]

D. Option D: [SentenceD21, SentenceD22, SentenceD23, ...]

Answer: C

Explanation:

The best way to distribute the training examples across the train-test-eval subsets while maintaining the 80-10-10 proportion is to use option C. This option ensures that each subset contains a balanced and representative sample of the different classes (Democrat and Republican) and the different authors. This way, the model can learn from a diverse and comprehensive set of articles and avoid overfitting or underfitting. Option C also avoids the problem of data leakage, which occurs when the same author appears in more than one subset, potentially biasing the model and inflating its

performance. Therefore, option C is the most suitable technique for this use case.

Question 8

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?

Options:

- 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 experiments. Collect the results in a shared Google Sheets file, and query the results using the Google Sheets API

Answer: C

Explanation:

AI Platform Training is a service that allows you to run your machine learning experiments on Google Cloud using various features, model architectures, and hyperparameters. You can use AI Platform Training to scale up your experiments, leverage distributed training, and access specialized hardware such as GPUs and TPUs¹. Cloud Monitoring is a service that collects and analyzes metrics, logs, and traces from Google Cloud, AWS, and other sources. You can use Cloud Monitoring to create dashboards, alerts, and reports based on your data². The Monitoring API is an interface that allows you to programmatically access and manipulate your monitoring data³.

By using AI Platform Training and Cloud Monitoring, you can track and report your experiments while minimizing manual effort. You can write the accuracy metrics from your experiments to Cloud Monitoring using the AI Platform Training Python package⁴. You can then query the results using the Monitoring API and compare the performance of different experiments. You can also visualize the metrics in the Cloud Console or create custom dashboards and alerts⁵. Therefore, using AI Platform Training and Cloud Monitoring is the best option for this use case.

Reference:

[AI Platform Training documentation](#)

[Cloud Monitoring documentation](#)

[Monitoring API overview](#)

[Using Cloud Monitoring with AI Platform Training](#)

[Viewing evaluation metrics](#)

Question 9

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?

Options:

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

Answer: B

Explanation:

Federated learning is a machine learning technique that enables organizations to train AI models on decentralized data without centralizing or sharing it¹. It allows data privacy, continual learning, and better performance on end-user devices². Federated learning works by sending the model parameters to the devices, where they are updated locally on the device's data, and then aggregating the updated parameters on a central server to form a global model³. This way, the data never leaves the device and the model can learn from a large and diverse dataset.

Federated learning is suitable for the use case of building an ML-based biometric authentication for the bank's mobile 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. By using federated learning, the bank can train and deploy an ML model that can recognize fingerprints without compromising the data privacy of the customers. The model can also adapt to the variations and changes in the fingerprints over time and improve its accuracy and reliability. Therefore, federated learning is the best learning strategy for this use case.

Question 10

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?

Options:

- 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: B

Explanation:

One-hot encoding is a technique that converts categorical variables into numerical variables by creating dummy variables for each possible category. Each dummy variable has a value of 1 if the original variable belongs to that category, and 0 otherwise¹. One-hot encoding can help linear regression models to capture the effect of different categories on the target variable without imposing any ordinal relationship among them². Dataprep is a service that allows you to explore, clean, and transform your data for analysis and machine learning. You can use Dataprep to apply one-hot encoding to your city name variable and make each city a column with binary values³. This way, you can prepare your data using the least amount of coding while maintaining the predictive variables. Therefore, using Dataprep to transform the state column using a one-hot encoding method is the best option for this use case.

Reference:

One Hot Encoding: A Beginner's Guide

One-Hot Encoding in Linear Regression Models

Dataprep documentation

Would you like to see more? Don't miss our Google Professional-Machine-Learning-Engineer PDF file at:

<https://www.certification-questions.com/google-pdf/professional-machine-learning-engineer-pdf.html>