

Google Professional Machine Learning Engineer Study Training Dumps Grasped the Core Knowledge of Professional-Machine-Learning-Engineer Exam



DOWNLOAD the newest PassSureExam Professional-Machine-Learning-Engineer PDF dumps from Cloud Storage for free: https://drive.google.com/open?id=1U4tJzXXcMTJ8IZglgsf_IH5EtP5jbJkV

Since One of the significant factors to judge whether one is competent or not is his or her Professional-Machine-Learning-Engineer certificates. So to get Professional-Machine-Learning-Engineer real exam and pass the Professional-Machine-Learning-Engineer exam is important. Generally speaking, certificates function as the fundamental requirement when a company needs to increase manpower in its start-up stage. In this respect, our Professional-Machine-Learning-Engineer practice materials can satisfy your demands if you are now in preparation for a certificate. We will be your best friend to help you achieve success!

Google Professional Machine Learning Engineer certification is a valuable credential for individuals seeking to demonstrate their expertise in machine learning. Professional-Machine-Learning-Engineer exam covers a wide range of topics and requires candidates to have a solid understanding of machine learning algorithms, statistical analysis, and data visualization. Achieving this certification can help individuals differentiate themselves in the job market and open up new career opportunities.

Google Professional-Machine-Learning-Engineer Certification Exam is intended for machine learning engineers, data scientists, and software developers who want to demonstrate their expertise in building and deploying machine learning models on Google Cloud Platform. Professional-Machine-Learning-Engineer exam covers a wide range of topics, including data preparation and analysis, feature engineering, model selection and training, model evaluation and optimization, and deploying and managing machine learning models on Google Cloud Platform.

>> **Professional-Machine-Learning-Engineer Dumps Collection** <<

Frequent Professional-Machine-Learning-Engineer Update | Valid Professional-Machine-Learning-Engineer Exam Forum

If you want to pass the Professional-Machine-Learning-Engineer exam in the least time with the least efforts, then you only need to purchase our Professional-Machine-Learning-Engineer learning guide. You can own the most important three versions of our Professional-Machine-Learning-Engineer practice materials if you buy the Value Pack! Also you can only choose the one you like best. As you know, the best for yourself is the best. Choosing the best product for you really saves a lot of time! Professional-Machine-Learning-Engineer Actual Exam look forward to be your best partner.

Preparation Process

The candidates for the Google Professional Machine Learning Engineer certification can find everything they need to efficiently prepare for the qualifying test on the official website. The most recommended resource offered by the vendor is the Machine Learning Engineer learning path. It contains both lessons and practical labs for a comprehensive understanding of the exam content. Moreover, the students can take advantage of the sample questions designed to help the potential test takers familiarize themselves with the possible exam questions. Finally, the applicants can opt for the Machine Learning Engineer Prep Webinar to join the Google

experts and recently certified professionals for the tips and insights on the Machine Learning models, data processing systems, solution quality, and more.

Google Professional Machine Learning Engineer Sample Questions (Q74-Q79):

NEW QUESTION # 74

You are pre-training a large language model on Google Cloud. This model includes custom TensorFlow operations in the training loop. Model training will use a large batch size, and you expect training to take several weeks. You need to configure a training architecture that minimizes both training time and compute costs. What should you do?

- A. Implement 8 workers of a2-megagpu-16g machines by using `tf.distribute.MultiWorkerMirroredStrategy`.
- B. Implement 16 workers of a2-highgpu-8g machines by using `tf.distribute.MultiWorkerMirroredStrategy`.
- C. Implement 16 workers of c2d-highcpu-32 machines by using `tf.distribute.MirroredStrategy`.
- D. Implement a TPU Pod slice with `-accelerator-type=v4-128` by using `tf.distribute.TPUStrategy`.

Answer: D

NEW QUESTION # 75

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

Answer: A

Explanation:

Dataflow is a fully managed service for executing Apache Beam pipelines that can process streaming or batch data¹.

AI Platform is a unified platform that enables you to build and run machine learning applications across Google Cloud².

BigQuery is a serverless, highly scalable, and cost-effective cloud data warehouse designed for business agility³.

These services are suitable for building an ML model to detect anomalies in real-time sensor data, as they can handle large-scale data ingestion, preprocessing, training, serving, storage, and visualization. The other options are not as suitable because:

DataProc is a service for running Apache Spark and Apache Hadoop clusters, which are not optimized for streaming data processing⁴.

AutoML is a suite of machine learning products that enables developers with limited machine learning expertise to train high-quality models specific to their business needs⁵. However, it does not support custom models or real-time predictions.

Cloud Bigtable is a scalable, fully managed NoSQL database service for large analytical and operational workloads. However, it is not designed for ad hoc queries or interactive analysis.

Cloud Functions is a serverless execution environment for building and connecting cloud services. However, it is not suitable for storing or visualizing data.

Cloud Storage is a service for storing and accessing data on Google Cloud. However, it is not a data warehouse and does not support SQL queries or visualization tools.

NEW QUESTION # 76

You work on a data science team at a bank and are creating an ML model to predict loan default risk. You have collected and cleaned hundreds of millions of records worth of training data in a BigQuery table, and you now want to develop and compare multiple models on this data using TensorFlow and Vertex AI. You want to minimize any bottlenecks during the data ingestion state while considering scalability. What should you do?

- A. Export data to CSV files in Cloud Storage, and use `tf.data.TextLineDataset()` to read them.
- B. Convert the data into TFRecords, and use `tf.data.TFRecordDataset()` to read them.
- C. Use the BigQuery client library to load data into a dataframe, and use `tf.data.Dataset.from_tensor_slices()` to read it.
- **D. Use TensorFlow I/O's BigQuery Reader to directly read the data.**

Answer: D

Explanation:

The best option for developing and comparing multiple models on a large-scale BigQuery table using TensorFlow and Vertex AI is to use TensorFlow I/O's BigQuery Reader to directly read the data. This option has the following advantages:

* It minimizes any bottlenecks during the data ingestion stage, as the BigQuery Reader can stream data from BigQuery to TensorFlow in parallel and in batches, without loading the entire table into memory or disk. The BigQuery Reader can also perform data transformations and filtering using SQL queries, reducing the need for additional preprocessing steps in TensorFlow.

* It leverages the scalability and performance of BigQuery, as the BigQuery Reader can handle hundreds of millions of records worth of training data efficiently and reliably. BigQuery is a serverless, fully managed, and highly scalable data warehouse that can run complex queries over petabytes of data in seconds.

* It simplifies the integration with Vertex AI, as the BigQuery Reader can be used with both custom and pre-built TensorFlow models on Vertex AI. Vertex AI is a unified platform for machine learning that provides various tools and features for data ingestion, data labeling, data preprocessing, model training, model tuning, model deployment, model monitoring, and model explainability.

The other options are less optimal for the following reasons:

* Option A: Using the BigQuery client library to load data into a dataframe, and using `tf.data.Dataset.from_tensor_slices()` to read it, introduces memory and performance issues. This option requires loading the entire BigQuery table into a Pandas dataframe, which can consume a lot of memory and cause out-of-memory errors. Moreover, using `tf.data.Dataset.from_tensor_slices()` to read the dataframe can be slow and inefficient, as it creates one slice per row of the dataframe, resulting in a large number of small tensors.

* Option B: Exporting data to CSV files in Cloud Storage, and using `tf.data.TextLineDataset()` to read them, introduces additional steps and complexity. This option requires exporting the BigQuery table to one or more CSV files in Cloud Storage, which can take a long time and consume a lot of storage space.

Moreover, using `tf.data.TextLineDataset()` to read the CSV files can be slow and error-prone, as it requires parsing and decoding each line of text, handling missing values and invalid data, and applying data transformations and validations.

* Option C: Converting the data into TFRecords, and using `tf.data.TFRecordDataset()` to read them, introduces additional steps and complexity. This option requires converting the BigQuery table into one or more TFRecord files, which are binary files that store serialized TensorFlow examples. This can take a long time and consume a lot of storage space. Moreover, using `tf.data.TFRecordDataset()` to read the TFRecord files requires defining and parsing the schema of the TensorFlow examples, which can be tedious and error-prone.

References:

* [TensorFlow I/O documentation]

* [BigQuery documentation]

* [Vertex AI documentation]

NEW QUESTION # 77

You have recently created a proof-of-concept (POC) deep learning model. You are satisfied with the overall architecture, but you need to determine the value for a couple of hyperparameters. You want to perform hyperparameter tuning on Vertex AI to determine both the appropriate embedding dimension for a categorical feature used by your model and the optimal learning rate.

You configure the following settings:

For the embedding dimension, you set the type to INTEGER with a `minValue` of 16 and `maxValue` of 64.

For the learning rate, you set the type to DOUBLE with a `minValue` of 10e-05 and `maxValue` of 10e-02.

You are using the default Bayesian optimization tuning algorithm, and you want to maximize model accuracy.

Training time is not a concern. How should you set the hyperparameter scaling for each hyperparameter and the maxParallelTrials?

- A. Use UNIT_LINEAR_SCALE for the embedding dimension, UNIT_LOG_SCALE for the learning rate, and a small number of parallel trials.
- **B. Use UNIT_LINEAR_SCALE for the embedding dimension, UNIT_LOG_SCALE for the learning rate, and a large number of parallel trials.**
- C. Use UNIT_LOG_SCALE for the embedding dimension, UNIT_LINEAR_SCALE for the learning rate, and a small number of parallel trials.
- D. Use UNIT_LOG_SCALE for the embedding dimension, UNIT_LINEAR_SCALE for the learning rate, and a large number of parallel trials.

Answer: B

Explanation:

The best option for performing hyperparameter tuning on Vertex AI to determine the appropriate embedding dimension and the optimal learning rate is to use UNIT_LINEAR_SCALE for the embedding dimension, UNIT_LOG_SCALE for the learning rate, and a large number of parallel trials. This option has the following advantages:

* It matches the appropriate scaling type for each hyperparameter, based on their range and distribution.

The embedding dimension is an integer hyperparameter that varies linearly between 16 and 64, so using UNIT_LINEAR_SCALE makes sense. The learning rate is a double hyperparameter that varies exponentially between 10e-05 and 10e-02, so using UNIT_LOG_SCALE is more suitable.

* It maximizes the exploration of the hyperparameter space, by using a large number of parallel trials.

Since training time is not a concern, using more trials can help find the best combination of hyperparameters that maximizes model accuracy. The default Bayesian optimization tuning algorithm can efficiently sample the hyperparameter space and converge to the optimal values.

The other options are less optimal for the following reasons:

* Option B: Using UNIT_LINEAR_SCALE for the embedding dimension, UNIT_LOG_SCALE for the learning rate, and a small number of parallel trials, reduces the exploration of the hyperparameter space, by using a small number of parallel trials. Since training time is not a concern, using fewer trials can miss some potentially good combinations of hyperparameters that maximize model accuracy. The default Bayesian optimization tuning algorithm can benefit from more trials to sample the hyperparameter space and converge to the optimal values.

* Option C: Using UNIT_LOG_SCALE for the embedding dimension, UNIT_LINEAR_SCALE for the learning rate, and a large number of parallel trials, mismatches the appropriate scaling type for each hyperparameter, based on their range and distribution. The embedding dimension is an integer hyperparameter that varies linearly between 16 and 64, so using UNIT_LOG_SCALE is not suitable.

The learning rate is a double hyperparameter that varies exponentially between 10e-05 and 10e-02, so using UNIT_LINEAR_SCALE makes less sense.

* Option D: Using UNIT_LOG_SCALE for the embedding dimension, UNIT_LINEAR_SCALE for the learning rate, and a small number of parallel trials, combines the drawbacks of option B and option C. It mismatches the appropriate scaling type for each hyperparameter, based on their range and distribution, and reduces the exploration of the hyperparameter space, by using a small number of parallel trials.

:

[Vertex AI: Hyperparameter tuning overview]

[Vertex AI: Configuring the hyperparameter tuning job]

NEW QUESTION # 78

Your company stores a large number of audio files of phone calls made to your customer call center in an on-premises database. Each audio file is in wav format and is approximately 5 minutes long. You need to analyze these audio files for customer sentiment. You plan to use the Speech-to-Text API. You want to use the most efficient approach. What should you do?

- A. 1 Upload the audio files to Cloud Storage
2. Call the speech: longrunningrecognize API endpoint to generate transcriptions
3. Call the predict method of an AutoML sentiment analysis model to analyze the transcriptions
- B. 1 Iterate over your local Tiles in Python
2. Use the Speech-to-Text Python library to create a speech.RecognitionAudio object and set the content to the audio file data
3. Call the speech: recognize API endpoint to generate transcriptions
4. Call the predict method of an AutoML sentiment analysis model to analyze the transcriptions
- C. 1 Iterate over your local files in Python
2 Use the Speech-to-Text Python Library to create a speech.RecognitionAudio object, and set the content to the audio file

P.S. Free & New Professional-Machine-Learning-Engineer dumps are available on Google Drive shared by PassSureExam:
https://drive.google.com/open?id=1U4tJzXXcMTJ8Zg1gsf_IH5EtP5jbJkV