Professional-Machine-Learning-Engineer Dumps

Google Professional Machine Learning Engineer

https://www.certleader.com/Professional-Machine-Learning-Engineer-dumps.html

NEW QUESTION 1
You are developing a Kubeflow pipeline on Google Kubernetes Engine. The first step in the pipeline is to issue a query against BigQuery. You plan to use the results of that query as the input to the next step in your pipeline. You want to achieve this in the easiest way possible. What should you do?

A. Use the BigQuery console to execute your query and then save the query results Into a new BigQuery table.
B. Write a Python script that uses the BigQuery API to execute queries against BigQuery Execute this script as the first step in your Kubeflow pipeline
C. Use the Kubeflow Pipelines domain-specific language to create a custom component that uses the Python BigQuery client library to execute queries
D. Locate the Kubeflow Pipelines repository on GitHub Find the BigQuery Query Component, copy that component's URL, and use it to load the component into your pipelin
E. Use the component to execute queries against BigQuery

Answer: A

NEW QUESTION 2
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

Answer: A

NEW QUESTION 3
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 customers1 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 thata 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 4
You are building a real-time prediction engine that streams files which may contain Personally Identifiable Information (Pll) to Google Cloud. You want to use the Cloud Data Loss Prevention (DLP) API to scan the files. How should you ensure that the Pll 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 5
You recently joined a machine learning team that will soon release a new project. As a lead on the project, you are asked to determine the production readiness of the ML components. The team has already tested features and data, model development, and infrastructure. Which additional readiness check should you recommend to the team?

A. Ensure that training is reproducible
B. Ensure that all hyperparameters are tuned
C. Ensure that model performance is monitored
D. Ensure that feature expectations are captured in the schema

Answer: B

NEW QUESTION 6
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 7
You are developing ML models with AI Platform for image segmentation on CT scans. You frequently update your model architectures based on the newest available research papers, and have to rerun training on the same dataset to benchmark their performance. You want to minimize computation costs and manual intervention while having version control for your code. What should you do?

A. Use Cloud Functions to identify changes to your code in Cloud Storage and trigger a retraining job
B. Use the gcloud command-line tool to submit training jobs on AI Platform when you update your code
C. Use Cloud Build linked with Cloud Source Repositories to trigger retraining when new code is pushed to the repository
D. Create an automated workflow in Cloud Composer that runs daily and looks for changes in code in Cloud Storage using a sensor.

Answer: A

NEW QUESTION 8
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 9
You have a functioning end-to-end ML pipeline that involves tuning the hyperparameters of your ML model using AI Platform, and then using the best-tuned parameters for training. Hypertuning is taking longer than expected and is delaying the downstream processes. You want to speed up the tuning job without significantly compromising its effectiveness. Which actions should you take?
Choose 2 answers

A. Decrease the number of parallel trials
B. Decrease the range of floating-point values
C. Set the early stopping parameter to TRUE
D. Change the search algorithm from Bayesian search to random search.
E. Decrease the maximum number of trials during subsequent training phases.

Answer: DE

NEW QUESTION 10
You have trained a deep neural network model on Google Cloud. The model has low loss on the training data, but is performing worse on the validation data. You want the model to be resilient to overfitting. Which strategy should you use when retraining the model?

A. Apply a dropout parameter of 0 2, and decrease the learning rate by a factor of 10
B. Apply a 12 regularization parameter of 0.4, and decrease the learning rate by a factor of 10.
C. Run a hyperparameter tuning job on AI Platform to optimize for the L2 regularization and dropout parameters
D. Run a hyperparameter tuning job on AI Platform to optimize for the learning rate, and increase the number of neurons by a factor of 2.

Answer: A

NEW QUESTION 10
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

Answer: D

NEW QUESTION 13
You are training a TensorFlow model on a structured data set with 100 billion records stored in several CSV files. You need to improve the input/output execution performance. What should you do?

A. Load the data into BigQuery and read the data from BigQuery.
B. Load the data into Cloud Bigtable, and read the data from Bigtable
C. Convert the CSV files into shards of TFRecords, and store the data in Cloud Storage
D. Convert the CSV files into shards of TFRecords, and store the data in the Hadoop Distributed File System (HDFS)

Answer: B

NEW QUESTION 17
You work for an advertising company and want to understand the effectiveness of your company's latest advertising campaign. You have streamed 500 MB of campaign data into BigQuery. You want to query the table, and then manipulate the results of that query with a pandas dataframe in an AI Platform notebook. What should you do?

A. Use AI Platform Notebooks' BigQuery cell magic to query the data, and ingest the results as a pandas dataframe
B. Export your table as a CSV file from BigQuery to Google Drive, and use the Google Drive API to ingest the file into your notebook instance
C. Download your table from BigQuery as a local CSV file, and upload it to your AI Platform notebook instance Use panda
D. read_csv to ingest the file as a pandas dataframe
E. From a bash cell in your AI Platform notebook, use the bq extract command to export the table as a CSV file to Cloud Storage, and then use gsutii cp to copy the data into the notebook Use panda
F. read_csv to ingest the file as a pandas dataframe

Answer: B

NEW QUESTION 21
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:

```
AuthorA:Political Party A
    TextA1: [SentenceA11, SentenceA12, SentenceA13, ...]
    TextA2: [SentenceA21, SentenceA22, SentenceA23, ...]
    ...
AuthorB:Political Party B
    TextB1: [SentenceB11, SentenceB12, SentenceB13, ...]
    TextB2: [SentenceB21, SentenceB22, SentenceB23, ...]
    ...
AuthorC:Political Party B
    TextC1: [SentenceC11, SentenceC12, SentenceC13, ...]
    TextC2: [SentenceC21, SentenceC22, SentenceC23, ...]
    ...
AuthorD:Political Party A
    TextD1: [SentenceD11, SentenceD12, SentenceD13, ...]
    TextD2: [SentenceD21, SentenceD22, SentenceD23, ...]
    ...
...
```

A)
Distribute texts randomly across the train-test-eval subsets:
    Train set: [TextA1, TextB2, ...]
    Test set: [TextA2, TextC1, TextD2, ...]
    Eval set: [TextB1, TextC2, TextD1, ...]

B)
Distribute authors randomly across the train-test-eval subsets: (*)
    Train set: [TextA1, TextA2, TextD1, TextD2, ...]
    Test set: [TextB1, TextB2, ...]
    Eval set: [TextC1, TextC2, ...]

C)
Distribute sentences randomly across the train-test-eval subsets:
    Train set: [SentenceA11, SentenceA21, Sentence B11, SentenceB21, SentenceC11, SentenceD21, ...]
    Test set: [SentenceA12, SentenceA22, Sentence B12, SentenceC22, SentenceC12, SentenceD22, ...]
    Eval set: [SentenceA13, SentenceA23, Sentence B13, SentenceC23, SentenceC13, SentenceD31, ...]

D)
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, ...]

A. Option A
B. Option B
C. Option C
D. Option D

Answer: D

NEW QUESTION 25
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 AI Platform's continuous evaluation service to ensure that the models have high accuracy on your test data set. 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.

Answer: C

NEW QUESTION 30
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?

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 experiment
E. Collect the results in a shared Google Sheetsfile, and query the results using the Google Sheets API

Answer: A

NEW QUESTION 31
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 model every month. You want to follow Google-recommended best practices. How should you configure the end-to-end architecture of the predictive model?

A. Configure Kubeflow Pipelines to schedule your multi-step workflow from training to deploying your model.
B. Use a model trained and deployed on BigQuery ML and trigger retraining with the scheduled query feature in BigQuery
C. Write a Cloud Functions script that launches a training and deploying job on Ai Platform that is triggered by Cloud Scheduler
D. Use Cloud Composer to programmatically schedule a Dataflow job that executes the workflow from training to deploying your model

Answer: B

NEW QUESTION 33
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?

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: C

NEW QUESTION 35
You built and manage a production system that is responsible for predicting sales numbers. Model accuracy is crucial, because the production model is required to keep up with market changes. Since being deployed to production, the model hasn't changed; however the accuracy of the model has steadily deteriorated. What issue is most likely causing the steady decline in model accuracy?

A. Poor data quality
B. Lack of model retraining
C. Too few layers in the model for capturing information
D. Incorrect data split ratio during model training, evaluation, validation, and test

Answer: D

NEW QUESTION 40
You are an ML engineer in the contact center of a large enterprise. You need to build a sentiment analysis tool that predicts customer sentiment from recorded phone conversations. You need to identify the best approach to building a model while ensuring that the gender, age, and cultural differences of the customers who called the contact center do not impact any stage of the model development pipeline and results. What should you do?

A. Extract sentiment directly from the voice recordings
B. Convert the speech to text and build a model based on the words
C. Convert the speech to text and extract sentiments based on the sentences
D. Convert the speech to text and extract sentiment using syntactical analysis

Answer: C

NEW QUESTION 41

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: ['driversjicense', '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 46
You have written unit tests for a Kubeflow Pipeline that require custom libraries. You want to automate the execution of unit tests with each new push to your development branch in Cloud Source Repositories. What should you do?

A. Write a script that sequentially performs the push to your development branch and executes the unit tests on Cloud Run
B. Using Cloud Build, set an automated trigger to execute the unit tests when changes are pushed to your development branch.
C. Set up a Cloud Logging sink to a Pub/Sub topic that captures interactions with Cloud Source Repositories Configure a Pub/Sub trigger for Cloud Run, and execute the unit tests on Cloud Run.
D. Set up a Cloud Logging sink to a Pub/Sub topic that captures interactions with Cloud Source Repositorie
E. Execute the unit tests using a Cloud Function that is triggered when messages are sent to the Pub/Sub topic

Answer: B

NEW QUESTION 49
......

Thank You for Trying Our Product

* 100% Pass or Money Back

    All our products come with a 90-day Money Back Guarantee.

* One year free update

    You can enjoy free update one year. 24x7 online support.

* Trusted by Millions

    We currently serve more than 30,000,000 customers.

* Shop Securely

    All transactions are protected by VeriSign!

100% Pass Your Professional-Machine-Learning-Engineer Exam with Our Prep Materials Via below:

https://www.certleader.com/Professional-Machine-Learning-Engineer-dumps.html