Google

Exam Questions Professional-Machine-Learning-Engineer

Google Professional Machine Learning Engineer

NEW QUESTION 1
You recently designed and built a custom neural network that uses critical dependencies specific to your organization's framework. You need to train the model using a managed training service on Google Cloud. However, the ML framework and related dependencies are not supported by AI Platform Training. Also, both your model and your data are too large to fit in memory on a single machine. Your ML framework of choice uses the scheduler, workers, and servers distribution structure. What should you do?

A. Use a built-in model available on AI Platform Training
B. Build your custom container to run jobs on AI Platform Training
C. Build your custom containers to run distributed training jobs on AI Platform Training
D. Reconfigure your code to a ML framework with dependencies that are supported by AI Platform Training

Answer: C

NEW QUESTION 2
You are an ML engineer at a large grocery retailer with stores in multiple regions. You have been asked to create an inventory prediction model. Your models features include region, location, historical demand, and seasonal popularity. You want the algorithm to learn from new inventory data on a daily basis. Which algorithms should you use to build the model?

A. Classification
B. Reinforcement Learning
C. Recurrent Neural Networks (RNN)
D. Convolutional Neural Networks (CNN)

Answer: B

NEW QUESTION 3
You need to design a customized deep neural network in Keras that will predict customer purchases based on their purchase history. You want to explore model performance using multiple model architectures, store training data, and be able to compare the evaluation metrics in the same dashboard. What should you do?

A. Create multiple models using AutoML Tables
B. Automate multiple training runs using Cloud Composer
C. Run multiple training jobs on AI Platform with similar job names
D. Create an experiment in Kubeflow Pipelines to organize multiple runs

Answer: C

NEW QUESTION 4
You work for a toy manufacturer that has been experiencing a large increase in demand. You need to build an ML model to reduce the amount of time spent by quality control inspectors checking for product defects. Faster defect detection is a priority. The factory does not have reliable Wi-Fi. Your company wants to implement the new ML model as soon as possible. Which model should you use?

A. AutoML Vision model
B. AutoML Vision Edge mobile-versatile-1 model
C. AutoML Vision Edge mobile-low-latency-1 model
D. AutoML Vision Edge mobile-high-accuracy-1 model

Answer: A

NEW QUESTION 5
You are training a Resnet model on AI Platform using TPUs to visually categorize types of defects in automobile engines. You capture the training profile using the Cloud TPU profiler plugin and observe that it is highly input-bound. You want to reduce the bottleneck and speed up your model training process. Which modifications should you make to the tf .data dataset?
Choose 2 answers

A. Use the interleave option for reading data
B. Reduce the value of the repeat parameter
C. Increase the buffer size for the shuffle option.
D. Set the prefetch option equal to the training batch size
E. Decrease the batch size argument in your transformation

Answer: AD

NEW QUESTION 6
You are going to train a DNN regression model with Keras APIs using this code:

```
model = tf.keras.Sequential()
model.add(tf.keras.layers.Dense(
    256,
    use_bias=True,
    activation='relu',
    kernel_initializer=None,
    kernel_regularizer=None,
    input_shape=(500,)))
model.add(tf.keras.layers.Dropout(rate=0.25))
model.add(tf.keras.layers.Dense(
    128, use_bias=True,
    activation='relu',
    kernel_initializer='uniform',
    kernel_regularizer='l2'))
model.add(tf.keras.layers.Dropout(rate=0.25))
model.add(tf.keras.layers.Dense(
    2, use_bias=False,
    activation='softmax'))
    model.compile(loss='mse')
```

How many trainable weights does your model have? (The arithmetic below is correct.)

A. 501*256+257*128+2 = 161154
B. 500*256+256*128+128*2 = 161024
C. 501*256+257*128+128*2=161408
D. 500*256*0 25+256*128*0 25+128*2 = 40448

Answer: D

NEW QUESTION 7
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. 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: C

NEW QUESTION 8
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 9
You have deployed multiple versions of an image classification model on AI Platform. You want to monitor the performance of the model versions overtime. 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

Answer: B

NEW QUESTION 10

You are building a real-time prediction engine that streams files which may contain Personally Identifiable Information (PII) to Google Cloud. You want to use the Cloud Data Loss Prevention (DLP) API to scan the files. How should you ensure that the PII 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 10
You have been asked to develop an input pipeline for an ML training model that processes images from disparate sources at a low latency. You discover that your input data does not fit in memory. How should you create a dataset following Google-recommended best practices?

A. Create a tf.data.Dataset.prefetch transformation
B. Convert the images to tf .Tensor Objects, and then run Datase
C. from_tensor_slices{).
D. Convert the images to tf .Tensor Objects, and then run t
E. dat
F. Datase
G. from_tensors ().
H. Convert the images Into TFRecords, store the images in Cloud Storage, and then use the t
I. data API to read the images for training

Answer: D

NEW QUESTION 11
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 15
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 17
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 20
You have trained a text classification model in TensorFlow using AI Platform. You want to use the trained model for batch predictions on text data stored in BigQuery while minimizing computational overhead. What should you do?

A. Export the model to BigQuery ML.
B. Deploy and version the model on AI Platform.
C. Use Dataflow with the SavedModel to read the data from BigQuery
D. Submit a batch prediction job on AI Platform that points to the model location in Cloud Storage.

Answer: A

NEW QUESTION 25
You are designing an ML recommendation model for shoppers on your company's ecommerce website. You will use Recommendations AI to build, test, and deploy your system. How should you develop recommendations that increase revenue while following best practices?

A. Use the "Other Products You May Like" recommendation type to increase the click-through rate
B. Use the "Frequently Bought Together' recommendation type to increase the shopping cart size for each order.
C. Import your user events and then your product catalog to make sure you have the highest quality event stream
D. Because it will take time to collect and record product data, use placeholder values for the product catalog to test the viability of the model.

Answer: C

NEW QUESTION 28
You are developing models to classify customer support emails. You created models with TensorFlow Estimators using small datasets on your on-premises system, but you now need to train the models using large datasets to ensure high performance. You will port your models to Google Cloud and want to minimize code refactoring and infrastructure overhead for easier migration from on-prem to cloud. What should you do?

A. Use AI Platform for distributed training
B. Create a cluster on Dataproc for training
C. Create a Managed Instance Group with autoscaling
D. Use Kubeflow Pipelines to train on a Google Kubernetes Engine cluster.

Answer: D

NEW QUESTION 29
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 33
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 35
You are training an LSTM-based model on AI Platform to summarize text using the following job submission script:

```
gcloud ai-platform jobs submit training $JOB_NAME \
        --package-path $TRAINER_PACKAGE_PATH \
        --module-name $MAIN_TRAINER_MODULE \
        --job-dir $JOB_DIR \
        --region $REGION \
        --scale-tier basic \
        -- \
        --epochs 20 \
        --batch_size=32 \
        --learning_rate=0.001 \
```

You want to ensure that training time is minimized without significantly compromising the accuracy of your model. What should you do?

A. Modify the 'epochs' parameter
B. Modify the 'scale-tier' parameter
C. Modify the batch size' parameter
D. Modify the 'learning rate' parameter

Answer: A

NEW QUESTION 36
You work for a global footwear retailer and need to predict when an item will be out of stock based on historical inventory data. 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?

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

NEW QUESTION 39
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 40
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 45
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 49
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 54
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 56
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 59
You are an ML engineer at a global shoe store. You manage the ML models for the company's website. You are asked to build a model that will recommend new products to the user based on their purchase behavior and similarity with other users. What should you do?

A. Build a classification model
B. Build a knowledge-based filtering model
C. Build a collaborative-based filtering model
D. Build a regression model using the features as predictors

Answer: C

NEW QUESTION 63
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 I AM 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 categorie
D. Apply a label to each created resource so that users can filter the results by label when viewing or monitoring the resources
E. 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: B

NEW QUESTION 67
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 69
......

Thank You for Trying Our Product

We offer two products:

1st - We have Practice Tests Software with Actual Exam Questions

2nd - Questons and Answers in PDF Format

Professional-Machine-Learning-Engineer Practice Exam Features:

* Professional-Machine-Learning-Engineer Questions and Answers Updated Frequently

* Professional-Machine-Learning-Engineer Practice Questions Verified by Expert Senior Certified Staff

* Professional-Machine-Learning-Engineer Most Realistic Questions that Guarantee you a Pass on Your FirstTry

* Professional-Machine-Learning-Engineer Practice Test Questions in Multiple Choice Formats and Updatesfor 1 Year

100% Actual & Verified — Instant Download, Please Click
Order The Professional-Machine-Learning-Engineer Practice Test Here