



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 need to train a computer vision model that predicts the type of government ID present in a given image using a GPU-powered virtual machine on Compute Engine. You use the following parameters:

- Optimizer: SGD
- Image shape = 224x224
- Batch size = 64
- Epochs = 10
- Verbose = 2

During training you encounter the following error: ResourceExhaustedError: out of Memory (oom) when allocating tensor. What should you do?

- A. Change the optimizer
- B. Reduce the batch size
- C. Change the learning rate
- D. Reduce the image shape

Answer: A

NEW QUESTION 3

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?

- 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: D

NEW QUESTION 4

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 5

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

Answer: B

NEW QUESTION 6

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 7

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?

- 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: C

NEW QUESTION 8

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 9

You work with a data engineering team that has developed a pipeline to clean your dataset and save it in a Cloud Storage bucket. You have created an ML model and want to use the data to refresh your model as soon as new data is available. As part of your CI/CD workflow, you want to automatically run a Kubeflow Pipelines training job on Google Kubernetes Engine (GKE). How should you architect this workflow?

- A. Configure your pipeline with Dataflow, which saves the files in Cloud Storage After the file is saved, start the training job on a GKE cluster
- B. Use App Engine to create a lightweight python client that continuously polls Cloud Storage for new files As soon as a file arrives, initiate the training job
- C. Configure a Cloud Storage trigger to send a message to a Pub/Sub topic when a new file is available in a storage bucket
- D. Use a Pub/Sub-triggered Cloud Function to start the training job on a GKE cluster
- E. Use Cloud Scheduler to schedule jobs at a regular interval
- F. For the first step of the job
- G. check the timestamp of objects in your Cloud Storage bucket If there are no new files since the last run, abort the job.

Answer: A

NEW QUESTION 10

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 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 15

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 19

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 22

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 L2 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 25

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 26

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 30

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 gsutil 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 32

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 36

You have a demand forecasting pipeline in production that uses Dataflow to preprocess raw data prior to model training and prediction. During preprocessing, you employ Z-score normalization on data stored in BigQuery and write it back to BigQuery. New training data is added every week. You want to make the process more efficient by minimizing computation time and manual intervention. What should you do?

- A. Normalize the data using Google Kubernetes Engine
- B. Translate the normalization algorithm into SQL for use with BigQuery
- C. Use the `normalizer_fn` argument in TensorFlow's Feature Column API
- D. Normalize the data with Apache Spark using the Dataproc connector for BigQuery

Answer: B

NEW QUESTION 38

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 40

You manage a team of data scientists who use a cloud-based backend system to submit training jobs. This system has become very difficult to administer, and you want to use a managed service instead. The data scientists you work with use many different frameworks, including Keras, PyTorch, theano, Scikit-team, and custom libraries. What should you do?

- A. Use the AI Platform custom containers feature to receive training jobs using any framework
- B. Configure Kubeflow to run on Google Kubernetes Engine and receive training jobs through TFJob
- C. Create a library of VM images on Compute Engine; and publish these images on a centralized repository
- D. Set up Slurm workload manager to receive jobs that can be scheduled to run on your cloud infrastructure.

Answer: D

NEW QUESTION 43

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 47

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: `['driverslicense', '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 49

.....

Thank You for Trying Our Product

We offer two products:

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

2nd - Questions 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 First Try
- * Professional-Machine-Learning-Engineer Practice Test Questions in Multiple Choice Formats and Updates for 1 Year

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