PROFESSIONAL-MACHINE-LEARNING-ENGINEER Q&As

Professional Machine Learning Engineer

Pass Google PROFESSIONAL-MACHINE-LEARNING-ENGINEER Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

https://www.pass2lead.com/professional-machine-learning-engineer.html

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by Google Official Exam Center

QUESTION 1

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 Cloud, 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. Periodically conduct a bulk scan of that bucket using the DLP API, and move the data to either the Sensitive or Non-Sensitive bucket.

Correct Answer: D

https://cloud.google.com/architecture/automating-classification-of-data-uploaded-to-cloud-storage#building_the_quarantine_and_classification_pipeline

QUESTION 2

You have deployed a model on Vertex AI for real-time inference. During an online prediction request, you get an "Out of Memory" error. What should you do?

A. Use batch prediction mode instead of online mode.

B. Send the request again with a smaller batch of instances.

C. Use base64 to encode your data before using it for prediction.

D. Apply for a quota increase for the number of prediction requests.

Correct Answer: B

https://cloud.google.com/ai-platform/training/docs/troubleshooting

QUESTION 3

You recently deployed an ML model. Three months after deployment, you notice that your model is underperforming on certain subgroups, thus potentially leading to biased results. You suspect that the inequitable performance is due to class imbalances in the training data, but you cannot collect more data. What should you do? (Choose two.)

A. Remove training examples of high-performing subgroups, and retrain the model.

B. Add an additional objective to penalize the model more for errors made on the minority class, and retrain the model

C. Remove the features that have the highest correlations with the majority class.

D. Upsample or reweight your existing training data, and retrain the model

E. Redeploy the model, and provide a label explaining the model\\'s behavior to users.

Correct Answer: BD

https://machinelearningmastery.com/tactics-to-combat-imbalanced-classes-in-your-machine-learning-dataset/

QUESTION 4

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 Vertex AI Pipelines to execute the experiments. Query the results stored in MetadataStore using the Vertex AI API.

B. Use Vertex AI Training to execute the experiments. Write the accuracy metrics to BigQuery, and query the results using the BigQuery API.

C. Use Vertex AI Training to execute the experiments. Write the accuracy metrics to Cloud Monitoring, and query the results using the Monitoring API.

D. Use Vertex AI Workbench user-managed notebooks to execute the experiments. Collect the results in a shared Google Sheets file, and query the results using the Google Sheets API.

Correct Answer: A

The Vertex AI Pipelines provide a powerful tool for automating machine learning workflows, including data preparation, training, and deployment. MetadataStore can be used to track the performance of different models by logging accuracy metrics and other important information. The Vertex AI API can then be used to query the metadata store and retrieve the results of different experiments.

QUESTION 5

You have successfully deployed to production a large and complex TensorFlow model trained on tabular data. You want to predict the lifetime value (LTV) field for each subscription stored in the BigQuery table named subscription. subscriptionPurchase in the project named my-fortune500-company-project.

You have organized all your training code, from preprocessing data from the BigQuery table up to deploying the validated model to the Vertex AI endpoint, into a TensorFlow Extended (TFX) pipeline. You want to prevent prediction drift, i.e., a situation when a feature data distribution in production changes significantly over time. What should you do?

A. Implement continuous retraining of the model daily using Vertex AI Pipelines.

B. Add a model monitoring job where 10% of incoming predictions are sampled 24 hours.

C. Add a model monitoring job where 90% of incoming predictions are sampled 24 hours.

D. Add a model monitoring job where 10% of incoming predictions are sampled every hour.

Correct Answer: B

https://cloud.google.com/vertex-ai/docs/model-monitoring/overview https://cloud.google.com/vertex-ai/docs/model-monitoring/using-model-monitoring#drift-detection

QUESTION 6

You are training an object detection machine learning model on a dataset that consists of three million X-ray images, each roughly 2 GB in size. You are using Vertex AI Training to run a custom training application on a Compute Engine instance with 32-cores, 128 GB of RAM, and 1 NVIDIA P100 GPU. You notice that model training is taking a very long time. You want to decrease training time without sacrificing model performance. What should you do?

A. Increase the instance memory to 512 GB and increase the batch size.

B. Replace the NVIDIA P100 GPU with a v3-32 TPU in the training job.

C. Enable early stopping in your Vertex AI Training job.

D. Use the tf.distribute.Strategy API and run a distributed training job.

Correct Answer: D

QUESTION 7

You deployed an ML model into production a year ago. Every month, you collect all raw requests that were sent to your model prediction service during the previous month. You send a subset of these requests to a human labeling service to evaluate your model\\'s performance. After a year, you notice that your model\\'s performance sometimes degrades significantly after a month, while other times it takes several months to notice any decrease in performance. The labeling service is costly, but you also need to avoid large performance degradations. You want to determine how often you should retrain your model to maintain a high level of performance while minimizing cost. What should you do?

A. Train an anomaly detection model on the training dataset, and run all incoming requests through this model. If an anomaly is detected, send the most recent serving data to the labeling service.

B. Identify temporal patterns in your model\\'s performance over the previous year. Based on these patterns, create a schedule for sending serving data to the labeling service for the next year.

C. Compare the cost of the labeling service with the lost revenue due to model performance degradation over the past year. If the lost revenue is greater than the cost of the labeling service, increase the frequency of model retraining; otherwise, decrease the model retraining frequency.

D. Run training-serving skew detection batch jobs every few days to compare the aggregate statistics of the features in the training dataset with recent serving data. If skew is detected, send the most recent serving data to the labeling service.

Correct Answer: D

https://cloud.google.com/blog/topics/developers-practitioners/monitor-models-training-serving-skew-vertex-aiew-vertex-aiandved=2ahUKEwiRg_aoj9n8AhWb7TgGHcGCDREQFnoECAwQAQandusg=AOvVaw197NneIJM0ra7fLq2zsOin

QUESTION 8

You work for an online publisher that delivers news articles to over 50 million readers. You have built an AI model that recommends content for the company\\'s weekly newsletter. A recommendation is considered successful if the article is opened within two days of the newsletter\\'s published date and the user remains on the page for at least one minute.

All the information needed to compute the success metric is available in BigQuery and is updated hourly. The model is trained on eight weeks of data, on average its performance degrades below the acceptable baseline after five weeks, and training time is 12 hours. You want to ensure that the model\\'s performance is above the acceptable baseline while minimizing cost. How should you monitor the model to determine when retraining is necessary?

A. Use Vertex AI Model Monitoring to detect skew of the input features with a sample rate of 100% and a monitoring frequency of two days.

B. Schedule a cron job in Cloud Tasks to retrain the model every week before the newsletter is created.

C. Schedule a weekly query in BigQuery to compute the success metric.

D. Schedule a daily Dataflow job in Cloud Composer to compute the success metric.

Correct Answer: C

https://cloud.google.com/blog/topics/developers-practitioners/continuous-model-evaluation-bigquery-ml-stored-procedures-and-cloud-scheduler

---

QUESTION 9

You are the Director of Data Science at a large company, and your Data Science team has recently begun using the Kubeflow Pipelines SDK to orchestrate their training pipelines. Your team is struggling to integrate their custom Python code into the Kubeflow Pipelines SDK. How should you instruct them to proceed in order to quickly integrate their code with the Kubeflow Pipelines SDK?

A. Use the func_to_container_op function to create custom components from the Python code.

B. Use the predefined components available in the Kubeflow Pipelines SDK to access Dataproc, and run the custom code there.

C. Package the custom Python code into Docker containers, and use the load_component_from_file function to import the containers into the pipeline.

D. Deploy the custom Python code to Cloud Functions, and use Kubeflow Pipelines to trigger the Cloud Function.

Correct Answer: A

https://kubeflow-pipelines.readthedocs.io/en/stable/source/kfp.components.html?highlight=func_to_container_op%20#kfp.components.func_to_container_op

---

QUESTION 10

You lead a data science team at a large international corporation. Most of the models your team trains are large-scale models using high-level TensorFlow APIs on AI Platform with GPUs. Your team usually takes a few weeks or months to iterate on a new version of a model. You were recently asked to review your team\\'s spending. How should you reduce your Google Cloud compute costs without impacting the model\\'s performance?

A. Use AI Platform to run distributed training jobs with checkpoints.

B. Use AI Platform to run distributed training jobs without checkpoints.

C. Migrate to training with Kuberflow on Google Kubernetes Engine, and use preemptible VMs with checkpoints.

D. Migrate to training with Kuberflow on Google Kubernetes Engine, and use preemptible VMs without checkpoints.

Correct Answer: C

https://cloud.google.com/blog/products/ai-machine-learning/reduce-the-costs-of-ml-workflows-with-preemptible-vms-and-gpus?hl=en

QUESTION 11

You work for a gaming company that develops massively multiplayer online (MMO) games. You built a TensorFlow model that predicts whether players will make in-app purchases of more than $10 in the next two weeks. The model\\'s predictions will be used to adapt each user\\'s game experience. User data is stored in BigQuery. How should you serve your model while optimizing cost, user experience, and ease of management?

A. Import the model into BigQuery ML. Make predictions using batch reading data from BigQuery, and push the data to Cloud SQL

B. Deploy the model to Vertex AI Prediction. Make predictions using batch reading data from Cloud Bigtable, and push the data to Cloud SQL.

C. Embed the model in the mobile application. Make predictions after every in-app purchase event is published in Pub/Sub, and push the data to Cloud SQL.

D. Embed the model in the streaming Dataflow pipeline. Make predictions after every in-app purchase event is published in Pub/Sub, and push the data to Cloud SQL.

Correct Answer: A

https://cloud.google.com/bigquery-ml/docs/reference/standard-sql/bigqueryml-syntax-create-tensorflow

QUESTION 12

You are developing an ML model that uses sliced frames from video feed and creates bounding boxes around specific objects. You want to automate the following steps in your training pipeline: ingestion and preprocessing of data in Cloud Storage, followed by training and hyperparameter tuning of the object model using Vertex AI jobs, and finally deploying the model to an endpoint. You want to orchestrate the entire pipeline with minimal cluster management. What approach should you use?

A. Use Kubeflow Pipelines on Google Kubernetes Engine.

B. Use Vertex AI Pipelines with TensorFlow Extended (TFX) SDK.

C. Use Vertex AI Pipelines with Kubeflow Pipelines SDK.

D. Use Cloud Composer for the orchestration.

Correct Answer: C

From:

https://cloud.google.com/vertex-ai/docs/pipelines/build-pipeline#sdk

"1. If you use TensorFlow in an ML workflow that processes terabytes of structured data or text data, we recommend that you build your pipeline using TFX.

To learn more about building a TFX pipeline, follow the TFX getting started tutorials.

To learn more about using Vertex AI Pipelines to run a TFX pipeline, follow the TFX on Google Cloud tutorials.

2. For other use cases, we recommend that you build your pipeline using the Kubeflow Pipelines SDK. By building a pipeline with the Kubeflow Pipelines SDK, you can implement your workflow by building custom components or reusing prebuilt components, such as the Google Cloud Pipeline Components. Google Cloud Pipeline Components make it easier to use Vertex AI services like AutoML in your pipeline."

QUESTION 13

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?

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.

Correct Answer: A

https://cloud.google.com/ai-platform/prediction/docs/batch-predict

---

QUESTION 14

You recently developed a deep learning model using Keras, and now you are experimenting with different training strategies. First, you trained the model using a single GPU, but the training process was too slow. Next, you distributed the training across 4 GPUs using tf.distribute.MirroredStrategy (with no other changes), but you did not observe a decrease in training time. What should you do?

A. Distribute the dataset with tf.distribute.Strategy.experimental_distribute_dataset

B. Create a custom training loop.

C. Use a TPU with tf.distribute.TPUStrategy.

D. Increase the batch size.

Correct Answer: D

https://www.tensorflow.org/guide/gpu_performance_analysis

QUESTION 15

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 Dataset.from_tensor_slices().

C. Convert the images to tf.Tensor objects, and then run tf.data.Dataset.from_tensors().

D. Convert the images into TFRecords, store the images in Cloud Storage, and then use the tf.data API to read the images for training.

Correct Answer: D

Cite from Google Pag: to construct a Dataset from data in memory, use tf.data.Dataset.from_tensors() or tf.data.Dataset.from_tensor_slices(). When input data is stored in a file (not in memory), the recommended TFRecord format, you can use tf.data.TFRecordDataset().

tf.data.Dataset is for data in memory. tf.data.TFRecordDataset is for data in non-memory storage.