 Custom View Settings

Question #37   *Topic 1*

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.

Question #38   *Topic 1*

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.

Question #39   *Topic 1*

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. Use a Pub/Sub-triggered Cloud Function to start the training job on a GKE cluster.

D. Use Cloud Scheduler to schedule jobs at a regular interval. For the first step of the job, check the timestamp of objects in your Cloud Storage bucket. If there are no new files since the last run, abort the job.

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 two.)

    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.

← Previous Questions

Next Questions →