Custom View Settings

Question #29 *Topic 1*

You have trained a model on a dataset that required computationally expensive preprocessing operations. You need to execute the same preprocessing at prediction time. You deployed the model on AI Platform for high-throughput online prediction. Which architecture should you use?

A. Validate the accuracy of the model that you trained on preprocessed data. Create a new model that uses the raw data and is available in real time. Deploy the new model onto AI Platform for online prediction.

B. Send incoming prediction requests to a Pub/Sub topic. Transform the incoming data using a Dataflow job. Submit a prediction request to AI Platform using the transformed data. Write the predictions to an outbound Pub/Sub queue.

C. Stream incoming prediction request data into Cloud Spanner. Create a view to abstract your preprocessing logic. Query the view every second for new records. Submit a prediction request to AI Platform using the transformed data. Write the predictions to an outbound Pub/Sub queue.

D. Send incoming prediction requests to a Pub/Sub topic. Set up a Cloud Function that is triggered when messages are published to the Pub/Sub topic. Implement your preprocessing logic in the Cloud Function. Submit a prediction request to AI Platform using the transformed data. Write the predictions to an outbound Pub/Sub queue.

Question #30 *Topic 1*

Your team trained and tested a DNN regression model with good results. Six months after deployment, the model is performing poorly due to a change in the distribution of the input data. How should you address the input differences in production?

A. Create alerts to monitor for skew, and retrain the model.

B. Perform feature selection on the model, and retrain the model with fewer features.

C. Retrain the model, and select an L2 regularization parameter with a hyperparameter tuning service.

D. Perform feature selection on the model, and retrain the model on a monthly basis with fewer features.

Question #31

Topic 1

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 = 224—224
✑ 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.

Question #32

Topic 1

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.

← Previous Questions

Next Questions →