

[Custom View Settings](#)

Question #49

Topic 1

You work for an online travel agency that also sells advertising placements on its website to other companies. You have been asked to predict the most relevant web banner that a user should see next. Security is important to your company. The model latency requirements are 300ms@p99, the inventory is thousands of web banners, and your exploratory analysis has shown that navigation context is a good predictor. You want to implement the simplest solution. How should you configure the prediction pipeline?

- A. Embed the client on the website, and then deploy the model on AI Platform Prediction.
- B. Embed the client on the website, deploy the gateway on App Engine, and then deploy the model on AI Platform Prediction.
- C. Embed the client on the website, deploy the gateway on App Engine, deploy the database on Cloud Bigtable for writing and for reading the user's navigation context, and then deploy the model on AI Platform Prediction.
- D. Embed the client on the website, deploy the gateway on App Engine, deploy the database on Memorystore for writing and for reading the user's navigation context, and then deploy the model on Google Kubernetes Engine.

Question #50

Topic 1

Your team is building a convolutional neural network (CNN)-based architecture from scratch. The preliminary experiments running on your on-premises CPU-only infrastructure were encouraging, but have slow convergence. You have been asked to speed up model training to reduce time-to-market. You want to experiment with virtual machines (VMs) on Google Cloud to leverage more powerful hardware. Your code does not include any manual device placement and has not been wrapped in Estimator model-level abstraction. Which environment should you train your model on?

- A. AVM on Compute Engine and 1 TPU with all dependencies installed manually.
- B. AVM on Compute Engine and 8 GPUs with all dependencies installed manually.
- C. A Deep Learning VM with an n1-standard-2 machine and 1 GPU with all libraries pre-installed.
- D. A Deep Learning VM with more powerful CPU e2-highcpu-16 machines with all libraries pre-installed.

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 IAM 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 categories. Apply a label to each created resource so that users can filter the results by label when viewing or monitoring the resources.
- D. 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

You are training a deep learning model for semantic image segmentation with reduced training time. While using a Deep Learning VM Image, you receive the following error: The resource 'projects/deeplearning-platform/zones/europe-west4-c/acceleratorTypes/nvidia-tesla-k80' was not found. What should you do?

- A. Ensure that you have GPU quota in the selected region.
- B. Ensure that the required GPU is available in the selected region.
- C. Ensure that you have preemptible GPU quota in the selected region.
- D. Ensure that the selected GPU has enough GPU memory for the workload.

[← Previous Questions](#)[Next Questions →](#)