Free Questions for Professional-Machine-Learning-Engineer by go4braindumps

Shared by Charles on 06-06-2022

For More Free Questions and Preparation Resources

Check the Links on Last Page

Question 1

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?

Options:

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.

Answer:

A

Question 2

You work for a bank and are building a random forest model for fraud detection. You have a dataset that includes transactions, of which 1% are identified as fraudulent.

Which data transformation strategy would likely improve the performance of your classifier?

Options:

A- Write your data in TFRecords.

B- Z-normalize all the numeric features.

C- Oversample the fraudulent transaction 10 times.

D- Use one-hot encoding on all categorical features.

Answer:

C

Question 3

You are working on a Neural Network-based project. The dataset provided to you has columns with different ranges. While preparing the data for model training, you discover that gradient optimization is having difficulty moving weights to a good solution. What should you do?

Options:

A- Use feature construction to combine the strongest features.

B- Use the representation transformation (normalization) technique.

C- Improve the data cleaning step by removing features with missing values.

D- Change the partitioning step to reduce the dimension of the test set and have a larger training set.

Answer:

C

Question 4

Question Type: **MultipleChoice**

Your company manages a video sharing website where users can watch and upload videos. You need to create an ML model to predict which newly uploaded videos will be the most popular so that those videos can be prioritized on your company's website.

Which result should you use to determine whether the model is successful?

Answer:

C

Question 5

Question Type: MultipleChoice

You work for a credit card company and have been asked to create a custom fraud detection model based on historical data using AutoML Tables. You need to prioritize detection of fraudulent transactions while minimizing false positives. Which optimization objective

should you use when training the model?

Options:

A- An optimization objective that minimizes Log loss

B- An optimization objective that maximizes the Precision at a Recall value of 0.50

C- An optimization objective that maximizes the area under the precision-recall curve (AUC PR) value

D- An optimization objective that maximizes the area under the receiver operating characteristic curve (AUC ROC) value

Answer:

C

Question 6

Question Type: **MultipleChoice**

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?

Options:

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

Answer:

A

Question 7

Question Type: **MultipleChoice**

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.

Answer:

A

Question 8

Question Type: **MultipleChoice**

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?

Options:

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.

Answer:

B

Question 9

Question Type: MultipleChoice

You started working on a classification problem with time series data and achieved an area under the receiver operating characteristic curve (AUC ROC) value of 99% for training data after just a few experiments. You haven't explored using any sophisticated algorithms or spent any time on hyperparameter tuning. What should your next step be to identify and fix the problem?

Options:

A- Address the model overfitting by using a less complex algorithm.

B- Address data leakage by applying nested cross-validation during model training.

C- Address data leakage by removing features highly correlated with the target value.

D- Address the model overfitting by tuning the hyperparameters to reduce the AUC ROC value.

Answer:

B

Question 10

Question Type: MultipleChoice

You recently joined an enterprise-scale company that has thousands of datasets. You know that there are accurate descriptions for each table in BigQuery, and you are searching for the proper BigQuery table to use for a model you are building on AI Platform. How should

you find the data that you need?

Options:

A- Use Data Catalog to search the BigQuery datasets by using keywords in the table description.

B- Tag each of your model and version resources on AI Platform with the name of the BigQuery table that was used for training.

C- Maintain a lookup table in BigQuery that maps the table descriptions to the table ID. Query the lookup table to find the correct table ID for the data that you need.

D- Execute a query in BigQuery to retrieve all the existing table names in your project using the INFORMATION_SCHEMA metadata tables that are native to BigQuery. Use the result o find the table that you need.

Answer:

B

Question 11

Question Type: MultipleChoice

You need to build classification workflows over several structured datasets currently stored in BigQuery. Because you will be performing the classification several times, you want to complete the following steps without writing code: exploratory data analysis, feature

selection, model building, training, and hyperparameter tuning and serving. What should you do?

Options:

A- Configure AutoML Tables to perform the classification task

B- Run a BigQuery ML task to perform logistic regression for the classification

C- Use AI Platform Notebooks to run the classification model with pandas library

D- Use AI Platform to run the classification model job configured for hyperparameter tuning

Answer:

B

Explanation:

BigQuery ML supports supervised learning with thelogistic regressionmodel type.

Question 12

Question Type: **MultipleChoice**

You are an ML engineer at a regulated insurance company. You are asked to develop an insurance approval model that accepts or rejects insurance applications from potential customers. What factors should you consider before building the model?

Options:

A- Redaction, reproducibility, and explainability

B- Traceability, reproducibility, and explainability

C- Federated learning, reproducibility, and explainability

D- Differential privacy federated learning, and explainability

Answer:

A