KTest

更に上のクオリティ　更に上のサービス

問題集

Exam : **Professional Machine Learning Engineer**

Title : Google Professional Machine Learning Engineer

Version : DEMO

1.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.

Answer: A

Explanation:

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

2.You work for a global footwear retailer and need to predict when an item will be out of stock based on historical inventory data. Customer behavior is highly dynamic since footwear demand is influenced by many different factors. You want to serve models that are trained on all available data, but track your performance on specific subsets of data before pushing to production.

What is the most streamlined and reliable way to perform this validation?

A. Use the TFX Model Validator tools to specify performance metrics for production readiness

B. Use k-fold cross-validation as a validation strategy to ensure that your model is ready for production.

C. Use the last relevant week of data as a validation set to ensure that your model is performing accurately on current data

D. Use the entire dataset and treat the area under the receiver operating characteristics curve (AUC ROC) as the main metric.

Answer: A

Explanation:

https://www.tensorflow.org/tfx/guide/evaluator

3.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 I AM 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: C

Explanation:

https://cloud.google.com/ai-platform/prediction/docs/resource-labels#overview_of_labels

You can add labels to your AI Platform Prediction jobs, models, and model versions, then use those

labels to organize resources into categories when viewing or monitoring the resources. For example, you can label jobs by team (such as engineering or research) and development phase (prod or test), then filter the jobs based on the team and phase. Labels are also available on operations, but these labels are derived from the resource to which the operation applies. You cannot add or update labels on an operation.

https://cloud.google.com/ai-platform/prediction/docs/sharing-models.

4.During batch training of a neural network, you notice that there is an oscillation in the loss.
How should you adjust your model to ensure that it converges?
A. Increase the size of the training batch
B. Decrease the size of the training batch
C. Increase the learning rate hyperparameter
D. Decrease the learning rate hyperparameter
Answer: D
Explanation:
https://developers.google.com/machine-learning/crash-course/introduction-to-neural-networks/playground-exercises

5.You are building a linear model with over 100 input features, all with values between -1 and 1. You suspect that many features are non-informative. You want to remove the non-informative features from your model while keeping the informative ones in their original form.
Which technique should you use?
A. Use Principal Component Analysis to eliminate the least informative features.
B. Use L1 regularization to reduce the coefficients of uninformative features to 0.
C. After building your model, use Shapley values to determine which features are the most informative.
D. Use an iterative dropout technique to identify which features do not degrade the model when removed.
Answer: B
Explanation:
https://cloud.google.com/ai-platform/prediction/docs/ai-explanations/overview#sampled-shapley