PROFESSIONAL-MACHINE-LEARNING-ENGINEER Q&As

Professional Machine Learning Engineer

Pass Google PROFESSIONAL-MACHINE-LEARNING-ENGINEER Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

https://www.pass2lead.com/professional-machine-learning-engineer.html

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by Google Official Exam Center

QUESTION 1

You are building an ML model to predict trends in the stock market based on a wide range of factors. While exploring the data, you notice that some features have a large range. You want to ensure that the features with the largest magnitude don\\'t overfit the model. What should you do?

A. Standardize the data by transforming it with a logarithmic function.

B. Apply a principal component analysis (PCA) to minimize the effect of any particular feature.

C. Use a binning strategy to replace the magnitude of each feature with the appropriate bin number.

D. Normalize the data by scaling it to have values between 0 and 1.

Correct Answer: A

https://developers.google.com/machine-learning/data-prep/transform/normalization

QUESTION 2

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.

Correct Answer: A

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

QUESTION 3

You are developing an image recognition model using PyTorch based on ResNet50 architecture. Your code is working fine on your local laptop on a small subsample. Your full dataset has 200k labeled images. You want to quickly scale your training workload while minimizing cost. You plan to use 4 V100 GPUs. What should you do?

A. Create a Google Kubernetes Engine cluster with a node pool that has 4 V100 GPUs. Prepare and submit a TFJob operator to this node pool.

B. Create a Vertex AI Workbench user-managed notebooks instance with 4 V100 GPUs, and use it to train your model.

C. Package your code with Setuptools, and use a pre-built container. Train your model with Vertex AI using a custom tier that contains the required GPUs.

D. Configure a Compute Engine VM with all the dependencies that launches the training. Train your model with Vertex AI using a custom tier that contains the required GPUs.

Correct Answer: C

"Vertex AI provides flexible and scalable hardware and secured infrastructure to train PyTorch based deep learning models with pre-built containers and custom containers. (...) use PyTorch ResNet-50 as the example model and train it on ImageNet validation data (50K images) to measure the training performance for different training strategies": https://cloud.google.com/blog/products/ai-machine-learning/efficient-pytorch-training-with-vertex-ai

QUESTION 4

You work for an online retail company that is creating a visual search engine. You have set up an end-to-end ML pipeline on Google Cloud to classify whether an image contains your company\\'s product. Expecting the release of new products in the near future, you configured a retraining functionality in the pipeline so that new data can be fed into your ML models. You also want to use AI Platform\\'s continuous evaluation service to ensure that the models have high accuracy on your test dataset. What should you do?

A. Keep the original test dataset unchanged even if newer products are incorporated into retraining.

B. Extend your test dataset with images of the newer products when they are introduced to retraining.

C. Replace your test dataset with images of the newer products when they are introduced to retraining.

D. Update your test dataset with images of the newer products when your evaluation metrics drop below a pre-decided threshold.

Correct Answer: B

You need to correctly classify newer products, so you need the new training data ==> A is wrong;

You need to keep doing a good job on older dataset, you can\\'t just ignore it ==> C is wrong;

You know when you are introducing new products, there is no need to wait for a drop in preformaces ==> D is wrong;

B is correct

QUESTION 5

You work for a retailer that sells clothes to customers around the world. You have been tasked with ensuring that ML models are built in a secure manner. Specifically, you need to protect sensitive customer data that might be used in the models. You have identified four fields containing sensitive data that are being used by your data science team: AGE, IS_EXISTING_CUSTOMER, LATITUDE_LONGITUDE, and SHIRT_SIZE. What should you do with the data before it is made available to the data science team for training purposes?

A. Tokenize all of the fields using hashed dummy values to replace the real values.

B. Use principal component analysis (PCA) to reduce the four sensitive fields to one PCA vector.

C. Coarsen the data by putting AGE into quantiles and rounding LATITUDE_LONGTTUDE into single precision. The other two fields are already as coarse as possible.

D. Remove all sensitive data fields, and ask the data science team to build their models using non-sensitive data.

Correct Answer: A

https://cloud.google.com/blog/products/identity-security/take-charge-of-your-data-how-tokenization-makes-data-usable-without-sacrificing-privacy