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.leads4pass.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 work for a retail company. You have been tasked with building a model to determine the probability of churn for each customer. You need the predictions to be interpretable so the results can be used to develop marketing campaigns that target at-risk customers. What should you do?

A. Build a random forest regression model in a Vertex AI Workbench notebook instance. Configure the model to generate feature importances after the model is trained.

B. Build an AutoML tabular regression model. Configure the model to generate explanations when it makes predictions.

C. Build a custom TensorFlow neural network by using Vertex AI custom training. Configure the model to generate explanations when it makes predictions.

D. Build a random forest classification model in a Vertex AI Workbench notebook instance. Configure the model to generate feature importances after the model is trained.

Correct Answer: B

QUESTION 2

You are an ML engineer at a mobile gaming company. A data scientist on your team recently trained a TensorFlow model, and you are responsible for deploying this model into a mobile application. You discover that the inference latency of the current model doesn\\'t meet production requirements. You need to reduce the inference time by 50%, and you are willing to accept a small decrease in model accuracy in order to reach the latency requirement. Without training a new model, which model optimization technique for reducing latency should you try first?

A. Weight pruning

B. Dynamic range quantization

C. Model distillation

D. Dimensionality reduction

Correct Answer: B

https://www.tensorflow.org/lite/performance/post_training_quantization#dynamic_range_quantization

QUESTION 3

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?

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

Correct Answer: C

https://stats.stackexchange.com/questions/262616/roc-vs-precision-recall-curves-on-imbalanced-dataset

https://neptune.ai/blog/f1-score-accuracy-roc-auc-pr-auc

QUESTION 4

You work at a leading healthcare firm developing state-of-the-art algorithms for various use cases. You have unstructured textual data with custom labels. You need to extract and classify various medical phrases with these labels. What should you do?

A. Use the Healthcare Natural Language API to extract medical entities

B. Use a BERT-based model to fine-tune a medical entity extraction model

C. Use AutoML Entity Extraction to train a medical entity extraction model

D. Use TensorFlow to build a custom medical entity extraction model

Correct Answer: C

QUESTION 5

Your data science team has requested a system that supports scheduled model retraining, Docker containers, and a service that supports autoscaling and monitoring for online prediction requests. Which platform components should you choose for this system?

A. Vertex AI Pipelines and App Engine

B. Vertex AI Pipelines, Vertex AI Prediction, and Vertex AI Model Monitoring

C. Cloud Composer, BigQuery ML, and Vertex AI Prediction

D. Cloud Composer, Vertex AI Training with custom containers, and App Engine

Correct Answer: B

https://cloud.google.com/vertex-ai/docs/training/containers-overview

QUESTION 6

You need to design a customized deep neural network in Keras that will predict customer purchases based on their purchase history. You want to explore model performance using multiple model architectures, store training data, and be able to compare the evaluation metrics in the same dashboard. What should you do?

A. Create multiple models using AutoML Tables.

B. Automate multiple training runs using Cloud Composer.

C. Run multiple training jobs on AI Platform with similar job names.

D. Create an experiment in Kubeflow Pipelines to organize multiple runs.

Correct Answer: D

https://www.kubeflow.org/docs/about/use-cases/

QUESTION 7

You recently used BigQuery ML to train an AutoML regression model. You shared results with your team and received positive feedback. You need to deploy your model for online prediction as quickly as possible. What should you do?

A. Retrain the model by using BigQuery ML, and specify Vertex AI as the model registry. Deploy the model from Vertex AI Model Registry to a Vertex AI endpoint,

B. Retrain the model by using Vertex AI Deploy the model from Vertex AI Model. Registry to a Vertex AI endpoint.

C. Alter the model by using BigQuery ML, and specify Vertex AI as the model registry. Deploy the model from Vertex AI Model Registry to a Vertex AI endpoint.

D. Export the model from BigQuery ML to Cloud Storage. Import the model into Vertex AI Model Registry. Deploy the model to a Vertex AI endpoint.

Correct Answer: C

QUESTION 8

You work at a gaming startup that has several terabytes of structured data in Cloud Storage. This data includes gameplay time data user metadata and game metadata. You want to build a model that recommends new games to users that

requires the least amount of coding.

What should you do?

A. Load the data in BigQuery Use BigQuery ML to tram an Autoencoder model.

B. Load the data in BigQuery Use BigQuery ML to train a matrix factorization model.

C. Read data to a Vertex AI Workbench notebook Use TensorFlow to train a two-tower model.

D. Read data to a Vertex AI Workbench notebook Use TensorFlow to train a matrix factorization model.

Correct Answer: B

QUESTION 9

You want to train an AutoML model to predict house prices by using a small public dataset stored in BigQuery. You

need to prepare the data and want to use the simplest, most efficient approach. What should you do?

A. Write a query that preprocesses the data by using BigQuery and creates a new table. Create a Vertex AI managed dataset with the new table as the data source.

B. Use Dataflow to preprocess the data. Write the output in TFRecord format to a Cloud Storage bucket.

C. Write a query that preprocesses the data by using BigQuery. Export the query results as CSV files, and use those files to create a Vertex AI managed dataset.

D. Use a Vertex AI Workbench notebook instance to preprocess the data by using the pandas library. Export the data as CSV files, and use those files to create a Vertex AI managed dataset.

Correct Answer: A

QUESTION 10

You work for a small company that has deployed an ML model with autoscaling on Vertex AI to serve online predictions in a production environment. The current model receives about 20 prediction requests per hour with an average response time of one second. You have retrained the same model on a new batch of data, and now you are canary testing it, sending ~10% of production traffic to the new model. During this canary test, you notice that prediction requests for your new model are taking between 30 and 180 seconds to complete. What should you do?

A. Submit a request to raise your project quota to ensure that multiple prediction services can run concurrently.

B. Turn off auto-scaling for the online prediction service of your new model. Use manual scaling with one node always available.

C. Remove your new model from the production environment. Compare the new model and existing model codes to identify the cause of the performance bottleneck.

D. Remove your new model from the production environment. For a short trial period, send all incoming prediction requests to BigQuery. Request batch predictions from your new model, and then use the Data Labeling Service to validate your model\\'s performance before promoting it to production.

Correct Answer: B

QUESTION 11

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 then TFX ModelValidator 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.

Correct Answer: C

https://cloud.google.com/learn/what-is-time-series

QUESTION 12

You are training a custom language model for your company using a large dataset. You plan to use the Reduction Server strategy on Vertex AI. You need to configure the worker pools of the distributed training job. What should you do?

A. Configure the machines of the first two worker pools to have GPUs, and to use a container image where your training code runs. Configure the third worker pool to have GPUs, and use the reductionserver container image.

B. Configure the machines of the first two worker pools to have GPUs and to use a container image where your training code runs. Configure the third worker pool to use the reductionserver container image without accelerators, and choose a machine type that prioritizes bandwidth.

C. Configure the machines of the first two worker pools to have TPUs and to use a container image where your training code runs. Configure the third worker pool without accelerators, and use the reductionserver container image without accelerators, and choose a machine type that prioritizes bandwidth.

D. Configure the machines of the first two pools to have TPUs, and to use a container image where your training code runs. Configure the third pool to have TPUs, and use the reductionserver container image.

Correct Answer: B

QUESTION 13

You are an ML engineer at a retail company. You have built a model that predicts a coupon to offer an ecommerce customer at checkout based on the items in their cart. When a customer goes to checkout, your serving pipeline, which is hosted on Google Cloud, joins the customer\\'s existing cart with a row in a BigQuery table that contains the customers\\' historic purchase behavior and uses that as the model\\'s input. The web team is reporting that your model is returning predictions too slowly to load the coupon offer with the rest of the web page. How should you speed up your model\\'s predictions?

A. Attach an NVIDIA P100 GPU to your deployed model\\'s instance.

B. Use a low latency database for the customers' historic purchase behavior.

C. Deploy your model to more instances behind a load balancer to distribute traffic.

D. Create a materialized view in BigQuery with the necessary data for predictions.

Correct Answer: D

QUESTION 14

You need to analyze user activity data from your company\\'s mobile applications. Your team will use BigQuery for data analysis, transformation, and experimentation with ML algorithms. You need to ensure real-time ingestion of the user activity data into BigQuery. What should you do?

A. Configure Pub/Sub to stream the data into BigQuery.

B. Run an Apache Spark streaming job on Dataproc to ingest the data into BigQuery.

C. Run a Dataflow streaming job to ingest the data into BigQuery.

D. Configure Pub/Sub and a Dataflow streaming job to ingest the data into BigQuery,

Correct Answer: A

Previously Google pattern was Pub/Sub -> Dataflow -> BQ but now it looks as there is new Pub/Sub -> BQ
https://cloud.google.com/blog/products/data-analytics/pub-sub-launches-direct-path-to-bigquery-for-streaming-analytics

QUESTION 15

You have a custom job that runs on Vertex AI on a weekly basis. The job is implemented using a proprietary ML workflow that produces the datasets, models, and custom artifacts, and sends them to a Cloud Storage bucket. Many different versions of the datasets and models were created. Due to compliance requirements, your company needs to track which model was used for making a particular prediction, and needs access to the artifacts for each model. How should you configure your workflows to meet these requirements?

A. Use the Vertex AI Metadata API inside the custom job to create context, execution, and artifacts for each model, and use events to link them together.

B. Create a Vertex AI experiment, and enable autologging inside the custom job.

C. Configure a TensorFlow Extended (TFX) ML Metadata database, and use the ML Metadata API.

D. Register each model in Vertex AI Model Registry, and use model labels to store the related dataset and model information.

Correct Answer: A