Google Cloud

Machine Learning in the Enterprise

Quiz Question Answers

Module 1: Understanding the ML Enterprise Workflow

Question 1
Which two activities are involved in ML development?

A: Training formalization and training operationalization
Feedback: This answer is partially correct, please review the module again.

B: Version control and training operationalization
Feedback: This answer is partially correct, please review the module again.

C: Experimentation and version control
Feedback: This answer is partially correct, please review the module again.

*D: Experimentation and training operationalization
Feedback: This answer is correct.

Question 2
Which process covers algorithm selection, model training, hyperparameter tuning, and model evaluation in the Experimentation and Prototyping activity?

A: Data exploration
Feedback: This answer is incorrect, please review the module again.

*B: Model prototyping
Feedback: This answer is correct.

C: Model validation
Feedback: This answer is incorrect, please review the module again.

D: Feature engineering
Feedback: This answer is incorrect, please review the module again.

Question 3
What is the correct process that data scientists use to develop the models on an experimentation platform?

A: Problem definition > Data exploration > Data selection > Feature engineering > Model prototyping > Model validation
Feedback: This answer is incorrect, please review the module again.

*B: Problem definition > Data selection > Data exploration > Feature engineering > Model prototyping > Model validation
Feedback: This answer is correct.

C: Problem definition > Data selection > Data exploration > Model prototyping > Feature engineering > Model validation
Feedback: This answer is incorrect, please review the module again.

D: Problem definition > Data selection > Data exploration > Model prototyping > Model validation > Feature engineering
Feedback: This answer is incorrect, please review the module again.

Question 4
If the model needs to be repeatedly retrained in the future, an automated training pipeline is also developed. Which task do we use for this?

A: Training formalization
Feedback: This answer is incorrect, please review the module again.

B: Training implementation
Feedback: This answer is incorrect, please review the module again.

*C: Training operationalization
Feedback: This answer is correct.

D: Experimentation & prototyping
Feedback: This answer is incorrect, please review the module again.

Module 2: Data in the Enterprise

Question 1
Which of the following is correct for Online serving?

*A: Online serving is for low-latency data retrieval of small batches of data for real-time processing.
Feedback: This answer is correct.

B: Online serving is for high-latency data retrieval of small batches of data for real-time processing.
Feedback: This answer is incorrect, please review the module again.

C: Online serving is for high throughput and serving large volumes of data for offline processing.
Feedback: This answer is incorrect, please review the module again.

D: Online serving is for low throughput and serving large volumes of data for offline processing.
Feedback: This answer is incorrect, please review the module again.

Question 2
What does the Aggregation Values contain in any feature?

A: The min, median, and Std.dev values for each features
Feedback: This answer is incorrect, please review the module again.

B: The min, zeros, and Std.dev values for each features
Feedback: This answer is incorrect, please review the module again.

*C: The min, median, and max values for each features
Feedback: This answer is correct.

D: The Count, median, and max values for each features
Feedback: This answer is incorrect, please review the module again.

Question 3
Which of the following is not a part of Google's enterprise data management and governance tool?

A: Feature Store
Feedback: This answer is incorrect, please review the module again.

B: Data Catalog
Feedback: This answer is incorrect, please review the module again.

C: Dataplex
Feedback: This answer is incorrect, please review the module again.

*D: Analytics Catalog
Feedback: This answer is correct.

Question 4
Which of the following statements is not a feature of Analytics Hub?

A: Analytics Hub efficiently and securely exchanges data analytics assets across organizations to address challenges of data reliability and cost.
Feedback: This answer is incorrect, please review the module again.

B: You can create and access a curated library of internal and external assets, including unique datasets like Google Trends, backed by the power of BigQuery.
Feedback: This answer is incorrect, please review the module again.

*C: Analytics Hub requires batch data pipelines that extract data from databases, store it in flat files, and transmit them to the consumer where they are ingested into another database.
Feedback: This answer is correct.

D: There are three roles in Analytics Hub - A Data Publisher, Exchange Administrator, and a Data Subscriber.
Feedback: This answer is incorrect, please review the module again.

Question 5
Which Data processing option can be used for transforming large unstructured data in Google Cloud?

A: Hadoop proc
Feedback: This answer is incorrect, please review the module again.

*B: Dataflow
Feedback: This answer is correct.

C: Beam proc
Feedback: This answer is incorrect, please review the module again.

D: Apache prep
Feedback: This answer is incorrect, please review the module again.

Module 3: Science of Machine Learning and Custom Training

Question 1

Model complexity often refers to the number of features or terms included in a given predictive model. What happens when the complexity of the model increases?

A: Model is more likely to overfit.
Feedback: This answer is partially correct, please review the module again.

B: Model will not figure out general relationships in the data.
Feedback: This answer is partially correct, please review the module again.

C: Model performance on a test set is going to be poor.
Feedback: This answer is partially correct, please review the module again.

*D: All of the options are correct.
Feedback: This answer is correct.

Question 2

The learning rate is a hyperparameter that controls how much to change the model in response to the estimated error each time the model weights are updated. Choosing the learning rate is challenging. What can happen if the value is too small?

*A: Training may take a long time.
Feedback: This answer is correct.

B: If the learning rate value is too small, then the model will diverge.
Feedback: This answer is incorrect, please review the module again.

C: The model will train more quickly.
Feedback: This answer is incorrect, please review the module again.

D: Smaller learning rates require less training epochs given the smaller changes made to the weights each update.
Feedback: This answer is incorrect, please review the module again.

Question 3

The learning rate is a hyperparameter that controls how much to change the model in response to the estimated error each time the model weights are updated. Choosing the learning rate is challenging. What can happen if the value is too large?

A: Training may take a long time.
Feedback: This answer is incorrect, please review the module again.

B: If the learning rate value is too large, then the model will converge.
Feedback: This answer is incorrect, please review the module again.

C: The model will not train..
Feedback: This answer is incorrect, please review the module again.

*D: A large learning rate value may result in the model learning a sub-optimal set of weights too fast or an unstable training process.
Feedback: This answer is correct.

Question 4
Which of the following is true?

*A: Larger batch sizes require smaller learning rates.
Feedback: This answer is correct.

B: Smaller batch sizes require smaller learning rates.
Feedback: This answer is incorrect, please review the module again.

C: Larger batch sizes require larger learning rates.
Feedback: This answer is incorrect, please review the module again.

D: Smaller batch sizes require larger learning rates.
Feedback: This answer is incorrect, please review the module again.

Question 5
The learning rate is a configurable hyperparameter used in the training of neural networks that has a small positive value, often in the range between _____

A: 1.0 and 3.0.
Feedback: This answer is incorrect, please review the module again.

*B: 0.0 and 1.0.
Feedback: This answer is correct.

C: > 0.0 and < 1.00.
Feedback: This answer is incorrect, please review the module again.

D: < 0.0 and > 1.00.
Feedback: This answer is incorrect, please review the module again.

Question 6
What is "data parallelism" in distributed training?

A: Run the same model & computation on every device, but train each of them using the same training samples.
Feedback: This answer is incorrect, please review the module again.

B: Run different models & computation on every device, but train each of them using only one training sample.
Feedback: This answer is incorrect, please review the module again.

*C: Run the same model & computation on every device, but train each of them using different training samples.
Feedback: This answer is correct.

D: Run different models & computation on a single device, but train each of them using different training samples.
Feedback: This answer is incorrect, please review the module again.

Module 4: Vertex Vizier Hyperparameter Tuning

Question 1
Bayesian optimization takes into account past evaluations when choosing the hyperparameter set to evaluate next. By choosing its parameter combinations in an informed way, it enables itself to focus on those areas of the parameter space that it believes will bring the most promising validation scores. Therefore it _____.

A: enables itself to focus on those areas of the parameter space that it believes will bring the most promising validation scores.
Feedback: This answer is partially correct, please review the module again.

B: requires less iterations to get to the optimal set of hyperparameter values.
Feedback: This answer is partially correct, please review the module again.

C: limits the number of times a model needs to be trained for validation.
Feedback: This answer is partially correct, please review the module again.

*D: All of the options are correct.

Feedback: This answer is correct.

Question 2
Which of the following is a black-box optimization service?

A: Manual Search
Feedback: This answer is incorrect, please review the module again.

*B: Vertex Vizier
Feedback: This answer is correct.

C: AutoML
Feedback: This answer is incorrect, please review the module again.

D: Early stopping
Feedback: This answer is incorrect, please review the module again.

Question 3
Which of the following algorithms is useful, if you want to specify a quantity of trials that is greater than the number of points in the feasible space?

*A: Grid Search
Feedback: This answer is correct.

B: Bayesian Optimization
Feedback: This answer is incorrect, please review the module again.

C: Random Search
Feedback: This answer is incorrect, please review the module again.

D: Manual Search
Feedback: This answer is incorrect, please review the module again.

Question 4
Black box optimization algorithms find the best operating parameters for any system whose _____?

A: iterations to get to the optimal set of hyperparameter values are less.
Feedback: This answer is incorrect, please review the module again.

B: execution time is less.

Feedback: This answer is incorrect, please review the module again.

*C: performance can be measured as a function of adjustable parameters.
Feedback: This answer is correct.

D: number of iterations is limited to train a model for validation.
Feedback: This answer is incorrect, please review the module again.

Question 5
Which of the following can make a huge difference in model quality?

A: Increasing the learning rate.
Feedback: This answer is incorrect, please review the module again.

*B: Setting hyperparameters to their optimal values for a given dataset.
Feedback: This answer is correct.

C: Decreasing the number of epochs.
Feedback: This answer is incorrect, please review the module again.

D: Increasing the training time.
Feedback: This answer is incorrect, please review the module again.

Module 5: Prediction and Model Monitoring Using Vertex AI

Question 1
Which statements are correct for serving predictions using Pre-built containers?

A: Vertex AI provides Docker container images that you run as pre-built containers for serving predictions.
Feedback: This answer is incorrect, please review the module again.

B: Pre-built containers provide HTTP prediction servers that you can use to serve prediction using minimal configurations.
Feedback: This answer is incorrect, please review the module again.

C: Pre-built containers are organized by Machine learning framework and framework version.
Feedback: This answer is incorrect, please review the module again.

*D: All of the options are correct.

Feedback: This answer is correct.

Question 2
Which statement is correct regarding the maximum size for a CSV file during batch prediction?

A: The data source file must be no larger than 100 GB.
Feedback: This answer is incorrect, please review the module again.

*B: Each data source file must not be larger than 10 GB. You can include multiple files, up to a maximum amount of 100 GB.
Feedback: This answer is correct.

C: The data source file must be no larger than 50 GB. You can not include multiple files.
Feedback: This answer is incorrect, please review the module again.

D: Each data source file must include multiple files, up to a maximum amount of 50 GB.
Feedback: This answer is incorrect, please review the module again.

Question 3
What should be done if the source table is in a different project?

*A: You should provide the BigQuery Data Editor role to the Vertex AI service account in that project.
Feedback: This answer is correct.

B: You should provide the BigQuery Data Viewer role to the Vertex AI service account in that project.
Feedback: This answer is incorrect, please review the module again.

C: You should provide the BigQuery Data Editor role to the Vertex AI service account in your project.
Feedback: This answer is incorrect, please review the module again.

D: You should provide the BigQuery Data Viewer role to the Vertex AI service account in your project.
Feedback: This answer is incorrect, please review the module again.

Question 4
Which of the following statements is invalid for a data source file in batch prediction?

A: The first line of the data source CSV file must contain the name of the columns.

Feedback: This answer is incorrect, please review the module again.

B: If the Cloud Storage bucket is in a different project than where you use Vertex AI, you must provide the Storage Object Creator role to the Vertex AI service account in that project.
Feedback: This answer is incorrect, please review the module again.

C: BigQuery data source tables must be no larger than 100 GB.
Feedback: This answer is incorrect, please review the module again.

*D: You must use a regional BigQuery dataset.
Feedback: This answer is correct.

Question 5
What are the features of Vertex AI model monitoring?

A: Drift in data quality
Feedback: This answer is partially correct, please review the module again.

B: Skew in training vs. serving data
Feedback: This answer is partially correct, please review the module again.

C: Feature Attribution and UI visualizations
Feedback: This answer is partially correct, please review the module again.

*D: All of the options are correct.
Feedback: This answer is correct.

Question 6
For which, the baseline is the statistical distribution of the feature's values seen in production in the recent past.

A: Categorical features
Feedback: This answer is incorrect, please review the module again.

B: Numerical features
Feedback: This answer is incorrect, please review the module again.

*C: Drift detection
Feedback: This answer is correct.

D: Skew detection
Feedback: This answer is incorrect, please review the module again.

Module 6: Vertex AI Pipelines

Question 1
Which package is used to define and interact with pipelines and components?

*A: kfp.dsl package
Feedback: This answer is correct.

B: kfp.compiler
Feedback: This answer is incorrect, please review the module again.

C: kfp.components
Feedback: This answer is incorrect, please review the module again.

D: kfp.containers
Feedback: This answer is incorrect, please review the module again.

Question 2
How can you define the pipeline's workflow as a graph?

A: By using different inputs for each component.
Feedback: This answer is incorrect, please review the module again.

B: Use the previous pipeline's output as an input for the current pipeline.
Feedback: This answer is incorrect, please review the module again.

*C: By using the outputs of a component as an input of another component
Feedback: This answer is correct.

D: By using predictive input for each component.
Feedback: This answer is incorrect, please review the module again.

Question 3
What can you use to compile the pipeline?

A: compiler.Compiler
Feedback: This answer is incorrect, please review the module again.

B: kfp.v2.compiler
Feedback: This answer is incorrect, please review the module again.

C: kfp.Compiler
Feedback: This answer is incorrect, please review the module again.

*D: kfp.v2.compiler.Compiler
Feedback: This answer is correct.

Question 4
What can you use to create a pipeline run on Vertex AI Pipelines?

*A: Vertex AI python client
Feedback: This answer is correct.

B: Pipeline root path
Feedback: This answer is incorrect, please review the module again.

C: kfp.v2.compiler.Compiler
Feedback: This answer is incorrect, please review the module again.

D: Service account
Feedback: This answer is incorrect, please review the module again.