Google Cloud

Launching into Machine Learning

Quiz Question Answers

Module 1: Improve Data Quality and Exploratory Data Analysis

Question 1

Which of the following are categories of data quality tools?

A: Cleaning tools
Feedback: This answer is partially correct, please review the module again.

B: Monitoring tools
Feedback: This answer is partially correct, please review the module again.

*C: Both A and B
Feedback: This answer is correct.

D: None of the above
Feedback: This is incorrect, please review the module again

Question 2

What are the features of low data quality?

A: Unreliable info
Feedback: This answer is partially correct, please review the module again.

B: Incomplete data
Feedback: This answer is partially correct, please review the module again.

C: Duplicated data
Feedback: This answer is partially correct, please review the module again.

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

Question 3
What are the objectives of exploratory data analysis?

A: Check for missing data and other mistakes.
Feedback: This answer is partially correct, please review the module again.

B: Gain maximum insight into the data set and its underlying structure.
Feedback: This answer is partially correct, please review the module again.

C: Uncover a parsimonious model, one which explains the data with a minimum number of predictor variables.
Feedback: This answer is partially correct, please review the module again.

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

Question 4
Exploratory Data Analysis is majorly performed using the following methods:

A: Univariate
Feedback: This answer is partially correct, please review the module again.

B: Bivariate
Feedback: This answer is partially correct, please review the module again.

*C: Both A and B
Feedback: This answer is correct.

D: None of the above
Feedback: This answer is incorrect, please review the module again.

Question 5
Which of the following is not a component of Exploratory Data Analysis?

A: Accounting and Summarizing
Feedback: This answer is incorrect, please review the module again.

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

C: Statistical Analysis and Clustering

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

*D: Hyperparameter tuning
Feedback: This answer is correct.

Module 2: ML in Practice

Question 1
Which of the following machine learning models have labels, or in other words, the correct answers to whatever it is that we want to learn to predict?

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

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

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

D: None of the above.
Feedback: This answer is incorrect, please review the module again.

Question 2
Which model would you use if your problem required a discrete number of values or classes?

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

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

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

*D: Classification Model
Feedback: This answer is correct.

Question 3

To predict the continuous value of our label, which of the following algorithms is used?

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

*B: Regression
Feedback: This answer is correct.

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

D: None of the above
Feedback: This answer is incorrect, please review the module again.

Question 4
What is the most essential metric a regression model uses?

*A: Mean squared error as their loss function
Feedback: This answer is correct.

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

C: Both a & b
Feedback: This answer is incorrect, please review the module again.

D: None of the above
Feedback: This answer is incorrect, please review the module again.

Question 5
Why is regularization important in logistic regression?

*A: Avoids overfitting.
Feedback: This answer is correct.

B: Keeps training time down by regulating the time allowed.
Feedback: This answer is incorrect, please review the module again.

C: Finds errors in the algorithm.
Feedback: This answer is incorrect, please review the module again.

D: Encourages the use of large weights.

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

Module 3: Training AutoML models using Vertex AI

Question 1
What is the main benefit of using an automated Machine Learning workflow?

A: It makes the model perform better.
Feedback: This answer is incorrect, please review the module again.

B: It makes the model run faster.
Feedback: This answer is incorrect, please review the module again.

C: It deploys the model into production.
Feedback: This answer is incorrect, please review the module again.

D: It reduces the time it takes to develop trained models and assess their performance
Feedback: This answer is correct

Question 2
For a user who can use SQL, has little Machine Learning experience and wants a 'Low-Code' solution, which Machine Learning framework should they use?

A: Scikit-Learn
Feedback: This answer is incorrect, please review the module again.

*B: BigQuery ML
Feedback: This answer is correct

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

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

Question 3
If a dataset is presented in a Comma Separated Values (CSV) file, which is the correct data type to choose in Vertex AI?

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

*B: Tabular
Feedback: This answer is correct

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

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

Question 4
If the business case is to predict fraud detection, which is the correct Objective to choose in Vertex AI?

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

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

*C: Regression/Classification
Feedback: This answer is correct

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

Question 5
Which of the following are stages of the Machine Learning workflow that can be managed with Vertex AI?

A: Create a dataset and upload data.
Feedback: This answer is partially correct, please review the module again.

B: Train an ML model on your data
Feedback: This answer is partially correct, please review the module again.

C: Deploy your trained model to an endpoint for serving predictions.
Feedback: This answer is partially correct, please review the module again.

*D: All of the above

Feedback: This answer is correct

Question 6
What is the default setting in AutoML Tables for the data split in model evaluation?

A: 70% Training, 20% Validation, 10% Testing
Feedback: This answer is incorrect, please review the module again.

B: 80% Training, 15% Validation, 5% Testing
Feedback: This answer is incorrect, please review the module again.

C: 80% Training, 5% Validation, 15% Testing
Feedback: This answer is incorrect, please review the module again.

*D: 80% Training 10% Validation, 10% Testing
Feedback: This answer is correct

Question 7
MAE, MAPE, RMSE, RMSLE and R^2 are all available as test examples in the Evaluate section of Vertex AI and are common examples of what type of metric?

A: Forecasting Regression Metrics
Feedback: This answer is incorrect, please review the module again.

*B: Linear Regression Metrics
Feedback: This answer is correct

C: Clustering Regression Metrics
Feedback: This answer is incorrect, please review the module again.

D: Decision Trees Progression Metrics
Feedback: This answer is incorrect, please review the module again.

Question 8
What does the Feature Importance attribution in Vertex AI display?

*A: How much each feature impacts the model, expressed as a percentage
Feedback: This answer is correct

B: How much each feature impacts the model, expressed as a ratio
Feedback: This answer is incorrect, please review the module again.

C: How much each feature impacts the model, expressed as a ranked list
Feedback: This answer is incorrect, please review the module again.

D: How much each feature impacts the model, expressed as a decimal
Feedback: This answer is incorrect, please review the module again.

Question 9
Which of the following metrics can be used to find a suitable balance between precision and recall in a model?

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

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

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

*D: F1 Score
Feedback: This answer is correct

Module 4: BigQuery ML

Question 1
Which of the following are advantages of BigQuery ML when compared to Python based ML frameworks?

A: BigQuery ML custom models can be created without the use of multiple tools
Feedback: This answer is partially correct, please review the module again.

B: BigQuery ML automates multiple steps in the ML workflow
Feedback: This answer is partially correct, please review the module again.

C: Moving and formatting large amounts of data takes longer with Python based models compared to model training in BigQuery
Feedback: This answer is partially correct, please review the module again.

D: All of the above
Feedback: This answer is correct

Question 2
Which of these BigQuery supported classification models is most relevant for predicting binary results, such as True/False?

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

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

*C: Logistic Regression
Feedback: This answer is correct

D: DNN Classifier (TensorFlow)
Feedback: This answer is incorrect, please review the module again.

Question 3
For Classification or Regression problems with decision trees, which of the following models is most relevant?

A: Wide and Deep NNs
Feedback: This answer is incorrect, please review the module again.

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

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

*D: XGBoost
Feedback: This answer is correct

Question 4
Where labels are not available, for example where customer segmentation is required, which of the following BigQuery supported models is useful?

A: Time Series Forecasting
Feedback: This answer is incorrect, please review the module again.

B: Recommendation: Matrix Factorization
Feedback: This answer is incorrect, please review the module again.

*C: K-Means Clustering
Feedback: This answer is correct

D: Time Series Anomaly Detection
Feedback: This answer is incorrect, please review the module again.

Question 5
What are the 3 key steps for creating a Recommendation System with BigQuery ML?

*A: Prepare training data in BigQuery, train a recommendation system with BigQuery ML, use the predicted recommendations in production
Feedback: This answer is correct

B: Import training data to BigQuery, train a recommendation system with BigQuery ML, tune the hyperparameters
Feedback: This answer is incorrect, please review the module again.

C: Prepare training data in BigQuery, select a recommendation system from BigQuery ML, deploy and test the model
Feedback: This answer is incorrect, please review the module again.

D: Prepare training data in BigQuery, specify the model options in BigQuery ML, export the predictions to Google Analytics
Feedback: This answer is incorrect, please review the module again.

Module 5: Optimization

Question 1
For the formula used to model the relationship i.e. y = mx + b, what does ‘m’ stand for?

*A: It captures the amount of change we've observed in our label in response to a small change in our feature.
Feedback: This answer is correct.

B: It refers to a bias term which can be used for regression.
Feedback: This answer is incorrect, please review the module again.

C: It refers to a bias term which can be used for regression and it captures the amount of change we've observed in our label in response to a small change in our feature.

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

D: None of the options are correct.
Feedback: This answer is incorrect, please review the module again.

Question 2
What are the basic steps in an ML workflow (or process)?

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

B: Check for anomalies, missing data and clean the data
Feedback: This answer is partially correct, please review the module again.

C: Perform statistical analysis and initial visualization
Feedback: This answer is partially correct, please review the module again.

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

Question 3
Which of the following loss functions is used for classification problems?

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

*B: Cross entropy
Feedback: This answer is correct.

C: Both MSE & Cross entropy
Feedback: This answer is incorrect, please review the module again.

D: None of the options are correct.
Feedback: This answer is incorrect, please review the module again.

Question 4

Which of the following gradient descent methods is used to compute the entire dataset?

*A: Batch gradient descent
Feedback: This answer is correct.

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

C: Mini-batch gradient descent
Feedback: This answer is incorrect, please review the module again.

D: None of the options are correct.
Feedback: This answer is incorrect, please review the module again.

Question 5

Which of the following are benefits of Performance metrics over loss functions?

A: Performance metrics are easier to understand.
Feedback: This answer is partially correct, please review the module again.

B: Performance metrics are directly connected to business goals.
Feedback: This answer is partially correct, please review the module again.

*C: Performance metrics are easier to understand and are directly connected to business goals.
Feedback: This answer is correct.

D: None of the options are correct.
Feedback: This answer is incorrect, please review the module again.

Module 6: Generalization and Sampling

Question 1

Which is the best way to assess the quality of a model?

*A: Observing how well a model performs against a new dataset that it hasn't seen before.
Feedback: This answer is correct.

B: Observing how well a model performs against an existing known dataset.
Feedback: This answer is incorrect, please review the module again.

C: Observing how well a model performs against a new dataset that it hasn't seen before and observing how well a model performs against an existing known dataset.

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

D: None of the options are correct.
Feedback: This answer is incorrect, please review the module again.

Question 2

How do you decide when to stop training a model?

*A: When your loss metrics start to increase
Feedback: This answer is correct.

B: When your loss metrics start to decrease
Feedback: This answer is incorrect, please review the module again.

C: When your loss metrics start to both increase and decrease
Feedback: This answer is incorrect, please review the module again.

D: None of the options are correct
Feedback: This answer is incorrect, please review the module again.

Question 3

Which of the following actions can you perform on your model when it is trained and validated?

*A: You can write it once, and only once, against the independent test dataset.
Feedback: This answer is correct.

B: You can write it once, and only once against the dependent test dataset.
Feedback: This answer is incorrect, please review the module again.

C: You can write it multiple times against the independent test dataset.
Feedback: This answer is incorrect, please review the module again.

D: You can write it multiple times against the dependent test dataset.
Feedback: This answer is incorrect, please review the module again.

Question 4

Which of the following allows you to create repeatable samples of your data?

*A: Use the last few digits of a hash function on the field that you're using to split or bucketize your data.
Feedback: This answer is correct.

B: Use the first few digits of a hash function on the field that you're using to split or bucketize your data.
Feedback: This answer is incorrect, please review the module again.

C: Use the first few digits or the last few digits of a hash function on the field that you're using to split or bucketize your data.
Feedback: This answer is incorrect, please review the module again.

D: None of the options are correct.
Feedback: This answer is incorrect, please review the module again.

Question 5

Which of the following allows you to split the dataset based upon a field in your data?

*A: FARM_FINGERPRINT, an open-source hashing algorithm that is implemented in BigQuery SQL.
Feedback: This answer is correct.

B: BUCKETIZE, an open-source hashing algorithm that is implemented in BigQuery SQL.
Feedback: This answer is incorrect, please review the module again.

C: ML_FEATURE FINGERPRINT, an open-source hashing algorithm that is implemented in BigQuery SQL.
Feedback: This answer is incorrect, please review the module again.

D: None of the options are correct.
Feedback: This answer is incorrect, please review the module again.