[Full-Version 2023 Updated Google Study Guide Professional-Machine-Learning-Engineer Dumps Questions [Q49-Q67

[Full-Version] 2023 Updated Google Study Guide Professional-Machine-Learning-Engineer Dumps Questions
Newest Professional-Machine-Learning-Engineer Exam Dumps Achieve Success in Actual
Professional-Machine-Learning-Engineer Exam

Exam Details

The Google Professional Machine Learning Engineer exam is two hours long. The candidates can expect multiple-choice as well as multiple-select questions in their delivery of the certification test. The exam is currently given to the learners in the English language. To register for and schedule it, you need to pay $200 (plus applicable taxes). While registering for the test, the potential applicants will be offered to select the convenient mode of exam delivery: an online proctored session from a remote location or an in-person proctored session at the nearest testing center.

The Google Professional Machine Learning Engineer Certification Exam is a comprehensive test that validates the expertise of individuals in the field of machine learning. The certification exam is designed to test the individual's ability to design, build, and deploy scalable machine learning models using Google Cloud Platform. Individuals who pass the exam will receive a certificate that is recognized by Google Cloud Platform and can be used to advance one's career in the field of machine learning.

QUESTION 49

You are an ML engineer at a large grocery retailer with stores in multiple regions. You have been asked to create an inventory prediction model. Your models features include region, location, historical demand, and seasonal popularity. You want the algorithm to learn from new inventory data on a daily basis. Which algorithms should you use to build the model?
* Classification
* Reinforcement Learning
* Recurrent Neural Networks (RNN)
* Convolutional Neural Networks (CNN)

QUESTION 50

You work for a toy manufacturer that has been experiencing a large increase in demand. You need to build an ML model to reduce the amount of time spent by quality control inspectors checking for product defects. Faster defect detection is a priority. The factory does not have reliable Wi-Fi. Your company wants to implement the new ML model as soon as possible. Which model should you use?
* AutoML Vision model
* AutoML Vision Edge mobile-versatile-1 model
* AutoML Vision Edge mobile-low-latency-1 model
* AutoML Vision Edge mobile-high-accuracy-1 model

QUESTION 51

A data scientist needs to identify fraudulent user accounts for a company's ecommerce platform. The company wants the ability to determine if a newly created account is associated with a previously known fraudulent user.

The data scientist is using AWS Glue to cleanse the company’s application logs during ingestion.

Which strategy will allow the data scientist to identify fraudulent accounts?
* Execute the built-in FindDuplicates Amazon Athena query.
* Create a FindMatches machine learning transform in AWS Glue.
* Create an AWS Glue crawler to infer duplicate accounts in the source data.
* Search for duplicate accounts in the AWS Glue Data Catalog.
Explanation/Reference: https://docs.aws.amazon.com/glue/latest/dg/machine-learning.html

QUESTION 52

A Machine Learning Specialist built an image classification deep learning model. However, the Specialist ran into an overfitting problem in which the training and testing accuracies were 99% and 75%, respectively.

How should the Specialist address this issue and what is the reason behind it?
* The learning rate should be increased because the optimization process was trapped at a local minimum.
* The dropout rate at the flatten layer should be increased because the model is not generalized enough.
* The dimensionality of dense layer next to the flatten layer should be increased because the model is not complex enough.
* The epoch number should be increased because the optimization process was terminated before it reached the global minimum.
Explanation/Reference: https://www.tensorflow.org/tutorials/keras/overfit_and_underfit

QUESTION 53

You are an ML engineer at a bank. You have developed a binary classification model using AutoML Tables to predict whether a customer will make loan payments on time. The output is used to approve or reject loan requests. One customer’s loan request has been rejected by your model, and the bank’s risks department is asking you to provide the reasons that contributed to the model’s decision. What should you do?
* Use local feature importance from the predictions.
* Use the correlation with target values in the data summary page.
* Use the feature importance percentages in the model evaluation page.
* Vary features independently to identify the threshold per feature that changes the classification.

QUESTION 54

You trained a text classification model. You have the following SignatureDefs:

```
signature_def['serving_default']:
  The given SavedModel SignatureDef contains the following input(s):
    inputs['text'] tensor_info:
        dtype: DT_STRING
        shape: (-1, 2)
        name: serving_default_text:0
  The given SavedModel SignatureDef contains the following output(s):
    outputs['Softmax'] tensor_info:
        dtype: DT_FLOAT
        shape: (-1, 2)
        name: StatefulPartitionedCall:0
  Method name is: tensorflow/serving/predict
```

You started a TensorFlow-serving component server and tried to send an HTTP request to get a prediction using:

```
headers = {"content-type": "application/json"}
json_response = requests.post('http://localhost:8501/v1/models/text_model:predict', data=d
headers=headers)
```

What is the correct way to write the predict request?

* data = json.dumps({"signature_name": "serving_default' "instances":
[fab', 'be1, 'cd']]})

* data = json dumps({"signature_name": "serving_default"! "instances":
[['a', 'b', "c", 'd', 'e', 'f']]})

* data = json.dumps({"signature_name": "serving_default, "instances": [['a',
'b 'c'1, [d 'e T]]})

* data = json dumps({"signature_name": f,serving_default", "instances":
[['a', 'b'], [c 'd'], ['e T]]})

https://stackoverflow.com/questions/37956197/what-is-the-negative-index-in-shape-arrays-used-for-tensorflow

QUESTION 55

You were asked to investigate failures of a production line component based on sensor readings. After receiving the dataset, you discover that less than 1% of the readings are positive examples representing failure incidents. You have tried to train several classification models, but none of them converge. How should you resolve the class imbalance problem?
* Use the class distribution to generate 10% positive examples
* Use a convolutional neural network with max pooling and softmax activation
* Downsample the data with upweighting to create a sample with 10% positive examples
* Remove negative examples until the numbers of positive and negative examples are equal
https://developers.google.com/machine-learning/data-prep/construct/sampling-splitting/imbalanced-data#downsampling-and-upweig
hting

https://developers.google.com/machine-learning/data-prep/construct/sampling-splitting/imbalanced-data

QUESTION 56

Given the following confusion matrix for a movie classification model, what is the true class frequency for Romance and the predicted class frequency for Adventure?
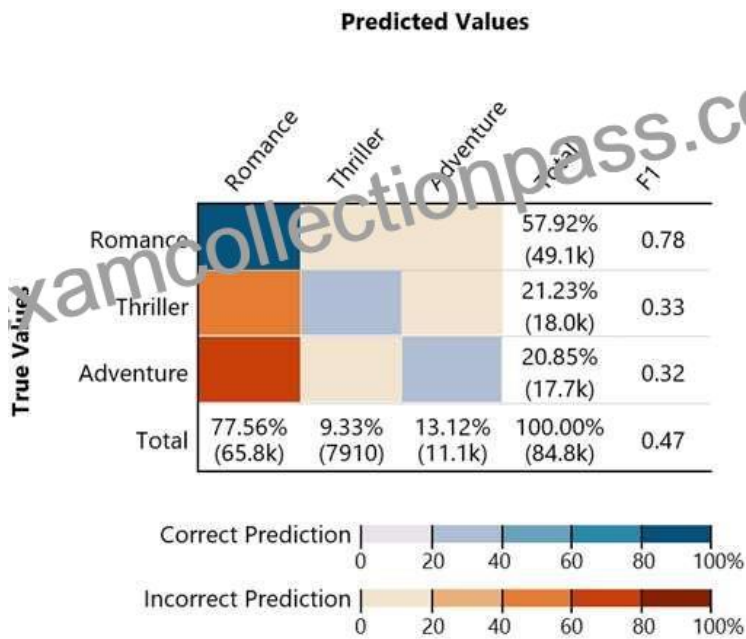
* The true class frequency for Romance is 77.56% and the predicted class frequency for Adventure is

20.85%
* The true class frequency for Romance is 57.92% and the predicted class frequency for Adventure is

13.12%
* The true class frequency for Romance is 0.78 and the predicted class frequency for Adventure is (0.47-

0.32)
* The true class frequency for Romance is 77.56% * 0.78 and the predicted class frequency for Adventure is

20.85%*0.32

QUESTION 57

Your organization’s call center has asked you to develop a model that analyzes customer sentiments in each call. The call center receives over one million calls daily, and data is stored in Cloud Storage. The data collected must not leave the region in which the call originated, and no Personally Identifiable Information (Pll) can be stored or analyzed. The data science team has a third-party tool for visualization and access which requires a SQL ANSI-2011 compliant interface. You need to select components for data processing and for analytics. How should the data pipeline be designed?
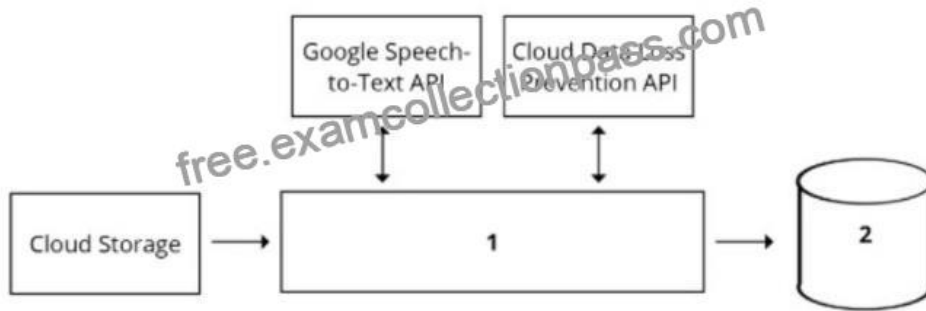
* 1 = Dataflow, 2 = BigQuery
* 1 = Pub/Sub, 2 = Datastore
* 1 = Dataflow, 2 = Cloud SQL
* 1 = Cloud Function, 2 = Cloud SQL

QUESTION 58

A Machine Learning Specialist kicks off a hyperparameter tuning job for a tree-based ensemble model using Amazon SageMaker with Area Under the ROC Curve (AUC) as the objective metric. This workflow will eventually be deployed in a pipeline that retrains and tunes hyperparameters each night to model click-through on data that goes stale every 24 hours.

With the goal of decreasing the amount of time it takes to train these models, and ultimately to decrease costs, the Specialist wants to reconfigure the input hyperparameter range(s).

Which visualization will accomplish this?
* A histogram showing whether the most important input feature is Gaussian.
* A scatter plot with points colored by target variable that uses t-Distributed Stochastic Neighbor Embedding (t-SNE) to visualize the large number of input variables in an easier-to-read dimension.
* A scatter plot showing the performance of the objective metric over each training iteration.
* A scatter plot showing the correlation between maximum tree depth and the objective metric.

QUESTION 59

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?
* An optimization objective that minimizes Log loss
* An optimization objective that maximizes the Precision at a Recall value of 0.50
* An optimization objective that maximizes the area under the precision-recall curve (AUC PR) value
* An optimization objective that maximizes the area under the receiver operating characteristic curve (AUC ROC) value

QUESTION 60

You work for a company that is developing a new video streaming platform. You have been asked to create a recommendation system that will suggest the next video for a user to watch. After a review by an AI Ethics team, you are approved to start development. Each video asset in your company’s catalog has useful metadata (e.g., content type, release date, country), but you do not have any historical user event dat a. How should you build the recommendation system for the first version of the product?

* Launch the product without machine learning. Present videos to users alphabetically, and start collecting user event data so you can develop a recommender model in the future.
* Launch the product without machine learning. Use simple heuristics based on content metadata to recommend similar videos to users, and start collecting user event data so you can develop a recommender model in the future.
* Launch the product with machine learning. Use a publicly available dataset such as MovieLens to train a model using the Recommendations AI, and then apply this trained model to your data.
* Launch the product with machine learning. Generate embeddings for each video by training an autoencoder on the content metadata using TensorFlow. Cluster content based on the similarity of these embeddings, and then recommend videos from the same cluster.

QUESTION 61

You work for a magazine distributor and need to build a model that predicts which customers will renew their subscriptions for the upcoming year. Using your company's historical data as your training set, you created a TensorFlow model and deployed it to AI Platform. You need to determine which customer attribute has the most predictive power for each prediction served by the model. What should you do?

* Use AI Platform notebooks to perform a Lasso regression analysis on your model, which will eliminate features that do not provide a strong signal.
* Stream prediction results to BigQuery. Use BigQuery's CORR(X1, X2) function to calculate the Pearson correlation coefficient between each feature and the target variable.
* Use the AI Explanations feature on AI Platform. Submit each prediction request with the 'explain' keyword to retrieve feature attributions using the sampled Shapley method.
* Use the What-If tool in Google Cloud to determine how your model will perform when individual features are excluded. Rank the feature importance in order of those that caused the most significant performance drop when removed from the model.

QUESTION 62

A Data Scientist is training a multilayer perception (MLP) on a dataset with multiple classes. The target class of interest is unique compared to the other classes within the dataset, but it does not achieve and acceptable recall metric. The Data Scientist has already tried varying the number and size of the MLP's hidden layers, which has not significantly improved the results. A solution to improve recall must be implemented as quickly as possible.

Which techniques should be used to meet these requirements?
* Gather more data using Amazon Mechanical Turk and then retrain
* Train an anomaly detection model instead of an MLP
* Train an XGBoost model instead of an MLP
* Add class weights to the MLP's loss function and then retrain

QUESTION 63

Which of the following metrics should a Machine Learning Specialist generally use to compare/evaluate machine learning classification models against each other?
* Recall
* Misclassification rate
* Mean absolute percentage error (MAPE)
* Area Under the ROC Curve (AUC)

QUESTION 64

You recently joined a machine learning team that will soon release a new project. As a lead on the project, you are asked to

determine the production readiness of the ML components. The team has already tested features and data, model development, and infrastructure. Which additional readiness check should you recommend to the team?

* Ensure that training is reproducible
* Ensure that all hyperparameters are tuned
* Ensure that model performance is monitored
* Ensure that feature expectations are captured in the schema

QUESTION 65

A company wants to predict the sale prices of houses based on available historical sales data. The target variable in the company’s dataset is the sale price. The features include parameters such as the lot size, living area measurements, non-living area measurements, number of bedrooms, number of bathrooms, year built, and postal code. The company wants to use multi-variable linear regression to predict house sale prices.

Which step should a machine learning specialist take to remove features that are irrelevant for the analysis and reduce the model’s complexity?

* Plot a histogram of the features and compute their standard deviation. Remove features with high variance.
* Plot a histogram of the features and compute their standard deviation. Remove features with low variance.
* Build a heatmap showing the correlation of the dataset against itself. Remove features with low mutual correlation scores.
* Run a correlation check of all features against the target variable. Remove features with low target variable correlation scores.

QUESTION 66

You built and manage a production system that is responsible for predicting sales numbers. Model accuracy is crucial, because the production model is required to keep up with market changes. Since being deployed to production, the model hasn’t changed; however the accuracy of the model has steadily deteriorated. What issue is most likely causing the steady decline in model accuracy?

* Poor data quality
* Lack of model retraining
* Too few layers in the model for capturing information
* Incorrect data split ratio during model training, evaluation, validation, and test

QUESTION 67

You have trained a deep neural network model on Google Cloud. The model has low loss on the training data, but is performing worse on the validation dat a. You want the model to be resilient to overfitting. Which strategy should you use when retraining the model?

* Apply a dropout parameter of 0 2, and decrease the learning rate by a factor of 10
* Apply a 12 regularization parameter of 0.4, and decrease the learning rate by a factor of 10.
* Run a hyperparameter tuning job on Al Platform to optimize for the L2 regularization and dropout parameters
* Run a hyperparameter tuning job on Al Platform to optimize for the learning rate, and increase the number of neurons by a factor of 2.

Updated Google Professional-Machine-Learning-Engineer Dumps – Check Free

Professional-Machine-Learning-Engineer Exam Dumps:

https://www.examcollectionpass.com/Google/Professional-Machine-Learning-Engineer-practice-exam-dumps.html]