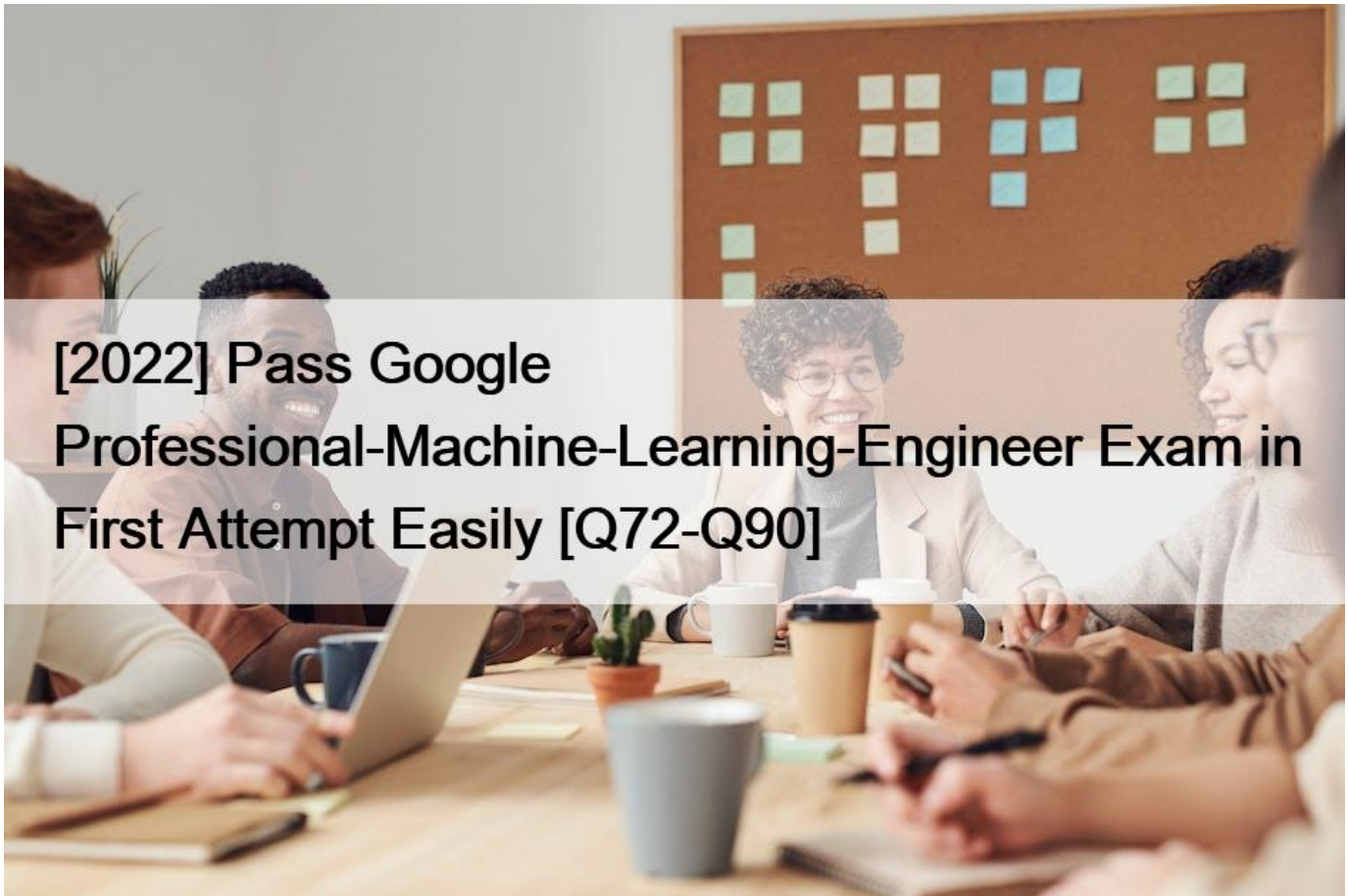


[2022 Pass Google Professional-Machine-Learning-Engineer Exam in First Attempt Easily [Q72-Q90]



[2022] Pass Google Professional-Machine-Learning-Engineer Exam in First Attempt Easily [Q72-Q90]

[2022] Pass Google Professional-Machine-Learning-Engineer Exam in First Attempt Easily
The Most Efficient Professional-Machine-Learning-Engineer Pdf Dumps For Assured Success

Difficulty in Writing Professional Machine Learning Engineer - Google

This exam may go hard for you if you had not done its preparation properly. There are many websites that are offering the latest Google Machine Learning Professional questions and answers but these questions are not verified by Google certified experts and that's why many are failed in their just first attempt. ExamcollectionPass is the best platform which provides the candidate with the necessary Google Machine Learning Professional exam questions that will help him to pass the Google Machine Learning Professional on the first time. Candidate will not have to take the Google Machine Learning Professional twice because with the help of **Google Professional-Machine-Learning-Engineer exam dumps** candidate will have every valuable material required to pass the Google Machine Learning Professional. We are providing the latest and actual questions and that is the reason why this is the one that he needs to use and there are no chances to fail when a candidate will have valid exam dumps from ExamcollectionPass. We have the guarantee that the questions that we have will be the ones that will pass candidate in the Google Machine Learning Professional in the very first attempt.

The aim is to keep candidates up-to-date and we shall automatically amend the material when and when the Offensive Protection reports any changes in the **Google Professional-Machine-Learning-Engineer exam dumps**.

The benefit of obtaining the Professional Machine Learning Engineer - Google Certification - 87% of Google Cloud certified individuals are more confident about their cloud skills- More than 1 in 4 of Google Cloud certified individuals took on more responsibility or leadership roles at work- Professional Cloud Architect was the highest paying certification of 2020 and 2019

Topics of Professional Machine Learning Engineer - Google

Candidates must know the exam topics before they start preparation. Because it will help them in hitting the core. **Google Professional-Machine-Learning-Engineer exam dumps pdf** will include the following topics:

- ML Problem Framing- ML Solution Monitoring, Optimization, and Maintenance- ML Pipeline Automation & Orchestration
- ML Model Development- ML Solution Architecture

Q72. A Machine Learning Specialist trained a regression model, but the first iteration needs optimizing. The Specialist needs to understand whether the model is more frequently overestimating or underestimating the target.

What option can the Specialist use to determine whether it is overestimating or underestimating the target value?

- * Root Mean Square Error (RMSE)
- * Residual plots
- * Area under the curve
- * Confusion matrix

Q73. You work on a growing team of more than 50 data scientists who all use AI Platform. You are designing a strategy to organize your jobs, models, and versions in a clean and scalable way. Which strategy should you choose?

- * Set up restrictive IAM permissions on the AI Platform notebooks so that only a single user or group can access a given instance.
- * Separate each data scientist's work into a different project to ensure that the jobs, models, and versions created by each data scientist are accessible only to that user.
- * Use labels to organize resources into descriptive categories. Apply a label to each created resource so that users can filter the results by label when viewing or monitoring the resources
- * Set up a BigQuery sink for Cloud Logging logs that is appropriately filtered to capture information about AI Platform resource usage In BigQuery create a SQL view that maps users to the resources they are using.

https://cloud.google.com/ai-platform/prediction/docs/resource-labels#overview_of_labels You can add labels to your AI Platform Prediction jobs, models, and model versions, then use those labels to organize resources into categories when viewing or monitoring the resources. For example, you can label jobs by team (such as engineering or research) and development phase (prod or test), then filter the jobs based on the team and phase. Labels are also available on operations, but these labels are derived from the resource to which the operation applies. You cannot add or update labels on an operation.

<https://cloud.google.com/ai-platform/prediction/docs/sharing-models>.

Q74. You are building a model to predict daily temperatures. You split the data randomly and then transformed the training and test datasets. Temperature data for model training is uploaded hourly. During testing, your model performed with 97% accuracy; however, after deploying to production, the model's accuracy dropped to 66%. How can you make your production model more accurate?

- * Normalize the data for the training, and test datasets as two separate steps.
- * Split the training and test data based on time rather than a random split to avoid leakage
- * Add more data to your test set to ensure that you have a fair distribution and sample for testing
- * Apply data transformations before splitting, and cross-validate to make sure that the transformations are applied to both the training and test sets.

Q75. You have been asked to build a model using a dataset that is stored in a medium-sized (~10 GB) BigQuery table. You need to quickly determine whether this data is suitable for model development. You want to create a one-time report that includes both informative visualizations of data distributions and more sophisticated statistical analyses to share with other ML engineers on your team. You require maximum flexibility to create your report. What should you do?

- * Use Vertex AI Workbench user-managed notebooks to generate the report.
- * Use the Google Data Studio to create the report.
- * Use the output from TensorFlow Data Validation on Dataflow to generate the report.
- * Use Dataprep to create the report.

Q76. You have deployed multiple versions of an image classification model on AI Platform. You want to monitor the performance of the model versions overtime. How should you perform this comparison?

- * Compare the loss performance for each model on a held-out dataset.
- * Compare the loss performance for each model on the validation data
- * Compare the receiver operating characteristic (ROC) curve for each model using the What-If Tool
- * Compare the mean average precision across the models using the Continuous Evaluation feature

<https://cloud.google.com/ai-platform/prediction/docs/continuous-evaluation/view-metrics>

Q77. You are an ML engineer on an agricultural research team working on a crop disease detection tool to detect leaf rust spots in images of crops to determine the presence of a disease. These spots, which can vary in shape and size, are correlated to the severity of the disease. You want to develop a solution that predicts the presence and severity of the disease with high accuracy. What should you do?

- * Create an object detection model that can localize the rust spots.
- * Develop an image segmentation ML model to locate the boundaries of the rust spots.
- * Develop a template matching algorithm using traditional computer vision libraries.
- * Develop an image classification ML model to predict the presence of the disease.

Q78. Machine Learning Specialist is training a model to identify the make and model of vehicles in images. The Specialist wants to use transfer learning and an existing model trained on images of general objects. The Specialist collated a large custom dataset of pictures containing different vehicle makes and models.

What should the Specialist do to initialize the model to re-train it with the custom data?

- * Initialize the model with random weights in all layers including the last fully connected layer.
- * Initialize the model with pre-trained weights in all layers and replace the last fully connected layer.
- * Initialize the model with random weights in all layers and replace the last fully connected layer.
- * Initialize the model with pre-trained weights in all layers including the last fully connected layer.

Explanation/Reference:

Q79. A company ingests machine learning (ML) data from web advertising clicks into an Amazon S3 data lake. Click data is added to an Amazon Kinesis data stream by using the Kinesis Producer Library (KPL). The data is loaded into the S3 data lake from the data stream by using an Amazon Kinesis Data Firehose delivery stream.

As the data volume increases, an ML specialist notices that the rate of data ingested into Amazon S3 is relatively constant. There also is an increasing backlog of data for Kinesis Data Streams and Kinesis Data Firehose to ingest.

Which next step is MOST likely to improve the data ingestion rate into Amazon S3?

- * Increase the number of S3 prefixes for the delivery stream to write to.
- * Decrease the retention period for the data stream.
- * Increase the number of shards for the data stream.
- * Add more consumers using the Kinesis Client Library (KCL).

Explanation/Reference:

Q80. Your company manages an application that aggregates news articles from many different online sources and sends them to users. You need to build a recommendation model that will suggest articles to readers that are similar to the articles they are currently reading. Which approach should you use?

- * Create a collaborative filtering system that recommends articles to a user based on the user's past behavior.
- * Encode all articles into vectors using word2vec, and build a model that returns articles based on vector similarity.
- * Build a logistic regression model for each user that predicts whether an article should be recommended to a user.
- * Manually label a few hundred articles, and then train an SVM classifier based on the manually classified articles that categorizes additional articles into their respective categories.

Q81. You manage a team of data scientists who use a cloud-based backend system to submit training jobs. This system has become very difficult to administer, and you want to use a managed service instead. The data scientists you work with use many different frameworks, including Keras, PyTorch, theano, Scikit-team, and custom libraries. What should you do?

- * Use the AI Platform custom containers feature to receive training jobs using any framework
 - * Configure Kubeflow to run on Google Kubernetes Engine and receive training jobs through TFJob
 - * Create a library of VM images on Compute Engine; and publish these images on a centralized repository
 - * Set up Slurm workload manager to receive jobs that can be scheduled to run on your cloud infrastructure.
- because AI platform supported all the frameworks mentioned. And Kubeflow is not managed service in GCP.
- <https://cloud.google.com/ai-platform/training/docs/getting-started-pytorch>

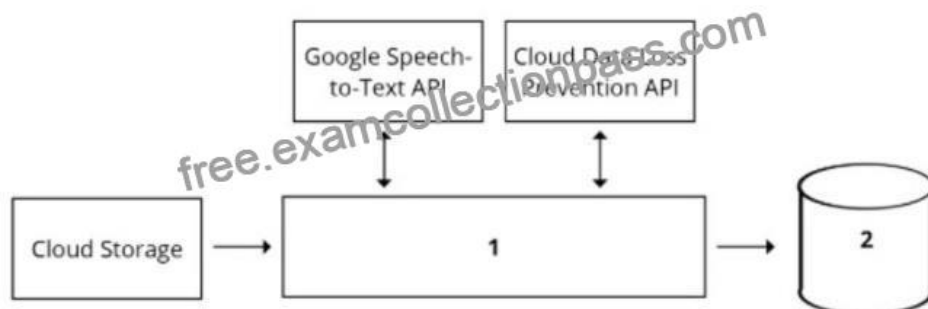
https://cloud.google.com/ai-platform/training/docs/containers-overview#advantages_of_custom_containers Use the ML framework of your choice. If you can't find an AI Platform Training runtime version that supports the ML framework you want to use, then you can build a custom container that installs your chosen framework and use it to run jobs on AI Platform Training.

Q82. You have trained a model on a dataset that required computationally expensive preprocessing operations. You need to execute the same preprocessing at prediction time. You deployed the model on AI Platform for high-throughput online prediction. Which architecture should you use?

- * Validate the accuracy of the model that you trained on preprocessed data
- * Create a new model that uses the raw data and is available in real time
- * Deploy the new model onto AI Platform for online prediction
- * Send incoming prediction requests to a Pub/Sub topic
- * Transform the incoming data using a Dataflow job
- * Submit a prediction request to AI Platform using the transformed data
- * Write the predictions to an outbound Pub/Sub queue
- * Stream incoming prediction request data into Cloud Spanner
- * Create a view to abstract your preprocessing logic.
- * Query the view every second for new records
- * Submit a prediction request to AI Platform using the transformed data
- * Write the predictions to an outbound Pub/Sub queue.
- * Send incoming prediction requests to a Pub/Sub topic
- * Set up a Cloud Function that is triggered when messages are published to the Pub/Sub topic.
- * Implement your preprocessing logic in the Cloud Function

- * Submit a prediction request to AI Platform using the transformed data
- * Write the predictions to an outbound Pub/Sub queue

Q83. 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 (PII) 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

Q84. You work for an advertising company and want to understand the effectiveness of your company's latest advertising campaign. You have streamed 500 MB of campaign data into BigQuery. You want to query the table, and then manipulate the results of that query with a pandas dataframe in an AI Platform notebook. What should you do?

- * Use AI Platform Notebooks's BigQuery cell magic to query the data, and ingest the results as a pandas dataframe
- * Export your table as a CSV file from BigQuery to Google Drive, and use the Google Drive API to ingest the file into your notebook instance
- * Download your table from BigQuery as a local CSV file, and upload it to your AI Platform notebook instance Use pandas.read_csv to ingest the file as a pandas dataframe
- * From a bash cell in your AI Platform notebook, use the bq extract command to export the table as a CSV file to Cloud Storage, and then use gsutil cp to copy the data into the notebook Use pandas.read_csv to ingest the file as a pandas dataframe

Q85. You are building a model to predict daily temperatures. You split the data randomly and then transformed the training and test datasets. Temperature data for model training is uploaded hourly. During testing, your model performed with 97% accuracy; however, after deploying to production, the model's accuracy dropped to 66%. How can you make your production model more accurate?

- * Normalize the data for the training, and test datasets as two separate steps.
- * Split the training and test data based on time rather than a random split to avoid leakage
- * Add more data to your test set to ensure that you have a fair distribution and sample for testing
- * Apply data transformations before splitting, and cross-validate to make sure that the transformations are applied to both the training and test sets.

Q86. You are going to train a DNN regression model with Keras APIs using this code:


```
model = tf.keras.Sequential()
model.add(tf.keras.layers.Dense(
    256,
    use_bias=True,
    activation='relu',
    kernel_initializer=None,
    kernel_regularizer=None,
    input_shape=(500,)))
model.add(tf.keras.layers.Dropout(rate=0.25))
model.add(tf.keras.layers.Dense(
    128, use_bias=True,
    activation='relu',
    kernel_initializer='uniform',
    kernel_regularizer='l2'))
model.add(tf.keras.layers.Dropout(rate=0.25))
model.add(tf.keras.layers.Dense(
    2, use_bias=False,
    activation='softmax'))
model.compile(loss='mse')
```

How many trainable weights does your model have? (The arithmetic below is correct.)

- * $501 * 256 + 257 * 128 + 2 = 161154$
- * $500 * 256 + 256 * 128 + 128 * 2 = 161024$
- * $501 * 256 + 257 * 128 + 128 * 2 = 161408$
- * $500 * 256 * 0.25 + 256 * 128 * 0.25 + 128 * 2 = 40448$

Q87. A Machine Learning Specialist is configuring Amazon SageMaker so multiple Data Scientists can access notebooks, train models, and deploy endpoints. To ensure the best operational performance, the Specialist needs to be able to track how often the Scientists are deploying models, GPU and CPU utilization on the deployed SageMaker endpoints, and all errors that are generated when an endpoint is invoked.

Which services are integrated with Amazon SageMaker to track this information? (Choose two.)

- * AWS CloudTrail
- * AWS Health
- * AWS Trusted Advisor
- * Amazon CloudWatch
- * AWS Config

Explanation/Reference: <https://aws.amazon.com/sagemaker/faqs/>

Q88. A Machine Learning Specialist is developing a custom video recommendation model for an application. The dataset used to train this model is very large with millions of data points and is hosted in an Amazon S3 bucket.

The Specialist wants to avoid loading all of this data onto an Amazon SageMaker notebook instance because it would take hours to move and will exceed the attached 5 GB Amazon EBS volume on the notebook instance.

Which approach allows the Specialist to use all the data to train the model?

- * Load a smaller subset of the data into the SageMaker notebook and train locally. Confirm that the training code is executing and the model parameters seem reasonable. Initiate a SageMaker training job using the full dataset from the S3 bucket using Pipe input mode.
- * Launch an Amazon EC2 instance with an AWS Deep Learning AMI and attach the S3 bucket to the instance. Train on a small amount of the data to verify the training code and hyperparameters. Go back to Amazon SageMaker and train using the full dataset
- * Use AWS Glue to train a model using a small subset of the data to confirm that the data will be compatible with Amazon SageMaker. Initiate a SageMaker training job using the full dataset from the S3 bucket using Pipe input mode.
- * Load a smaller subset of the data into the SageMaker notebook and train locally. Confirm that the training code is executing and the model parameters seem reasonable. Launch an Amazon EC2 instance with an AWS Deep Learning AMI and attach the S3 bucket to train the full dataset.

Q89. You work on an operations team at an international company that manages a large fleet of on-premises servers located in few data centers around the world. Your team collects monitoring data from the servers, including CPU/memory consumption. When an incident occurs on a server, your team is responsible for fixing it. Incident data has not been properly labeled yet. Your management team wants you to build a predictive maintenance solution that uses monitoring data from the VMs to detect potential failures and then alerts the service desk team. What should you do first?

- * Train a time-series model to predict the machines' performance values. Configure an alert if a machine's actual performance values significantly differ from the predicted performance values.
- * Implement a simple heuristic (e.g., based on z-score) to label the machines' historical performance data. Train a model to predict anomalies based on this labeled dataset.
- * Develop a simple heuristic (e.g., based on z-score) to label the machines' historical performance data. Test this heuristic in a production environment.
- * Hire a team of qualified analysts to review and label the machines' historical performance data. Train a model based on this manually labeled dataset.

Q90. 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 model's 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)

algorithm to learn from new inventory data on a daily basis = time series model, best option to deal with time series is recurrent RNN

<https://builtin.com/data-science/recurrent-neural-networks-and-lstm>

We offer you the latest free online Professional-Machine-Learning-Engineer dumps to practice:

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