

← Google Discussions



## Exam Professional Machine Learning Engineer All Questions

View all questions & answers for the Professional Machine Learning Engineer exam

Go to Exam

### EXAM PROFESSIONAL MACHINE LEARNING ENGINEER TOPIC 1 QUESTION 232 DISCUSSI...

Actual exam question from Google's Professional Machine Learning Engineer

Question #: 232

Topic #: 1

[\[All Professional Machine Learning Engineer Questions\]](#)

You need to use TensorFlow to train an image classification model. Your dataset is located in a Cloud Storage directory and contains millions of labeled images. Before training the model, you need to prepare the data. You want the data preprocessing and model training workflow to be as efficient, scalable, and low maintenance as possible. What should you do?

- A. 1. Create a Dataflow job that creates sharded TFRecord files in a Cloud Storage directory.  
2. Reference `tf.data.TFRecordDataset` in the training script.  
3. Train the model by using Vertex AI Training with a V100 GPU.
- B. 1. Create a Dataflow job that moves the images into multiple Cloud Storage directories, where each directory is named according to the corresponding label  
2. Reference `tfds.folder_dataset.ImageFolder` in the training script.  
3. Train the model by using Vertex AI Training with a V100 GPU.
- C. 1. Create a Jupyter notebook that uses an `n1-standard-64` V100 GPU Vertex AI Workbench instance.  
2. Write a Python script that creates sharded TFRecord files in a directory inside the instance.  
3. Reference `tf.data.TFRecordDataset` in the training script.  
4. Train the model by using the Workbench instance.
- D. 1. Create a Jupyter notebook that uses an `n1-standard-64`, V100 GPU Vertex AI Workbench instance.  
2. Write a Python script that copies the images into multiple Cloud Storage directories, where each. directory is named according to the corresponding label.  
3. Reference `tfds.foladr_dataset.ImageFolder` in the training script.  
4. Train the model by using the Workbench instance.


Show Suggested Answer

by  pikachu007 at Jan. 13, 2024, 8:01 a.m.

## Comments

Type your comment...




Submit

  **pinimichele01** Highly Voted  6 months, 2 weeks ago

Selected Answer: A

millions of labeled images -> dataflow  
tfrecord faster than folder-based

   upvoted 5 times

  **AzureDP900** Most Recent  3 months, 2 weeks ago

A is correct Here's why  
You need to prepare the data before training an image classification model.  
Using TFRecord files allows you to store your data in a format that can be efficiently read and processed by TensorFlow.  
Sharding the data into multiple files allows for parallel processing and scalability.  
Dataflow is a Google Cloud service that provides a scalable and reliable way to process large datasets.  
By using Vertex AI Training with a V100 GPU, you can train your model in an efficient and cost-effective manner.


   upvoted 1 times

  **b1a8fae** 9 months, 1 week ago

Selected Answer: A

Ideally you want to export your data in TFRecords (most efficient image format) in Cloud Storage, and not in the instance (to improve scalability)

   upvoted 3 times

  **pikachu007** 9 months, 1 week ago

Selected Answer: A

B. Folder-Based Structure: While viable, it's less efficient for large datasets compared to TFRecord files, potentially leading to slower I/O during training.  
C. Workbench Processing: Local preprocessing on a single instance can be less scalable and efficient for millions of images, potentially introducing bottlenecks.  
D. Workbench Training: While Workbench offers a Jupyter environment, Vertex AI Training is specifically designed for scalable model training, providing optimized hardware and infrastructure.

   upvoted 2 times

Start Learning for free





## Social Media

[Facebook](#) , [Twitter](#)

[YouTube](#) , [Reddit](#)

[Pinterest](#)



We are the biggest and most updated IT certification exam material website.

Using our own resources, we strive to strengthen the IT professionals community for free.



© 2024 ExamTopics

ExamTopics doesn't offer Real Microsoft Exam Questions. ExamTopics doesn't offer Real Amazon Exam Questions. ExamTopics Materials do not contain actual questions and answers from Cisco's Certification Exams.

CFA Institute does not endorse, promote or warrant the accuracy or quality of ExamTopics. CFA® and Chartered Financial Analyst® are registered trademarks owned by CFA Institute.