

# Google

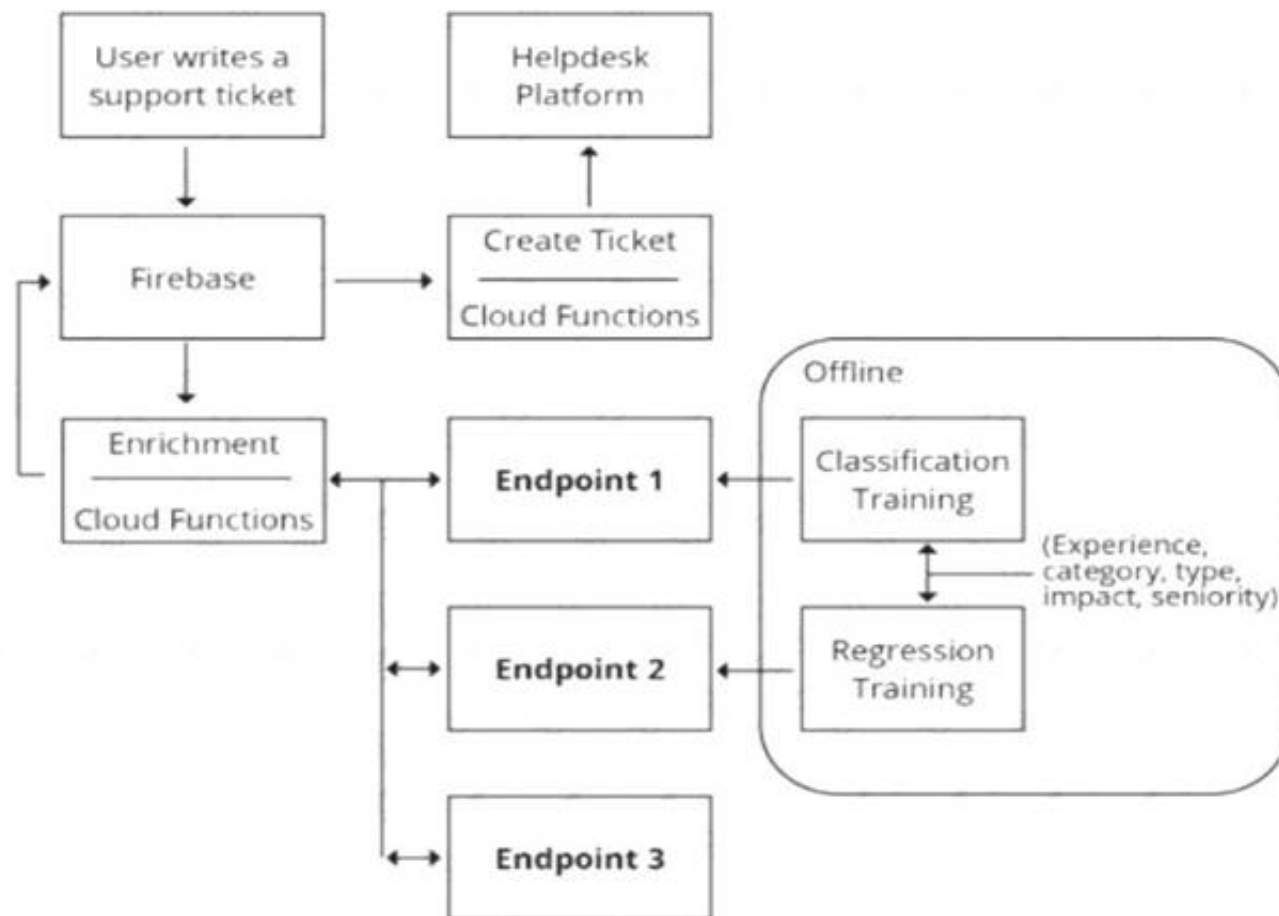
## Exam Questions Professional-Machine-Learning-Engineer

Google Professional Machine Learning Engineer



### NEW QUESTION 1

You are designing an architecture with a serverless ML system to enrich customer support tickets with informative metadata before they are routed to a support agent. You need a set of models to predict ticket priority, predict ticket resolution time, and perform sentiment analysis to help agents make strategic decisions when they process support requests. Tickets are not expected to have any domain-specific terms or jargon. The proposed architecture has the following flow:



Which endpoints should the Enrichment Cloud Functions call?

- A. 1 = AI Platform, 2 = AI Platform, 3 = AutoML Vision
- B. 1 = AI Platform, 2 = AI Platform, 3 = AutoML Natural Language
- C. 1 = AI Platform, 2 = AI Platform, 3 = Cloud Natural Language API
- D. 1 = cloud Natural Language API, 2 = AI Platform, 3 = Cloud Vision API

**Answer: B**

### NEW QUESTION 2

Your organization wants to make its internal shuttle service route more efficient. The shuttles currently stop at all pick-up points across the city every 30 minutes between 7 am and 10 am. The development team has already built an application on Google Kubernetes Engine that requires users to confirm their presence and shuttle station one day in advance. What approach should you take?

- A. 1. Build a tree-based regression model that predicts how many passengers will be picked up at each shuttle station.\* 2. Dispatch an appropriately sized shuttle and provide the map with the required stops based on the prediction.
- B. 1. Build a tree-based classification model that predicts whether the shuttle should pick up passengers at each shuttle station.\* 2. Dispatch an available shuttle and provide the map with the required stops based on the prediction
- C. 1. Define the optimal route as the shortest route that passes by all shuttle stations with confirmed attendance at the given time under capacity constraints.\* 2 Dispatch an appropriately sized shuttle and indicate the required stops on the map
- D. 1. Build a reinforcement learning model with tree-based classification models that predict the presenceof passengers at shuttle stops as agents and a reward function around a distance-based metric\* 2. Dispatch an appropriately sized shuttle and provide the map with the required stops based on the simulated outcome.

**Answer: D**

### NEW QUESTION 3

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?

- A. Classification
- B. Reinforcement Learning
- C. Recurrent Neural Networks (RNN)
- D. Convolutional Neural Networks (CNN)

**Answer: B**

### NEW QUESTION 4

You work for a large hotel chain and have been asked to assist the marketing team in gathering predictions for a targeted marketing strategy. You need to make predictions about user lifetime value (LTV) over the next 30 days so that marketing can be adjusted accordingly. The customer dataset is in BigQuery, and you are preparing the tabular data for training with AutoML Tables. This data has a time signal that is spread across multiple columns. How should you ensure that AutoML fits the best model to your data?

- A. Manually combine all columns that contain a time signal into an array Allow AutoML to interpret this array appropriatelyChoose an automatic data split across

the training, validation, and testing sets

- B. Submit the data for training without performing any manual transformations Allow AutoML to handle the appropriate transformations Choose an automatic data split across the training, validation, and testing sets
- C. Submit the data for training without performing any manual transformations, and indicate an appropriate column as the Time column Allow AutoML to split your data based on the time signal provided, and reserve the more recent data for the validation and testing sets
- D. Submit the data for training without performing any manual transformations Use the columns that have a time signal to manually split your data Ensure that the data in your validation set is from 30 days after the data in your training set and that the data in your testing set is from 30 days after your validation set

**Answer:** D

#### NEW QUESTION 5

You need to design a customized deep neural network in Keras that will predict customer purchases based on their purchase history. You want to explore model performance using multiple model architectures, store training data, and be able to compare the evaluation metrics in the same dashboard. What should you do?

- A. Create multiple models using AutoML Tables
- B. Automate multiple training runs using Cloud Composer
- C. Run multiple training jobs on AI Platform with similar job names
- D. Create an experiment in Kubeflow Pipelines to organize multiple runs

**Answer:** C

#### NEW QUESTION 6

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?

- A. AutoML Vision model
- B. AutoML Vision Edge mobile-versatile-1 model
- C. AutoML Vision Edge mobile-low-latency-1 model
- D. AutoML Vision Edge mobile-high-accuracy-1 model

**Answer:** A

#### NEW QUESTION 7

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.)

- A.  $501 \times 256 + 257 \times 128 + 2 = 161154$
- B.  $500 \times 256 + 256 \times 128 + 128 \times 2 = 161024$
- C.  $501 \times 256 + 257 \times 128 + 128 \times 2 = 161408$
- D.  $500 \times 256 + 256 \times 128 + 128 \times 2 = 40448$

**Answer:** D

**NEW QUESTION 8**

During batch training of a neural network, you notice that there is an oscillation in the loss. How should you adjust your model to ensure that it converges?

- A. Increase the size of the training batch
- B. Decrease the size of the training batch
- C. Increase the learning rate hyperparameter
- D. Decrease the learning rate hyperparameter

**Answer:** C

**NEW QUESTION 9**

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?

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

**Answer:** B

**NEW QUESTION 10**

Your team is building an application for a global bank that will be used by millions of customers. You built a forecasting model that predicts customers' account balances 3 days in the future. Your team will use the results in a new feature that will notify users when their account balance is likely to drop below \$25. How should you serve your predictions?

- A. 1. Create a Pub/Sub topic for each user\* 2. Deploy a Cloud Function that sends a notification when your model predicts that a user's account balance will drop below the \$25 threshold.
- B. 1. Create a Pub/Sub topic for each user\* 2. Deploy an application on the App Engine standard environment that sends a notification when your model predicts that a user's account balance will drop below the \$25 threshold
- C. 1. Build a notification system on Firebase\* 2. Register each user with a user ID on the Firebase Cloud Messaging server, which sends a notification when the average of all account balance predictions drops below the \$25 threshold
- D. 1. Build a notification system on Firebase\* 2. Register each user with a user ID on the Firebase Cloud Messaging server, which sends a notification when your model predicts that a user's account balance will drop below the \$25 threshold

**Answer:** B

**NEW QUESTION 10**

You are building a real-time prediction engine that streams files which may contain Personally Identifiable Information (PII) to Google Cloud. You want to use the Cloud Data Loss Prevention (DLP) API to scan the files. How should you ensure that the PII is not accessible by unauthorized individuals?

- A. Stream all files to Google Cloud and then write the data to BigQuery. Periodically conduct a bulk scan of the table using the DLP API.
- B. Stream all files to Google Cloud, and write batches of the data to BigQuery. While the data is being written to BigQuery, conduct a bulk scan of the data using the DLP API.
- C. Create two buckets of data: Sensitive and Non-sensitive. Write all data to the Non-sensitive bucket. Periodically conduct a bulk scan of that bucket using the DLP API, and move the sensitive data to the Sensitive bucket.
- D. Create three buckets of data: Quarantine, Sensitive, and Non-sensitive. Write all data to the Quarantine bucket.
- E. Periodically conduct a bulk scan of that bucket using the DLP API, and move the data to either the Sensitive or Non-Sensitive bucket.

**Answer:** A

**NEW QUESTION 13**

You have been asked to develop an input pipeline for an ML training model that processes images from disparate sources at a low latency. You discover that your input data does not fit in memory. How should you create a dataset following Google-recommended best practices?

- A. Create a `tf.data.Dataset.prefetch` transformation
- B. Convert the images to `tf.Tensor` Objects, and then run `Dataset`
- C. `from_tensor_slices()`.
- D. Convert the images to `tf.Tensor` Objects, and then run `tf.data`
- E. `tf.data`
- F. `Dataset`
- G. `from_tensors()`.
- H. Convert the images into `TFRecords`, store the images in Cloud Storage, and then use the `tf.data`
- I. `tf.data` API to read the images for training

**Answer:** D

**NEW QUESTION 18**

You are an ML engineer at a bank that has a mobile application. Management has asked you to build an ML-based biometric authentication for the app that verifies a customer's identity based on their fingerprint. Fingerprints are considered highly sensitive personal information and cannot be downloaded and stored into the bank databases. Which learning strategy should you recommend to train and deploy this ML model?

- A. Differential privacy
- B. Federated learning
- C. MD5 to encrypt data
- D. Data Loss Prevention API

**Answer:** B

**NEW QUESTION 20**

You are responsible for building a unified analytics environment across a variety of on-premises data marts. Your company is experiencing data quality and security challenges when integrating data across the servers, caused by the use of a wide range of disconnected tools and temporary solutions. You need a fully managed, cloud-native data integration service that will lower the total cost of work and reduce repetitive work. Some members on your team prefer a codeless interface for building Extract, Transform, Load (ETL) process. Which service should you use?

- A. Dataflow
- B. Dataprep
- C. Apache Flink
- D. Cloud Data Fusion

**Answer: D**

**NEW QUESTION 21**

As the lead ML Engineer for your company, you are responsible for building ML models to digitize scanned customer forms. You have developed a TensorFlow model that converts the scanned images into text and stores them in Cloud Storage. You need to use your ML model on the aggregated data collected at the end of each day with minimal manual intervention. What should you do?

- A. Use the batch prediction functionality of AI Platform
- B. Create a serving pipeline in Compute Engine for prediction
- C. Use Cloud Functions for prediction each time a new data point is ingested
- D. Deploy the model on AI Platform and create a version of it for online inference.

**Answer: D**

**NEW QUESTION 25**

You are developing models to classify customer support emails. You created models with TensorFlow Estimators using small datasets on your on-premises system, but you now need to train the models using large datasets to ensure high performance. You will port your models to Google Cloud and want to minimize code refactoring and infrastructure overhead for easier migration from on-prem to cloud. What should you do?

- A. Use AI Platform for distributed training
- B. Create a cluster on Dataproc for training
- C. Create a Managed Instance Group with autoscaling
- D. Use Kubeflow Pipelines to train on a Google Kubernetes Engine cluster.

**Answer: D**

**NEW QUESTION 30**

You work for a social media company. You need to detect whether posted images contain cars. Each training example is a member of exactly one class. You have trained an object detection neural network and deployed the model version to AI Platform Prediction for evaluation. Before deployment, you created an evaluation job and attached it to the AI Platform Prediction model version. You notice that the precision is lower than your business requirements allow. How should you adjust the model's final layer softmax threshold to increase precision?

- A. Increase the recall
- B. Decrease the recall.
- C. Increase the number of false positives
- D. Decrease the number of false negatives

**Answer: D**

**NEW QUESTION 33**

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 data. You want the model to be resilient to overfitting. Which strategy should you use when retraining the model?

- A. Apply a dropout parameter of 0.2, and decrease the learning rate by a factor of 10
- B. Apply a L2 regularization parameter of 0.4, and decrease the learning rate by a factor of 10.
- C. Run a hyperparameter tuning job on AI Platform to optimize for the L2 regularization and dropout parameters
- D. Run a hyperparameter tuning job on AI Platform to optimize for the learning rate, and increase the number of neurons by a factor of 2.

**Answer: A**

**NEW QUESTION 38**

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?

- A. Use the class distribution to generate 10% positive examples
- B. Use a convolutional neural network with max pooling and softmax activation
- C. Downsample the data with upweighting to create a sample with 10% positive examples
- D. Remove negative examples until the numbers of positive and negative examples are equal

**Answer: D**

**NEW QUESTION 40**

You are an ML engineer at a regulated insurance company. You are asked to develop an insurance approval model that accepts or rejects insurance applications from potential customers. What factors should you consider before building the model?



- A. Redaction, reproducibility, and explainability
- B. Traceability, reproducibility, and explainability
- C. Federated learning, reproducibility, and explainability
- D. Differential privacy federated learning, and explainability

**Answer: B**

#### NEW QUESTION 43

You are training an LSTM-based model on AI Platform to summarize text using the following job submission script:

```
gcloud ai-platform jobs submit training $JOB_NAME \
  --package-path $TRAINER_PACKAGE_PATH \
  --module-name $MAIN_TRAINER_MODULE \
  --job-dir $JOB_DIR \
  --region $REGION \
  --scale-tier basic \
  -- \
  --epochs 20 \
  --batch_size=32 \
  --learning_rate=0.001 \
```

You want to ensure that training time is minimized without significantly compromising the accuracy of your model. What should you do?

- A. Modify the 'epochs' parameter
- B. Modify the 'scale-tier' parameter
- C. Modify the batch size' parameter
- D. Modify the 'learning rate' parameter

**Answer: A**

#### NEW QUESTION 46

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=data,
headers=headers)
```

What is the correct way to write the predict request?

- A. data = json.dumps({"signature\_name": "serving\_default\ "instances": [fab', 'be1, 'cd']})
- B. data = json.dumps({"signature\_name": "serving\_default"! "instances": [['a', 'b', 'c', 'd', 'e', 'f']])
- C. data = json.dumps({"signature\_name": "serving\_default", "instances": [['a', 'b\ 'c'1, [d\ 'e\ T']])
- D. data = json.dumps({"signature\_name": f,serving\_default", "instances": [['a', 'b'], [c\ 'd'], [e\ T']])

**Answer: B**

#### NEW QUESTION 47

You work for a large technology company that wants to modernize their contact center. You have been asked to develop a solution to classify incoming calls by product so that requests can be more quickly routed to the correct support team. You have already transcribed the calls using the Speech-to-Text API. You want to minimize data preprocessing and development time. How should you build the model?

- A. Use the AI Platform Training built-in algorithms to create a custom model

- B. Use AutoML Natural Language to extract custom entities for classification
- C. Use the Cloud Natural Language API to extract custom entities for classification
- D. Build a custom model to identify the product keywords from the transcribed calls, and then run the keywords through a classification algorithm

**Answer:** A

#### NEW QUESTION 52

Your team is working on an NLP research project to predict political affiliation of authors based on articles they have written. You have a large training dataset that is structured like this:

```
AuthorA:Political Party A
  TextA1: [SentenceA11, SentenceA12, SentenceA13, ...]
  TextA2: [SentenceA21, SentenceA22, SentenceA23, ...]
  ...
AuthorB:Political Party B
  TextB1: [SentenceB11, SentenceB12, SentenceB13, ...]
  TextB2: [SentenceB21, SentenceB22, SentenceB23, ...]
  ...
AuthorC:Political Party B
  TextC1: [SentenceC11, SentenceC12, SentenceC13, ...]
  TextC2: [SentenceC21, SentenceC22, SentenceC23, ...]
  ...
AuthorD:Political Party A
  TextD1: [SentenceD11, SentenceD12, SentenceD13, ...]
  TextD2: [SentenceD21, SentenceD22, SentenceD23, ...]
  ...
...
```

A)

Distribute texts randomly across the train-test-eval subsets:

```
Train set: [TextA1, TextB2, ...]
Test set: [TextA2, TextC1, TextD2, ...]
Eval set: [TextB1, TextC2, TextD1, ...]
```

B)

Distribute authors randomly across the train-test-eval subsets: (\*)

```
Train set: [TextA1, TextA2, TextD1, TextD2, ...]
Test set: [TextB1, TextB2, ...]
Eval set: [TextC1, TextC2, ...]
```

C)

Distribute sentences randomly across the train-test-eval subsets:

```
Train set: [SentenceA11, SentenceA21, Sentence B11, SentenceB21, SentenceC11, SentenceD21, ...]
Test set: [SentenceA12, SentenceA22, Sentence B12, SentenceC22, SentenceC12, SentenceD22, ...]
Eval set: [SentenceA13, SentenceA23, Sentence B13, SentenceC23, SentenceC13, SentenceD31, ...]
```

D)

Distribute paragraphs of texts (i.e., chunks of consecutive sentences) across the train-test-eval subsets:

```
Train set: [SentenceA11, SentenceA12, Sentence D11, SentenceD12, ...]
Test set: [SentenceA13, SentenceB13, Sentence B21, SentenceD23, SentenceC12, SentenceD13, ...]
Eval set: [SentenceA11, SentenceA22, Sentence B13, SentenceD22, SentenceC23, SentenceD11, ...]
```

- A. Option A
- B. Option B
- C. Option C
- D. Option D

**Answer:** D

#### NEW QUESTION 56

You work for an online retail company that is creating a visual search engine. You have set up an end-to-end ML pipeline on Google Cloud to classify whether an image contains your company's product. Expecting the release of new products in the near future, you configured a retraining functionality in the pipeline so that new data can be fed into your ML models. You also want to use AI Platform's continuous evaluation service to ensure that the models have high accuracy on your test data set. What should you do?

- A. Keep the original test dataset unchanged even if newer products are incorporated into retraining
- B. Extend your test dataset with images of the newer products when they are introduced to retraining
- C. Replace your test dataset with images of the newer products when they are introduced to retraining.
- D. Update your test dataset with images of the newer products when your evaluation metrics drop below a pre-decided threshold.

**Answer:** C

#### NEW QUESTION 58

You are building a linear model with over 100 input features, all with values between -1 and 1. You suspect that many features are non-informative. You want to remove the non-informative features from your model while keeping the informative ones in their original form. Which technique should you use?

- A. Use Principal Component Analysis to eliminate the least informative features.
- B. Use L1 regularization to reduce the coefficients of uninformative features to 0.
- C. After building your model, use Shapley values to determine which features are the most informative.
- D. Use an iterative dropout technique to identify which features do not degrade the model when removed.

**Answer:** C

#### NEW QUESTION 62

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?

- A. Poor data quality
- B. Lack of model retraining
- C. Too few layers in the model for capturing information
- D. Incorrect data split ratio during model training, evaluation, validation, and test

**Answer:** D

#### NEW QUESTION 63

You are an ML engineer at a global shoe store. You manage the ML models for the company's website. You are asked to build a model that will recommend new products to the user based on their purchase behavior and similarity with other users. What should you do?

- A. Build a classification model
- B. Build a knowledge-based filtering model
- C. Build a collaborative-based filtering model
- D. Build a regression model using the features as predictors

**Answer:** C

#### NEW QUESTION 67

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?

- A. Use the AI Platform custom containers feature to receive training jobs using any framework
- B. Configure Kubeflow to run on Google Kubernetes Engine and receive training jobs through TFJob
- C. Create a library of VM images on Compute Engine; and publish these images on a centralized repository
- D. Set up Slurm workload manager to receive jobs that can be scheduled to run on your cloud infrastructure.

**Answer:** D

#### NEW QUESTION 71

Your team needs to build a model that predicts whether images contain a driver's license, passport, or credit card. The data engineering team already built the pipeline and generated a dataset composed of 10,000 images with driver's licenses, 1,000 images with passports, and 1,000 images with credit cards. You now have to train a model with the following label map: ['driverslicense', 'passport', 'credit\_card']. Which loss function should you use?

- A. Categorical hinge
- B. Binary cross-entropy
- C. Categorical cross-entropy
- D. Sparse categorical cross-entropy

**Answer:** B

#### NEW QUESTION 72

.....



## Thank You for Trying Our Product

### We offer two products:

1st - We have Practice Tests Software with Actual Exam Questions

2nd - Questions and Answers in PDF Format

### Professional-Machine-Learning-Engineer Practice Exam Features:

- \* Professional-Machine-Learning-Engineer Questions and Answers Updated Frequently
- \* Professional-Machine-Learning-Engineer Practice Questions Verified by Expert Senior Certified Staff
- \* Professional-Machine-Learning-Engineer Most Realistic Questions that Guarantee you a Pass on Your FirstTry
- \* Professional-Machine-Learning-Engineer Practice Test Questions in Multiple Choice Formats and Updatesfor 1 Year

**100% Actual & Verified — Instant Download, Please Click**  
**[Order The Professional-Machine-Learning-Engineer Practice Test Here](#)**