
# Amazon-Web-Services

## Exam Questions MLA-C01

AWS Certified Machine Learning Engineer - Associate

**NEW QUESTION 1**
An ML engineer is using Amazon SageMaker to train a deep learning model that requires distributed training. After some training attempts, the ML engineer observes that the instances are not performing as expected. The ML engineer identifies communication overhead between the training instances.
What should the ML engineer do to MINIMIZE the communication overhead between the instances?

A. Place the instances in the same VPC subne
B. Store the data in a different AWS Region from where the instances are deployed.
C. Place the instances in the same VPC subnet but in different Availability Zone
D. Store the data in a different AWS Region from where the instances are deployed.
E. Place the instances in the same VPC subne
F. Store the data in the same AWS Region and Availability Zone where the instances are deployed.
G. Place the instances in the same VPC subne
H. Store the data in the same AWS Region but in a different Availability Zone from where the instances are deployed.

**Answer:** C

**NEW QUESTION 2**
A company has historical data that shows whether customers needed long-term support from company staff. The company needs to develop an ML model to predict whether new customers will require long-term support.
Which modeling approach should the company use to meet this requirement?

A. Anomaly detection
B. Linear regression
C. Logistic regression
D. Semantic segmentation

**Answer:** C

**NEW QUESTION 3**
An ML engineer needs to use data with Amazon SageMaker Canvas to train an ML model. The data is stored in Amazon S3 and is complex in structure. The ML engineer must use a file format that minimizes processing time for the data.
Which file format will meet these requirements?

A. CSV files compressed with Snappy
B. JSON objects in JSONL format
C. JSON files compressed with gzip
D. Apache Parquet files

**Answer:** D

**NEW QUESTION 4**
A company is using an AWS Lambda function to monitor the metrics from an ML model. An ML engineer needs to implement a solution to send an email message when the metrics breach a threshold.
Which solution will meet this requirement?

A. Log the metrics from the Lambda function to AWS CloudTrai
B. Configure a CloudTrail trail to send the email message.
C. Log the metrics from the Lambda function to Amazon CloudFron
D. Configure an Amazon CloudWatch alarm to send the email message.
E. Log the metrics from the Lambda function to Amazon CloudWatc
F. Configure a CloudWatch alarm to send the email message.
G. Log the metrics from the Lambda function to Amazon CloudWatc
H. Configure an Amazon CloudFront rule to send the email message.

**Answer:** D

**NEW QUESTION 5**
A company has an application that uses different APIs to generate embeddings for input text. The company needs to implement a solution to automatically rotate the API tokens every 3 months.
Which solution will meet this requirement?

A. Store the tokens in AWS Secrets Manage
B. Create an AWS Lambda function to perform the rotation.
C. Store the tokens in AWS Systems Manager Parameter Stor
D. Create an AWS Lambda function to perform the rotation.
E. Store the tokens in AWS Key Management Service (AWS KMS). Use an AWS managed key to perform the rotation.
F. Store the tokens in AWS Key Management Service (AWS KMS). Use an AWS owned key to perform the rotation.

**Answer:** A

**NEW QUESTION 6**
A company is running ML models on premises by using custom Python scripts and proprietary datasets. The company is using PyTorch. The model building requires unique domain knowledge. The company needs to move the models to AWS.
Which solution will meet these requirements with the LEAST effort?

A. Use SageMaker built-in algorithms to train the proprietary datasets.

B. Use SageMaker script mode and premade images for ML frameworks.
C. Build a container on AWS that includes custom packages and a choice of ML frameworks.
D. Purchase similar production models through AWS Marketplace.

**Answer:** B

**NEW QUESTION 7**
A company has an ML model that needs to run one time each night to predict stock values. The model input is 3 MB of data that is collected during the current day. The model produces the predictions for the next day. The prediction process takes less than 1 minute to finish running.
How should the company deploy the model on Amazon SageMaker to meet these requirements?

A. Use a multi-model serverless endpoin
B. Enable caching.
C. Use an asynchronous inference endpoin
D. Set the InitialInstanceCount parameter to 0.
E. Use a real-time endpoin
F. Configure an auto scaling policy to scale the model to 0 when the model is not in use.
G. Use a serverless inference endpoin
H. Set the MaxConcurrency parameter to 1.

**Answer:** D

**NEW QUESTION 8**
A company has trained an ML model in Amazon SageMaker. The company needs to host the model to provide inferences in a production environment.
The model must be highly available and must respond with minimum latency. The size of each request will be between 1 KB and 3 MB. The model will receive unpredictable bursts of requests during the day. The inferences must adapt proportionally to the changes in demand.
How should the company deploy the model into production to meet these requirements?

A. Create a SageMaker real-time inference endpoin
B. Configure auto scalin
C. Configure the endpoint to present the existing model.
D. Deploy the model on an Amazon Elastic Container Service (Amazon ECS) cluste
E. Use ECS scheduled scaling that is based on the CPU of the ECS cluster.
F. Install SageMaker Operator on an Amazon Elastic Kubernetes Service (Amazon EKS) cluste
G. Deploy the model in Amazon EK
H. Set horizontal pod auto scaling to scale replicas based on the memory metric.
I. Use Spot Instances with a Spot Fleet behind an Application Load Balancer (ALB) for inference
J. Use the ALBRequestCountPerTarget metric as the metric for auto scaling.

**Answer:** A

**NEW QUESTION 9**
A financial company receives a high volume of real-time market data streams from an external provider. The streams consist of thousands of JSON records every second.
The company needs to implement a scalable solution on AWS to identify anomalous data points.
Which solution will meet these requirements with the LEAST operational overhead?

A. Ingest real-time data into Amazon Kinesis data stream
B. Use the built-in RANDOM_CUT_FOREST function in Amazon Managed Service for Apache Flink to process the data streams and to detect data anomalies.
C. Ingest real-time data into Amazon Kinesis data stream
D. Deploy an Amazon SageMaker endpoint for real-time outlier detectio
E. Create an AWS Lambda function to detect anomalie
F. Use the data streams to invoke the Lambda function.
G. Ingest real-time data into Apache Kafka on Amazon EC2 instance
H. Deploy an Amazon SageMaker endpoint for real-time outlier detectio
I. Create an AWS Lambda function to detect anomalie
J. Use the data streams to invoke the Lambda function.
K. Send real-time data to an Amazon Simple Queue Service (Amazon SQS) FIFO queu
L. Create an AWS Lambda function to consume the queue message
M. Program the Lambda function to start an AWS Glue extract, transform, and load (ETL) job for batch processing and anomaly detection.

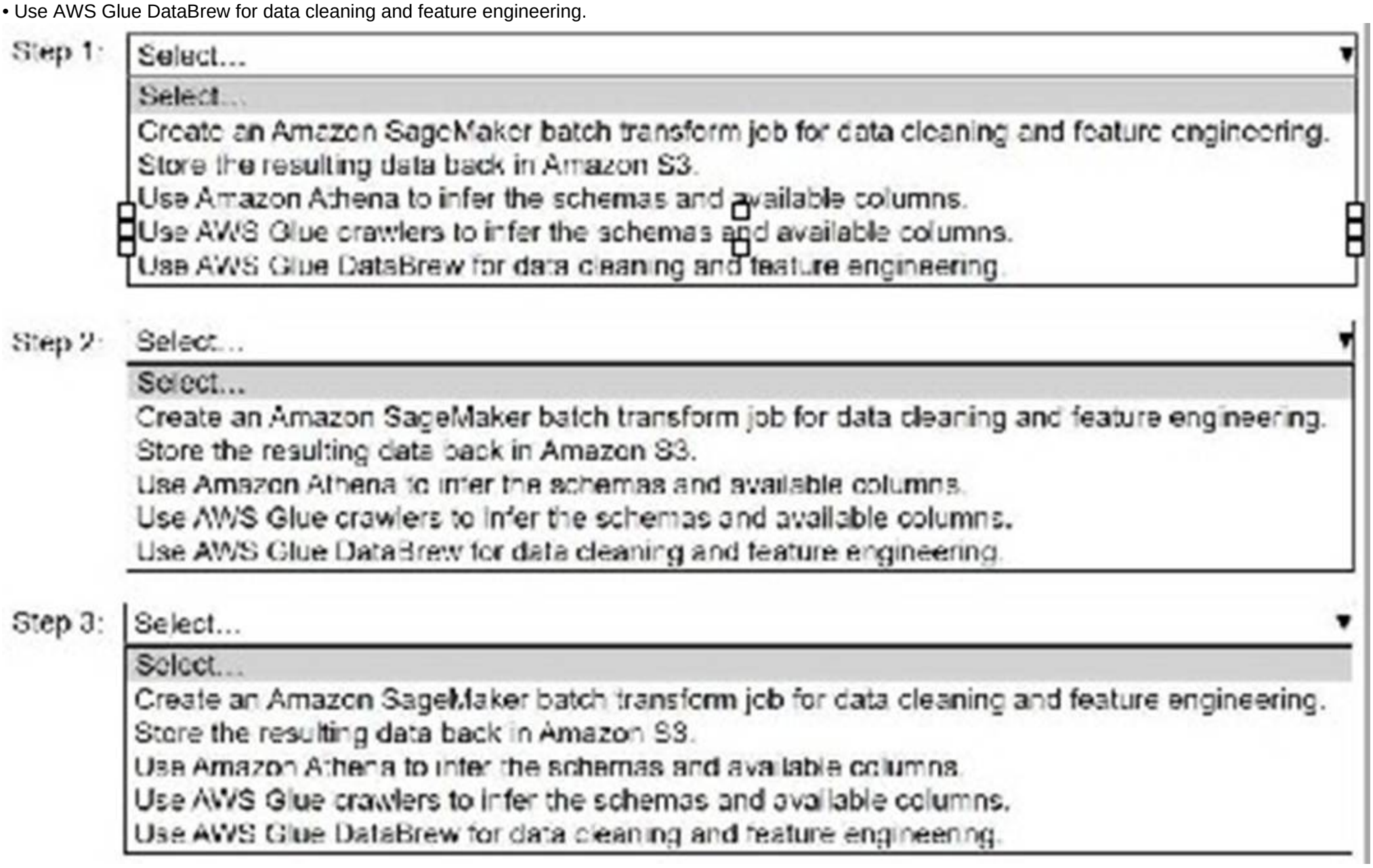
**Answer:** A

**NEW QUESTION 10**
A company has trained and deployed an ML model by using Amazon SageMaker. The company needs to implement a solution to record and monitor all the API call events for the SageMaker endpoint. The solution also must provide a notification when the number of API call events breaches a threshold.
Use SageMaker Debugger to track the inferences and to report metrics. Create a custom rule to provide a notification when the threshold is breached.
Which solution will meet these requirements?

A. Use SageMaker Debugger to track the inferences and to report metric
B. Create a custom rule to provide a notification when the threshold is breached.
C. Use SageMaker Debugger to track the inferences and to report metric
D. Use the tensor_variance built-in rule to provide a notification when the threshold is breached.
E. Log all the endpoint invocation API events by using AWS CloudTrai
F. Use an Amazon CloudWatch dashboard for monitorin
G. Set up a CloudWatch alarm to provide notification when the threshold is breached.
H. Add the Invocations metric to an Amazon CloudWatch dashboard for monitorin
I. Set up a CloudWatch alarm to provide notification when the threshold is breached.

**Answer:** D

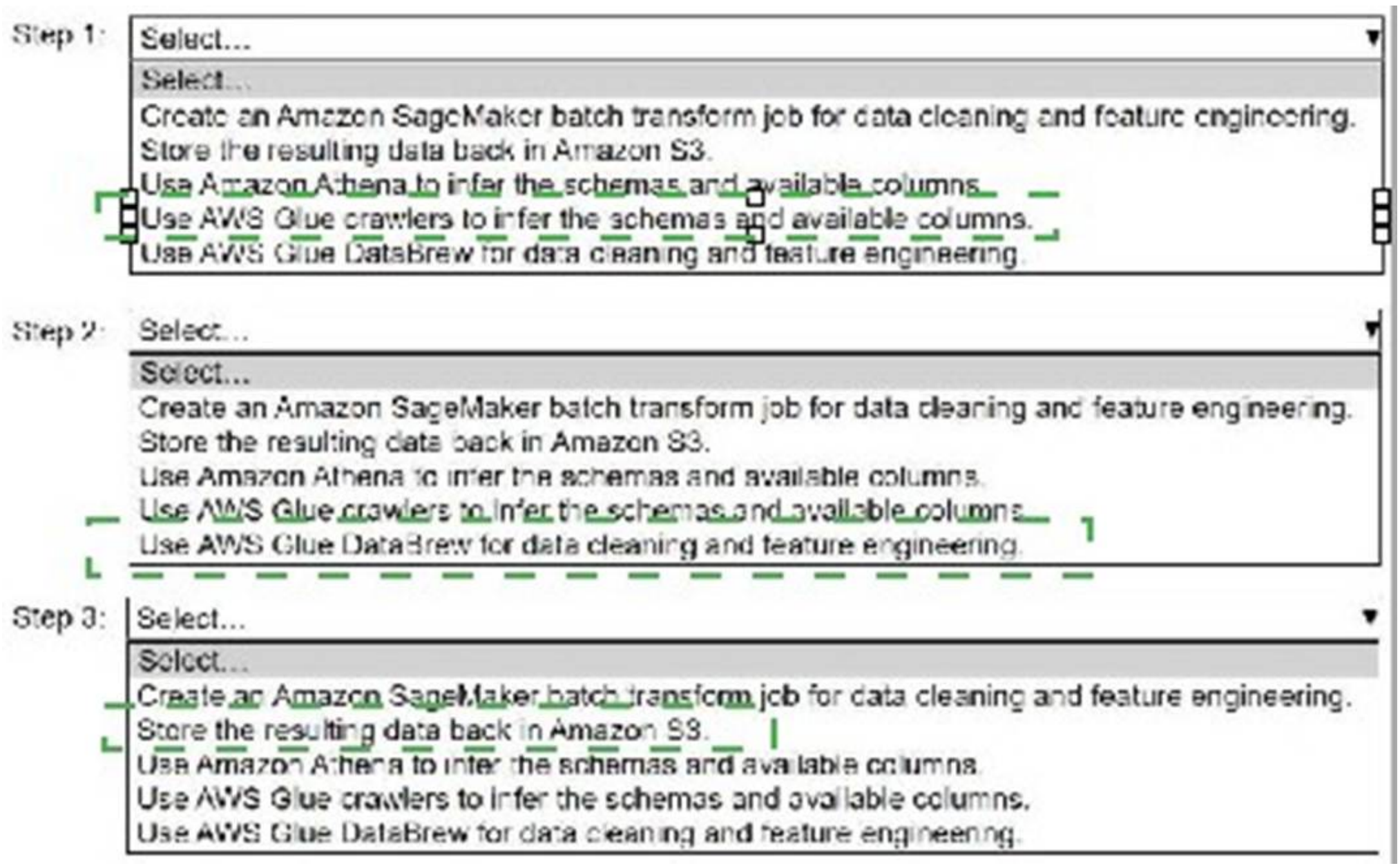
**NEW QUESTION 10**
HOTSPOT
A company stores historical data in .csv files in Amazon S3. Only some of the rows and columns in the .csv files are populated. The columns are not labeled. An ML engineer needs to prepare and store the data so that the company can use the data to train ML models.
Select and order the correct steps from the following list to perform this task. Each step should be selected one time or not at all. (Select and order three.)
• Create an Amazon SageMaker batch transform job for data cleaning and feature engineering.
• Store the resulting data back in Amazon S3.
• Use Amazon Athena to infer the schemas and available columns.
• Use AWS Glue crawlers to infer the schemas and available columns.
• Use AWS Glue DataBrew for data cleaning and feature engineering.




A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**

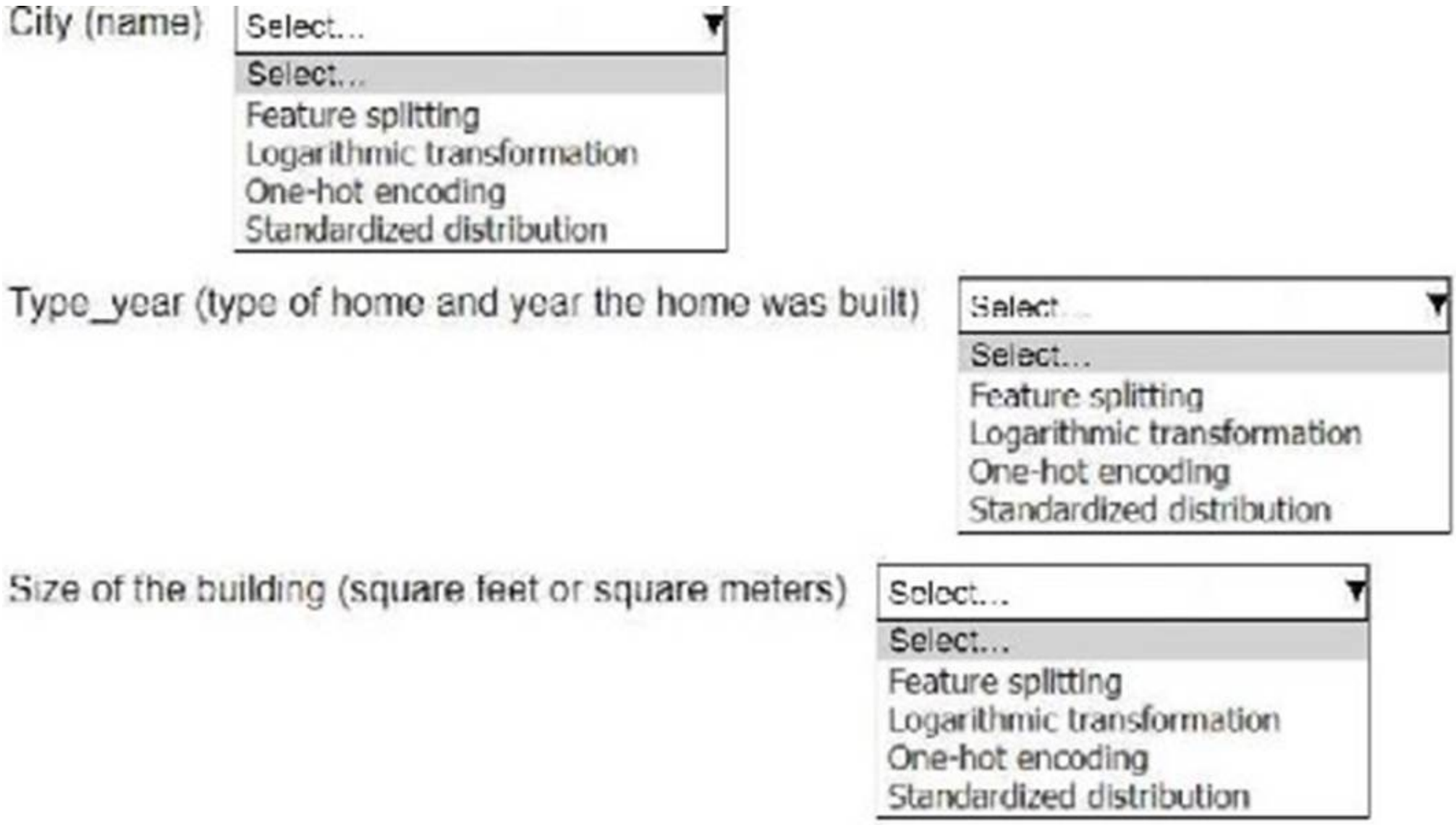



**NEW QUESTION 14**
An ML engineer is using a training job to fine-tune a deep learning model in Amazon SageMaker Studio. The ML engineer previously used the same pre-trained model with a similar dataset. The ML engineer expects vanishing gradient, underutilized GPU, and overfitting problems.
The ML engineer needs to implement a solution to detect these issues and to react in predefined ways when the issues occur. The solution also must provide comprehensive real-time metrics during the training.
Which solution will meet these requirements with the LEAST operational overhead?

A. Use TensorBoard to monitor the training jo
B. Publish the findings to an Amazon Simple Notification Service (Amazon SNS) topi
C. Create an AWS Lambda function to consume the findings and to initiate the predefined actions.
D. Use Amazon CloudWatch default metrics to gain insights about the training jo
E. Use themetrics to invoke an AWS Lambda function to initiate the predefined actions.
F. Expand the metrics in Amazon CloudWatch to include the gradients in each training ste
G. Use the metrics to invoke an AWS Lambda function to initiate the predefined actions.
H. Use SageMaker Debugger built-in rules to monitor the training jo
I. Configure the rules to initiate the predefined actions.

**Answer:** D

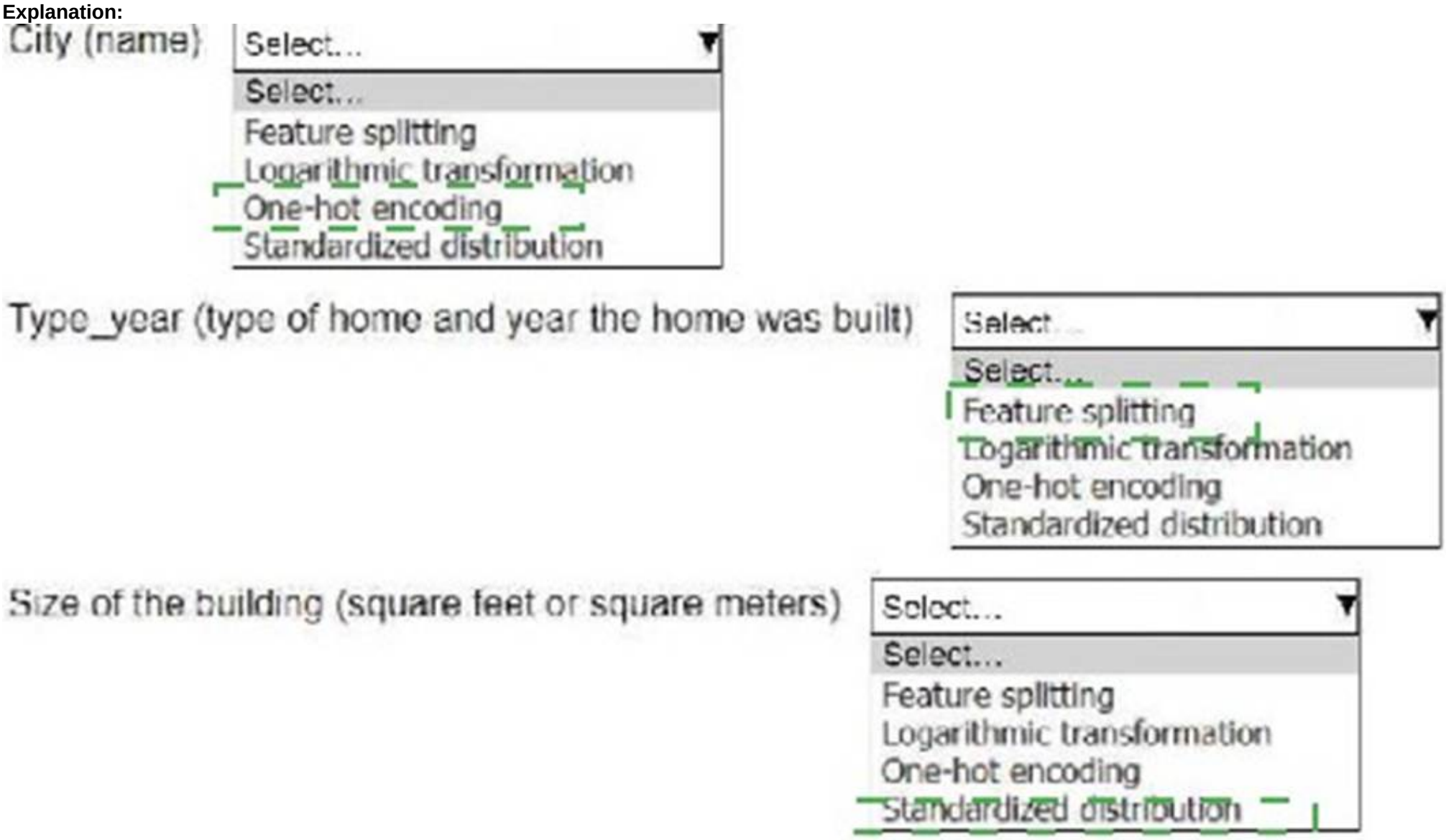
**NEW QUESTION 19**
HOTSPOT
An ML engineer is working on an ML model to predict the prices of similarly sized homes. The model will base predictions on several features The ML engineer will use the following feature engineering techniques to estimate the prices of the homes:
• Feature splitting
• Logarithmic transformation
• One-hot encoding
• Standardized distribution
Select the correct feature engineering techniques for the following list of features. Each feature engineering technique should be selected one time or not at all (Select three.)




A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**




**NEW QUESTION 21**
An ML engineer is developing a fraud detection model by using the Amazon SageMaker XGBoost algorithm. The model classifies transactions as either fraudulent or legitimate.
During testing, the model excels at identifying fraud in the training dataset. However, the model is inefficient at identifying fraud in new and unseen transactions.
What should the ML engineer do to improve the fraud detection for new transactions?

A. Increase the learning rate.
B. Remove some irrelevant features from the training dataset.
C. Increase the value of the max_depth hyperparameter.
D. Decrease the value of the max_depth hyperparameter.

**Answer:** D

**NEW QUESTION 23**
An ML engineer needs to create data ingestion pipelines and ML model deployment pipelines on AWS. All the raw data is stored in Amazon S3 buckets.
Which solution will meet these requirements?

A. Use Amazon Data Firehose to create the data ingestion pipeline
B. Use Amazon SageMaker Studio Classic to create the model deployment pipelines.
C. Use AWS Glue to create the data ingestion pipeline
D. Use Amazon SageMaker Studio Classic to create the model deployment pipelines.
E. Use Amazon Redshift ML to create the data ingestion pipeline
F. Use Amazon SageMaker Studio Classic to create the model deployment pipelines.
G. Use Amazon Athena to create the data ingestion pipeline
H. Use an Amazon SageMaker notebook to create the model deployment pipelines.

**Answer:** B

**NEW QUESTION 24**
A company has developed a new ML model. The company requires online model validation on 10% of the traffic before the company fully releases the model in production. The company uses an Amazon SageMaker endpoint behind an Application Load Balancer (ALB) to serve the model.
Which solution will set up the required online validation with the LEAST operational overhead?

A. Use production variants to add the new model to the existing SageMaker endpoin
B. Set the variant weight to 0.1 for the new mode
C. Monitor the number of invocations by using Amazon CloudWatch.
D. Use production variants to add the new model to the existing SageMaker endpoin
E. Set the variant weight to 1 for the new mode
F. Monitor the number of invocations by using Amazon CloudWatch.
G. Create a new SageMaker endpoin
H. Use production variants to add the new model to the new endpoin
I. Monitor the number of invocations by using Amazon CloudWatch.
J. Configure the ALB to route 10% of the traffic to the new model at the existing SageMaker endpoin
K. Monitor the number of invocations by using AWS CloudTrail.

**Answer:** A

**NEW QUESTION 26**
Case Study
A company is building a web-based AI application by using Amazon SageMaker. The application will provide the following capabilities and features: ML experimentation, training, a central model registry, model deployment, and model monitoring.
The application must ensure secure and isolated use of training data during the ML lifecycle. The training data is stored in Amazon S3.
The company needs to use the central model registry to manage different versions of models in the application.
Which action will meet this requirement with the LEAST operational overhead?

A. Create a separate Amazon Elastic Container Registry (Amazon ECR) repository for each model.
B. Use Amazon Elastic Container Registry (Amazon ECR) and unique tags for each model version.
C. Use the SageMaker Model Registry and model groups to catalogthe models.
D. Use the SageMaker Model Registry and unique tags for each model version.

**Answer:** C

**NEW QUESTION 28**
An ML engineer needs to use Amazon SageMaker to fine-tune a large language model (LLM) for text summarization. The ML engineer must follow a low-code no-code (LCNC) approach.
Which solution will meet these requirements?

A. Use SageMaker Studio to fine-tune an LLM that is deployed on Amazon EC2 instances.
B. Use SageMaker Autopilot to fine-tune an LLM that is deployed by a custom API endpoint.
C. Use SageMaker Autopilot to fine-tune an LLM that is deployed on Amazon EC2 instances.
D. Use SageMaker Autopilot to fine-tune an LLM that is deployed by SageMaker JumpStart.

**Answer:** D

**NEW QUESTION 29**
A company has deployed an ML model that detects fraudulent credit card transactions in real time in a banking application. The model uses Amazon SageMaker Asynchronous Inference. Consumers are reporting delays in receiving the inference results.
An ML engineer needs to implement a solution to improve the inference performance. The solution also must provide a notification when a deviation in model quality occurs.
Which solution will meet these requirements?

A. Use SageMaker real-time inference for inferenc
B. Use SageMaker Model Monitor for notifications about model quality.
C. Use SageMaker batch transform for inferenc

D. Use SageMaker Model Monitor for notifications about model quality.
E. Use SageMaker Serverless Inference for inferenc
F. Use SageMaker Inference Recommender for notifications about model quality.
G. Keep using SageMaker Asynchronous Inference for inferenc
H. Use SageMaker Inference Recommender for notifications about model quality.

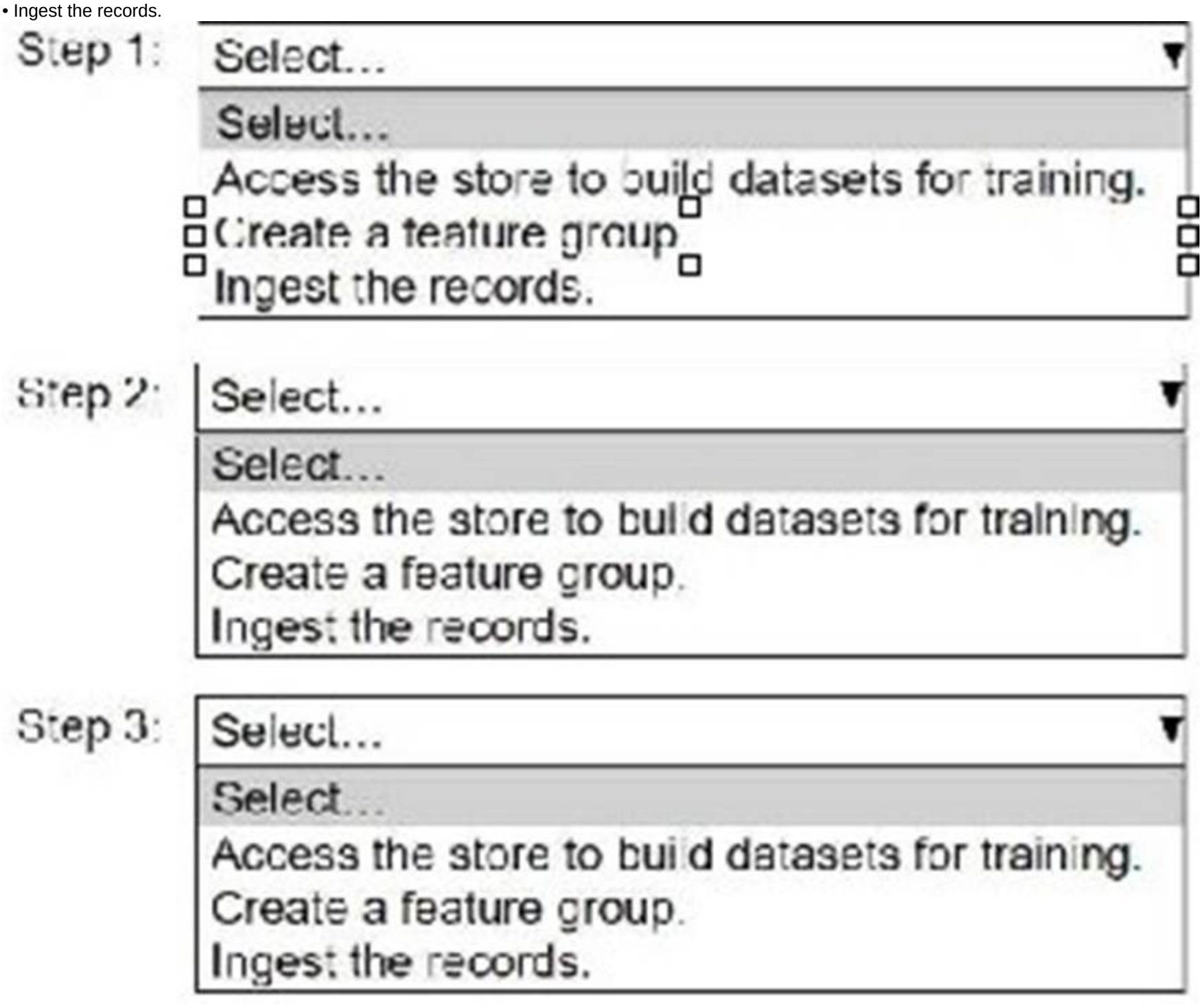
**Answer:** A

**NEW QUESTION 32**
An ML engineer has trained a neural network by using stochastic gradient descent (SGD). The neural network performs poorly on the test set. The values for training loss and validation loss remain high and show an oscillating pattern. The values decrease for a few epochs and then increase for a few epochs before repeating the same cycle.
What should the ML engineer do to improve the training process?

A. Introduce early stopping.
B. Increase the size of the test set.
C. Increase the learning rate.
D. Decrease the learning rate.

**Answer:** D

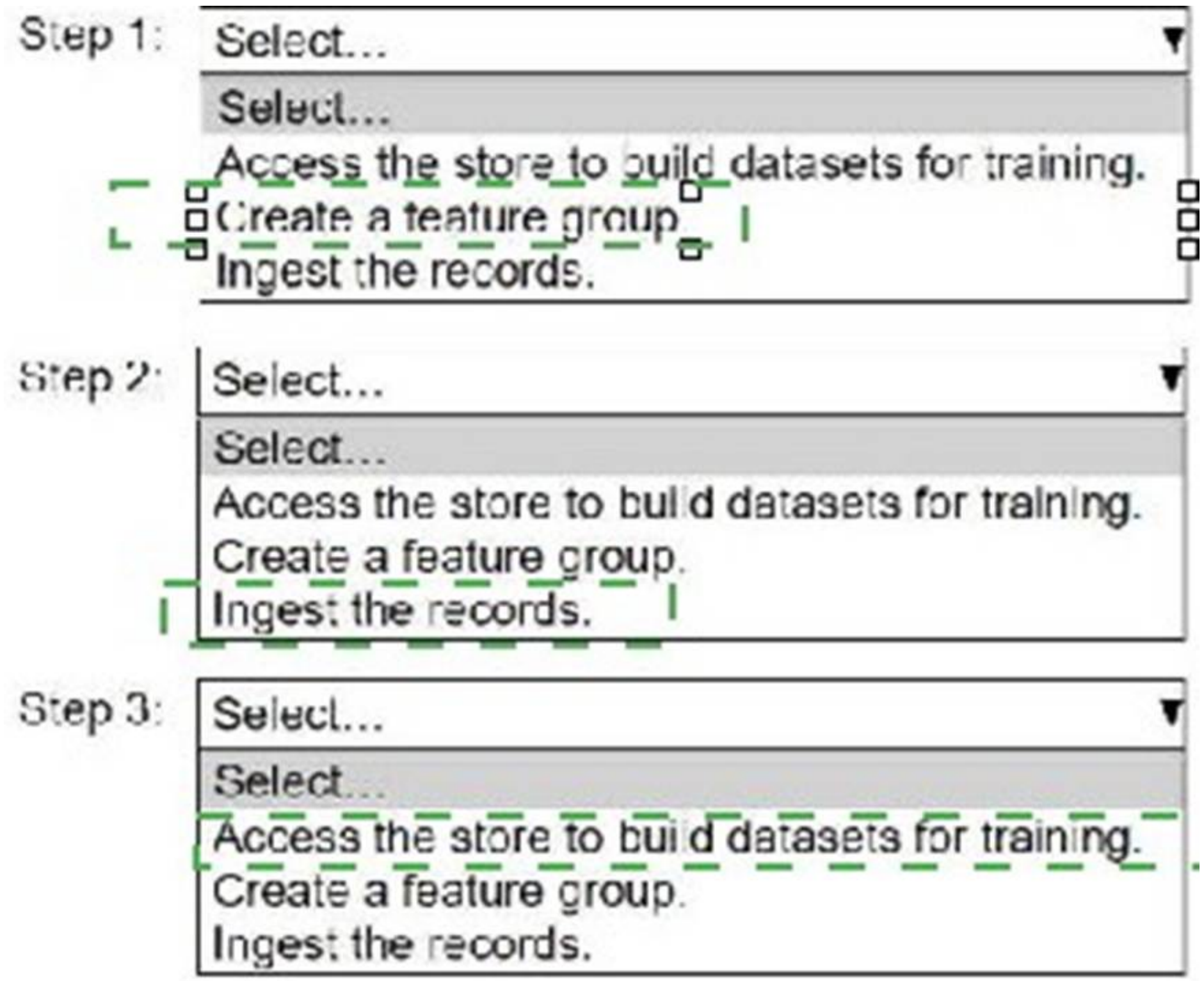
**NEW QUESTION 34**
HOTSPOT
An ML engineer needs to use Amazon SageMaker Feature Store to create and manage features to train a model.
Select and order the steps from the following list to create and use the features in Feature Store. Each step should be selected one time. (Select and order three.)
• Access the store to build datasets for training.
• Create a feature group.
• Ingest the records.




A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**




**NEW QUESTION 36**
A company has a binary classification model in production. An ML engineer needs to develop a new version of the model.
The new model version must maximize correct predictions of positive labels and negative labels. The ML engineer must use a metric to recalibrate the model to meet these requirements.
Which metric should the ML engineer use for the model recalibration?

A. Accuracy
B. Precision
C. Recall
D. Specificity

**Answer:** A

**NEW QUESTION 41**
An ML engineer trained an ML model on Amazon SageMaker to detect automobile accidents from dosed-circuit TV footage. The ML engineer used SageMaker Data Wrangler to create a training dataset of images of accidents and non-accidents.
The model performed well during training and validation. However, the model is underperforming in production because of variations in the quality of the images from various cameras.
Which solution will improve the model's accuracy in the LEAST amount of time?

A. Collect more images from all the camera
B. Use Data Wrangler to prepare a new trainingdataset.
C. Recreate the training dataset by using the Data Wrangler corrupt image transfor
D. Specify the impulse noise option.
E. Recreate the training dataset by using the Data Wrangler enhance image contrast transfor
F. Specify the Gamma contrast option.
G. Recreate the training dataset by using the Data Wrangler resize image transfor
H. Crop all images to the same size.

**Answer:** B

**NEW QUESTION 43**
An ML engineer normalized training data by using min-max normalization in AWS Glue DataBrew. The ML engineer must normalize the production inference data in the same way as the training data before passing the production inference data to the model for predictions.

Which solution will meet this requirement?

A. Apply statistics from a well-known dataset to normalize the production samples.
B. Keep the min-max normalization statistics from the training se
C. Use these values to normalize the production samples.
D. Calculate a new set of min-max normalization statistics from a batch of production sample
E. Use these values to normalize all the production samples.
F. Calculate a new set of min-max normalization statistics from each production sampl
G. Use these values to normalize all the production samples.

**Answer:** B

**NEW QUESTION 47**
A company regularly receives new training data from the vendor of an ML model. The vendor delivers cleaned and prepared data to the company's Amazon S3 bucket every 3-4 days.
The company has an Amazon SageMaker pipeline to retrain the model. An ML engineer needs to implement a solution to run the pipeline when new data is uploaded to the S3 bucket.
Which solution will meet these requirements with the LEAST operational effort?

A. Create an S3 Lifecycle rule to transfer the data to the SageMaker training instance and to initiate training.
B. Create an AWS Lambda function that scans the S3 bucke
C. Program the Lambda function to initiate the pipeline when new data is uploaded.
D. Create an Amazon EventBridge rule that has an event pattern that matches the S3 uploa
E. Configure the pipeline as the target of the rule.
F. Use Amazon Managed Workflows for Apache Airflow (Amazon MWAA) to orchestrate the pipeline when new data is uploaded.A company wants to develop an ML model by using tabular data from its customer
G. The data contains meaningful ordered features with sensitive information that should not be discarde
H. An ML engineer must ensure that the sensitive data is masked before another team starts to build the model.Which solution will meet these requirements?
I. Use Amazon Made to categorize the sensitive data.
J. Prepare the data by using AWS Glue DataBrew.
K. Run an AWS Batch job to change the sensitive data to random values.
L. Run an Amazon EMR job to change the sensitive data to random values.

**Answer:** B

**NEW QUESTION 50**
A company has implemented a data ingestion pipeline for sales transactions from its ecommerce website. The company uses Amazon Data Firehose to ingest data into Amazon OpenSearch Service. The buffer interval of the Firehose stream is set for 60 seconds. An OpenSearch linear model generates real-time sales forecasts based on the data and presents the data in an OpenSearch dashboard.
The company needs to optimize the data ingestion pipeline to support sub-second latency for the real-time dashboard.
Which change to the architecture will meet these requirements?

A. Use zero buffering in the Firehose strea
B. Tune the batch size that is used in the PutRecordBatch operation.
C. Replace the Firehose stream with an AWS DataSync tas
D. Configure the task with enhanced fan-out consumers.
E. Increase the buffer interval of the Firehose stream from 60 seconds to 120 seconds.
F. Replace the Firehose stream with an Amazon Simple Queue Service (Amazon SQS) queue.

**Answer:** A

**NEW QUESTION 51**
A company needs to host a custom ML model to perform forecast analysis. The forecast analysis will occur with predictable and sustained load during the same 2-hour period every day.
Multiple invocations during the analysis period will require quick responses. The company needs AWS to manage the underlying infrastructure and any auto scaling activities.
Which solution will meet these requirements?

A. Schedule an Amazon SageMaker batch transform job by using AWS Lambda.
B. Configure an Auto Scaling group of Amazon EC2 instances to use scheduled scaling.
C. Use Amazon SageMaker Serverless Inference with provisioned concurrency.
D. Run the model on an Amazon Elastic Kubernetes Service (Amazon EKS) cluster on Amazon EC2 with pod auto scaling.

**Answer:** C

**NEW QUESTION 52**
Case study
An ML engineer is developing a fraud detection model on AWS. The training dataset includes transaction logs, customer profiles, and tables from an on-premises MySQL database. The transaction logs and customer profiles are stored in Amazon S3.
The dataset has a class imbalance that affects the learning of the model's algorithm. Additionally, many of the features have interdependencies. The algorithm is not capturing all the desired underlying patterns in the data.
The ML engineer needs to use an Amazon SageMaker built-in algorithm to train the model. Which algorithm should the ML engineer use to meet this requirement?

A. LightGBM
B. Linear learner
C. -means clustering
D. Neural Topic Model (NTM)

**Answer:** B

**NEW QUESTION 54**
An ML engineer needs to deploy ML models to get inferences from large datasets in an asynchronous manner. The ML engineer also needs to implement scheduled monitoring of the data quality of the models. The ML engineer must receive alerts when changes in data quality occur.
Which solution will meet these requirements?

A. Deploy the models by using scheduled AWS Glue job
B. Use Amazon CloudWatch alarms to monitor the data quality and to send alerts.
C. Deploy the models by using scheduled AWS Batch job
D. Use AWS CloudTrail to monitor the data quality and to send alerts.
E. Deploy the models by using Amazon Elastic Container Service (Amazon ECS) on AWS Fargat
F. Use Amazon EventBridge to monitor the data quality and to send alerts.
G. Deploy the models by using Amazon SageMaker batch transfor
H. Use SageMaker Model Monitor to monitor the data quality and to send alerts.

**Answer:** D

**NEW QUESTION 59**
A company wants to improve the sustainability of its ML operations.
Which actions will reduce the energy usage and computational resources that are associated with the company's training jobs? (Choose two.)

A. Use Amazon SageMaker Debugger to stop training jobs when non-converging conditions are detected.
B. Use Amazon SageMaker Ground Truth for data labeling.
C. Deploy models by using AWS Lambda functions.
D. Use AWS Trainium instances for training.
E. Use PyTorch or TensorFlow with the distributed training option.

**Answer:** AD

**NEW QUESTION 60**
An ML engineer has developed a binary classification model outside of Amazon SageMaker. The ML engineer needs to make the model accessible to a SageMaker Canvas user for additional tuning.
The model artifacts are stored in an Amazon S3 bucket. The ML engineer and the Canvas user are part of the same SageMaker domain.
Which combination of requirements must be met so that the ML engineer can share the model with the Canvas user? (Choose two.)

A. The ML engineer and the Canvas user must be in separate SageMaker domains.
B. The Canvas user must have permissions to access the S3 bucket where the model artifacts are stored.
C. The model must be registered in the SageMaker Model Registry.
D. The ML engineer must host the model on AWS Marketplace.
E. The ML engineer must deploy the model to a SageMaker endpoint.

**Answer:** BC

**NEW QUESTION 61**
A company that has hundreds of data scientists is using Amazon SageMaker to create ML models. The models are in model groups in the SageMaker Model Registry.
The data scientists are grouped into three categories: computer vision, natural language processing (NLP), and speech recognition. An ML engineer needs to implement a solution to organize the existing models into these groups to improve model discoverability at scale. The solution must not affect the integrity of the model artifacts and their existing groupings.
Which solution will meet these requirements?

A. Create a custom tag for each of the three categorie
B. Add the tags to the model packages in the SageMaker Model Registry.
C. Create a model group for each categor
D. Move the existing models into these category model groups.
E. Use SageMaker ML Lineage Tracking to automatically identify and tag which model groups should contain the models.
F. Create a Model Registry collection for each of the three categorie
G. Move the existing model groups into the collections.

**Answer:** A

**NEW QUESTION 63**
.....

# Thank You for Trying Our Product

## We offer two products:

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

2nd - Questons and Answers in PDF Format

## MLA-C01 Practice Exam Features:

* MLA-C01 Questions and Answers Updated Frequently

* MLA-C01 Practice Questions Verified by Expert Senior Certified Staff

* MLA-C01 Most Realistic Questions that Guarantee you a Pass on Your FirstTry

* MLA-C01 Practice Test Questions in Multiple Choice Formats and Updatesfor 1 Year