# Exam Questions MLA-C01

AWS Certified Machine Learning Engineer - Associate

**https://www.2passeasy.com/dumps/MLA-C01/**

**NEW QUESTION 1**
An ML engineer needs to process thousands of existing CSV objects and new CSV objects that are uploaded. The CSV objects are stored in a central Amazon S3 bucket and have the same number of columns. One of the columns is a transaction date. The ML engineer must query the data based on the transaction date.
Which solution will meet these requirements with the LEAST operational overhead?

A. Use an Amazon Athena CREATE TABLE AS SELECT (CTAS) statement to create a table based on the transaction date from data in the central S3 bucke
B. Query the objects from the table.
C. Create a new S3 bucket for processed dat
D. Set up S3 replication from the central S3 bucket to the new S3 bucke
E. Use S3 Object Lambda to query the objects based on transaction date.
F. Create a new S3 bucket for processed dat
G. Use AWS Glue for Apache Spark to create a job to query the CSV objects based on transaction dat
H. Configure the job to store the results in the new S3 bucke
I. Query the objects from the new S3 bucket.
J. Create a new S3 bucket for processed dat
K. Use Amazon Data Firehose to transfer the data from the central S3 bucket to the new S3 bucke
L. Configure Firehose to run an AWS Lambda function to query the data based on transaction date.

**Answer:** A

**NEW QUESTION 2**
A company is using an Amazon Redshift database as its single data source. Some of the data is sensitive.
A data scientist needs to use some of the sensitive data from the database. An ML engineer must give the data scientist access to the data without transforming the source data and without storing anonymized data in the database.
Which solution will meet these requirements with the LEAST implementation effort?

A. Configure dynamic data masking policies to control how sensitive data is shared with the data scientist at query time.
B. Create a materialized view with masking logic on top of the databas
C. Grant the necessary read permissions to the data scientist.
D. Unload the Amazon Redshift data to Amazon S3. Use Amazon Athena to create schema-on-read with masking logi
E. Share the view with the data scientist.
F. Unload the Amazon Redshift data to Amazon S3. Create an AWS Glue job to anonymize the dat
G. Share the dataset with the data scientist.

**Answer:** A

**NEW QUESTION 3**
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 4**
A company has a large collection of chat recordings from customer interactions after a product release. An ML engineer needs to create an ML model to analyze the chat data. The ML engineer needs to determine the success of the product by reviewing customer sentiments about the product.
Which action should the ML engineer take to complete the evaluation in the LEAST amount of time?

A. Use Amazon Rekognition to analyze sentiments of the chat conversations.
B. Train a Naive Bayes classifier to analyze sentiments of the chat conversations.
C. Use Amazon Comprehend to analyze sentiments of the chat conversations.
D. Use random forests to classify sentiments of the chat conversations.

**Answer:** C

**NEW QUESTION 5**
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 6**
A company has an ML model that generates text descriptions based on images that customers upload to the company's website. The images can be up to 50 MB in total size.
An ML engineer decides to store the images in an Amazon S3 bucket. The ML engineer must implement a processing solution that can scale to accommodate changes in demand.
Which solution will meet these requirements with the LEAST operational overhead?

A. Create an Amazon SageMaker batch transform job to process all the images in the S3 bucket.
B. Create an Amazon SageMaker Asynchronous Inference endpoint and a scaling polic
C. Run a script to make an inference request for each image.
D. Create an Amazon Elastic Kubernetes Service (Amazon EKS) cluster that uses Karpenter for auto scalin
E. Host the model on the EKS cluste
F. Run a script to make an inference request for each image.
G. Create an AWS Batch job that uses an Amazon Elastic Container Service (Amazon ECS) cluste
H. Specify a list of images to process for each AWS Batch job.

**Answer:** B

**NEW QUESTION 7**
An ML engineer is training a simple neural network model. The ML engineer tracks the performance of the model over time on a validation dataset. The model's performance improves substantially at first and then degrades after a specific number of epochs.
Which solutions will mitigate this problem? (Choose two.)

A. Enable early stopping on the model.
B. Increase dropout in the layers.
C. Increase the number of layers.
D. Increase the number of neurons.
E. Investigate and reduce the sources of model bias.

**Answer:** AB

**NEW QUESTION 8**
A company wants to reduce the cost of its containerized ML applications. The applications use ML models that run on Amazon EC2 instances, AWS Lambda functions, and an Amazon Elastic Container Service (Amazon ECS) cluster. The EC2 workloads and ECS workloads use Amazon Elastic Block Store (Amazon EBS) volumes to save predictions and artifacts.
An ML engineer must identify resources that are being used inefficiently. The ML engineer also must generate recommendations to reduce the cost of these resources.
Which solution will meet these requirements with the LEAST development effort?

A. Create code to evaluate each instance's memory and compute usage.
B. Add cost allocation tags to the resource
C. Activate the tags in AWS Billing and Cost Management.
D. Check AWS CloudTrail event history for the creation of the resources.
E. Run AWS Compute Optimizer.

**Answer:** D

**NEW QUESTION 9**
A company uses Amazon SageMaker Studio to develop an ML model. The company has a single SageMaker Studio domain. An ML engineer needs to implement a solution that provides an automated alert when SageMaker compute costs reach a specific threshold.
Which solution will meet these requirements?

A. Add resource tagging by editing the SageMaker user profile in the SageMaker domai
B. Configure AWS Cost Explorer to send an alert when the threshold is reached.
C. Add resource tagging by editing the SageMaker user profile in the SageMaker domai
D. Configure AWS Budgets to send an alert when the threshold is reached.
E. Add resource tagging by editing each user's IAM profil
F. Configure AWS Cost Explorer to send an alert when the threshold is reached.
G. Add resource tagging by editing each user's IAM profil
H. Configure AWS Budgets to send an alert when the threshold is reached.

**Answer:** B

**NEW QUESTION 10**
A company has used Amazon SageMaker to deploy a predictive ML model in production. The company is using SageMaker Model Monitor on the model. After a model update, an ML engineer notices data quality issues in the Model Monitor checks.

What should the ML engineer do to mitigate the data quality issues that Model Monitor has identified?

A. Adjust the model's parameters and hyperparameters.
B. Initiate a manual Model Monitor job that uses the most recent production data.
C. Create a new baseline from the latest datase
D. Update Model Monitor to use the new baseline for evaluations.
E. Include additional data in the existing training set for the mode
F. Retrain and redeploy the model.

**Answer:** C

**NEW QUESTION 10**
A company has a Retrieval Augmented Generation (RAG) application that uses a vector database to store embeddings of documents. The company must migrate the application to AWS and must implement a solution that provides semantic search of text files. The company has already migrated the text repository to an Amazon S3 bucket.
Which solution will meet these requirements?

A. Use an AWS Batch job to process the files and generate embedding
B. Use AWS Glue to store the embedding
C. Use SQL queries to perform the semantic searches.
D. Use a custom Amazon SageMaker notebook to run a custom script to generate embedding
E. Use SageMaker Feature Store to store the embedding
F. Use SQL queries to perform the semantic searches.
G. Use the Amazon Kendra S3 connector to ingest the documents from the S3 bucket into Amazon Kendr
H. Query Amazon Kendra to perform the semantic searches.
I. Use an Amazon Textract asynchronous job to ingest the documents from the S3 bucke
J. Query Amazon Textract to perform the semantic searches.

**Answer:** C

**NEW QUESTION 11**
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 14**
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.

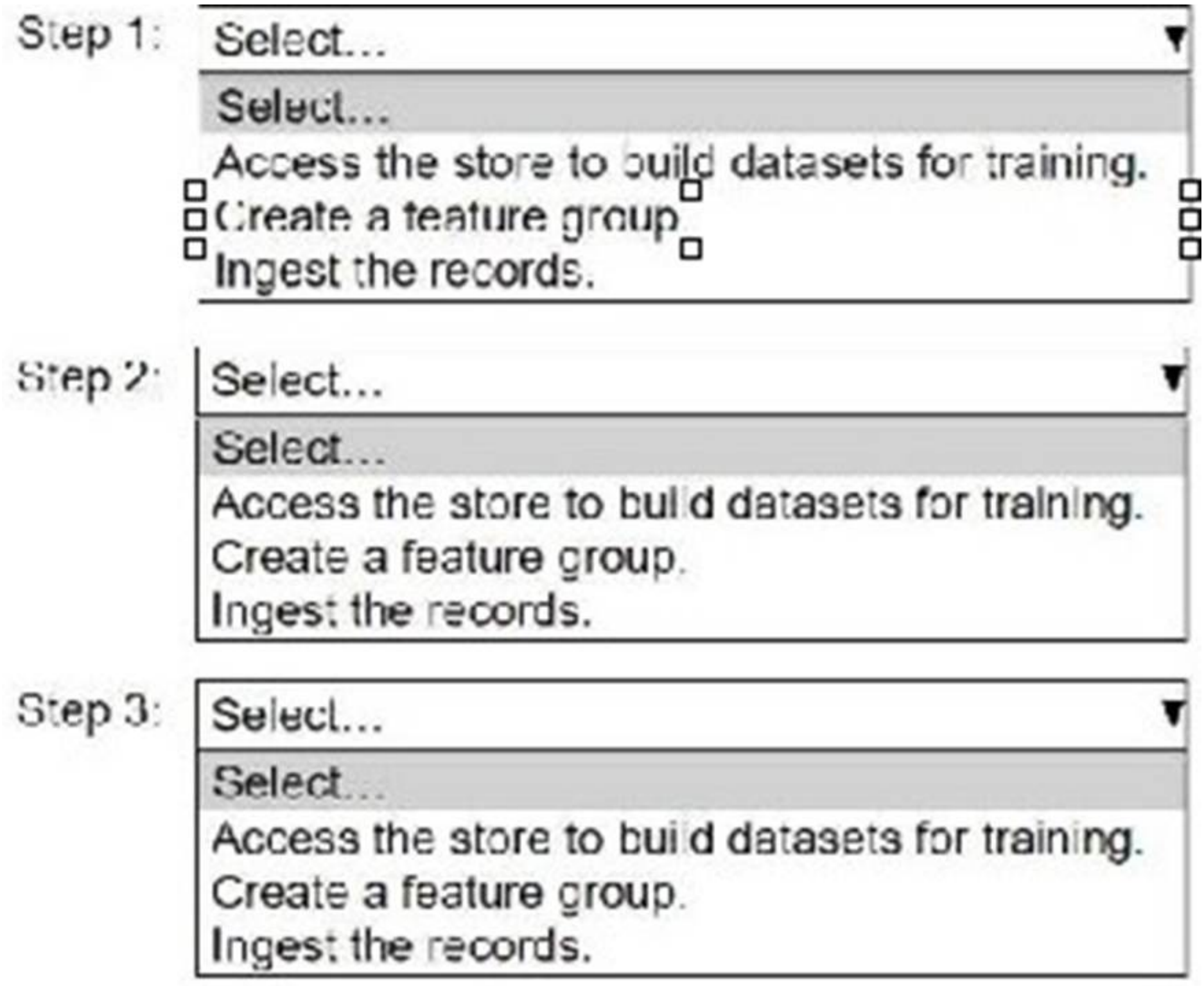
**Answer:** D

**NEW QUESTION 19**
A company is using ML to predict the presence of a specific weed in a farmer's field. The company is using the Amazon SageMaker linear learner built-in algorithm with a value of multiclass_dassifier for the predictorjype hyperparameter.
What should the company do to MINIMIZE false positives?

A. Set the value of the weight decay hyperparameter to zero.
B. Increase the number of training epochs.
C. Increase the value of the target_precision hyperparameter.
D. Change the value of the predictorjype hyperparameter to regressor.

**Answer:** C

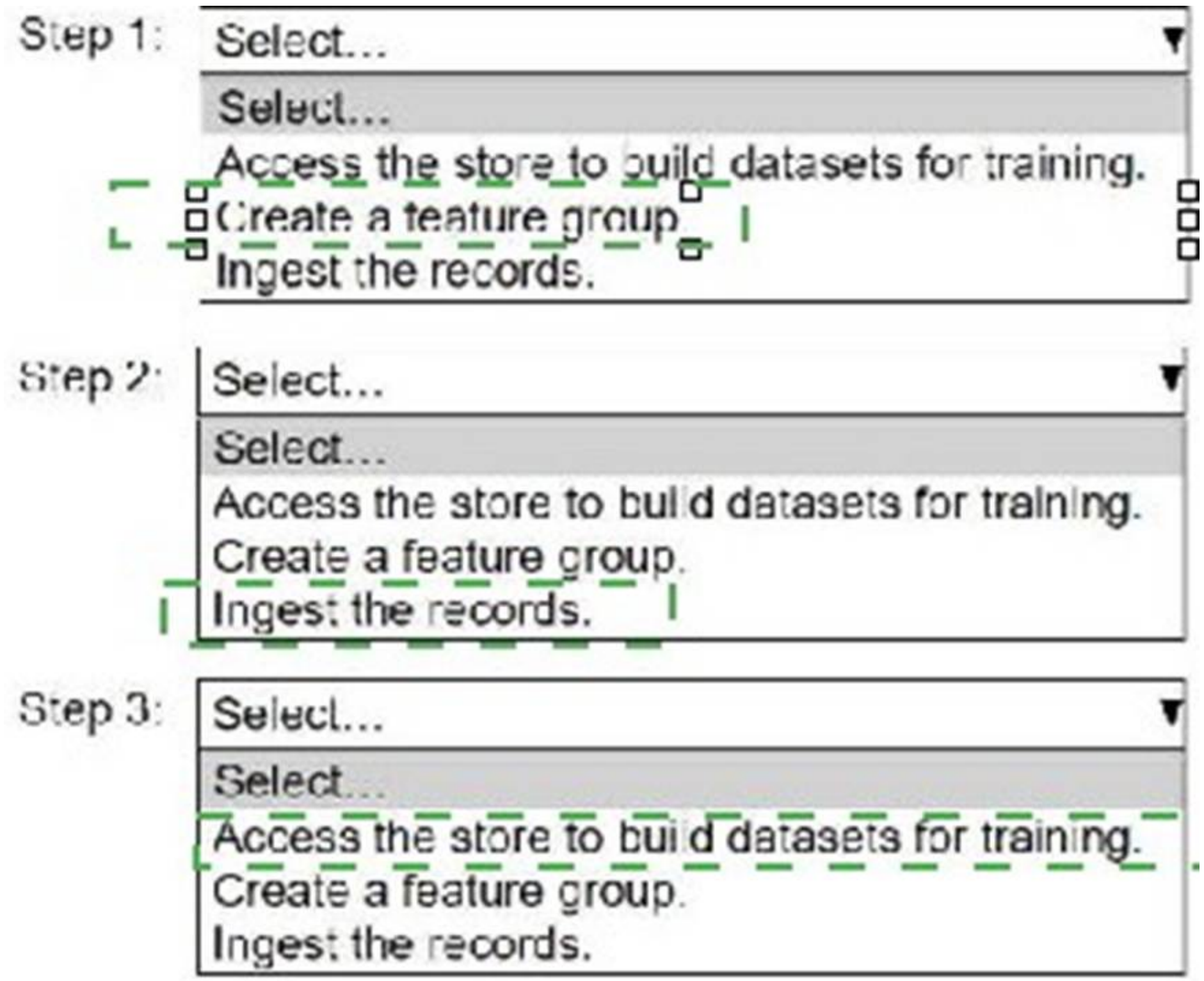
**NEW QUESTION 23**
HOTSPOT
An ML engineer needs to use Amazon SageMaker Feature Store to create and manage features to train a model.
Select and order the steps from the following list to create and use the features in Feature Store. Each step should be selected one time. (Select and order three.)
• 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 25**
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 28**
An ML engineer needs to use AWS services to identify and extract meaningful unique keywords from documents.
Which solution will meet these requirements with the LEAST operational overhead?

A. Use the Natural Language Toolkit (NLTK) library on Amazon EC2 instances for text pre- processin
B. Use the Latent Dirichlet Allocation (LDA) algorithm to identify and extract relevant keywords.
C. Use Amazon SageMaker and the BlazingText algorith
D. Apply custom pre-processing steps for stemming and removal of stop word
E. Calculate term frequency-inverse document frequency (TF-IDF) scores to identify and extract relevant keywords.
F. Store the documents in an Amazon S3 bucke
G. Create AWS Lambda functions to process the documents and to run Python scripts for stemming and removal of stop word
H. Use bigram and trigram techniques to identify and extract relevant keywords.
I. Use Amazon Comprehend custom entity recognition and key phrase extraction to identify and extract relevant keywords.

**Answer:** D

**NEW QUESTION 31**
An ML engineer is evaluating several ML models and must choose one model to use in production. The cost of false negative predictions by the models is much higher than the cost of false positive predictions.
Which metric finding should the ML engineer prioritize the MOST when choosing the model?

A. Low precision
B. High precision
C. Low recall
D. High recall

**Answer:** D

**NEW QUESTION 33**
An ML engineer needs to use AWS CloudFormation to create an ML model that an Amazon SageMaker endpoint will host.
Which resource should the ML engineer declare in the CloudFormation template to meet this requirement?

A. AWS::SageMaker::Model
B. AWS::SageMaker::Endpoint
C. AWS::SageMaker::NotebookInstance
D. AWS::SageMaker::Pipeline

**Answer:** A

**NEW QUESTION 38**
An ML engineer needs to use an ML model to predict the price of apartments in a specific location.
Which metric should the ML engineer use to evaluate the model's performance?

A. Accuracy
B. Area Under the ROC Curve (AUC)
C. F1 score
D. Mean absolute error (MAE)

**Answer:** D

**NEW QUESTION 40**
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 42**
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 run an on-demand workflow to monitor bias drift for models that are deployed to real-time endpoints from the application.
Which action will meet this requirement?

A. Configure the application to invoke an AWS Lambda function that runs a SageMaker Clarify job.
B. Invoke an AWS Lambda function to pull the sagemaker-model-monitor-analyzer built-in SageMaker image.
C. Use AWS Glue Data Quality to monitor bias.
D. Use SageMaker notebooks to compare the bias.

**Answer:** A

**NEW QUESTION 47**
An ML engineer needs to use an Amazon EMR cluster to process large volumes of data in batches. Any data loss is unacceptable.
Which instance purchasing option will meet these requirements MOST cost-effectively?

A. Run the primary node, core nodes, and task nodes on On-Demand Instances.
B. Run the primary node, core nodes, and task nodes on Spot Instances.
C. Run the primary node on an On-Demand Instanc
D. Run the core nodes and task nodes on Spot Instances.
E. Run the primary node and core nodes on On-Demand Instance
F. Run the task nodes on Spot Instances.

**Answer:** D

**NEW QUESTION 52**
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 55**
.....

# THANKS FOR TRYING THE DEMO OF OUR PRODUCT

Visit Our Site to Purchase the Full Set of Actual MLA-C01 Exam Questions With Answers.

We Also Provide Practice Exam Software That Simulates Real Exam Environment And Has Many Self-Assessment Features. Order the MLA-C01 Product From:

**https://www.2passeasy.com/dumps/MLA-C01/**

## Money Back Guarantee

## 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