Amazon-Web-Services

Exam Questions MLA-C01

AWS Certified Machine Learning Engineer - Associate

NEW QUESTION 1
A company's ML engineer has deployed an ML model for sentiment analysis to an Amazon SageMaker endpoint. The ML engineer needs to explain to company stakeholders how the model makes predictions.
Which solution will provide an explanation for the model's predictions?

A. Use SageMaker Model Monitor on the deployed model.
B. Use SageMaker Clarify on the deployed model.
C. Show the distribution of inferences from A/ testing in Amazon CloudWatch.
D. Add a shadow endpoin
E. Analyze prediction differences on samples.

Answer: B

NEW QUESTION 2
A company is using Amazon SageMaker and millions of files to train an ML model. Each file is several megabytes in size. The files are stored in an Amazon S3 bucket. The company needs to improve training performance.
Which solution will meet these requirements in the LEAST amount of time?

A. Transfer the data to a new S3 bucket that provides S3 Express One Zone storag
B. Adjust the training job to use the new S3 bucket.
C. Create an Amazon FSx for Lustre file syste
D. Link the file system to the existing S3 bucke
E. Adjust the training job to read from the file system.
F. Create an Amazon Elastic File System (Amazon EFS) file syste
G. Transfer the existing data to the file syste
H. Adjust the training job to read from the file system.
I. Create an Amazon ElastiCache (Redis OSS) cluste
J. Link the Redis OSS cluster to the existing S3 bucke
K. Stream the data from the Redis OSS cluster directly to the training job.

Answer: B

NEW QUESTION 3
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 4
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 5
A company is planning to use Amazon SageMaker to make classification ratings that are based on images. The company has 6 of training data that is stored on an Amazon FSx for NetApp ONTAP system virtual machine (SVM). The SVM is in the same VPC as SageMaker.
An ML engineer must make the training data accessible for ML models that are in the SageMaker environment.
Which solution will meet these requirements?

A. Mount the FSx for ONTAP file system as a volume to the SageMaker Instance.
B. Create an Amazon S3 bucke
C. Use Mountpoint for Amazon S3 to link the S3 bucket to the FSx for ONTAP file system.
D. Create a catalog connection from SageMaker Data Wrangler to the FSx for ONTAP file system.
E. Create a direct connection from SageMaker Data Wrangler to the FSx for ONTAP file system.

Answer: A

NEW QUESTION 6
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 7
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 8
A company has a team of data scientists who use Amazon SageMaker notebook instances to test ML models. When the data scientists need new permissions, the company attaches the permissions to each individual role that was created during the creation of the SageMaker notebook instance.
The company needs to centralize management of the team's permissions. Which solution will meet this requirement?

A. Create a single IAM role that has the necessary permission
B. Attach the role to each notebook instance that the team uses.
C. Create a single IAM grou
D. Add the data scientists to the grou
E. Associate the group with each notebook instance that the team uses.
F. Create a single IAM use
G. Attach the AdministratorAccess AWS managed IAM policy to the use
H. Configure each notebook instance to use the IAM user.
I. Create a single IAM grou
J. Add the data scientists to the grou
K. Create an IAM rol
L. Attach the AdministratorAccess AWS managed IAM policy to the rol
M. Associate the role with the grou
N. Associate the group with each notebook instance that the team uses.

Answer: A

NEW QUESTION 9
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 10
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 10
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 15
A company is gathering audio, video, and text data in various languages. The company needs to use a large language model (LLM) to summarize the gathered data that is in Spanish.
Which solution will meet these requirements in the LEAST amount of time?

A. Train and deploy a model in Amazon SageMaker to convert the data into English tex
B. Train and deploy an LLM in SageMaker to summarize the text.
C. Use Amazon Transcribe and Amazon Translate to convert the data into English tex
D. Use Amazon Bedrock with the Jurassic model to summarize the text.
E. Use Amazon Rekognition and Amazon Translate to convert the data into English tex
F. Use Amazon Bedrock with the Anthropic Claude model to summarize the text.
G. Use Amazon Comprehend and Amazon Translate to convert the data into English tex
H. Use Amazon Bedrock with the Stable Diffusion model to summarize the text.

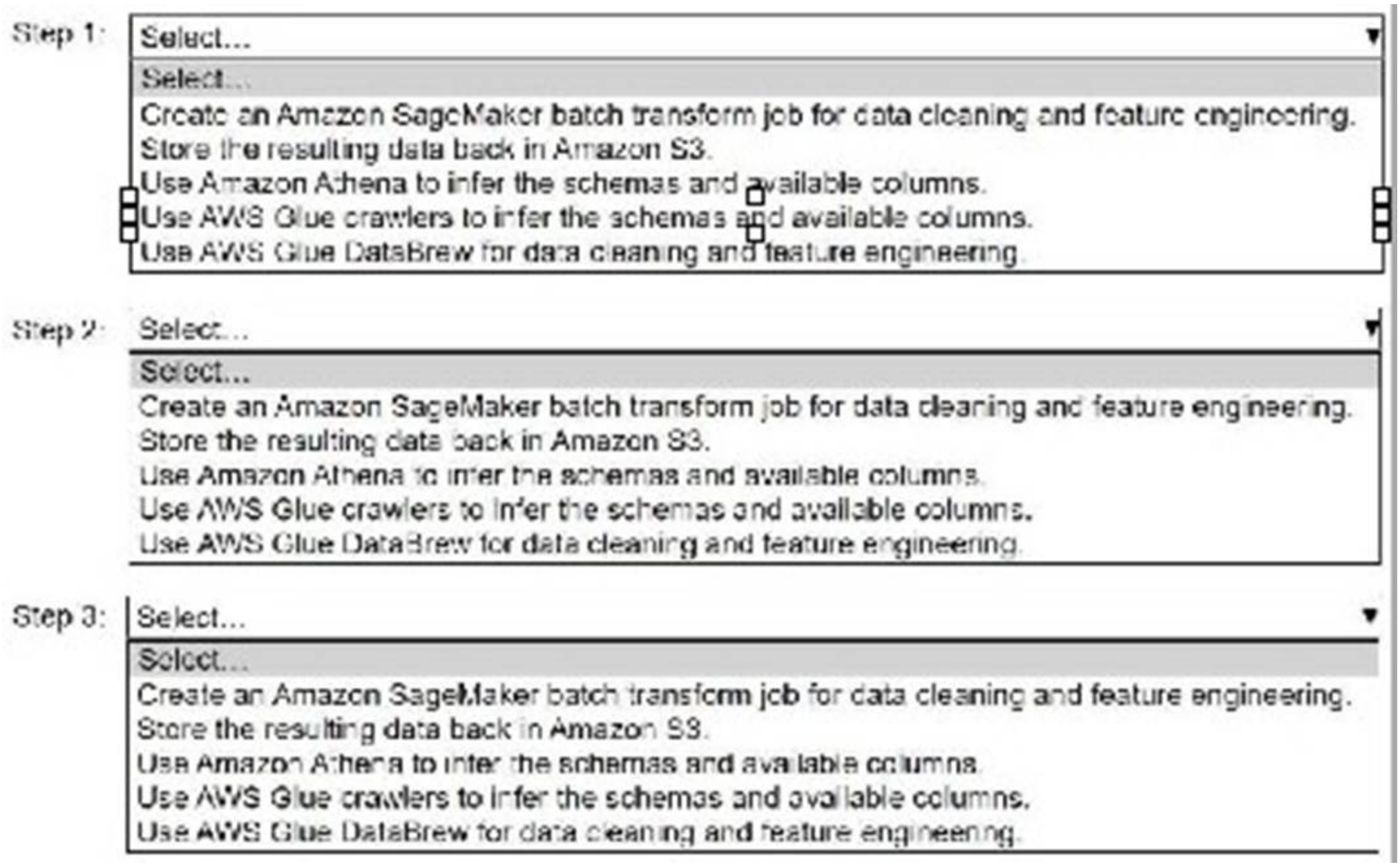
Answer: B

NEW QUESTION 19
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

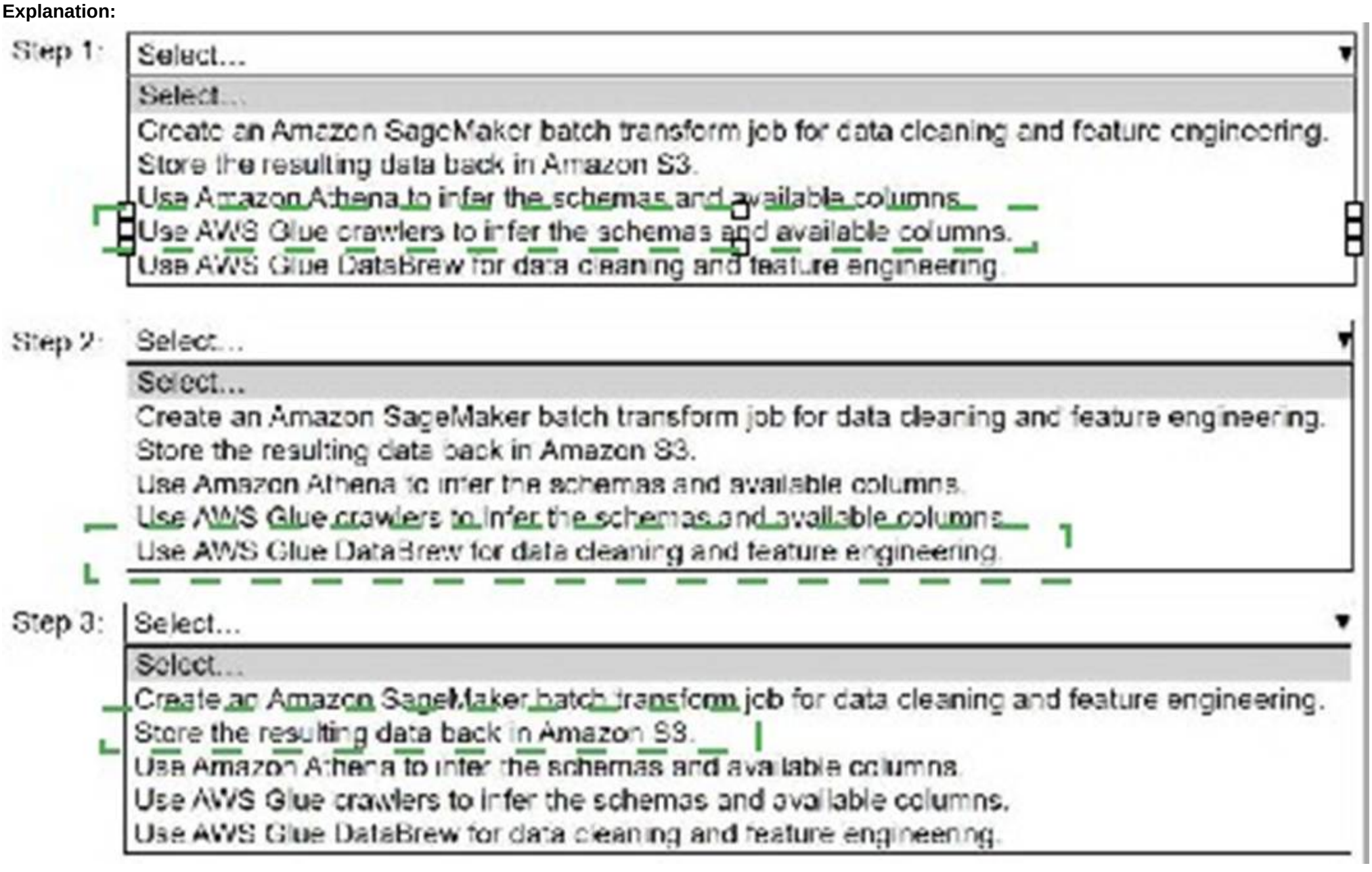
NEW QUESTION 24
HOTSPOT
A company stores historical data in .csv files in Amazon S3. Only some of the rows and columns in the .csv files are populated. The columns are not labeled. An ML engineer needs to prepare and store the data so that the company can use the data to train ML models.
Select and order 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:



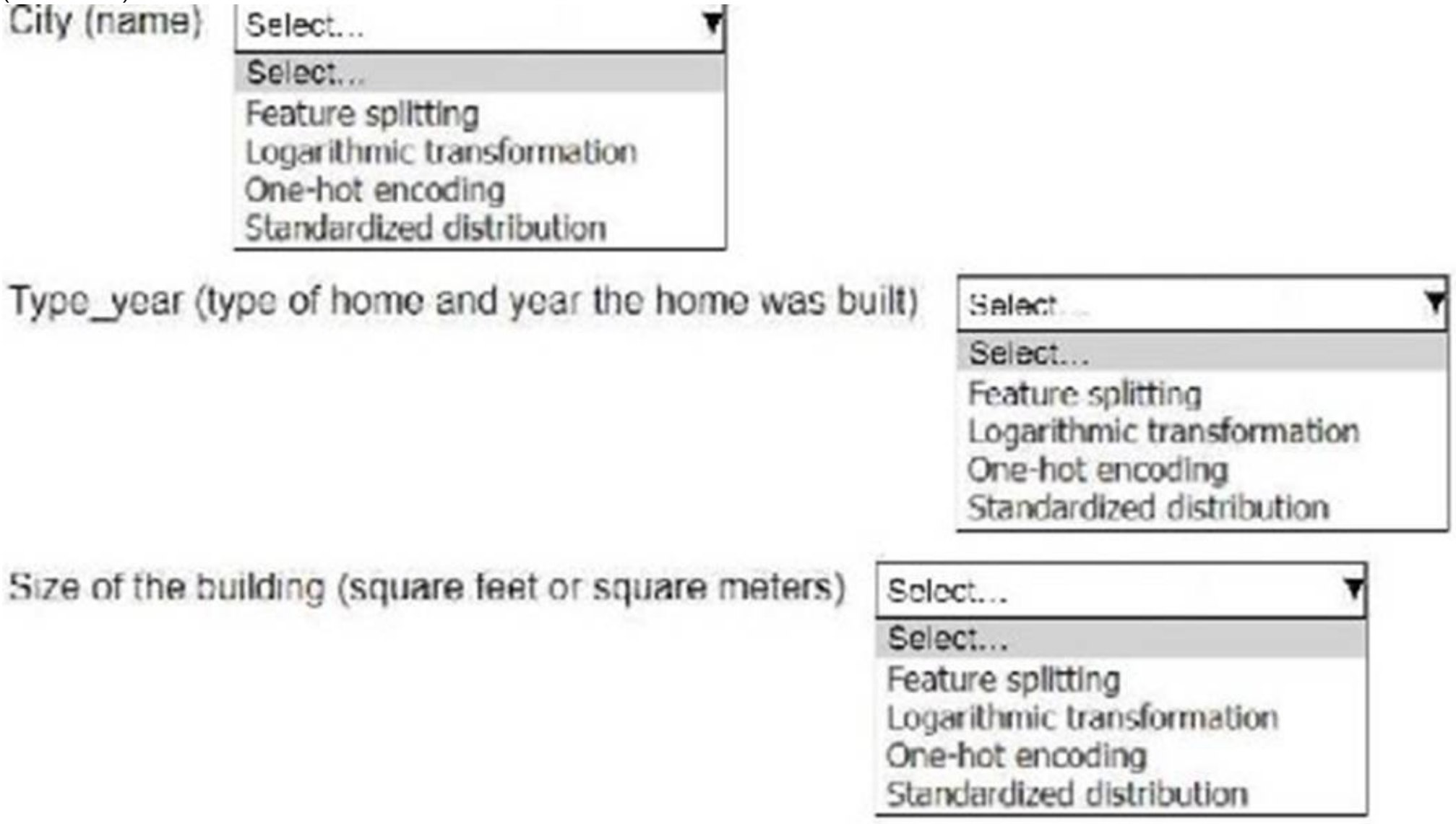
NEW QUESTION 25

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

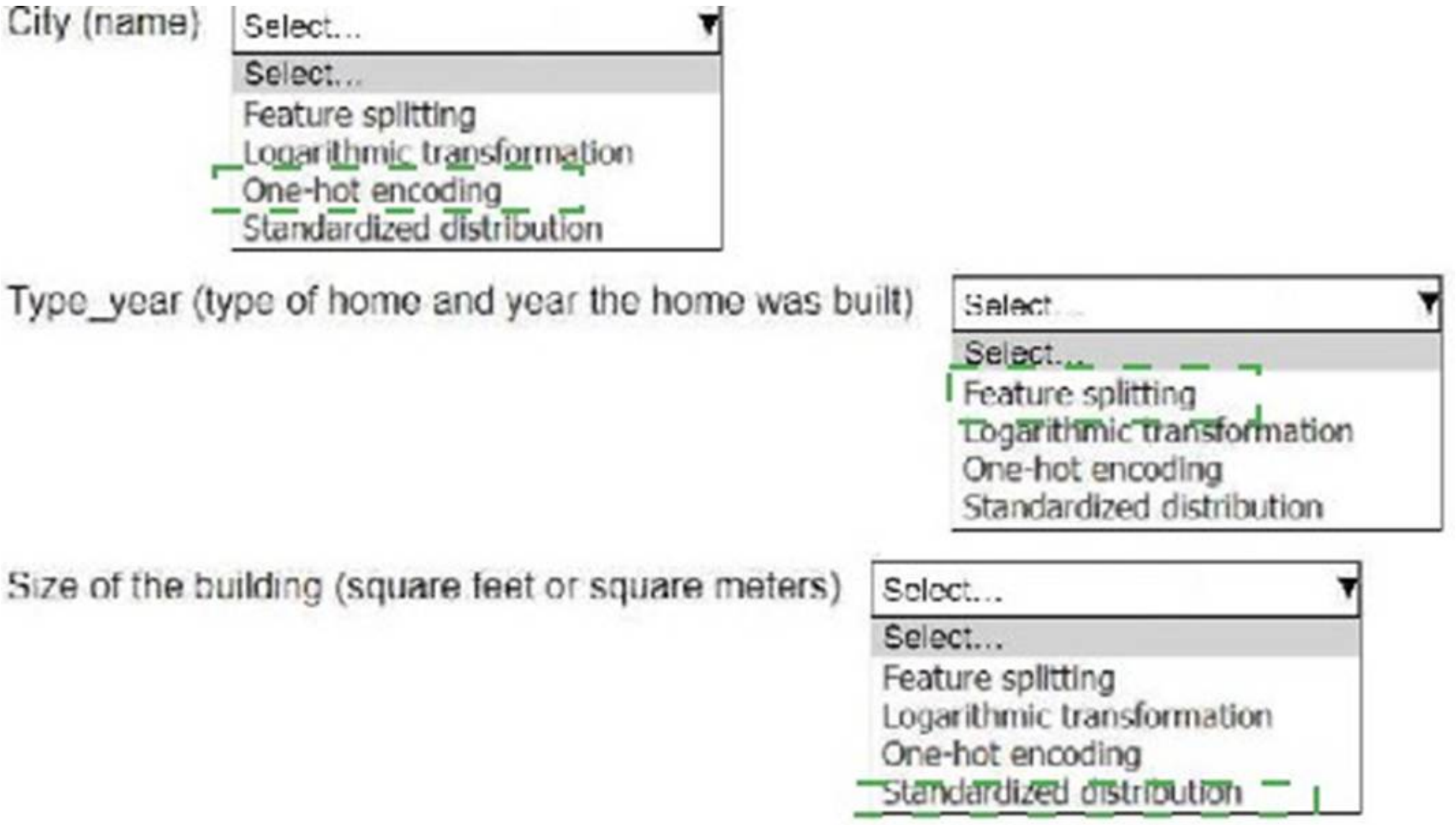
NEW QUESTION 29
HOTSPOT
An ML engineer is working on an ML model to predict the prices of similarly sized homes. The model will base predictions on several features The ML engineer will use the following feature engineering techniques to estimate the prices of the homes:
• Feature splitting
• Logarithmic 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 33
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 35
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 37
A company is using Amazon SageMaker to create ML models. The company's data scientists need fine-grained control of the ML workflows that they orchestrate. The data scientists also need the ability to visualize SageMaker jobs and workflows as a directed acyclic graph (DAG). The data scientists must keep a running history of model discovery experiments and must establish model governance for auditing and compliance verifications.
Which solution will meet these requirements?

A. Use AWS CodePipeline and its integration with SageMaker Studio to manage the entire ML workflow
B. Use SageMaker ML Lineage Tracking for the running history of experiments and for auditing and compliance verifications.
C. Use AWS CodePipeline and its integration with SageMaker Experiments to manage the entire ML workflow
D. Use SageMaker Experiments for the running history of experiments and for auditing and compliance verifications.
E. Use SageMaker Pipelines and its integration with SageMaker Studio to manage the entire ML workflow
F. Use SageMaker ML Lineage Tracking for the running history of experiments and for auditing and compliance verifications.
G. Use SageMaker Pipelines and its integration with SageMaker Experiments to manage the entire ML workflow
H. Use SageMaker Experiments for the running history of experiments and for auditing and compliance verifications.

Answer: C

NEW QUESTION 40
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 44
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 46
A company needs to give its ML engineers appropriate access to training data. The ML engineers must access training data from only their own business group. The ML engineers must not be allowed to access training data from other business groups.
The company uses a single AWS account and stores all the training data in Amazon S3 buckets. All ML model training occurs in Amazon SageMaker.
Which solution will provide the ML engineers with the appropriate access?

A. Enable S3 bucket versioning.
B. Configure S3 Object Lock settings for each user.
C. Add cross-origin resource sharing (CORS) policies to the S3 buckets.
D. Create IAM policie
E. Attach the policies to IAM users or IAM roles.

Answer: D

NEW QUESTION 50
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 55
FILL IN THE BLANK
A company stores time-series data about user clicks in an Amazon S3 bucket. The raw data consists of millions of rows of user activity every day. ML engineers access the data to develop their ML models.
The ML engineers need to generate daily reports and analyze click trends over the past 3 days by using Amazon Athena. The company must retain the data for 30 days before archiving the data.
Which solution will provide the HIGHEST performance for data retrieval?

A. Keep all the time-series data without partitioning in the S3 bucke
B. Manually move data that is older than 30 days to separate S3 buckets.
C. Create AWS Lambda functions to copy the time-series data into separate S3 bucket
D. Apply S3 Lifecycle policies to archive data that is older than 30 days to S3 Glacier Flexible Retrieval.
E. Organize the time-series data into partitions by date prefix in the S3 bucke
F. Apply S3 Lifecycle policies to archive partitions that are older than 30 days to S3 Glacier Flexible Retrieval.
G. Put each day's time-series data into its own S3 bucke

H. Use S3 Lifecycle policies to archive S3 buckets that hold data that is older than 30 days to S3 Glacier Flexible Retrieval.

Answer: C

NEW QUESTION 58
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 62
A company uses Amazon SageMaker for its ML workloads. The company's ML engineer receives a 50 MB Apache Parquet data file to build a fraud detection model. The file includes several correlated columns that are not required.
What should the ML engineer do to drop the unnecessary columns in the file with the LEAST effort?

A. Download the file to a local workstatio
B. Perform one-hot encoding by using a custom Python script.
C. Create an Apache Spark job that uses a custom processing script on Amazon EMR.
D. Create a SageMaker processing job by calling the SageMaker Python SDK.
E. Create a data flow in SageMaker Data Wrangle
F. Configure a transform step.

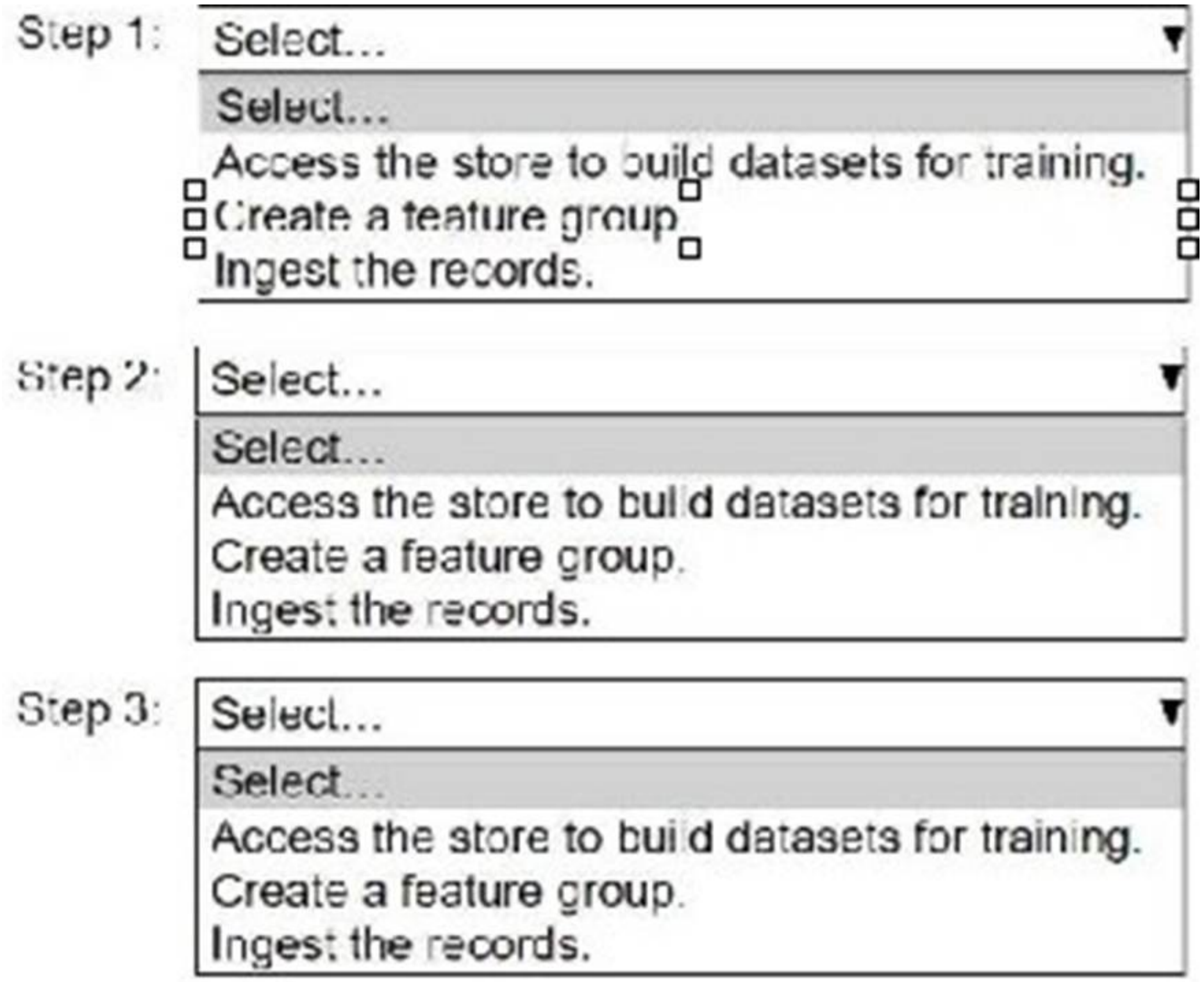
Answer: D

NEW QUESTION 64
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

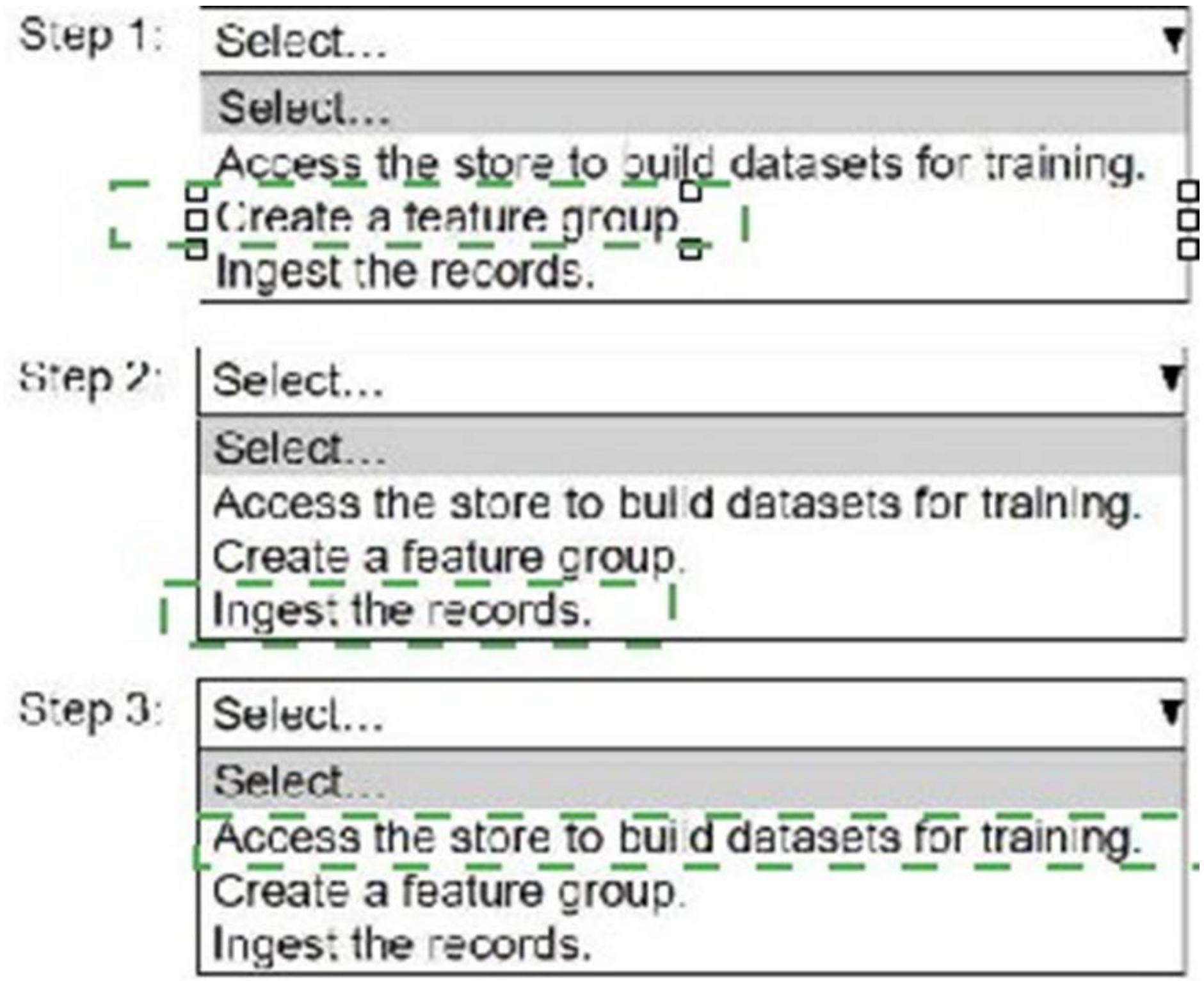
NEW QUESTION 65
HOTSPOT
An ML engineer needs to use Amazon SageMaker Feature Store to create and manage features to train a model.
Select and order the steps from the following list to create and use the features in Feature Store. Each step should be selected one time. (Select and order three.)
• Access the store to build datasets for training.
• Create a feature group.
• Ingest the records.



A. Mastered
B. Not Mastered

Answer: A

Explanation:



NEW QUESTION 69
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 72
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

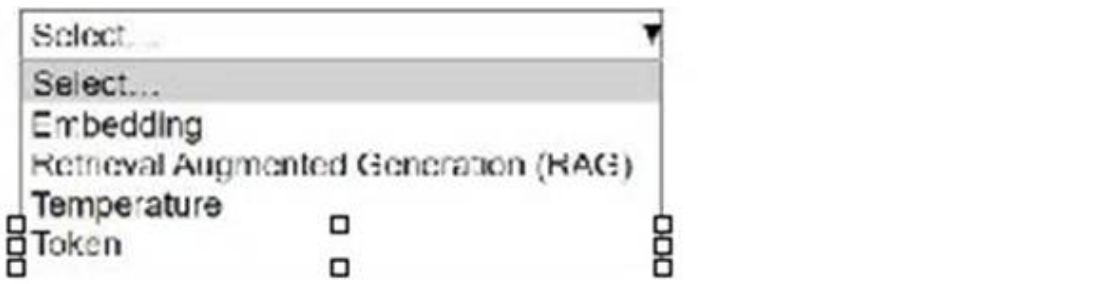
NEW QUESTION 74
HOTSPOT
An ML engineer is building a generative AI application on Amazon Bedrock by using large language models (LLMs).
Select the correct generative AI term from the following list for each description. Each term should be selected one time or not at all. (Select three.)
• Embedding
• Retrieval Augmented Generation (RAG)

• Temperature
• Token

Text representation of basic units of data processed by LLMs — Select...

High-dimensional vectors that contain the semantic meaning of text — Select...

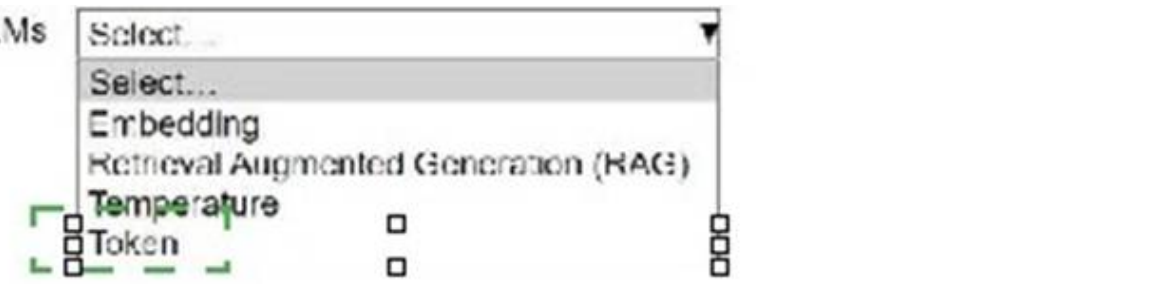
Enrichment of information from additional data sources to improve a generated response — Select...

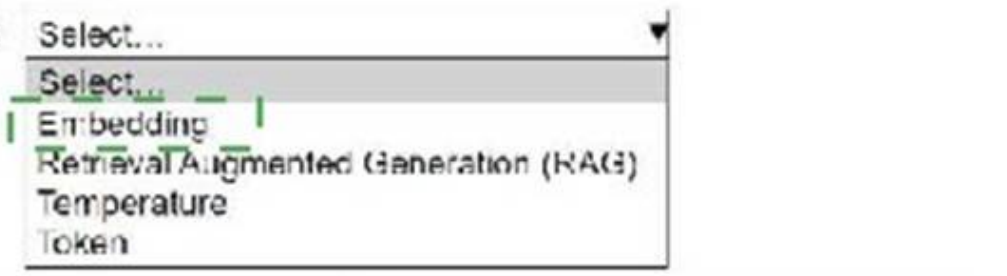


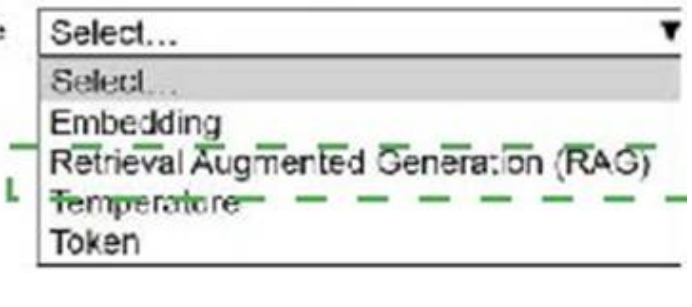
A. Mastered
B. Not Mastered

Answer: A

Explanation:







NEW QUESTION 79
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 82
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 86
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 88
A company has AWS Glue data processing jobs that are orchestrated by an AWS Glue workflow. The AWS Glue jobs can run on a schedule or can be launched manually.
The company is developing pipelines in Amazon SageMaker Pipelines for ML model development. The pipelines will use the output of the AWS Glue jobs during the data processing phase of model development. An ML engineer needs to implement a solution that integrates the AWS Glue jobs with the pipelines.
Which solution will meet these requirements with the LEAST operational overhead?

A. Use AWS Step Functions for orchestration of the pipelines and the AWS Glue jobs.
B. Use processing steps in SageMaker Pipeline
C. Configure inputs that point to the Amazon Resource Names (ARNs) of the AWS Glue jobs.
D. Use Callback steps in SageMaker Pipelines to start the AWS Glue workflow and to stop the pipelines until the AWS Glue jobs finish running.
E. Use Amazon EventBridge to invoke the pipelines and the AWS Glue jobs in the desired order.

Answer: C

NEW QUESTION 93
A company is planning to use Amazon Redshift ML in its primary AWS account. The source data is in an Amazon S3 bucket in a secondary account.
An ML engineer needs to set up an ML pipeline in the primary account to access the S3 bucket in the secondary account. The solution must not require public IPv4 addresses.
Which solution will meet these requirements?

A. Provision a Redshift cluster and Amazon SageMaker Studio in a VPC with no public access enabled in the primary accoun
B. Create a VPC peering connection between the account
C. Update the VPC route tables to remove the route to 0.0.0.0/0.
D. Provision a Redshift cluster and Amazon SageMaker Studio in a VPC with no public access enabled in the primary accoun
E. Create an AWS Direct Connect connection and a transit gatewa
F. Associate the VPCs from both accounts with the transit gatewa
G. Update the VPC route tables to remove the route to 0.0.0.0/0.
H. Provision a Redshift cluster and Amazon SageMaker Studio in a VPC in the primary accoun
I. Create an AWS Site-to-Site VPN connection with two encrypted IPsec tunnels between the account
J. Set up interface VPC endpoints for Amazon S3.
K. Provision a Redshift cluster and Amazon SageMaker Studio in a VPC in the primary accoun
L. Create an S3 gateway endpoin
M. Update the S3 bucket policy to allow IAM principals from the primary accoun
N. Set up interface VPC endpoints for SageMaker and Amazon Redshift.

Answer: D

NEW QUESTION 98
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 103
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 108
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 110
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 112
An ML engineer receives datasets that contain missing values, duplicates, and extreme outliers. The ML engineer must consolidate these datasets into a single data frame and must prepare the data for ML.
Which solution will meet these requirements?

A. Use Amazon SageMaker Data Wrangler to import the datasets and to consolidate them into a single data fram
B. Use the cleansing and enrichment functionalities to prepare the data.
C. Use Amazon SageMaker Ground Truth to import the datasets and to consolidate them into a single data fram
D. Use the human-in-the-loop capability to prepare the data.
E. Manually import and merge the dataset
F. Consolidate the datasets into a single data fram
G. Use Amazon Q Developer to generate code snippets that will prepare the data.
H. Manually import and merge the dataset
I. Consolidate the datasets into a single data fram
J. Use Amazon SageMaker data labeling to prepare the data.

Answer: A

NEW QUESTION 115
A credit card company has a fraud detection model in production on an Amazon SageMaker endpoint. The company develops a new version of the model. The company needs to assess the new model's performance by using live data and without affecting production end users.
Which solution will meet these requirements?

A. Set up SageMaker Debugger and create a custom rule.
B. Set up blue/green deployments with all-at-once traffic shifting.
C. Set up blue/green deployments with canary traffic shifting.
D. Set up shadow testing with a shadow variant of the new model.

Answer: D

NEW QUESTION 119

.....

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