Amazon-Web-Services

Exam Questions MLA-C01

AWS Certified Machine Learning Engineer - Associate

NEW QUESTION 1
An ML engineer needs to implement a solution to host a trained ML model. The rate of requests to the model will be inconsistent throughout the day.
The ML engineer needs a scalable solution that minimizes costs when the model is not in use. The solution also must maintain the model's capacity to respond to requests during times of peak usage.
Which solution will meet these requirements?

A. Create AWS Lambda functions that have fixed concurrency to host the mode
B. Configure the Lambda functions to automatically scale based on the number of requests to the model.
C. Deploy the model on an Amazon Elastic Container Service (Amazon ECS) cluster that uses AWS Fargat
D. Set a static number of tasks to handle requests during times of peak usage.
E. Deploy the model to an Amazon SageMaker endpoin
F. Deploy multiple copies of the model to the endpoin
G. Create an Application Load Balancer to route traffic between the different copies of the model at the endpoint.
H. Deploy the model to an Amazon SageMaker endpoin
I. Create SageMaker endpoint auto scaling policies that are based on Amazon CloudWatch metrics to adjust the number of instances dynamically.

Answer: D

NEW QUESTION 2
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 3
A company is building a deep learning model on Amazon SageMaker. The company uses a large amount of data as the training dataset. The company needs to optimize the model's hyperparameters to minimize the loss function on the validation dataset.
Which hyperparameter tuning strategy will accomplish this goal with the LEAST computation time?

A. Hyperbaric!
B. Grid search
C. Bayesian optimization
D. Random search

Answer: A

NEW QUESTION 4
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 5
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 6
A company uses Amazon Athena to query a dataset in Amazon S3. The dataset has a target variable that the company wants to predict.
The company needs to use the dataset in a solution to determine if a model can predict the target variable.
Which solution will provide this information with the LEAST development effort?

A. Create a new model by using Amazon SageMaker Autopilo
B. Report the model's achieved performance.
C. Implement custom scripts to perform data pre-processing, multiple linear regression, and performance evaluatio
D. Run the scripts on Amazon EC2 instances.
E. Configure Amazon Macie to analyze the dataset and to create a mode
F. Report the model's achieved performance.
G. Select a model from Amazon Bedroc
H. Tune the model with the dat
I. Report the model's achieved performance.

Answer: A

NEW QUESTION 7
An advertising company uses AWS Lake Formation to manage a data lake. The data lake contains structured data and unstructured data. The company's ML engineers are assigned to specific advertisement campaigns.
The ML engineers must interact with the data through Amazon Athena and by browsing the data directly in an Amazon S3 bucket. The ML engineers must have access to only the resources that are specific to their assigned advertisement campaigns.
Which solution will meet these requirements in the MOST operationally efficient way?

A. Configure IAM policies on an AWS Glue Data Catalog to restrict access to Athena based on the ML engineers' campaigns.
B. Store users and campaign information in an Amazon DynamoDB tabl
C. Configure DynamoDB Streams to invoke an AWS Lambda function to update S3 bucket policies.
D. Use Lake Formation to authorize AWS Glue to access the S3 bucke
E. Configure Lake Formation tags to map ML engineers to their campaigns.
F. Configure S3 bucket policies to restrict access to the S3 bucket based on the ML engineers' campaigns.

Answer: C

NEW QUESTION 8
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 9
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 10
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 10
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 14
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 18
A company runs an Amazon SageMaker domain in a public subnet of a newly created VPC. The network is configured properly, and ML engineers can access the SageMaker domain.
Recently, the company discovered suspicious traffic to the domain from a specific IP address. The company needs to block traffic from the specific IP address.
Which update to the network configuration will meet this requirement?

A. Create a security group inbound rule to deny traffic from the specific IP addres
B. Assign the security group to the domain.
C. Create a network ACL inbound rule to deny traffic from the specific IP addres
D. Assign the rule to the default network Ad for the subnet where the domain is located.
E. Create a shadow variant for the domai
F. Configure SageMaker Inference Recommender to send traffic from the specific IP address to the shadow endpoint.
G. Create a VPC route table to deny inbound traffic from the specific IP addres
H. Assign the route table to the domain.

Answer: B

NEW QUESTION 19
A company is creating an application that will recommend products for customers to purchase. The application will make API calls to Amazon Q Business. The company must ensure that responses from Amazon Q Business do not include the name of the company's main competitor.
Which solution will meet this requirement?

A. Configure the competitor's name as a blocked phrase in Amazon Q Business.
B. Configure an Amazon Q Business retriever to exclude the competitor??s name.
C. Configure an Amazon Kendra retriever for Amazon Q Business to build indexes that exclude the competitor's name.
D. Configure document attribute boosting in Amazon Q Business to deprioritize the competitor's name.

Answer: A

NEW QUESTION 20
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 25
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 28
An ML engineer has an Amazon Comprehend custom model in Account A in the us-east-1 Region. The ML engineer needs to copy the model to Account in the same Region.
Which solution will meet this requirement with the LEAST development effort?

A. Use Amazon S3 to make a copy of the mode
B. Transfer the copy to Account B.
C. Create a resource-based IAM polic
D. Use the Amazon Comprehend ImportModel API operation to copy the model to Account B.
E. Use AWS DataSync to replicate the model from Account A to Account B.
F. Create an AWS Site-to-Site VPN connection between Account A and Account to transfer the model.

Answer: B

NEW QUESTION 33
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 34
A company has a conversational AI assistant that sends requests through Amazon Bedrock to an Anthropic Claude large language model (LLM). Users report that when they ask similar questions multiple times, they sometimes receive different answers. An ML engineer needs to improve the responses to be more consistent and less random.
Which solution will meet these requirements?

A. Increase the temperature parameter and the top_k parameter.
B. Increase the temperature paramete
C. Decrease the top_k parameter.
D. Decrease the temperature paramete
E. Increase the top_k parameter.
F. Decrease the temperature parameter and the top_k parameter.

Answer: D

NEW QUESTION 37
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 40
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 41
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 44
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 49
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 50
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 52
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 must implement a manual approval-based workflow to ensure that only approved models can be deployed to production endpoints.
Which solution will meet this requirement?

A. Use SageMaker Experiments to facilitate the approval process during model registration.
B. Use SageMaker ML Lineage Tracking on the central model registr
C. Create tracking entities for the approval process.
D. Use SageMaker Model Monitor to evaluate the performance of the model and to manage the approval.
E. Use SageMaker Pipeline
F. When a model version is registered, use the AWS SDK to change the approval status to "Approved."

Answer: D

NEW QUESTION 53
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 57
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.
Before the ML engineer trains the model, the ML engineer must resolve the issue of the imbalanced data.
Which solution will meet this requirement with the LEAST operational effort?

A. Use Amazon Athena to identify patterns that contribute to the imbalanc
B. Adjust the dataset accordingly.
C. Use Amazon SageMaker Studio Classic built-in algorithms to process the imbalanced dataset.
D. Use AWS Glue DataBrew built-in features to oversample the minority class.
E. Use the Amazon SageMaker Data Wrangler balance data operation to oversample the minority class.

Answer: D

NEW QUESTION 59
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 64
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.

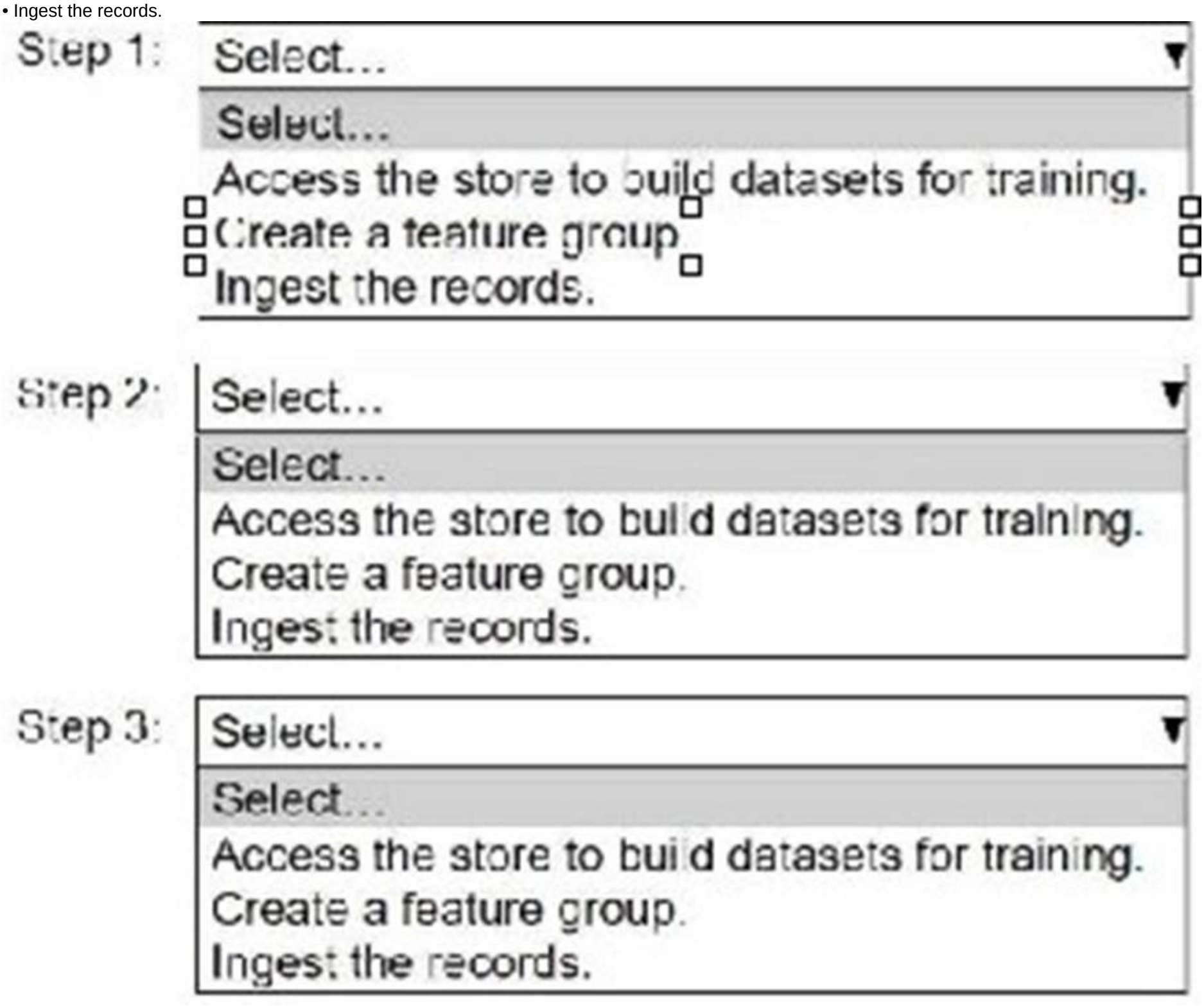
Answer: D

NEW QUESTION 67
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

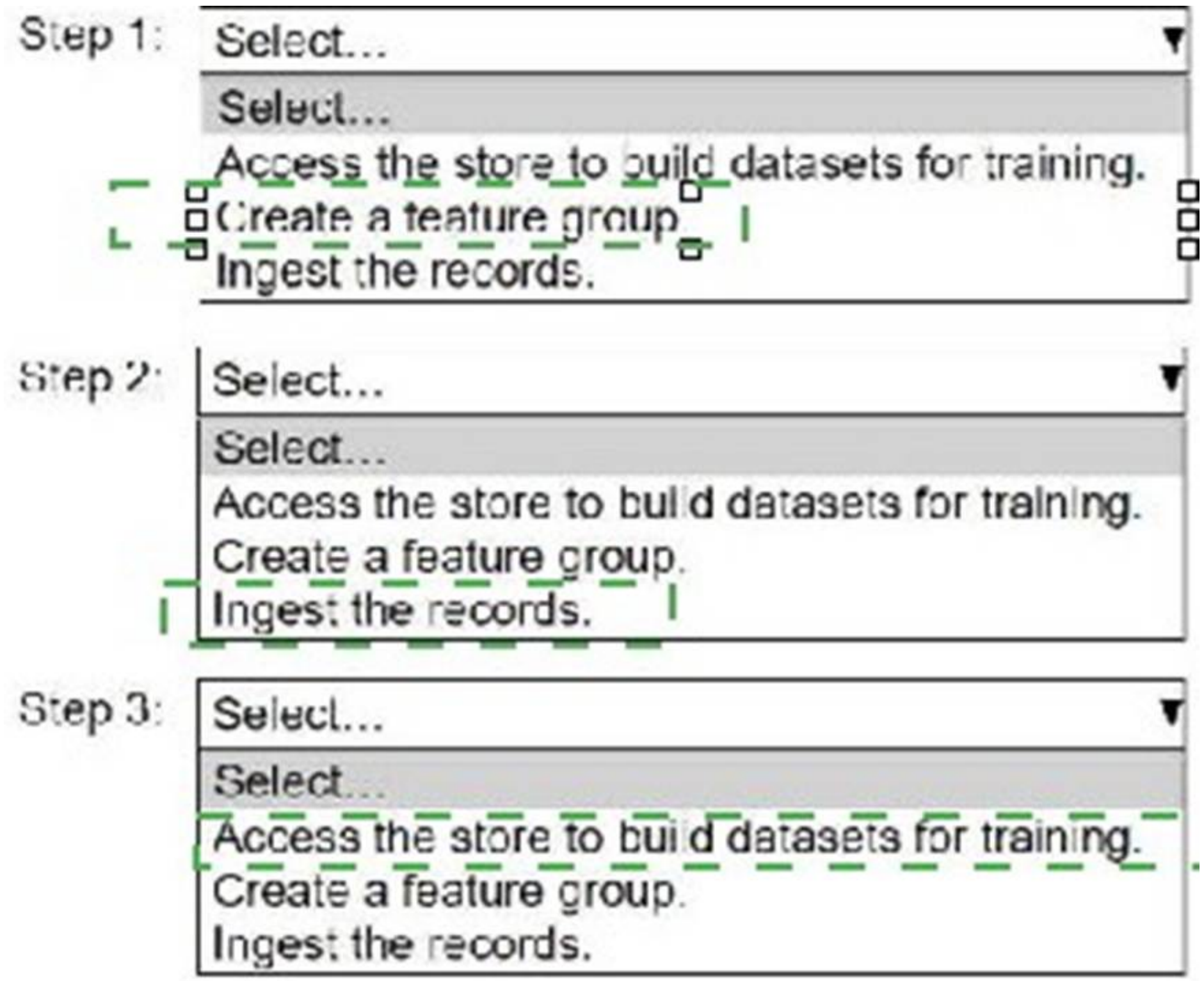
NEW QUESTION 71
HOTSPOT
An ML engineer needs to use Amazon SageMaker Feature Store to create and manage features to train a model.
Select and order the steps from the following list to create and use the features in Feature Store. Each step should be selected one time. (Select and order three.)
• Access the store to build datasets for training.
• Create a feature group.
• Ingest the records.



A. Mastered
B. Not Mastered

Answer: A

Explanation:



NEW QUESTION 72
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 76
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 80
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

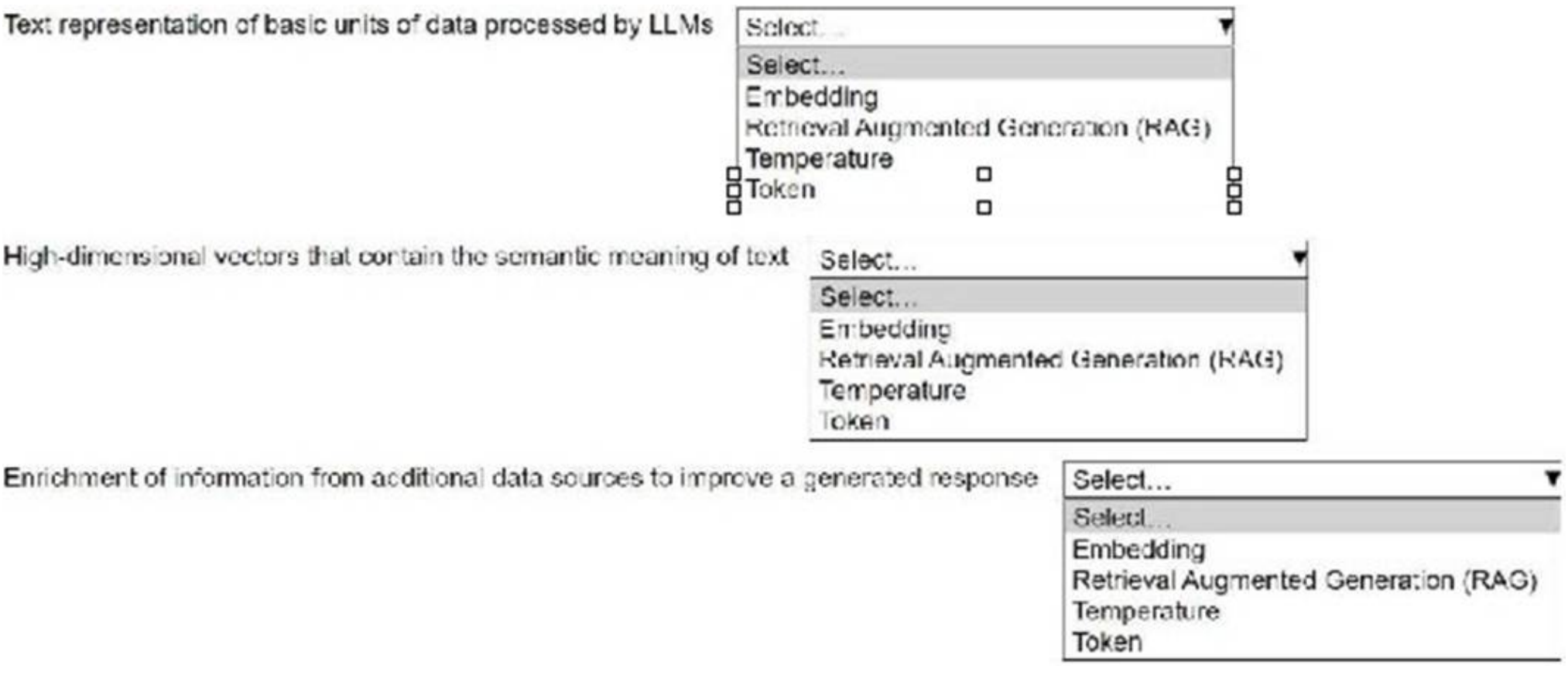
Answer: D

NEW QUESTION 84
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

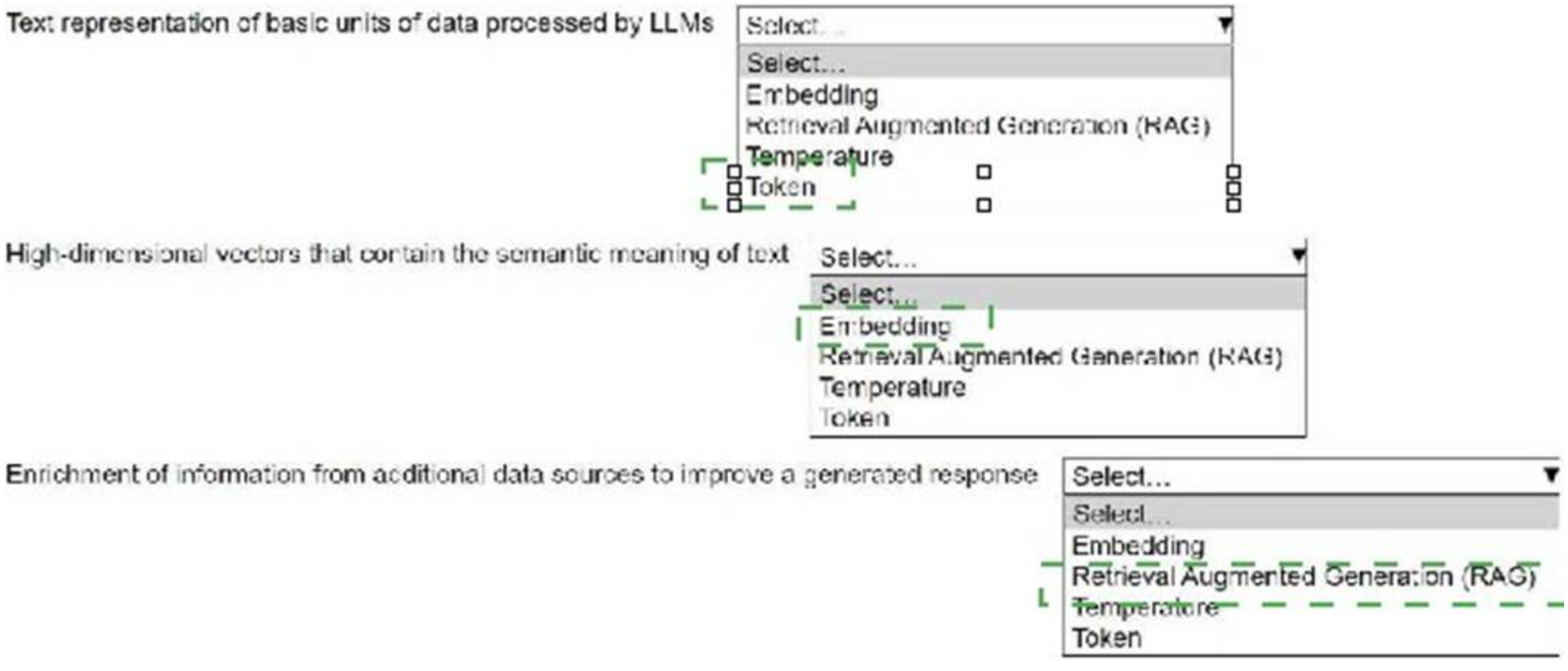
NEW QUESTION 88
HOTSPOT
An ML engineer is building a generative AI application on Amazon Bedrock by using large language models (LLMs).
Select the correct generative AI term from the following list for each description. Each term should be selected one time or not at all. (Select three.)
• Embedding
• Retrieval Augmented Generation (RAG)
• Temperature
• Token



A. Mastered
B. Not Mastered

Answer: A

Explanation:



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 94
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 99
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 100
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.
After the data is aggregated, the ML engineer must implement a solution to automatically detect anomalies in the data and to visualize the result.
Which solution will meet these requirements?

A. Use Amazon Athena to automatically detect the anomalies and to visualize the result.
B. Use Amazon Redshift Spectrum to automatically detect the anomalie
C. Use Amazon QuickSight to visualize the result.
D. Use Amazon SageMaker Data Wrangler to automatically detect the anomalies and to visualize the result.
E. Use AWS Batch to automatically detect the anomalie
F. Use Amazon QuickSight to visualize the result.

Answer: C

NEW QUESTION 102
.....

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