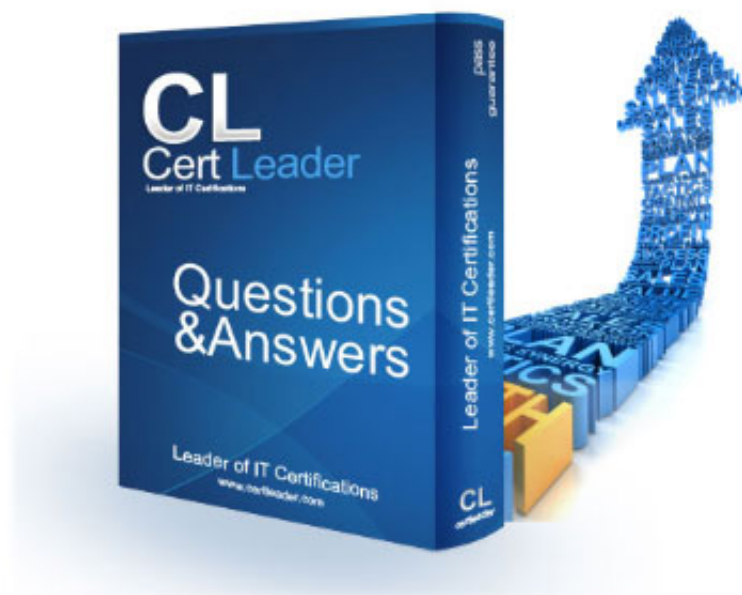


## MLA-C01 Dumps

### AWS Certified Machine Learning Engineer - Associate

<https://www.certleader.com/MLA-C01-dumps.html>



**NEW QUESTION 1**

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 2**

An ML engineer needs to use data with Amazon SageMaker Canvas to train an ML model. The data is stored in Amazon S3 and is complex in structure. The ML engineer must use a file format that minimizes processing time for the data.

Which file format will meet these requirements?

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

**Answer:** D

**NEW QUESTION 3**

A company is planning to create several ML prediction models. The training data is stored in Amazon S3. The entire dataset is more than 5 in size and consists of CSV, JSON, Apache Parquet, and simple text files.

The data must be processed in several consecutive steps. The steps include complex manipulations that can take hours to finish running. Some of the processing involves natural language processing (NLP) transformations. The entire process must be automated.

Which solution will meet these requirements?

- A. Process data at each step by using Amazon SageMaker Data Wrangle
- B. Automate the process by using Data Wrangler jobs.
- C. Use Amazon SageMaker notebooks for each data processing ste
- D. Automate the process by using Amazon EventBridge.
- E. Process data at each step by using AWS Lambda function
- F. Automate the process by using AWS Step Functions and Amazon EventBridge.
- G. Use Amazon SageMaker Pipelines to create a pipeline of data processing step
- H. Automate the pipeline by using Amazon EventBridge.

**Answer:** D

**NEW QUESTION 4**

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 5**

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 6

A company has an application that uses different APIs to generate embeddings for input text. The company needs to implement a solution to automatically rotate the API tokens every 3 months.

Which solution will meet this requirement?

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

**Answer: A**

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

**Answer: D**

#### NEW QUESTION 9

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 10

A company has deployed an XGBoost prediction model in production to predict if a customer is likely to cancel a subscription. The company uses Amazon SageMaker Model Monitor to detect deviations in the F1 score.

During a baseline analysis of model quality, the company recorded a threshold for the F1 score. After several months of no change, the model's F1 score decreases significantly.

What could be the reason for the reduced F1 score?

- A. Concept drift occurred in the underlying customer data that was used for predictions.
- B. The model was not sufficiently complex to capture all the patterns in the original baseline data.
- C. The original baseline data had a data quality issue of missing values.
- D. Incorrect ground truth labels were provided to Model Monitor during the calculation of the baseline.

**Answer:** A

#### NEW QUESTION 15

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 is experimenting with consecutive training jobs.

How can the company MINIMIZE infrastructure startup times for these jobs?

- A. Use Managed Spot Training.
- B. Use SageMaker managed warm pools.
- C. Use SageMaker Training Compiler.
- D. Use the SageMaker distributed data parallelism (SMDDP) library.

**Answer:** B

#### NEW QUESTION 16

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 CloudTrail
- F. Use an Amazon CloudWatch dashboard for monitoring
- 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 monitoring
- I. Set up a CloudWatch alarm to provide notification when the threshold is breached.

**Answer:** D

#### NEW QUESTION 17

##### HOTSPOT

A company wants to host an ML model on Amazon SageMaker. An ML engineer is configuring a continuous integration and continuous delivery (CI/CD) pipeline in AWS CodePipeline to deploy the model. The pipeline must run automatically when new training data for the model is uploaded to an Amazon S3 bucket.

Select and order the pipeline's correct steps from the following list. Each step should be selected one time or not at all. (Select and order three.)

- An S3 event notification invokes the pipeline when new data is uploaded.
- S3 Lifecycle rule invokes the pipeline when new data is uploaded.
- SageMaker retrains the model by using the data in the S3 bucket.
- The pipeline deploys the model to a SageMaker endpoint.
- The pipeline deploys the model to SageMaker Model Registry.

Step 1:

Select...

Select...

An S3 event notification invokes the pipeline when new data is uploaded.  
An S3 Lifecycle rule invokes the pipeline when new data is uploaded.  
SageMaker retrains the model by using the data in the S3 bucket.  
The pipeline deploys the model to a SageMaker endpoint.  
The pipeline deploys the model to SageMaker Model Registry.

Step 2:

Select...

Select...

An S3 event notification invokes the pipeline when new data is uploaded.  
An S3 Lifecycle rule invokes the pipeline when new data is uploaded.  
SageMaker retrains the model by using the data in the S3 bucket.  
The pipeline deploys the model to a SageMaker endpoint.  
The pipeline deploys the model to SageMaker Model Registry.

Step 3:

Select...

Select...

An S3 event notification invokes the pipeline when new data is uploaded.  
An S3 Lifecycle rule invokes the pipeline when new data is uploaded.  
SageMaker retrains the model by using the data in the S3 bucket.  
The pipeline deploys the model to a SageMaker endpoint.  
The pipeline deploys the model to SageMaker Model Registry.

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Step 1: Select...

Select...

An S3 event notification invokes the pipeline when new data is uploaded.

An S3 Lifecycle rule invokes the pipeline when new data is uploaded.

SageMaker retrains the model by using the data in the S3 bucket.

The pipeline deploys the model to a SageMaker endpoint.

The pipeline deploys the model to SageMaker Model Registry.

Step 2: Select...

Select...

An S3 event notification invokes the pipeline when new data is uploaded.

An S3 Lifecycle rule invokes the pipeline when new data is uploaded.

SageMaker retrains the model by using the data in the S3 bucket.

The pipeline deploys the model to a SageMaker endpoint.

The pipeline deploys the model to SageMaker Model Registry.

Step 3: Select...

Select...

An S3 event notification invokes the pipeline when new data is uploaded.

An S3 Lifecycle rule invokes the pipeline when new data is uploaded.

SageMaker retrains the model by using the data in the S3 bucket.

The pipeline deploys the model to a SageMaker endpoint.

The pipeline deploys the model to SageMaker Model Registry.

#### NEW QUESTION 20

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 23

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 parameter.
- C. Decrease the top\_k parameter.
- D. Decrease the temperature parameter.
- E. Increase the top\_k parameter.
- F. Decrease the temperature parameter and the top\_k parameter.

**Answer:** D

#### NEW QUESTION 24

##### 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.)

City (name)

|                            |
|----------------------------|
| Select...                  |
| Select...                  |
| Feature splitting          |
| Logarithmic transformation |
| One-hot encoding           |
| Standardized distribution  |

Type\_year (type of home and year the home was built)

|                            |
|----------------------------|
| Select...                  |
| Select...                  |
| Feature splitting          |
| Logarithmic transformation |
| One-hot encoding           |
| Standardized distribution  |

Size of the building (square feet or square meters)

|                            |
|----------------------------|
| Select...                  |
| Select...                  |
| Feature splitting          |
| Logarithmic transformation |
| One-hot encoding           |
| Standardized distribution  |

- A. Mastered  
B. Not Mastered

Answer: A

Explanation:

City (name)

|                            |
|----------------------------|
| Select...                  |
| Select...                  |
| Feature splitting          |
| Logarithmic transformation |
| One-hot encoding           |
| Standardized distribution  |

Type\_year (type of home and year the home was built)

|                            |
|----------------------------|
| Select...                  |
| Select...                  |
| Feature splitting          |
| Logarithmic transformation |
| One-hot encoding           |
| Standardized distribution  |

Size of the building (square feet or square meters)

|                            |
|----------------------------|
| Select...                  |
| Select...                  |
| Feature splitting          |
| Logarithmic transformation |
| One-hot encoding           |
| Standardized distribution  |

#### NEW QUESTION 29

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 31

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 endpoint
- 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 endpoint
- 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 endpoint
- H. Use production variants to add the new model to the new endpoint
- 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 endpoint
- K. Monitor the number of invocations by using AWS CloudTrail.

**Answer: A**

#### NEW QUESTION 34

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 39

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 imbalance
- 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 40

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 inference
- B. Use SageMaker Model Monitor for notifications about model quality.
- C. Use SageMaker batch transform for inference
- D. Use SageMaker Model Monitor for notifications about model quality.
- E. Use SageMaker Serverless Inference for inference
- F. Use SageMaker Inference Recommender for notifications about model quality.
- G. Keep using SageMaker Asynchronous Inference for inference
- H. Use SageMaker Inference Recommender for notifications about model quality.

**Answer: A**

NEW QUESTION 41

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 44

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 46

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...

Select...

Embedding

Retrieval Augmented Generation (RAG)

Temperature

Token

High-dimensional vectors that contain the semantic meaning of text

Select...

Select...

Embedding

Retrieval Augmented Generation (RAG)

Temperature

Token

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

Select...

Select...

Embedding

Retrieval Augmented Generation (RAG)

Temperature

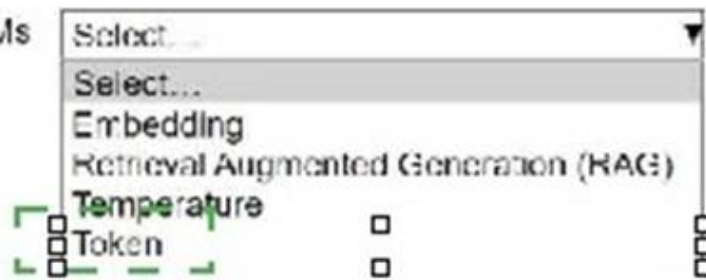
Token

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

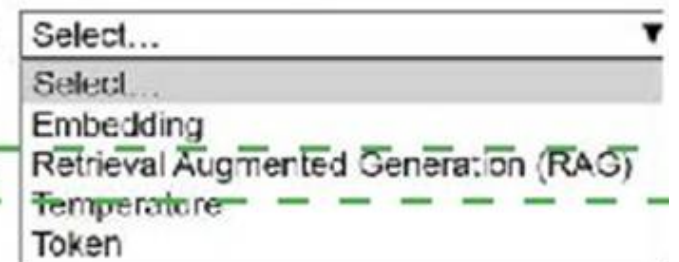
Text representation of basic units of data processed by LLMs



High-dimensional vectors that contain the semantic meaning of text



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



#### NEW QUESTION 51

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 55

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 60

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.
- 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.
- 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 61

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 62

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

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

**Answer:** D

#### NEW QUESTION 67

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 training dataset includes categorical data and numerical data. The ML engineer must prepare the training dataset to maximize the accuracy of the model.

Which action will meet this requirement with the LEAST operational overhead?

- A. Use AWS Glue to transform the categorical data into numerical data.
- B. Use AWS Glue to transform the numerical data into categorical data.
- C. Use Amazon SageMaker Data Wrangler to transform the categorical data into numerical data.
- D. Use Amazon SageMaker Data Wrangler to transform the numerical data into categorical data.

**Answer:** C

#### NEW QUESTION 71

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 74

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 frame
- 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 frame
- 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 frame
- 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 frame
- J. Use Amazon SageMaker data labeling to prepare the data.

**Answer:** A

#### NEW QUESTION 78

A company that has hundreds of data scientists is using Amazon SageMaker to create ML models. The models are in model groups in the SageMaker Model Registry.

The data scientists are grouped into three categories: computer vision, natural language processing (NLP), and speech recognition. An ML engineer needs to implement a solution to organize the existing models into these groups to improve model discoverability at scale. The solution must not affect the integrity of the model artifacts and their existing groupings.

Which solution will meet these requirements?

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

**Answer:** A

#### **NEW QUESTION 82**

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 84**

.....

## Thank You for Trying Our Product

\* 100% Pass or Money Back

All our products come with a 90-day Money Back Guarantee.

\* One year free update

You can enjoy free update one year. 24x7 online support.

\* Trusted by Millions

We currently serve more than 30,000,000 customers.

\* Shop Securely

All transactions are protected by VeriSign!

**100% Pass Your MLA-C01 Exam with Our Prep Materials Via below:**

<https://www.certleader.com/MLA-C01-dumps.html>