

Amazon-Web-Services

Exam Questions MLA-C01

AWS Certified Machine Learning Engineer - Associate



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 storage
- B. Adjust the training job to use the new S3 bucket.
- C. Create an Amazon FSx for Lustre file system
- D. Link the file system to the existing S3 bucket
- E. Adjust the training job to read from the file system.
- F. Create an Amazon Elastic File System (Amazon EFS) file system
- G. Transfer the existing data to the file system
- H. Adjust the training job to read from the file system.
- I. Create an Amazon ElastiCache (Redis OSS) cluster
- J. Link the Redis OSS cluster to the existing S3 bucket
- K. Stream the data from the Redis OSS cluster directly to the training job.

Answer: B

NEW QUESTION 2

A company is using an Amazon Redshift database as its single data source. Some of the data is sensitive.

A data scientist needs to use some of the sensitive data from the database. An ML engineer must give the data scientist access to the data without transforming the source data and without storing anonymized data in the database.

Which solution will meet these requirements with the LEAST implementation effort?

- A. Configure dynamic data masking policies to control how sensitive data is shared with the data scientist at query time.
- B. Create a materialized view with masking logic on top of the database
- 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 logic
- 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 data
- G. Share the dataset with the data scientist.

Answer: A

NEW QUESTION 3

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

Which file format will meet these requirements?

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

Answer: D

NEW QUESTION 4

A company is planning to create several ML prediction models. The training data is stored in Amazon S3. The entire dataset is more than 5 TB 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 Wrangler
- B. Automate the process by using Data Wrangler jobs.
- C. Use Amazon SageMaker notebooks for each data processing step
- 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 steps
- H. Automate the pipeline by using Amazon EventBridge.

Answer: D

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 CloudTrail
- B. Configure a CloudTrail trail to send the email message.
- C. Log the metrics from the Lambda function to Amazon CloudFront
- D. Configure an Amazon CloudWatch alarm to send the email message.
- E. Log the metrics from the Lambda function to Amazon CloudWatch
- F. Configure a CloudWatch alarm to send the email message.
- G. Log the metrics from the Lambda function to Amazon CloudWatch

H. Configure an Amazon CloudFront rule to send the email message.

Answer: D

NEW QUESTION 6

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 endpoint
- B. Configure auto scaling
- C. Configure the endpoint to present the existing model.
- D. Deploy the model on an Amazon Elastic Container Service (Amazon ECS) cluster
- 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) cluster
- G. Deploy the model in Amazon EKS
- 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 7

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 8

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 detection
- E. Create an AWS Lambda function to detect anomalies
- 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 detection
- I. Create an AWS Lambda function to detect anomalies
- J. Use the data streams to invoke the Lambda function.
- K. Send real-time data to an Amazon Simple Queue Service (Amazon SQS) FIFO queue
- 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 9

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 B 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 model
- B. Transfer the copy to Account B.
- C. Create a resource-based IAM policy
- 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 B to transfer the model.

Answer: B

NEW QUESTION 10

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 domain
- 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 domain.
- D. Configure AWS Budgets to send an alert when the threshold is reached.
- E. Add resource tagging by editing each user's IAM profile.
- F. Configure AWS Cost Explorer to send an alert when the threshold is reached.
- G. Add resource tagging by editing each user's IAM profile.
- H. Configure AWS Budgets to send an alert when the threshold is reached.

Answer: B

NEW QUESTION 10

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

Which solution will meet these requirements?

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

Answer: C

NEW QUESTION 11

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 workstation.
- 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 Wrangler.
- F. Configure a transform step.

Answer: D

NEW QUESTION 16

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.

Step 1: Select...
Select...
Access the store to build datasets for training.
 Create a feature group
 Ingest the records.

Step 2: Select...
Select...
Access the store to build datasets for training.
Create a feature group.
Ingest the records.

Step 3: Select...
Select...
Access the store to build datasets for training.
Create a feature group.
Ingest the records.

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Step 1: Select...
 Select...
 Access the store to build datasets for training.
 Create a feature group.
 Ingest the records.

Step 2: Select...
 Select...
 Access the store to build datasets for training.
 Create a feature group.
 Ingest the records.

Step 3: Select...
 Select...
 Access the store to build datasets for training.
 Create a feature group.
 Ingest the records.

NEW QUESTION 18

A company needs to create a central catalog for all the company's ML models. The models are in AWS accounts where the company developed the models initially. The models are hosted in Amazon Elastic Container Registry (Amazon ECR) repositories. Which solution will meet these requirements?

- A. Configure ECR cross-account replication for each existing ECR repository
- B. Ensure that each model is visible in each AWS account.
- C. Create a new AWS account with a new ECR repository as the central catalog
- D. Configure ECR cross-account replication between the initial ECR repositories and the central catalog.
- E. Use the Amazon SageMaker Model Registry to create a model group for models hosted in Amazon EC
- F. Create a new AWS account
- G. In the new account, use the SageMaker Model Registry as the central catalog
- H. Attach a cross-account resource policy to each model group in the initial AWS accounts.
- I. Use an AWS Glue Data Catalog to store the model
- J. Run an AWS Glue crawler to migrate the models from the ECR repositories to the Data Catalog
- K. Configure cross-account access to the Data Catalog.

Answer: C

NEW QUESTION 22

An ML engineer trained an ML model on Amazon SageMaker to detect automobile accidents from doped-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 training dataset.
- C. Recreate the training dataset by using the Data Wrangler corrupt image transform
- D. Specify the impulse noise option.
- E. Recreate the training dataset by using the Data Wrangler enhance image contrast transform
- F. Specify the Gamma contrast option.
- G. Recreate the training dataset by using the Data Wrangler resize image transform
- H. Crop all images to the same size.

Answer: B

NEW QUESTION 25

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 set.
- C. Use these values to normalize the production samples.
- D. Calculate a new set of min-max normalization statistics from a batch of production samples.
- E. Use these values to normalize all the production samples.
- F. Calculate a new set of min-max normalization statistics from each production sample.
- G. Use these values to normalize all the production samples.

Answer: B

NEW QUESTION 30

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 34

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.

Which AWS service or feature can aggregate the data from the various data sources?

- A. Amazon EMR Spark jobs
- B. Amazon Kinesis Data Streams
- C. Amazon DynamoDB
- D. AWS Lake Formation

Answer: A

NEW QUESTION 38

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 40

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 45

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 47

.....

Thank You for Trying Our Product

We offer two products:

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

2nd - Questions 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

100% Actual & Verified — Instant Download, Please Click
[Order The MLA-C01 Practice Test Here](#)