



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 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

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

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

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

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

Answer: A

NEW QUESTION 3

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 gathering audio, video, and text data in various languages. The company needs to use a large language model (LLM) to summarize the gathered data that is in Spanish.

Which solution will meet these requirements in the LEAST amount of time?

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

Answer: B

NEW QUESTION 5

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 6

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 7

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 8

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 policy.
- 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 scaling.
- E. Host the model on the EKS cluster.
- 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) cluster.
- H. Specify a list of images to process for each AWS Batch job.

Answer: B

NEW QUESTION 9

HOTSPOT

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

Select and order the correct steps from the following list to perform this task. Each step should be selected one time or not at all. (Select and order three.)

- Create an Amazon SageMaker batch transform job for data cleaning and feature engineering.
- Store the resulting data back in Amazon S3.
- Use Amazon Athena to infer the schemas and available columns.
- Use AWS Glue crawlers to infer the schemas and available columns.
- Use AWS Glue DataBrew for data cleaning and feature engineering.

Step 1: Select...

Select...

Create an Amazon SageMaker batch transform job for data cleaning and feature engineering.
 Store the resulting data back in Amazon S3.
 Use Amazon Athena to infer the schemas and available columns.
 Use AWS Glue crawlers to infer the schemas and available columns.
 Use AWS Glue DataBrew for data cleaning and feature engineering.

Step 2: Select...

Select...

Create an Amazon SageMaker batch transform job for data cleaning and feature engineering.
 Store the resulting data back in Amazon S3.
 Use Amazon Athena to infer the schemas and available columns.
 Use AWS Glue crawlers to infer the schemas and available columns.
 Use AWS Glue DataBrew for data cleaning and feature engineering.

Step 3: Select...

Select...

Create an Amazon SageMaker batch transform job for data cleaning and feature engineering.
 Store the resulting data back in Amazon S3.
 Use Amazon Athena to infer the schemas and available columns.
 Use AWS Glue crawlers to infer the schemas and available columns.
 Use AWS Glue DataBrew for data cleaning and feature engineering.

- A. Mastered
 B. Not Mastered

Answer: A

Explanation:

Step 1: Select...

Select...

Create an Amazon SageMaker batch transform job for data cleaning and feature engineering.
 Store the resulting data back in Amazon S3.
 Use Amazon Athena to infer the schemas and available columns.
 Use AWS Glue crawlers to infer the schemas and available columns.
 Use AWS Glue DataBrew for data cleaning and feature engineering.

Step 2: Select...

Select...

Create an Amazon SageMaker batch transform job for data cleaning and feature engineering.
 Store the resulting data back in Amazon S3.
 Use Amazon Athena to infer the schemas and available columns.
 Use AWS Glue crawlers to infer the schemas and available columns.
 Use AWS Glue DataBrew for data cleaning and feature engineering.

Step 3: Select...

Select...

Create an Amazon SageMaker batch transform job for data cleaning and feature engineering.
 Store the resulting data back in Amazon S3.
 Use Amazon Athena to infer the schemas and available columns.
 Use AWS Glue crawlers to infer the schemas and available columns.
 Use AWS Glue DataBrew for data cleaning and feature engineering.

NEW QUESTION 10

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 10

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 12

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 13

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 17

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 20

A company needs to run a batch data-processing job on Amazon EC2 instances. The job will run during the weekend and will take 90 minutes to finish running. The processing can

handle interruptions. The company will run the job every weekend for the next 6 months.

Which EC2 instance purchasing option will meet these requirements MOST cost- effectively?

- A. Spot Instances
- B. Reserved Instances
- C. On-Demand Instances
- D. Dedicated Instances

Answer: A

NEW QUESTION 21

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 ECR
- 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

A company wants to predict the success of advertising campaigns by considering the color scheme of each advertisement. An ML engineer is preparing data for a neural network model. The dataset includes color information as categorical data. Which technique for feature engineering should the ML engineer use for the model?

- A. Apply label encoding to the color categories
- B. Automatically assign each color a unique integer.
- C. Implement padding to ensure that all color feature vectors have the same length.
- D. Perform dimensionality reduction on the color categories.
- E. One-hot encode the color categories to transform the color scheme feature into a binary matrix.

Answer: D

NEW QUESTION 25

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 30

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 categories
- B. Add the tags to the model packages in the SageMaker Model Registry.
- C. Create a model group for each category
- 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 categories
- G. Move the existing model groups into the collections.

Answer: A

NEW QUESTION 34

.....

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

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