



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

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 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 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 6

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 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

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 9

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 10

A company has used Amazon SageMaker to deploy a predictive ML model in production. The company is using SageMaker Model Monitor on the model. After a model update, an ML engineer notices data quality issues in the Model Monitor checks.

What should the ML engineer do to mitigate the data quality issues that Model Monitor has identified?

- A. Adjust the model's parameters and hyperparameters.
- B. Initiate a manual Model Monitor job that uses the most recent production data.
- C. Create a new baseline from the latest dataset
- D. Update Model Monitor to use the new baseline for evaluations.
- E. Include additional data in the existing training set for the mode
- F. Retrain and redeploy the model.

Answer: C

NEW QUESTION 10

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 14

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 17

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 18

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 19

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 22

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 23

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-processing
- B. Use the Latent Dirichlet Allocation (LDA) algorithm to identify and extract relevant keywords.
- C. Use Amazon SageMaker and the BlazingText algorithm
- 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 bucket
- 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 27

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 28

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 account
- B. Create a VPC peering connection between the accounts
- 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 account
- E. Create an AWS Direct Connect connection and a transit gateway
- F. Associate the VPCs from both accounts with the transit gateway
- 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 account
- I. Create an AWS Site-to-Site VPN connection with two encrypted IPsec tunnels between the accounts
- J. Set up interface VPC endpoints for Amazon S3.
- K. Provision a Redshift cluster and Amazon SageMaker Studio in a VPC in the primary account
- L. Create an S3 gateway endpoint
- M. Update the S3 bucket policy to allow IAM principals from the primary account
- N. Set up interface VPC endpoints for SageMaker and Amazon Redshift.

Answer: D

NEW QUESTION 32

A company has implemented a data ingestion pipeline for sales transactions from its ecommerce website. The company uses Amazon Data Firehose to ingest data into Amazon OpenSearch Service. The buffer interval of the Firehose stream is set for 60 seconds. An OpenSearch linear model generates real-time sales forecasts based on the data and presents the data in an OpenSearch dashboard.

The company needs to optimize the data ingestion pipeline to support sub-second latency for the real-time dashboard.

Which change to the architecture will meet these requirements?

- A. Use zero buffering in the Firehose stream
- B. Tune the batch size that is used in the PutRecordBatch operation.

- C. Replace the Firehose stream with an AWS DataSync tas
- D. Configure the task with enhanced fan-out consumers.
- E. Increase the buffer interval of the Firehose stream from 60 seconds to 120 seconds.
- F. Replace the Firehose stream with an Amazon Simple Queue Service (Amazon SQS) queue.

Answer: A

NEW QUESTION 34

An ML engineer needs to use an Amazon EMR cluster to process large volumes of data in batches. Any data loss is unacceptable. Which instance purchasing option will meet these requirements MOST cost-effectively?

- A. Run the primary node, core nodes, and task nodes on On-Demand Instances.
- B. Run the primary node, core nodes, and task nodes on Spot Instances.
- C. Run the primary node on an On-Demand Instanc
- D. Run the core nodes and task nodes on Spot Instances.
- E. Run the primary node and core nodes on On-Demand Instance
- F. Run the task nodes on Spot Instances.

Answer: D

NEW QUESTION 39

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 42

.....

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](#)