

Amazon-Web-Services

Exam Questions MLA-C01

AWS Certified Machine Learning Engineer - Associate



NEW QUESTION 1

A company has historical data that shows whether customers needed long-term support from company staff. The company needs to develop an ML model to predict whether new customers will require long-term support.

Which modeling approach should the company use to meet this requirement?

- A. Anomaly detection
- B. Linear regression
- C. Logistic regression
- D. Semantic segmentation

Answer: C

NEW QUESTION 2

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 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 3

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 4

A company has a large, unstructured dataset. The dataset includes many duplicate records across several key attributes.

Which solution on AWS will detect duplicates in the dataset with the LEAST code development?

- A. Use Amazon Mechanical Turk jobs to detect duplicates.
- B. Use Amazon QuickSight ML Insights to build a custom deduplication model.
- C. Use Amazon SageMaker Data Wrangler to pre-process and detect duplicates.
- D. Use the AWS Glue FindMatches transform to detect duplicates.

Answer: D

NEW QUESTION 5

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 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

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 8

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 9

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 10

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 sample
- 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 10

An ML engineer needs to use AWS CloudFormation to create an ML model that an Amazon SageMaker endpoint will host.

Which resource should the ML engineer declare in the CloudFormation template to meet this requirement?

- A. AWS::SageMaker::Model
- B. AWS::SageMaker::Endpoint
- C. AWS::SageMaker::NotebookInstance
- D. AWS::SageMaker::Pipeline

Answer: A

NEW QUESTION 13

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 15

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 18

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 ML engineer needs to use an Amazon SageMaker built-in algorithm to train the model. Which algorithm should the ML engineer use to meet this requirement?

- A. LightGBM
- B. Linear learner
- C. -means clustering
- D. Neural Topic Model (NTM)

Answer: B

NEW QUESTION 20

Case study

An ML engineer is developing a fraud detection model on AWS. The training dataset includes transaction logs, customer profiles, and tables from an on-premises MySQL database. The transaction logs and customer profiles are stored in Amazon S3.

The dataset has a class imbalance that affects the learning of the model's algorithm. Additionally, many of the features have interdependencies. The algorithm is not capturing all the desired underlying patterns in the data.

After the data is aggregated, the ML engineer must implement a solution to automatically detect anomalies in the data and to visualize the result.

Which solution will meet these requirements?

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

Answer: C

NEW QUESTION 23

An ML engineer has developed a binary classification model outside of Amazon SageMaker. The ML engineer needs to make the model accessible to a SageMaker Canvas user for additional tuning.

The model artifacts are stored in an Amazon S3 bucket. The ML engineer and the Canvas user are part of the same SageMaker domain.

Which combination of requirements must be met so that the ML engineer can share the model with the Canvas user? (Choose two.)

- A. The ML engineer and the Canvas user must be in separate SageMaker domains.
- B. The Canvas user must have permissions to access the S3 bucket where the model artifacts are stored.
- C. The model must be registered in the SageMaker Model Registry.
- D. The ML engineer must host the model on AWS Marketplace.
- E. The ML engineer must deploy the model to a SageMaker endpoint.

Answer: BC

NEW QUESTION 24

.....

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](#)