

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

An ML engineer needs to implement a solution to host a trained ML model. The rate of requests to the model will be inconsistent throughout the day.

The ML engineer needs a scalable solution that minimizes costs when the model is not in use. The solution also must maintain the model's capacity to respond to requests during times of peak usage.

Which solution will meet these requirements?

- A. Create AWS Lambda functions that have fixed concurrency to host the model
- B. Configure the Lambda functions to automatically scale based on the number of requests to the model.
- C. Deploy the model on an Amazon Elastic Container Service (Amazon ECS) cluster that uses AWS Fargate
- D. Set a static number of tasks to handle requests during times of peak usage.
- E. Deploy the model to an Amazon SageMaker endpoint
- F. Deploy multiple copies of the model to the endpoint
- G. Create an Application Load Balancer to route traffic between the different copies of the model at the endpoint.
- H. Deploy the model to an Amazon SageMaker endpoint
- I. Create SageMaker endpoint auto scaling policies that are based on Amazon CloudWatch metrics to adjust the number of instances dynamically.

Answer: D

NEW QUESTION 3

A company is building a deep learning model on Amazon SageMaker. The company uses a large amount of data as the training dataset. The company needs to optimize the model's hyperparameters to minimize the loss function on the validation dataset.

Which hyperparameter tuning strategy will accomplish this goal with the LEAST computation time?

- A. Hyperband
- B. Grid search
- C. Bayesian optimization
- D. Random search

Answer: A

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 table
- 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 bucket
- 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 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 Manager
- B. Create an AWS Lambda function to perform the rotation.
- C. Store the tokens in AWS Systems Manager Parameter Store
- 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 6

A company has a team of data scientists who use Amazon SageMaker notebook instances to test ML models. When the data scientists need new permissions, the company attaches the permissions to each individual role that was created during the creation of the SageMaker notebook instance. The company needs to centralize management of the team's permissions. Which solution will meet this requirement?

- A. Create a single IAM role that has the necessary permission
- B. Attach the role to each notebook instance that the team uses.
- C. Create a single IAM group
- D. Add the data scientists to the group
- E. Associate the group with each notebook instance that the team uses.
- F. Create a single IAM user
- G. Attach the AdministratorAccess AWS managed IAM policy to the user
- H. Configure each notebook instance to use the IAM user.
- I. Create a single IAM group
- J. Add the data scientists to the group
- K. Create an IAM role
- L. Attach the AdministratorAccess AWS managed IAM policy to the role
- M. Associate the role with the group
- N. Associate the group with each notebook instance that the team uses.

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 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

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

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 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 11

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 14

A company is using Amazon SageMaker to create ML models. The company's data scientists need fine-grained control of the ML workflows that they orchestrate. The data scientists also need the ability to visualize SageMaker jobs and workflows as a directed acyclic graph (DAG). The data scientists must keep a running history of model discovery experiments and must establish model governance for auditing and compliance verifications.

Which solution will meet these requirements?

- A. Use AWS CodePipeline and its integration with SageMaker Studio to manage the entire ML workflow
- B. Use SageMaker ML Lineage Tracking for the running history of experiments and for auditing and compliance verifications.
- C. Use AWS CodePipeline and its integration with SageMaker Experiments to manage the entire ML workflow
- D. Use SageMaker Experiments for the running history of experiments and for auditing and compliance verifications.
- E. Use SageMaker Pipelines and its integration with SageMaker Studio to manage the entire ML workflow
- F. Use SageMaker ML Lineage Tracking for the running history of experiments and for auditing and compliance verifications.
- G. Use SageMaker Pipelines and its integration with SageMaker Experiments to manage the entire ML workflow
- H. Use SageMaker Experiments for the running history of experiments and for auditing and compliance verifications.

Answer: C

NEW QUESTION 15

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 19

A company needs to give its ML engineers appropriate access to training data. The ML engineers must access training data from only their own business group. The ML engineers must not be allowed to access training data from other business groups.

The company uses a single AWS account and stores all the training data in Amazon S3 buckets. All ML model training occurs in Amazon SageMaker.

Which solution will provide the ML engineers with the appropriate access?

- A. Enable S3 bucket versioning.
- B. Configure S3 Object Lock settings for each user.
- C. Add cross-origin resource sharing (CORS) policies to the S3 buckets.
- D. Create IAM policies
- E. Attach the policies to IAM users or IAM roles.

Answer: D

NEW QUESTION 20

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_classifier for the predictor_type hyperparameter.

What should the company do to MINIMIZE false positives?

- A. Set the value of the weight decay hyperparameter to zero.

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 27

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 28

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

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

NEW QUESTION 31

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 36

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 transfor
- H. Use SageMaker Model Monitor to monitor the data quality and to send alerts.

Answer: D

NEW QUESTION 38

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 43

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 44

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

Answer: A

NEW QUESTION 45

.....

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](#)