

MLA-C01 Dumps

AWS Certified Machine Learning Engineer - Associate

<https://www.certleader.com/MLA-C01-dumps.html>



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

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 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 Autopilot
- B. Report the model's achieved performance.
- C. Implement custom scripts to perform data pre-processing, multiple linear regression, and performance evaluation
- D. Run the scripts on Amazon EC2 instances.
- E. Configure Amazon Macie to analyze the dataset and to create a model
- F. Report the model's achieved performance.
- G. Select a model from Amazon Bedrock
- H. Tune the model with the data
- I. Report the model's achieved performance.

Answer: A

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

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 buckets
- 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 6

A company is running ML models on premises by using custom Python scripts and proprietary datasets. The company is using PyTorch. The model building requires unique domain knowledge. The company needs to move the models to AWS. Which solution will meet these requirements with the LEAST effort?

- A. Use SageMaker built-in algorithms to train the proprietary datasets.
- B. Use SageMaker script mode and premade images for ML frameworks.
- C. Build a container on AWS that includes custom packages and a choice of ML frameworks.
- D. Purchase similar production models through AWS Marketplace.

Answer: B

NEW QUESTION 7

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 8

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 CloudTrail
- F. Use an Amazon CloudWatch dashboard for monitoring
- 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 monitoring
- I. Set up a CloudWatch alarm to provide notification when the threshold is breached.

Answer: D

NEW QUESTION 9

A company has a conversational AI assistant that sends requests through Amazon Bedrock to an Anthropic Claude large language model (LLM). Users report that when they ask similar questions multiple times, they sometimes receive different answers. An ML engineer needs to improve the responses to be more consistent and less random. Which solution will meet these requirements?

- A. Increase the temperature parameter and the top_k parameter.
- B. Increase the temperature parameter
- C. Decrease the top_k parameter.
- D. Decrease the temperature parameter
- E. Increase the top_k parameter.
- F. Decrease the temperature parameter and the top_k parameter.

Answer: D

NEW QUESTION 10

An ML engineer is using a training job to fine-tune a deep learning model in Amazon SageMaker Studio. The ML engineer previously used the same pre-trained model with a similar dataset. The ML engineer expects vanishing gradient, underutilized GPU, and overfitting problems. The ML engineer needs to implement a solution to detect these issues and to react in predefined ways when the issues occur. The solution also must provide comprehensive real-time metrics during the training. Which solution will meet these requirements with the LEAST operational overhead?

- A. Use TensorBoard to monitor the training job
- B. Publish the findings to an Amazon Simple Notification Service (Amazon SNS) topic
- C. Create an AWS Lambda function to consume the findings and to initiate the predefined actions.
- D. Use Amazon CloudWatch default metrics to gain insights about the training job
- E. Use the metrics to invoke an AWS Lambda function to initiate the predefined actions.
- F. Expand the metrics in Amazon CloudWatch to include the gradients in each training step
- G. Use the metrics to invoke an AWS Lambda function to initiate the predefined actions.
- H. Use SageMaker Debugger built-in rules to monitor the training job
- I. Configure the rules to initiate the predefined actions.

Answer: D

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 15

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 18

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 21

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 inferenc
- B. Use SageMaker Model Monitor for notifications about model quality.
- C. Use SageMaker batch transform for inferenc
- D. Use SageMaker Model Monitor for notifications about model quality.
- E. Use SageMaker Serverless Inference for inferenc
- F. Use SageMaker Inference Recommender for notifications about model quality.
- G. Keep using SageMaker Asynchronous Inference for inferenc
- H. Use SageMaker Inference Recommender for notifications about model quality.

Answer: A

NEW QUESTION 22

FILL IN THE BLANK

A company stores time-series data about user clicks in an Amazon S3 bucket. The raw data consists of millions of rows of user activity every day. ML engineers access the data to develop their ML models.

The ML engineers need to generate daily reports and analyze click trends over the past 3 days by using Amazon Athena. The company must retain the data for 30 days before archiving the data.

Which solution will provide the HIGHEST performance for data retrieval?

- A. Keep all the time-series data without partitioning in the S3 bucke
- B. Manually move data that is older than 30 days to separate S3 buckets.
- C. Create AWS Lambda functions to copy the time-series data into separate S3 bucket
- D. Apply S3 Lifecycle policies to archive data that is older than 30 days to S3 Glacier Flexible Retrieval.
- E. Organize the time-series data into partitions by date prefix in the S3 bucke
- F. Apply S3 Lifecycle policies to archive partitions that are older than 30 days to S3 Glacier Flexible Retrieval.
- G. Put each day's time-series data into its own S3 bucke
- H. Use S3 Lifecycle policies to archive S3 buckets that hold data that is older than 30 days to S3 Glacier Flexible Retrieval.

Answer: C

NEW QUESTION 24

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 25

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.
- B. Increase the number of training epochs.
- C. Increase the value of the target_precision hyperparameter.
- D. Change the value of the predictor_type hyperparameter to regressor.

Answer: C

NEW QUESTION 28

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...

Access the store to build datasets for training.

Create a feature group.

Ingest the records.

Step 2:

Select...

Access the store to build datasets for training.

Create a feature group.

Ingest the records.

Step 3:

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 33

A company has a binary classification model in production. An ML engineer needs to develop a new version of the model. The new model version must maximize correct predictions of positive labels and negative labels. The ML engineer must use a metric to recalibrate the model to meet these requirements. Which metric should the ML engineer use for the model recalibration?

- A. Accuracy
- B. Precision
- C. Recall
- D. Specificity

Answer: A

NEW QUESTION 37

An ML engineer is evaluating several ML models and must choose one model to use in production. The cost of false negative predictions by the models is much higher than the cost of false positive predictions. Which metric finding should the ML engineer prioritize the MOST when choosing the model?

- A. Low precision
- B. High precision
- C. Low recall
- D. High recall

Answer: D

NEW QUESTION 42

An ML engineer trained an ML model on Amazon SageMaker to detect automobile accidents from closed-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 trainingdataset.
- 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 43

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 47

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 52

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 57

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 61

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

Answer: D

NEW QUESTION 66

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 69

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 70

.....

Thank You for Trying Our Product

* 100% Pass or Money Back

All our products come with a 90-day Money Back Guarantee.

* One year free update

You can enjoy free update one year. 24x7 online support.

* Trusted by Millions

We currently serve more than 30,000,000 customers.

* Shop Securely

All transactions are protected by VeriSign!

100% Pass Your MLA-C01 Exam with Our Prep Materials Via below:

<https://www.certleader.com/MLA-C01-dumps.html>