DBCLOUD
EXAMS

QUESTIONS & ANSWERS

verified



Technology Provider: Amazon

Exam: AWS Certified Machine Learning Engineer - Associate MLA-C01

Number of questions in the database: 275

Exam Version: NOV.2026

Topic 1

Question #1Topic 1

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 use the central model registry to manage different versions of models in the application.

Which action will meet this requirement with the LEAST operational overhead?

A. Create a separate Amazon Elastic Container Registry (Amazon ECR) repository for each model.

B. Use Amazon Elastic Container Registry (Amazon ECR) and unique tags for each model version.

C. Use the SageMaker Model Registry and model groups to catalog the models.

D. Use the SageMaker Model Registry and unique tags for each model version.

Correct Answer: C

Question #2Topic 1

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 is experimenting with consecutive training jobs.

How can the company MINIMIZE infrastructure startup times for these jobs?

https://dbcloud.academy/

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.

Correct Answer: B

Question #3Topic 1

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 registry. Create tracking entities for the approval process.

C. Use SageMaker Model Monitor to evaluate the performance of the model and to manage the approval.

D. Use SageMaker Pipelines. When a model version is registered, use the AWS SDK to change the approval status to "Approved."

Correct Answer: D

Question #4Topic 1

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.

https://dbcloud.academy/

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.

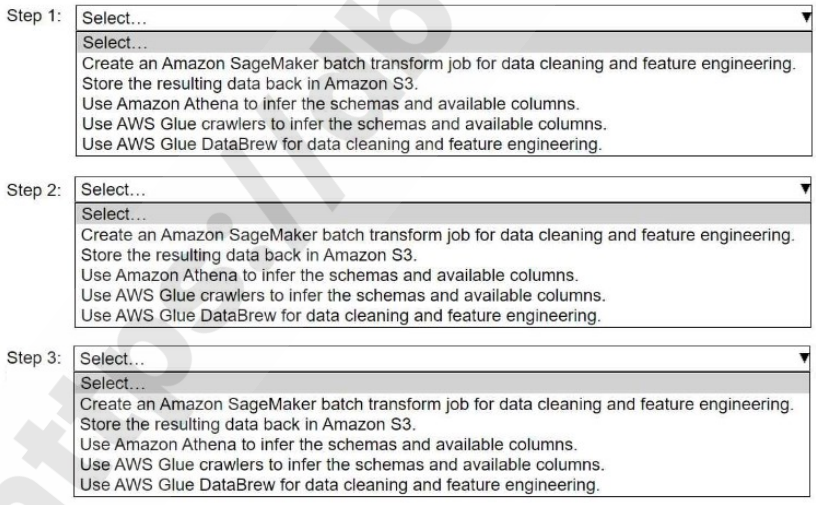
C. Use AWS Glue Data Quality to monitor bias.

D. Use SageMaker notebooks to compare the bias.

Correct Answer: A

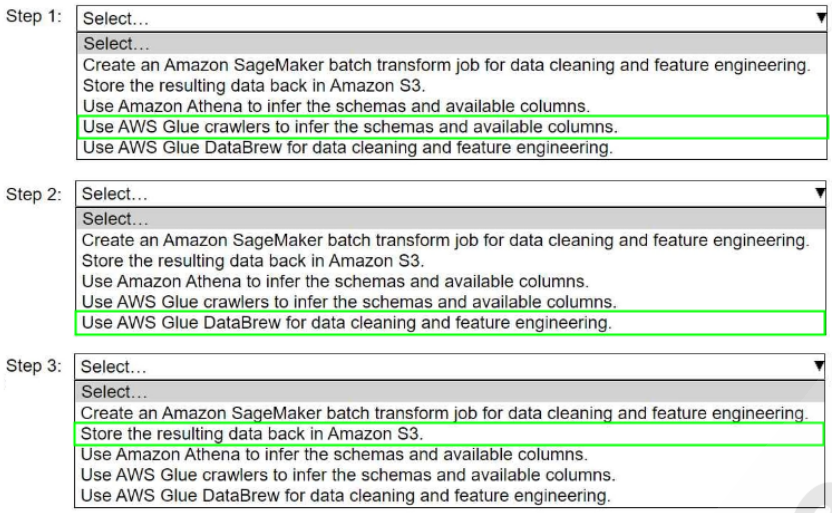
Question #5*Topic 1*

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.



Correct Answer:

https://dbcloud.academy/



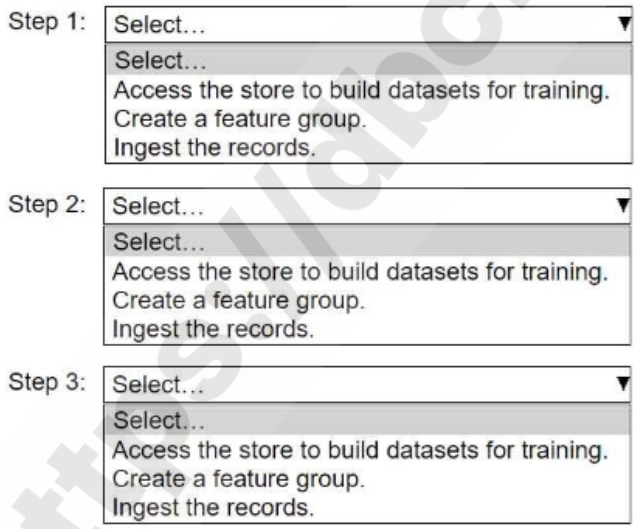
Question #6*Topic 1*

HOTSPOT -

An ML engineer needs to use Amazon SageMaker Feature Store to create and manage features to train a model.

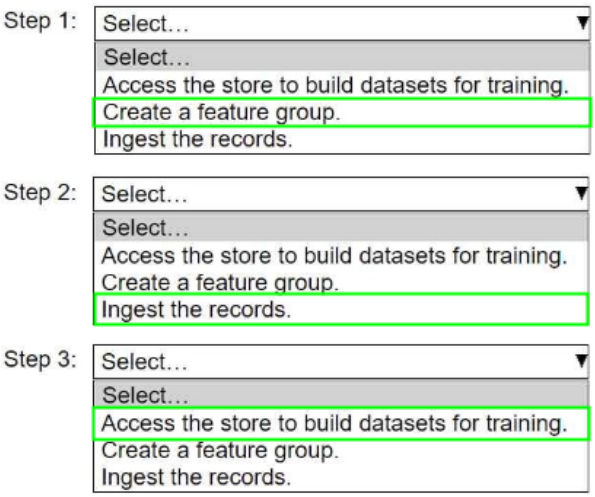
Select and order the steps from the following list to create and use the features in Feature Store. Each step should be selected one time. (Select and order three.)

- Access the store to build datasets for training.
- Create a feature group.
- Ingest the records.



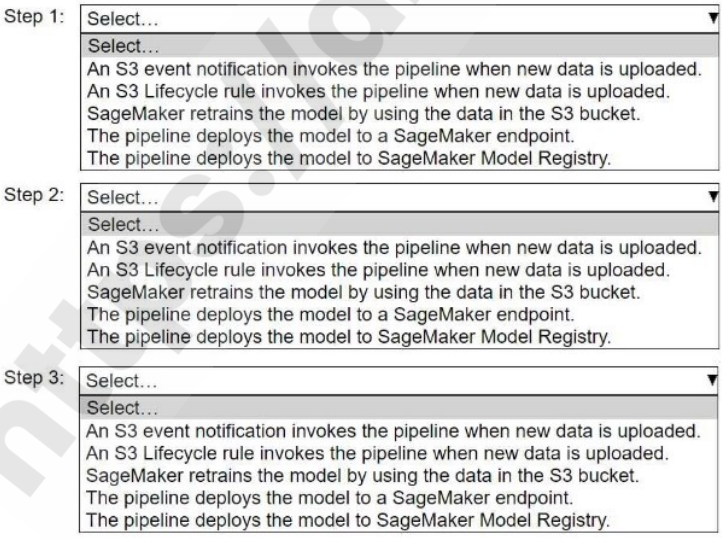
Correct Answer:

https://dbcloud.academy/



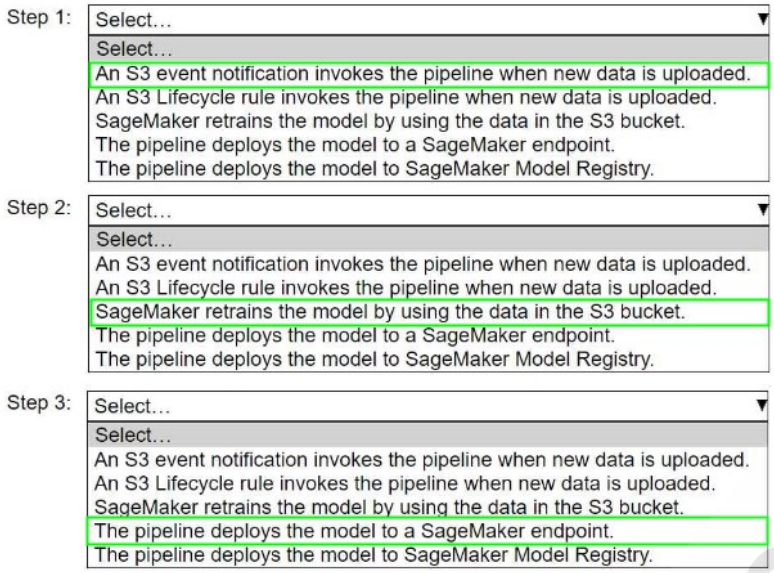
Question #7*Topic 1*

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.



Correct Answer:

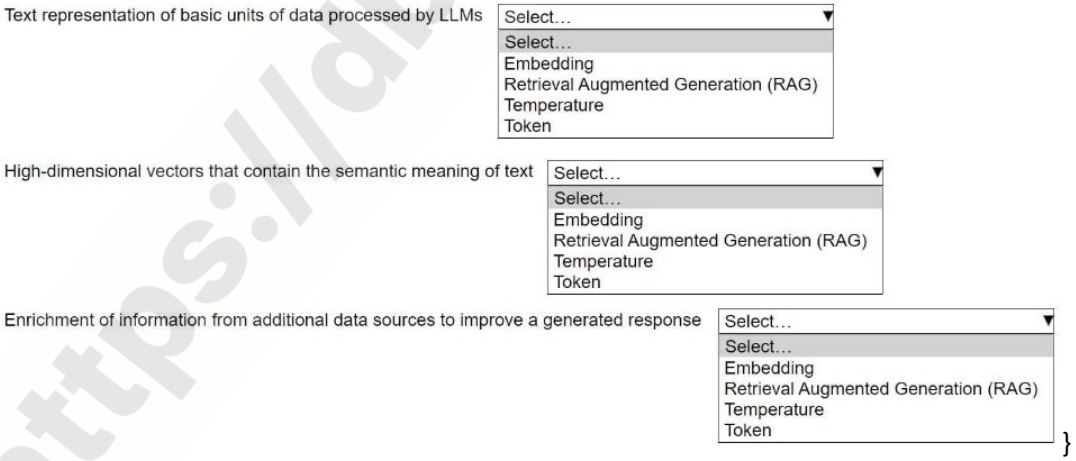
https://dbcloud.academy/



Question #8*Topic 1*

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.)

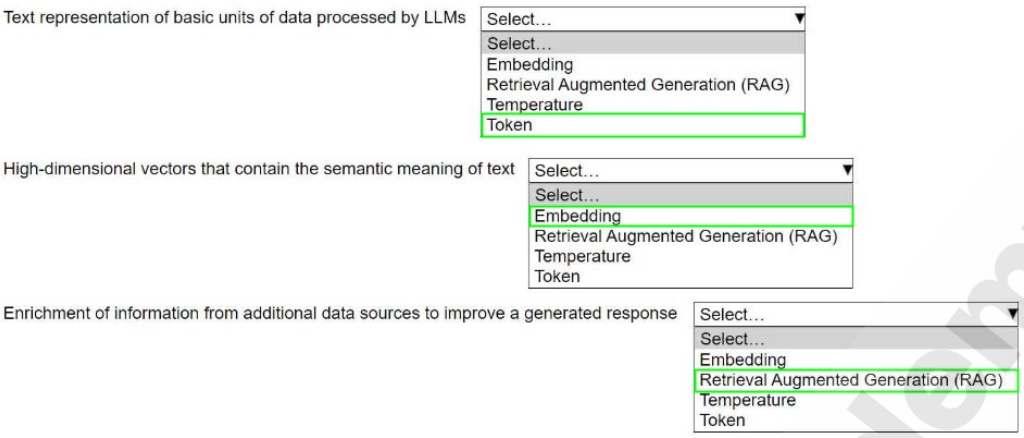
- Embedding
- Retrieval Augmented Generation (RAG)
- Temperature
- Token



}

Correct Answer:

https://dbcloud.academy/



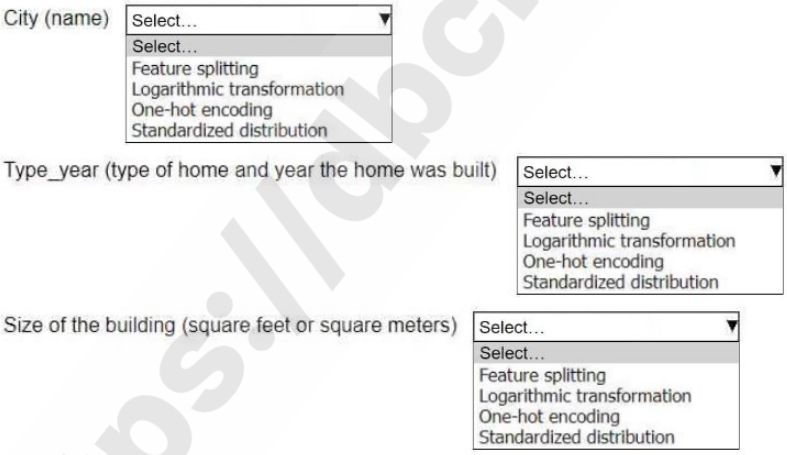
Question #9*Topic 1*

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:

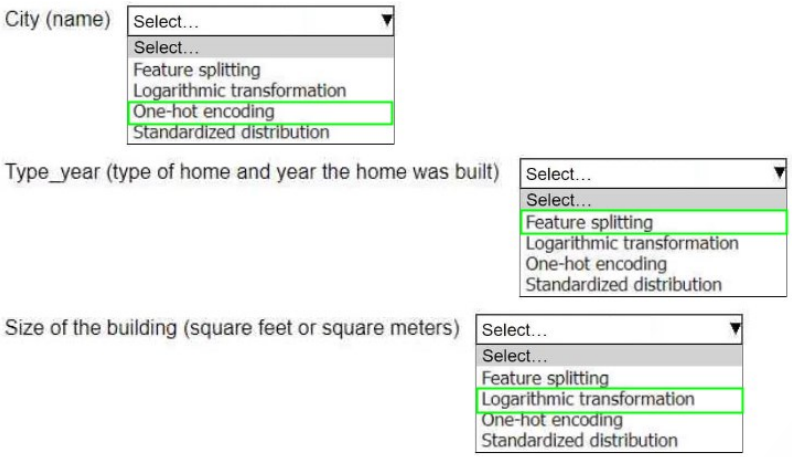
- Feature splitting
- Logarithmic transformation
- One-hot encoding
- Standardized distribution

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.)



Correct Answer:

https://dbcloud.academy/



Question #10Topic 1

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

Correct Answer: D

Question #11Topic 1

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.

https://dbcloud.academy/

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. Use Amazon QuickSight to visualize the result.

C. Use Amazon SageMaker Data Wrangler to automatically detect the anomalies and to visualize the result.

D. Use AWS Batch to automatically detect the anomalies. Use Amazon QuickSight to visualize the result.

Correct Answer: C

Question #12Topic 1

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.

Correct Answer: C

https://dbcloud.academy/