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
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 mode
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 Fargat
D. Set a static number of tasks to handle requests during times of peak usage.
E. Deploy the model to an Amazon SageMaker endpoin
F. Deploy multiple copies of the model to the endpoin
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 endpoin
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
An ML engineer is using Amazon SageMaker to train a deep learning model that requires distributed training. After some training attempts, the ML engineer observes that the instances are not performing as expected. The ML engineer identifies communication overhead between the training instances.
What should the ML engineer do to MINIMIZE the communication overhead between the instances?

A. Place the instances in the same VPC subne
B. Store the data in a different AWS Region from where the instances are deployed.
C. Place the instances in the same VPC subnet but in different Availability Zone
D. Store the data in a different AWS Region from where the instances are deployed.
E. Place the instances in the same VPC subne
F. Store the data in the same AWS Region and Availability Zone where the instances are deployed.
G. Place the instances in the same VPC subne
H. Store the data in the same AWS Region but in a different Availability Zone from where the instances are deployed.

Answer: C

NEW QUESTION 4
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. Hyperbaric!
B. Grid search
C. Bayesian optimization
D. Random search

Answer: A

NEW QUESTION 5
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 6
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 ste
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 step
H. Automate the pipeline by using Amazon EventBridge.

Answer: D

NEW QUESTION 7
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 grou
D. Add the data scientists to the grou
E. Associate the group with each notebook instance that the team uses.
F. Create a single IAM use
G. Attach the AdministratorAccess AWS managed IAM policy to the use
H. Configure each notebook instance to use the IAM user.
I. Create a single IAM grou
J. Add the data scientists to the grou
K. Create an IAM rol
L. Attach the AdministratorAccess AWS managed IAM policy to the rol
M. Associate the role with the grou
N. Associate the group with each notebook instance that the team uses.

Answer: A

NEW QUESTION 8
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 9
A company uses a hybrid cloud environment. A model that is deployed on premises uses data in Amazon 53 to provide customers with a live conversational engine.
The model is using sensitive data. An ML engineer needs to implement a solution to identify and remove the sensitive data.
Which solution will meet these requirements with the LEAST operational overhead?

A. Deploy the model on Amazon SageMake
B. Create a set of AWS Lambda functions to identify and remove the sensitive data.
C. Deploy the model on an Amazon Elastic Container Service (Amazon ECS) cluster that uses AWS Fargat
D. Create an AWS Batch job to identify and remove the sensitive data.
E. Use Amazon Macie to identify the sensitive dat
F. Create a set of AWS Lambda functions to remove the sensitive data.
G. Use Amazon Comprehend to identify the sensitive dat
H. Launch Amazon EC2 instances to remove the sensitive data.

Answer: C

NEW QUESTION 10
A company has deployed an XGBoost prediction model in production to predict if a customer is likely to cancel a subscription. The company uses Amazon SageMaker Model Monitor to detect deviations in the F1 score.
During a baseline analysis of model quality, the company recorded a threshold for the F1 score. After several months of no change, the model's F1 score

decreases significantly.
What could be the reason for the reduced F1 score?

A. Concept drift occurred in the underlying customer data that was used for predictions.
B. The model was not sufficiently complex to capture all the patterns in the original baseline data.
C. The original baseline data had a data quality issue of missing values.
D. Incorrect ground truth labels were provided to Model Monitor during the calculation of the baseline.

Answer: A

NEW QUESTION 10
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 11
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 polic
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 scalin
E. Host the model on the EKS cluste
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) cluste
H. Specify a list of images to process for each AWS Batch job.

Answer: B

NEW QUESTION 16
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 18
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?

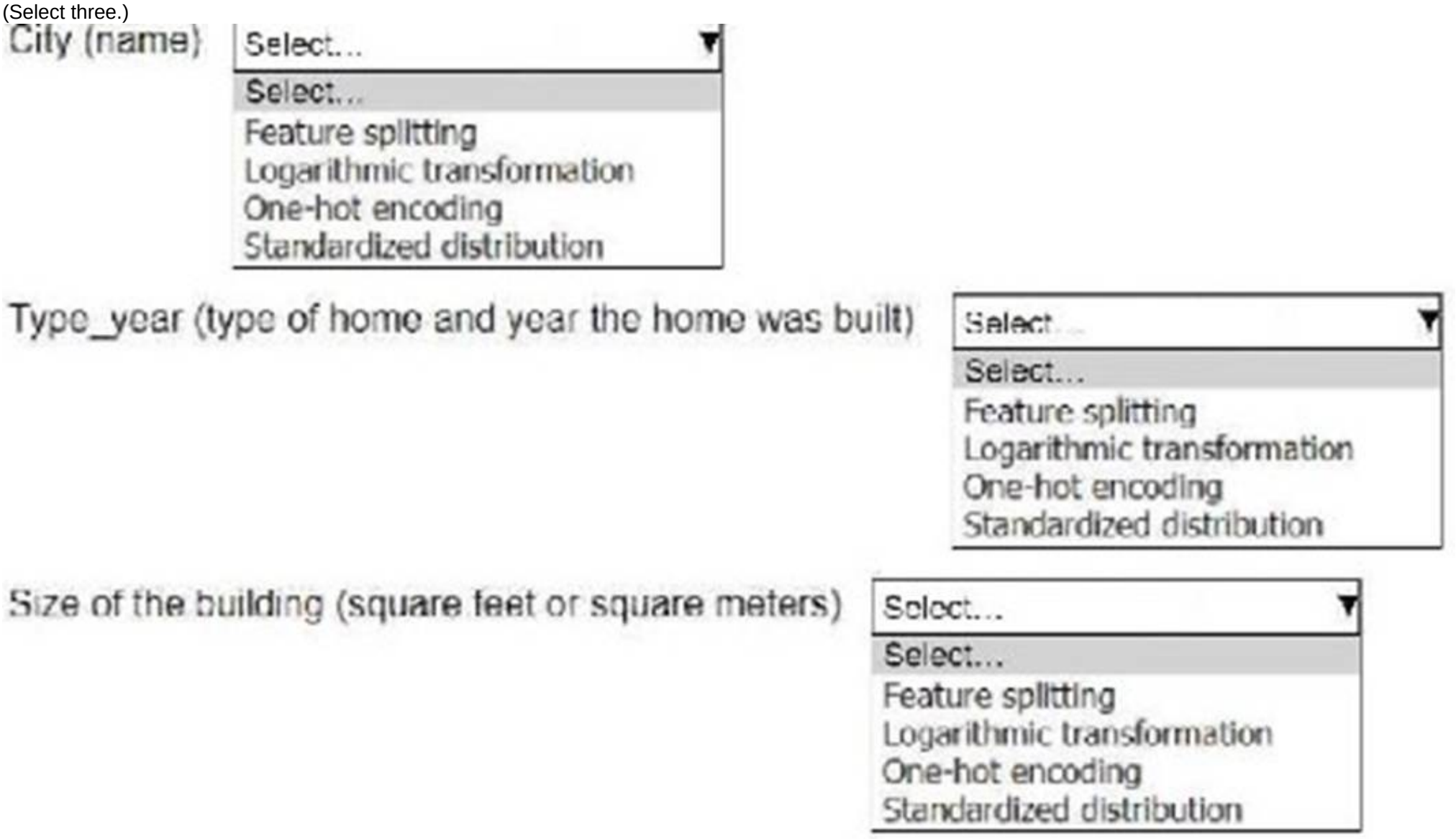
A. Use TensorBoard to monitor the training jo
B. Publish the findings to an Amazon Simple Notification Service (Amazon SNS) topi
C. Create an AWS Lambda function to consume the findings and to initiate the predefined actions.
D. Use Amazon CloudWatch default metrics to gain insights about the training jo
E. Use themetrics 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 ste
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 jo
I. Configure the rules to initiate the predefined actions.

Answer: D

NEW QUESTION 21
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

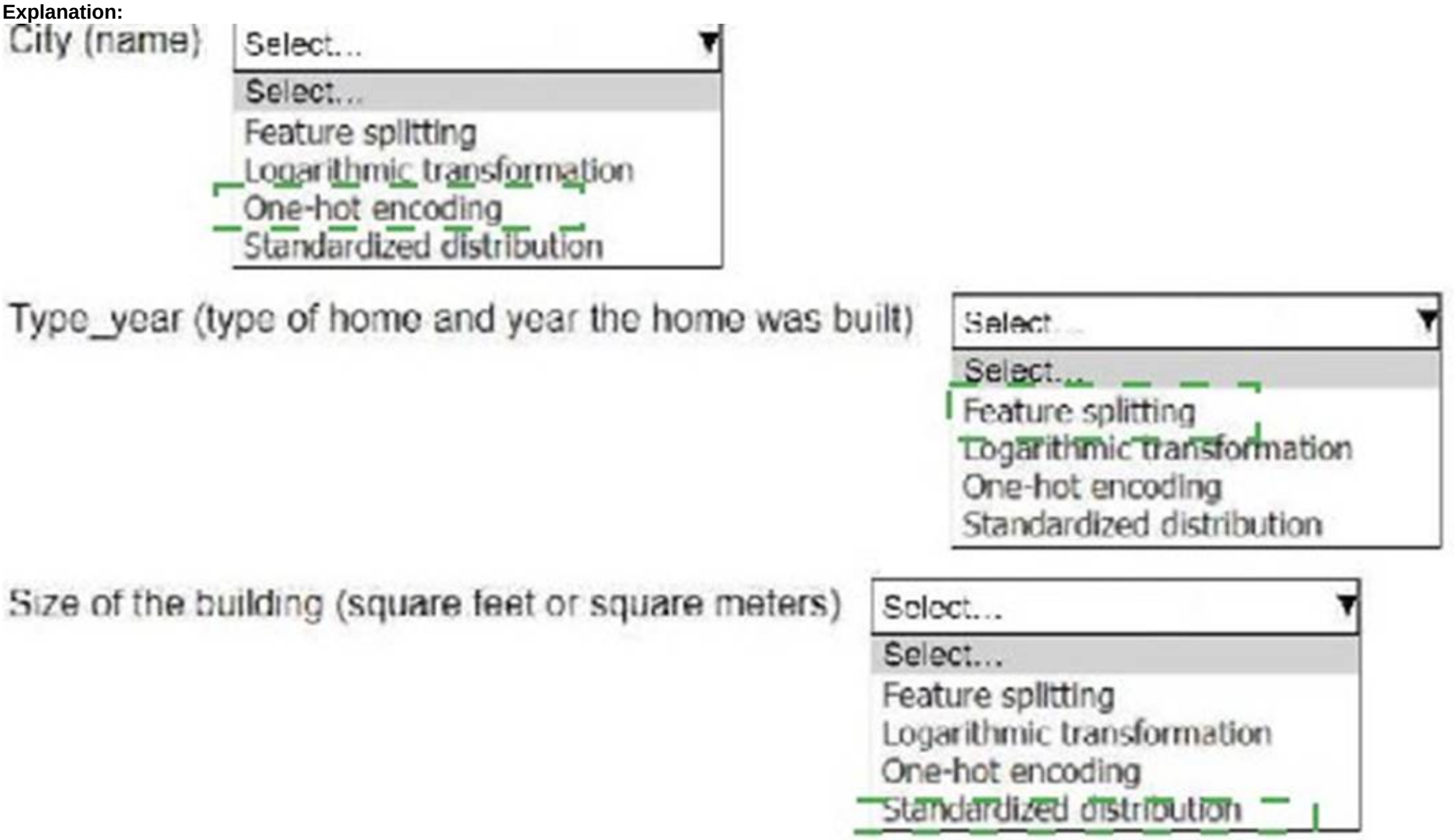
Select the correct feature engineering techniques for the following list of features. Each feature engineering technique should be selected one time or not at all (Select three.)



A. Mastered
B. Not Mastered

Answer: A

Explanation:



NEW QUESTION 26
A company has a Retrieval Augmented Generation (RAG) application that uses a vector database to store embeddings of documents. The company must migrate the application to AWS and must implement a solution that provides semantic search of text files. The company has already migrated the text repository to an Amazon S3 bucket.
Which solution will meet these requirements?

A. Use an AWS Batch job to process the files and generate embedding
B. Use AWS Glue to store the embedding
C. Use SQL queries to perform the semantic searches.
D. Use a custom Amazon SageMaker notebook to run a custom script to generate embedding
E. Use SageMaker Feature Store to store the embedding
F. Use SQL queries to perform the semantic searches.
G. Use the Amazon Kendra S3 connector to ingest the documents from the S3 bucket into Amazon Kendr
H. Query Amazon Kendra to perform the semantic searches.
I. Use an Amazon Textract asynchronous job to ingest the documents from the S3 bucke
J. Query Amazon Textract to perform the semantic searches.

Answer: C

NEW QUESTION 28
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 33
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 catalogthe models.
D. Use the SageMaker Model Registry and unique tags for each model version.

Answer: C

NEW QUESTION 35
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.
Before the ML engineer trains the model, the ML engineer must resolve the issue of the imbalanced data.
Which solution will meet this requirement with the LEAST operational effort?

A. Use Amazon Athena to identify patterns that contribute to the imbalanc
B. Adjust the dataset accordingly.
C. Use Amazon SageMaker Studio Classic built-in algorithms to process the imbalanced dataset.
D. Use AWS Glue DataBrew built-in features to oversample the minority class.
E. Use the Amazon SageMaker Data Wrangler balance data operation to oversample the minority class.

Answer: D

NEW QUESTION 39
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 42
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 45
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 49
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 52
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 53
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 58
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 60
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 anomalie
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 anomalie
F. Use Amazon QuickSight to visualize the result.

Answer: C

NEW QUESTION 62
A company that has hundreds of data scientists is using Amazon SageMaker to create ML models. The models are in model groups in the SageMaker Model Registry.
The data scientists are grouped into three categories: computer vision, natural language processing (NLP), and speech recognition. An ML engineer needs to implement a solution to organize the existing models into these groups to improve model discoverability at scale. The solution must not affect the integrity of the model artifacts and their existing groupings.
Which solution will meet these requirements?

A. Create a custom tag for each of the three categorie
B. Add the tags to the model packages in the SageMaker Model Registry.
C. Create a model group for each categor
D. Move the existing models into these category model groups.
E. Use SageMaker ML Lineage Tracking to automatically identify and tag which model groups should contain the models.
F. Create a Model Registry collection for each of the three categorie
G. Move the existing model groups into the collections.

Answer: A

NEW QUESTION 63
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 65
.....

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