

MLS-C01^{Q&As}

AWS Certified Machine Learning - Specialty (MLS-C01)

Pass Amazon MLS-C01 Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

<https://www.leads4pass.com/aws-certified-machine-learning-specialty.html>

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by Amazon
Official Exam Center

- ⚙️ **Instant Download** After Purchase
- ⚙️ **100% Money Back** Guarantee
- ⚙️ **365 Days** Free Update
- ⚙️ **800,000+** Satisfied Customers



QUESTION 1

A Machine Learning Specialist needs to create a data repository to hold a large amount of time-based training data for a new model. In the source system, new files are added every hour. Throughout a single 24-hour period, the volume of hourly updates will change significantly. The Specialist always wants to train on the last 24 hours of the data. Which type of data repository is the MOST cost-effective solution?

- A. An Amazon EBS-backed Amazon EC2 instance with hourly directories
- B. An Amazon RDS database with hourly table partitions
- C. An Amazon S3 data lake with hourly object prefixes
- D. An Amazon EMR cluster with hourly hive partitions on Amazon EBS volumes

Correct Answer: C

QUESTION 2

A machine learning (ML) specialist is training a multilayer perceptron (MLP) on a dataset with multiple classes. The target class of interest is unique compared to the other classes in the dataset, but it does not achieve an acceptable recall metric. The ML specialist varies the number and size of the MLP's hidden layers, but the results do not improve significantly.

Which solution will improve recall in the LEAST amount of time?

- A. Add class weights to the MLP's loss function, and then retrain.
- B. Gather more data by using Amazon Mechanical Turk, and then retrain.
- C. Train a k-means algorithm instead of an MLP.
- D. Train an anomaly detection model instead of an MLP.

Correct Answer: A

QUESTION 3

A global financial company is using machine learning to automate its loan approval process. The company has a dataset of customer information. The dataset contains some categorical fields, such as customer location by city and housing

status. The dataset also includes financial fields in different units, such as account balances in US dollars and monthly interest in US cents.

The company's data scientists are using a gradient boosting regression model to infer the credit score for each customer. The model has a training accuracy of 99% and a testing accuracy of 75%. The data scientists want to improve the model's testing accuracy.

Which process will improve the testing accuracy the MOST?

- A. Use a one-hot encoder for the categorical fields in the dataset. Perform standardization on the financial fields in the dataset. Apply L1 regularization to the data.
- B. Use tokenization of the categorical fields in the dataset. Perform binning on the financial fields in the dataset. Remove the outliers in the data by using the z-score.
- C. Use a label encoder for the categorical fields in the dataset. Perform L1 regularization on the financial fields in the dataset. Apply L2 regularization to the data.
- D. Use a logarithm transformation on the categorical fields in the dataset. Perform binning on the financial fields in the dataset. Use imputation to populate missing values in the dataset.

Correct Answer: A

QUESTION 4

A data scientist is using an Amazon SageMaker notebook instance and needs to securely access data stored in a specific Amazon S3 bucket.

How should the data scientist accomplish this?

- A. Add an S3 bucket policy allowing GetObject, PutObject, and ListBucket permissions to the Amazon SageMaker notebook ARN as principal.
- B. Encrypt the objects in the S3 bucket with a custom AWS Key Management Service (AWS KMS) key that only the notebook owner has access to.
- C. Attach the policy to the IAM role associated with the notebook that allows GetObject, PutObject, and ListBucket operations to the specific S3 bucket.
- D. Use a script in a lifecycle configuration to configure the AWS CLI on the instance with an access key ID and secret.

Correct Answer: C

QUESTION 5

A company has raw user and transaction data stored in AmazonS3 a MySQL database, and Amazon RedShift A Data Scientist needs to perform an analysis by joining the three datasets from Amazon S3, MySQL, and Amazon RedShift, and then calculating the average-of a few selected columns from the joined data

Which AWS service should the Data Scientist use?

- A. Amazon Athena
- B. Amazon Redshift Spectrum
- C. AWS Glue
- D. Amazon QuickSight

Correct Answer: A

[MLS-C01 Practice Test](#)

[MLS-C01 Study Guide](#)

[MLS-C01 Exam Questions](#)