

## MLS-C01<sup>Q&As</sup>

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**

During mini-batch training of a neural network for a classification problem, a Data Scientist notices that training accuracy oscillates. What is the MOST likely cause of this issue?

- A. The class distribution in the dataset is imbalanced
- B. Dataset shuffling is disabled
- C. The batch size is too big
- D. The learning rate is very high

Correct Answer: D

Reference: <https://towardsdatascience.com/deep-learning-personal-notes-part-1-lesson-2-8946fe970b95>

---

**QUESTION 2**

A Machine Learning Specialist working for an online fashion company wants to build a data ingestion solution for the company's Amazon S3-based data lake.

The Specialist wants to create a set of ingestion mechanisms that will enable future capabilities comprised of:

1.  
Real-time analytics
2.  
Interactive analytics of historical data
3.  
Clickstream analytics
4.  
Product recommendations

Which services should the Specialist use?

- A. AWS Glue as the data catalog; Amazon Kinesis Data Streams and Amazon Kinesis Data Analytics for real-time data insights; Amazon Kinesis Data Firehose for delivery to Amazon ES for clickstream analytics; Amazon EMR to generate personalized product recommendations
- B. Amazon Athena as the data catalog; Amazon Kinesis Data Streams and Amazon Kinesis Data Analytics for near-realtime data insights; Amazon Kinesis Data Firehose for clickstream analytics; AWS Glue to generate personalized product recommendations
- C. AWS Glue as the data catalog; Amazon Kinesis Data Streams and Amazon Kinesis Data Analytics for historical data insights; Amazon Kinesis Data Firehose for delivery to Amazon ES for clickstream analytics; Amazon EMR to generate personalized product recommendations

D. Amazon Athena as the data catalog; Amazon Kinesis Data Streams and Amazon Kinesis Data Analytics for historical data insights; Amazon DynamoDB streams for clickstream analytics; AWS Glue to generate personalized product recommendations

Correct Answer: A

AWS Glue is a fully managed extract, transform, and load (ETL) service that makes it easy to move data between data stores. It can be used as a data catalog to store metadata information about the data in the data lake. Amazon Kinesis Data Streams and Amazon Kinesis Data Analytics can be used together to collect, process, and analyze real-time streaming data. Amazon Kinesis Data Firehose can be used to deliver streaming data to destinations such as Amazon ES for clickstream analytics. Finally, Amazon EMR can be used to run big data frameworks such as Apache Spark and Apache Hadoop to generate personalized product recommendations.

---

### QUESTION 3

A machine learning (ML) specialist wants to secure calls to the Amazon SageMaker Service API. The specialist has configured Amazon VPC with a VPC interface endpoint for the Amazon SageMaker Service API and is attempting to secure traffic from specific sets of instances and IAM users. The VPC is configured with a single public subnet.

Which combination of steps should the ML specialist take to secure the traffic? (Choose two.)

- A. Add a VPC endpoint policy to allow access to the IAM users.
- B. Modify the users' IAM policy to allow access to Amazon SageMaker Service API calls only.
- C. Modify the security group on the endpoint network interface to restrict access to the instances.
- D. Modify the ACL on the endpoint network interface to restrict access to the instances.
- E. Add a SageMaker Runtime VPC endpoint interface to the VPC.

Correct Answer: AC

Reference: <https://aws.amazon.com/blogs/machine-learning/securing-all-amazon-sagemaker-api-calls-with-aws-privatelink/>

---

### QUESTION 4

A company is converting a large number of unstructured paper receipts into images. The company wants to create a model based on natural language processing (NLP) to find relevant entities such as date, location, and notes, as well as some custom entities such as receipt numbers.

The company is using optical character recognition (OCR) to extract text for data labeling. However, documents are in different structures and formats, and the company is facing challenges with setting up the manual workflows for each document type. Additionally, the company trained a named entity recognition (NER) model for custom entity detection using a small sample size. This model has a very low confidence score and will require retraining with a large dataset.

Which solution for text extraction and entity detection will require the LEAST amount of effort?

- A. Extract text from receipt images by using Amazon Textract. Use the Amazon SageMaker BlazingText algorithm to train on the text for entities and custom entities.
- B. Extract text from receipt images by using a deep learning OCR model from the AWS Marketplace. Use the NER deep

learning model to extract entities.

C. Extract text from receipt images by using Amazon Textract. Use Amazon Comprehend for entity detection, and use Amazon Comprehend custom entity recognition for custom entity detection.

D. Extract text from receipt images by using a deep learning OCR model from the AWS Marketplace. Use Amazon Comprehend for entity detection, and use Amazon Comprehend custom entity recognition for custom entity detection.

Correct Answer: C

Reference: <https://aws.amazon.com/blogs/machine-learning/building-an-nlp-powered-search-index-with-amazon-textract-and-amazon-comprehend/>

---

## QUESTION 5

A machine learning (ML) specialist wants to create a data preparation job that uses a PySpark script with complex window aggregation operations to create data for training and testing. The ML specialist needs to evaluate the impact of the number of features and the sample count on model performance.

Which approach should the ML specialist use to determine the ideal data transformations for the model?

A. Add an Amazon SageMaker Debugger hook to the script to capture key metrics. Run the script as an AWS Glue job.

B. Add an Amazon SageMaker Experiments tracker to the script to capture key metrics. Run the script as an AWS Glue job.

C. Add an Amazon SageMaker Debugger hook to the script to capture key parameters. Run the script as a SageMaker processing job.

D. Add an Amazon SageMaker Experiments tracker to the script to capture key parameters. Run the script as a SageMaker processing job.

Correct Answer: D

<https://docs.aws.amazon.com/sagemaker/latest/dg/experiments-create.html#:~:text=CreateTrainingJob-,Processing,-Processor.run>

[Latest MLS-C01 Dumps](#)

[MLS-C01 PDF Dumps](#)

[MLS-C01 Braindumps](#)