



# 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.lead4pass.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

An agency collects census information within a country to determine healthcare and social program needs by province and city. The census form collects responses for approximately 500 questions from each citizen

Which combination of algorithms would provide the appropriate insights? (Select TWO )

- A. The factorization machines (FM) algorithm
- B. The Latent Dirichlet Allocation (LDA) algorithm
- C. The principal component analysis (PCA) algorithm
- D. The k-means algorithm
- E. The Random Cut Forest (RCF) algorithm

Correct Answer: CD

The PCA and K-means algorithms are useful in collection of data using census form.

---

### QUESTION 2

An online reseller has a large, multi-column dataset with one column missing 30% of its data A Machine Learning Specialist believes that certain columns in the dataset could be used to reconstruct the missing data

Which reconstruction approach should the Specialist use to preserve the integrity of the dataset?

- A. Listwise deletion
- B. Last observation carried forward
- C. Multiple imputation
- D. Mean substitution

Correct Answer: C

Reference: <https://worldwidescience.org/topicpages/i/imputing+missing+values.html>

---

### QUESTION 3

A Machine Learning Specialist is creating a new natural language processing application that processes a dataset comprised of 1 million sentences. The aim is to then run Word2Vec to generate embeddings of the sentences and enable different types of predictions.

Here is an example from the dataset:

"The quck BROWN FOX jumps over the lazy dog."

Which of the following are the operations the Specialist needs to perform to correctly sanitize and prepare the data in a



repeatable manner? (Choose three.)

- A. Perform part-of-speech tagging and keep the action verb and the nouns only
- B. Normalize all words by making the sentence lowercase
- C. Remove stop words using an English stopwords dictionary.
- D. Correct the typography on "quck" to "quick."
- E. One-hot encode all words in the sentence
- F. Tokenize the sentence into words.

Correct Answer: ABD

---

#### QUESTION 4

A company is using Amazon Textract to extract textual data from thousands of scanned text-heavy legal documents daily. The company uses this information to process loan applications automatically. Some of the documents fail business validation and are returned to human reviewers, who investigate the errors. This activity increases the time to process the loan applications.

What should the company do to reduce the processing time of loan applications?

- A. Configure Amazon Textract to route low-confidence predictions to Amazon SageMaker Ground Truth. Perform a manual review on those words before performing a business validation.
- B. Use an Amazon Textract synchronous operation instead of an asynchronous operation.
- C. Configure Amazon Textract to route low-confidence predictions to Amazon Augmented AI (Amazon A2I). Perform a manual review on those words before performing a business validation.
- D. Use Amazon Rekognition's feature to detect text in an image to extract the data from scanned images. Use this information to process the loan applications.

Correct Answer: C

---

#### QUESTION 5

A company sells thousands of products on a public website and wants to automatically identify products with potential durability problems. The company has 1,000 reviews with date, star rating, review text, review summary, and customer email fields, but many reviews are incomplete and have empty fields. Each review has already been labeled with the correct durability result.

A machine learning specialist must train a model to identify reviews expressing concerns over product durability. The first model needs to be trained and ready to review in 2 days.

What is the MOST direct approach to solve this problem within 2 days?

- A. Train a custom classifier by using Amazon Comprehend.
- B. Build a recurrent neural network (RNN) in Amazon SageMaker by using Gluon and Apache MXNet.



- C. Train a built-in BlazingText model using Word2Vec mode in Amazon SageMaker.
- D. Use a built-in seq2seq model in Amazon SageMaker.

Correct Answer: B

### QUESTION 6

Machine Learning Specialist is working with a media company to perform classification on popular articles from the company's website. The company is using random forests to classify how popular an article will be before it is published. A sample of the data being used is below.

Article_Title	Author	Top_Keywords	Day_Of_Week	URL_of_Article	Page_Views
Building a Big Data Platform	Jane Doe	Big Data, Spark, Hadoop	Tuesday	<a href="http://examplecorp.com/data_platform.html">http://examplecorp.com/data_platform.html</a>	1303456
Getting Started with Deep Learning	John Doe	Deep Learning, Machine Learning, Spark	Tuesday	<a href="http://examplecorp.com/started_deep_learning.html">http://examplecorp.com/started_deep_learning.html</a>	1230661
MXNet ML Guide	Jane Doe	Machine Learning, MXNet, Logistic Regression	Thursday	<a href="http://examplecorp.com/mxnet_guide.html">http://examplecorp.com/mxnet_guide.html</a>	937291
Intro to NoSQL Databases	Mary Major	NoSQL, Operations, Database	Monday	<a href="http://examplecorp.com/nosql_intro_guide.html">http://examplecorp.com/nosql_intro_guide.html</a>	407312

Given the dataset, the Specialist wants to convert the Day\_Of\_Week column to binary values.

What technique should be used to convert this column to binary values?

- A. Binarization
- B. One-hot encoding
- C. Tokenization
- D. Normalization transformation

Correct Answer: B

### QUESTION 7

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 8

A Machine Learning Specialist wants to determine the appropriate SageMakerVariantInvocationsPerInstance setting for an endpoint automatic scaling configuration. The Specialist has performed a load test on a single instance and determined that peak requests per second (RPS) without service degradation is about 20 RPS. As this is the first deployment, the Specialist intends to set the invocation safety factor to 0.5.

Based on the stated parameters and given that the invocations per instance setting is measured on a per-minute basis, what should the Specialist set as the SageMakerVariantInvocationsPerInstancesetting?

- A. 10
- B. 30
- C. 600
- D. 2,400

Correct Answer: C

---

#### QUESTION 9

A machine learning specialist works for a fruit processing company and needs to build a system that categorizes apples into three types. The specialist has collected a dataset that contains 150 images for each type of apple and applied transfer learning on a neural network that was pretrained on ImageNet with this dataset.

The company requires at least 85% accuracy to make use of the model.

After an exhaustive grid search, the optimal hyperparameters produced the following:

1.  
68% accuracy on the training set



2.

67% accuracy on the validation set

What can the machine learning specialist do to improve the system's accuracy?

- A. Upload the model to an Amazon SageMaker notebook instance and use the Amazon SageMaker HPO feature to optimize the model's hyperparameters.
- B. Add more data to the training set and retrain the model using transfer learning to reduce the bias.
- C. Use a neural network model with more layers that are pretrained on ImageNet and apply transfer learning to increase the variance.
- D. Train a new model using the current neural network architecture.

Correct Answer: B

---

#### QUESTION 10

A machine learning specialist needs to analyze comments on a news website with users across the globe. The specialist must find the most discussed topics in the comments that are in either English or Spanish.

What steps could be used to accomplish this task? (Choose two.)

- A. Use an Amazon SageMaker BlazingText algorithm to find the topics independently from language. Proceed with the analysis.
- B. Use an Amazon SageMaker seq2seq algorithm to translate from Spanish to English, if necessary. Use a SageMaker Latent Dirichlet Allocation (LDA) algorithm to find the topics.
- C. Use Amazon Translate to translate from Spanish to English, if necessary. Use Amazon Comprehend topic modeling to find the topics.
- D. Use Amazon Translate to translate from Spanish to English, if necessary. Use Amazon Lex to extract topics from the content.
- E. Use Amazon Translate to translate from Spanish to English, if necessary. Use Amazon SageMaker Neural Topic Model (NTM) to find the topics.

Correct Answer: B

Reference: <https://docs.aws.amazon.com/sagemaker/latest/dg/lda.html>

---

#### QUESTION 11

A Data Scientist is working on an application that performs sentiment analysis. The validation accuracy is poor, and the Data Scientist thinks that the cause may be a rich vocabulary and a low average frequency of words in the dataset.

Which tool should be used to improve the validation accuracy?

- A. Amazon Comprehend syntax analysts and entity detection



- B. Amazon SageMaker BlazingText allow mode
- C. Natural Language Toolkit (NLTK) stemming and stop word removal
- D. Scikit-learn term frequency-inverse document frequency (TF-IDF) vectorizers

Correct Answer: D

Reference: <https://monkeylearn.com/sentiment-analysis/>

---

## QUESTION 12

A Data Scientist wants to gain real-time insights into a data stream of GZIP files. Which solution would allow the use of SQL to query the stream with the LEAST latency?

- A. Amazon Kinesis Data Analytics with an AWS Lambda function to transform the data.
- B. AWS Glue with a custom ETL script to transform the data.
- C. An Amazon Kinesis Client Library to transform the data and save it to an Amazon ES cluster.
- D. Amazon Kinesis Data Firehose to transform the data and put it into an Amazon S3 bucket.

Correct Answer: A

Reference: <https://aws.amazon.com/big-data/real-time-analytics-featured-partners/>

[Latest MLS-C01 Dumps](#)

[MLS-C01 Study Guide](#)

[MLS-C01 Braindumps](#)





To Read the [Whole Q&As](#), please purchase the [Complete Version](#) from [Our website](#).

## Try our product !

100% Guaranteed Success

100% Money Back Guarantee

365 Days Free Update

Instant Download After Purchase

24x7 Customer Support

Average 99.9% Success Rate

More than 800,000 Satisfied Customers Worldwide

Multi-Platform capabilities - [Windows](#), [Mac](#), [Android](#), [iPhone](#), [iPod](#), [iPad](#), [Kindle](#)

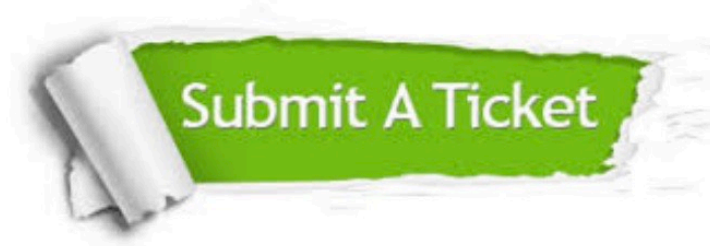
We provide exam PDF and VCE of Cisco, Microsoft, IBM, CompTIA, Oracle and other IT Certifications. You can view Vendor list of All Certification Exams offered:

<https://www.lead4pass.com/allproducts>

## Need Help

Please provide as much detail as possible so we can best assist you.

To update a previously submitted ticket:



 <p><b>One Year Free Update</b> Free update is available within One Year after your purchase. After One Year, you will get 50% discounts for updating. And we are proud to boast a 24/7 efficient Customer Support system via Email.</p>	 <p><b>Money Back Guarantee</b> To ensure that you are spending on quality products, we provide 100% money back guarantee for 30 days from the date of purchase.</p>	 <p><b>Security &amp; Privacy</b> We respect customer privacy. We use McAfee's security service to provide you with utmost security for your personal information &amp; peace of mind.</p>
---	---	--

Any charges made through this site will appear as Global Simulators Limited.

All trademarks are the property of their respective owners.

Copyright © lead4pass, All Rights Reserved.