AWS MLS-C01 Study Guide PDF

GRAB THE AWS MACHINE LEARNING SPECIALTY CERTIFICATION PDF QUESTIONS & ANSWERS

Exam Summary – Syllabus –Questions

MLS-C01

AWS Certified Machine Learning - Specialty
65 Questions Exam – Duration of 180 minutes

www.VMExam.com

Table of Contents

Get an Overview of the MLS-C01 Certification:

Who should take the **MLS-C01 exam**? This is the first question that comes to a candidate's mind when preparing for the Machine Learning Specialty certification. The MLS-C01 certification is suitable for candidates who are keen to earn knowledge on the Specialty and grab their AWS Certified Machine Learning - Specialty. When it is about starting the preparation, most candidates get confused regarding the study materials and study approach. But MLS-C01 study guide PDF is here to solve the problem. MLS-C01 PDF combines some effective sample questions and offers valuable tips to pass the exam with ease.

Why Should You Earn the AWS MLS-C01 Certification?

There are several reasons why one should grab the MLS-C01 certification.

- The Machine Learning Specialty certification proves to be one of the most recognized certifications.
- The certification badge proves the knowledge of the candidate regarding subject matters and makes his resume presentable to potential candidates.
- Thus earning the **AWS Certified Machine Learning - Specialty** is a powerful qualification for a prosperous career.

What Is the AWS MLS-C01 Machine Learning Specialty Certification Exam Structure?

Exam Name	AWS Certified Machine Learning - Specialty (Machine Learning Specialty)
Exam Code	MLS-C01
Exam Price	$300 USD
Duration	180 minutes
Number of Questions	65
Passing Score	750 / 1000
Recommended Training / Books	**Practical Data Science with Amazon SageMaker** **The Machine Learning Pipeline on AWS** **Deep Learning on AWS**
Schedule Exam	**PEARSON VUE**
Sample Questions	**AWS MLS-C01 Sample Questions**
Recommended Practice	**AWS Certified Machine Learning - Specialty Practice Test**

Enhance Knowledge with MLS-C01 Sample Questions:

Question: 1

A Data Scientist is evaluating different binary classification models. A false positive result is 5 times more expensive (from a business perspective) than a false negative result.

The models should be evaluated based on the following criteria:

1) Must have a recall rate of at least 80%
2) Must have a false positive rate of 10% or less
3) Must minimize business costs

After creating each binary classification model, the Data Scientist generates the corresponding confusion matrix.

Which confusion matrix represents the model that satisfies the requirements?

a) TN = 91, FP = 9 FN = 22, TP = 78
b) TN = 99, FP = 1 FN = 21, TP = 79
c) TN = 96, FP = 4 FN = 10, TP = 90
d) TN = 98, FP = 2 FN = 18, TP = 82

Answer: d

Question: 2

An insurance company needs to automate claim compliance reviews because human reviews are expensive and error-prone. The company has a large set of claims and a compliance label for each.

Each claim consists of a few sentences in English, many of which contain complex related information. Management would like to use Amazon SageMaker built-in algorithms to design a machine learning supervised model that can be trained to read each claim and predict if the claim is compliant or not.

Which approach should be used to extract features from the claims to be used as inputs for the downstream supervised task?

a) Derive a dictionary of tokens from claims in the entire dataset. Apply one-hot encoding to tokens found in each claim of the training set. Send the derived features space as inputs to an Amazon SageMaker builtin supervised learning algorithm.
b) Apply Amazon SageMaker BlazingText in Word2Vec mode to claims in the training set. Send the derived features space as inputs for the downstream supervised task.
c) Apply Amazon SageMaker BlazingText in classification mode to labeled claims in the training set to derive features for the claims that correspond to the compliant and non-compliant labels, respectively.
d) Apply Amazon SageMaker Object2Vec to claims in the training set. Send the derived features space as inputs for the downstream supervised task.

Answer: d

Question: 3

A Data Scientist uses logistic regression to build a fraud detection model. While the model accuracy is 99%, 90% of the fraud cases are not detected by the model.

What action will definitively help the model detect more than 10% of fraud cases?

- a) Using oversampling to balance the dataset
- b) Using regularization to reduce overfitting
- c) Decreasing the class probability threshold
- d) Using undersampling to balance the dataset

Answer: c

Question: 4

A Machine Learning Engineer is preparing a data frame for a supervised learning task with the Amazon SageMaker Linear Learner algorithm.

The ML Engineer notices the target label classes are highly imbalanced and multiple feature columns contain missing values. The proportion of missing values across the entire data frame is less than 5%.

What should the ML Engineer do to minimize bias due to missing values?

- a) Replace each missing value by the mean or median across non-missing values in same row.
- b) Delete observations that contain missing values because these represent less than 5% of the data.
- c) Replace each missing value by the mean or median across non-missing values in the same column.
- d) For each feature, approximate the missing values using supervised learning based on other features.

Answer: d

Question: 5

A company is interested in building a fraud detection model. Currently, the Data Scientist does not have a sufficient amount of information due to the low number of fraud cases.

Which method is MOST likely to detect the GREATEST number of valid fraud cases?

- a) Oversampling using bootstrapping
- b) Undersampling
- c) Oversampling using SMOTE
- d) Class weight adjustment

Answer: c

Question: 6

A Machine Learning team has several large CSV datasets in Amazon S3. Historically, models built with the Amazon SageMaker Linear Learner algorithm have taken hours to train on similar-sized datasets. The team's leaders need to accelerate the training process.

What can a Machine Learning Specialist do to address this concern?

a) Use Amazon SageMaker Pipe mode.
b) Use Amazon Machine Learning to train the models.
c) Use Amazon Kinesis to stream the data to Amazon SageMaker.
d) Use AWS Glue to transform the CSV dataset to the JSON format.

Answer: a

Question: 7

A company has collected customer comments on its products, rating them as safe or unsafe, using decision trees. The training dataset has the following features:

id, date, full review, full review summary, and a binary safe/unsafe tag. During training, any data sample with missing features was dropped. In a few instances, the test set was found to be missing the full review text field.

For this use case, which is the most effective course of action to address test data samples with missing features?

a) Drop the test samples with missing full review text fields, and then run through the test set.
b) Copy the summary text fields and use them to fill in the missing full review text fields, and then run through the test set.
c) Use an algorithm that handles missing data better than decision trees.
d) Generate synthetic data to fill in the fields that are missing data, and then run through the test set.

Answer: b

Question: 8

A term frequency–inverse document frequency (tf–idf) matrix using both unigrams and bigrams is built from a text corpus consisting of the following two sentences:

1. Please call the number below.

2. Please do not call us.

What are the dimensions of the tf–idf matrix?

a) (2, 16)
b) (2, 8)
c) (2, 10)
d) (8, 10)

Answer: a

Question: 9

A company is setting up a system to manage all of the datasets it stores in Amazon S3.

The company would like to automate running transformation jobs on the data and maintaining a catalog of the metadata concerning the datasets. The solution should require the least amount of setup and maintenance.

Which solution will allow the company to achieve its goals?

  a) Create an Amazon EMR cluster with Apache Hive installed. Then, create a Hive metastore and a script to run transformation jobs on a schedule.
  b) Create an AWS Glue crawler to populate the AWS Glue Data Catalog. Then, author an AWS Glue ETL job, and set up a schedule for data transformation jobs.
  c) Create an Amazon EMR cluster with Apache Spark installed. Then, create an Apache Hive metastore and a script to run transformation jobs on a schedule.
  d) Create an AWS Data Pipeline that transforms the data. Then, create an Apache Hive metastore and a script to run transformation jobs on a schedule.

Answer: b

Question: 10

A Data Scientist is working on optimizing a model during the training process by varying multiple parameters. The Data Scientist observes that, during multiple runs with identical parameters, the loss function converges to different, yet stable, values.

What should the Data Scientist do to improve the training process?

  a) Increase the learning rate. Keep the batch size the same.
  b) Reduce the batch size. Decrease the learning rate.
  c) Keep the batch size the same. Decrease the learning rate.
  d) Do not change the learning rate. Increase the batch size.

Answer: b

What Study Guide Works Best in Acing the AWS MLS-C01 Machine Learning Specialty Certification?

The MLS-C01 study guide is a combination of some proven study tips and the combination of all valuable study materials like sample questions, syllabus and practice tests in one place.

Explore the Syllabus Topics and Learn from the Core:

If you are determined to earn success in the Machine Learning Specialty exam, getting in full touch of the **syllabus** is mandatory. During preparation, you might not like all syllabus sections or topics, but try to get at least the fundamental knowledge from the sections you don't like. The more you possess knowledge on all syllabus sections, the more is the chance to attempt maximum number of questions during the actual exam.

Make Your Schedule:

Studying and completing the syllabus becomes easier, if you work on the syllabus topics after making a schedule. Your syllabus must mention what areas you want to cover and within what time. Once you make a schedule and follow it regularly, syllabus completion becomes easier and preparation becomes smoother.

Get Expert Advice from the Training:

Do not forget to join the AWS MLS-C01 training if it is providing any. Training enhances the practical knowledge of a candidate, which helps them to work well in the practical field during projects.

Get Access to the PDF Sample Questions:

If your study material is in a **PDF format** or the materials are mobile-friendly, what could be better than that? Get access to the free sample questions and keep enhancing your knowledge beyond the syllabus.

Avoid Dumps and Utilize the AWS MLS-C01 Practice Test:

Why should you rely on practice tests? The reason is simple: you must get familiar with the exam pattern before reaching the exam hall. An aspirant aware of the exam structure and time management during the exam preparation can perform well in the actual exam and attempt the maximum number of questions during the exam.

Many aspirants prefer to read from dumps, but they miss out on the self assessment method. Therefore, MLS-C01 practice tests always stand out to be the better choice than dumps PDF.

Avail the Proven MLS-C01 Practice Test for Success!!!

Do you want to pass the MLS-C01 exam on your first attempt? Stop worrying; we, VMExam.com are here to provide you the best experience during your AWS Certified Machine Learning - Specialty preparation. Try out our free mock tests to get a glimpse of our quality study materials, and build your confidence with the premium **MLS-C01 practice tests**. Our expert-designed questions help you to improve performance and pass the exam on your first attempt.