Pass MLS-C01 Machine Learning Specialty Exam: Study Tips & Resources!

AWS MACHINE LEARNING SPECIALTY CERTIFICATION QUESTIONS & ANSWERS

Get Instant Access to Vital Exam Acing Materials | Study Guide | Sample Questions | Practice Test

MLS-C01

AWS Certified Machine Learning - Specialty

65 Questions Exam – 750 / 1000 Cut Score – Duration of 180 minutes

Table of Contents

Get Ready for the MLS-C01 Exam:

Prepare effectively for the MLS-C01 exam using reliable **study strategies and methods**. Enhance your preparedness, deepen your understanding of the Specialty, and enhance your likelihood of achieving success in the AWS Certified Machine Learning - Specialty with our comprehensive guide. Embark on your path to exam excellence today.

Know More About the AWS Certified Machine Learning - Specialty Certification:

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**
Schedule Exam	**AWS Certification**
Sample Questions	**AWS MLS-C01 Sample Questions**
Recommended Practice	**AWS Certified Machine Learning - Specialty Practice Test**

Learn More About the MLS-C01 Syllabus:

Section	Objectives
Data Engineering - 20%	
Create data repositories for ML.	- Identify data sources (for example, content and location, primary sources such as user data). - Determine storage mediums (for example, databases, Amazon S3, Amazon Elastic File System [Amazon EFS], Amazon Elastic Block Store [Amazon EBS]).

Section	Objectives
Identify and implement a data ingestion solution.	- Identify data job styles and job types (for example, batch load, streaming). - Orchestrate data ingestion pipelines (batch-based ML workloads and streaming-based ML workloads). • Amazon Kinesis • Amazon Data Firehose • Amazon EMR • AWS Glue • Amazon Managed Service for Apache Flink - Schedule jobs.
Identify and implement a data transformation solution.	- Transform data in transit (ETL, AWS Glue, Amazon EMR, AWS Batch). - Handle ML-specific data by using MapReduce (for example, Apache Hadoop, Apache Spark, Apache Hive).
Exploratory Data Analysis - 24%	
Sanitize and prepare data for modeling.	- Identify and handle missing data, corrupt data, and stop words. - Format, normalize, augment, and scale data. - Determine whether there is sufficient labeled data. • Identify mitigation strategies. • Use data labelling tools (for example, Amazon Mechanical Turk).
Perform feature engineering.	- Identify and extract features from datasets, including from data sources such as text, speech, image, public datasets. - Analyze and evaluate feature engineering concepts (for example, binning, tokenization, outliers, synthetic features, one-hot encoding, reducing dimensionality of data).
Analyze and visualize data for ML.	- Create graphs (for example, scatter plots, time series, histograms, box plots). - Interpret descriptive statistics (for example, correlation, summary statistics, p-value). - Perform cluster analysis (for example, hierarchical, diagnosis, elbow plot, cluster size).

Section	Objectives
Modeling - 36%	
Frame business problems as ML problems.	- Determine when to use and when not to use ML. - Know the difference between supervised and unsupervised learning. - Select from among classification, regression, forecasting, clustering, and recommendation and foundation models.
Select the appropriate model(s) for a given ML problem.	- XGBoost, logistic regression, k-means, linear regression, decision trees, random forests, RNN, CNN, ensemble, transfer learning and large language models (LLMs) - Express the intuition behind models.
Train ML models.	- Split data between training and validation (for example, cross validation). - Understand optimization techniques for ML training (for example, gradient descent, loss functions, convergence). - Choose appropriate compute resources (for example GPU or CPU, distributed or non-distributed). • Choose appropriate compute platforms (Spark or non-Spark). - Update and retrain models. • Batch or real-time/online
Perform hyperparameter optimization.	- Perform regularization. • Dropout • L1/L2 - Perform cross-validation. - Initialize models. - Understand neural network architecture (layers and nodes), learning rate, and activation functions. - Understand tree-based models (number of trees, number of levels). - Understand linear models (learning rate).
Evaluate ML models.	- Avoid overfitting or underfitting. • Detect and handle bias and variance. - Evaluate metrics (for example, area under curve [AUC]-receiver operating characteristics [ROC], accuracy, precision, recall, Root Mean Square Error [RMSE], F1

Section	Objectives
	score). - Interpret confusion matrices. - Perform offline and online model evaluation (A/B testing). - Compare models by using metrics (for example, time to train a model, quality of model, engineering costs). - Perform cross-validation.
Machine Learning Implementation and Operations - 20%	
Build ML solutions for performance, availability, scalability, resiliency, and fault tolerance.	- Log and monitor AWS environments. • AWS CloudTrail and Amazon CloudWatch • Build error monitoring solutions. - Deploy to multiple AWS Regions and multiple Availability Zones. - Create AMIs and golden images. - Create Docker containers. - Deploy Auto Scaling groups. - Rightsize resources (for example, instances, Provisioned IOPS, volumes). - Perform load balancing. - Follow AWS best practices.
Recommend and implement the appropriate ML services and features for a given problem.	- ML on AWS (application services), for example: • Amazon Polly • Amazon Lex • Amazon Transcribe - Understand AWS service quotas. - Determine when to build custom models and when to use Amazon SageMaker built-in algorithms. - Understand AWS infrastructure (for example, instance types) and cost considerations. • Use Spot Instances to train deep learning models by using AWS Batch.
Apply basic AWS security practices to ML solutions.	- AWS Identity and Access Management (IAM) - S3 bucket policies - Security groups - VPCs - Encryption and anonymization

Section	Objectives
Deploy and operationalize ML solutions.	- Expose endpoints and interact with them. - Understand ML models. - Perform A/B testing. - Retrain pipelines. - Debug and troubleshoot ML models. • Detect and mitigate drops in performance. • Monitor performance of the model.

Prepare with MLS-C01 Sample Questions:

Question: 1

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: 2

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: 3

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: 4

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: 5

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

Question: 6

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: 7

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: 8

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: 9

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: 10

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

Tips for Success in the AWS Certified Machine Learning - Specialty Exam:

Familiarize Yourself with the MLS-C01 Exam Format:

Before starting your study regimen, it's crucial to acquaint yourself with the structure of the MLS-C01 exam. Take a moment to **review the exam syllabus**, grasp the test format, and pinpoint the main areas of concentration. Having prior knowledge of the exam's layout will assist you in customizing your study strategy effectively.

Create A Study Timetable for the MLS-C01 Exam:

To prepare efficiently for the MLS-C01 exam, devise a study schedule that aligns with your lifestyle and preferred learning approach. Allocate dedicated time slots for studying each day, prioritizing topics according to their significance and your level of proficiency. Maintaining consistency by adhering to your schedule and steering clear of procrastination is imperative.

Diversify Your Study Sources:

Ensure you broaden your study material beyond just one source. Use various resources like textbooks, online courses, practice exams, and study guides to understand the MLS-C01 exam subjects thoroughly. Each resource provides distinct perspectives and explanations that can enrich your learning journey.

Regular Practice for the MLS-C01 Exam:

Consistent practice is essential for effective preparation for the MLS-C01 exam. Engaging in regular practice enables you to strengthen your grasp of essential concepts, improve your problem-solving abilities, and become accustomed to the exam format. Allocate dedicated time to solving practice questions and sample tests to assess your progress accurately.

Allow for Rest and Breaks:

While studying is crucial, taking breaks and rest is equally vital. Pushing yourself too hard without sufficient rest can result in burnout and reduced effectiveness. Incorporate short breaks into your study sessions to recharge and stay focused.

Maintain Organization Throughout Your MLS-C01 Exam Preparation:

Keep yourself organized as you prepare for the MLS-C01 exam by monitoring your progress and managing your materials effectively. Ensure your study area remains neat, utilize folders or digital aids to arrange your notes and resources, and develop a checklist of topics to review. Employing an organized approach will assist you in staying focused and reducing stress levels.

Seek Guidance from Mentors:

Feel free to ask for clarification when you come across confusing or difficult concepts during your study sessions. Seek support from peers, instructors, or online forums to address any uncertainties. Addressing doubts will prevent misunderstandings and ensure you develop a strong **understanding of the material**.

Regular Review is Crucial for the MLS-C01 Exam:

Frequent revisiting of material is paramount for retaining information over the long term. Revisit topics you've already covered to strengthen your comprehension and pinpoint areas that need further focus. Regular review sessions will **solidify your understanding** and enhance your confidence.

Master Time Management for the MLS-C01 Exam:

Skillful time management is essential on the exam day to ensure you finish all sections within the designated time limits. During your practice sessions, replicate the conditions of the MLS-C01 exam and practice managing your time accordingly. Formulate strategies for efficiently addressing each section to optimize your score.

Have A Positive Mindset:

Finally, maintain a positive attitude and have faith in your capabilities. Stay confident in your preparation and trust that you are well-prepared to handle the MLS-C01 exam. Envision success, remain focused, and approach the exam calmly and objectively.

Benefits of Passing the MLS-C01 Exam:

- Completing the MLS-C01 exam unlocks pathways to fresh career prospects and progression within your industry.
- The extensive preparation needed for the MLS-C01 certification equips you with comprehensive knowledge and practical expertise applicable to your field.

- Possessing the MLS-C01 certification showcases your mastery and dedication to excellence, garnering acknowledgment from both peers and employers.
- Certified professionals often command higher salaries and have greater potential for earning than those without certification.
- Acquiring the MLS-C01 certification validates your competence and trustworthiness, fostering confidence among clients, employers, and peers.

Explore the Trusted Practice Exam for the MLS-C01 Certification:

At vmexam.com, you'll find comprehensive resources for the MLS-C01 exam. Our platform offers authentic practice exams tailored specifically for the MLS-C01 certification. What advantages do these practice exams provide? You'll encounter genuine exam-style questions expertly crafted by industry professionals, allowing you to improve your performance in the exam. Rely on vmexam.com for rigorous, unlimited access to **MLS-C01 practice exams** for two months, allowing you to boost your confidence steadily. Through focused practice, numerous candidates have successfully streamlined their path to achieving the AWS Certified Machine Learning - Specialty.

Final Remarks:

Preparing for the MLS-C01 examination demands commitment, strategic planning, and efficient study methods. Implementing these study suggestions can enrich your preparation, elevate your self-assurance, and increase your likelihood of excelling in the exam. Keep your focus sharp, maintain organization, and believe in your abilities. Best of luck!

Here Is the Trusted Practice Test for the MLS-C01 Certification

VMExam.Com is here with all the necessary details regarding the MLS-C01 exam. We provide authentic practice tests for the MLS-C01 exam. What do you gain from these practice tests? You get to experience the real exam-like questions made by industry experts and get a scope to improve your performance in the actual exam. Rely on VMExam.Com for rigorous, unlimited two-month attempts on the **MLS-C01 practice tests**, and gradually build your confidence. Rigorous practice made many aspirants successful and made their journey easy towards grabbing the AWS Certified Machine Learning - Specialty.

Start Online Practice of MLS-C01 Exam by Visiting URL

https://www.vmexam.com/aws/mls-c01-aws-certified-machine-learning-specialty