AWS Certified Machine Learning Specialty (MLS-C01) Practice Test (Sample)

Study Guide

Everything you need from our exam experts!

This is a sample study guide. To access the full version with hundreds of questions,

Copyright © 2026 by Examzify - A Kaluba Technologies Inc. product.

ALL RIGHTS RESERVED.

No part of this book may be reproduced or transferred in any form or by any means, graphic, electronic, or mechanical, including photocopying, recording, web distribution, taping, or by any information storage retrieval system, without the written permission of the author.

Notice: Examzify makes every reasonable effort to obtain from reliable sources accurate, complete, and timely information about this product.

SAMPLE

Table of Contents

SAMPLE

Introduction

Preparing for a certification exam can feel overwhelming, but with the right tools, it becomes an opportunity to build confidence, sharpen your skills, and move one step closer to your goals. At Examzify, we believe that effective exam preparation isn't just about memorization, it's about understanding the material, identifying knowledge gaps, and building the test-taking strategies that lead to success.

This guide was designed to help you do exactly that.

Whether you're preparing for a licensing exam, professional certification, or entry-level qualification, this book offers structured practice to reinforce key concepts. You'll find a wide range of multiple-choice questions, each followed by clear explanations to help you understand not just the right answer, but why it's correct.

The content in this guide is based on real-world exam objectives and aligned with the types of questions and topics commonly found on official tests. It's ideal for learners who want to:

- **Practice answering questions under realistic conditions,**
- **Improve accuracy and speed,**
- **Review explanations to strengthen weak areas, and**
- **Approach the exam with greater confidence.**

We recommend using this book not as a stand-alone study tool, but alongside other resources like flashcards, textbooks, or hands-on training. For best results, we recommend working through each question, reflecting on the explanation provided, and revisiting the topics that challenge you most.

Remember: successful test preparation isn't about getting every question right the first time, it's about learning from your mistakes and improving over time. Stay focused, trust the process, and know that every page you turn brings you closer to success.

Let's begin.

SAMPLE

How to Use This Guide

This guide is designed to help you study more effectively and approach your exam with confidence. Whether you're reviewing for the first time or doing a final refresh, here’s how to get the most out of your Examzify study guide:

1. Start with a Diagnostic Review

Skim through the questions to get a sense of what you know and what you need to focus on. Don’t worry about getting everything right, your goal is to identify knowledge gaps early.

2. Study in Short, Focused Sessions

Break your study time into manageable blocks (e.g. 30 – 45 minutes). Review a handful of questions, reflect on the explanations, and take breaks to retain information better.

3. Learn from the Explanations

After answering a question, always read the explanation, even if you got it right. It reinforces key points, corrects misunderstandings, and teaches subtle distinctions between similar answers.

4. Track Your Progress

Use bookmarks or notes (if reading digitally) to mark difficult questions. Revisit these regularly and track improvements over time.

5. Simulate the Real Exam

Once you're comfortable, try taking a full set of questions without pausing. Set a timer and simulate test-day conditions to build confidence and time management skills.

6. Repeat and Review

Don’t just study once, repetition builds retention. Re-attempt questions after a few days and revisit explanations to reinforce learning.

7. Use Other Tools

Pair this guide with other Examzify tools like flashcards, and digital practice tests to strengthen your preparation across formats.

SAMPLE

There’s no single right way to study, but consistent, thoughtful effort always wins. Use this guide flexibly — adapt the tips above to fit your pace and learning style. You've got this!

SAMPLE

Questions

SAMPLE

1. What function does the SageMaker Ground Truth service provide?

A. It offers automated data preprocessing.

B. It manages labeled datasets using human labeling workflows.

C. It optimizes model training without human intervention.

D. It generates synthetic datasets for training.

2. What is a defining feature of AWS Identity and Access Management (IAM)?

A. It provides block storage for virtual machines

B. It offers object storage without server management

C. It centrally manages user permissions and credentials

D. It helps to analyze data through machine learning

3. What is the significance of the Precision-Recall curve in machine learning?

A. It helps in evaluating the trade-off between precision and recall.

B. It is used to compare different models based on accuracy.

C. It visualizes the loss function over epochs.

D. It determines the best fit line in regression analysis.

4. Which service is best suited for analyzing streaming data in real-time using Apache Flink?

A. Amazon Kinesis Data Streams

B. Amazon Kinesis Data Analytics

C. Amazon Kinesis Firehose

D. AWS Glue

5. In which scenario would transfer learning be most beneficial?

A. When you have an abundance of labeled data.

B. When adapting a pre-trained model to a related task.

C. When there is no available data for the new task.

D. When building a model from scratch with no prior knowledge.

SAMPLE

6. What type of business intelligence service is Amazon QuickSight?

A. Data storage service with analytics capabilities

B. Machine learning-powered BI service

C. ETL service for data management

D. Monitoring service for cloud resources

7. Which value is utilized in the Stochastic Gradient Descent algorithm to control the learning speed?

A. Momentum

B. Learning rate

C. Batch size

D. Epochs

8. Which service is primarily focused on providing a fully managed data lake storage solution?

A. Amazon Kinesis

B. Amazon EMR

C. AWS Glue

D. AWS Lambda

9. What is the developer kit that helps learn machine learning concepts with computer vision?

A. AWS RoboMaker

B. AWS DeepLens

C. Amazon Rekognition

D. Amazon SageMaker

10. What key advantage does AWS Lambda offer for machine learning applications?

A. It allows for continuous data storage

B. It enables real-time updates to data sets

C. It facilitates serverless architecture and reduces operational overhead

D. It standardizes model accuracy across different applications

SAMPLE

Answers

SAMPLE

1. B
2. C
3. A
4. B
5. B
6. B
7. B
8. C
9. B
10. C

SAMPLE

Explanations

SAMPLE

1. What function does the SageMaker Ground Truth service provide?

A. It offers automated data preprocessing.

B. It manages labeled datasets using human labeling workflows.

C. It optimizes model training without human intervention.

D. It generates synthetic datasets for training.

SageMaker Ground Truth is specifically designed to manage the process of creating labeled datasets, which are crucial for training machine learning models. This service facilitates human labeling workflows, allowing users to efficiently create high-quality training datasets by leveraging both automated labeling and human annotators. Ground Truth supports various labeling tasks such as image, audio, and text data, and it includes features for managing and monitoring the labeling process. This capability is particularly important in machine learning, where the quality of the dataset directly impacts model performance. By streamlining and organizing the labeling process with workflows, Ground Truth helps ensure that datasets are accurately labeled, thus improving the training phase of machine learning models. The other options, while related to aspects of machine learning and data processing, do not capture the primary function of SageMaker Ground Truth. Automated data preprocessing, model training optimization, and synthetic dataset generation are valuable services in the machine learning pipeline but are not the main focus of SageMaker Ground Truth.

2. What is a defining feature of AWS Identity and Access Management (IAM)?

A. It provides block storage for virtual machines

B. It offers object storage without server management

C. It centrally manages user permissions and credentials

D. It helps to analyze data through machine learning

A defining feature of AWS Identity and Access Management (IAM) is its ability to centrally manage user permissions and credentials. IAM allows administrators to create and manage AWS users and groups, assigning specific permissions to control what actions different users can perform on AWS resources. This centralized management ensures that access to sensitive data and services is properly controlled and monitored, contributing to enhanced security. By providing a systematic approach to user management, IAM facilitates the principle of least privilege, where users only have the access necessary to perform their job functions. This helps reduce the risk of unauthorized access or accidental changes to critical AWS resources. IAM also supports various authentication mechanisms, such as multi-factor authentication (MFA), and integrates with other AWS services, making it a foundational component of AWS security management.

SAMPLE

3. What is the significance of the Precision-Recall curve in machine learning?

A. It helps in evaluating the trade-off between precision and recall.

B. It is used to compare different models based on accuracy.

C. It visualizes the loss function over epochs.

D. It determines the best fit line in regression analysis.

The Precision-Recall curve plays a crucial role in evaluating the performance of machine learning models, particularly for problems where the class distribution is imbalanced. This curve provides a graphical representation of the trade-off between precision and recall for different threshold settings. Precision measures the accuracy of positive predictions, that is, the proportion of true positive results among all the instances classified as positive. Recall, on the other hand, assesses how well the model identifies actual positive instances by calculating the proportion of true positives out of the total actual positive instances. As the threshold for classifying an instance as positive is varied, both precision and recall can change, sometimes inversely. The Precision-Recall curve illustrates this relationship, allowing a practitioner to identify the optimal balance of precision and recall depending on the specific needs of a task, such as whether false positives or false negatives carry a heavier cost. This visual representation is particularly important in scenarios like medical diagnosis or fraud detection, where the consequences of false positives and false negatives differ significantly. Therefore, understanding the Precision-Recall trade-off enables practitioners to make informed decisions tailored to the demands of their specific use cases.

4. Which service is best suited for analyzing streaming data in real-time using Apache Flink?

A. Amazon Kinesis Data Streams

B. Amazon Kinesis Data Analytics

C. Amazon Kinesis Firehose

D. AWS Glue

Amazon Kinesis Data Analytics is the most suitable service for analyzing streaming data in real-time using Apache Flink. This service is specifically designed to process and analyze large streams of data in real-time, leveraging Apache Flink as its underlying engine. It allows users to write SQL queries on live data streams, facilitating immediate insights and enabling prompt decision-making. With Kinesis Data Analytics, you can easily integrate data from the Kinesis Data Streams or other sources, apply complex processing and transformations with Flink, and then either store the results in databases or stream them to other services for further analysis or action. This makes it highly effective for scenarios where real-time data processing is required. Other services, while related to streaming data, serve different primary functions. For example, while Amazon Kinesis Data Streams is ideal for collecting and processing large streams of data, it does not provide the built-in capability for sophisticated data analysis that Kinesis Data Analytics does. Amazon Kinesis Firehose is focused on delivering data streams to destinations like Amazon S3 or Redshift, rather than the analysis of the data itself. AWS Glue is primarily a data integration service that prepares data for analytics but does not specialize in real-time analysis like Kinesis Data Analytics. Thus, Kinesis Data Analytics

SAMPLE

5. In which scenario would transfer learning be most beneficial?

A. When you have an abundance of labeled data.

B. When adapting a pre-trained model to a related task.

C. When there is no available data for the new task.

D. When building a model from scratch with no prior knowledge.

Transfer learning is particularly advantageous when adapting a pre-trained model to a related task because it leverages the knowledge that the model has already gained from being trained on a large dataset. This allows for faster training and often leads to improved performance, especially when the new task does not have a large amount of labeled data available. By starting with an existing model, you can utilize the learned representations and features that are likely to be relevant to the new task, even if the specific data or labels differ. In cases of having an abundance of labeled data, building a model from scratch may often yield better performance since you can tailor the model specifically to the nuances of the new dataset without the biases that a pre-trained model might carry. Similarly, when there is no available data for the new task, transfer learning becomes ineffective because there is no knowledge or features to adapt from the pre-trained model. Lastly, building a model from scratch when there is no prior knowledge provides the opportunity to create a unique architecture for a specific problem, but it does not benefit from the efficiencies and capabilities gained from transfer learning.

6. What type of business intelligence service is Amazon QuickSight?

A. Data storage service with analytics capabilities

B. Machine learning-powered BI service

C. ETL service for data management

D. Monitoring service for cloud resources

Amazon QuickSight is a machine learning-powered business intelligence (BI) service that enables users to visualize and analyze their data. It leverages the capabilities of machine learning to provide deeper insights into data trends and patterns, making it easier for users to make data-driven decisions. QuickSight includes features such as anomaly detection and forecasting, which are driven by machine learning algorithms that automatically analyze the underlying data to draw conclusions and highlight key metrics. This capacity to automatically generate insights through machine learning sets QuickSight apart from traditional BI tools, which typically rely on static reporting or manual analysis. Users can quickly create interactive dashboards that are powered by machine learning to understand their data better and to retrieve actionable insights without extensive setup. The other options highlight different types of services that do not encompass the full functionality of QuickSight: data storage services focus on storing data without inherent analytics capabilities; ETL (Extract, Transform, Load) services are designed for data integration and management rather than visualization; and monitoring services pertain to tracking cloud resources rather than offering BI functionalities. Therefore, the classification of Amazon QuickSight as a machine learning-powered BI service accurately reflects its core functionality.

SAMPLE

7. Which value is utilized in the Stochastic Gradient Descent algorithm to control the learning speed?

A. Momentum

B. Learning rate

C. Batch size

D. Epochs

The learning rate is a crucial hyperparameter in the Stochastic Gradient Descent (SGD) algorithm because it determines how much to change the model's weights in response to the estimated error each time the model weights are updated. Specifically, the learning rate controls the size of the steps taken towards the minimum of the loss function. A higher learning rate can lead to faster convergence but may also risk overshooting the optimal solution. In contrast, a lower learning rate allows for finer updates, which can lead to better convergence but may slow down the training process. Setting an appropriate learning rate is vital for effective training; if set too high, the optimization may diverge or oscillate, while a rate that is too low can lead to unnecessarily long training times and a risk of getting stuck in local minima. The learning rate, therefore, directly affects both the speed and stability of the convergence during the training process.

8. Which service is primarily focused on providing a fully managed data lake storage solution?

A. Amazon Kinesis

B. Amazon EMR

C. AWS Glue

D. AWS Lambda

The correct choice is AWS Glue because it is specifically designed to facilitate the creation, management, and transformation of data lakes in AWS. AWS Glue provides a fully managed environment that makes it easier for users to prepare and load data from various sources into a data lake, enabling query, analytics, and machine learning workflows. It integrates with other AWS services, such as Amazon S3, to store data, while also offering features like data cataloging, job scheduling, and automated data transformation. AWS Glue is particularly focused on the ETL (Extract, Transform, Load) process, making it ideal for users looking to organize and manage large volumes of data in a data lake environment. This service automates much of the labor-intensive processes involved in preparing data for analytics, thus streamlining the setup of a data lake storage solution. In contrast, the other services listed have different primary functionalities. Amazon Kinesis is primarily used for real-time data processing and analytics. Amazon EMR is used for big data processing frameworks, such as Apache Hadoop and Apache Spark, but is more oriented toward processing data than managing a data lake. AWS Lambda is a serverless compute service that helps run code in response to events but does not specifically address the needs of data lake management.

SAMPLE

9. What is the developer kit that helps learn machine learning concepts with computer vision?

A. AWS RoboMaker

B. AWS DeepLens

C. Amazon Rekognition

D. Amazon SageMaker

The developer kit that helps learn machine learning concepts with a focus on computer vision is AWS DeepLens. This device is equipped with a built-in camera and runs deep learning models locally, allowing users to experiment with real-time image and video analysis directly on the device. DeepLens supports frameworks such as TensorFlow and MXNet, enabling users to deploy pre-trained models or create their own to perform various computer vision tasks, such as object recognition, image classification, and facial recognition. Engaging with AWS DeepLens provides a practical, hands-on approach to understanding machine learning concepts by integrating hardware and software, making it an ideal choice for learning. The ability to process video feeds in real time at the edge also enhances the learning experience, allowing developers to see tangible results immediately. Other options, while relevant to machine learning, do not specifically address the learning of computer vision concepts in the same way. AWS RoboMaker is focused on robotics and simulation; Amazon Rekognition is a service for image and video analysis but does not provide a hands-on learning kit, and Amazon SageMaker is a comprehensive service for building, training, and deploying machine learning models but is not specifically tailored to computer vision learning in a practical sense.

10. What key advantage does AWS Lambda offer for machine learning applications?

A. It allows for continuous data storage

B. It enables real-time updates to data sets

C. It facilitates serverless architecture and reduces operational overhead

D. It standardizes model accuracy across different applications

AWS Lambda provides significant benefits for machine learning applications primarily through its serverless architecture, which reduces operational overhead. With AWS Lambda, developers can run code without provisioning or managing servers, allowing for a more agile development process. This means you can focus on writing and deploying your machine learning models and applications without worrying about the complexities of server management, scaling, or infrastructure maintenance. By leveraging a serverless framework, you can easily scale your machine learning applications in response to changing workloads, such as processing data in real-time or handling variable user requests. This flexibility allows developers to better allocate resources and optimize costs, ensuring that they only pay for the compute resources they actually use while the application runs. In contrast, data storage and updates, while important in a machine learning workflow, do not directly relate to the primary advantage of using AWS Lambda. Standardizing model accuracy is also not a feature of AWS Lambda, as it primarily focuses on execution rather than model performance across applications. Therefore, the key advantage of AWS Lambda for machine learning applications lies in its ability to facilitate a serverless architecture that minimizes operational overhead.

SAMPLE

Next Steps

Congratulations on reaching the final section of this guide. You've taken a meaningful step toward passing your certification exam and advancing your career.

As you continue preparing, remember that consistent practice, review, and self-reflection are key to success. Make time to revisit difficult topics, simulate exam conditions, and track your progress along the way.

If you need help, have suggestions, or want to share feedback, we'd love to hear from you. Reach out to our team at hello@examzify.com.

Or visit your dedicated course page for more study tools and resources:

https://awscertmachinelearningspecialty.examzify.com

We wish you the very best on your exam journey. You've got this!

SAMPLE