

考试指南 (MLS-C01)

# AWS Certified Machine Learning - Specialty



# AWS Certified Machine Learning - Specialty: 考试指南 (MLS-C01)

Copyright © 2026 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon's trademarks and trade dress may not be used in connection with any product or service that is not Amazon's, in any manner that is likely to cause confusion among customers, or in any manner that disparages or discredits Amazon. All other trademarks not owned by Amazon are the property of their respective owners, who may or may not be affiliated with, connected to, or sponsored by Amazon.

---

# Table of Contents

AWS Certified Machine Learning - Specialty (MLS-C01)	1
简介	1
目标考生说明	2
建议掌握的 AWS 知识	2
超出目标考生考试范围的工作任务	2
考试内容	2
内容大纲	3
服务参考资料	3
内容领域 1：数据工程	3
任务 1.1：为 ML 创建数据存储库	4
任务 1.2：确定并实施数据摄取解决方案	4
任务 1.3：确定并实施数据转换解决方案	4
内容领域 2：探索性数据分析	4
任务 2.1：清理和准备数据进行建模	5
任务 2.2：执行特征工程	5
任务 2.3：分析和可视化数据，进行 ML 处理	5
内容领域 3：建模	5
任务 3.1：将业务问题转换为 ML 问题	5
任务 3.2：为给定 ML 问题选择相应的模型	6
任务 3.3：训练 ML 模型	6
任务 3.4：执行超参数优化	6
任务 3.5：评估 ML 模型。	6
内容领域 4：机器学习实施和操作	7
任务 4.1：针对性能、可用性、可扩展性、韧性和容错能力构建 ML 解决方案	7
任务 4.2：针对给定问题建议和实施相应的 ML 服务和功能	7
任务 4.3：将基本 AWS 安全实践应用于 ML 解决方案	8
任务 4.4：部署和运行 ML 解决方案	8
考试范围内的 AWS 服务	8
分析	9
计算	9
容器	10
数据库	10
物联网	10
机器学习	10

---

管理和监管 .....	11
联网和内容分发 .....	11
安全性、身份与合规性 .....	11
存储 .....	11
考试范围外的 AWS 服务 .....	11
分析 .....	11
机器学习 .....	12
技术和概念 .....	12
调查问卷 .....	12

# AWS Certified Machine Learning - Specialty (MLS-C01)

AWS Certified Machine Learning - Specialty (MLS-C01) 考试面向担任人工智能和机器学习 (AI/ML) 开发或数据科学角色的个人。本考试考查考生使用 AWS 云针对给定业务问题设计、构建、部署、优化、训练、调整和维护 ML 解决方案的能力。

## 主题

- [简介](#)
- [目标考生说明](#)
- [考试内容](#)
- [内容大纲](#)
- [服务参考资料](#)
- [内容领域 1：数据工程](#)
- [内容领域 2：探索性数据分析](#)
- [内容领域 3：建模](#)
- [内容领域 4：机器学习实施和操作](#)
- [考试范围内的 AWS 服务](#)
- [考试范围外的 AWS 服务](#)
- [技术和概念](#)
- [调查问卷](#)

## 简介

[AWS Certified Machine Learning - Specialty \(MLS-C01\)](#) 考试面向担任人工智能和机器学习 (AI/ML) 开发或数据科学角色的个人。本考试考查考生使用 AWS 云针对给定业务问题设计、构建、部署、优化、训练、调整和维护 ML 解决方案的能力。

本考试还考查考生完成以下任务的能力：

- 为给定业务问题选择相应的 ML 方法并证明其合理性。
- 确定相应的 AWS 服务来实施 ML 解决方案。
- 设计和实施可扩展、成本优化、可靠且安全的 ML 解决方案。

## 目标考生说明

目标考生应具备 2 年或更长时间在 AWS 云中开发、架构和运行 ML 或深度学习工作负载的经验。

## 建议掌握的 AWS 知识

目标考生应具备以下 AWS 知识：

- 执行基本超参数优化的经验
- ML 和深度学习框架的经验

## 超出目标考生考试范围的工作任务

下表列出了不要求目标考生必须掌握的知识。此列表并非详尽无遗。以下领域的知识超出考试范围：

- 广泛或复杂的算法开发
- 广泛的超参数优化
- 复杂的数学证明和计算
- 高级联网和网络设计
- 高级数据库、安全性和 DevOps 概念
- Amazon EMR 的 DevOps 相关任务

## 考试内容

考试具有两种类型的试题：

- 单选题：具有一个正确答案和三个错误答案（干扰项）
- 多选题：在五个或更多答案选项中具有两个或更多正确答案

选择一个或多个最准确表述或回答试题的答案。干扰项或错误答案是知识或技能不全面的考生可能会选择的答案选项。干扰项通常是与内容领域相符的看似合理的答案。

未回答的试题将计为回答错误；猜答案不会扣分。本考试包括 50 道试题，这些试题将影响您的分数。

本考试包括 15 道不计分试题，这些试题不影响您的分数。AWS 收集这些不计分试题的答题情况进行评估，以便将来将这些试题作为计分试题。在考试中不会标明这些不计分试题。

AWS Certified Machine Learning - Specialty (MLS-C01) 考试结果分为及格和不及格两种。本考试按照 AWS 专业人员根据认证行业最佳实践和准则制定的最低标准进行评分。

您的考试结果换算分数为 100-1000 分。最低及格分数为 750 分。您的分数表明您的总体考试答题情况以及是否通过考试。标准分模型有助于将难易程度可能略有不同的多种考试形式中的分数进行公平处理。

您的成绩单可能包含一个分类表，其中列出您在每个部分的考试成绩。本考试采用补偿评分模型，这意味着您无需在每个部分都达到及格分数。您只需通过整体考试即可。

考试的每个部分具有特定的权重，因此，某些部分的试题比其他部分多。分类表包含常规信息，用于突出显示您的强项和弱项。在解读各个部分的反馈时，请务必小心谨慎。

## 内容大纲

本考试指南包括考试的权重、内容领域和任务说明，并未列出考试的全部内容。不过，每个任务说明都提供有额外的背景信息，有助于您备考。

考试中考查的内容领域和相应的权重如下：

- [内容领域 1：数据工程（占计分内容的 20%）](#)
- [内容领域 2：探索性数据分析（占计分内容的 24%）](#)
- [内容领域 3：建模（占计分内容的 36%）](#)
- [内容领域 4：机器学习实施和操作（占计分内容的 20%）](#)

## 服务参考资料

以下各节提供与本认证考试相关的 AWS 服务、技术和概念的详细信息：

- [考试范围内的 AWS 服务](#)
- [考试范围外的 AWS 服务](#)
- [技术和概念](#)

## 内容领域 1：数据工程

任务

- [任务 1.1：为 ML 创建数据存储器](#)

- [任务 1.2：确定并实施数据摄取解决方案](#)
- [任务 1.3：确定并实施数据转换解决方案](#)

## 任务 1.1：为 ML 创建数据存储库

- 确定数据来源（例如：内容和位置，用户数据等主要来源）。
- 确定存储介质（例如：数据库、Amazon S3、Amazon Elastic File System [Amazon EFS]、Amazon Elastic Block Store [Amazon EBS]）。

## 任务 1.2：确定并实施数据摄取解决方案

- 确定数据任务方式和任务类型（例如：批量加载、流式处理）。
- 编排数据摄取管道（基于批处理的 ML 工作负载和基于流式处理的 ML 工作负载）。
  - Amazon Kinesis
  - Amazon Data Firehose
  - Amazon EMR
  - AWS Glue
  - 适用于 Apache Flink 的亚马逊托管服务
- 安排任务。

## 任务 1.3：确定并实施数据转换解决方案

- 转换传输中的数据（ETL、AWS Glue、Amazon EMR、AWS Batch）。
- 使用 MapReduce 处理特定于 ML 的数据（例如：Apache Hadoop、Apache Spark、Apache Hive）。

## 内容领域 2：探索性数据分析

### 任务

- [任务 2.1：清理和准备数据进行建模](#)
- [任务 2.2：执行特征工程](#)
- [任务 2.3：分析和可视化数据，进行 ML 处理](#)

## 任务 2.1：清理和准备数据进行建模

- 确定和处理缺少的数据、损坏的数据和筛选停用词等。
- 格式化、标准化、扩充和扩缩数据。
- 确定是否有足够的已标记数据。
  - 确定缓解策略。
  - 使用数据标记工具（例如：Amazon Mechanical Turk）。

## 任务 2.2：执行特征工程

- 从数据集（包括从文本、语音、图像和公开数据集等数据来源）中识别和提取特征。
- 分析和评估特征工程概念（例如：分箱、词元化、异常值、合成特征、独热编码、降低数据维度）。

## 任务 2.3：分析和可视化数据，进行 ML 处理

- 创建图形（例如：散点图、时间序列图、直方图、箱线图）。
- 解释描述性统计数据（例如：相关性、汇总统计数据、p 值）。
- 执行聚类分析（例如：分层分析、诊断、肘图、聚类大小）。

## 内容领域 3：建模

### 任务

- [任务 3.1：将业务问题转换为 ML 问题](#)
- [任务 3.2：为给定 ML 问题选择相应的模型](#)
- [任务 3.3：训练 ML 模型](#)
- [任务 3.4：执行超参数优化](#)
- [任务 3.5：评估 ML 模型。](#)

## 任务 3.1：将业务问题转换为 ML 问题

- 确定何时使用以及何时不使用 ML。
- 了解有监督学习和无监督学习的区别。

- 从分类、回归、预测、聚类、建议以及基础模型中进行选择。

## 任务 3.2：为给定 ML 问题选择相应的模型

- XGBoost、逻辑回归、k-means、线性回归、决策树、随机森林、RNN、CNN、集成学习、迁移学习，以及大型语言模型 (LLM)
- 表达模型背后的直觉知识。

## 任务 3.3：训练 ML 模型

- 拆分数据分别用于训练和验证（例如：交叉验证）。
- 了解 ML 训练的优化技术（例如：梯度下降、损失函数、收敛）。
- 选择合适的计算资源（例如：GPU 或 CPU，分布式或非分布式）。
  - 选择合适的计算平台（Spark 或非 Spark）。
- 更新和重新训练模型。
  - 批量或实时/线上

## 任务 3.4：执行超参数优化

- 执行正则化。
  - 随机失活
  - L1/L2
- 执行交叉验证。
- 初始化模型。
- 了解神经网络架构（层和节点）、学习率、激活函数。
- 了解基于树的模型（树数量、级别数量）。
- 了解线性模型（学习率）。

## 任务 3.5：评估 ML 模型。

- 避免过拟合或欠拟合。
  - 检测和处理偏差和方差。

- 评估指标 ( 例如 : 曲线下面积 [AUC] - 接收方操作特性 [ROC]、准确率、查准率、查全率、均方根误差 [RMSE]、F1 分数 )。
- 解释混淆矩阵。
- 执行离线和线上模型评估 ( A/B 测试 )。
- 使用指标 ( 例如 : 训练模型的时间、模型质量、工程成本 ) 比较模型。
- 执行交叉验证。

## 内容领域 4 : 机器学习实施和操作

### 任务

- [任务 4.1 : 针对性能、可用性、可扩展性、韧性和容错能力构建 ML 解决方案](#)
- [任务 4.2 : 针对给定问题建议和实施相应的 ML 服务和功能](#)
- [任务 4.3 : 将基本 AWS 安全实践应用于 ML 解决方案](#)
- [任务 4.4 : 部署和运行 ML 解决方案](#)

### 任务 4.1 : 针对性能、可用性、可扩展性、韧性和容错能力构建 ML 解决方案

- 记录和监控 AWS 环境。
  - AWS CloudTrail 和 Amazon CloudWatch
  - 构建错误监控解决方案。
- 部署到多个 AWS 区域和多个可用区。
- 创建 AMI 和黄金映像。
- 创建 Docker 容器。
- 部署自动扩缩组。
- 合理调整资源大小 ( 例如 : 实例、预调配 IOPS、卷 )。
- 执行负载均衡。
- 遵循 AWS 最佳实践。

### 任务 4.2 : 针对给定问题建议和实施相应的 ML 服务和功能

- AWS 上的 ML ( 应用程序服务 ) , 例如 :

- Amazon Polly
- Amazon Lex
- Amazon Transcribe
- Amazon Q
- 了解 AWS 服务配额。
- 确定何时构建自定义模型以及何时使用 Amazon SageMaker 内置算法。
- 了解 AWS 基础设施（例如：实例类型）以及与成本相关的注意事项。
  - 使用竞价型实例通过 AWS Batch 训练深度学习模型。

## 任务 4.3：将基本 AWS 安全实践应用于 ML 解决方案

- AWS Identity and Access Management (IAM)
- S3 存储桶策略
- 安全组
- VPC
- 加密和匿名化

## 任务 4.4：部署和运行 ML 解决方案

- 公开终端节点并与之进行交互。
- 了解 ML 模型。
- 执行 A/B 测试。
- 重新训练管道。
- 对 ML 模型进行调试和故障排除。
  - 检测和缓解性能下降。
  - 监控模型性能。

## 考试范围内的 AWS 服务

以下列表包含 AWS Certified Machine Learning - Specialty (MLS-C01) 考试范围内的 AWS 服务和功能。此列表并非详尽无遗，且可能随时会更改。AWS 各项产品和服务按其进行主要功能进行分类。

## 主题

- [分析](#)
- [计算](#)
- [容器](#)
- [数据库](#)
- [物联网](#)
- [机器学习](#)
- [管理和监管](#)
- [联网和内容分发](#)
- [安全性、身份与合规性](#)
- [存储](#)

## 分析

- Amazon Athena
- Amazon Data Firehose
- Amazon EMR
- AWS Glue
- Amazon Kinesis
- Amazon Kinesis Data Streams
- AWS Lake Formation
- 适用于 Apache Flink 的亚马逊托管服务
- Amazon OpenSearch Service
- Amazon QuickSight

## 计算

- AWS Batch
- Amazon EC2
- AWS Lambda

## 容器

- Amazon Elastic Container Registry (Amazon ECR)
- Amazon Elastic Container Service (Amazon ECS)
- Amazon Elastic Kubernetes Service (Amazon EKS)
- AWS Fargate

## 数据库

- Amazon Redshift

## 物联网

- AWS IoT Greengrass

## 机器学习

- Amazon Bedrock
- Amazon Comprehend
- AWS Deep Learning AMI (DLAMI)
- Amazon Forecast
- Amazon Fraud Detector
- Amazon Lex
- Amazon Kendra
- Amazon Mechanical Turk
- Amazon Polly
- Amazon Q
- Amazon Rekognition
- Amazon SageMaker
- Amazon Textract
- Amazon Transcribe
- Amazon Translate

## 管理和监管

- AWS CloudTrail
- Amazon CloudWatch

## 联网和内容分发

- Amazon VPC

## 安全性、身份与合规性

- AWS Identity and Access Management (IAM)

## 存储

- Amazon Elastic Block Store (Amazon EBS)
- Amazon Elastic File System (Amazon EFS)
- Amazon FSx
- Amazon S3

## 考试范围外的 AWS 服务

以下列表包含 AWS Certified Machine Learning - Specialty (MLS-C01) 考试范围外的 AWS 服务和功能。此列表并非详尽无遗，并且可能随时会更改。与考试目标职位完全无关的 AWS 产品和服务的内容，均未包含在本表中。

### 主题

- [分析](#)
- [机器学习](#)

## 分析

- AWS Data Pipeline

## 机器学习

- AWS DeepRacer
- Amazon Machine Learning (Amazon ML)

## 技术和概念

以下列表包含考试中可能出现的技术和概念。此列表并非详尽无遗，并且可能随时会更改。表中各项的顺序与位置，并不代表考试中的相对权重或重要性：

- 摄取和收集
- 处理和 ETL
- 数据分析和可视化
- 模型训练
- 模型部署和推理
- 运行 ML
- AWS 机器学习应用服务
- 与 ML 相关的语言 ( 例如 : Python、Java、Scala、R、SQL )
- 笔记本和集成开发环境 (IDE)

## 调查问卷

本考试指南对您有帮助吗？欢迎填写[调查问卷](#)，与我们分享您的建议。