

AW 认证 AI 实践者考试旨在验证您对人工智能、机器学习、生成式 AI 概念及其使用案例的知识。

1. 一家公司正在开发一个模型，根据各种特征（如尺寸、重量、品牌和制造日期）来预测产品价格。哪种机器学习方法最适合此任务？

- A. 分类
- B. 回归
- C. 聚类
- D. 维度缩减

正确答案为选项 B. 回归。回归用于预测连续数值，而分类则用于预测分类结果；聚类将类似数据点分组，而维度缩减则是减少特征数量。

---

2. 一家公司正在扩大对人工智能的应用，应该优先考虑哪项核心原则，以建立 AI 开发和使用的清晰指导、监督和问责机制？

- A. 偏见预防
- B. 准确性与可靠性
- C. 数据保护与安全
- D. 治理

正确答案为选项 D. 治理。治理涵盖管理 AI 系统的整体框架，包括政策、流程和决策机制。尽管其他选项也很重要，但从监督的角度来看，治理是最佳答案。

---

3. 一家公司开始在 AWS 上使用生成式 AI，为确保负责任的 AI 实践，哪种工具可以为其提供指导和信息？

- A. AWS 市场
- B. AWS AI 服务卡

C. SageMaker

D. Bedrock

正确答案为选项 B. AWS AI 服务卡。AI 服务卡提供有关特定 AWS 服务的详细信息，包括使用案例、限制及负责任的设计考虑。

---

4. 机器学习中的特征工程的主要目的是什么？

A. 确保模型的稳定性能

B. 评估模型的性能

C. 收集和预处理数据特征

D. 转换数据并创建模型所需的变量或特征

正确答案为选项 D. 转换数据并创建变量或特征供模型使用。

---

5. 一家小公司希望使用机器学习来预测客户流失，但缺乏专门的数据科学团队。哪种 AWS 工具可以帮助他们轻松地构建模型而无需大量编码？

A. Amazon SageMaker JumpStart

B. SageMaker Studio

C. SageMaker Canvas

D. SageMaker Data Wrangler

正确答案为选项 C. SageMaker Canvas。SageMaker Canvas 提供可视化界面，无需编码即可构建机器学习模型。

---

6. 一家金融机构正在开发欺诈检测模型，项目负责人宣布将使用 MLOps。在此项目背景下，如何解释 MLOps？

- A. 用于可视化机器学习模型性能的工具
- B. 用于管理机器学习系统全生命周期的一组实践
- C. 用于在生产环境中部署和维护机器学习模型的流程
- D. 构建和训练机器学习模型的框架

正确答案为选项 B. MLOps 是管理机器学习模型全生命周期的一组实践，涵盖了模型的开发、部署、监控和维护。

---

7. 哪种 AWS 服务可用于创建基于知识的聊天机器人，能够根据公司内部文件回答有关产品和服务的问题？

- A. Amazon SageMaker
- B. Amazon Q Business
- C. Amazon Polly
- D. Amazon Rekognition

正确答案为选项 B. Amazon Q Business。该服务允许企业创建智能聊天机器人，并能够基于公司内部文件回答问题。

---

8. 开发团队需要选择一种服务来存储和查询向量嵌入，哪种 AWS 服务最适合此需求？

- A. Glue Data Catalog
- B. Amazon S3
- C. Redshift
- D. OpenSearch Service

正确答案为选项 D. OpenSearch Service。OpenSearch 支持基于向量的存储和查询。

---

9. 一家公司希望评估 AWS 服务的安全性和合规性实践，这些服务是供销售 AI 产品的厂商使用的。哪项 AWS 服务可以帮助他们访问 AWS 合规性报告和证书？

- A. AWS Organizations
- B. Amazon Inspector
- C. AWS CloudTrail
- D. AWS Artifact

正确答案为选项 D. AWS Artifact。Artifact 提供集中式的 AWS 合规性报告和证书。

---

10. 一个机器学习模型在训练数据上表现良好，但在新数据上表现不佳，可能的问题是什么？

- A. 过拟合
- B. 欠拟合
- C. 训练数据不足
- D. 数据质量差

正确答案为选项 A. 过拟合。过拟合意味着模型过于复杂，以至于在新数据上无法泛化。

---

11. 一家公司希望通过访问外部信息来提高大型语言模型的响应质量，哪种方法所需的开发工作量最少？

- A. 小样本学习 (Few-shot Learning)

B. 零样本学习 (Zero-shot Learning)

C. 检索增强生成 (RAG)

D. 微调 (Fine-tuning)

正确答案为选项 C. 检索增强生成 (RAG)。RAG 能够在提示中整合外部信息，以提高生成结果的准确性。

---

12. 一个模型已被训练识别手写数字，但准确率不高。专家建议增加 Epoch 值。在机器学习中，Epoch 是什么？

A. 测量模型在训练期间的准确性

B. 模型通过整个训练数据集的一次遍历

C. 将数据集划分为训练集和测试集的过程

D. 神经网络中的层数

正确答案为选项 B. 模型通过整个训练数据集的一次遍历。

---

13. 在机器学习模型中，以下哪一项被认为是超参数？

A. 神经网络的权重

B. 优化算法的学习率

C. 激活函数的输出

D. 模型的预测结果

正确答案为选项 B. 学习率是调整模型在训练过程中步伐大小的超参数。

---

14. 一个模型在输入稍微改变时，仍然给出非常相似的输出。可以调整哪个推理时间参数以增加创造性？

- A. 学习率
- B. 批次大小
- C. 温度
- D. Epoch

正确答案为选项 C. 温度。温度控制生成时的随机性，更高的温度增加随机性和多样性。

---

15. 您正在评估语言生成模型在文本生成相关任务中的质量。哪个评估指标最能衡量其与人类书写文本的语义相似度？

- A. BERT Score
- B. BLEU
- C. Perplexity
- D. ROUGE

正确答案为选项 A. BERT Score。此指标使用预训练语言模型来计算文本间的语义相似度。

---

16. 一名开发人员正在设计 AI 系统，需要一个提供全面工具来分析和解释模型预测的解决方案。哪项 AWS 服务专门用于提高透明度和可解释性？

- A. SageMaker Clarify
- B. SageMaker Debugger
- C. SageMaker Autopilot
- D. SageMaker Data Wrangler

正确答案为选项 A. SageMaker Clarify。Clarify 提供分析模型预测和检测偏见的工具。

---

17. 一家公司计划训练并构建自己的基础模型，这种方法与使用预训练基础模型相比有什么潜在缺点？

- A. 更复杂的实施过程
- B. 性能降低
- C. 更高的错觉风险
- D. 增加的开发成本

正确答案为选项 A 和 D。即更复杂的实施过程和增加的开发成本。

---

18. 一家公司希望使用现有的流行预训练 AI 模型生成内容，但没有丰富的 AI 专业知识，且不想自行管理模型。哪种 AWS 服务最适合他们的需求？

- A. Amazon Textract
- B. Amazon Comprehend
- C. Amazon Bedrock
- D. Amazon SageMaker

正确答案为选项 C. Amazon Bedrock。Bedrock 提供对预训练基础模型的访问，无需广泛的技术知识或基础设施。

---

19. 要微调模型以响应某种格式和风格的问题，最适合的训练数据类型是什么？

- A. 列格式数据集

- B. 卷标数据
- C. 转录日志
- D. 包含提示和回应的文本对

正确答案为选项 D. 包含提示和回应的文本对。

---

20. 公司需要记录 Amazon Bedrock 的 API 调用信息以确保合规性，包括 API 调用的详细信息、用户及时间戳。哪项 AWS 服务可以帮助他们完成此操作？

- A. CloudTrail
- B. CloudWatch
- C. IAM
- D. Security Hub

正确答案为选项 A. CloudTrail。CloudTrail 记录 AWS 服务的 API 调用，适用于合规性跟踪。

---

21. 一个数据科学团队希望提高模型性能，他们想增加训练数据的数量和多样性，并调整算法的学习率。哪组 ML 流程步骤可以满足这些要求？

- A. 数据增强
- B. 模型监控
- C. 特征工程
- D. 超参数调整

正确答案为选项 A 和 D。数据增强增加数据多样性，超参数调整可控制模型的学习率。

---



22. 一家公司希望确保其 Amazon Bedrock 驱动的应用生成的内容符合道德准则，避免有害或冒犯性内容。哪项 AWS 服务可以帮助他们实现这些安全措施？

- A. SageMaker
- B. Comprehend
- C. Textract
- D. Bedrock Guardrails

正确答案为选项 D. Bedrock Guardrails。它可防止生成有害、偏见或不当的内容。

---

23. 公司正在 AWS 上使用 Amazon SageMaker 训练机器学习模型，其数据集存储在 S3 中且包含敏感的客户信息。如何确保在训练模型之前移除或匿名化数据中的敏感信息？

- A. 使用 S3 加密来保护静态数据
- B. 使用 Amazon Macie 识别数据集中敏感信息
- C. 使用 S3 访问控制限制授权人员访问
- D. 实施数据屏蔽技术来替换敏感信息

正确答案为选项 B 和 D。使用 Macie 识别敏感信息并实施数据屏蔽技术。

---

24. 一家公司希望使用生成式 AI 为其产品创建营销标语。为什么公司需要仔细审查所有生成的标语？

- A. 生成式 AI 可能生成过长且难以记住的标语
- B. 生成式 AI 可能难以捕捉公司的独特品牌形象
- C. 生成式 AI 可能生成不当或误导性的标语

D. 生成式 AI 可能需要大量训练数据来生成有效的标语

正确答案为选项 C. 生成式 AI 有时会生成带有偏见或冒犯性的内容，因此需要人为审查以确保质量和适当性。

---

25. 您的公司正在 EC2 实例上训练机器学习模型，担心模型的安全性并希望识别基础设施中的潜在漏洞。哪项 AWS 服务可以帮助您扫描 EC2 实例的漏洞？

A. AWS X-Ray

B. Amazon CloudWatch

C. Amazon Inspector

D. AWS Config

正确答案为选项 C. Amazon Inspector。它是一项安全服务，可帮助识别 EC2 实例中的漏洞。

---

26. 一个机器学习模型在城市地区的申请人中表现更佳，因为训练数据中有更多来自城市地区的批准样本。这是哪种类型的偏见的例子？

A. 抽样偏差

B. 算法偏差

C. 观察者偏差

D. 时间偏差

正确答案为选项 A. 抽样偏差。抽样偏差发生在训练数据不具备全面代表性时。

---

27. 在具有多个用户之间关系的社交网络数据集中，哪种机器学习算法最适合将这些相互关联的关系分类到预定义的类别中？

- A. 线性回归
- B. 判定树
- C. 图神经网络 (Graph Neural Networks)
- D. 逻辑回归

正确答案为选项 C. 图神经网络。图神经网络可以处理互相关联的数据并捕捉节点之间的关系。

---

28. 一个机器人被指派要在迷宫中找到目标。哪种机器学习范式最适合训练机器人通过自学和试错找到最佳路径？

- A. 监督学习
- B. 无监督学习
- C. 随机学习
- D. 强化学习

正确答案为选项 D. 强化学习。强化学习让模型根据获得的奖励或惩罚进行决策。

---

29. 一位研究人员希望使一个预训练的机器学习模型在新的特定领域任务上表现良好，但只有有限的标记数据。以下哪种方法最有效且最合适？

- A. 使用额外的未标记数据继续预训练
- B. 使用来自新领域的标记数据进行微调
- C. 使用预训练模型而不做进一步调整
- D. 使用新的标记数据从头开始训练

正确答案为选项 B. 微调。微调在特定任务上有效率且适合有限的标记数据。

---

30. 如果您是一家具有不稳定工作负载的小型初创公司，并希望在 Amazon Bedrock 上实验不同的基础模型，哪种计价模式最适合您？

- A. 按需模式
- B. 预置吞吐量
- C. 模型定制
- D. 自定义合同

正确答案为选项 A. 按需模式。按需模式适合不稳定的工作负载，按使用量计费。

---

31. 在自然语言处理的背景下，以下哪项是用于表示单词或子词的基本单位？

- A. 标记 (Token)
- B. 向量嵌入 (Vector Embedding)
- C. N-gram
- D. 词汇表

正确答案为选项 A. 标记。标记是表示单词或子词的基本单位。

---

32. 开发人员正在创建一个预测客户流失的系统。为了确保透明性，他们需要记录有关模型的关键细节。哪项 AWS 工具最适合此任务？

- A. Amazon SageMaker Clarify
- B. AWS AI 服务卡
- C. SageMaker 模型卡
- D. SageMaker JumpStart

正确答案为选项 C. SageMaker 模型卡。模型卡用于记录和共享机器学习模型的相关信息。

---

33. 工程师正在训练一个机器学习模型，以防止欠拟合或过拟合，应如何对模型进行数据训练？

- A. 高偏差和高方差
- B. 低偏差和低方差
- C. 高偏差和低方差
- D. 低偏差和高方差

正确答案为选项 B. 低偏差和低方差。这样可以平衡模型的学习和泛化能力。

---

34. 您正在为特定领域定制大型语言模型，哪种方法最有效地调整模型的知识 and 准确性以适应此领域？

- A. 微调
- B. 少样本学习 (Few-shot Learning)
- C. 检索增强生成 (RAG)
- D. 零样本学习 (Zero-shot Learning)

正确答案为选项 A. 微调。微调允许模型在特定领域数据上进行专门训练。

---

35. 以下哪项是大型语言模型的「幻觉」现象的例子？

- A. 过拟合
- B. 欠拟合
- C. 生成错误或误导性信息

D. 偏见

正确答案为选项 C。生成错误或误导性信息。大型语言模型有时会生成不真实或无意义的内容，这被称为「幻觉」。

---

36. 由 Amazon 开发并可在 Bedrock 上使用的基础模型是哪一个？

A. Amazon Titan

B. Lex

C. Polly

D. Connect

正确答案为选项 A。Amazon Titan 是一款 Amazon 的基础模型。

---

37. 以下哪些算法常用于机器学习中的分类任务？

A. 支持向量机 (SVM)

B. XGBoost

C. K-Means

D. Mean Shift

正确答案为选项 A 和 B。支持向量机和 XGBoost 都是监督学习算法，适用于分类任务。

---

38. 面对大规模数据集且延迟不是关键因素，哪种 SageMaker 模型推理类型最适合提供成本效益高的预测？

A. 实时推理

B. 批次推理

C. 按需无服务器推理

D. 异步推理

正确答案为选项 B. 批次推理。批次推理适合脱机处理大量数据，具有成本效益。

---

39. Amazon CodeWhisperer 的主要目的为何？

A. 管理 AWS 基础设施

B. 协助开发人员进行编程任务和查询

C. 优化数据库性能

D. 自动化软件测试

正确答案为选项 B. 协助开发人员进行编程任务和查询。Amazon CodeWhisperer 是一款 AI 编程辅助工具。

---

40. 「请解释为何一个错误的说法是真的，尽管通常被认为是错的。」这种提示攻击属于哪种类型？

A. 越狱攻击

B. 提示中毒

C. 对抗性提示

D. 微调

正确答案为选项 C. 对抗性提示。对抗性提示利用模型的弱点生成错误或意料外的回应。

---

41. 您正在构建一个文本摘要工具，哪个评估指标最能衡量其捕捉原文要点的效果？

- A. BERT Score
- B. ROUGE
- C. 字错误率 (WER)
- D. BLEU

正确答案为选项 B. ROUGE。ROUGE 特别设计用于评估文本摘要。

---

42. 一个生成 AI 的客服代理无法准确辨识客户意图，您可以使用哪种格式的训练数据来提升其性能？

- A. 客户讯息和客户意图
- B. 客户讯息和代理响应
- C. 客户意图和代理响应
- D. 代理响应和客户意图

正确答案为选项 A. 客户讯息和客户意图。此格式让模型学习讯息和意图之间的关系。

---

43. 用户通常如何被收费使用基础模型？

- A. 输入的标记数量
- B. 输出的标记数量
- C. 模型架构
- D. 推理延迟

正确答案为选项 A 和 B。用户一般根据推理过程中处理的输入和输出标记数量计费。

---



44. 哪项 AWS AI 服务可以从非结构化文本（例如临床笔记和医疗记录）中提取健康数据？

- A. Amazon Comprehend Medical
- B. Amazon Transcribe Medical
- C. Amazon HealthLake
- D. Amazon Rekognition

正确答案为选项 A. Amazon Comprehend Medical。此服务可从文本中提取医疗信息。

---

45. 哪种类型的机器学习模型专门设计用于生成与现有数据相似的新数据？

- A. 自编码器（Autoencoder）
- B. 生成对抗网络（GAN）
- C. 判定树
- D. 支持向量机（SVM）

正确答案为选项 B. 生成对抗网络。GAN 专为生成新数据设计，例如创建新图像或文本。

---

46. 用户将使用长提示来向大型语言模型提问，选择 LLM 时应考虑哪个关键方面？

- A. 推理延迟
- B. 最大上下文窗口
- C. 模型大小
- D. 训练数据

正确答案为选项 B. 最大上下文窗口。较大的上下文窗口可以让模型处理和理解更多提示中的信息。

---

47. Amazon SageMaker Feature Store 的主要用途是什么？

- A. 自动训练和部署机器学习模型
- B. 存储和管理机器学习工作流的特征
- C. 提供预训练机器学习模型的市场
- D. 优化 SageMaker 训练任务的性能

正确答案为选项 B. 存储和管理机器学习工作流的特征。Feature Store 是一个专门管理机器学习特征的储存库。

---

48. 一家医疗保健机构正在开发一个 AI 驱动的诊断工具，以帮助早期检测罕见疾病。就合规性而言，以下哪项最不相关？

- A. 确保 AI 系统对某些患者群体不存偏见
- B. 最小化 AI 系统的运营费用
- C. 确保 AI 系统的决策过程透明
- D. 防止 AI 系统被用于未经授权的用途

正确答案为选项 B. 最小化运营费用。合规性主要关注于公平性、透明性和安全性，而运营费用相对不重要。

---

49. 您是一家大型企业，拥有大量分散在各内部系统中的非结构化数据，想为员工提供一个强大的搜索工具，以理解自然语言查询并返回准确、相关的结果。哪项 AWS 服务最适合满足此需求？

- A. Amazon Redshift

B. Amazon Lex

C. Amazon Kendra

D. Amazon DynamoDB

正确答案为选项 C. Amazon Kendra。Kendra 是一项管理式服务，支持使用自然语言查询进行非结构化数据搜索。

---

50. 一名数据科学家正进行一个项目，该项目需要快速构建原型并使用各种机器学习算法进行实验。哪项 AWS 服务最适合此任务？

A. Amazon SageMaker Ground Truth

B. Amazon EC2

C. Amazon SageMaker Autopilot

D. Amazon Bedrock

正确答案为选项 C. Amazon SageMaker Autopilot。Autopilot 自动执行模型的构建、训练和部署过程，适合快速原型构建。

---

51. 一家大型公司希望为销售经理创建一个应用程序，能够进行推理、执行多步骤任务并从企业数据中提供有见解的回应。哪项 AWS 服务最适合此任务？

A. Amazon Lex

B. SageMaker

C. Bedrock Knowledge Bases

D. Bedrock Agents

正确答案为选项 D. Bedrock Agents。它可以利用大型语言模型（LLM）和知识库提供基于企业数据的回答。

---

52. 一家公司希望分析客户评论，以识别常见主题和情绪。哪项 AWS 服务可以帮助公司达到此目标？

- A. Amazon Connect
- B. Amazon Comprehend
- C. Amazon Translate
- D. Amazon Transcribe

正确答案为选项 B. Amazon Comprehend。Comprehend 可从文本中提取情绪、实体和关键短语。

---

53. 一家公司希望将数据从一种格式转换为另一种格式，以便进行机器学习任务。哪项 AWS 服务最适合进行此类数据转换？

- A. AWS Glue
- B. Amazon Translate
- C. AWS Config
- D. Amazon Kinesis

正确答案为选项 A. AWS Glue。Glue 是一项完全管理的 ETL 服务，用于将数据从一种格式转换为另一种格式。

---

54. 一家公司希望为实时推理部署一个训练好的机器学习模型。哪项 AWS 服务最适合此用途？

- A. Amazon SageMaker JumpStart
- B. Amazon Personalize
- C. Amazon EC2
- D. SageMaker Endpoints

正确答案为选项 D. SageMaker Endpoints。该服务专为实时推理设计，允许部署训练好的模型并进行实时预测。

---

55. 一家公司已部署一个用于客户情绪分析的机器学习模型。为了确保模型的准确性和可靠性，应使用哪些 AWS 服务进行监控和人工审查？

- A. Amazon Bedrock
- B. SageMaker Model Monitor
- C. SageMaker Ground Truth
- D. Amazon A2I

正确答案为选项 B 和 D。SageMaker Model Monitor 可监控模型性能，而 Amazon A2I 支持在机器学习工作流程中引入人工审查。

---

56. 一位机器学习专家正在 Amazon SageMaker 上对一个大型深度学习模型进行训练，但单个 GPU 无法高效处理此任务。哪个 SageMaker 功能可以帮助优化大模型和大数据集的训练过程？

- A. 增量训练
- B. 超参数调整
- C. 管道模式 (Pipe Mode)
- D. 模型并行

正确答案为选项 D. 模型并行。此功能将大模型分割到多个 GPU 或实例上进行处理。

---

57. 您正在处理包含大量特征的数据集，为了提升模型性能和计算效率，您需要在不丢失重要信息的情况下简化数据。哪种技术最有效？

- A. 降维

B. 特征工程

C. 数据增强

D. 数据清理

正确答案为选项 A. 降维。降维技术可减少数据特征数量，并保留关键信息。

---

58. 您希望根据文本描述生成高质量的图像，哪种 AI 模型最适合生成任务并能产生高质量、多样化的输出？

A. 生成对抗网络（GAN）

B. 循环神经网络（RNN）

C. 卷积神经网络（CNN）

D. 稳定扩散（Stable Diffusion）

正确答案为选项 D. 稳定扩散。此模型特别适合生成图像任务。

---

59. 一家公司有一个系统可以从产品数据生成向量嵌入，他们希望提高寻找相似产品的速度和准确性。哪项 AWS 服务最适合实现向量搜索以优化系统？

A. OpenSearch Service

B. Redshift

C. Neptune

D. DocumentDB

正确答案为选项 A、C 和 D。OpenSearch、Neptune 和 DocumentDB 支持处理向量数据并进行相似性搜索。

---

60. 某银行每天收到大量贷款申请，贷款处理团队手动从这些申请中提取信息，这非常耗时。使用 AI 工具来自动化此过程的哪项 AWS 服务会有帮助？

- A. Amazon Rekognition
- B. Textract
- C. Translate
- D. Transcribe

正确答案为选项 B. Textract。Textract 可以从扫描的文件中提取文本。

---

61. 一家医疗公司希望开发一个机器学习模型，用于根据各种健康指标预测患者发展糖尿病的可能性。以下哪项指标最适合评估该模型的性能？

- A. 准确率
- B. 精确率
- C. F1 分数
- D. 召回率（灵敏度）
- E. ROC 曲线下面积（AUC）

正确答案为选项 D 和 E。召回率和 ROC 曲线下面积适合评估分类模型的性能。

---

62. 某组织已在大型一般图像数据集上训练了一个深度学习模型，现在他们希望使用该模型来分类医学图像，但数据集较小。哪种机器学习技术最适合这种情境？

- A. 强化学习
- B. 迁移学习

C. 监督学习

D. 无监督学习

正确答案为选项 B. 迁移学习。迁移学习允许模型在新的但相关的任务上进行微调。

---

63. 您正在 AWS 上构建一个机器学习模型，并希望安全地与第三方合作伙伴共享该模型。哪项 AWS 服务可用于在您的 VPC 和合作伙伴的 VPC 之间建立私有连接，确保数据不暴露在公网上？

A. Direct Connect

B. PrivateLink

C. Transit Gateway

D. VPN

正确答案为选项 B. PrivateLink。PrivateLink 支持在 VPC 之间建立安全的私有连接。

---

64. 您正在 AWS SageMaker 上使用敏感的客户数据进行模型训练。根据 AWS 的共享责任模型，以下哪项主要由您负责？

A. 确保 SageMaker 基础设施的安全

B. 保护 SageMaker 实例的操作系统

C. 确保存储在 S3 中的客户数据的安全

D. 为 AWS SageMaker 软件打补丁

正确答案为选项 C. 确保存储在 S3 中的客户数据的安全。根据 AWS 的共享责任模型，用户负责其客户数据的安全。

---



65. 在实施生成式 AI 安全范围矩阵时，以下哪项因素应用于评估与生成式 AI 项目相关的风险？

- A. 模型的计算效率
- B. 用于训练模型的数据敏感性
- C. 推理延迟
- D. 模型的参数数量

正确答案为选项 B. 用于训练模型的数据敏感性。数据的敏感性是评估生成式 AI 项目风险的重要因素。