



NVIDIA-Certified Associate: AI Infrastructure and Operations 学习指南



NVIDIA-Certified Associate: AI Infrastructure and Operations 学习指南

目录

AI 基础知识: 考试权重 38%	2
AI 基础架构: 考试权重 40%	3
AI 运营: 考试权重 22%	4

针对 NVIDIA AI Infrastructure and Operations (NCA-AIIO) 认证考试，本学习指南包含每个认证主题的介绍，并推荐相关培训课程和阅读资料。

查看所有 NVIDIA 认证，请[单击此处](#)。

考试适用人群

NCA-AIIO 是 Associate 级别的认证，用于验证与基础设施和运营相关的 AI 计算的基础理论。这项考试面向刚开始接触 AI 运营和基础架构，但需要了解和阐述在数据中心环境和本地环境中采用 AI 的不同组件和不同方面的 IT 专业人员。此认证适用于从技术售前到数据中心运营的各种工作角色。考生需要了解企业中的 AI 和 ML，能够描述 AI 运营的基本任务和要求，阐明 AI 部署中所需的各种软件和硬件，以及这些环境中将使用的 NVIDIA 软件、硬件和网络。

部分相关工作职责要求

1. 了解 AI 工作负载和用例
2. 区分 AI 和机器学习概念
3. 描述专门针对 AI 的数据中心运营的关键概念
4. 基本了解 AI 环境的网络需求
5. 深入了解 GPU 和 DPU 及其与 CPU 架构的区别
6. 与专业管理员协作，支持或负责 AI 数据中心的运营
7. 了解集群编排和管理，作业调度和监控要点
8. 说明如何在 AI 工作负载中使用虚拟化环境
9. 了解 AI 部署所使用的 NVIDIA 软件和硬件

建议具备的知识和经验

1. 具备企业级数据中心和本地计算环境方面的经验
2. 深入了解 AI 和机器学习概念
3. 理工科背景，如计算机科学、软件工程、AI 等

认证主题和参考资料

AI 基础知识：考试权重 38%

- 1.1 描述 AI 环境中使用的 NVIDIA 软件堆栈
 - 1.2 比较和对比训练和推理架构的要求和注意事项
 - 1.3 区分 AI、机器学习和深度学习的概念
 - 1.4 解释促成近期 AI 快速改进和采用的因素
 - 1.5 说明关键的 AI 用例和行业
 - 1.6 说明各种 NVIDIA 解决方案的用途和用例
 - 1.7 描述与 AI 开发和部署生命周期相关的软件组件
 - 1.8 比较和对比 GPU 和 CPU 架构
-

培训推荐 (可选)

课程参考：[AI 基础架构和运营基础](#)

- > 第 1 单元：各行各业的 AI 变革
- > 第 2 单元：AI 简介
- > 第 4 单元：通过 GPU 加速 AI
- > 第 5 单元：AI 软件生态系统
- > 第 7 单元：适用于 AI 的计算平台
- > 第 14 单元：编排、MLOps 和作业调度

阅读内容推荐

- > [NVIDIA® TensorRT™](#), NVIDIA Developer
- > [Deep Learning Training vs. Inference: Do You Know the Difference?](#), by AI TutorMaster, Medium
- > [Understanding Machine Learning Inference](#), Run:ai
- > [Tips on Scaling Storage for AI Training and Inferencing](#), NVIDIA Technical Blog
- > [What Is Machine Learning \(ML\)?](#), IBM
- > [Machine Learning: What It Is and Why It Matters](#), SAS
- > [What Are Large Language Models Used For?](#), NVIDIA Blog
- > [NVIDIA GPU Operator: Simplifying GPU Management in Kubernetes](#), NVIDIA Technical Blog
- > [CPU vs. GPU: What's the Difference?](#), Intel

AI 基础架构：考试权重 40%

检查、清理、转换和建模数据，以发现有用的信息、得出结论并支持决策。

2.1 理解使用数据挖掘、数据可视化和类似技术从大型数据集中提取见解的过程

2.2 使用统计性能指标 (例如损失函数或解释方差的比例) 比较模型

2.3 在高级团队成员的监督下进行数据分析

2.4 使用专用软件创建图形、图表或其他可视化效果，以传达数据分析结果

2.5 识别关系和趋势或任何可能影响研究结果的因素

培训推荐 (可选)

课程参考：[AI 基础架构和运营基础知识](#)

- > 第 4 单元：通过 GPU 加速 AI
- > 第 7.1 单元：数据中心平台
- > 第 10 单元：节能计算
- > 第 7.4 单元：通过 NVIDIA DPU 实现数据中心转型
- > 第 8 单元：面向 AI 的网络
- > 第 11 单元：参考架构
- > 第 12.4 单元：云端 AI 关键事项

阅读内容推荐

- > [Offloading and Isolating Data Center Workloads With NVIDIA Bluefield® DPU](#), NVIDIA Technical Blog
- > [NVIDIA DGX SuperPOD™ Reference Architecture](#), NVIDIA Docs Hub
- > [Power Constraints and AI Workloads: The Hidden Challenges of High-Density Data Centers](#), by Jeff Barber, NetZero News, LinkedIn
- > [High-Density Servers: Maximizing Efficiency and Performance in Data Centers](#), FS
- > [Introduction to the NVIDIA DGX™ H100 System](#), NVIDIA Docs Hub
- > [InfiniBand Key Features](#), NVIDIA Academy Vimeo
- > [Modernizing GPU Network Data Transfer With NVIDIA NVSwitch™](#), AMAX
- > [Accelerating IO in the Modern Data Center: Network IO](#), NVIDIA Technical Blog

AI 运营：考试权重 22%

- 3.1 说明 AI 数据中心管理和监控要点
 - 3.2 说明 AI 集群编排和作业调度要点
 - 3.3 阐明与监控 GPU 相关的关键措施和标准
 - 3.4 确定虚拟化加速基础设施的关键注意事项
-

培训推荐 (可选)

课程参考：[AI 基础架构和运营基础知识](#)

- > 第 5 单元：AI 软件生态系统
- > 第 8 单元：面向 AI 的网络
- > 第 13 单元：AI 数据中心管理和监控
- > 第 14 单元：编排、MLOps 和作业调度

阅读内容推荐

- > [Baseboard Management Controller](#), NVIDIA Docs Hub
- > [NVIDIA Base Command™](#), NVIDIA
- > [NVIDIA DCGM](#), NVIDIA Developer
- > [Out-of-Band Management Networks](#), Management Networks for Dell EMC Networking Configuration Guide, Dell Technologies
- > [Kubernetes Documentation](#), Kubernetes
- > [Let's Explore the Importance of Job Scheduling in a Cloud Environment](#), by Changju Lee, Samsung SDS
- > [What Is a Container?](#), Docker
- > [Slurm Workload Manager—Overview](#), SchedMD
- > [6 Reasons for Low GPU Utilization and How to Improve It](#), Run:ai
- > [NVIDIA Multi-Instance GPU \(MIG\)](#), NVIDIA

问题咨询，请发邮件至 dlichina@nvidia.com