MICROSOFT OFFICIAL CURRICULUM

Deploying and Optimizing AI Inference at Scale

NVIDIA Deep Learning Institute-aligned program for Deploying and Optimizing AI Inference at Scale.

NVIDIA NVIDIA

DURATION	LEVEL	FORMAT	CERTIFICATION
8 hours	Advanced	Virtual, On-site, or Hybrid	NVIDIA DLI Certificate of Competency

AUDIENCE PROFILE

Built for these roles

Built for AI Practitioner learners adopting Deploying and Optimizing AI Inference at Scale.

OVERVIEW

Executive overview

NVIDIA training program: Deploying and Optimizing AI Inference at Scale.

View official source

READINESS

Prerequisites

- Relevant foundational experience in the target technology area.
- Comfort with hands-on labs in a cloud or GPU-accelerated environment.

PROGRAM OUTCOMES

Capabilities your teams will gain

- Strengthen capability in deploying and optimizing ai inference at scale scenarios
- Strengthen capability in nvidia deep learning scenarios
- Strengthen capability in deep learning scenarios

ENTERPRISE CUSTOMIZATION

Enterprise customization

Tailor this program to your organization's priorities: Builds NVIDIA GPU-accelerated skills for deploying and optimizing ai inference at scale with hands-on DLI-style labs.

- Align labs to your production environment and platform priorities
- Add readiness reviews and instructor-led practice sessions
- Extend into project-specific architecture or delivery coaching

CURRICULUM

Curriculum roadmap

1 MODULE 1

Deploying and Optimizing AI Inference at Scale

Cover deploying and optimizing ai inference at scale skills and implementation practices aligned to Deploying and Optimizing AI Inference at Scale.

- Cover deploying
- optimizing ai inference at scale skills
- implementation practices aligned to Deploying
- Optimizing AI Inference at Scale

2

MODULE 2

NVIDIA Deep Learning

Cover nvidia deep learning skills and implementation practices aligned to Deploying and Optimizing AI Inference at Scale.

- Cover nvidia deep learning skills
- implementation practices aligned to Deploying
- Optimizing AI Inference at Scale

3

MODULE 3

Deep Learning

Cover deep learning skills and implementation practices aligned to Deploying and Optimizing AI Inference at Scale.

- Cover deep learning skills
- implementation practices aligned to Deploying
- Optimizing AI Inference at Scale

PRIVATE DELIVERY

Plan a private cohort for your team

Scope delivery format, role mix, and rollout timeline with our solutions team.

Plan Private Delivery

Book advisory call