Deploying a Model for Inference at Production Scale

This NVIDIA DLI course teaches teams how to deploy machine learning models on a GPU server using NVIDIA Triton Inference Server. It is especially useful for organizations that have moved beyond experimentation and need practical serving capability.

COURSE CODE
Program-aligned

DELIVERY
Virtual, On-site, or Hybrid

DURATION
4 hours

CERTIFICATION TRACK
Available on request

AUDIENCE PROFILE

Who This Program Is For

Built for practitioners who already train models and now need practical deployment and inference capability on GPU-based serving infrastructure.

PROGRAM SUMMARY

What This Course Covers

Official NVIDIA DLI program focused on deploying machine learning models to GPU servers with NVIDIA Triton Inference Server.

Tailored Delivery Available
This outline can be adapted for virtual, on-site, or hybrid delivery, with emphasis adjusted for your team's platform priorities, role mix, and implementation goals.

REGISTERED OFFICE
Wework, DLF Forum Infinity Tower C, Gurugram, INDIA

CONTACT
info@vnodeites.com
+91 9419 11 4792

DIRECT MESSAGING
+91 9419 11 4792
Serving enterprise clients across India and global markets

COMPLETE MODULE SEQUENCE

Module Flow and Topic Coverage

The structure below presents the current delivery flow for this program, including the associated topic areas covered under each module.

1 Build the foundation for production inference

Understand the core deployment patterns and operational considerations involved in moving trained models into production inference environments.

- Inference deployment foundations
- GPU-backed deployment workflows

2 Serve and manage models with Triton

Use NVIDIA Triton to expose models for inference while improving deployment readiness and scalability for AI applications.

- Serving models with Triton
- Production inference considerations

REGISTERED OFFICE
Wework, DLF Forum Infinity Tower C, Gurugram, INDIA

CONTACT
info@vnodeites.com
+91 9419 11 4792

DIRECT MESSAGING
+91 9419 11 4792
Serving enterprise clients across India and global markets