
Google Generative AI in Production

Duration: 1 Day **Course Code: GO8338**

Overview:

Traditional MLOps is a set of practices to productionize traditional ML systems for enterprise applications. Generative AI raises new challenges in managing and productionizing applications at scale. The field of generative AI operations seeks to address these new challenges.

In this course, you learn about the challenges that arise when deploying and productionizing generative AI-powered applications. You learn how to secure your generative AI-powered applications. Finally, you will discuss best practices for logging and monitoring your generative AI-powered applications in production.

Products:

- Vertex AI

Updated August 2026

Target Audience:

Developers, DevOps engineers and machine learning engineers who wish to operationalize GenAI-based applications

Objectives:

- | | |
|--|--|
| ■ After this course participants should be able to: | ■ Maintain and update LLM models |
| ■ Understand generative AI operations | ■ Test and evaluate gen AI-powered apps |
| ■ Compare traditional MLOps and GenAI Ops | ■ Deploy CI/CD pipelines for gen AI-powered apps |
| ■ Analyze the components of an LLM system | ■ Identify security challenges for gen AI applications |
| ■ Define and compare RAG and ReAct | ■ Understand prompt security issues |
| ■ Evaluate application deployment options | ■ Apply sensitive data protection and DLP API |
| ■ Deploy, package, and version apps | ■ Implement Model Armor |
-

Prerequisites:

Completion of the "Application Development with LLMs on Google Cloud" or equivalent knowledge.

Testing and Certification

None

Follow-on-Courses:

None

Content:

Module 01: Introduction to Generative AI in Production

- Generative AI operations
- Traditional MLOps vs. GenAIOps
- Components of an LLM system
- RAG/ReAct architecture

Module 02: Generative AI Application Deployment

- Application deployment options
- Deployment, packaging, and versioning
- Lab: Deploying an Agentic Application on Cloud Run

Module 03: Productionizing Generative AI

- Maintenance and updates
- Testing and evaluation
- CI/CD pipelines for gen AI-powered apps
- Lab: Tracking Versions of Generative AI Applications

Module 04: Securing Generative AI Applications

- Security challenges
- Prompt security
- Sensitive Data Protection and DLP API
- Model Armor
- Lab: Securing Generative AI-Powered Applications

Further Information:

For More information, or to book your course, please call us on 0800/84.009

info@globalknowledge.be

www.globalknowledge.com/en-be/