Google Cloud

Vertex AI and Generative AI Security

This course is designed to empower your organization to fully harness the transformative potential of Google's Vertex AI and generative AI (gen AI) technologies, with a strong emphasis on security. Tailored for AI practitioners and security engineers, it provides targeted knowledge and hands-on skills to navigate and adopt AI safely and effectively. Participants will gain practical insights and develop a security-conscious approach, ensuring a secure and responsible integration of gen AI within their organization.

DURATION
2 days

LEVEL
Intermediate

FORMAT
Instructor-led

What you'll learn

- Establish foundational knowledge of Vertex AI and its security challenges.
- Implement identity and access control measures to restrict access to Vertex AI resources.
- Configure encryption strategies and protect sensitive information.
- Enable logging, monitoring, and alerting for real-time security oversight of Vertex AI operations.
- Identify and mitigate unique security threats associated with generative AI.
- Apply testing techniques to validate and secure generative AI model responses.
- Implement best practices for securing data sources and responses within Retrieval-Augmented Generation (RAG) systems.
- Establish foundational knowledge of AI Safety.

Overview	8 modules · 10 labs
Who this course is for	AI practitioners, security professionals, and cloud architects
Products	• Vertex AI • Gemini • Cloud IAM • Cloud VPC • Cloud KMS • Cloud Operations • Sensitive Data Protection
Prerequisites	Fundamental knowledge of machine learning, in particular generative AI, and basic understanding of security on Google Cloud.

Module 01 Introduction to Vertex AI Security Principles

Topics	• Google Cloud Security • Vertex AI components • Vertex AI Security concerns
Objectives	• Review Google Cloud Security fundamentals. • Establish a foundational understanding of Vertex AI. • Enumerate the security concerns related to Vertex AI features and components.
Activities	Lab: Vertex AI: Training and Serving a Custom Model

Module 02 Identity and Access Management (IAM) in Vertex AI

Topics	Overview of IAM in Google Cloud
Objectives	• Control access with Identity Access Management. • Simplify permission using organization hierarchies and policies. • Use service accounts for least privileged access.
Activities	Lab: Service Accounts and Roles: Fundamentals

Module 03 Data Security and Privacy

Topics
- Data encryption
- Protecting Sensitive Data
- VPC Service Controls
- Disaster recovery planning

Objectives
- Configure encryption at rest and in-transit.
- Encrypt data using customer-managed encryption keys.
- Protect sensitive data using the Data Loss Prevention service.
- Prevent exfiltration of data using VPC Service Controls.
- Architect systems with disaster recovery in mind.

Activities
- Lab: Getting Started with Cloud KMS
- Lab: Creating a De-identified Copy of Data in Cloud Storage

---

Module 04 Securing Vertex AI Endpoints and model deployment

Topics
- Network security
- Securing model endpoints

Objectives
- Deploy ML models using model endpoints.
- Secure model endpoints.

Activities
Lab: Configuring Private Google Access and Cloud NAT

Module 05 Monitoring and logging in Vertex AI

Topics
- Logging
- Monitoring

Objectives
- Write to and analyze logs.
- Set up monitoring and alerting.

Module 06 Security risks in generative AI applications

Topics
- Overview of gen AI security risks
- Overview of AI Safety
- Prompt security
- LLM safeguards

Objectives

- Identify security risks specific to LLMs and gen AI applications.
- Understand methods for mitigating prompt hacking and injection attacks.
- Explore the fundamentals of securing generative AI models and applications.
- Introduce fundamentals of AI Safety.

Activities

- Lab: Safeguarding with Vertex AI Gemini API
- Lab: Gen AI & LLM Security for Developers

Module 07 Testing and evaluating generative AI model responses

Topics

- Testing generative AI model responses.
- Evaluating model responses.
- Fine-Tuning LLMs.

Objectives

- Implement best practices for testing model responses.
- Apply techniques for improving response security in gen AI applications.

Activities

- Lab: Measure Gen AI Performance with the Generative AI Evaluation Service
- Lab: Unit Testing Generative AI Applications

Module 08 Securing Retrieval-Augmented Generation (RAG) systems

Topics

- Fundamentals of Retrieval-Augmented Generation
- Security in RAG systems

Objectives

- Understand RAG architecture and security implications.
- Implement best practices for grounding and securing data sources in RAG systems.

Activities

- Lab: Multimodal Retrieval Augmented Generation (RAG) Using the Vertex AI Gemini API
- Lab: Introduction to Function Calling with Gemini