Course description

Learn how to secure Generative AI applications using Google Cloud Model Armor. Discover how Model Armor protects Large Language Models (LLMs) against prompt injection, jailbreaking, malicious URLs, sensitive data exposure, and unsafe outputs while enabling organizations to implement responsible and secure AI deployments at scale.

Participants will learn how Model Armor mitigates common LLM security threats, configure security guardrails, customize protection templates, enable detection mechanisms, review audit logs, and integrate Model Armor APIs into AI applications. Through demonstrations, labs, and practical examples, learners gain the knowledge required to build secure and compliant AI solutions.

- Course level: Introductory
- Duration: 1 day

Activities

This course includes instructor-led training, security demonstrations, hands-on lab, api configuration exercises, scenario-based learning, threat analysis activities, knowledge checks and quizzes, and security best practices reviews to reinforce learning.

Course objectives

In this course, you will learn to:

- Understand the purpose and architecture of Model Armor.
- Identify common LLM security threats and vulnerabilities.
- Map Model Armor features to AI security risks.
- Configure floor settings and security guardrails.
- Create and customize Model Armor templates.
- Enable and configure detection mechanisms.
- Set up and use the Model Armor API.
- Analyze audit logs and flagged violations.
- Secure AI prompts and responses using Model Armor controls.

Intended audience

This course is intended for:

- Security Engineers
- AI/ML Developers
- Cloud Architects
- Security Architects
- Platform Engineers
- AI Security Professionals
- Risk and Compliance Teams
- Cloud Security Practitioners

Prerequisites

We recommend that attendees of this course have:

- Working knowledge of APIs.
- Working knowledge of Google Cloud CLI.
- Understanding of cloud security foundational principles.
- Familiarity with the Google Cloud Console.

Products covered

- Model Armor
- Google Cloud Security Services

- Security Command Center (SCC)
- Large Language Models (LLMs)

Course outline

Module 01 - Course Overview

- What's in it for Me?
- Course Learning Objectives

Module 02 - Model Armor Overview

- About Model Armor
- LLM Security Risks
- Model Armor Architecture
- OWASP LLM Vulnerabilities

Module 03 - Customize Model Armor

- Model Armor Customization
- Floor Settings
- Guardrails and Confidence Levels
- Templates
- Detection Configuration

Module 04 - Use Model Armor

- Model Armor API Setup
- API Prerequisites
- Flagged Violations
- Audit Logging
- Security Command Center Integration

Module 05 - Put It All Together

- Prompts and Responses
- Application Code Integration
- Prompt Sanitization
- Response Filtering

Module 06 - Course Conclusion

- What Did I Learn?
- Course Summary

Certification details

Participants receive a course completion certificate upon successfully completing the training program.

FAQs

Who should attend this course?

This course is designed for security engineers, AI/ML developers, cloud architects, and professionals responsible for securing AI applications.

What is Model Armor?

Model Armor is a Google Cloud security solution that protects AI applications and LLMs from threats such as prompt injection, jailbreaking, malicious URLs, sensitive data leaks, and unsafe outputs.

Is this a hands-on course?

Yes. The course includes hands-on labs, demonstrations, API setup exercises, and security configuration activities.

What security threats are covered?

The course covers prompt injection, jailbreaking, malicious URLs, sensitive data exposure, improper output handling, and selected OWASP LLM vulnerabilities.

Will I learn how to configure Model Armor?

Yes. Participants learn how to configure floor settings, templates, guardrails, detections, logging, and API integrations.

Does the course cover audit logging?

Yes. Learners explore audit logs, flagged violations, Security Command Center integration, and violation management workflows.

How long is the course?

The course is delivered in a 3-hour instructor-led format.

What business benefits can organizations expect?

Organizations can strengthen AI security, reduce risk exposure, improve compliance, protect sensitive information, and establish secure AI deployment practices using Model Armor.