Securing Machine Learning Models Training Course

Professional Development Training Course Outline

Category	Data Security	**Duration**	5 Days
Base Fee	$1,500 USD	**Delivery Mode**	Online / On-site Hybrid

Course Introduction

The rapid AI adoption across critical sectors has made Machine Learning security a paramount concern, fundamentally shifting the cybersecurity threat landscape. Securing Machine Learning Models Training Course addresses the critical need to harden ML systems against novel and sophisticated attacks that target the entire ML lifecycle, from data poisoning in the training phase to adversarial attacks during inference. Professionals must master specialized techniques like robust model training, data integrity validation, and threat modeling to ensure the confidentiality, integrity, and availability of production AI systems. Our program provides hands-on labs and real-world case studies focusing on MLOps security, empowering security engineers and data scientists to build secure-by-design AI solutions and achieve AI trust and compliance.

This intensive training is crucial for those tasked with safeguarding proprietary models and sensitive training data. It dives deep into Adversarial Machine Learning, teaching defense mechanisms against model evasion, poisoning, and extraction attacks. Graduates will be equipped with the trending keywords and advanced skills necessary to implement a Zero Trust security posture for their AI infrastructure, covering everything from secure data pipelines and federated learning to Explainable AI for forensic analysis. Secure your organization's most valuable AI assets by mastering the next generation of cyber defense for intelligent systems.

Programme Curriculum

Securing Machine Learning Models Training Course

Introduction

The rapid AI adoption across critical sectors has made Machine Learning security a paramount concern, fundamentally shifting the cybersecurity threat landscape. Securing Machine Learning Models Training Course addresses the critical need to harden ML systems against novel and sophisticated attacks that target the entire

ML lifecycle, from data poisoning in the training phase to adversarial attacks during inference. Professionals must master specialized techniques like robust model training, data integrity validation, and threat modeling to ensure the confidentiality, integrity, and availability of production AI systems. Our program provides hands-on labs and real-world case studies focusing on MLOps security, empowering security engineers and data scientists to build secure-by-design AI solutions and achieve AI trust and compliance.

This intensive training is crucial for those tasked with safeguarding proprietary models and sensitive training data. It dives deep into Adversarial Machine Learning, teaching defense mechanisms against model evasion, poisoning, and extraction attacks. Graduates will be equipped with the trending keywords and advanced skills necessary to implement a Zero Trust security posture for their AI infrastructure, covering everything from secure data pipelines and federated learning to Explainable AI for forensic analysis. Secure your organization's most valuable AI assets by mastering the next generation of cyber defense for intelligent systems.

Course Duration

5 days

Course Objectives

Upon completion, participants will be able to:

1. Threat Model the complete MLSecOps lifecycle to identify unique AI vulnerabilities.
2. Design and implement secure data pipelines to prevent data poisoning and data leakage.
3. Execute and defend against adversarial evasion attacks on deep neural networks.
4. Apply robust model training techniques like Adversarial Training to improve model resilience.
5. Analyze and mitigate model inversion and membership inference attacks against sensitive data.
6. Establish ML model integrity and authenticity using cryptographic signing and blockchain methods.
7. Integrate Explainable AI (XAI) tools for post-attack forensic analysis and model debugging.
8. Implement federated learning security protocols to protect distributed training data privacy.
9. Secure cloud-native ML deployments leveraging Kubernetes and Confidential Computing.
10. Develop AI governance and risk management frameworks that comply with emerging AI regulations.
11. Implement input validation and output sanitization specifically for ML inference endpoints.
12. Use the MITRE ATLAS framework to map and categorize Adversarial Machine Learning tactics.
13. Apply Zero Trust principles to the entire AI/ML infrastructure, from data storage to API access.

Target Audience

1. Security Engineers and Cybersecurity Architects
2. Data Scientists and Machine Learning Engineers
3. MLOps Engineers and AI Infrastructure Teams
4. Application Security Professionals
5. Chief Information Security Officers and Security Leadership
6. IT/DevOps Professionals migrating to MLOps
7. Ethical Hackers and Penetration Testers specializing in AI
8. Risk Management and AI Governance Specialists

Course Modules

Module 1: Introduction to ML Threat Modeling and Attack Vectors

- Machine Learning (ML) Lifecycle overview and its unique Attack Surface.
- The CIA Triad in the context of ML models and data.

- ML Threat Modeling.
- Introduction to the MITRE ATLAS Framework for classifying AI threats.
- Case Study: Analyzing the Tay Chatbot incident to understand data integrity and availability attacks.

Module 2: Securing the ML Data Pipeline

- Mitigating Data Poisoning Attacks during training.
- Techniques for Data Integrity validation and anomaly detection in training sets.
- Preventing Training Data Leakage and addressing Data Privacy concerns.
- Implementing Secure Data Storage and access control for large datasets.
- Case Study: The Google reCAPTCHA poisoning attempt and how data provenance could have mitigated it.

Module 3: Adversarial Evasion Attacks (Inference Security)

- Understanding Adversarial Examples and their generation techniques
- Black-Box and White-Box attack scenarios and their practical implementation.
- Defenses against evasion.
- Gradient Masking and other defensive distillation techniques.
- Case Study: Tesla Autopilot.

Module 4: Model Integrity and Intellectual Property Theft

- Model Extraction/Stealing Attacks.
- Defenses against model extraction, including rate-limiting and query analysis.
- Model Inversion Attacks to reconstruct training data samples from model output.
- Protecting Model Intellectual Property and utilizing model watermarking.
- Case Study: A public cloud provider's API abuse led to a Model Extraction event, revealing the need for robust API security and access controls.

Module 5: Secure ML Deployment and Cloud-Native MLOps

- Container Security for ML model serving endpoints.
- Implementing Confidential Computing for secure inference environments.
- Secure API Design for ML inference services and proper Input Validation.
- Model version control and rollback strategies for quick recovery from compromise.
- Case Study: Vulnerabilities in a major bank?ÇÖs Kubernetes deployment exposed an ML model API, highlighting container hardening necessity.

Module 6: Advanced Privacy-Preserving ML Techniques

- Introduction to Differential Privacy and its application to training data.
- Implementing Federated Learning architectures for decentralized, private training.
- Mitigating Membership Inference Attacks using privacy-enhancing technologies.
- Securing Large Language Models against prompt injection and sensitive data disclosure.
- Case Study: Apple's Differential Privacy implementation for user data analytics and its trade-offs with data utility.

Module 7: AI Governance, Risk, and Compliance (GRC)

- Establishing an AI Risk Management Framework.
- Role of Explainable AI in security forensics, auditing, and compliance.

- Developing an AI Incident Response Plan.
- Regulatory landscape compliance for secure AI.
- Case Study: A healthcare AI system was successfully audited due to its XAI capabilities, demonstrating transparency required by patient data regulations.

Module 8: Building a Secure ML Ecosystem (MLSecOps)

- Integrating security checks and testing into CI/CD/CT pipelines.
- Implementing security-as-code for infrastructure and model configuration.
- Automated Vulnerability Scanning and drift detection for production models.
- The concept of AI Zero Trust never trusts, always verify for ML components.
- Case Study: Implementation of an end-to-end MLSecOps pipeline at a major tech firm, reducing deployment-to-vulnerability-fix time by 80%.

Training Methodology

This course employs a participatory and hands-on approach to ensure practical learning, including:

- Interactive lectures and presentations.
- Group discussions and brainstorming sessions.
- Hands-on exercises using real-world datasets.
- Role-playing and scenario-based simulations.
- Analysis of case studies to bridge theory and practice.
- Peer-to-peer learning and networking.
- Expert-led Q&A sessions.
- Continuous feedback and personalized guidance.

Register as a group from 3 participants for a Discount

Send us an email: info@fineskilltrainingcenter.org or call +254769199797

Certification

Upon successful completion of this training, participants will be issued with a globally- recognized certificate.

Tailor-Made Course

We also offer tailor-made courses based on your needs.

Key Notes

a. The participant must be conversant with English.

b. Upon completion of training the participant will be issued with an Authorized Training Certificate

c. Course duration is flexible and the contents can be modified to fit any number of days.

d. The course fee includes facilitation training materials, 2 coffee breaks, buffet lunch and A Certificate upon successful completion of Training.

e. One-year post-training support Consultation and Coaching provided after the course.

f. Payment should be done at least a week before commence of the training, to Fineskill Training Center account, as indicated in the invoice so as to enable us prepare better for you.

Upcoming Scheduled Sessions

Start Date	End Date	Location	Price
23 Nov 2026	27 Nov 2026	Online	$1,000
23 Nov 2026	27 Nov 2026	Physical	$1,500
23 Nov 2026	27 Nov 2026	Physical	$3,000
23 Nov 2026	27 Nov 2026	Physical	$3,000
23 Nov 2026	27 Nov 2026	Physical	$6,000
23 Nov 2026	27 Nov 2026	Physical	$4,500
23 Nov 2026	27 Nov 2026	Physical	$6,000
23 Nov 2026	27 Nov 2026	Physical	$6,000

Ready to enroll or request a customized program? Book online or contact us:

Register Online Now

Email: info@fineskilltrainingcenter.com | Website: www.fineskilltrainingcenter.com