AI Security Deep Dive (TTAI2800) Eğitimi

Süre: **3 Gün**

BİLGİNÇ IT Academy
info@bilginc.com
bilginc.com

Yapay zekâ ve makine öğrenmesi sistemlerine yönelik gelişmiş saldırıları analiz edin, AI güvenlik açıklarını keşfedin ve güvenli AI uygulamaları geliştirmeyi öğrenin.

AI Security Deep Dive eğitimi, yapay zekâ ve makine öğrenmesi sistemlerinde ortaya çıkan modern güvenlik tehditlerini anlamak ve bu tehditlere karşı etkili savunma stratejileri geliştirmek isteyen profesyoneller için hazırlanmış kapsamlı ve uygulamalı bir eğitim programıdır.

Kuruluşlar giderek daha fazla **AI**, **Machine Learning**, **LLM sistemleri** ve AI destekli uygulamalar kullanmaya başladıkça; geleneksel güvenlik yaklaşımlarının yetersiz kaldığı yeni saldırı yüzeyleri ortaya çıkmaktadır. Bu eğitim, AI sistemlerine özel güvenlik risklerini anlamak ve gerçek dünya ortamlarında güvenli AI entegrasyonu sağlamak için gerekli teknik bilgi ve uygulamalı deneyimi sunmaktadır.

Program boyunca katılımcılar; **data poisoning**, **adversarial attacks**, **model inversion**, **membership inference attacks**, **prompt injection**, **LLM security**, **secure AI APIs**, **privacy protection**, **AI governance** ve enterprise AI güvenliği gibi kritik konuları uygulamalı laboratuvarlar ile deneyimleyecektir.

Eğitimin ilk iki günü interaktif**Jupyter Notebook** laboratuvarları üzerinden ilerlerken, üçüncü gün gerçek dünya AI entegrasyonlarına ve güvenli AI mimarilerine odaklanmaktadır.

Katılımcılar ayrıca **Python/Flask**, **Java/Spring**, **ASP.Net** ve **Node.js** ortamlarında güvenli AI entegrasyonu geliştirmeyi öğrenecektir.

Bu yoğun eğitim programı; AI güvenliği, offensive AI testing, secure AI deployment ve enterprise AI risk yönetimi alanlarında uzmanlaşmak isteyen profesyoneller için güçlü bir teknik temel sunmaktadır.

Teknik Yetkinlikler

Katılımcılar eğitim boyunca aşağıdaki teknolojiler ve araçlarla çalışacaktır:

- **Python**
- **Jupyter Notebook**
- **Flask**
- **Java/Spring**
- **ASP.Net**
- **Node.js**
- **Hugging Face Inference API**
- **LLM Security Frameworkleri**
- **OWASP AI Security**
- **Differential Privacy**

Önkoşullar

Katılımcıların aşağıdaki konularda temel bilgi sahibi olması önerilir:

- Temel programlama bilgisi

AI Security Deep Dive (TTAI2800) Eğitimi

Süre: **3 Gün**

BİLGİNÇ IT Academy
info@bilginc.com
bilginc.com

- Kod okuyabilme ve küçük kod değişiklikleri yapabilme
- Siber güvenlik farkındalığı
- HTTP ve web uygulamaları hakkında temel bilgi
- Veri yapıları ve temel istatistik bilgisi
- Yazılım geliştirme yaşam döngüsü (SDLC)

Kimler Katılmalı

Bu eğitim aşağıdaki profesyoneller için idealdir:

- **Programmers & Developers**
- **Security Analysts**
- **Cybersecurity Professionals**
- **Technical Leads**
- **DevSecOps Engineers**
- **Security Architects**
- **Incident Response Teams**
- **AI Security Engineers**
- AI destekli uygulamalar geliştiren yazılım ekipleri

Bu program özellikle **AI güvenliği**, **LLM security**, **AI risk management** ve **secure AI development** alanlarında uzmanlaşmak isteyen teknik ekipler için uygundur.

Neler Öğreneceksiniz

Eğitim sonunda katılımcılar şunları öğrenir:

- **AI/ML security fundamentals**
- Yapay zekâ sistemlerine yönelik saldırı vektörlerini analiz etmek
- **Data poisoning** saldırıları gerçekleştirmek ve önlemek
- **Adversarial attacks** oluşturmak ve savunma geliştirmek
- **Membership inference** ve **model inversion attacks** uygulamak
- AI tabanlı uygulamalarda güvenli API entegrasyonu sağlamak
- **Prompt injection saldırılarını önlemek**
- Differential privacy tekniklerini uygulamak
- Enterprise AI security governance süreçleri oluşturmak
- AI incident response stratejileri geliştirmek

Eğitim İçeriği

1. AI/ML Security Foundations

- Yapay zekâ ve makine öğrenmesine güvenlik perspektifiyle yaklaşım
- **OWASP Top 10** güvenlik açıkları

BİLGİNÇ IT Academy

info@bilginc.com

bilginc.com

- AI sistem mimarileri
- Training vs inference süreçleri
- AI attack surface analizi
- AI sistemlerinde geleneksel güvenlik yaklaşımlarının eksikleri

Hands-On Lab

- AI security testing environment kurulumu
- Vulnerable ML model analizi

2. Data Poisoning and Training Attacks

- **Targeted vs untargeted attacks**
- Clean-label attacks
- Training data güvenlik açıkları
- AI supply chain security
- Backdoor attacks
- Compromised pre-trained model riskleri

Hands-On Lab

- Image classifier poisoning attacks
- Text model poisoning
- Data validation pipeline geliştirme
- Poisoning detection systems

Day 2: Adversarial Attacks and Model Security

3. Adversarial Examples and Model Manipulation

- **FGSM ve PGD saldırıları**
- Black-box vs white-box attacks
- Model evasion techniques
- Adversarial patches
- Transferability attacks

Hands-On Lab

- Adversarial example generation
- Production AI robustness testing

4. Privacy Attacks and Information Extraction

- **Membership inference attacks**
- **Model inversion attacks**
- Property inference
- Model stealing techniques
- Sensitive data extraction

Hands-On Lab

BİLGİNÇ IT Academy

info@bilginc.com

bilginc.com

- Membership inference uygulamaları
- Model inversion techniques
- Differential privacy implementations

Day 3: Secure AI Integration and Enterprise Defense

5. Securing AI-Integrated Applications

- Secure AI API integration
- Authentication & rate limiting
- **Prompt injection mitigation**
- Input/output validation
- Secure AI microservices
- Containerization ve monitoring

Hands-On Lab

- Hugging Face Inference API integration
- Python/Flask, Java/Spring, ASP.Net, Node.js entegrasyonları
- Secure AI-powered web applications

6. Enterprise AI Security Strategy

- AI governance frameworks
- AI risk assessment
- Continuous monitoring
- AI red teaming
- Incident response for AI breaches
- AI security dashboards

Hands-On Lab

- AI monitoring system kurulumu
- AI security assessment çalışmaları
- Remediation planning

7. Advanced Topics and Emerging Threats

- **LLM jailbreaking**
- Instruction-following exploits
- Multi-modal AI attacks
- AI supply chain security
- GDPR ve AI compliance
- Algorithmic auditing