Certified AI Security Professional Eğitimi

Süre: **4 Gün**

BİLGİNÇ IT Academy
info@bilginc.com
bilginc.com

Bu Eğitim Size Ne Kazandırır?

- **Yapay zekâ sistemlerini geliştirme, dağıtım ve işletim süreçleri boyunca** karşılaşabilecekleri siber güvenlik tehditlerine karşı korumayı öğrenebilirsiniz.
- **Makine öğrenimi, derin öğrenme, büyük dil modelleri (LLM) ve AI ajanlarının** mimarilerini ve güvenlik gereksinimlerini anlayabilirsiniz.
- **Prompt Injection, Jailbreaking, Data Poisoning ve Model Extraction** gibi yapay zekâya özgü saldırı tekniklerini tanıyabilirsiniz.
- **MITRE ATLAS** ve **OWASP LLM Top 10** çerçevelerini kullanarak AI tehditlerini ve saldırı yollarını değerlendirebilirsiniz.
- **AI Red Teaming, güvenlik testleri, izleme ve olay müdahalesi** süreçlerinin nasıl yürütüldüğünü öğrenebilirsiniz.
- **Güvenli LLM uygulamaları, RAG mimarileri, AI ajan izinleri ve tedarik zinciri güvenliği** için koruma yaklaşımlarını keşfedebilirsiniz.
- **PECB Certified Artificial Intelligence Security Professional** sertifikasyon sınavına yönelik bilgi ve hazırlık kazanabilirsiniz.

Certified AI Security Professional (CAISP) Eğitimi, yapay zekâ sistemlerinin yaşam döngüsü boyunca güvenliğini sağlamak isteyen **siber güvenlik ve yapay zekâ profesyonelleri** için tasarlanmış **4 günlük bir eğitim programıdır**.

Yapay zekâ çözümleri, geleneksel yazılım sistemlerinden farklı güvenlik riskleri barındırır. **Model eğitimi, veri işleme hatları, çıkarım (Inference) servisleri, büyük dil modelleri, harici araç bağlantıları ve otonom ajanlar**, korunması gereken yeni varlıklar ve saldırı yüzeyleri oluşturur.

Bu eğitim, **makine öğrenimi (Machine Learning), derin öğrenme (Deep Learning), büyük dil modelleri (LLM) ve AI ajanlarının** nasıl tasarlandığını, geliştirildiğini, dağıtıldığını, saldırıya uğrayabildiğini ve savunulabildiğini ele alır.

Katılımcılar, yapay zekâ sistemlerine yönelik **tehdit aktörlerini, saldırı hedeflerini, güven sınırlarını ve güvenlik kontrollerini** inceler. **MITRE ATLAS** ve **OWASP LLM Top 10** gibi çerçeveler aracılığıyla AI tehditlerinin nasıl sınıflandırılabileceğini ve değerlendirilebileceğini öğrenir.

Programın teknik bölümlerinde **Prompt Injection, Jailbreaking, veri zehirleme, model çıkarma, hassas bilgi ifşası, güvenli olmayan çıktı işleme ve AI araçlarının kötüye kullanılması** gibi saldırı teknikleri açıklanır.

Savunma odaklı bölümlerde ise **güvenli AI mimarisi, derinlemesine savunma (Defence-in-Depth), erişim kontrolleri, Guardrails, güvenli RAG, ajan yetkilendirmesi ve insan denetimi** üzerinde durulur.

Eğitim ayrıca **AI tedarik zinciri güvenliği, güvenlik testleri, AI Red Teaming, güvenlik izleme, SIEM entegrasyonu, olay müdahalesi ve AI destekli siber savunma** konularını kapsar.

Certified AI Security Professional Eğitimi

Süre: **4 Gün**

BİLGİNÇ IT Academy
info@bilginc.com
bilginc.com

Son bölümlerde katılımcılar, **AI güvenlik yönetişimi, risk değerlendirmesi ve kurumsal AI güvenlik programı oluşturma** yaklaşımlarını öğrenir.

Uygulamalı çalışmalar, laboratuvarlar, bilgi kontrolleri ve kısa değerlendirmeler, teknik konuların pekiştirilmesini ve katılımcıların **PECB Certified Artificial Intelligence Security Professional** sınavına hazırlanmasını destekler.

Sınav ve Değerlendirme

Eğitimin ardından katılımcılar, **PECB tarafından gerçekleştirilen resmî Certified Artificial Intelligence Security Professional sertifikasyon sınavına** girer.

Önkoşullar

Bu eğitime katılacak kişilerin aşağıdaki bilgi ve deneyimlere sahip olması beklenir:

- **Temel siber güvenlik ve bilgi güvenliği kavramları** hakkında sağlam bir bilgi birikimi
- **Tehdit modelleme, kimlik ve erişim kontrolü, olay müdahalesi, güvenlik izleme ve genel güvenlik mimarisi** konularında bilgi veya deneyim sahibi olunması önerilir
- **Yapay zekâ ve makine öğreniminin temel kavramları**, model eğitimi, veri işleme hatları ve **API tabanlı AI servisleri** hakkında temel anlayış

Certified AI Security Professional Eğitimi

Süre: **4 Gün**

BİLGİNÇ IT Academy
info@bilginc.com
bilginc.com

Kimler Katılmalı

Bu eğitim, yapay zekâ çözümlerinin güvenliğinden sorumlu olan veya**AI güvenliği alanında uzmanlaşmak** isteyen profesyonellere yöneliktir.

- **Siber güvenlik ve bilgi güvenliği uzmanları**
- **Yapay zekâ ve makine öğrenimi mühendisleri**
- **AI güvenlik mühendisleri**
- **Güvenlik mimarları ve danışmanları**
- **SOC analistleri ve olay müdahale uzmanları**
- **Tehdit istihbaratı profesyonelleri**
- **Veri bilimcileri ve AI geliştiricileri**
- **AI risk ve uyumluluk uzmanları**
- AI güvenliğinden sorumlu **teknoloji yöneticileri ve teknik ekip liderleri**
- **Yapay zekâ destekli uygulama ve hizmetlerin güvenliğini** sağlamaktan sorumlu profesyoneller

Certified AI Security Professional Eğitimi

Süre: **4 Gün**

BİLGİNÇ IT Academy
info@bilginc.com
bilginc.com

Neler Öğreneceksiniz

Bu eğitimi tamamladığınızda:

- **Makine öğrenimi, derin öğrenme, LLM ve AI ajanlarının** mimarilerini ve çalışma prensiplerini açıklayabileceksiniz.
- **AI sistemlerinin geliştirme, eğitim, dağıtım ve işletim aşamalarındaki** güvenlik risklerini belirleyebileceksiniz.
- **AI saldırı yüzeylerini, tehdit aktörlerini ve güven sınırlarını** analiz edebileceksiniz.
- **MITRE ATLAS** ve **OWASP LLM Top 10** çerçeveleriyle AI tehditlerini değerlendirebileceksiniz.
- **Prompt Injection, Jailbreaking, Data Poisoning, Model Extraction ve Model Denial of Service** gibi saldırı tekniklerini tanıyabileceksiniz.
- **AI Red Teaming** ve güvenlik testi yaklaşımlarının amaçlarını ve kullanım alanlarını açıklayabileceksiniz.
- **AI modelleri, veriler, uygulamalar ve altyapı** için uygun güvenlik kontrollerini belirleyebileceksiniz.
- **Güvenli RAG, Guardrails, bağlam izolasyonu ve vektör veritabanı güvenliği** yaklaşımlarını değerlendirebileceksiniz.
- **AI ajanlarının araç erişimlerini ve yetkilerini** en az ayrıcalık ilkesiyle ele alabileceksiniz.
- **Üçüncü taraf modeller, açık kaynak bileşenler ve eğitim veri kümeleri** üzerinden oluşabilecek tedarik zinciri risklerini tanıyabileceksiniz.
- **AI güvenlik izleme, telemetri, loglama, SIEM entegrasyonu ve tespit mühendisliği** yaklaşımlarını açıklayabileceksiniz.
- **AI güvenlik olaylarının tespit, araştırma, sınırlandırma ve kurtarma** süreçlerini değerlendirebileceksiniz.
- **Tehdit istihbaratı, tehdit avcılığı, zararlı yazılım analizi ve tersine mühendislikte** AI kullanımının fırsatlarını ve sınırlarını anlayabileceksiniz.
- **AI güvenlik yönetişimi, risk değerlendirmesi ve güvenlik politikalarının** kurumsal güvenlik programındaki rolünü açıklayabileceksiniz.
- **Sürekli güvenlik güvencesi ve tedarik zinciri yönetişimini** kapsayan bir AI güvenlik programının temel bileşenlerini tanımlayabileceksiniz.

Eğitim İçeriği

1. Yapay Zekâ Güvenliğinin Temelleri

Certified AI Security Professional Eğitimi

Süre: **4 Gün**

BİLGİNÇ IT Academy
info@bilginc.com
bilginc.com

Bu bölümde yapay zekâ sistemlerinin teknik yapısı ve güvenlik değerlendirmeleri için gerekli temel kavramlar incelenir.

- **Yapay zekâ kavramları ve terminolojisi**
- **Makine öğreniminin temelleri**
- **Denetimli ve denetimsiz öğrenme**
- **Derin öğrenme ve sinir ağları**
- **AI model geliştirme ve eğitimi**
- **Eğitim, doğrulama ve test verileri**
- **Veri işleme hatları ve model yaşam döngüsü**
- **AI çıkarımı (Inference) ve dağıtım**
- **Yapay zekâ teknoloji yığınının (AI Technology Stack)** anlaşılması

2. Yapay Zekâ Sistemleri için Güvenlik

Bu bölüm, AI güvenliğinin geleneksel uygulama güvenliğinden ayrılan yönlerine ve sistemlerin korunması gereken bileşenlerine odaklanır.

- **AI güvenliği ile geleneksel uygulama güvenliği** arasındaki farklar
- **AI saldırı yüzeyleri (Attack Surfaces)**
- **Tehdit aktörleri ve saldırgan hedefleri**
- Korunması gereken **AI varlıkları**
- AI ortamlarında **gizlilik, bütünlük ve erişilebilirlik (CIA)**
- **Model ve veri bütünlüğü**
- **AI sistemlerinin güven sınırları (Trust Boundaries)**
- **AI yaşam döngüsü boyunca güvenlik**
- **Yapay zekâ sistemlerinde tehdit modelleme**

3. AI Mimarisi ve Güvenlik

Bu bölümde AI çözümlerinin altyapı, platform ve entegrasyon bileşenlerinin güvenlik üzerindeki etkileri ele alınır.

- **AI altyapısı ve platformları**
- **Model barındırma (Model Hosting) ve çıkarım servisleri**

BİLGİNÇ IT Academy

info@bilginc.com

bilginc.com

- **API'ler ve uygulama entegrasyonu**
- **Bulut tabanlı AI hizmetleri**
- **Üçüncü taraf AI bağımlılıkları**
- **AI tedarik zinciri** ile ilgili güvenlik değerlendirmeleri
- **Model kökeni ve geçmişinin (Model Provenance)** güvenlik üzerindeki etkileri

4. Büyük Dil Modeli (LLM) Mimarisi

Bu bölümde büyük dil modellerinin çalışma prensipleri ve LLM tabanlı uygulamaların temel mimari bileşenleri incelenir.

- **LLM mimarisi ve çalışma biçimi**
- **Transformer modelleri ve tokenizasyon**
- **Model eğitimi ve ince ayar (Fine-Tuning)**
- **Temel modeller (Foundation Models)**
- **Retrieval-Augmented Generation (RAG)**
- **Embedding'ler ve vektör veritabanları**
- **LLM uygulama mimarileri**
- **Sistem istemleri (System Prompts) ve bağlam (Context)**
- **AI ajanları ve otonom sistemler**

5. LLM ve Üretken Yapay Zekâ Saldırıları

Bu bölümde büyük dil modellerine ve üretken AI uygulamalarına yönelik başlıca saldırı yöntemleri incelenir.

- **Prompt Injection**
- **Doğrudan ve dolaylı Prompt Injection**
- **Jailbreaking**
- **Hassas bilgilerin ifşa edilmesi (Sensitive Information Disclosure)**
- **Model manipülasyonu**
- **Model çıkarma (Model Extraction)**
- **Eğitim verilerine yönelik saldırılar**
- **Veri zehirleme (Data Poisoning)**

- **Adversarial Inputs (Saldırgan Amaçlı Girdiler)**
- **Model hizmet engelleme saldırıları (Model Denial of Service)**
- **Güvenli olmayan çıktı işleme (Insecure Output Handling)**
- **Aşırı yetkilendirme / Excessive Agency**
- **Araç ve eklentilerin kötüye kullanılması (Tool and Plugin Abuse)**

6. AI Tehdit Çerçeveleri

Bu bölümde AI saldırılarının sistematik biçimde analiz edilmesi için kullanılan tehdit çerçeveleri ele alınır.

- **AI saldırı yaşam döngüsü**
- **MITRE ATLAS**
- **OWASP LLM Top 10**
- **AI saldırı tekniklerinin eşleştirilmesi**
- **Tehdit senaryoları ve saldırı yolları**
- Farklı **AI mimarilerindeki güvenlik zayıflıklarının** belirlenmesi

7. AI Red Teaming

Bu bölümde yapay zekâ sistemlerinin saldırgan bakış açısıyla değerlendirilmesine yönelik test yaklaşımları incelenir.

- **AI Red Teaming çalışmalarının amacı**
- **AI güvenlik testi metodolojileri**
- **Prompt tabanlı güvenlik testleri**
- **Model sınırlarının ve koruma mekanizmalarının** test edilmesi
- **İstismar edilebilir davranışların** belirlenmesi
- **Güvenlik bulgularının belgelenmesi ve paylaşılması**

8. Yapay Zekâ Sistemlerinin Güvenliğinin Sağlanması

Bu bölüm, AI modellerinin, verilerinin, uygulamalarının ve destekleyici altyapının korunmasına yönelik kontrolleri kapsar.

- **Güvenli AI mimarisi**
- **AI için derinlemesine savunma (Defence-in-Depth)**

BİLGİNÇ IT Academy

info@bilginc.com

bilginc.com

- **Girdi doğrulama (Input Validation)**
- **Çıktı filtreleme (Output Filtering)**
- **Model erişim kontrolleri**
- **Kimlik doğrulama ve yetkilendirme**
- **Ayrıcalıklı AI işlemleri**
- **Gizli bilgiler ve kimlik bilgilerinin korunması**
- **Veri koruma**
- **Model koruma**

9. LLM Uygulamalarının Güvenliğinin Sağlanması

Bu bölümde LLM tabanlı uygulamalar, RAG mimarileri ve AI ajanları için güvenlik önlemleri incelenir.

- **Prompt Injection saldırılarına karşı önlemler**
- **Guardrails (Koruma Sınırları)**
- **Sistem istemlerinin korunması**
- **Bağlam izolasyonu (Context Isolation)**
- **Güvenli RAG mimarileri**
- **Vektör veritabanlarının korunması**
- **Araç erişiminin kontrol edilmesi**
- **AI ajan izinleri**
- **AI ajanları için en az ayrıcalık ilkesi (Least Privilege)**
- **İnsan denetimi ve onay kontrolleri**

10. AI Tedarik Zinciri Güvenliği

Bu bölümde AI sistemlerinin geliştirilmesi ve işletilmesinde kullanılan harici bileşenlerden kaynaklanabilecek riskler ele alınır.

- **Üçüncü taraf modeller**
- **Açık kaynak modeller**
- **Model depoları (Model Repositories)**
- **AI kütüphaneleri ve bağımlılıkları**
- **Eğitim veri kümeleri**

- **Model kökeni (Model Provenance)**
- **Tedarikçi güvencesi (Supplier Assurance)**
- **Tedarik zincirinin ele geçirilmesi (Supply Chain Compromise)**
- **Bağımlılık ve bileşen yönetimi**

11. AI Güvenlik Testleri

Bu bölümde yapay zekâ sistemlerindeki güvenlik zayıflıklarını belirlemek ve kontrollerin etkinliğini değerlendirmek için kullanılan yöntemler incelenir.

- **Zafiyet değerlendirmesi (Vulnerability Assessment)**
- **Adversarial Testing**
- **Model güvenlik testleri**
- **AI penetrasyon testleri**
- **AI Red Teaming**
- **Guardrails ve güvenlik kontrollerinin test edilmesi**
- **Dağıtım öncesi güvenlik güvencesi (Security Assurance)**

12. AI Güvenlik Operasyonları

Bu bölümde yapay zekâ sistemlerinin izlenmesi ve şüpheli faaliyetlerin belirlenmesi için gerekli güvenlik operasyonları ele alınır.

- **AI sistemlerine özgü telemetri**
- **AI faaliyetlerinin loglanması**
- **Model etkileşimlerinin izlenmesi**
- **Anormal AI davranışlarının tespiti**
- **Güvenlik izleme sistemleriyle entegrasyon**
- **SIEM entegrasyonu**
- **AI sistemleri için tespit mühendisliği (Detection Engineering)**
- **Saldırgan faaliyetlerin belirlenmesi**

13. AI Güvenlik Olaylarına Müdahale

Bu bölümde yapay zekâ sistemlerini etkileyen güvenlik olaylarının tespit edilmesi, araştırılması ve yönetilmesi incelenir.

BİLGİNÇ IT Academy
info@bilginc.com
bilginc.com

- **AI güvenlik olayı senaryoları**
- **AI sistemlerinin ele geçirildiğinin tespit edilmesi**
- **Olay ön değerlendirmesi (Incident Triage)**
- **Olay araştırması (Investigation)**
- **Olayın sınırlandırılması (Containment)**
- **Model izolasyonu**
- **Kurtarma (Recovery)**
- **Kanıt toplama (Evidence Collection)**
- **Olay sonrası analiz (Post-Incident Analysis)**

14. Yapay Zekâ Destekli Siber Güvenlik

Bu bölümde yapay zekânın savunma amaçlı siber güvenlik faaliyetlerinde nasıl kullanılabileceği değerlendirilir.

- **AI destekli güvenlik izleme**
- **AI destekli tehdit avcılığı (Threat Hunting)**
- **Tehdit istihbaratı**
- **Güvenlik analizi**
- **Zararlı yazılım analizi (Malware Analysis)**
- **AI destekli tersine mühendislik (Reverse Engineering)**
- **Savunma amaçlı AI kullanımının fırsatları ve sınırlamaları**

15. AI Güvenlik Yönetimi

Bu bölümde kurumsal AI güvenliğinin yönetilmesine yönelik yönetişim, politika ve risk yönetimi süreçleri ele alınır.

- **AI güvenlik yönetişimi**
- **Roller ve sorumluluklar**
- **AI güvenlik politikaları**
- **AI risk değerlendirmesi**
- **Güvenlik güvencesi (Security Assurance)**
- **Düzenleyici gereklilikler ve uyumluluk**

BİLGİNÇ IT Academy

info@bilginc.com

bilginc.com

- **Üçüncü taraf AI risklerinin yönetimi**
- **AI modellerinin ve ajanlarının yönetişimi**

16. AI Güvenlik Programının Oluşturulması

Bu bölümde AI güvenliğinin mevcut kurumsal siber güvenlik programlarına nasıl dahil edilebileceği incelenir.

- **AI varlıklarının belirlenmesi**
- **AI güvenlik stratejisi**
- **Güvenli mimari ilkeleri**
- **AI güvenlik kontrolleri**
- **Güvenlik testleri ve Red Teaming**
- **İzleme ve olay müdahalesi**
- **Tedarik zinciri yönetişimi**
- **Sürekli güvenlik güvencesi (Continuous Assurance)**
- **AI güvenliğinin mevcut siber güvenlik programlarıyla bütünleştirilmesi**