# Siber Savunmada Yapay Zeka: Güncel Uygulamalar ve Teknolojiler

Ahsen Beyza Özkul

- 1) Tehdit Tespiti ve Olay Yanıtı'nda Yapay Zeka
- Twitter'ın Anomali Tespit Aracı Twitter, veri üzerinde alışılmadık desenleri tespit etmek için bir araç geliştirdi. Bu araç, ağ trafiği ve kullanıcı aktiviteleri gibi şeyleri gerçek zamanlı olarak izler ve makine öğrenimi kullanarak garip davranışları hızlıca tespit eder ve uyarı verir.

# twitter/ **AnomalyDetection**



Anomaly Detection with R

A 9
Contributors

**೪ 777**Forks

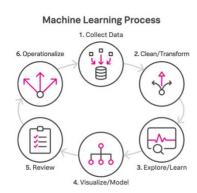


 Splunk'ın Makine Öğrenimi Araç Seti Splunk, sistem günlükleri ve operasyonlarda anomali tespiti için makine öğrenimi kullanan bir dizi araç sunar. Bu araçlar, IT dünyasında genellikle alışılmadık aktiviteleri belirleyip büyük sorunlara dönüşmeden önce uyarı gönderir.

# Splunk Machine Learning Toolkit (MLTK)

The MLTK is available to all Splunk Cloud or Splunk Enterprise customers and extends the value of the Splunk platform by enabling users to easily apply machine learning to their data.

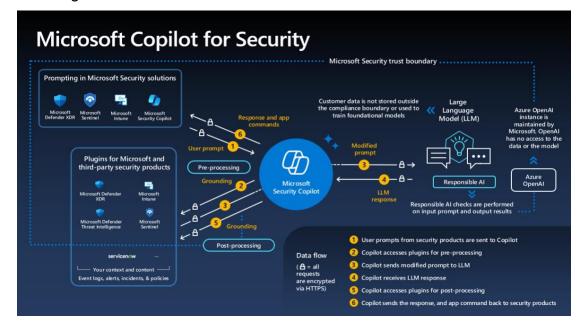
- Guide investigations by using machine learning to discover hidden meaningful patterns in your data
- Investigate your expanding data universe and avoid costly downtime
- Analyze and monitor at machine speed with purpose-built machine learning algorithms
- Automate action with trained models for alerts in real time



• Darktrace Al ile Olay Yanıtı Darktrace, siber tehditlere otomatik olarak yanıt veren bir yapay zeka kullanır. Bir saldırı gerçekleştiğinde, Darktrace etkilenmiş sistem parçalarını izole eder, böylece güvenlik ekibiniz sorunun yayılmaması için odaklanabilir ve sorunu çözmeye yönelik çalışabilir.



- Microsoft Security Copilot ile Olay Yanıtı Microsoft Security Copilot, siber saldırılarla başa çıkma sürecini otomatikleştirir. Olaylara hızlıca yanıt verebilir, problemi izole edebilir, gerekli bilgileri sağlayabilir ve sistemleri normale döndürebilir. Bu, güvenlik ekiplerinin zaman ve enerjisini korur.
- Microsoft Security Copilot Microsoft Security Copilot, ağı sürekli tarayarak alışılmadık aktiviteleri ve potansiyel tehditleri tespit eden bir yapay zeka aracıdır. Bu sayede şirketlerin verilerini korumak daha kolay hale gelir, insan gözetimine olan bağımlılık azalır.



 SentinelOne'ın Singularity Platformu SentinelOne'ın Singularity platformu, güvenlik ekiplerinin tüm ağda tehditleri aramasına yardımcı olur. Sistemdeki tüm verileri bir araya getirir ve tehditleri hızlıca bulup çözmeye yönelik net bir görünüm sağlar.



 IBM QRadar ve Watson ile Siber Güvenlik IBM QRadar güvenlik sistemi, Watson'ın yapay zekasını kullanarak güvenlik verilerini otomatik olarak analiz eder. Bu sayede güvenlik ekipleri potansiyel tehditleri hızlıca tespit edip yanıt verebilir, zaman ve çaba tasarrufu sağlar.



 AWS GuardDuty AWS GuardDuty, AWS hizmetlerinden gelen verileri izleyip analiz eder, alışılmadık aktiviteleri tespit eder. Anormal API çağrıları veya yetkisiz erişim girişimleri gibi potansiyel güvenlik ihlallerini belirler, tehdit tespitini geliştirir.



# 2) Güvenlik Açığı Yönetimi ve Güvenlik Artırımı'nda Yapay Zeka

• **IBM Adversarial Robustness Toolbox (ART)** IBM'in ART, Al sistemlerini saldırılardan korumak için tasarlanmış bir araç kitidir. Veri zehirlenmesi veya Al zayıflıklarını hedef alan siber saldırılara karşı savunma sağlar.



#### **Extended Support**

ART supports all popular machine learning frameworks (TensorFlow, Keras, PyTorch, MXNet, scikit-learn, XGBoost, LightGBM, CatBoost, GPy, etc.), all data types (images, tables, audio, video, etc.) and machine learning tasks (classification, object detection, generation, certification, etc.).



#### 39 Attack Modules

On a high level, ART supports 4 attack modules: Evasion, Poisoning, Extraction, and Inference.

Detailed information about the supported attack modules can be found <u>here</u>.



#### 29 Defense Modules

On a high level, ART supports 5 attack modules:
Preprocessor, Postprocessor, Trainer, Transformer, and Detector. Detailed information about the supported defense modules can be found here.



#### **Estimators and Metrics**

ART supports 3 robustness metrics, 1 certification and 1 verification metric. It also supports multiple estimators and details about the same can be found here.

- TensorFlow Privacy TensorFlow Privacy, TensorFlow platformunun bir uzantısıdır ve Al modellerinin gizlilik korumalarıyla eğitilmesini sağlar. Bu, Al sistemlerinin saldırganlar tarafından sömürülebilir veya manipüle edilebilir olmasını azaltır.
- OWASP ZAP ve Makine Öğrenimi OWASP ZAP, web uygulamalarında güvenlik açıklarını test eden bir araçtır. Makine öğrenimi eklenerek, SQL enjeksiyonları

veya çapraz site betikleme (XSS) gibi sorunları otomatikleştirerek bulma konusunda daha etkili hale gelir.

- HackerOne ve Al Destekli Tarama HackerOne, güvenlik araştırmacılarının yazılım hatalarını bulduğu bir platformdur. Al kullanarak güvenlik açıklarını bulma sürecini otomatikleştirir, uygulamalardaki zayıflıkları daha hızlı ve kolay bir şekilde keşfeder.
- GitHub Copilot ve CodeQL GitHub Copilot, geliştiricilere kod yazmada yardımcı olan bir Al aracıdır. CodeQL ile kullanıldığında, yazılan kodda güvenlik sorunlarını da tarar, sorunları erken aşamada yakalayarak büyük problemlere dönüşmelerini önler.

```
import datetime
monalisa
Write unit tests for this function
                                                                           def parse expenses (expenses string):
                                                                               """Parse the list of expenses and return the list of t Ignore lines starting with \#.
GitHub Copilot
                                                                               Example expenses_string:
  import unittest
                                                                                   2023-01-03 2.59 DKK
  import datetime
                                                                              2023-01-03 -2.72 EUR
  class TestParseExpenses(unittest.TestCase):
     def test_parse_expenses_with_valid_input(self):
                                                                               expenses = []
         expenses_string = '2023-01-02 -34.01 USD\n2023-01
expected_output = [(datetime.datetime(2023, 1, 2)
                                                                            for line in expenses_string.splitlines():
          self.assertEqual(parse_expenses(expenses_string),
                                                                               date, value, currency = line.split (" ")
     def test_parse_expenses_with_empty_input(self):
                                                                                 expenses.append((datetime.datetime.strptime (date,
          expenses_string =
          expected_output = []
          self.assertEqual(parse expenses(expenses string),
                                                                           expenses_data = '''2023-01-02 -34.01 USD
                                                                                            2023-01-03 -2.72 EUR'''
```

- SonarQube ve Al Eklentileri SonarQube, kod kalitesi ve güvenliğini kontrol
  eden bir araçtır. Al eklentileri ile, güvenlik açıklarını bulmada daha etkili olur ve
  yanlış alarm sayısını azaltır, böylece geliştiricilerin gerçek sorunları daha kolay
  çözmelerini sağlar.
- **IBM Guardium ve Yaman Yönetimi** IBM Guardium, verilerinizi güvence altına almak için Al kullanır, hem yerel hem de bulut ortamlarında güvenlik açıklarını bulur ve düzeltir. Yeni tehditlere uyum sağlar ve verilerinizi korumak için gerekli yamaları zamanında uygular.
- **Tenable'in Exposure Al** Tenable'in Exposure Al, sisteminizdeki zayıf noktaları tespit eder ve bunları hackerlar tarafından kullanılmadan önce düzeltir. Al kullanarak, zayıflıkları tarar, hangi açıkların daha kritik olduğunu önceliklendirir ve hızlıca yamalar, sisteminizin daha güvenli olmasını sağlar.



 Zscaler Veri Koruma Zscaler Veri Koruma, hassas bilgileri korumak için Al kullanır, belgeleri, e-postaları ve görüntüleri tarar. Verileri sınıflandırır, yetkisiz erişimleri tespit eder ve veri ihlallerini önler, hassas bilgilerinizi yönetmeyi ve korumayı kolaylaştırır.

### Kimlik ve Erişim Yönetimi'nde Yapay Zeka

• **IBM Verify** IBM Verify, kimlik ve erişim yönetimini geliştirmek için Al kullanır, kullanıcı davranışlarını analiz eder ve kimlik doğrulama gereksinimlerini ayarlar. Anomalileri tespit edebilir ve otomatik olarak çok faktörlü kimlik doğrulama uygular, güvenliği artırırken kullanıcı erişim yönetimini basitleştirir.



## Kaynaklar

- IBM. (t.y.). Tehdit tespiti ve yanıt. Erişim adresi: https://www.ibm.com/services/threat-detection-response
- IBM. (t.y.). Guardium. Erişim adresi: <a href="https://www.ibm.com/guardium">https://www.ibm.com/guardium</a>
- IBM. (t.y.). Watsonx yönetimi. Erişim adresi: https://www.ibm.com/products/watsonx-governance
- Watkins, O. (2024, 19 Nisan). Siber güvenlikte Al kullanımı için 4 vaka. Red Hat. Erişim adresi: <a href="https://www.redhat.com/en/blog/4-use-cases-ai-cyber-security">https://www.redhat.com/en/blog/4-use-cases-ai-cyber-security</a>
- Goss, A. (2024, 13 Mayıs). Siber güvenlikte Al örnekleri. Station X. Erişim adresi: https://www.stationx.net/examples-of-ai-in-cyber-security/
- Shutenko, V. (2024, 8 Ağustos). Siber güvenlikte AI. TechMagic. Erişim adresi: https://www.techmagic.co/blog/ai-in-cybersecurity/
- (t.y.). Siber güvenlik projelerinde yapay zeka. Network Simulation Tools. Erişim adresi: <a href="https://networksimulationtools.com/artificial-intelligence-in-cyber-security-projects/">https://networksimulationtools.com/artificial-intelligence-in-cyber-security-projects/</a>
- Daivi. (2024, 19 Mart). Siber güvenlik makine öğrenimi projeleri. ProjectPro. Erişim adresi: <a href="https://www.projectpro.io/article/cybersecurity-machine-learning-projects/631">https://www.projectpro.io/article/cybersecurity-machine-learning-projects/631</a>
- Caniszczyk. (t.y.). AnomalyDetection. GitHub. Erişim adresi: https://github.com/twitter/AnomalyDetection
- Fier, J., & Kenyon Grant, S. (t.y.). *Splunk makine öğrenimi aracı*. Splunk. Erişim adresi: <a href="https://www.splunk.com/en\_us/resources/splunk-machine-learning-toolkit.html">https://www.splunk.com/en\_us/resources/splunk-machine-learning-toolkit.html</a>
- Darktrace. (2022, 12 Nisan). Darktrace'in Siber Al Analisti, ABD federal hükümetine olay raporlamalarını nasıl hızlandırır. Erişim adresi: <a href="https://darktrace.com/blog/how-darktraces-cyber-ai-analyst-accelerates-reporting-incidents-to-the-us-federal-government">https://darktrace.com/blog/how-darktraces-cyber-ai-analyst-accelerates-reporting-incidents-to-the-us-federal-government</a>
- Microsoft. (2024, 18 Temmuz). Microsoft güvenlik yardımcı pilotu. Erişim adresi: https://learn.microsoft.com/en-us/copilot/security/microsoft-security-copilot
- (t.y.). *Singularity platformu*. Kidan. Erişim adresi: https://kidan.co/partners/singularity-platform/
- (t.y.). *IBM'in QRadar danışmanı ile Watson'un çalışma aşamaları*. ResearchGate. Erişim adresi: <a href="https://www.researchgate.net/figure/Stages-involved-in-the-working-of-IBMs-QRadar-advisor-with-Watson-64">https://www.researchgate.net/figure/Stages-involved-in-the-working-of-IBMs-QRadar-advisor-with-Watson-64</a> fig3 373712758
- Amazon Web Services. (t.y.). Amazon GuardDuty artık Amazon EKS çalışma zamanı izlemeyi destekliyor. Erişim adresi: <a href="https://aws.amazon.com/tr/blogs/aws/amazon-guardduty-now-supports-amazon-eks-runtime-monitoring/">https://aws.amazon.com/tr/blogs/aws/amazon-guardduty-now-supports-amazon-eks-runtime-monitoring/</a>
- Hajra, A. (2023, 4 Mayıs). IBM ART karşıt saldırganlık testi bir film öneri sistemi için. Medium. Erişim adresi: <a href="https://medium.com/@asmitahajra/ibm-art-adversarial-robustness-check-for-a-movie-recommendation-system-649ba46e9e8a">https://medium.com/@asmitahajra/ibm-art-adversarial-robustness-check-for-a-movie-recommendation-system-649ba46e9e8a</a>
- (t.y.). Copilot. GitHub. Erişim adresi: <a href="https://github.com/features/copilot">https://github.com/features/copilot</a>
- (t.y.). *Maruziyet yönetimi: Tenable One nedir?* E-SPIN. Erişim adresi: <a href="https://www.e-spincorp.com/exposure-management-what-is-tenable-one/">https://www.e-spincorp.com/exposure-management-what-is-tenable-one/</a>

• (t.y.). *IBM Verify incelemeleri*. TrustRadius. Erişim adresi: <a href="https://www.trustradius.com/products/ibm-verify/reviews?qs=pros-and-cons#comparisons">https://www.trustradius.com/products/ibm-verify/reviews?qs=pros-and-cons#comparisons</a>