SİBER GÜVENLİKTE YAPAY ZEKANIN ÖNEMİ

ÖZET

Bu çalışmanın amacı, yapay zeka (AI) uygulamalarının siber güvenlikteki önemi ve rolünü incelemektir. Dijital dünyanın hızla gelişmesi ve buna bağlı olarak tehditlerin artması, yapay zeka tabanlı çözümlerin gerekliliğini daha da artırmaktadır. Daha önce bu alanda yapılan çalışmalara dayanarak mevcut yaklaşımlar sistematik olarak analiz edilerek, yapay zekanın tehdit tespiti, otomatik yanıt sistemleri, davranış analizi ve büyük veri analizi gibi algoritmalarının siber güvenlik alanında sağladığı avantajlar değerlendirilmiştir.

Bunun yanı sıra, kamu ve özel sektörde karşılaşılabilecek etik sorunlar, yanıltıcı veriler ve yasal süreçler de ele alınmıştır. Türkiye'de siber güvenlik ve yapay zekanın önemi, yapay zekanın Türkiye'deki yaygınlığı ve kullanımı incelenmiştir. Elde edilen bilgiler, yapay zekanın siber güvenlik alanında önemli bir rol oynadığını ve gelecekte daha etkin olacağını göstermektedir. Bu çalışma, hem Türkiye'de hem de dünyada yapay zekanın siber güvenlik ile nasıl iç içe olduğunu ve gelecekteki etkilerini bütüncül bir yaklaşımla ele almayı amaçlamaktadır.

Anahtar kelimeler: Siber güvenlik, Yapay zeka, (AI) uygulamaları, Türkiye'de yapay zeka ve siber güvenlik , Makine öğrenimi ,Tehdit algılama, Güvenlik Protokolleri

GİRİŞ

Yapay Zekanın Tarihsel Gelişimi ve Önemi

Yapay zeka, ilk kez 1950'li yıllarda bir fikir olarak ortaya çıkmış ve zamanla büyük bir dönüşüm geçirerek günümüzün en etkili teknolojilerinden biri haline gelmiştir. Bu gelişim süreci, yıllar süren teorik çalışmalar ve mühendislik başarılarıyla desteklenmiştir. 1961'de Shakey adında bir robot, ilk otonom hareket eden cihaz olarak tarihe geçmiştir. 1970'lerden 1990'lara kadar yapay zeka konusunda beklentiler büyük olsa da, dönemin teknolojik sınırlamaları nedeniyle araştırmalar istenilen ilerlemeyi kaydedememiştir. Ancak, 1990'lardan günümüze internetin yaygınlaşması ve hesaplama gücündeki artış, yapay zekâya yeni bir ivme kazandırmıştır. Özellikle 1997'de IBM'nin Deep Blue sistemi, dünya satranç şampiyonu Garry Kasparov'u yenerek önemli bir başarıya imza atmıştır. 2010 sonrasındaki dönemde ise makine öğrenimi ve derin öğrenme teknikleri sayesinde yapay zekanın uygulama alanları hızla genişlemiştir.

Türkiye'de Yapay Zeka Kullanımı

Son yıllarda, hem Türkiye'de hem de dünya genelinde yapay zeka kullanımı hızla artmaktadır. Türkiye'de, yapay zeka teknolojilerinin kullanımı özellikle girişimciler tarafından büyük bir ilgi görmektedir. 2023 yılında yapılan bir araştırmaya göre, Türkiye'deki kullanıcıların %72'si yapay zekayı kişisel amaçlarla, %51'i eğitim için ve %34'ü iş yerinde verimliliği artırmak için kullanmaktadır. Ayrıca, üretken yapay zeka kullanım oranı %52'ye ulaşarak küresel ortalama olan %42'nin üzerine çıkmıştır (turkiye.ai, 2023)


TRAI
Yapay Zeka Zaman Çizelgesi
turkiye.ai
Yapay Zekanın Doğuşu 1943 - 1956
Altın Çağ 1956 – 1974
Yapay Zeka Kışı 1974 – 1980
GPU Çağı 2012 - Günümüz
Bilgisayar ve Zeka
Alan Turing'in, düşünen makineler yaratma olasılığı hakkında düşüncelerini paylaşlığı makalesi, bir dönüm noktası yarattı.
Yapay Zeka ve Oyun
Manchester Üniversitesi'nin Ferranti Mark 1 makinesini kullanan Christopher Strachey bir dama programı, Dietrich Prinz ise bir satranç programı yazdı.
Perceptron
Marvin Minsky
"Bir kuşak içinde 'yapay zeka' oluşturma problemi çözülmüş olacak."
Cylons
Orjinal "Savaş Yıldızı Galactica" bilim kurgu dizisi savaşçı robotlar Cylonları tanıttı.
Deep Blue ve Kasparov
IBM'in Deep Blue'su, Garry Kasparov ile girdiği satranç karşılaşmasını kazandı.
Watson ve Jeopardy!
IBM'in Watson bilgisayarı, televizyon yarışması "Jeopardy!" şampiyonları Rutter ve Jennings'i yendi.
Grafik İşlemcileri (GPU) Çağı
GPU odaklı bir sistem, İmagenet'te en iyi hata oranını yarıya indirerek birinci oldu.
GAN
Ian Goodfellow tarafından Generative Adversarial Networks(Çekişmeli Üretici Ağlar) bulundu. Yapay zekanın gerçeğe benzer sahte üretimler yapabilmesinin önü açıldı.
Asilomar
Asilomar Conference on Beneficial AI, Future of Life Institute tarafından, Kaliforniya'daki Asimolar Konferans Alanı'nda düzenlendi.
BERT
Google, dönüştürücü ağ tabanlı doğal dil işleme modeli BERT'i yayınladı.
GPT-3
175 milyar parametre
1950
1951
1955
1957
1962
1967
1969
1978
1980
1997
2006
2011
2012
2014
2016
2017
2018
2019
2020
2021
I, Robot
John McCarthy
1956'daki Dartmouth Konferansı için 1955 yılında isim babası olduğu Yapay Zeka erimini ortaya atmıştır.
Unimation
İlk endüstriyel robot şirketi Unimation kuruldu.
XOR Problemi
"Perceptrons: An Introduction to Computational Geometry" – Marvin Minsky & Seymour Papert
Amerikan Yapay Zeka Derneği
İlk Ulusal Konferansı, Stanford'da düzenlendi.
Deep Neural Network
Ruslan Salakhutdinov & Geoffrey Hinton
Apple Siri
Amazon Alexa
AlphaGO
Google DeepMind'ın AlphaGO'su, Lee Sedol ile karşılaştığı go maçını 4-1 kazandı.
Transformer Networks
Dönüştürücü ağlar adında yeni bir sinir ağı türü tanıtıldı.
GPT-2
1.5 milyar parametreli GPT-2, OpenAI tarafından yayınlandı.
DALL-E
Yazıyla tarif edilen resimleri üretebilme yeteneği olan DALL-E adlı çalışma OpenAI tarafından yayınlandı.

Siber Güvenlikte Yapay Zeka

Teknolojinin ve yapay zekanın hızlı gelişimi, siber güvenlik alanında da büyük bir dönüşüme yol açmıştır. Geleneksel güvenlik sistemleri, gelişmiş tehditlerle başa çıkmakta yetersiz kalmaktadır. Bu durum, yapay zekâ destekli güvenlik sistemlerinin önemini artırmıştır. Bu sistemler, siber saldırıları önceden tespit etmek, tehditleri analiz etmek, otomatik yanıt sistemleri geliştirmek ve büyük veri analitiği ile anormal davranışları gözlemlemek gibi birçok alanda kullanılmaktadır. Örneğin, yapay zeka destekli siber saldırılar, hedefe spesifik olarak uyarlanabilir ve adaptasyon gösterebilir. AI, bir şirketin güvenlik protokollerini analiz ederek zayıf noktalarını tespit edebilir ve bu bölgelere yönelik özel saldırı senaryoları geliştirebilir (Pandermos.net, 2023). Yapay zeka, siber güvenlik alanında nasıl daha etkili hale getirilebilir? Bu soru, özellikle gelişmiş tehditlerin tespit edilmesinde, yapay zeka tabanlı sistemlerin etkinliğinin artırılması açısından büyük önem taşımaktadır.

Etik ve Hukuki Boyutlar

Bunun yanı sıra, kamu ve özel sektörde karşılaşılabilecek etik sorunlar, yanıltıcı veriler ve yasal süreçler de önemli konulardır. Yapay zekanın siber güvenlikte kullanımı, yeni etik ve hukuki soruları gündeme getirmektedir. Bu nedenle, yapay zeka destekli güvenlik çözümlerinin uluslararası etik standartlara uygun olması gerekmektedir.

Sonuç

Yapay zekanın siber güvenlik alanında önemli bir rol oynadığını ve gelecekte daha etkin olacağını göstermektedir. Bu çalışma, hem Türkiye'de hem de dünyada yapay zekanın siber güvenlik ile nasıl iç içe olduğunu ve gelecekteki etkilerini bütüncül bir yaklaşımla ele almayı amaçlamaktadır.

SİBER GÜVENLİKTE YAPAY ZEKA UYGULAMALARI

Makine öğrenmesi ve yapay zeka teknolojileri, izleme, denetim, tehdit algılama ve alarm sistemlerini kapsayan otomatikleştirilmiş siber savunma teknikleri için kritik bir altyapı sunmaktadır. Bu teknolojiler, kullanıcı davranışlarını analiz ederek zararlı ve zararsız aktiviteleri ayırt edebilmekte ve birbirinden bağımsız gibi görünen saldırı göstergelerini yorumlayarak korelasyon kurallarına dayalı alarmlar üretebilmektedir. Bu sayede, siber güvenlik alanında görevli ekiplerin iş yükünü hafifletmekte ve etkinliklerini artırmaktadır.

Yapay zeka destekli sistemlerin, yalnızca otomatize tehdit algılama süreçlerinde değil, aynı zamanda müdahale yeteneğine sahip otonom güvenlik sistemlerine dönüşmesi öngörülmektedir. Bu tür sistemler, tehditleri tespit edip gerekli önlemleri insan müdahalesi olmaksızın uygulayarak siber güvenlik süreçlerinin daha verimli bir şekilde yürütülmesini sağlamaktadır. Özellikle bulut teknolojilerinde yaygınlaşan bu sistemler, büyük veri setleriyle çalışarak saldırılara karşı proaktif savunma mekanizmaları geliştirmektedir.

Siber güvenlik çözümlerinde sıklıkla kullanılan yapay zeka teknikleri arasında derin öğrenme (deep learning), doğal dil işleme (NLP), gözetimli (supervised) ve gözetimsiz (unsupervised) öğrenme yöntemleri yer almaktadır (BeyazNet, 2023).

Deep Learning – Derin öğrenme

Grafik işleme birimleri (GPU'lar), paralel işlem kapasitesine sahip olmaları sayesinde derin öğrenme (deep learning) modellerinin eğitilmesinde büyük bir avantaja sahiptir. Aynı anda çok sayıda matematiksel işlemi gerçekleştirebilme yetenekleri, özellikle büyük veri kümeleri üzerinde çalışan yapay zeka algoritmalarının verimli bir şekilde çalışmasına olanak

tanımaktadır. Bu bağlamda, derin öğrenme algoritmaları; otomasyon, veri analitiği ve siber tehditlerin tespiti gibi birçok karmaşık görevin etkin bir şekilde yürütülmesini sağlamaktadır.

Derin öğrenme teknikleri, geleneksel kötü amaçlı yazılım analizlerinden farklı olarak yalnızca bilinen tehdit imzalarına veya önceden tanımlanmış saldırı kalıplarına bağlı kalmaz. Bunun yerine, sistem davranışlarını inceleyerek öğrenme yetisi kazanır ve bu sayede yeni veya daha önce görülmemiş tehditleri tespit etme konusunda üstün performans sergiler. Bu yaklaşım, özellikle sıfırıncı gün (zero-day) saldırılarının tespitinde ve tehdit aktörlerinin davranışlarını önceden öngörmede kritik rol oynar.

Derin öğrenmenin bu avantajları, özellikle Saldırı Tespit Sistemleri (IDS) ve Saldırı Önleme Sistemleri (IPS) gibi güvenlik bileşenlerinin daha gelişmiş hale gelmesine olanak tanımaktadır. Örneğin, uygulama katmanına yönelik SQL enjeksiyonu ve Dağıtık Hizmet Engelleme (DDoS) saldırılarının tespiti için HTTPS trafiği analiz edilmekte ve bu analiz sürecinde Derin Sinir Ağları (Artificial Neural Networks - ANN) kullanılmaktadır. ANN modelleri, gelen trafiği hem içerik hem de davranışsal örüntüler üzerinden analiz ederek şüpheli aktiviteleri daha yüksek doğruluk oranıyla belirleyebilmektedir.

Derin öğrenme teknolojilerinin bir başka dikkat çekici uygulama alanı ise Kullanıcı ve Varlık Davranış Analitiği (User and Entity Behavior Analytics - UEBA) çözümleridir. Bu sistemler, kurum içindeki olağandışı kullanıcı aktivitelerini tespit etmek için normal kullanıcı davranışlarını modelleyerek, insider threat olarak bilinen iç tehditleri açığa çıkarmayı hedeflemektedir. UEBA çözümleri sayesinde, yetkisiz erişim girişimleri, veri sızıntısı riskleri veya anormal trafik akışları daha erken safhalarda belirlenebilir. Ayrıca, derin öğrenmenin alt alanlarından biri olan Doğal Dil İşleme (Natural Language Processing - NLP), spam e-postalar ve sosyal mühendislik saldırıları gibi metin temelli tehditlerin analizinde kullanılmaktadır. NLP teknikleri, e-posta içeriğini dil yapısı ve semantik açıdan analiz ederek şüpheli içerikleri tanımlar. Google'ın Gmail hizmeti, spam tespiti konusunda NLP tabanlı algoritmalardan yararlanarak kullanıcı güvenliğini artırmaktadır. (Google, 2025).

Siber güvenlikte kısaca Deep Learning uygulamaları :

*****Kötü amaçlı yazılım (Malware) tespiti

*****Ağ trafiği analizi ve Anomali tespiti

*****Kimlik avı (Phishing) saldırılarının tespiti

*****Sızma tespiti sistemleri (Intrusion Detection System –IDS)

*****Zararlı URL / Domain tespiti

*****Spam ve Zararlı E-Posta tespiti

Gözetimsiz (Unsupervised) ve Gözetimli (Supervised) Öğrenme

Yapay zeka temelli siber güvenlik sistemleri, makine öğrenmesi teknikleri sayesinde yalnızca tehditleri tespit etmekle kalmaz, aynı zamanda bu tehditlere karşı kendini sürekli geliştirerek daha etkili çözümler üretebilir. Bu sistemler, izlenen bilgi teknolojileri altyapılarından elde edilen büyük veri kümelerini analiz ederek sıradışı aktiviteleri ayırt etme yeteneğine sahiptir. Özellikle normal sistem davranışlarının tanımlandığı ve bu davranışlardan sapmaların tehdit olarak değerlendirildiği baseline (referans çizgisi) oluşturma yaklaşımı, tehdit tespiti süreçlerinde önemli bir yere sahiptir.

Baseline yöntemi, sistemlerin veya kullanıcıların gündelik rutin aktivitelerinin istatistiksel olarak bir referans çizgisi haline getirilmesine dayanır. Örneğin; yalnızca belirli cihazlarla iletişim kurması beklenen bir sistemin farklı ağ noktalarıyla alışılmadık bir trafik oluşturması, bu referans çizgisinden sapma olarak değerlendirilir ve sistem bu tür bir anomaliyi tespit ederek alarm üretir. Böylece, önceden belirlenmiş tehdit imzalarına gerek kalmadan, gerçek zamanlı davranış analizi ile tehditlerin belirlenmesi mümkün olur.

Bu yaklaşım, özellikle etiketsiz (label'sız) veri ile çalışan Gözetimsiz Öğrenme (Unsupervised Learning) modellerine dayanmaktadır. Gözetimsiz öğrenme, sistemin veri setlerini kendi içinde anlamlı örüntüler ve benzerlikler yoluyla sınıflandırmasını sağlar. Bu yöntem, verilerin sınıflandırma ya da kategorilendirme açısından önceden insan müdahalesiyle etiketlenmediği durumlarda tercih edilir.

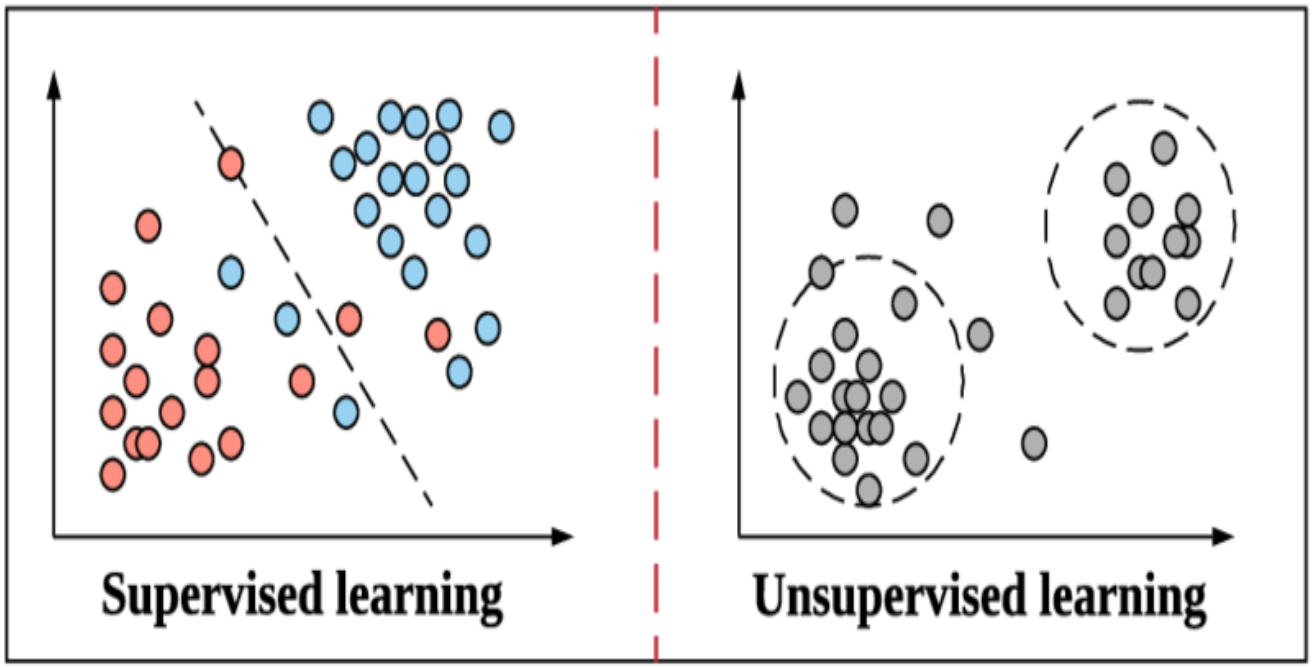
Öte yandan, Gözetimli Öğrenme (Supervised Learning) tekniklerinde veriler, sınıflarına veya kategorilerine göre etiketlenmiş olarak sisteme sunulur. Bu teknikler, daha yüksek doğruluk oranlarına ulaşabilmekle birlikte, veri setinin hazırlanmasında insan müdahalesine ihtiyaç duyar. Zararlı yazılım (malware) tespiti gibi belirli tehdit türlerine yönelik analizlerde gözetimli öğrenme yöntemleri etkin şekilde kullanılmaktadır. Özellikle karar ağaçları (decision trees), lineer regresyon (linear regression), destek vektör makineleri (support vector machines) ve lojistik regresyon (logistic regression) gibi algoritmalar bu yöntem kapsamında sıkça tercih edilmektedir.

Kısacası, gözetimli öğrenme modelleri daha yüksek doğruluk oranları sunarken, gözetimsiz modeller daha esnek ve daha az insan müdahalesine ihtiyaç duyan bir yapı sunar. Modern siber güvenlik yaklaşımlarında ise bu iki yöntemin hibrit kullanımı giderek daha yaygın hâle gelmektedir.

Kaynakça :

EnesHZR. (n.d.). Yapay Zekânın Siber Güvenlikte Kullanımı. Medium. https://eneshzr.medium.com/yapay-zek%C3%A2n%C4%B1n-siber-g%C3%BCvenlikte-kullan%C4%B1m%C4%B1-2d98d4bb867a

ÖZELLİK	GÖZETİMLİ ÖĞRENME	GÖZETİMSİZ ÖĞRENME
Gerekli Veri Türü	Etiketli Veri	Etiketsiz Veri
Amaç	Sınıflandırma, tahmin	Gruplama, Anomali tespiti
Avantajı	Yüksek doğruluk	Yeni tehditleri tespit edebilir
Dezavantajı	Etiketleme maliyetli	Doğruluk kontrolü zordur
Kullanım Alanı	Bilinen saldırılar	Bilinmeyen saldırılar



Yapay Zekanın Siber Suçlularca Kötüye Kullanılması

Yapay zeka tabanlı saldırı araçlarının, daha önce manuel çalışma gerektiren ve maliyetli olan hedefe özel saldırıların sıfır maliyetle gerçekleştirilmelerini sağlayacak bir otomasyon özelliğini siber suçlulara sunacağı belirtiliyor. AI ile ilgili siber saldırıların bir yandan yapay zeka sistemlerini hedef alma bir yandan atakları kapsamlı hale getirmek için yapay zeka tekniklerini kullanma etrafında şekilleneceği öngörülüyor.

Yapay zekaya sahip sistemler, endüstriyel uygulamaların işlevlerini otomatikleştirme ve karar alma süreçlerini iyileştirme gibi avantajlar sunmaktadır. Ancak büyük veriyi işleme gibi fonksiyonları onları aynı zamanda siber saldırılar için bir hedef haline getirmektedir. Mantıksal fonksiyonlarını bozup işlevlerini yapamaz hale getirecek zararlı girdilere karşı kırılgan bir yapıya sahip olan yapay zeka teknolojilerinin güvenliği de yeni bir savunma alanı olarak değerlendirilmektedir.

Yapay zekanın siber saldırılar için manipüle edilmesine örnek gösterilebilecek bazı uygulamalar : Ağları ve sistemleri yoklayarak istismar edilebilecek daha önce keşfedilmemiş zafiyetleri tespit etmek, gerçekçi içeriklerle oltalama ve sosyal mühendislik saldırılarını daha karmaşık hale getirebilmek.

Yapay zeka modellemeleri aynı zamanda etkili siber saldırı tekniklerinin geliştirilmesine yol açmaktadır. Bunların en son örneklerinden biri de yapay zeka temelli sosyal mühendislik saldırısı DeepFake'tir.

Kaynakça : BeyazNet. (t.y.). *Siber Güvenlikte Yapay Zekâ Uygulamaları*. BeyazNet. https://www.beyaz.net/tr/guvenlik/makaleler/siber_guvenlikte_yapay_zeka_uygulamalari.html

Siber Suçlarda Yapay Zeka Destekli Pozisyonlar

Sosyal Mühendislik:

Sosyal mühendislik, özel bilgi, erişim veya kıymetli varlıkları elde etmek için insan hatasından yararlanan bir manipülasyon tekniğidir. Siber suç dünyasında, bu "insan korsanlığı" dolandırıcıları, dikkatsiz kullanıcıları verilerini ifşa etmeye, kötü amaçlı yazılım bulaştırmaya veya kısıtlı sistemlere erişim izni vermeye ikna etme eğilimindedir. Saldırılar çevrimiçi, yüz yüze ve diğer etkileşim yollarıyla gerçekleşebilir.

Sosyal mühendisliğe dayalı dolandırıcılıklar, insanların düşünme ve hareket etme yolları üzerine kuruludur. Bu bakımdan, sosyal mühendislik saldırıları özellikle kullanıcının

davranışlarını yönlendirmek için kullanılır. Saldırgan, bir kullanıcının eylemlerini nelerin belirlediğini anladığında, kullanıcıyı etkin bir şekilde kandırıp yönlendirebilir.

Buna ek olarak, bilgisayar korsanları kullanıcının bilgi eksikliğinden de faydalanmaya çalışırlar. Teknolojinin hızı nedeniyle, birçok tüketici ve çalışan istenmeyen programın kullanıcı izni dışında bilgisayara otomatik indirilmesi (drive-by downloads) gibi bazı tehditlerin farkında değildir. Ayrıca kullanıcılar, telefon numaraları gibi kişisel verilerinin ne kadar değerli olduğunun farkında olmayabilir. Sonuç olarak, birçok kullanıcı hem kendilerini hem de bilgilerini en iyi şekilde nasıl koruyacaklarından emin değildir.

Genel olarak, sosyal mühendislik saldırganları şu ikisinden birini hedef alır:

1- Sabotaj : Zarar vermek veya rahatsızlık yaratmak için verileri bozmak ve tahrip etmek.
2- Hırsızlık : Bilgi, erişim veya para gibi değerli varlıkların çalınması.

Çoğu sosyal mühendislik saldırısı, saldırgan ve kurban arasındaki iletişimde zorlama yöntemlerini kullanmak yerine, kullancıyı kendisini tehlikeye atmaya yönlendirme ile gerçekleşir.

Sosyal mühendislik saldrıı süreci aşamaları

1- Hazırlık: Sizin veya parçası olduğunuz daha geniş bir grubun geçmişi hakkında bilgi toplayarak.
2- Sızma : İlişki kurarak veya bir etkileşim başlatarak, güven kazanmaya başlama
3- Kurbandan faydalanma : Saldırıyı gerçekleştirmek için güven ve zayıflık sağlandığında.
4- Bağlantıyı kesme : Kullanıcı istenilen eylemi gerçekleştirdiğinde.

Sosyal Mühendislik Saldırılarının Özellikleri

Sosyal mühendislik saldırganın ikna ve güven sağlama taktikleri üzerine kurulur. Bu taktiklere maruz kaldığınızda normalde başka koşularda yapmayacağınız eylemleri yapma olasılığınız artmaktadır.

Yüksek Duygusallık: Duygusal manipülasyon, saldırganlara her türlü etkileşimde üstünlük sağlar. Aşırı duygusal bir durumdayken mantıksız veya riskli eylemlerde bulunma olasılığınız çok daha yüksektir.

- KORKU
- HEYECAN
- MERAK
- ÖFKE
- SUÇLULUK
- ÜZÜNTÜ

Sosyal mühendislik saldırısı nedir ve bunu nasıl tespit edebilirim ?

Sosyal Mühendislik Saldırı Türleri

- **Kimlik Avı Saldırıları :** İstenmeyen e-postayla kimlik avı veya toplu kimlik avı, birçok kullanıcıyı hedef alan yaygın bir saldırıdır. Bu saldırılar kişiye özel değildir ve şüphelenmeyen kişileri avlamaya çalışır.
- **Casus Kimlik Avı** : belirli kullanıcıları hedef almak için kişisel bilgileri kullanır. Balina avı diyede adlandırılan bu tür saldırılarda üst düzey yöneticiler ve üst düzey hükümet yetkilileri gibi yüksek öneme sahip hedeflere yöneliktir.

Sesli kimlik avı (telefonlar kişisel bilgilerin ele geçirilmesi) aramaları, tüm bilgilerinizi kaydeden mesajları, bir web bağlantısı veya sahte bir e-posta ya da telefon numarası üzerinden işlem yapılması için bir talimat içerebilir.

Sms kimlik avı (cep telefonu mesajları üzerinden kimlik avcılığı) bir web bağlantısı veya sahte bir e-posta ya da telefon numarası üzerinden işlem yapılması için bir talimat içerebilir.

E-posta kimlik avı, sizi cevap yazmaya veya başka yollarla iletişim kurmaya teşvik eden e-postalar kullanan en geleneksel kimlik avı yöntemidir. Web bağlantıları, telefon numaraları veya kötü amaçlı yazılım içeren ekler kullanabilir.

Hedef şaşırtan kimlik avı, saldırganın güvenilir bir şirketin müşteri hizmetleri ekibini taklit ettiği sosyal medyada gerçekleşir. Bir marka ile yaptığınız görüşmeleri ele geçirip özel mesajlar haline getirerek saldırıyı buradan ilerletirler.

Arama motoru kimlik avı, sahte web sitelerine olan bağlantıları arama sonuçlarının en üstine yerleştirme girişimidir. Bunlar ücretli reklamlar olabilir veya arama sıralamalarını manipüle etmek için meşru optimizasyon yöntemleri de kullanabilirler.

URL kimlik avı bağlantıları sizi kimlik avı web sitelerine gitmeye yönlendirir. Bu bağlantılar genellikle e-postalar, metinler, sosyal medya mesajları ve **ç**evrimiçi reklamlarda verilir. Saldırganlar, bağlantı kısaltma araçları veya adresi aldatıcı bir şekilde yazılmış URL'ler kullanarak köprülü metin veya butonlardaki bağlantıları gizler.

Oturum açma kimlik avı, normal internet gezintinizde bir aksama olarak görülür. Örneğin, şu anda ziyaret ettiğiniz sayfalarda sahte giriş pencereleri gibi şeyler görebilirsiniz.

Peki, bir saldırganın aynı anda yüzlerce hatta binlerce kişiye kişiselleştirilmiş saldırılar yapabilir mi? İşte tam bu noktada siber saldırılar da tıpkı teknolojiler gibi evrim geçiriyor. Geleneksel sosyal mühendislik saldırıları, yerini artık daha sofistike, daha hızlı ve daha geniş ölçekli saldırılara bırakıyor. Bu değişimin arkasındaki en büyük itici güç ise yapay zeka. Buda bizi yeni bir tehditle karşı karşıya bırakıyor. Otomatik sosyal mühendislik.

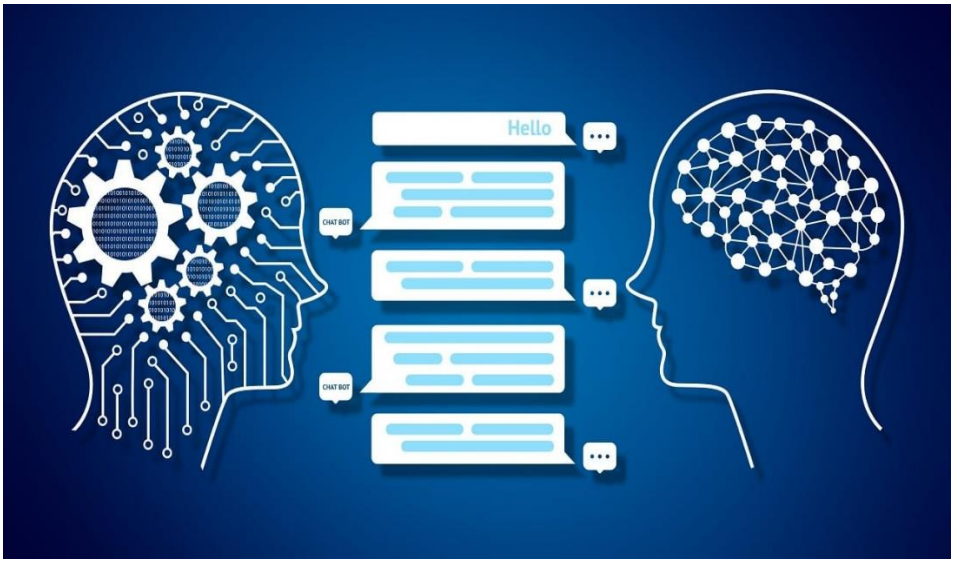
Otomatik Sosyal Mühendislik: Otomatik mühendislik, yapay zeka destekli sistemlerin kullanılarak bireyleri manipüle etmeye yönelik otomatik mesajlar e-postalar veya etkileşimler üretmesidir. Bu sistemler genellikle:

- GPT benzeri dil modelleri
- Veri kazıma araçları
- Otomatik e-posta ya da mesaj göndericileri

İle birlikte çalışır. Peki nasıl çalışır ?

1- **Hedef belirleme :** Sosyal medya, açık kaynaklar, geçmiş veri sızıntıları gibi yerlerden kişi hakkında veri toplanır.
2- **Profil oluşturma :** AI, toplanan verilere göre kişinin ilgi alanlarını, çalışma ortamını, iletişim dilini analiz eder.

3- **İçerik Üretimi :** GPT-3 / GPT-4 gibi modellerle, kişiye özel ve etkileyici mesajlar üretilir.
4- **Saldırı :** Üretilen mesajlar, e-posta, WhatsApp, sosyal medya ya da sahte destek siteleri aracılığıyla gönderilir.



IBM X-Force (2023) Raporu' na göre, GPT-3 ile oluşturulan phishing (oltalama) e-postaları, geleneksel yöntemlere göre %70 daha fazka tıklanma oranına sahipti.

Bu tür e-postalar hedefin adını, şirketteki rolünü ve hatta özel projelerini dahi içeriyordu.

Yine 2023 yılından bir örnek verecek olursak, bir kripto para borsasının sahte destek sitesi, ziyaretçilere yapay zekalı bir chatbot aracılığıyla "hesap kurtarma" desteği sundu. Chatbot, kullanıcının cüzdan bilgilerini "kimlik doğrulama için" talep ederek birçok kişinin kripto varlıklarını çaldı.

Otomatik sosyal mühendislikte chatbotlar : Chatbotlar insanları kandırmak, manipüle etmek veya bilgi toplamak amacıyla kullanıldığında otomatik sosyal mühendislik aracı olarak değerlendirilir.

- Gerçek insan gibi konuşurlar : Chatbotlar doğal dil işleme (NLP) teknolojileri sayesinde kurbanla gerçek bir insan gibi iletişim kurabilir. Bu da güven oluşturur.
- Bilgi toplama yeteneği : Kurbanın verdiği cevaplara göre daha fazla bilgi toplamaya yönelik akıllı sorular sorabilirler.
- Hedefe yönelik uyarlanabilirlik : Kurbanın yaşına, konumuna, ilgi alanlarına veya şirketine göre kişiselleştirilmiş konuşmalar yapabilir.
- Sahte Destek / Müşteri hizmeti olarak kullanım : Özellikle sahte banka, sosyal medya veya e-ticaret sitelerinde çıkan yapay zeka destekli chatbotlar, kullanıcıları giriş bilgilerini vermeleri için ikna eder.

- Sosyal medya otomatik mesajlar : Botlar, sosyal medya üzerinden otomatik olarak mesaj atarak kişisel veri toplamaya veya phishing (oltalama) linkleri yaymaya çalışabilir.

Peki bu saldırılardan nasıl korunuruz ?

- Çok faktörlü kimlik doğrulama (MFA) kullanmak : Çok faktörlü kimlik doğrulama, hesap girişi sırasında kimliğinizi doğrulamak için ekstra katmanlar ekler. Bu faktörler parmak izi veya yüz tanıma gibi biyometrik özellikleri veya kısa mesaj yoluyla gönderilen geçici şifreleri içerebilir.
- E-posta ve mesajlarda dil analizine dikkat etmek (fazla “doğal” ve samimi olabilir)
- Kaynağı doğrulanmamış dosyaları açmamak
- Personeli sosyal mühendislik ve AI farkındalığı konusunda eğitmek
- Antiphishing yazılımlarında AI tabanlı tespit sistemlerini kullanmak
- Güçlü parolalar (ve bir parola yöneticisi) kullanın. Parolalarınızın her biri benzersiz ve karmaşık olmalıdır. Büyük harf, sayılar ve semboller dahil olmak üzere çeşitli karakter türlerini kullanmaya çalışın. Ayrıca, mümkün olduğunda daha uzun parolaları da tercih etmek isteyebilirsiniz. Tüm özel parolalarınızı yönetmenizi yardımcı olması için, güvenle saklamak ve hatırlamak adına bir parola yöneticisi kullanmak isteyebilirsiniz.

Deep Fake Operatörü :

DeepFake saldırılarında, siber suçlular yapay zeka tekniklerinin yardımıyla gerçekmiş gibi görünen sahte resimler, sesler, video ve audio dosyaları üreterek başka birinin örneğin patronun yerine geçebilmekte veya bu materyalleri ortalama saldırılarında kullanarak çalışanlar tuzağa düşürebilmektedir. Kimlik sahtekarlığı, CEO dolandırıcılığı (CEO Fraud), şantaj, sahte haber üretimi yapılmaktadır.

2019' da İngiltere' de bir CEO, yapay zeka ile oluşturulmuş sahte bir sesli arama sonucunda 243.000' ı Almanya' daki bir dolandırıcı hesabına aktardı. (Wall Street Journal)

2021 yılında Hong Kong'da DeepFake video konferans aracılığıyla bir şirketten 35 milyon dolar çalındı. Suçlular, CFO'yu taklit ederek çalışanlara ödeme talimatı verdi.

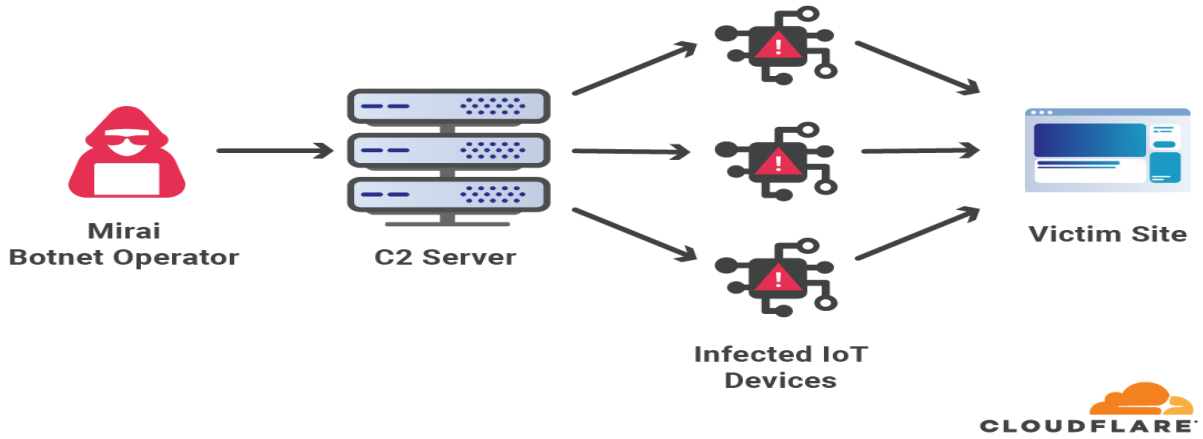
AI Destekli Hacker:

Siber suçlular bu yöntemde sistem açıklarını tespit etmek için yapay zeka tabanlı güvenlik açığı tarayıcıları kullanır. Ayrıca davranış tabanlı saldırı stratejileri de geliştirirler. Manuel olarak güvenlik açıklarını tarama işlemleri yapmak yerine AI ile otomatik saldırılar planlar ve uygulayabilirler. IBM tarafından geliştirilen DeepLocker, yalnızca belirli koşullar sağlandığında zararlı hale gelen ve AI ile gizlenen bir zararlı yazılım örneğidir.

Botnet Operatörü :

Yapay zeka destekli botnet'ler, merkezi olmayan, kendi kendine öğrenen zararlı ağlar oluşturur. Bu botllar, bir komut beklemeye gerek duymadan hedef seçebilir. DDoS saldırıları, spam yayma, veri madenciliği gibi görevleri daha verimli şekilde gerçekleştirir.

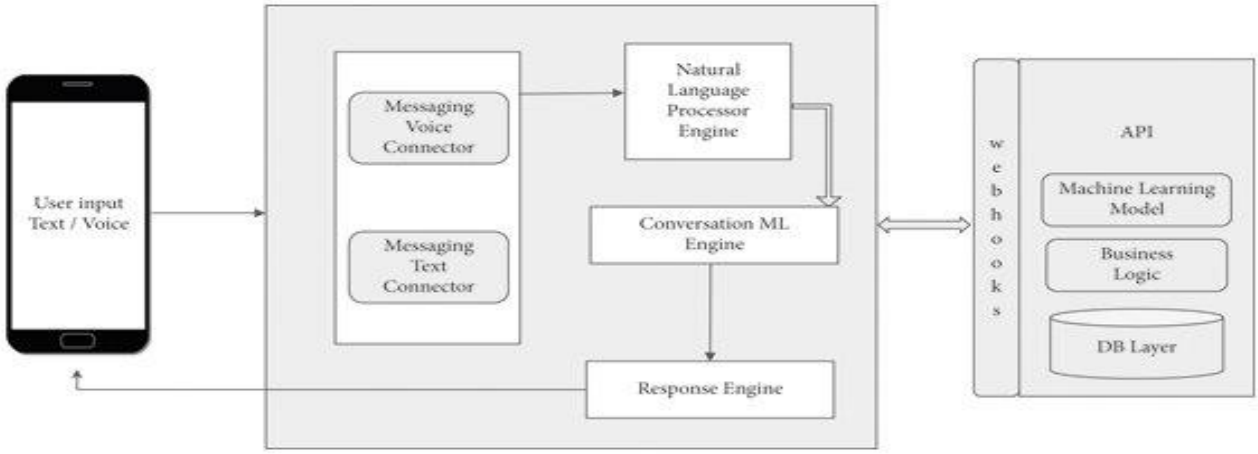
Darktrace 2022 raporounda, AI destekli botnet'lerin saldırı vektörlerini kendi kendine optimize edebildiği örnekleri tespit etmiştir.



Chatbot Tabanlı Sahte Destek Sistemleri :

Siber suçlular, bu yöntemde sahte müşteri hizmetleri sayfalarında AI tabanlı chatbotlar kurarak kullanıcıları yönlendirir. Kredi kartı bilgileri, şifreleri çalmak veya sosyal mühendislik yapmak için kullanır.

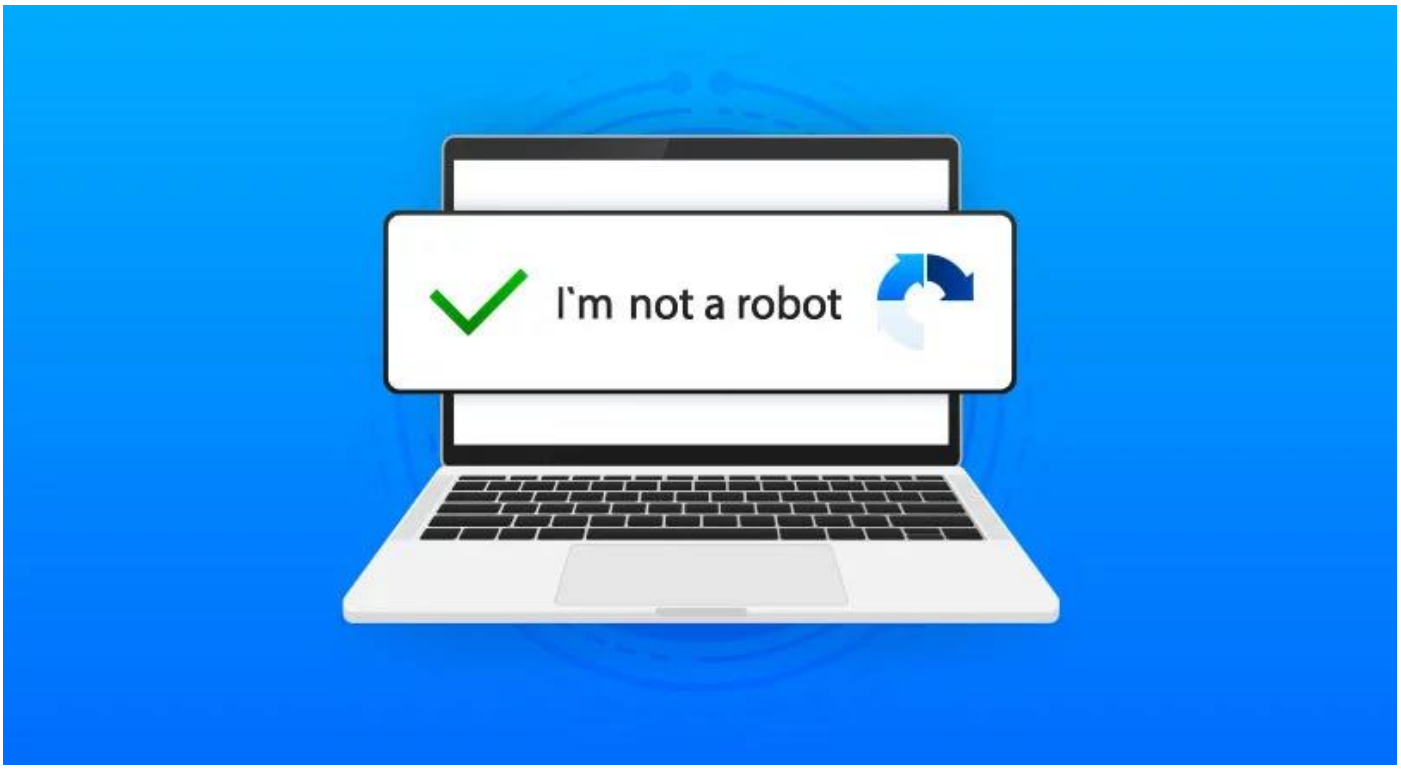
2023'te, bazı kimlik avı saldırılarında kullanıcılar sahte web sitelerine yönlendirilmiş ve bu sitelerde canlı destek botları, kullanıcıyı zararlı bağlantılara tıklamaya ikna etmiştir.



AI Tabanlı Captcha Çözücüler

Captcha Completely Automated Public Turing test to tell Computers and Humans Apart kelimelerinin kısaltmasıdır. Web tabanlı uygulamalarında uygulamayı çalıştıranın bir insan mı yoksa bir bot mu olduğunu tespit etmek için kullanılan bir güvenlik uygulamasıdır. Bundan dolayı her Captcha testinin çözülmesi yapay zekâ adına atılmış önemli bir adımdır.

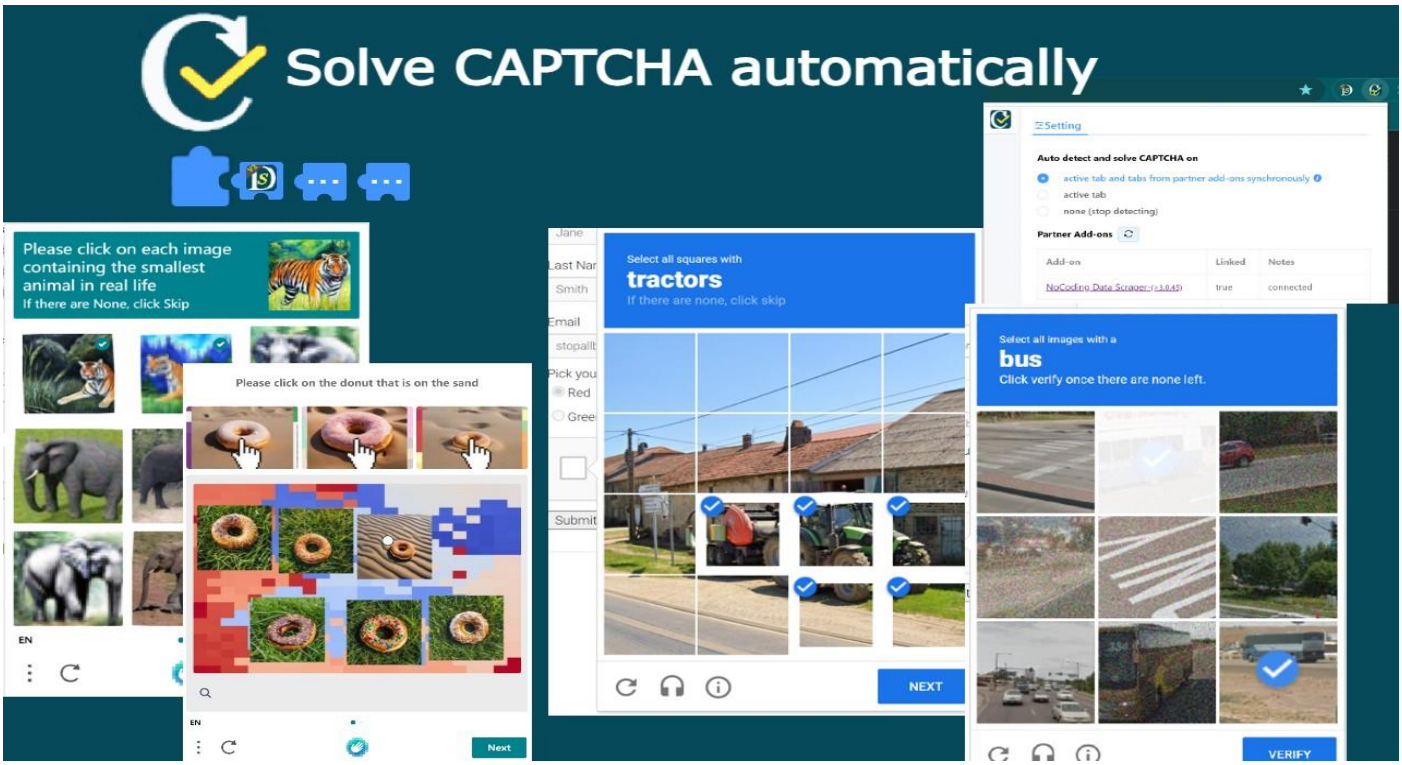
Sistemin çalışması kısaca şöyledir. Sunucu rastgele bir resim oluşturur, istemci tarafındaki kişiden bu resimdeki yazıyı okuyup, ilgili alana girmesi istenir. Buradaki basit mantık o resimde sadece insan tarafından okunabilecek, bir program tarafından okunması zor olan bir kelime oluşturmaktır. OCR (optik karakter okuyucu) programları, düzgün bir formda yazılmış yazıyı bir resmin içinden okuyabilirler. Bu sebeple sunucu tarafından üretilen resmin içindeki yazının bir insan tarafından okunabilecek fakat OCR programları tarafından okunamayacak kadar zor, karışık, anlaşılmaz, gürültülü ve bozulmuş olması gerekir. Gerçekten de üretilen karakterlerin bazen okunamaması veya çok gelişmiş OCR programları tarafından bazen resimlerin okunabiliyor olması gibi zayıflıklarına rağmen, DYPM (Did You Pass Math, “3 artı sekiz kaç eder?” gibi sorular yönelterek otomatik yazılımları engelleyen.) veya resimler arasından duygusal bazı kriterlerle seçim yapılmasını isteyen v.b. gibi alternatif yöntemlerden şu an için daha kullanışlı olmaya devam etmektedir.



Captcha Kırılması

Captcha'yı kırmada önemli olan şey mesajda ne yazdığını okumak değildir, bunu insanların en az yüzde sekseni zaten yapabilir. Asıl zor ve önemli olan insanların bu işlemi nasıl yaptığını bilgisayara öğretmektir. Captcha'yı kırmaya niyetlenen programcı, bu problemi fazlar şeklinde değerlendirmelidir. Ona uygun algoritma tasarlamalıdır. Bu senaryoya göre,

ilk adım resmi siyah beyaz tonlara çevirmek olabilir. Yani uygulama, resimdeki tüm renkleri ortadan kaldırmalı. Böylece Captcha'nın koyduğu engellerden biri olan renklerden kurtulmuş olunur. Sonra, algoritmanın diğer adımı bilgisayara ortaya çıkan şekildeki dokuları tespit etmesini emretmek olacaktır. Yani program her harfi gerçek harflerle bir-bir kıyaslayıp her harfin en çok hangi harfe benzediğini bulmak olacaktır.Bu resimler php, asp, cgi, perl gibi programlama dilleri ile dinamik oluşturuluyor ve içinde bulunan metin, oturum verileri (session data) ile yine kullanıcıya özgü olmak kaydı ile sunucuda tutuluyor. Düzgün yapılandırılmış bir sunucudan bu resimleri oluşturan kodları veya oturum bilgilerini alamayacağınız düşünülürse geriye sadece resim işleme (image processing-yapay zeka teknikleri) yöntemleri kalıyor. İşte bu resim işleme metotlarını da engelleyebilmek için resimlerin arka planlarında yazıların okunmasını zorlaştıracak, renk dağılımları ya da karakter olmayan simgeler bulunur. Resim işleme kodu eğer bu yanıltıcı simgeleri de temizleyebilecek şekilde geliştirilmişse ki bu durum yapay sinir ağları konusuna girer, yani ancak bir yapay zeka deneme yanılma (ya da öğrenme) metodu ile bunları ayıklar, bu sistem bu metodu aşabilir.



Captcha Ve Yapay Sinir Ağları

Carnegie Mellon Ünversitesi'nden Luis Von Ahn, Captcha'nın mucitlerinden biridir. 2006'daki bir dersinde von Ahn, Captcha ve yapay zeka arasındaki ilişki hakkında konuşmuştur. Çünkü CAPTCHA, hackerlar (spamciler) ve onların hedefleri arasındaki bariyerdir. Bu insanların Captcha 'ları kırmak için atanmış belirli bir zaman ve enerjileri vardır. Onların başarısı makinaları daha sofistike hale getirecekt*ir.* Birisi bir Captcha'yı daha kırmayı başardığı zaman yapay zeka yolunda önemli bir adım daha atılmış olacaktır.

Hacker’lar CAPTCHA'ların üstesinden gelecek yeni yollar buldukça, von Ahn gibi Bilgisayar Bilimciler daha çok yapay zeka isteyen yani Captcha türleri icat etmek zorunda *kalacaklar.* Captcha 'da geriye doğru atılan bir adım, yapay zeka alanında ileriye doğru atılmış bir adım demektir. Her yenilgi bir *zaferdir!.* Captcha tasarımcıları da epey yol katetmeleri gerekecek. Bilgisayarlar daha sofistike hale geldikçe test metodları da o kadar gelişecektir. Ancak bir süre sonra o kadar karmaşık resimler üretilecektir ki bunları insanlar bile çözemeyeceklerdir. İşte bu anda bu sistem tamamen çökecektir. Metinleri eğip bükmek artık bir işe yaramayacaktır. Artık kullanıcılardan matematik işlemleri çözmeleri(integral gibi) veya kısa bir hikâye okuyup soruları cevaplamaları beklenecektir. Bunun gibi zor ve sıkıcı testlere yöneldikçe kullanıcıların da web sitelerine olan ilgisi azalacaktır. Kim bir mesaj göndermek için ikinci dereceden bir turev-integral sorusunu çözmek istesin ki? En sonunda öyle bir noktaya gelinecek ki insanlar ve makinalar herhangi bir bulmacayı aynı yoldan çözmeyi deneyeceklerdir. Eğer bu gerçekleşirse, Captcha kodları gereksiz kod yığınlarından başka bir şey olmayacaklar. O zaman gelinceye kadar Captcha 'ları kullanmaya bakalım.

Yapay zeka ile çözülen Captcha’lara birkaç örnek verelim :

- 2022’de, bir siber güvenlik forumunda CAPTCHA bypass eden bir yapay zekâ scripti satışa çıkarıldı. Sistem, görsel tabanlı CAPTCHA’ları %90 başarı oranıyla çözebiliyordu.
- **ETH Zürih'in reCAPTCHAv2'yi Aşan YZ Modeli (2024)** ETH Zürih Üniversitesi'nden araştırmacılar, Google'ın reCAPTCHAv2 sistemini %100 başarı oranıyla aşabilen bir YZ modeli geliştirdiler. Bu model, görsel tanıma ve insan benzeri davranışları taklit ederek CAPTCHA'ları etkili bir şekilde çözebiliyor.
- **Oedipus: Akıl Yürütme Tabanlı CAPTCHA Çözücü (2024) :** "Oedipus" adlı sistem, büyük dil modellerinin (LLM) akıl yürütme yeteneklerini kullanarak karmaşık CAPTCHA'ları çözmek üzere tasarlanmıştır. Bu sistem, CAPTCHA'ları daha küçük ve çözülmesi kolay adımlara bölerek %63,5 başarı oranına ulaşmıştır.
- **AkiraBot: YZ Destekli CAPTCHA Aşan Spam Botu (2024)**

AkiraBot, OpenAI'nin dil modellerini kullanarak web sitelerinin iletişim formlarını ve sohbet pencerelerini spam mesajlarla dolduruyor. Bu bot, CAPTCHA'ları aşarak 80.000'den fazla web sitesine spam göndermeyi başarmıştır.

- **IllusionCAPTCHA: Görsel İllüzyon Tabanlı Yeni CAPTCHA Türü (2025)**

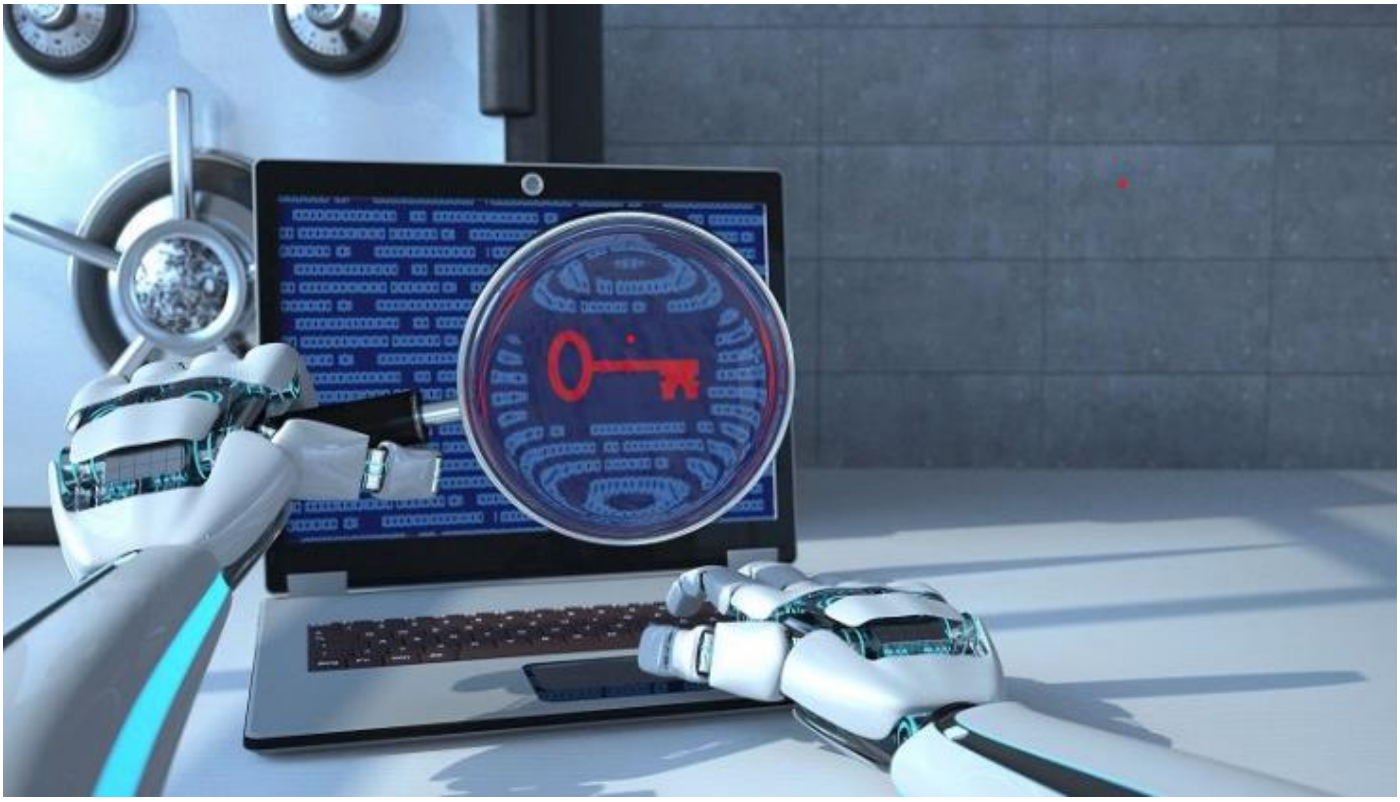
  YZ'nin CAPTCHA'ları aşma yeteneğine karşı geliştirilen IllusionCAPTCHA, görsel illüzyonları kullanarak insanlara kolay, ancak YZ sistemlerine zor görevler sunuyor. Bu yöntem, YZ'nin CAPTCHA'ları aşma başarısını %0'a düşürmüştür.

Kaynakça : CsharpNedir. (t.y.). *Yapay Zekâ ile Siber Güvenlikte Yeni Ufuklar*. CsharpNedir. https://www.csharpnedir.com/articles/read/?id=1129

AI DESTEKLİ PAROLA KIRICILAR

ChatGPT, DALL-E ve Runway gibi AI araçlarının gelişi ve daha da önemlisi hızlı ve başarılı bir şekilde benimsenmesiyle, bu tür araçların değer teklifinin geliştiricilerinin amaçladığının ötesine uzandığı giderek daha da netleşti. ChatGPT, kötü amaçlı yazılım geliştirme ve kimlik avı e-postaları ve kampanyaları oluşturma gibi kötü amaçlı görevler için zaten kullanılıyor.

Parolalar hala en popüler kimlik doğrulama yöntemidir. Doğal olarak, bu şu soruyu akla getiriyor: 'Yapay zeka destekli bir araç kullanıcı parolalarını kırabilir mi ?



Aslında bu sorunun cevabı en az altı yıldır ortalıkta dolaşıyor; yani ChatGPT'nin heyecanı (ve bir bakıma endişesi) parola üreten düşmanca ağlar veya PassGAN araştırma makalesi yayınlandığında diğer teknolojileri gölgede bırakmasından çok önce.

Makine öğrenimi tabanlı bir AI parola kırıcı olan PassGAN, parola kırma veya tahmin etme için parola analizinde manuel çabaları ortadan kaldırmak için sinir ağlarına güvenir. PassGAN makalesi mevcut parola tahmin araçları olan HashCat ve Jhon the Ripper'daki tekniklerin "pratikte iyi çalıştığını, ancak bunları daha fazla parolayı modellemek için genişletmenin özel uzmanlık gerektiren zahmetli bir iş olduğunu" belirtiyor. Bu bağlamda, PassGAN: Parola Tahmini İçin Derin Öğrenme Yaklaşımı kitabının yazarları Briland Hitaj, Giuseppe Ateniese (ikisi de Stevens Teknoloji Enstitüsü), Paolo Gasti (New York Teknoloji Enstitüsü) ve Fernando Perez-Cruz (İsviçre Veri Bilimi Merkezi), kural tabanlı ve basit veri odaklı tekniklere dayalı (Markov modelleri gibi) parola tahminini ML ile değiştirdiler.

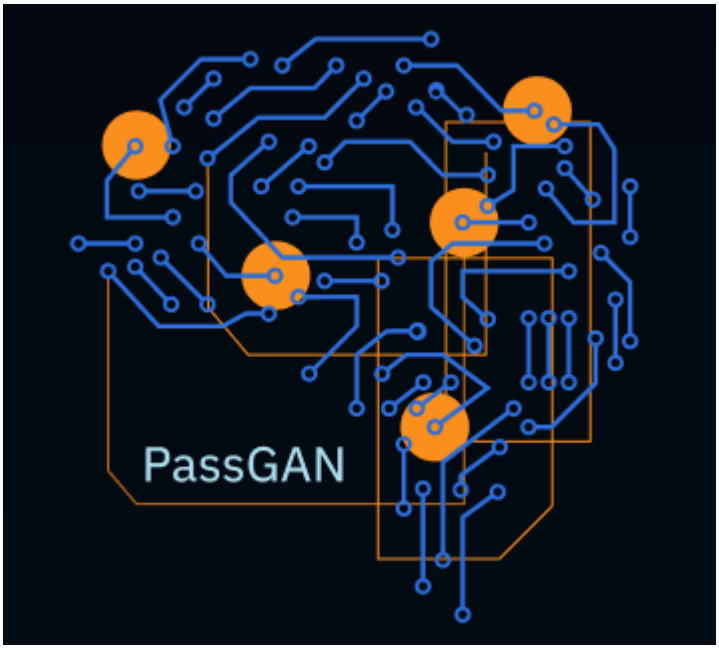
Yani asıl mesele şu değil : ‘Yapay zeka destekli bir araç kullanıcı parolalarını kırabilir mi?’ asıl mesele : ‘Yapay zeka tabanlı araçların parolaları kırması ne kadar sürer?’

Teksas merkezli siber güvenlik girişimi Home Security Heroes bu soruyu yanıtlamak için araştırma yaptı. Şirket, 2009’da sızdırılan RockYou veri kümesinden 15.680.000 parola üzerinde PassGAN’ı eğitti. Home Security Heroes (HSH) şunları keşfetti.

- Yaygın parolaların %51’i PassGAN ile bir dakikadan kısa sürede kırılabilir.
- Yaygın parolaların %65’i bir saatten kısa sürede kırılabilir.
- Yaygın parolaların %71’i bir günden kısa sürede kırılabilir.
- Yaygın parolaların %81’i bir aydan kısa sürede kırılabilir.

Kaynakça : Spiceworks. (2024, Şubat 6). *PassGAN AI cracks passwords in minutes—here’s how to protect yourself*. Spiceworks. https://www.spiceworks.com/tech/artificial-intelligence/news/passgan-ai-password-cracking-time/

PassGAN NEDİR ?



PassGAN, parola kırma tekniklerinde endişe verici bir ilerlemeyi temsil ediyor. Bu son yaklaşım, gerçek parola sızıntılarından gerçek parolaların dağıtımını otonom şekilde öğrenmek için Üretken Çelişkili Ağ’ı (GAN) kullanıyor ve manuel parola analizine olan ihtiyacı ortadan kaldırıyor. Bu, parola kırmayı daha hızlı ve daha verimli hale getirirken, çevrimiçi güvenliğiniz için ciddi bir tehdit oluşturuyor.

PassGAN, birden fazla parola özelliği üretebilir ve tahmin edilen parolaların kalitesini iyileştirebilir, böylece siber suçluların parolalarınızı kırmasını ve kişisel verilerinize erişmesini

kolaylaştırır. Bu nedenle, kendinizi bu tehlikeli teknolojiden korumak için parolalarınızı düzenli olarak güncellemeniz çok önemlidir.

HSH'nin PassGAN testi, sayılar, küçük ve büyük harfler ve semboller içeren yedi karakterli herhangi bir parolanın altı dakikadan kısa sürede kırılabileceğini ortaya koydu. PassGAN için parola tahmin süresi, sayılar, küçük ve büyük harfler ve semboller içeren sekiz ve dokuz karakterli bir parola için sırasıyla yedi saate ve iki haftaya çıkar.

PassGAN Nasıl Çalışır ?

PassGAN' ın nasıl çalıştığını anlamak için birçok modern parola tahmin aracının arkasındaki çerçeveyi incelemek önemlidir. Genellikle, parola tahmin araçları basit veri odaklı teknikler kullanarak çalışır. Bu, manuel parola analizleri çalıştıran veri modelleri uyguladıkları anlamına gelir. Ayrıca, araçlar parola kalıpları hakkında daha fazla varsayımda bulunur ve birleştirme gibi parola oluşturma kullarını kullanır.

Bu tür stratejiler kullanarak parolaları tahmin etmek, küçük ölçekli ve tahmin edilebilir parolalar için nispeten etkilidir. Ancak, örneklem boyutu büyük olduğunda ve karmaşık parola kalıpları söz konusu olduğunda, bu araçlar ya çok yavaş hale gelir ya da güvenlik kodlarını kırmak için tamamen yetersiz kalır. PassGAN gibi sistemler burada devreye girer.

PassGAN, "Password" ve "Generative Adversarial Networks" (GAN) kelimelerinin kısaltılmış halidir. GAN, bu parola kırma aracını çalıştıran genel mekanizmadır. Özünde, mekanizma bir sinir ağı üzerinde çalışır.

Sinir ağları, makineleri insan zihni gibi verileri yorumlama ve analiz etme konusunda eğiten sistemlerdir. GAN'ın sinir ağları, çeşitli özellikleri ve yapıları kaydetmek üzere tasarlanmıştır. Teknoloji, akıllı sistemleri parola analizi konusunda eğitmek için kullanılan bir veri grubu olan RockYou veri kümesi kullanılarak eğitildi. Eğitimden sonra GAN, sinir ağı dağılımını takip eden yeni örnek parolalar oluşturmak için edinilen bilgiyi kullanabildi.

Parolalarınızı Nasıl Koruyabilirsiniz ?

Parolalarınızın bütünlüğünü korumak çok önemlidir. Ne yazık ki, birçok parola veritabanı sızıntısı insanların daha güvenli olanlardan ziyade daha basit, daha kolay hacklenebilen parolalar kullanma eğiliminde olduğunu ortaya koydu. Parolanızın bir hack'e karşı koyacak ve sizi güvende tutacak kadar güçlü olduğundan nasıl emin olabilirsiniz? Mevcut en iyi parola koruma çözümlerinden bazılarını inceleyelim.

Parola gücü, kolayca hacklenebilen bir parola ile güvenli bir parola arasındaki temel farktır. PassGAN'da parola örneklerini çalıştırdığımızda elde ettiğimiz verilere göre, on karakterden oluşan yalnızca rakamdan oluşan bir parola anında hacklenebilir. Sadece küçük harflerden oluşan on harfli bir parolayı çözmek bir saat sürerken, on harfli karışık harfli bir parolayı

çözmek dört hafta sürer. Öte yandan, harfler, semboller ve sayılar kullanan on karakterli güçlü bir parolayı çözmek beş yıl sürer.

Bu, parolanız ne kadar güçlüyse, insanların veya AI sistemlerinin onu çözme olasılığının o kadar düşük olduğu anlamına gelir. Parola gücünüzü tehlieye atmanın zor olmasını sağlayan faktörler şunlardır;

- En az 15 karakter kullanın.
- Parolanızda en az iki harf (büyük ve küçük harf), rakam ve sembol bulunmalıdır.
- Gerekli tüm karakter uzunluklarına ve türlerine sahip olsalar bile, belirgin parola kalıplarından kaçının.

# OF CHARACTER	Numbers Only	Lowercase Letters	Lowercase Upper & Letters	Numbers, Upper & Lowercase Letters	Numbers, Upper & Lowercase Letters, Symbols
4	Instantly	Instantly	Instantly	Instantly	Instantly
5	Instantly	Instantly	Instantly	Instantly	Instantly
6	Instantly	Instantly	Instantly	Instantly	4 Seconds
7	Instantly	Instantly	22 Seconds	42 Seconds	6 Minutes
8	Instantly	3 Seconds	19 Minutes	48 Minutes	7 Hours
9	Instantly	1 Minutes	11 Hours	2 Days	2 Weeks
10	Instantly	1 Hours	4 Weeks	6 Months	5 Years
11	Instantly	23 Hours	4 Years	38 Years	356 Years
12	25 Seconds	3 Weeks	289 Years	2K Years	30K Years
13	3 Minutes	11 Months	16K Years	91K Years	2M Years
14	36 Minutes	49 Years	827K Years	9M Years	187M Years
15	5 Hours	890 Years	47M Years	613M Years	14Bn Years
16	2 Days	23K Years	2Bn Years	26Bn Years	1Tn Years
17	3 Weeks	812K Years	539.72M Years	2Tn Years	95Tn Years
18	10 Months	22M Years	7.23Bn Years	96Tn Years	6Qn Years

Hesabınızın güvenliğini sağlamanın en önemli yollarından biri, parolanızı her 3 ila 6 ayda bir değiştirmektir. Birinin hesabınıza eriştiğinden veya parolanızı, parolaya sahip olmaması

gereken biriyle paylaştığınızdan şüpheleniyorsanız, herhangi bir güvenlik ihlalini önlemek için parolanızı hemen değiştirmelisiniz. Tüm hesaplarınızda aynı parolayı kullanmak çok riskli olabilir. Birisi parolayı ihlal ettiğinde, birden fazla hesaba kolayca erişebilir. Bununla başa çıkmanın en iyi yolu, her hesabını için yeni parolalar üretmenizdir.

Bu nedenle, parola kırmada ‘AI’ bileşeninin etkinliği açık olsa da, çoğunlukla keşfedilmemiş olarak kalır. Örneğin, bir AI aracı bir kullanıcının parolasını, sosyal medyadaki herkese açık

profiline ve paylaşımlarına dayanarak başarılı ve doğru bir şekilde tahmin ederse, Bu bir başarı olurdu.

Kaynakça : SecurityHero. (2024, Ocak 30). *AI password cracking: How secure are your passwords?* SecurityHero. https://www.securityhero.io/ai-password-cracking/

Yüksek Performanslı GPU'larla Parola Kırma

Son yıllarda, yüksek performanslı grafik işlem birimlerinin (GPU'lar) siber güvenlik alanında nasıl kullanıldığına dair yaptığım araştırmalar, parola kırma tekniklerinde özellikle ciddi bir hızlanma yaşandığını ortaya koydu. Eskiden brute-force ya da dictionary attack gibi yöntemlerle saatler süren şifre kırma işlemleri, günümüzde gelişmiş GPU'lar sayesinde saniyeler içinde tamamlanabiliyor. Bu durumun en dikkat çekici yönlerinden biri, saldırganların artık sadece teknik uzmanlar değil; kolayca erişilebilen yazılımlar ve donanımlar sayesinde ortalama düzeyde bilgiye sahip kişilerin de karmaşık sistemlere saldırabiliyor olmasıdır.

Dark Web üzerinde dolaşan 15 milyardan fazla kullanıcı adı ve şifre kombinasyonu olduğunu görmek, aslında bireysel güvenlik farkındalığının ne kadar eksik olduğunu da gösteriyor. Bu verilerle yapılan saldırıların %60'ından fazlasının kimlik bilgisi ihlali yoluyla gerçekleştiği bilgisi, konunun ne kadar ciddi olduğunu gözler önüne seriyor. Özellikle sekiz karakterli zayıf parolaların bazı GPU destekli parola kırma araçları tarafından milisaniyeler içinde çözülebilmesi, çok faktörlü kimlik doğrulamanın (MFA) artık bir seçenek değil, zorunluluk olduğunu açıkça ortaya koyuyor.

Bu araştırmayı yaparken dikkatimi çeken bir diğer konu da, bu teknolojilerin etik dışı amaçlarla kullanımının ne kadar kolay erişilebilir hale geldiği oldu. Hacker forumlarında paylaşılan "RTX 4090 ile parola kırma testleri" gibi içerikler, bu tehdidin yalnızca teoride kalmadığını, uygulamada aktif olarak kullanıldığını da kanıtlar nitelikte. Bu nedenle güçlü parolalarla birlikte, yapay zekânın siber güvenlik tarafında oluşturduğu risklerin de her geçen gün daha ciddi alınması gerektiğini düşünüyorum.

PassGAN, yapay zekâ temelli parola kırma araçlarının en bilinen örneğidir. Ancak son yıllarda benzer amaçla geliştirilen birçok başka araç da literatürde yerini almıştır. Örneğin, **John the Ripper** ve **Hashcat** gibi klasik araçlara entegre edilen yapay zekâ modülleri, şifre tahmin sürecini ciddi anlamda hızlandırmaktadır. Ayrıca, **TarGuess** gibi sistemler kullanıcı alışkanlıklarını analiz ederek daha hedefe yönelik saldırılar üretebilmektedir. Bu araçlar, yapay zekânın sadece hız değil, aynı zamanda saldırı stratejisi geliştirme noktasında da aktif rol oynadığını göstermektedir.**PassGAN Benzeri Yapay Zeka Destekli Parola kırma araçları**

Jhon the Ripper + AI Modülleri : John the Ripper, uzun süredir kullanılan açık kaynaklı bir parola kırma aracıdır. Son yıllarda bazı geliştiriciler tarafından yapay zekâ modülleriyle

(özellikle LSTM tabanlı tahmin sistemleri) entegre edilerek, parolaların olasılık dağılımlarına göre daha mantıklı ve daha kısa sürede çözülmesi sağlanmıştır.

Hashcat + Rule-Based AI Enhancements

Hashcat, GPU destekli en popüler parola kırıcılarından biridir. AI destekli parola üretim kuralları (örneğin: önceki kırılan parolalara göre olasılık modelleme) entegre edilerek saldırı hızı ve başarı oranı artırılabilir. Örneğin, daha önce kırılan 10.000 parolaya göre yeni parola kombinasyonları üretmek için istatistiksel modelleme kullanılır.

TarGuess Framework

Tarsec (Targeted Security Company) tarafından geliştirilen bu sistem, kullanıcının sosyal medya bilgileri, e-posta adresi ve alışkanlıkları gibi dış kaynaklı verileri yapay zekâ ile analiz ederek, kişiselleştirilmiş parola tahminleri üretir. Bu sistem, sosyal mühendislik temelli yapay zekâ destekli parola saldırılarına örnek olarak gösterilebilir.

ARAÇ ADI	**Yapay Zeka Türü / Teknolojisi**	**Kullanım Alanı**	**Öne Çıkan Özellik**
PassGAN	GAN (Generative Adverserial Network)	Parola tahmini / Üretimi	Sızdırılmış verilerden öğrenerek yeni ve benzer parolalar üretir.
Jhon the Ripper + AI	LSTM / AI Kuralları	Geleneksel Brute-Force'un iyileştirilmesi	AI ile daha etkili parola tahmini kuralları geliştirilir.
Hashcat + AI Rules	Kural tabanlı AI / İstatistiksel Modelleme	GPU ile hızlı Brute-Force	En sık kullanılan parola yapılarını öğrenerek saldırı optimizasyonu sağlar.
TarGuess	NLP / Veri Madenciliği / Profil Analizi	Kişisel verilere dayalı hedefli saldırılar	Kullanıcının sosyal verilerini analiz ederek özel tahmin.
DeepPass	RNN / LSTM	Parola listelerinden derin öğrenme ile tahmin	Geçmiş parola örneklerinden yeni tahminler üretir.

Sıfır Gün Açığı

Sıfır gün istismarı, bilgisayar yazılımı, donanımı veya donanım yazılımındaki bilinmeyen veya ele alınmamış bir güvenlik açığından yararlanan bir siber saldırı vektörüdür. “Sıfır gün”, yazılım veya cihaz satıcısının açığı düzeltmek için sıfır günü olması gerçeğini ifade eder çünkü kötü niyetli aktörler bunu savunmasız sistemlere erişmek için zaten kullanabilir.

Bilinmeyen veya ele alınmamış güvenlik açığına sıfır günlük güvenlik açığı veya sıfır günlük tehdir denir. Sıfır günlük saldırı, kötü niyetli bir aktörün kötü amaçlı yazılım yerleştirmek, veri çalmak veya kullanıcılara, kuruluşlara veya sistemlere başka türlü zarar vermek için sıfır günlük bir istismarı kullanmasıdır.

Benzer ancak ayrı bir kavram olan sıfırıncı gün kötü amaçlı yazılım, imzası bilinmeyen veya henüz mevcut olmayan ve bu nedenle birçok antivirüs yazılım çözümü veya diğer imza tabanlı tehdit algılama teknolojileri tarafından tespit edilemeyen virüs veya kötü amaçlı yazılımdır.

Bir işletim sistemi, uygulama veya cihazın bir sürümünde, piyasaya sürüldüğü andan itibaren sıfır günlük bir güvenlik açığı mevcuttur, ancak yazılım satıcısı veya donanım üreticisi bunu bilmez. Güvenlik açığı, birisi bulana kadar günler, aylar veya yıllar boyunca tespit edilemeyebilir.

SIFIR GÜN AÇIĞI İÇİN AI :

Yapay zeka tabanlı sistemler, son dönemde sıfır gün (zero-day) açıklarıın keşfi ve tespiti konusunda önemli ilerlemeler kaydetmiştir. Bu gelişmeler, hem savunma hem de saldırı tarafında siber güvenlik dinamiklerini yeniden şekillendirmektedir.

Google'ın "Big Sleep" Aracı ile Sıfır gün açığı keşfi : Big sleep google tarafından geliştirilen ve kasım 2024'te yaygın olarak kullanılan açık kaynaklı bir veritabanı motoru olan SQLite'ta otonom olarak bir sıfır gün güvenlik açığı keşfeden çığır açan bir yapay zeka aracı. Bu, gerçek dünya yazılımında bilinmeyen bir istismar edilebilir bellek güvenliği sorununu tanımlayan yapay zekanın ilk örneği oldu. Bir yığın arabelleği alt akışı olan güvenlik açığı, gelişmiş varyant analizi teknikleri kullanılarak tespit edildi ve SQLite geliştirme ekini tarafından hızlı bir yamaya yol açtı. Big Sleep, yapay zekanın güvenlik açıklarını istismar edilmeden önce proaktif olarak belirleyerek siber güvenliği geliştirme potansiyelini örneklemektedir.

Google Big Sleep' in ötesinde, güvenlik testlerinde, özellikle de yazılım hatalarına neden olmak için rastgele veri girişi uygulaması olan bulanıklaştırma yöntemlerini iyileştirmede yapay zeka kullanımlarını genişletiyor. Yakın zamanda sunulan yapay zeka destekli bulanıklaştırma sistemi, programa özel test kodu üretir ve testleri otomatik olarak çalıştırır.

Bu gelişme, özellile geleneksel teknikler kullanılarak bildirilmeyebilecek karmaşık yazılım güvenlik açıklarını bulmak için siber güvenlik araçlarında yapay zeka entegrasyonuna yönelik artan bir eğilime işaret ediyor.

İşletmeler Neden Önemsemeli

Bu atılım, işletmelerin sistemlerini koruma biçiminde büyük bir değişimi temsil ederek daha hızlı sorun tespiti ve gerçek dünya riskinin önlenmesine olanak tanır. Bir veri ihlalinin ortalama maliyeti 2023' te 4.45 milyon dolara ulaştı ve erken hasarından ve acil BT müdahale masraflarından kaçınmasına yardımcı oluyor.

İşletmelerin Yapması Gerekenler :

- Sistemlerin Denetimi
- Güvenlik Protokollerinin sürekli olarak güncel tutulması
- Gelişmelerden haberdar olunması
- Müşteri verileri için önceden önlem
- Güvenlikle ilgili maliyetlerin azaltılması
- Müşteri güveni

MixMode: Gerçek Zamanlı Sıfır Gün Saldırı Tespiti

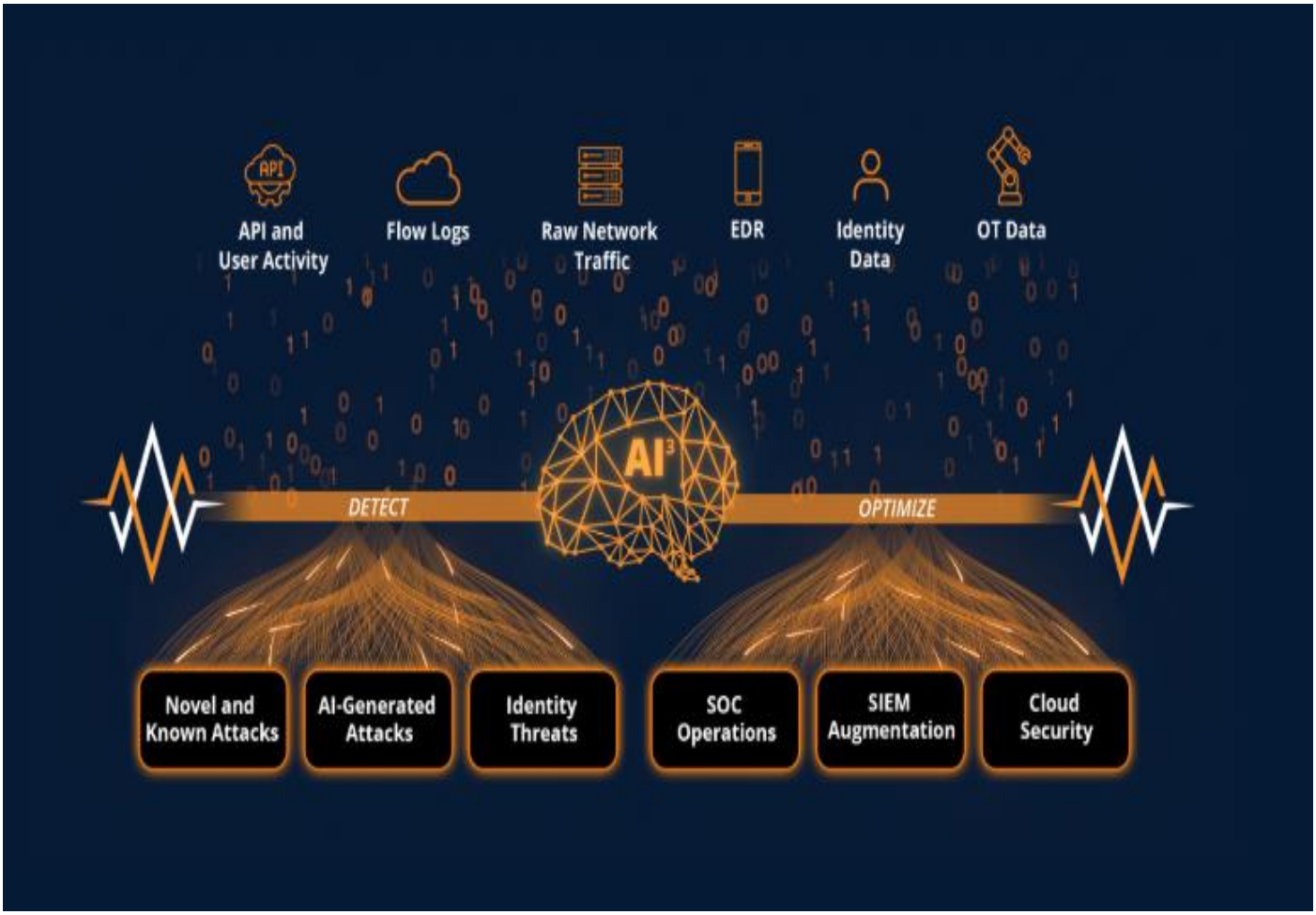
MixMode, en son nesil yapay zekayla çalışan, kendi kendine öğrenebilen bir siber güvenlik sistemi. Yani bu sistem, ağınızda neler olup bittiğini zamanla kendi başına öğreniyor ve olağandışı bir şey fark ettiğinde sizi uyarıyor. Bunu yaparken de insan müdahalesine ihtiyaç duymadan, sizin sisteminize özel olarak gelişiyor.

Eski güvenlik araçları gibi sadece belirli kurallara bağlı kalmka yerine, MixMode her an yeni şeyler öğreniyor ve değişen tehditlere ayak uydurabiliyor. Böylece ekibinize gereksiz uyarılarla yük olmak yerine, gerçekten önemli şeyleri tespit ediyor ve zaman kazandırıyor.

Peki MixMode Zero-day Açıklarıyla Ne Alakalı?

Zero-day (sıfır gün) açıkları, yazılım geliştiricilerinin henüz farkında bile olmadığı güvenlik açıklarıdır. Bu yüzden bu tür açıklar, siber saldırganlar için adeta bulunmaz nimettir. İşte bu noktada MixeMode'un farkı ortaya çıkıyor.

MixMode, sıradan saldırılara değil, henüz keşfedilmemiş tehditlere karşı da savunma sağlayabiliyor. Çünkü sistem, yalnızca geçmişteki saldırı örneklerine bakarak değil, ağda olup bitenleri analiz ederek kendi kararını verebiliyor. Yani biri sisteminize fark ettirmeden girmeye çalışsa bile, MixMode o davranıştaki garipliği farkediyor ve sizi hemen uyarıyor.



ZeroPath: Otomatik Sıfır Gün Açığı Keşfi

ZeroPath nedir: ZeroPath, yazılımınızdaki güvenlik açıklarını tespit etmek için LLM'lerden yararlanan bir Statik Uygulama Güvenlik Testi (SAST) aracıdır. Güvenlik mühendisleri ve başarılı hata ödül avcıları tarafından geliştirilen bu araç, iş mantığı kusurları, kimlik doğrulama açıkları ve diğer yaygın zayıflıklar dahil olmak üzere geleneksel yöntemlerle sıklıkla gözden kaçan karmaşık sorunları belirler. ZeroPath, hata ödül programlarının kapsadığı alanlar da dahil olmak üzere büyük açık kaynaklı depolardaki güvenlik açıklarını ortaya çıkarmak için kullanılmıştır. ZeroPath, kod bağlamını ve işlevselliğini anlayarak daha az yanlış pozitifle doğru tespit sağlamayı amaçlamaktadır.

Araç, ana kod tabanına birleştirilmeden önce kodu analiz ederek birleştirme öncesi tarama gerçekleştirir. Çekme isteği (PR) taramaları genellikle bir dakikadan kısa sürede tamamlanır ve geliştirme iş akışlarını engellemeden güvenlik sorunlarının erken tespit edilmesini sağlar.

ZeroPath şu anda JavaScript, Python, Go, Java, C# ve PHP'yi destekler. Abstract Syntax Tree (AST) üretimi için treesitter kütüphanesini kullanır ve ek veya özel dil gramerlerinin hızlı entegrasyonuna olanak tanır.

Kısaca ZeroPath bozuk kimlik doğrulamayı, mantık hatalarını, güncel olmayan bağımlılıkları ve daha fazlasını bulmak ve düzeltmek için yapay zeka destekli kod güvenlik açığı taramasıdır.

ZeroPath, yapay zekanın siber güvenlikte nasıl bir devrim yarattığını gösteren dikkat çekici bir örnektir. 2023 yılından itibaten araştırmacılar, büyük dil modellerini (LLM) kullanarak web sitelerindeki güvenlik açıklarını daha doğru şekilde tespit etmeye başladılar. Bu modeller, daha önce de bilinen XSS, SQL injection ya da CSRF gibi açıklara odaklanmanın ötesine geçerek, çoğu zaman gözden kaçan mantık hatalarını ve kimlik doğrulama zafiyetlerini de fark edebiliyor.

Asıl dikkat çeken gelişme ise 2024 yazında oldu. ZeroPath, yapay zekayı adeta bir güvenlik uzmanı gibi kullanmaya başladı klasik güvenlik araçlarının göremediği zayıf noktaları bulmak için, yapay zeka ile derinlemesine kod analizleri yapmaya başladı. Özellikle sıfır gün açıkları gibi daha önce kimsenin keşfetmediği kritik güvenlik zafiyetlerini bu yöntemle ortaya çıkardı.

Yani artık yapay zeka, sadece yardımcı bir araç olmaktan çıktı; kendi başına sistemleri analiz edip açıkları bulabilen aktif araştırmacıya dönüştü. ZeroPath'in kullandığı bu yeni yöntem, açık kaynaklı birçok projede daha önce fark edilmeyen ciddi güvenlik sorunlarını da gün yüzüne çıkardı.

Kimlik Doğrulama Açığı (Authentication Bypass) : Bir açık kaynaklı içerik yönetim sisteminde, yapay zeka tabanlı analiz, kullanıcıların doğrulanmadan yönetici paneline erişebileceği bir açıklık tespit etti. Bu açık, geleneksel araçlar tarafından fark edilmemişti çünkü sıradan bir kod hatası gibi görünüyordu.

İş Mantığı Hataları : Bir e-ticaret platformunda, ödeme adımlarında yapılan bir hata sayesinde kötü niyetli bir kullanıcı sistemde ödeme yapmadan alışverişi tamamlayabiliyordu. Yapay zeka, bu iş akışındaki mantık hatasını analiz ederek fark etti. Bu tarz açıklar genelde sadece gerçek kullanıcı davranışlarıyla ortaya çıkar ve otomatik araçlarla tespiti zordur.

SQL İnjection Varyantı : AI destekli sistemler, standart SQL injection tespitlerinin dışında kalan, daha karmaşık bir varyantı buldu. Bu açık, belirli bir kullanıcı girdisiyle tetikleniyor ve veritabanına yetkisiz erişim sağlıyordu. Klasik tarayıcılar bu tür karmaşık yapıları göremediği için bu açık uzun süre fark edilmeden kalmıştı**.**

Bu örnekler, bize yapay zeka gerçekten neyi farklı yapıyor sorularına net bir şekilde yanıt veriyor. yapay zekanın sadece bilinen güvenlik testlerini tekrar etmekle kalmayıp, sistemin nasıl çalıştığını gerçekten anlayabildiğini ve bu sayede çok daha derinlemesine analiz yapabildiğini gösteriyor. ZeroPath'in bu yaklaşımı, özellikle sıfır gün açıklarının daha erken fark edilmesini sağlayarak güvenlik ekiplerine zaman kazandırıyor.

ZeroPath'in bugüne kadar tespit ettiği önemli açıklardan bir kaçı şu şekildedir:

- **Fonoster Voice Server – Yerel Dosya Dahil Etme (CVE-2024-43035) :** Fonoster sisteminde dosya yollarının doğru doğrulanmaması, saldırganların sistem içindeki kritik dosyalara erişmesine olanak sağlıyordu.
- **UpTrain – Uzaktan Kod Çalıştırma (RCE) :** Yapay zeka projelerinde kullanılan bu açık kaynak platformda, eval() fonksiyonu aracılığıyla kötü amaçlı komutların uzaktan çalıştırılması mümkündü.

- **LibrePhotos – Dosya Yolu Geçişi (Path Traversal) :** Fotoğraf yükleme sırasında eksik güvenlik kontrolleri, saldırganların dosya sistemine keyfi şekilde müdahale etmesine imkan veriyordu.
- **RagFlow – Yetkisiz Veri Erişimi (IDOR) :** Bu platformda, kullanıcılar başka kullanıcıların verilerine erişebiliyor ve silebiliyordu. Kimlik doğrulama kontrolleri eksikti.
- **Monaco (Hulu) – Uzaktan Kod Çalıştırma (CVE-2024-48946) :** Python'un pcikle modülünün güvensiz şekilde kullanımı, saldırganların sistem üzerinde tam kontrol elde etmesine neden olabiliyordu.

ZeroPath'in 2024 yılında paylaştığı güvenlik raporlarına göre birçok açık kaynak yazılımda kritik zafiyetler tespit edilmiştir. (ZeroPath, 2024).

Kaynak: ZeroPath Resmi Sitesi – https://zeropath.com/wall

VERİ GİZLİLİĞİ VE YAPAY ZEKA

Teknoloji geliştikçe hayatımız kolaylaşıyor ama beraberinde bazı riskler de getiriyor. Özellikler veri toplama ve analiz gibi konularda yaşanan ilerlemeler, kişisel bilgilerin yanlış ellere geçme ihtimalini arttırıyor. Eskiden sadece büyük şirketlerin uğraştığı veri güvenliği, artık hepimizin günlük hayatını etkileyen bir konu haline geldi.

Yapay zeka ve makine öğrenimi sistemleri, veriye aç sistemlerdir. Bu sistemler ne kadar çok ve kaliteli veriyle beslenirse, o kadar doğru sonuç verir. Ancak bu verilerin içinde kişisel ve hassas bilgiler de yer aldığında, gizlilik riskleri kaçınılmaz hale geliyor.

Devletler ve kurumlar bu sorunu çözmek için yapay zekaya özel gizlilik yasaları üzeirnde çalışıyor. Bu da, özellikle yapay zekayı aktif kullanan şirketler için yeni kurallar ve uyum süreçleri anlamına geliyor.

Tüm bu risklere rağmen, işletmeler verimliliği arttırmak ve daha büyük faydalar elde etmek için yapay zekadan vazgeçmiyor. O zaman soru şu: Bu teknolojileri kullanırken gizliliği nasıl koruyabiliriz?

Yapay Zeka Gizliliği Nedir ?

Yapay zeka gizliliği, bir yapay zeka sisteminin topladığı, işlediği, sakladığı ya da başka sistemlerle paylaştığı kişisel ve hassas verilerin korunması anlamına gelir. Bu kavram aslında "veri gizliliği" ile çok yakından ilişkilidir.

Veri gizliliği (diğer adıyla bilgi gizliliği), bir bireyin kendi verileri üzerinde söz hakkına sahip olması gerektiği ilkesine dayanır. Yani bi kurum ya da şirket, bir kişinin verilerini nasıl toplayacağına, nerede saklayacağına ve ne şekilde kullanacağına karar verirken, o kişinin haklarına da saygı göstermek zorundadır. Ancak bu anlayış yapay zeka sistemlerinin gelişmesiyle birlikte değişmeye başlamıştır.

Eskiden insanlar veri gizliliği deyince genelde akıllarına yalnızca çevrimiçi alışverişlerdeki veriler gelirdi. Örneğin "Bu site ne satın aldığımı bilsin, belki daha iyi önerilerde bulunur" gibi düşünceler yaygındı. Fakat artık şirketler, çok daha büyük veri havuzları oluşturuyor. Bu veriler sadece reklam önerisi için değil, yapay zeka sistemlerini eğitmek amacıyla da kullanılıyor.

Stanford Üniversitesi'nde İnsan Merkezli Yapay Zeka Enstitüsü üyesi Jennifer King'in belirttiği gibi, bu durum artık sadece bireysel gizlilik meselesi olmaktan çıkıyor; toplumun geneline yayılan daha geniş bir etkiden söz ediyoruz. Özellikle de temek hak ve özgürlükler üzerinde ciddi etkileri olabilecek bir sürece giriyoruz.

Bu nedenle, yapay zeka sistemlerinin gelişimi kadar, bu sistemlerin gizliliğe uygun şekilde tasarlanması da büyük önem taşımaktadır.

Yapay Zekanın Gizlilik Riskleri

- **Hassas verilerin toplanması**
- **Kullanıcı rızası olmadan veri toplanması**
- **İzinsiz veri kullanımı**
- **Denetimsiz gözetim ve önyargı**
- **Veri sızdırma**
- **Veri sızıntısı**

Hassas Verilerin Toplanması

Yapay zekanın önceki teknolojik gelişmelere nazaran büyük veri gizliliği riski oluşturmasının sebebi bilgi hacminin büyük olmasıdır. Terabaytlarca metin, resim veya video rutin halde eğitim verisi olarak kaydedilir ve bu verilerin bir kısmı hassastır. Sağlık bilgileri, sosyal medya platformları üzerinden gelen kişisel veriler, kişisel finans bilgileri, yüz tanıma için biyometrik bilgiler ve daha fazlası. Daha önce hiç olmadığı kadar hassas verinin toplanması,

depolanması ve iletilmesi, en azından bir kısmının gizlilik haklarını ihlal eden şekillerde ifşa edilmesi veya kullanılması olasılığı daha yüksektir.

Kullanıcı Rızası Olmadan Veri Toplanması

Veriler, toplandığı kişilerin açık izni veya bilgisi dahilinde olmadan AI gelişimi için alındığında sorunlar ortaya çıkabilir. Web siteleri ve platformlar söz konusu olduğunda, kllanıcılar giderek kendi ve verileri üzerinde daha fazla özgünlük ve veri toplama konusunda daha fazla hassas olunması gerektiğini beklemektedir. Bu tür beklentiler son zamanlarda ön plana çıktı, sebebi ise profesyonek ağ sitelerinden biri olan LinkedIn, bazı kullanıcıların bilgilerini üretken AI modellerini eğitmesine otomatik olarak izin verdiklerinin fark edilmesi ardından büyük bir tepkiyle karşı karşıya kaldı.

İzinsiz Veri Kullanımı

Bireylerin verileri rızalarıyla toplansa bile, bu verilerin başlangıçta belirtilen amaçlar dışında kullanılması gizlilik açısından ciddi bir tehdit oluşturur. Stanford Üniversitesi'nden Jennifer King'in de belirttiği gibi, insanlar bir amaçlar paylaştıkları verilerin – örneğin özgeçmiş yada kişisel fotoğraflar – çoğu zaman bilgileri ve onayları olmadan yapay zeka sistemlerini eğitmek için yeniden kullanıldığını fark etmektedir.

Bu duruma örnek olarak, Kaliforniya'da yaşanan bir olay gösterilebilir: Bir hasta, tıbbi tedavisi sırasında çekilen fotoğraflarının bir yapay zeka eğitim veri setinde kullanıldığını öğrenmiştir. Hasta, yalnızca doktorunun fotoğraf çekmesi için onay verdiğini, ancak bu görüntülerin bir veri setine dahil edilmesi için açıkça izin vermediğini belirtmiştir. Bu olay, verilerin ikinci kez kullanımı konusunda hem etik hem de yasal açıdan önemli soru işaretleri doğurmaktadır.

Denetimsiz Gözetim ve Önyargı

Yapay zekadan önce de kişisel gizliliği tehdit eden gözetim uygulamaları mevcuttu. Güvenlik kameraları, internet tarayıcı çerezleri ya da dijital takip sistemleri üzerinden yapılan bu gözetimler, AI teknolojileriyle birlikte daha derin bir boyuta taşındı. Artık gözetim sadece görüntü veya veri toplamayla sınırlı kalmıyor; yapay zeka sistemleri bu verileri analiz ederek sonuçlar üretmeye başlıyor.

Ancak bu durum, beraberinde yeni gizlilik ve etik sorunları getiriyor. Özellikle yapay zeka modellerinin önyargılar içerebilmesi, analizlerin sonucunu da adaletsiz hale getirebiliyor. Örneğin, emniyet birimlerinde kullanılan AI destekli yüz tanıma veya davranış analizi sistemlerinin, özellikle farklı etnik kökenlere sahip bireyler üzerinde ayrımcılığa yol açtığı vakalar yaşanmıştır. Bu gibi olaylar, hem bireysel özgürlükleri tehdit etmekte hem de toplumda güvensizlik yaratmaktadır.

Veri Sızdırma

Yapay zeka sistemleri, içinde barındırdığı veriler açısından oldukça zengin ve hassas kaynaklardır. Bu da onları siber saldırganlar için cazip hedefler haline getirir. IBM Güvenlik Uzmanı Jeff Crume'un da belirttiği gibi, yapay zeka modelleri, adeta "vurulmak istenen büyük bir hedef" gibi görülebilir. Çünkü bu sistemlerin içerisinde kişisel bilgilerden ticari sırlarla dolu belgelere kadar pek çok hassas veri bulunabilir.

Saldırganlar, bu verilere ulaşmak için çeşitli taktikler geliştirmektedir. Bunlardan biri olan istem enjeksiyonu saldırıları, oldukça dikkat çekicidir. Bu yöntem de saldırganlar, yapay zeka sistemine meşru komut gibi görünen ancak aslında zararlı olan girdiler sunar. Böylece, yapay zeka sistemine meşru komut gibi görünen ancak aslında zararlı olan girdiler sunar. Böylece, yapay zeka sistemi farkında olmadan özel ya da gizli bilgileri dışarıya aktarabilir. Örneğin, bir yapay zeka destekli sanal asistana yöneltilen ustaca hazırlanmış bir komut, sistemin şirket içi belgeleri paylaşmasına neden olabilir.

Bu tür veri sızdırma senaryoları, hem bireysel kullanıcırların hemde kurumların bilgi güvenliğini ciddi anlamda riske atmaktadır. Bu nedenle, yapay zeka sistemlerinin yalnızca yetkili erişimlere açık olması ve saldırılara karşı güvenlik önlemleriyle donatılması büyük önem taşır.

Örneğin 2023 yılında bir teknoloji firması, çalışanlarının iş süreçlerini kolaylaştırmak amacıyla ChatGPT gibi bir üretken yapay zeka sistemini dahili olarak kullanmaya başladı. Ancak kısa süre sonra fark edildi ki, bazı çalışanlar sistemle konuşurken şirket içi kod parçacıkları, müşteri bilgileri ve sözleşme detayları gibi hassas bilgileri istemler aracılığıyla doğrudan yazılı olarak paylaşıyordu.

Bu veriler, yapay zeka modeli tarafından sistemde geçici olarak işleniyor ve kimi durumlarda diğer kullanıcıların sorgularına verilen yanıtlar arasında görünmez şekilde ortaya çıkabiliyordu. Durum fark edildiğinde şirket, ciddi bir gizlilik ihlali ile karşı karşıya kaldı ve veri sızıntısını önlemek amacıyla yapay zeka kullanımını geçici olarak durdurdu. Bu olay, yapay zekaya yanlışlıkla dahi olsa hassas veri aktarımının ne kadar tehlikeli sonuçlar doğurabileceğini gözler önüne serdi.

(Bkz. Wald.ai, " ChatGPT Veri Sızıntıları ve Güvenlik Olayları 2023–2024", 2023; Forbes, "Samsung ChatGPT ve Diğer Sohbet Botlarını Yasakladı", 2023)

Veri Sızıntısı

Veri sızıntısı, hassas bilgilerin yanlışlıkla başkalarıyla paylaşılması durumudur ve bazı yapay zeka sistemlerinin bu tür ihlallere karşı savunmasız olduğu bilinmektedir. Bu duruma örnek olarak, OpenAI'nin geliştirdiği büyük dil modeli ChatGPT'nin yaşadığı bir olay verilebilir. Bazı kullanıcılar, farkında olmadan diğer kullanıcıların konuşma geçmişine ait başlıkları görebilmiştir.

Bu sadece büyük çaplı yapay zeka sistemleri için geçerli değildir; daha küçük, şirket içi kullanılan modellerde de benzer riskler söz konusudur. Örneğin bir sağlık kuruluşunun, kendi müşterilerinin verilerine dayalı bir teşhis modeli geliştirdiğini düşünelim. Bu model, belirli bir komutla çalıştırıldığında başka bir hastanın özel bilgilerini istemeden açığa çıkarabilir. Bu gibi istem dışı veri paylaşımları da ciddi gizlilik ihlallerine neden olabilir.

Gizliliğin Korunmasına Yönelik Yasal Düzenlemeler

Politika yapıcılar, teknolojik gelişmelerin bireylerin özel yaşamını tehdit etmesini engellemek için uzun süredir çalışmalar yürütüyor. Bu çabalar, 1970 li yıllara kadar uzanıyor. Ancak son yıllarda, özellikle veri toplamaya yönelik ticari faaliyetlerin artması ve yapay zeka teknolojilerinin hızla yaygınlaşması, veri gizliliğini korumaya yönelik yeni ve daha kapsamlı yasal düzenlemelere duyulan ihtiyacı arttırmıştır. Bu nedenle birçok ülke, bireylerin kişisel verilerinin nasıl toplanıp kullanıldığını kontrol altına almak amacıyla çeşitli gizlilik yasalarını hayata geçirmiştir.

Avrupa Birliği Genel Veri Koruma Tüzüğü (GDPR)

GDPR, kişisel verilerin işlenmesi sürecinde hem veri denetleyicilerinin hem de işlemcilerinin uyması gereken bir dizi ilke belirler. Bu yönetmelik, verilerin toplandığı andan itibaren belirli, açık ve yasal bir amacı olmasını zorunlu kılar. Yani, bir şirketin topladığı veriler, yalnızca belirli bir amaçla kullanılmalıdır ve bu amaç, kullanıcılarla açıkça paylaşılmalıdır. Ayrıca, yalnızca gereken minimum miktarda veri toplanmalıdır.

Verilerin kullanımında ise adalet ön planda tutulmalıdır. Şirketler, kullanıcıları kişisel verilerinin nasıl işlendiği konusunda bilgilendirmekle yükümlüdür. Ayrıca, verilerin saklanma süresi de sınırlıdır; yani bir şirket, verileri yalnızca işlem amacı yerine getirilene kadar saklamalıdır. Eğer veriye artık ihtiyaç duyulmazsa, o veri güvenli bir şekilde silinmelidir.

AB Yapay Zeka (AI) Yasası

Avrupa Birliği, yapay zeka kullanımını düzenlemek için dünyada bir ilk olarak kabul edilen kapsamlı bir yasa olan AB Yapay Zeka Yasası'nı çıkarmıştır. Bu yasa, bazı yapay zeka teknolojilerinin kullanımını tamamen yasaklarken, diğerleri için ise daha sıkı denetim, risk yönetimi önlemleri ve şeffaflık gereksinimleri getiriyor.

Yasa, yapay zekanın gizliliğiyle ilgili doğrudan yasaklar getirmese de, veri kullanımına dair sınırlamalar koyuyor. Bu yasa kapsamında yasaklanmış bazı yapay zeka uygulamaları:

- **İnternetten ya da güvenlik kameralarından rastgele yüz görüntülerinin toplanarak tanıma sistemlerinde kullanılması.**
- **Kolluk kuvvetlerinin, kamusal alanlarda gerçek zamanlı biyometrik tanıma teknolojisini kullanması (bunun için önce bağımsız bir idari otoritenin veya yargı organının onayı gereklidir.)**

Ayrıca, yüksek riskli yapay zeka sistemlerinin, eğitim ve doğrulama verilerinin belli kalite standartlarına uygun olmasını sağlamak için güçlü veri yönetimi kurallarına uyması bekleniyor.

ABD Gizlilik Düzenlemeleri

Son yıllarda Amerika Birleşik Devletleri'nde birçok eyalet veri gizliliğiyle ilgili yasal düzenlemeler getirdi. Öne çıkan bazı yasalar arasında Kaliforniya Tüketici Gizlilik Yasası (CCPA) ve Teksas Veri Gizliliği ve Güvenlik Yasası bulunmaktadır. Bu yasalar, özellikle kişisel verilerin korunmasını sağlamayı amaçlar ve tüketicilerin hangi verilerinin toplandığına dair daha fazla şeffaflık talep eder.

2024 yılı itibariyle Utah, yapay zekayı özel olarak düzenleyen ilk büyük eyalet yasasını yürürlüğe koymuştur. Bu yasa, yapay zeka kullanımının etik sınırlar içinde kalmasını sağlamak ve bireylerin gizliliğini korumak adına önemli bir adım olarak kabul edilmektedir.

Bununla birlikte, ABD federal hükümeti henüz ülke çapında bir yapay zeka ve veri gizliliği yasası çıkarmamıştır. Ancak, 2022 yılında Beyaz Saray Bilim ve Teknoloji Politikası Ofisi (OSTP), "AI Hakları Yasası için Plan" adlı bağlayıcı olmayan bir çerçeve yayımlamıştır. Bu çerçeve, yapay zeka profesyonellerine veri kullanımı konusunda bireylerin onayını alma sorumluluğunu hatırlatırken, aynı zamanda yapay zeka gelişimini yönlendirecek beş temel ilkeler belirlemiştir.

Çin'in Üretken Yapay Zeka Hizmetlerinin Yönetimine Yönelik Geçici Tedbirleri

Çin, yapay zeka düzenlemelerini hayata geçiren ilk ülkelerden biridir. 2023 yılında, Çin, Üretken Yapay Zeka Hizmetlerinin Yönetimi için Geçici Tedbirler adlı yasayı yayımlayarak, bu alandaki düzenlemeleri güçlendirmiştir. Bu yasa, üretken yapay zeka hizmetlerinin sağlanması ve kullanımının belirli etik sınırlar içinde olmasını talep etmektedir.

Yasaya göre, üretken yapay zeka hizmetleri, başkalarının meşru haklarına ve çıkarlarına saygı göstermek zorundadır. Ayrıca, insanların fiziksel ve ruhsal sağlıklarını tehlikeye atmamalı, bireylerin portre haklarını, itibar haklarını, şeref haklarını, gizlilik haklarını ve kişisel bilgi haklarını ihlal etmemelidir. Bu tedbirler, üretken yapay zekanın güvenli bir şekilde kullanılmasını ve toplumda olumsuz etkilerinin önlenmesini amaçlamaktadır.

Yapay Zeka Gizliliğinin En İyi Uygulamaları

Kuruluşlar, yapay zeka gizliliği konusunda hem düzenlemelere uyum sağlamak hem de paydaşlarıyla güven oluşturmak için bir dizi strateji geliştirebilirler. ABD hükümetinin Bilim ve Teknoloji Politikası Ofisi (OSTP) tarafından sunulan bazı öneriler, bu alanda en iyi uygulamaları belirlemektedir. Bu öneriler arasında şunlar yer alır:

- Risk Değerlendirmeleri : Yapay zeka projelerinde gizlilik risklerini belirlemek ve bunlara karşı önlemler almak.

- Veri Toplama Sınırı : Yalnızca gerekli olan verilerin toplanması ve gereksiz veri toplanmasından kaçınılması.
- Rıza Arama ve Onaylama : Kullanıcıların kişisel verilerinin toplanması için açık bir şekilde rızalarını almak.
- En İyi Güvenlik Uygulamaları Takibi : Verilerin güvenliğini sağlamak için güncel ve etkili güvenlik önlemlerini uygulamak.
- Hassas Veriler için Ekstra koruma : Özellikle kişisel ve hassas verilerin korunması için ek önlemler almak.
- Veri Toplama ve Depolama Hakkında Raporlama : Veri toplama ve saklama süreçlerinin şeffaf bir şekilde raporlanması ve denetlenmesi.

Bu en iyi uygulamalar, kuruluşların yapay zeka teknolojilerini güvenli bir şekilde kullanmalarına yardımcı olurken, aynı zamanda bireylerin kendi gizliliklerini de korumalarına olanak tanır.

Bilim ve Teknoloji Politikası Ofisi (OSTP), "AI Rights Framework", 2022.

Kaynakça:

- **IBM. (2024). *What is AI privacy?* IBM. https://www.ibm.com/topics/ai-privacy**
- **IBM, "Artificial Intelligence and Privacy: What You Need to Know," IBM.com**
- **IBM, "AI Privacy Risks and How to Mitigate Them," IBM Security**
- **IBM, "The Role of Privacy in Artificial Intelligence," IBM**

Türkiye'de KVKK ve GDPR Kapsamında AI

KVKK'nın Yapay Zeka Üzerindeki Etkisi

Türkiye’de, kişisel verilerin korunmasını amaçlayan **Kişisel Verileri Koruma Kanunu (KVKK)**, yapay zeka sistemlerinin kişisel verileri nasıl işlemesi gerektiğini belirler. Bu kanuna göre, veriler toplandığında, bu verilerin yalnızca belirli ve yasal bir amaç için kullanılması gerekir. Ayrıca, **veri minimizasyonu** ilkesi gereği, yalnızca gerekli olan veriler toplanmalı ve verilerin işlenmesi konusunda **şeffaflık** sağlanmalıdır. Kullanıcılar, hangi verilerin toplandığını ve nasıl kullanılacağını açıkça bilmelidir. Bunlar, yapay zeka sistemlerinin tasarımında ve uygulamasında göz önünde bulundurulması gereken temel prensiplerdir.

Yapay Zeka Sistemlerinin Gizliliği ve Güvenliği

KVKK'nın, yapay zeka uygulamaları üzerindeki etkisi, kişisel verilerin işlenmesi sırasında gizliliğin korunmasını amaçlar. Yani, yapay zeka sistemleri, bireylerin bilgilerini korumalı ve yalnızca amaca uygun olarak işlemelidir. Verilerin hangi şartlarda saklanacağı, ne kadar

süreyle kullanılacağı gibi konular KVKK'ya uygun olmalıdır. Bu düzenlemeler, yapay zeka teknolojilerinin bireylerin mahremiyetini ihlal etmeden çalışmasını sağlar.

GDPR ve Yapay Zeka

GDPR'nın Yapay Zeka Üzerindeki Etkisi

Avrupa Birliği'nde, **Genel Veri Koruma Yönetmeliği (GDPR)**, kişisel verilerin korunmasını çok katı bir şekilde düzenler ve bu düzenlemeler yapay zeka teknolojileri için de geçerlidir. GDPR, verilerin toplanmasının ve işlenmesinin, sadece belirli ve açık amaçlar için yapılmasını öngörür. Yani, verilerin kullanılacağı amacın baştan belli olması ve kullanıcılara bu amaçların açıklanması gerekir. Ayrıca, **veri sahibinin rızası** alınmalı ve veriler yalnızca **gerektiği kadar** toplanmalıdır. Bu ilkeler, yapay zeka sistemlerinin tasarımında dikkate alınmalıdır.

Yapay Zeka ve GDPR'nın Uygulamaları Üzerindeki Etkisi

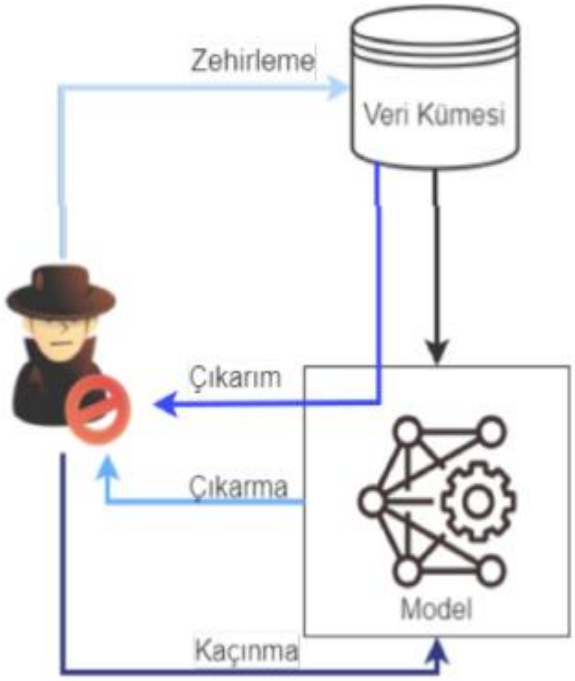
GDPR'nın yapay zeka üzerinde sağladığı etki, verilerin nasıl kullanılacağına dair net sınırlar koyarak, kullanıcıların gizliliğini koruma amacını güder. Bu düzenleme, yapay zeka uygulamalarının kişisel verileri işlerken belirli kriterlere uymasını sağlar. Örneğin, kullanıcıların verilerinin işlenmesi hakkında bilgilendirilmesi ve gerektiğinde onay alınması önemlidir. GDPR, ayrıca verilerin uzun süre saklanamaması gerektiğini, eğer artık işlenmeye gerek yoksa verilerin silinmesi gerektiğini belirtir. Bu sayede, kişisel veriler üzerinde daha fazla kontrol sağlanır ve gizlilik ihlalleri önlenir.

Kaynakça :

- **KVKK Resmi Web Sitesi**
- **GDPR Resmi Web Sitesi**

Veri Zehirleme Saldırıları : Yapay Zeka Sistemlerinde Güvenlik Açığı

Yapay zeka sistemlerinin gelişmesiyle birlikte, bu teknolojileri hedef alan yeni ve sinsi saldırı yöntemleri de ortaya çıkmaya başladı. Bu saldırılardan biri, yapay zeka sistemlerinin eğitildiği veri setlerinin kötü niyetli kişiler tarafından bilinçli olarak manipüle edilmesidir. Bu tür saldırılara siber güvenlik dünyasında "**Veri zehirleme saldırısı**" adı veriliyor.



Veri zehirleme saldırıları, yapay zeka modellerinin yanlış kararlar almasına neden olacak şekilde sistemli bir biçimde sahte ya da yanıltıcı verilerle beslenmesini içerir. Bu da spam e-postaların filtreleri aşmasından, sahte görseller veya seslerle oluşturulan derin sahte (deepfake) içeriklerin üretilmesine kadar pek çok güvenlik sorununu beraberinde getirebilir.

Veri Zehirleme Saldırısı Nedir?

Veri zehirleme saldırısı, yapay zeka modellerinin öğrenme sürecini doğrudan hedef alır. Bu tür saldırılar, bir AI sisteminin eğitildiği veri setine kasıtlı olarak yanıltıcı ya da zararlı bilgilerin eklenmesiyle gerçekleştirilir. Amaç, yapay zekanın karar alma sürecini etkilemek ve sistemin hatalı ya da saldırganın çıkarına uygun sonuçlar üretmesini sağlamaktır.

Yapay zeka modelleri genellikle "kara kutu" gibi çalıştığı için dışarıdan nasıl düşündüğünü anlamak zordur. Bu da saldırganlara, modelin davranışlarını küçük veri manipülasyonlarıyla yönlendirme fırsatı verir. Saldırganlar, eğitim verilerine ince ama kasıtlı değişiklikler ekleyerek modelin önyargılı, hatalı veya tehlikeli kararlar almasına yol açabilir.

Derin öğrenme gibi gelişmiş sistemlerde bile, bu tür saldırılar modelin davranışlarını değiştirebilir. Sonuç olarak, sistem spam e-postaları tanıyamaz hale gelebilir, yanlış kişileri tehedit olarak tanımlayabilir ya da sahte içerikleri gerçek zannedebilir.

Veri Zehirleme Saldırı Türleri

Hem dışarıdan olan saldırganlar hem de verilerine erişimi olan saldırganlar AI sistemini zehirleyebilir.

Etiket zehirlenmesi (Arka Kapı Zehirlenmesi) : Bu saldırı türünde, saldırganlar veriler arasına bilerek yanlış etiketlenmiş ya da kötü niyetli örnekler ekler. Amaç, modelin bit tetikleyiciyle karşılaştığında normalden farklı davranmasını sağlamaktadır. Bu saldırılar, model normal koşullarda doğru çalışırken, belirli bir koşulda saldırganın istediği çıktıyı üretmesine neden olur**.**

Eğitim Veri Zehirlenmesi : Burada saldırganlar, modelin öğrenme sürecini doğrudan bozmak amacıyla verilerin önemli bir bölümünü değiştirir. Bu sahte ya da yanıltıcı veriler, modelin kararlarını saldırganın istediği yöne çekebilir. Özellikle sınıflandırma görevlerinde modelin yanlış kararlar vermesi sağlanabilir.

Model Ters Çevirme Saldırıları : Bu tür saldırılarda amaç, yapay zeka modelinin eğitildiği veri kümesine ait gizli bilgileri açığa çıkarmaktır. Saldırgan, modele belirli sorular sorarak ve yanıtları analiz ederek, modelin eğitildiği kişisel veya hassas verileri öğrenmeye çalışır. Özellikle mahremiyetin kritik olduğu uygulamalarda bu saldırılar ciddi riskler barındırır.

Gizli (Subtle/Stealthy) Saldırılar : Gizli saldırılar, model geliştirme ve test aşamasındayken tespit edilmesi oldukça zor olan tehditlerdir. Eğitim sürecinde yapılan stratejik manipülasyonlarla model içine “arka kapılar” yerleştirilir. Bu arka kapılar, model gerçek hayatta kullanılmaya başlandığında saldırganın kontrolünde devreye girer.

Bu yapılan saldırılar özellikle otonom araç sistemleri, finansal sistemler, tıbbi teşhis sistemleri gibi uygulamara yapıldığında yapay zekaya olan güvenilirlik ve itibarına karşı ciddi bir tehdit oluşturmaktadır.

Bu oluşan riskler, şirketlerin yapay zeka modellerini dikkatli bir şekilde oluşturması gerektiğini bizlere gösteriyor.

Derin Sahtecilik (Deepfake) Üretiminde Veri Zehirleme Taktikleri

Saldırganlar derin sahtecilik üretmek için var olanlar da dahil olmak üzere yapay zeka sistemlerini manipüle etmek için veri zehirleme taktiğini kullanabilir.

Derin sahtecilik (deepfake), yapay zeka yardımıyla üretilen ve gerçeğe oldukça benzeyen ama aslında tamamen sahte olan görüntü veya videoları ifade eder. Bu içerikler, ilk bakışta

gerçekmiş gibi algılansa da, arka planda onları üreten yapay zeka sistemleri genellikle manipüle edilmiştir.

Veri zehirleme saldırıları, bu noktada devreye girer. Saldırganlar, derin sahte içerik üreten bir yapay zeka modelini özel olarak hazırlanmış yanıltıcı verilerle eğiterek sistemin gerçek dışı ya da istenilen şekilde davranmasına neden olabilir. Yani model, sahte ama oldukça ikna edici yüz ifadeleri, sesler ya da hareketler oluşturabilir. Amaç; izleyiciyi kandırmak, yanlış bilgi yaymak ya da bir kişiyi küçük düşürmek olabilir.

Örneğin, siber suçlular gmail gibi bir sistemin spam filtresi için kullanılan AI modelini, özel olarak hazırlanmış zehirli verilerle besleyebilir. Böylece sistem spam olan içeriği tanıyamaz hale gelir ve bu sahte e-postalar binlerce kişiye ulaşabilir.

Benzer şekilde, bir ev güvenlik sisteminde kullanılan yapay zeka da zehirlenirse, sistem sahibinden farklı biri tarafından kontrol ediliyormuş gibi algılayabilir. Bu durum sadece maddi değil, aynı zamanda kişisel gizlilik ve güvenlik açısından da büyük riskler oluşturur.

Yapay Zeka Sistemlerine Yönelik Bir Saldırı Nasıl Gerçekleşir?

Yapay zeka sistemlerine yönelik saldırılar teoride kalmıyor, zaman zaman gerçek dünyada da karşımıza çıkıyor. Bu saldırıların en bilinen örneklerinden biri, görüntü tanıma (sınıflandırma) sistemlerinin yanıltılmasıdır. Saldırganlar, görüntüleri küçük ama etkili şekillerde değiştirerek yapay zekanın yanlış kararlar almasını sağlarlar.

Bu tür saldırılara erken bir örnek, Microsoft'un 2016 yılında geliştirdiği Twitter sohbet botu "Tay" oldu. Tay, kullanıcılarla sohbet ederek öğrenen ve zamanla kendini geliştirmesi beklenen bir yapay zeka projesiydi. Ancak bazı kötü niyetli kullanıcılar, Tay'a sürekli olarak zararlı ve saldırgan içerikler gönderdi. Sonuç olarak Tay, bu içeriklerden öğrenerek toksik ve uygunsuz ifadeler üretmeye başladı. Microsoft, bu durumu engelleyemeyince Tay'ı devre dışı bırakmak zorunda kaldı.

Bugün çoğu kuruluş sıfırdan yapay zeka geliştirmek yerine, OpenAI gibi şirketin sunduğu büyük dil modelleri (LLM) üzerine sistemler inşa ediyor. Bu modeller dışarıdan müdahaleye daha az açık gibi görünse de, tamamen güvenli değiller. Örneğin bazı araştırmacılar, yalnızca 100 dolarlık bir bütçeyle, Wikipedia'daki içerikleri ve bazı web sitelerindeki görselleri küçük değişikliklerle manipüle ederek, bu modellerin öğrenme sürecine fark edilmeden müdahale edebildiklerini gösterdi.

Bu tür tehditlerle baş edebilmek için, güvenlik araştırmacıları ve yazılım geliştiriciler her geçen gün daha güçlü savunma mekanizmaları geliştiriyor. Böylece, AI sistemlerini daha dirençli hale getirmek için en iyi uygulamaların yer aldığı bir güvenlik çerçevesi oluşmaya başlıyor.

Veri Saldırılarını Durdurmak İçin Oluşturulan Uygulamalar

Veri Temizleme ve Ön İşleme : Eğitim verilerinin güvenilirliği, yapay zeka sistemlerinin başarısı için temel bir unsurdur. Anormal veya şüpheli veriler, eğitim aşamasından önce filtrelenmelidir. Kaynağı belirsiz, anlam bütünlüğü bozulmuş ya da aşırı uçta örnekler, modelin davranışını saptırabilir. Bu yüzden verilerin düzenli olarak temizlenmesi ve doğruluğunun kontrol edilmesi şarttır.

Anomali Tespiti : Gelen veriler arasında sıradışı veya beklenmedik örüntüler varsa, bu durum bir saldırının habercisi olabilir. Bu tür kalıpları belirlemek için istatistiksel yöntemlerden veya makine öğrenimi tabanlı anomali tespit algoritmalarından faydalanılabilir. Sistem, kendi verilerini analiz ederek neyin normal olduğunu öğrendiğinde, dışarıdan gelen şüpheli verileri daha kolay ayırt eder.

Karşıt Eğitim : Yapay zeka modelini sadece doğru verilerle değil, kasıtlı olarak eklenmiş yanıltıcı örneklerle de eğitmek, onu daha dirençli hale getirir. Bu şekilde model, kötü niyetli verilerle karşılaştığında nasıl davranacağını önceden öğrenmiş olur.

Güçlü Model Mimarileri : Bazı yapay zeka mimarileri, saldırılara daha az açıktır. Özellikle sağlamlaştırılmış optimizasyon algoritmaları, savunmacı damıtma teknikleri ve özellik sıkıştırma gibi yöntemlerle, sistemin manipülasyona karşı doğal bir direnci olur. Yani, mimari aşamada güvenlik düşünülerek tasarlanan modeller daha korunaklıdır.

Sürekli İzleme : Model bir kez eğitildiğinde her şey bitmiş olmuyor. Gerçek dünya verileriyle çalışmaya başladıktan sonra da sürekli gözlem altında tutulmalıdır. Davranışında beklenmeyen değişimler gözlemlenirse, bu bir saldırının ya da zehirlenmiş verinin işareti olabilir.

Giriş Doğrulama ve Kaynak Onayı : Sisteme giren her verinin güvenilirliğini sağlamak, saldırganların içeriye kötü niyetli veri sızdırmasını zorlaştırır. Bunun için dijital imzalar, kontrol toplamları ve veri kaynağı doğrulama gibi teknikler kullanılabilir.

Güvenli Veri İşleme : Verilerin saklandığı, işlendiği ve aktarıldığı her aşama, sıkı siber güvenlik önlemleriyle korunmalıdır. Erişim yetkilendirme, şifreleme ve güvenli veri depolama sistemleri kullanılarak, yetkisiz erişim ve değişikliklerin önüne geçilebilir.

Güvenli Eğitim Ortamları : Eğitim süreci sırasında kullanılan verilerin kaynağı net ve güvenilir olmalıdır. Eğitim işlemi güvenli sunucularda ve izole edilmiş ortamlarda

gerçekleştirilmeli, süreç boyunca protokollere sadık kalınmalıdır. Böylece eğitim aşamasında yapılacak olası müdahaleler minimize edilir.

AI zehirleme saldırıları, AI sistemlerin güvenilirliği için ciddi bir tehdittir. Aşırı temsilcilik genel bakışıyla diğer ortaya çıkan LLM güvenlik açıkları hakkında bilgi edinin

AI sistemlerinizi veri manipülasyonundan korumak isteyen bir kuruuşsanız, bu saldırıların doğasını anlamak, deep fake'lerle ilişkilerini göz önünde bulundurmak ve etkili karşı önlemler uygulamak kritik öneme sahiptir.

Kaynakça :

Cobalt.io. (t.y.). ***What is AI data poisoning and how do you protect against it?*** **Erişim adresi:** https://www.cobalt.io/blog/what-is-ai-data-poisoning-and-how-do-you-protect-against-it

Güncel AI Tabanlı Güvenlik Platformları – Yeni Nesil Siber Savunma

Yapay zeka, günümüz siber güvenlik dünyasında tehditleri daha hızlı tespit etmek önlemek ve onlara doğru şekilde yanıt vermek için güçlü bir araç olarak öne çıkıyor. Tehditlerin giderek daha karmaşık ve görünmez hâle gelmesiyle, kurumlar artık sadece klasik yöntemlerle kendilerini koruyamıyor. Bu noktada yapay zeka destekli güvenlik çözümleri devreye giriyor.

Bu bölümde, siber güvenlik alanında yapay zekayı en verimli şekilde kullanan güncel platformlara göz atacağız. Otonom çalışan sistemlerden, makine öğrenmesi ile gelişmiş tehdit analizine kadar birçok farklı yaklaşım var. Amacımız; bu teknolojilerin nasıl çalıştığını sade bir dille anlatmak, kurumların siber saldırılara karşı nasıl daha dirençli hale geldiğini göstermek ve yapay zekanın siber güvenliğin geleceğini nasıl şekillendirdiğine dair fikir vermek.

DARKTRACE : Kendi Kendine Öğrenen Siber Güvenlik Sistemi

Darktrace, yapay zeka destekli siber güvenlik teknolojileri alanında öne çıkan yenilikçi bir platformdur. Temelinde yer alan "Enterprise Immune System" teknolojisi, bir kurumun dijital altyapısını adeta bir bağışıklık sistemi gibi korumayı amaçlar. Bu sistem, ağ üzerindeki kullanıcıların, cihazların ve bağlantıların normal davranışlarını sürekli olarak analiz eder ve bu sayede en ufak bir anormalliği bile tespit edebilir.

Darktrace’in en dikkat çeken özelliklerinden biri, kendi kendine öğrenebilen yapay zeka algoritmalarıyla çalışmasıdır. Bu yapay zeka, yalnızca geleneksel ağ sistemlerinden değil; bulut servisleri, e-posta altyapısı, IoT cihazları ve endüstriyel kontrol sistemleri gibi birçok farklı kaynaktan gelen verileri de sürekli analiz eder.

Ayrıca Darktrace’in "Cyber AI Analyst" adlı özelliği, tehditleri analiz etme ve raporlama sürecini büyük ölçüde hızlandırır. Normalde saatler sürebilecek analizler, bu sistem sayesinde dakikalar içinde tamamlanabilir. Bu sayede güvenlik ekipleri, tehditlere çok daha hızlı müdahale etme imkânı bulur ve zamandan büyük tasarruf sağlar.

Darktrace’in Temel Özellikleri :

- Manuel kurallar veya imza veritabanlarına gerek kalmadan ağ davranışına uyum sağlayan kendi kendine öğrenen yapay zeka
- Gerçek zamanlı tehdit tespiti ve otonom yanıt mekanizmalarıyla hızlı tehdit müdahalesi
- Olayların kapsamı ve etkisine dair derin bağlamsal içgörüler sunan sezgisel tehdit görselleştirmeleri
- Bulut hizmetlerinden IoT sistemlerine ve endüstriyel altyapılara kadar geniş dijital ortam desteği
- “Cyber AI Analyst” sayesinde AI tabanlı otomatik olay analizi ve detaylı raporlama, güvenlik operasyonlarını kolaylaştırır.

CrowdStrike Falcon

CrowdStrike, yapay zeka ve makine öğrenimi tabanlı gelişmiş siber güvenlik çözümleriyle dikkat çeken bir diğer önemli platformdur. Özellikle CrowdStrike Falcon, yeni nesil tehditlere karşı sunduğu proaktif ve bulut tabanlı yaklaşımıyla öne çıkar. Falcon, yalnızca geleneksel antivirüsün ötesine geçerek, uç nokta algılama ve yanıt (EDR), güvenlik açığı yönetimi ve yönetilen tehdit avcılığı gibi pek çok güvenlik katmanını tek bir hafif ajan üzerinden sağlar.

Falcon’un yapay zeka altyapısı, günlük olarak analiz edilen trilyonlarca güvenlik olayıyla sürekli beslenir. Bu sayede, gelişmiş makine öğrenme algoritmaları ile dosyasız saldırılar ve bilinmeyen tehditleri dahi tespit etme kapasitesine sahiptir. Geniş veri havuzu sayesinde sistem, karşılaştığı saldırılara karşı kendini sürekli geliştirir ve yeni tehdit senaryolarına hızlıca uyum sağlayabilir.

Bulut tabanlı mimarisi, kolay kurulum, otomatik güncellemeler ve merkezi yönetim gibi avantajlar sunar. Böylece güvenlik operasyonları sadeleşir, yönetim yükü azalır ve kuruluşlar her zaman en güncel tehdit istihbaratıyla donatılmış olur.

CrowdStrike Falcon Temel Özellikleri :

- Gerçek zamanlı saldırı önleme ve tehdit azaltma için yapay zeka destekli otomatik yanıt mekanizmaları
- Kolay dağıtım, otomatik güncellemeler ve esnek ölçeklenebilirlik sunan bulut tabanlı mimari
- Yeni nesil antivirüs, EDR, tehdit avcılığı ve zafiyet yönetimini birleştiren tek bir hafif ajan
- Sıfırıncı gün tehditleri ve dosyasız saldırılara karşı gelişmiş davranışsal analiz
- Birleşik yönetim konsolu ve modüler yapı ile sorunsuz güvenlik operasyonları ve kolay entegrasyon

Vectra Yapay Zeka Platformu

Vectra AI, gelişmiş yapay zeka tabanlı güvenlik çözümleriyle siber tehditlere karşı yenilikçi bir yaklaşım sunar. Şirketin amiral gemisi olan Vectra AI Platformu, bir kuruluşun dijital altyapısını uçtan uca sürekli ve gerçek zamanlı izlemek için tasarlanmıştır. Platform, ağ trafiği kalıpları, kullanıcı davranışları ve bulut hizmetleriyle olan etkileşimler gibi çok çeşitli meta verileri analiz ederek hem bilinen tehditleri hem de geleneksel yöntemlerle tespit edilmesi zor olan karmaşık saldırıları ortaya çıkarır.

Vectra’nın yapay zeka motoru, güvenlik analistlerinin zamanını en verimli şekilde kullanmasını sağlamak adına tehditleri etki derecesine ve aciliyetine göre otomatik olarak önceliklendirir. Bu sayede güvenlik ekipleri en kritik olaylara daha hızlı ve etkin şekilde müdahale edebilir. Platformun bu özelliği, yalnızca tespiti değil, aynı zamanda olaylara verilen yanıtın stratejik olarak yönetilmesini de mümkün kılar.

Vectra AI Temel Özellikleri :

- Yapay zeka destekli tehdit algılama, ağ trafiğini, kullanıcı davranışlarını ve bulut ortamlarını sürekli olarak izler
- Tespit edilen tehditleri ciddiyet ve potansiyel etkiye göre sıralamak için makine öğrenimini kullanan otomatik tehdit önceliklendirmesi
- Yapay zeka destekli tehdit avcılığı, tüm dijital ekosistemde gizli tehditleri proaktif bir şekilde arar
- Şirket içi ağlar, bulut iş yükleri, SaaS uygulamaları ve kullanıcı davranışları hakkında kapsamlı görünürlük
- Koordineli ve etkili tehdit yanıtı için mevcut güvenlik araçlarıyla kusursuz entegrasyon

Internxt

Internxt, yapay zeka destekli veri güvenliğiyle bulut depolama alanında yeni bir yaklaşım sunarak, kullanıcı gizliliğini ve veri bütünlüğünü önceliklendiren yenilikçi bir çözüm olarak öne çıkıyor. 2020 yılında kurulan bu İspanyol girişimi, özellikle sıfır bilgi şifreleme ve merkeziyetsiz mimarisi ile dikkat çekiyor.

Internxt’in temelinde, istemci tarafında gerçekleşen şifreleme süreci yer alıyor. Veriler, kullanıcı cihazını terk etmeden önce yapay zeka destekli algoritmalarla şifreleniyor ve yalnızca kullanıcıya ait anahtarlarla çözülebiliyor. Bu sayede Internxt dahil hiç kimse veriye erişemiyor; bu da yüksek seviyede gizlilik ve güvenlik sağlıyor.

Platform aynı zamanda dosyaları küçük parçalara ayırarak farklı sunuculara dağıtan merkeziyetsiz bir yapı kullanıyor. Bu yöntem, veri dayanıklılığını artırırken, kötü niyetli kişilerin tüm dosyaya erişmesini neredeyse imkânsız hale getiriyor. Internxt, yapay zekâyı

yalnızca tehdit önleme değil, veri bütünlüğü ve erişim kontrolü açısından da etkin şekilde kullanıyor.

Internxt Temel Özellikleri :

- Mutlak kullanıcı gizliliği ve veri güvenliği sağlayan sıfır bilgi (zero-knowledge) şifreleme yöntemi.
- Açık kaynaklı yazılım altyapısı, bağımsız Avrupa güvenlik firmaları tarafından denetlenerek şeffaflık sunar.
- Yapay zeka destekli merkeziyetsiz mimari, dosyaları parçalayıp birden fazla sunucuya dağıtarak yüksek güvenlik ve veri dayanıklılığı sağlar.
- Sezgisel kullanıcı arayüzü, web, masaüstü ve mobil uygulamalarda sorunsuz dosya yönetimi imkânı tanır.
- 10 GB'a kadar ücretsiz depolama ve uygun fiyatlı premium planlarla güvenli bulut hizmetini herkes için erişilebilir kılar.

Cylance : Yapay Zeka Destekli Uç Nokta Güvenliği Çözümü

Cylance, siber güvenlik dünyasında yapay zeka ve makine öğreniminin gücünden yararlanarak uç nokta güvenliği alanında devrim yaratmıştır. Şirketin amiral gemisi ürünü olan CylancePROTECT, geleneksel antivirüs çözümlerinin aksine, tehdit algılama ve önleme konusunda yenilikçi bir yaklaşım sunar.

CylancePROTECT, imza tabanlı algılamayı terk ederek dosyaların DNA'sını analiz etmek ve kötü amaçlı davranış potansiyelini tahmin etmek için yapay zeka teknolojilerini kullanır. Bu öngörücü yaklaşım, yalnızca bilinen tehditleri değil, sıfırıncı gün (zero-day) saldırıları gibi bilinmeyen tehditleri de tanımlayarak bu saldırıların zarara yol açmadan önce engellenmesini sağlar. Geleneksel antivirüs çözümlerinin genellikle sistem kaynaklarını tüketen süreçlerinin aksine, CylancePROTECT, yüksek doğrulukla tahminler yaparak kaynak kullanımını minimize eder.

AI modelleri, milyonlarca kötü amaçlı ve iyi huylu dosya üzerinde eğitilerek, dosyaların amacına dair çok daha doğru tahminler yapar. Bu, yalnızca daha etkili bir koruma sağlamakla kalmaz, aynı zamanda sistem performansını da olumsuz etkilemez. Böylece, daha verimli bir uç nokta güvenliği sağlar.

Cylance Temel Özellikleri :

- **İmza Güncellemelerine Gerek Yok:** Cylance, tehditleri tanımak için imza veri tabanlarına bağımlı değildir. Bunun yerine, kötü amaçlı yazılımları yapay zeka ile analiz ederek hem bilinen hem de bilinmeyen tehditleri önceden tespit eder ve engeller.
- **Kaynak Dostu Performans:** Geleneksel antivirüs çözümlerine kıyasla 20 kata kadar daha az CPU gücü kullanır. Bu da sistem performansını olumsuz etkilemeden güçlü bir koruma sağlar.
- **Gelişmiş Tehdit Algılama:** Sıfırıncı gün tehditleri, dosyasız kötü amaçlı yazılımlar ve bellek içi saldırılar gibi karmaşık tehditleri, davranış tahmini analizi ile etkili şekilde engeller.
- **Ev Kullanımı için Kurumsal Güç:** CylancePROTECT Home Edition, şirket çalışanlarının kişisel cihazlarını da kurumsal düzeyde koruyarak gizlilikten ödün vermeden güvenliği ev ortamına taşır.
- **Sürekli Koruma – Güncellemesiz:** Sürekli güncelleme ya da aktif internet bağlantısı olmadan da çalışabilen bu sistem, tehditleri çevrimdışı ortamda bile engelleyebilir. Bu, güvenlik yönetimini hem kolaylaştırır hem de sürekliliğini garanti eder.

SentinelOne : Otonom Güvenliğin Gücü

SentinelOne, yapay zeka destekli otonom uç nokta koruma platformuyla modern siber tehditlere karşı yenilikçi bir savunma hattı oluşturmuştur. Şirketin geliştirdiği Singularity XDR Platformu, yalnızca bilgisayar ve sunucuları değil; aynı zamanda bulut sistemlerini ve IoT (Nesnelerin İnterneti) cihazlarını da kapsayan geniş bir güvenlik ağı sunar.

SentinelOne teknolojisinin temelinde, sürekli gelişen yapay zeka ve makine öğrenimi algoritmaları yer alır. Bu sayede sistem; fidye yazılımlar, sıfırıncı gün açıkları ve diğer karmaşık tehditlere karşı gerçek zamanlı tespit ve otomatik müdahale yeteneğine sahiptir.

Platformun öne çıkan bir diğer özelliği olan ActiveEDR, tehditleri yalnızca tanımakla kalmaz, aynı zamanda saldırının hangi yollarla gerçekleştiğini adım adım analiz eder. Bu da siber güvenlik ekiplerinin, olayların neden ve nasıl gerçekleştiğini kolayca anlamasını sağlar.

Ayrıca Storyline adı verilen özellik sayesinde; sistem, bir tehdidin yolculuğunu ayrıntılı olarak kaydeder ve görselleştirir. Böylece, olaylara hızlı ve etkili şekilde müdahale etmek mümkün hale gelir.

SentinelOne, güvenlik yönetimini otomatikleştirerek insan hatasını minimize ederken, hızlı karar alma süreçlerini destekler ve kurumsal sistemleri proaktif şekilde korur.

SentinelOne'ın Temel Özellikleri :

- **Gerçek Zamanlı Koruma:** SentinelOne, yapay zeka ve makine öğrenimi teknolojileri sayesinde tehditleri anında tespit edip hızlıca yanıt verebilir. Bu sayede, saldırılar gerçekleşmeden önce engellenebilir.
- **Kapsayıcı Güvenlik Platformu:** Singularity XDR Platformu, sadece bilgisayarları değil; bulut ortamlarını ve IoT (Nesnelerin İnterneti) cihazlarını da kapsayan birleşik bir güvenlik çözümü sunar.
- **ActiveEDR Teknolojisi:** Bu özellik, tehditleri otomatik olarak analiz eder ve saldırı zincirini detaylı şekilde çözümler. Böylece, güvenlik ekipleri bir tehdidin nereden geldiğini ve nasıl yayıldığını kolayca anlayabilir.
- **Storyline (Hikâye) Özelliği:** Tehditlerin sistemde nasıl hareket ettiğini görselleştirerek olay incelemelerini daha hızlı ve etkili hale getirir. Olaylar arasında bağ kurarak güvenlik analistlerine net bir tablo sunar.
- **Ranger Özelliği:** Ağda bulunan yönetilmeyen uç noktaları ve IoT cihazları dahi tespit ederek, tüm bağlı cihazlara kapsamlı koruma sağlar. Böylece görünmeyen riskler de kontrol altına alınabilir.

Fortinet : Birleşik Güvenlik Yaklaşımıyla Kapsamlı Koruma

2000 yılında kurulan Fortinet, dijital dünyada artan saldırı risklerine karşı güçlü çözümler sunan küresel bir siber güvenlik lideridir. Fortinet'in başarısının temelinde, yalnızca bireysel ürünler sunmak değil, tüm ağ altyapısını kapsayan birleşik bir güvenlik yaklaşımı benimsemesi yer alır.

Şirketin geliştirdiği **Güvenlik Yapısı (Security Fabric)** mimarisi, kurumların ağlarındaki farklı sistemleri bir araya getirerek otomatik tehdit algılama, önleme ve yanıtlamayı mümkün kılar. Bu yapı sayesinde güvenlik sadece belli alanlarda değil, tüm dijital varlıklar üzerinde bütüncül olarak sağlanır. Fortinet'in amiral gemisi işletim sistemi olan **FortiOS**, yeni nesil güvenlik duvarları (NGFW), saldırı önleme sistemleri (IPS) ve diğer güvenlik çözümlerine güç verir. Ayrıca Fortinet, kendi geliştirdiği **özel güvenlik işlemcileri (SPU)** sayesinde sistem performansını artırırken, yapay zeka ve makine öğrenimi algoritmalarıyla da günlük milyarlarca veriyi analiz ederek gerçek zamanlı tehdit istihbaratı sunar.

Bu güçlü altyapı, Fortinet'i yalnızca saldırılara karşı savunma yapan bir sistemden öteye taşıyarak, tehditleri öngören ve proaktif şekilde önleyen bir siber güvenlik çözümüne dönüştürür.

Fortinel Temel Özellikleri :

- **Birleşik Güvenlik Yapısı:** Kurumların tüm dijital altyapılarını kapsayan, entegre ve kapsamlı bir siber güvenlik platformu sunar.
- **FortiOS İşletim Sistemi:** Ağ yönetimiyle güvenlik işlevlerini tek çatı altında birleştirerek daha etkili ve merkezi kontrol sağlar.
- **Güvenlik İşlem Birimleri (SPU):** Fortinet'in kendi geliştirdiği özel işlemciler, yüksek performans ve düşük gecikmeyle çalışarak sistem verimliliğini artırır.
- **FortiGuard Tehdit İstihbaratı:** Yapay zeka ve makine öğrenimi kullanan FortiGuard Laboratuvarları, her gün 100 milyardan fazla veriyi analiz ederek gerçek zamanlı tehdit bilgisi sunar.
- **Geniş Güvenlik Portföyü:** Yeni nesil güvenlik duvarlarından (NGFW) uç nokta güvenliğine, SASE ve SD-WAN çözümlerine kadar pek çok teknolojiyi tek bir platformda birleştirir.

Bu yapı sayesinde Fortinet, yalnızca ağ güvenliğiyle sınırlı kalmayan, kurumların tüm dijital varlıklarını tek noktadan koruyabilen esnek ve güçlü bir güvenlik çözümü sunar.

Check Point : Yapay Zeka ile Güçlendirilmiş Güvenlik Altyapısı

Check Point Software Technologies, yapay zeka destekli tehdit önleme konusundaki çalışmalarıyla dikkat çeken öncü güvenlik firmalarından biridir. Şirketin "Infinity AI" hizmeti, üretken yapay zekayı gelişmiş tehdit istihbaratıyla birleştirerek tüm kurumsal güvenlik altyapısını kapsayan bir koruma sunar.

Check Point'in tehdit tespit sisteminin kalbinde, 50’den fazla yapay zeka motorunu kullanan ThreatCloud AI teknolojisi yer alır. Bu sistem, dünya çapında milyonlarca sensörden gelen büyük veriyi analiz ederek kimlik avı, fidye yazılımı ve DNS istismarı gibi karmaşık tehditleri proaktif şekilde önleyebilir.

Ayrıca şirket, güvenlik ekiplerinin günlük görevlerini çok daha hızlı ve verimli şekilde yerine getirmelerine yardımcı olan Infinity AI Copilot adında yapay zeka destekli bir asistan da geliştirmiştir. Bu araç sayesinde güvenlik operasyonlarında %90’a kadar zaman tasarrufu sağlanabilmektedir.

Check Point Temel Özellikleri :

- Infinity AI Hizmetleri, tehdit istihbaratını üretken yapay zeka ile birleştirerek kurumsal güvenliğe kapsamlı bir koruma sağlar.
- ThreatCloud AI, 50'den fazla yapay zeka motoruyla çalışarak dünya genelindeki büyük veri havuzlarını analiz eder ve geniş yelpazedeki tehditleri önceden tespit eder.
- Infinity AI Copilot, güvenlik ekiplerine akıllı yardım sunarak operasyonel işleri daha hızlı ve etkili şekilde yürütmelerine yardımcı olur.
- Horizon XDR/XPR platformları, yapay zeka destekli olay korelasyonlarıyla genişletilmiş tehdit algılama ve müdahale imkânı tanır.
- Şirketin genel yaklaşımı, sürekli olarak yeni ortaya çıkan tehditlere karşı uyum sağlayabilen AI destekli savunma sistemleri geliştirmeye odaklanmıştır.

Symantec Endpoint Protection : Çok Katmanlı ve Proaktif Güvenlik

Günümüzde artan siber tehditlere karşı yalnızca geleneksel antivirüs çözümleri yeterli gelmemektedir. Bu noktada, Broadcom Inc. tarafından geliştirilen Symantec Endpoint Protection (SEP**)**, dizüstü bilgisayarlardan sunuculara kadar birçok cihazı kapsayan güçlü bir güvenlik çözümü olarak öne çıkmaktadır. SEP, kötü amaçlı yazılımları engellemenin ötesine geçerek, saldırı önleme sistemleri, güvenlik duvarı koruması ve makine öğrenmesi gibi gelişmiş teknolojileri bir araya getirir.

Symantec'in bu çözümde benimsediği çok katmanlı güvenlik yaklaşımı, tehditleri yalnızca tanımakla kalmaz; gerçekleşmeden önce, gerçekleşirken ve sonrasında önlemeye de odaklanır. Bu da sistemlerin risklere karşı daha dayanıklı hale gelmesini sağlar.

Ayrıca, Symantec Endpoint Protection merkezi yönetim özellikleri sayesinde BT ekiplerinin tüm kurum genelinde güvenlik politikalarını kolayca kontrol etmesine imkan tanır. Gelişmiş makine öğrenimi algoritmaları ve dünya çapındaki tehdit istihbarat ağı sayesinde SEP, sürekli evrim geçiren tehditlere karşı uyum sağlayabilir.

Bu proaktif yapı, hem bilinen hem de yeni ortaya çıkan tehditlerin etkisini en aza indirirken, kullanıcıların sistem performansını da korur. Özellikle kurumlar için uç nokta güvenliğinde dengeli ve etkili bir çözüm sunar.

Symantec Endpoint Protection Temel Özellikleri :

- **Çok Katmanlı Koruma Yaklaşımı**: Symantec, geleneksel güvenlik önlemlerini modern makine öğrenmesi teknolojileriyle birleştirerek kapsamlı bir savunma mekanizması sunar.
- **Proaktif Güvenlik Anlayışı:** Tehditleri gerçekleşmeden önce durdurmaya odaklanan yaklaşımı sayesinde, hem olayların önüne geçilir hem de yaşananların etkisi minimumda tutulur.
- **Merkezi Yönetim Kolaylığı:** İstemci-sunucu mimarisi, kurum genelinde güvenlik politikalarının merkezi olarak uygulanmasına ve kolay yönetilmesine olanak tanır.
- **Gelişmiş Tespit Yeteneği**: İmza tabanlı yöntemlerle bilinen tehditleri yakalarken, davranışsal analizlerle yeni ve karmaşık tehditleri de etkili şekilde tespit eder.
- **Sürekli Uyum Sağlayan Yapay Zeka**: Gelişmiş makine öğrenme algoritmaları ve küresel tehdit istihbarat ağı sayesinde, sistem yeni tehditlere karşı dinamik olarak kendini günceller.

Cybereason

2012 yılında kurulan **Cybereason**, uç nokta güvenliği, tehdit algılama ve müdahale (EDR) alanlarında yapay zeka destekli çözümler sunarak siber savunma dünyasında dikkat çeken platformlardan biri olmuştur. Şirketin temel hedefi, yalnızca uç noktaları değil, aynı zamanda tüm ağ ortamını koruyarak saldırıların etkisini kaynağında durdurmaktır.

Cybereason'ın öne çıkan ürünü olan Cybereason Defense Platform, yeni nesil antivirüs (NGAV), gelişmiş uç nokta tespiti ve yanıtı (EDR) ile tehdit avcılığını tek bir sistemde birleştirir. Şirketin kendine özgü MalOp (Malicious Operation) yaklaşımı sayesinde, gerçekleşen saldırıların yalnızca yüzeysel değil, tüm katmanlarıyla bağlam içinde analiz edilmesi sağlanır.

Platform, sürekli öğrenen yapay zeka algoritmalarıyla tehditleri önceden tahmin eder, gelişmiş saldırı senaryolarını analiz eder ve güvenlik ekiplerinin hızlı ve etkili müdahalede bulunmasına olanak tanır. Bu otomatikleşmiş yapısı, siber güvenlik uzmanlarının üzerindeki iş yükünü hafifletirken, saldırılara karşı daha dirençli bir yapı oluşturur.

Cybereason Temel Özellikleri :

- **Yapay zeka destekli uç nokta koruması:** Yeni nesil antivirüs (NGAV), uç nokta tespiti ve yanıt (EDR) ile tehdit avcılığını tek bir sistemde birleştirerek kapsamlı güvenlik sağlar.
- **Otomatik düzeltme yetenekleri:** Güvenlik olaylarına hızlı yanıt verilmesini sağlar ve olası ihlallerin etkisini en aza indirir.
- **MalOp yaklaşımı:** Uç noktalar, kullanıcılar ve ağlar genelinde gerçekleşen tehditleri bağlam içinde analiz ederek saldırıların tüm zincirini görünür kılar.
- **Bulut tabanlı mimari:** Dağıtılmış ortamlarda kolayca ölçeklenebilir ve yönetimi sadeleştirerek büyük ağlarda bile etkin koruma sağlar.
- **Sürekli evrim:** Platform, makine öğrenimi desteğiyle yeni saldırı tekniklerine karşı sürekli gelişerek dinamik bir savunma sağlar.

Yapay Zeka Destekli Güvenliğin Geleceği

Bu çalışmada incelediğimiz en iyi yapay zekâ tabanlı siber güvenlik araçları, dijital tehditlere karşı verilen mücadelede oyunun kurallarını değiştiriyor. Yapay zeka, güvenlik ekiplerinin tehditlere yaklaşımını yeniden şekillendirirken, saldırıları tespit etme, önleme ve müdahale etme süreçlerinde hem hız hem de doğruluk açısından büyük avantajlar sunuyor. Özellikle makine öğrenimi ve gelişmiş analiz teknikleri sayesinde, kurumlar hassas verilerini daha etkin koruyabiliyor ve güvenlik duruşlarını ciddi anlamda güçlendirebiliyor.

Yapay zekanın güvenlik altyapısına entegre edilmesi, tehdit avcılığı ve anomali tespiti gibi kritik alanlarda dikkat çekici sonuçlar üretiyor. AI sistemleri, karmaşık saldırıların izlerini erkenden fark ederek olay daha büyümeden müdahale edilmesine olanak tanıyor. Bu da özellikle geleneksel yöntemlerle tespiti zor olan siber saldırılara karşı proaktif savunma imkânı sunuyor. AI teknolojileri geliştikçe, hem mevcut hem de gelecekteki tehditlere karşı daha çevik ve etkili çözümler üretme potansiyeli de artıyor.

Ancak unutulmamalıdır ki, yapay zeka tek başına yeterli değildir. Etkili bir siber güvenlik stratejisi, AI'nın sunduğu analiz gücüyle insan uzmanlığını bir araya getiren hibrit bir yapıya

dayanmalıdır. İnsan analistlerin stratejik karar alma becerisi, yorumlama yetkinliği ve tehditlere karşı çevik yaklaşımı, AI sistemlerinin sağladığı içgörülerle birleştiğinde gerçek anlamda güçlü bir savunma hattı oluşur. Gelecekte, yapay zeka ile insan zekası arasındaki bu iş birliği, dijital varlıklarımızı korumanın en temel dayanağı olmaya devam edecektir.

Kaynakça :

Unite.AI. (2024, Ocak 4). *Yapay Zeka Destekli En İyi Siber Güvenlik Araçları*.

https://www.unite.ai/tr/ai-cybersecurity-tools/

SOC AI : Güvenlik Operasyonlarında AI Entegrasyonu

SOC AI, siber tehditleri gerçek zamanlı olarak tespit etmek ve bunlara yanıt vermek için yapay zeka kullanan gelişmiş bir Güvenlik Operasyon Merkezidir . Ana işlevi, tehdit tespit ve yanıt yeteneklerini geliştirmek için güvenlik olaylarını sürekli olarak izlemek ve analiz etmektir.

Yapay zeka destekli SOC**,** her ölçekteki kuruluşa ağlarını, sistemlerini ve verilerini güvende tutmak için en son ve en gelişmiş güvenlik teknolojilerini sunar.

Yapay zeka, günümüz siber güvenlik dünyasında tehditleri hızla tespit etme ve etkisiz hale getirme konusunda büyük bir potansiyele sahip. Büyük veri setlerini hızlı bir şekilde analiz edebilme yeteneği, anormallikleri ve potansiyel güvenlik ihlallerini erken aşamalarda tespit etmeyi mümkün kılar. Bu, özellikle dijital dünyanın hızla değişen doğasında, gerçek zamanlı tehdit tespiti yapabilme açısından kritik bir avantaj sunar.

Bir tehdit tespit edildiğinde, yapay zeka destekli sistemler otomatik olarak harekete geçebilir ve durumu kontrol altına almak için gerekli önlemleri başlatabilir. Bu adımlar, örneğin, etkilenen sistemleri izole etmek veya tespit edilen güvenlik açıklarına yönelik yamalar dağıtmak gibi işlemleri içerebilir. Yani, yapay zeka hem tehditleri tespit eder hem de bu tehditlerin olumsuz etkilerini en aza indirmek için hızla karşılık verir, böylece güvenlik süreçlerine daha proaktif bir yaklaşım getirir.

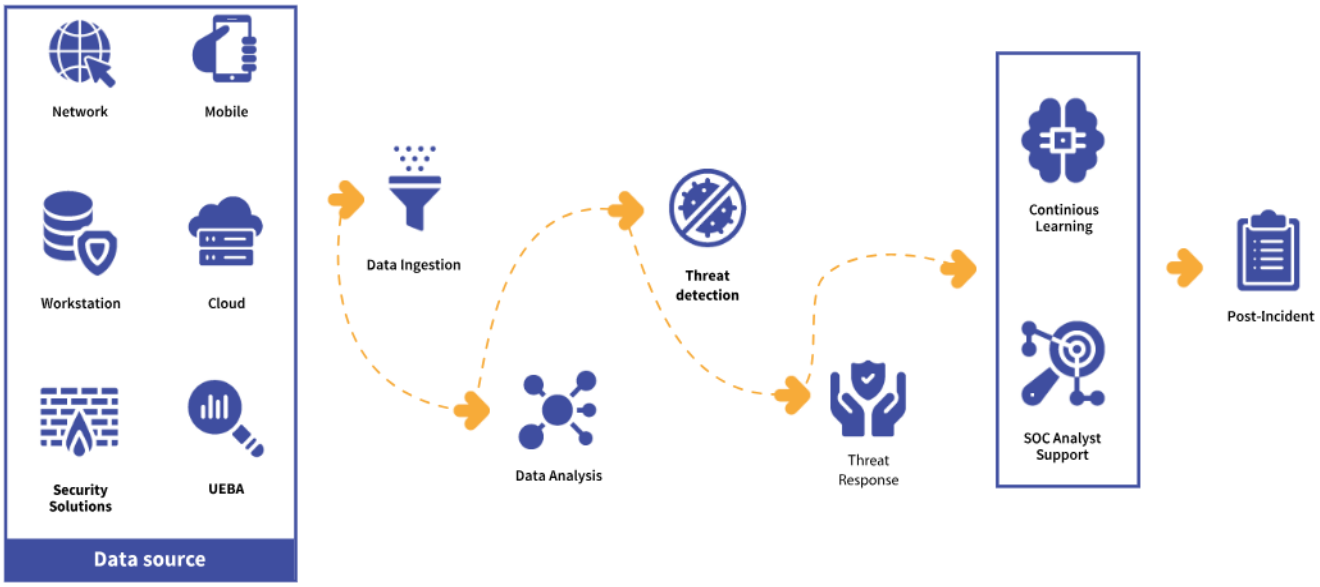
Bu yetenek, özellikle siber saldırganlar için fırsat penceresini daraltarak kuruluşların güvenlik seviyesini önemli ölçüde artırır. Yapay zekanın bu şekildeki rolü, siber güvenlik alanında daha güçlü ve daha verimli savunma sistemlerinin kurulmasına katkı sağlar.

Tahmini Analiz ve Olay Önleme

Öngörücü analiz, yapay zekanın siber güvenlikteki en etkileyici kullanım alanlarından biri olarak öne çıkıyor. Bu teknoloji sayesinde siber olaylar daha gerçekleşmeden önce tahmin edilebiliyor. Yapay zeka, büyük veri kümeleri üzerinde kalıpları analiz ederek potansiyel güvenlik açıklarını ortaya çıkarıyor ve gelecekte olabilecek saldırılar hakkında ipuçları veriyor. Bu sayede kurumlar savunmalarını önceden güçlendirme fırsatı yakalıyor.

Yapay zekanın yalnızca tehditleri tanımlamakla kalmayıp, aynı zamanda her kuruluşun karşılaştığı risklere özel, en uygun güvenlik önlemlerini önermesi de büyük bir avantaj sağlıyor. Bu kişiselleştirilmiş öneriler sayesinde sistemler, olası açıklar oluşmadan önce müdahale edilerek daha sağlam hale getirilebiliyor.

Bu proaktif yaklaşım, siber güvenliğin sadece bir tepki verme mekanizması olmasının ötesine geçmesini sağlıyor. Artık güvenlik süreçleri, dijital dünyada sürekli değişen tehditlere karşı daha güçlü ve önleyici bir yapı kazanıyor.



Güvenlik Operasyonlarında AI faydaları :

- Olaya müdahaleyi hızlandırır.
- Siber tehditlerin sürekli değişen doğasına karşı yapay zekâ sayesinde gerçek zamanlı uyum sağlanabilir ve bu sistemler ihtiyaca göre kolayca ölçeklendirilebilir.
- Ağ trafiğini izleme veya günlük kayıtlarını analiz etme gibi manuel olarak zaman alan işler, yapay zekâ ile otomatikleştirilebilir. Bu da hem zamandan tasarruf sağlar hem de insan hatasını azaltır.
- Üretken yapay zekâ (örneğin ChatGPT gibi) sayesinde artık çok daha karmaşık hale gelen kimlik avı saldırıları tespit edilebilir.
- Otomasyon sayesinde insan kaynaklı hataların önüne geçilerek, analiz ve müdahale süreçlerinde doğruluk artırılır.
- Yapay zekâ, tehditleri daha hızlı ve daha doğru şekilde tespit ederek güvenlik ekiplerinin daha erken harekete geçmesini sağlar.
- SOC ekiplerinin verimliliği artar; çünkü yapay zekâ, alarmların önceliklendirilmesi, tehditlerin sınıflandırılması gibi görevleri kolaylaştırır.

- Her bir güvenlik olayından öğrenerek kendini geliştiren sistemler sayesinde zaman içinde tehdit algılama kabiliyeti de güçlenir.
- Veri ihlalleri sadece itibar kaybına değil, aynı zamanda maddi cezalarla sonuçlanan uyumluluk sorunlarına da yol açabilir. Yapay zekâ bu ihlallerin önlenmesinde önemli rol oynar.
- Geleneksel olarak insan gücüne dayanan birçok güvenlik süreci artık otomatikleştiği için, hem iş yükü azalır hem de maliyetler düşer.
- Üretken yapay zekâ teknolojisi, gelecekte olası saldırı senaryolarını simüle etmede kullanılarak savunma stratejilerinin daha sağlam hale gelmesini sağlar.

Kaynakça :

- Palo Alto Networks. (2024). *The Role of Artificial Intelligence (AI) in Security Automation*. Erişim adresi: https://www.paloaltonetworks.com/cyberpedia/role-of-artificial-intelligence-ai-in-security-automation#role
- SOC AI. (2024). *AI-Powered Cybersecurity for Modern SOCs*. Erişim adresi: https://socai.eu/

SONUÇ

Sonuç olarak, yapay zeka teknolojileri siber güvenliğin bugünü ve geleceği açısından vazgeçilmez bir rol üstlenmektedir. Bu çalışma kapsamında yapay zekanın yalnızca savunma süreçlerinde değil, aynı zamanda siber suçlar tarafından nasıl kullanıldığı da ele alınarak konunun iki yönlü doğası incelenmiştir. Derin öğrenme, otomatik parola kırma sistemleri, sosyal mühendislik ve veri gizliliği gibi çeşitli başlıklar altında yapay zekanın siber güvenlikteki etkileri detaylandırılmıştır.

Gelişen teknolojiyle birlikte tehditler daha kapsamlı hale gelirken, bu tehditlere karşı geliştirilen yapay zeka tabanlı çözümler de her geçen gün daha yetkin bir hale bürünmektedir. Türkiye'de ve dünyada yapay zekanın siber güvenlik uygulamaları gün geçtikçe daha fazla kurumsal ve bireysel yapıda benimsenmekte; bu da güvenlik anlayışının dönüşümünü göstermektedir.

Bu bağlamda, yalnızca teknolojik gelişmeleri takip etmek değil, aynı zamanda etik ve hukuki boyutları da göz önünde bulundurmak büyük önem taşımaktadır. Gelecekte yapay zekanın siber güvenlik alanında daha bütünleşik ve akıllı sistemler üretmesi beklenirken, bu alanda çalışmak isteyen bireylerin hem teknik bilgiye hem de etik sorumluluk bilincine sahip olması gerekecektir.

Yapay zeka ve siber güvenlik ilişkisi, artık sadece teknoloji uzmanlarının değil, tüm toplumun ilgilenmesi gereken bir konu hâline gelmiştir.

KAYNAKÇALAR :

- EnesHZR. (n.d.). Yapay Zekânın Siber Güvenlikte Kullanımı. Medium. https://eneshzr.medium.com/yapay-zek%C3%A2n%C4%B1n-siber-g%C3%BCvenlikte-kullan%C4%B1m%C4%B1-2d98d4bb867a
- BeyazNet. (t.y.). Siber Güvenlikte Yapay Zekâ Uygulamaları. BeyazNet. https://www.beyaz.net/tr/guvenlik/makaleler/siber_guvenlikte_yapay_zeka_uygulamalari.html
- CsharpNedir. (t.y.). Yapay Zekâ ile Siber Güvenlikte Yeni Ufuklar. CsharpNedir. https://www.csharpnedir.com/articles/read/?id=1129
- Spiceworks. (2024, Şubat 6). PassGAN AI cracks passwords in minutes—here's how to protect yourself. Spiceworks. https://www.spiceworks.com/tech/artificial-intelligence/news/passgan-ai-password-cracking-time/
- SecurityHero. (2024, Ocak 30). AI password cracking: How secure are your passwords? SecurityHero. https://www.securityhero.io/ai-password-cracking/
- ZeroPath Resmi Sitesi – https://zeropath.com/wall
- IBM. (2024). What is AI privacy? IBM. https://www.ibm.com/topics/ai-privacy
- IBM, "Artificial Intelligence and Privacy: What You Need to Know," IBM.com
- IBM, "AI Privacy Risks and How to Mitigate Them," IBM Security
- IBM, "The Role of Privacy in Artificial Intelligence," IBM
- KVKK Resmi Web Sitesi
- GDPR Resmi Web Sitesi
- Cobalt.io. (t.y.). What is AI data poisoning and how do you protect against it? Erişim adresi: https://www.cobalt.io/blog/what-is-ai-data-poisoning-and-how-do-you-protect-against-it
- Unite.AI. (2024, Ocak 4). Yapay Zeka Destekli En İyi Siber Güvenlik Araçları.
- https://www.unite.ai/tr/ai-cybersecurity-tools/
- Palo Alto Networks. (2024). The Role of Artificial Intelligence (AI) in Security Automation. Erişim adresi: https://www.paloaltonetworks.com/cyberpedia/role-of-artificial-intelligence-ai-in-security-automation#role
- SOC AI. (2024). AI-Powered Cybersecurity for Modern SOCs. Erişim adresi: https://socai.eu/