AI Powered Cyberattacks

A key technology in enterprise IT toolbox becoming a weapon of cybercriminals

Edited by **MOHAN DAS VISWAM**

Modern Artificial Intelligence, powered by Machine Learning, is decorating the position that Computers occupied during the latter half of the last century. While the common man looked at both with wonder, experts and enthusiasts explored them to extract their potential for benefit of mankind. As the technologies matured, adversaries boarded the bandwagon to reap the benefits of the sweat on other's brows. Presently, threat actors utilize algorithms and techniques powered by AI to automate, accelerate, and enhance various phases of cyberattacks.

Characteristics of AI-Powered Cyberattacks

AI-powered cyberattacks have the following five main characteristics

Attack automation

Until very recently, most cyberattacks required significant hands-on support from a human adversary. However, growing access to AI and generative AI-enabled tools is allowing adversaries to automate attack research and execution.

Efficient data gathering

The first phase of every cyberattack is reconnaissance where the attackers search for targets, exploitable vulnerabilities, and assets that could be compromised. AI can automate or accelerate much of this legwork, enabling adversaries to drastically shorten the research phase and potentially improve the accuracy and completeness of their analysis.

Customization

One of the key capabilities of AI is data scraping, where the information from public sources, such as social media sites and corporate websites, is gathered and analyzed. In the context of a cyberattack, this information can be used to create hyper-personalized, relevant, and timely messages that serve as the foundation for phishing attacks and other attacks that leverage social engineering techniques.

A. Ramadevi
Scientist- D
rama.a@nic.in

AI-powered cyberattacks leverage AI and ML algorithms and techniques to automate, accelerate, and enhance various phases of a cyberattack. This includes identifying vulnerabilities, deploying campaigns along identified attack vectors, advancing attack paths, establishing backdoors within systems, exfiltrating or tampering with data, and interfering with system operations. The adaptability of AI algorithms to learn and evolve over time enables the threat actors to avoid detection by creating attack patterns that security systems can't distinguish.

Reinforcement learning

AI algorithms learn and adapt in real time. In the same way that these tools continuously evolve to provide more accurate insights for corporate users, they also evolve to help adversaries improve their techniques or avoid detection.

Employee targeting

Similar to attack customization, AI can be used to identify individuals within an organization that are high-value targets. These are people who may have access to sensitive data or broad system access, may appear to have lower technological aptitude, or have close relationships with other key targets.

Types of AI-Powered Cyberattacks

There are multiple types of cyberattacks enabled by AI and machine learning. Some include:

AI-Driven Social Engineering Attacks

Social engineering attack aims to manipulate human behavior to fulfill a purpose, such as sharing sensitive data, transferring money or ownership of high-value items, or granting access to a system, application, or database. AI driven attacks leverage AI algorithms to assist in the research, creative conception, and execution of a social engineering attack. It can identify ideal attack targets and develop online presence to communicate with them so that attention is generated through a realistic scenario.

AI-Driven Phishing Attacks

These attacks use generative AI to create highly personalized and realistic emails, SMS messages, phone communication, or social media outreach to achieve a desired result. AI-powered ChatBots can support interactions that make them nearly indistinguishable from humans. In most cases, the goals of these attacks are the same as that of a social engineering attack: to access sensitive information, gain access to a system, receive funds, or prompt a user to install a malicious file on their device.

DeepFakes

A DeepFake is an AI-generated video, image, or audio file that is meant to deceive people. DeepFakes commonly appear on the internet to enter-

tain and confuse. However, they can also be used more maliciously as part of disinformation campaigns, "fake news," smear campaigns of high-profile individuals, or cyberattacks. Attackers sometimes use existing voice/video footages of one person to create a doctored footage and instruct another person to take specific actions, such as transferring funds or granting system access.

Adversarial AI/ML

Adversarial AI or ML is a process used by an attacker to disrupt the performance or decrease the accuracy of AI/ML systems through manipulation or deliberate misinformation. Attackers use several adversarial AI/ML techniques that target different areas of model development and operation. Poisoning attacks, Evasion attacks and Model tampering are typical methodologies employed by adversaries to compromise the system and produce inaccurate outputs.

Malicious GPTs

A generative pre-trained transformer (GPT) is a type of AI model that can produce intelligent text in response to user prompts. A malicious GPT refers to an altered version of a GPT that produces harmful or deliberately misinformed outputs. In the context of cyberattacks, a malicious GPT can generate attack vectors (such as malware) or supporting attack materials (such as fraudulent emails or fake online content) to advance an attack.

Ransomware Attacks

AI-driven ransomware attacks leverages AI to improve its performance or automate some aspects of the attack path such as research targets, identify system vulnerabilities, or encrypt data. AI can also be used to adapt and modify the ransomware files over time, making them more difficult to detect with cybersecurity tools.

Mitigation of AI-Powered Cyberattacks

AI technology makes it potentially easier and faster for cybercriminals to carry out cyberattacks, effectively lowering the barrier to entry for some actors and increasing the level of sophistication of established players. AI-powered attacks are often more difficult to detect and prevent than attacks that use traditional techniques and manual processes, making them a significant security threat to all companies. Recommendations across four key categories to protect and defend against AI-powered cyberattacks are given below:

Regular Security Assessments

Deploying a comprehensive cybersecurity platform that offers continuous monitoring, intrusion detection, and endpoint protection is the preliminary step towards mitigation of any cyber-attacks. Baselines may be developed for system activity and user behavior to serve as a standard of comparison for future activity and establish user and entity behavior analytics (UEBA). Analyzing systems for abnormal user activity or unexpected changes within the environment may indicate onset of an attack. Real-time analysis of input and output data for the AI/ML system may be implemented to protect against adversarial AI attacks.

Incident Response

An incident response plan that outlines the procedures, steps, and responsibilities in the event of a cyberattack may be put in place within the organization. The plan should be capable of determining the type and severity when a security incident occurs. Restrict system use and operation to limit the spread and impact of the attack. Also execute remedial measures as envisaged in the plan to restore system usage. Patch the vulnerabilities detected and implement additional security measures to prevent similar attacks in the future and safeguard against a wider range of threats.

Awareness Training

Human beings being the weakest link in the cyber security chain, regular and updated training may be imparted to all stakeholders to prevent all kinds of cyber-attacks. Specific modules focusing on AI-powered attacks may be included in the training schedules. Teams may be trained to recognize suspicious activities, behaviours and outputs related to AI/ML-based systems. Awareness may be created on how realistic and convincing AI-enabled attack techniques can be, especially when it relates to social engineering techniques and DeepFake chat and audio-based attacks.

Use of AI-Powered Solutions

Organizations may use AI itself to counter AI-based attacks. AI-enabled tools may be leveraged to automate security-related tasks, such as monitoring, analysing datasets, identifying attack patterns, prevention, patching, and remediation. System parameters that alert teams to high-risk activity may be developed to device and prioritize the responses.

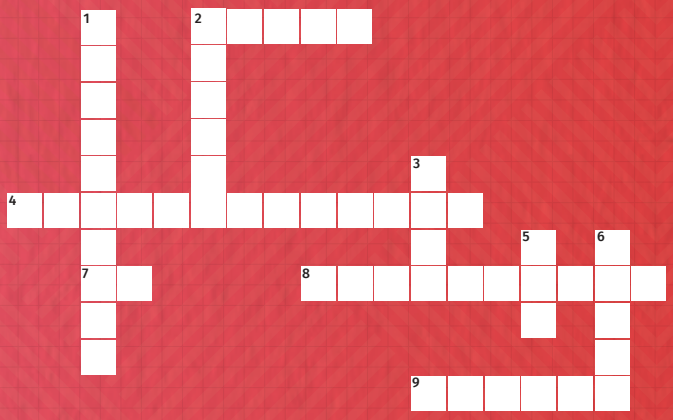
Conclusion

Organizations need to keep up to date and stay informed about the latest research and developments in the space of AI-powered security attacks and ways to prevent/remediate the exploits. Perform regular security audits to detect security vulnerabilities and make sure your infrastructure is compliant and secure. Proactively take measures to prevent these advanced security exploits. Invest in generative AI-powered security tools to take advantage of the benefits they offer in combating the fast-evolving cyber threats. Provide adequate training to your teams and create awareness about AI security risks and ways to take advantage of them securely.

Contact for more details

State Informatics Officer
NIC, Tamil Nadu State Centre
E2-A, Rajaji Bhavan, Besant Nagar
Chennai, Tamil Nadu 600090
Email: sio.tn@nic.in, Phone: 044-24917850

NICROSSWORD #1



Scan QR code for the answers

Across

2. Flagship Haryana initiative for proactive service delivery
4 F.ramework enabling modular design in software like SATHI
7. Technology that powers predictive governance and chatbots
8. Collaboration tools for government file management
9. For faster development and deployment of digital platforms

Down

1. Ensures secure record-keeping in land records and contracts
2. Website hosting service ensuring scalability and security
3. National e-Vidhan Application for legislative digitization
5. Used in disaster management for mapping and planning
6. Aadhaar-enabled Public Distribution System for ration distribution