March 2026 | Volume 2 | Issue 1 | Article No. 26202

Journal of Smart Sensors and Computing

Review Article | *Open Access* |

Artificial Intelligence and Machine Learning in Cybersecurity: A Review of Trends, Challenges, and Applications

Pooja Soni and Sanjay Gour*

Department of Computer Science & Engineering, Gandhinagar University, Gandhinagar, 382725, Gujarat, India
**Email:* sanjay.since@gmail.com (S. Gour)

Abstract

In the rapidly evolving digital environment, cybersecurity has become a critical component for protecting cloud infrastructures, Internet of Things (IoT) ecosystems, 5G networks, and sensitive digital assets. The increasing complexity and scale of cyber threats-including zero-day attacks, advanced persistent threats, ransomware, and social engineering-have revealed the inadequacy of traditional rule-based and signature-driven security systems. Consequently, Artificial Intelligence (AI) and Machine Learning (ML) have emerged as essential enablers in the advancement of modern cybersecurity solutions. This study provides a comprehensive overview of the evolving threat landscape, attack vectors, and the broader societal implications of cybersecurity. It presents an integrated theoretical and practical perspective on AI- and ML-driven cybersecurity frameworks, highlighting their roles in automated threat detection, behavioral analysis, predictive analytics, and security automation. The paper also examines key theoretical foundations, including behavioral modeling, information theory, adversarial learning, pattern recognition, and automation theory, while addressing emerging challenges such as adversarial machine learning, algorithmic bias, explainability, and data privacy. Furthermore, the review explores recent advancements, including federated learning and Explainable Artificial Intelligence (XAI), along with a statistical analysis of global adoption and utilization trends. It also discusses leading countries in the implementation of AI-driven cybersecurity solutions, supported by relevant data. By synthesizing current research and industry practices, the study proposes a structured roadmap for developing robust, adaptive, and ethically responsible AI-enabled cybersecurity systems capable of addressing both current and future cyber threats. Additionally, the study provides a statistical perspective on technology adoption, market growth, and the expanding role of AI and ML in cybersecurity applications.

***Keywords*:** Artificial intelligence; Machine learning; Cybersecurity; Explainable AI; Privacy.

Received: 31 December 2025; Revised: 22 February 2026; Accepted: 13 March 2026; Published Online: 16 March 2026.

1. Introduction

Owing to the hyperconnected advanced digital environment, cybersecurity has become a critical component of modern infrastructure, ensuring the protection of financial institutions, government systems, commercial operations, and individual privacy. The rapid expansion of the IoT, cloud computing, and 5G networks has significantly increased the attack surface, rendering traditional security mechanisms increasingly inadequate for addressing complex and large-scale threats.[1] These challenges are further intensified by the growing sophistication of cyber adversaries, who employ advanced tactics, techniques, and procedures (TTPs), including zero-day exploits, advanced persistent threats (APTs), social engineering, and ransomware, to evade conventional security measures. As a result of the limitations of static, rule-based, and signature-driven methods, the cybersecurity landscape has shifted toward Artificial Intelligence (AI) and Machine Learning (ML)-based adaptive defense mechanisms. AI and ML offer dynamic capabilities for threat detection, classification, and response

DOI: https://doi.org/10.64189/ssc.26201

by leveraging large-scale datasets and continuously adapting to emerging attack patterns, enabling a transition from reactive to proactive security strategies.[2] These technologies support real-time threat detection, behavioral analysis, and automated incident response, significantly enhancing threat intelligence. Their ability to identify previously unknown threats through anomaly detection, automate routine security tasks, and reduce response time demonstrates a level of adaptability unattainable with traditional systems.[3,4] Consequently, AI and ML have become integral to modern cybersecurity, supporting advanced applications such as intrusion detection, malware analysis, user behavior monitoring, and threat intelligence generation.[5] The review further extends its analysis to adversarial machine learning-an emerging and critical concern in cybersecurity-where malicious actors exploit AI models to evade or manipulate defensive systems, thereby introducing new risks and vulnerabilities into intelligent security frameworks.[6]

Fig. 1: AI & ML in cybersecurity.

Currently, AI- and ML-powered cybersecurity systems are increasingly incorporating Explainable Artificial Intelligence (XAI) to enhance transparency and trust. The integration of AI with quantum computing is emerging as a promising approach for addressing complex security challenges, while federated learning enables collaborative threat intelligence without compromising data privacy. Despite these advancements, AI-enabled security frameworks still face challenges in detecting zero-day attacks due to their reliance on historical datasets. Additionally, adversarial machine learning attacks can manipulate models through carefully crafted inputs, leading to misclassification and the evasion of security mechanisms. Furthermore, the limited interpretability of AI-driven decisions restricts their adoption in mission-critical applications. To address these challenges, this study analyzes state-of-the-art AI/ML techniques, evaluates their practical effectiveness, and proposes guidelines to enhance their performance, robustness, and interpretability.[8,9] Key application areas of AI and ML in cybersecurity include behavioral analysis, automated response, malware classification, threat intelligence, and intrusion detection.

1.1 Rising complexity of cyber threats

The evolution of cybersecurity represents a continuous cycle of adaptation, driven by rapid technological advancements and the parallel emergence of increasingly sophisticated cyber threats. In its early stages, cybersecurity efforts focused on protecting standalone systems from relatively simple forms of malware, such as viruses and worms, which were typically transmitted through physical media, including floppy disks and CDs, in the absence of widespread network connectivity.[8] With the advent of the internet in the late twentieth century, the proliferation of digital networks significantly expanded the attack surface. This transformation led to the emergence of network-based threats, particularly denial-of-service (DoS) and distributed denial-of-service (DDoS) attacks.

The 1990s and early 2000s marked a significant transformation in the cybersecurity threat landscape. The rapid growth of e-commerce, digital banking, and global internet access created lucrative opportunities for cybercriminals. During this period, attackers evolved from isolated individuals into organized groups driven by financial gain, political motives, and ideological objectives. Phishing attacks became increasingly prevalent, relying on deceptive communication techniques to obtain sensitive personal and financial information. Concurrently, botnets-large networks of compromised systems-were remotely controlled by attackers to conduct coordinated malicious activities. The emergence of sophisticated malware, including trojans and ransomware, further emphasized the need for more robust and advanced security solutions. As digital systems became increasingly interconnected and data gained strategic value, both state-sponsored actors and organized criminal groups began leveraging cyberattacks for espionage, disruption, and financial exploitation. Consequently, traditional security mechanisms such as antivirus software, firewalls, and intrusion detection systems (IDSs) proved increasingly inadequate in addressing these evolving threats, highlighting the limitations of static, signature-based defense approaches.[10]

In the modern digital age, cyber threats have reached unprecedented levels of scale and complexity, targeting critical infrastructure, mobile environments, and cloud platforms.[11] Contemporary attacks are characterized by stealth, automation, and persistence, with advanced persistent threats (APTs) posing significant concerns. These prolonged and targeted campaigns utilize social engineering, zero-day exploits, and cross-network techniques to evade traditional detection mechanisms. Ransomware has emerged as a dominant threat, often employing double-extortion strategies that combine data encryption with threats of public disclosure. Phishing attacks have also evolved into highly targeted spear-phishing campaigns, significantly increasing

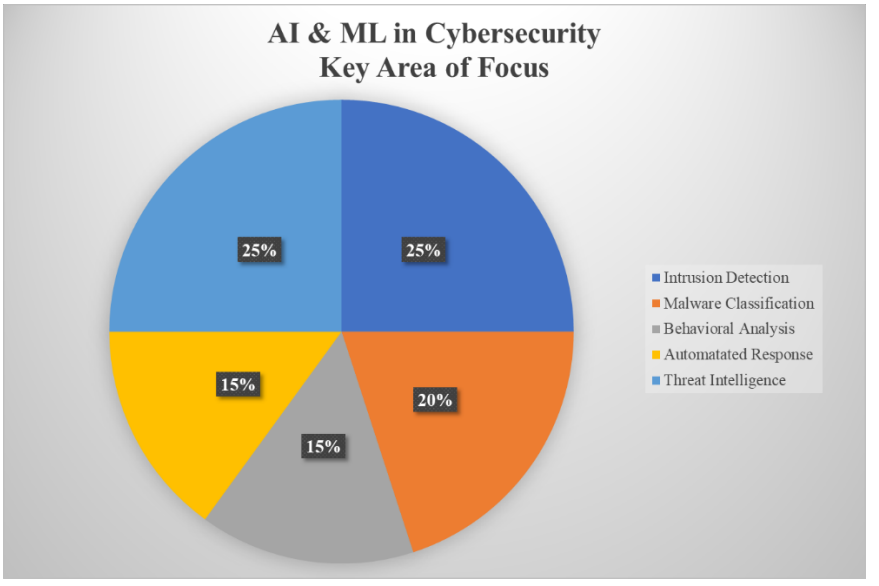

Fig. 2: AI & ML in cybersecurity: key area of focus. [7]

their success rates.[12] Furthermore, the rapid expansion of cloud computing and the Internet of Things (IoT) has broadened the attack surface, exposing vulnerabilities such as insecure devices and centralized data storage systems as attractive targets for attackers. The high level of interconnectivity and speed of modern digital systems has rendered traditional security approaches increasingly ineffective for real-time threat detection and mitigation.[13] To address these challenges, Artificial Intelligence (AI) and Machine Learning (ML) have become central to modern cybersecurity strategies. These approaches enable the analysis of large-scale datasets to detect anomalies and emerging attack patterns while continuously improving through adaptive learning mechanisms.[14] In contrast to traditional rule-based systems, AI/ML models can dynamically respond to novel threats, enhancing resilience against zero-day exploits and advanced attack techniques. They also improve incident response by automating detection, analysis, and mitigation processes, thereby significantly reducing response time. Despite challenges associated with techniques such as adversarial machine learning, AI and ML provide a scalable and adaptive framework for developing robust cybersecurity systems capable of addressing the complexities of the modern threat landscape.[15]

1.2 Limitations of traditional cybersecurity mechanisms

Traditional cybersecurity mechanisms have long served as the foundation of organizational defense strategies. Tools such as antivirus software, firewalls, and intrusion detection systems (IDSs) have been widely used to protect against known threats. However, as the threat landscape has evolved in scale and sophistication, these conventional methods have revealed significant limitations. One major weakness is their heavy reliance on signature-based detection techniques. Antivirus systems and many IDS solutions depend on predefined signatures to identify malicious code, making them ineffective against zero-day exploits and previously unseen malware variants.[16] As attackers increasingly employ polymorphic and metamorphic techniques to dynamically alter malicious code, signature-based tools struggle to keep pace.

Another limitation lies in the dependence on rule-based frameworks. Firewalls and traditional security management systems operate based on manually defined rules and policies. While these rules may block known attack patterns, they often generate high rates of false positives and false negatives, overwhelming security analysts and contributing to alert fatigue.[17] Furthermore, these systems lack contextual awareness and behavioral analysis capabilities, making it difficult to detect insider threats or sophisticated attacks that mimic legitimate activity. Scalability presents an additional challenge. With the exponential growth of mobile devices, cloud computing, and the Internet of Things (IoT), network traffic volumes have increased dramatically. Traditional systems were not designed to efficiently process large-scale, continuous data streams, leading to delayed detection and response.[18] Moreover, static defense mechanisms require frequent manual updates and patch management, creating windows of vulnerability between threat emergence and mitigation. Finally, traditional cybersecurity approaches are largely reactive rather than proactive. They respond to known threats only after signatures or patches have been developed, rather than anticipating or preventing emerging attack vectors. As cyber adversaries increasingly adopt automated and adaptive techniques, these limitations underscore the urgent need for more intelligent, scalable, and dynamic cybersecurity solutions.[19]

1.3 Why AI/ML is transformative in cybersecurity

Currently, Artificial Intelligence (AI) and Machine Learning (ML) are transforming cybersecurity by enabling adaptive,

data-driven, and proactive defense mechanisms. Traditional security systems rely heavily on static rules and signature-based detection, which are ineffective against zero-day attacks and rapidly evolving malware variants.[19] In contrast, AI/ML approaches analyze large volumes of both structured and unstructured data - such as system logs, network traffic, and user behavior - to detect anomalies and previously unseen attack patterns in real time.

A key transformative capability of AI/ML is behavioral analysis. Rather than relying solely on known signatures, ML models establish baseline patterns of normal activity and identify deviations that may indicate malicious behavior.[20] This capability significantly enhances the detection of advanced persistent threats (APTs), polymorphic malware, and insider threats. Moreover, AI-driven automation improves incident response by prioritizing alerts, reducing false positives, and accelerating threat mitigation processes.[21]

AI/ML systems also continuously learn from new threat intelligence, enabling dynamic adaptation to evolving attack techniques. This analytical and self-improving capability allows organizations to transition from reactive defense models to proactive security strategies. Despite challenges such as adversarial attacks on ML models, AI and ML provide scalable, intelligent, and resilient solutions that redefine modern cybersecurity frameworks.

Intrusion Detection Systems (IDSs) are critical security tools designed to detect a wide range of threats, including external attacks and insider misuse. Their primary function is to continuously monitor network traffic and system activities, analyze events, generate alerts, and respond to suspicious or malicious behavior.

Deep learning (DL), a subset of machine learning, has emerged as a powerful approach in cybersecurity. Techniques such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) are widely used to address complex problems, including malware classification and intrusion detection. These methods enable the development of proactive defense systems capable of identifying evolving threats, thereby providing a more dynamic and intelligent approach to cybersecurity.[22]

1.4 Unified taxonomy of AI/ML methods in cybersecurity

A unified taxonomy of Artificial Intelligence (AI) and Machine Learning (ML) approaches in cybersecurity provides a structured framework for understanding how diverse learning models are applied to various security challenges. These approaches can be broadly categorized based on learning paradigms and application domains. From a methodological perspective, the primary categories include supervised, unsupervised, semi-supervised, reinforcement, and deep learning.

Supervised learning algorithms, such as Support Vector Machines (SVMs), Random Forests (RFs), and Artificial Neural Networks (ANNs), rely on labeled datasets and are widely used for malware classification, spam filtering, and intrusion detection.[20] In contrast, unsupervised learning techniques, including clustering algorithms such as K-means and DBSCAN, are particularly useful for anomaly detection when labeled data are scarce or unavailable.[23]

Semi-supervised learning combines labeled and unlabeled data to improve detection accuracy in environments with limited annotated datasets. Reinforcement learning focuses on adaptive decision-making and is increasingly applied to dynamic threat response and automated defense mechanisms.

Deep learning methods, including Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), enhance feature extraction and pattern recognition in complex data sources such as system logs and network traffic.[24] In particular, RNNs and Long Short-Term Memory (LSTM) networks are well-suited for analyzing sequential data, enabling effective monitoring of user behavior over time. Unlike traditional machine learning models (e.g., SVM, RF, and K-Nearest Neighbors (KNN)), which treat data points independently, RNNs and LSTMs capture temporal dependencies, making them highly effective for behavioral analysis.

Reinforcement learning (RL) represents an advanced paradigm in which models learn through interaction with their environment, receiving feedback in the form of rewards or penalties. In cybersecurity, RL shows strong potential for adaptive defense systems, such as intrusion prevention systems (IPS), where models learn optimal response strategies against different types of attacks.

Furthermore, hybrid and ensemble learning approaches combine multiple models to improve overall performance and robustness. In cybersecurity contexts, where threats are diverse and evolving, such integrated approaches offer significant advantages by leveraging the strengths of different algorithms.[25,26]

From an application perspective, AI/ML techniques are employed across various domains, including network security, endpoint protection, IoT security, cloud security, and threat intelligence analytics. This unified taxonomy facilitates systematic evaluation and benchmarking in research while supporting the development of hybrid models that integrate multiple learning paradigms for enhanced performance and adaptability.[18]

1.5 Comparative analysis of AI/ML applications in cybersecurity

A comparison between traditional cybersecurity methods and AI/ML-enabled approaches clearly demonstrates the transformative potential of intelligent systems across various security domains. Traditional malware detection tools primarily rely on signature-based scanning, which matches files against known threat databases. Although effective against known malware, these systems struggle to detect novel, complex, or polymorphic variants.

Table 1: Comparison of traditional and AI-ML applications.

Application Area	Traditional Approach	AI ML-Based Approach	Advantages
Malware Detection	Signature scanning	Deep learning classification	Detects unknown variants
Intrusion Detection	Rule-based IDS	Anomaly based ML IDS	Lower false negatives
Phishing Detection	Blacklist filtering	NLP-based ML models	Detects zero-day phishing
Fraud Detection	Manual rule engines	Predictive ML models	Real-time fraud prevention
Threat Hunting	Manual log review	Automated anomaly detection	Faster response time

In contrast, AI-enabled approaches, particularly those based on deep learning, analyze behavioral characteristics and structural patterns, enabling the detection of previously unknown threats.

In intrusion detection, traditional rule-based systems rely on predefined signatures and static patterns, which limit their ability to detect sophisticated or evolving attacks. In contrast, machine learning-based anomaly detection systems overcome these limitations by learning normal network behavior and identifying deviations, thereby reducing false negatives and improving the detection of complex intrusions. Phishing detection further illustrates this transformation. Conventional blacklist-based filtering blocks known malicious domains but fails to identify newly created phishing websites. In contrast, Natural Language Processing (NLP)-based machine learning approaches analyze linguistic and contextual features to detect zero-day phishing attempts more effectively. Similarly, fraud detection has evolved from manual rule-based systems to data-driven machine learning models capable of analyzing transactions in real time. Automated anomaly detection has also replaced manual log analysis in threat hunting processes. Overall, AI-driven systems enhance detection accuracy, reduce response time, and support the development of adaptive and scalable cybersecurity defenses.

1.6 Open challenges and research gaps

Despite significant advancements, the integration of Artificial Intelligence (AI) and Machine Learning (ML) in cybersecurity presents several critical challenges. One of the primary concerns is adversarial attacks, in which malicious actors deliberately craft deceptive inputs to mislead machine learning models and evade detection systems. These attacks expose vulnerabilities in AI-driven defenses and highlight the need for robust and resilient model architectures.

Another major challenge is data imbalance. In real-world cybersecurity datasets, normal activity significantly outweighs malicious instances, which can bias models and reduce their ability to accurately detect rare but critical threats. This issue is closely related to the lack of standardized datasets, as many existing benchmark datasets are outdated, synthetic, or insufficiently representative of modern threat environments, thereby limiting reliable performance evaluation. The lack of explainability in deep learning models also raises significant concerns. Many AI systems operate as "black boxes," making it difficult for analysts to interpret decisions, justify alerts, or comply with regulatory requirements. Additionally, privacy concerns arise because effective ML training often requires large volumes of sensitive user and organizational data.

High computational costs further restrict deployment, particularly in resource-constrained environments. Moreover, model drift caused by continuously evolving cyber threats can degrade performance over time, necessitating regular retraining and maintenance to ensure sustained effectiveness.

2. Threat landscape, attack vector and security goals

An understanding of the threat landscape, attack vectors, and security objectives is essential for developing effective and resilient cybersecurity strategies. These foundational elements enable organizations to design robust defense mechanisms that anticipate the evolving tactics of adversaries. They also ensure alignment between security policies, organizational objectives, and risk tolerance levels. Furthermore, they support proactive approaches such as threat hunting, continuous monitoring, and AI-assisted detection systems. In today's dynamic digital environment, the integration of traditional and advanced security objectives enables organizations to develop comprehensive, adaptive, and future-ready cybersecurity frameworks.

2.1 Threat landscape in cybersecurity

In cybersecurity, the threat landscape refers to the dynamic and evolving ecosystem of threat actors, attack techniques, targets, and vulnerabilities that collectively define cyber risk. The present study highlights that the contemporary threat landscape is becoming increasingly complex due to rapid digital transformation, widespread adoption of cloud services, the proliferation of the Internet of Things (IoT), and the integration of Artificial Intelligence (AI) in both offensive and defensive cyber operations.

A key characteristic of the modern threat landscape is the professionalization of cybercrime. Organized cybercriminal groups now operate as structured enterprises, offering phishing kits, Ransomware-as-a-Service (RaaS), and exploit tools on underground markets. As reported by Dave et al., the commodification of cyberattack tools has significantly lowered the barrier to entry, enabling less-skilled actors to launch sophisticated attacks. This trend has led to a surge in ransomware campaigns targeting sectors such as finance, healthcare, academia, and critical infrastructure.[27]

Another major development is the rise of Advanced Persistent Threats (APTs). These groups, often state-

sponsored, conduct long-term, covert operations aimed at espionage, intellectual property theft, and geopolitical disruption. APT actors employ multi-stage attack strategies, including zero-day exploits, spear-phishing, and lateral movement techniques, to maintain persistent access within compromised systems.

The growing adoption of AI has also introduced AI-driven cyber threats. As highlighted by Erukude *et al.*, adversaries are leveraging machine learning for automated vulnerability scanning, adaptive malware development, and highly targeted phishing campaigns. Additionally, synthetic media and deepfake technologies further complicate the threat landscape by enabling identity spoofing and misinformation attacks, thereby increasing both the scalability and success rate of cyberattacks.[28]

The expansion of cloud computing and IoT ecosystems has further broadened the attack surface. Billions of interconnected devices often lack adequate security controls, making them attractive targets for botnets and distributed denial-of-service (DDoS) attacks. Similarly, cloud misconfigurations, insecure APIs, and weak identity and access management have emerged as common entry points for attackers.

Supply chain attacks have also become a significant concern. Instead of directly targeting an organization, attackers exploit vulnerabilities in trusted third-party vendors. This indirect approach enables large-scale compromise through legitimate software updates or dependencies. The SolarWinds incident remains a prominent example widely discussed in contemporary cybersecurity literature.[29]

From a risk perspective, the modern threat landscape is characterized by increased automation, higher attack frequency, and cross-border complexity. Threat actors exploit zero-day vulnerabilities and use encryption techniques to evade detection. Furthermore, geopolitical tensions have contributed to a rise in cyber warfare activities targeting national infrastructure.

Overall, the cybersecurity threat landscape has evolved beyond isolated malware incidents or individual hackers. It is now defined by a sophisticated, interconnected ecosystem driven by financial incentives, political motivations, and technological advancements. Continuous monitoring, threat intelligence sharing, AI-driven defense mechanisms, and global collaborative governance are essential for mitigating these evolving risks.[30]

2.2 Attack vectors

Attack vectors are the specific pathways or methods used by threat actors to gain unauthorized access to applications, systems, and networks by exploiting vulnerabilities. Understanding cyberattack vectors is fundamental to cybersecurity, as it defines the mechanisms through which threats materialize into actual security incidents.

Recent research and industry reports indicate that modern attack vectors are continuously evolving, driven by advancements in Artificial Intelligence (AI), cloud adoption, automation, and human-related factors. These developments have made attacks more frequent, sophisticated, and difficult to detect. The key attack vectors are as follows:

2.2.1 Malware

Malicious software (malware) refers to any program or code designed to disrupt system operations, compromise security, or exploit vulnerabilities in computer systems and networks. Threat actors develop and deploy malware for various purposes, including data theft, financial gain, system disruption, and espionage.[31]

Common types of malwares include the following:

Viruses: Self-replicating programs that attach themselves to legitimate files or applications. They spread when infected files are executed.

Worms: Standalone programs that can automatically replicate and spread across networks and devices without requiring user interaction.

Trojans (Trojan Horses): Malicious programs that disguise themselves as legitimate software while concealing harmful code.

Ransomware: A type of malware that encrypts user data and restricts access until a ransom is paid.

Spyware: Software that covertly monitors user activity and collects sensitive information, which is then transmitted to unauthorized parties.

Adware: Unwanted software that displays intrusive advertisements, often degrading system performance.

Rootkits: Stealthy programs that gain privileged access to a system while concealing their presence, often enabling persistent unauthorized control.

Botnets: Networks of compromised devices (referred to as "bots" or "zombies") controlled by a central command system, commonly used to launch coordinated attacks such as distributed denial-of-service (DDoS) attacks or spam campaigns.

2.2.2. Phishing

Phishing is a type of cyberattack in which individuals are deceived into disclosing sensitive information, such as login credentials, personal data, and credit card details, by impersonating a trusted entity. The term "phishing" is derived from "fishing," reflecting the idea of luring victims, while "ph" originates from "phone phreaking," a technique used to exploit telephone systems in the 1970s.[32,33]

Phishing attacks occur in various forms, each employing distinct techniques and targeting strategies. Common types of phishing attacks include the following:

1. Email phishing: A deceptive practice in which attackers send fraudulent emails that appear to originate from legitimate sources to trick recipients into revealing sensitive information.
2. Spear phishing: A targeted form of phishing in which

attackers tailor messages to specific individuals or organizations to increase the likelihood of success.

3. Whaling: A specialized form of spear phishing that targets high-profile individuals, such as executives or CEOs, with the aim of obtaining sensitive corporate information.
4. Vishing: Also known as voice phishing, this technique involves the use of phone calls to deceive victims into disclosing confidential information or performing actions such as fund transfers.
5. Smishing: A form of phishing conducted via SMS, where attackers send malicious links or request sensitive information through text messages.

2.2.3 Denial of Service (DoS) and Distributed Denial of Service (DDoS) attacks

Denial-of-Service (DoS) and Distributed Denial-of-Service (DDoS) attacks are malicious strategies designed to disrupt the functionality and availability of computer systems, networks, and websites. In a DoS attack, a single source overwhelms a targeted system with excessive traffic, rendering it unable to respond to legitimate user requests. In contrast, DDoS attacks involve multiple distributed sources that simultaneously flood the target, amplifying the scale and impact of the attack.

These attacks exploit limitations in system and network resources, such as bandwidth, memory, or processing power, resulting in service degradation or complete unavailability.[34]

Common types include:

1. Traditional DoS: An attack in which a single attacker or a small group generates excessive traffic toward a target system, often using automated tools or controlled systems.
2. Distributed Denial-of-Service (DDoS): A coordinated attack launched from numerous compromised devices (typically part of a botnet) to overwhelm a target system, making mitigation more challenging due to its distributed nature.

2.2.4 Zero-day exploits

Zero-day attacks are a class of cyberattacks that exploit previously unknown vulnerabilities in software or systems. These vulnerabilities, referred to as zero-day vulnerabilities, are not yet publicly disclosed or patched by vendors, making such attacks particularly difficult to detect and mitigate.

Because the vulnerability remains undiscovered, affected systems cannot be secured in advance, and traditional security solutions - such as signature-based antivirus systems - are unable to identify the threat.[35] Attackers exploit these weaknesses before developers have the opportunity to release updates or patches.

Zero-day vulnerabilities are often associated with targeted and sophisticated attacks, as evidenced by post-incident analyses linking them to advanced threat campaigns. These attacks are frequently leveraged by highly skilled cybercriminal groups and state-sponsored actors due to their ability to bypass conventional defenses and remain undetected for extended periods.

Key concepts include:

1. Zero-day vulnerability: A previously unknown and unpatched security flaw in software or systems that can be exploited by attackers.
2. Zero-day attack: The exploitation of a zero-day vulnerability to compromise systems before a fix or patch is available.

2.2.5 Man-in-the-Middle (MITM) Attacks

A Man-in-the-Middle (MITM) attack is a type of cyberattack

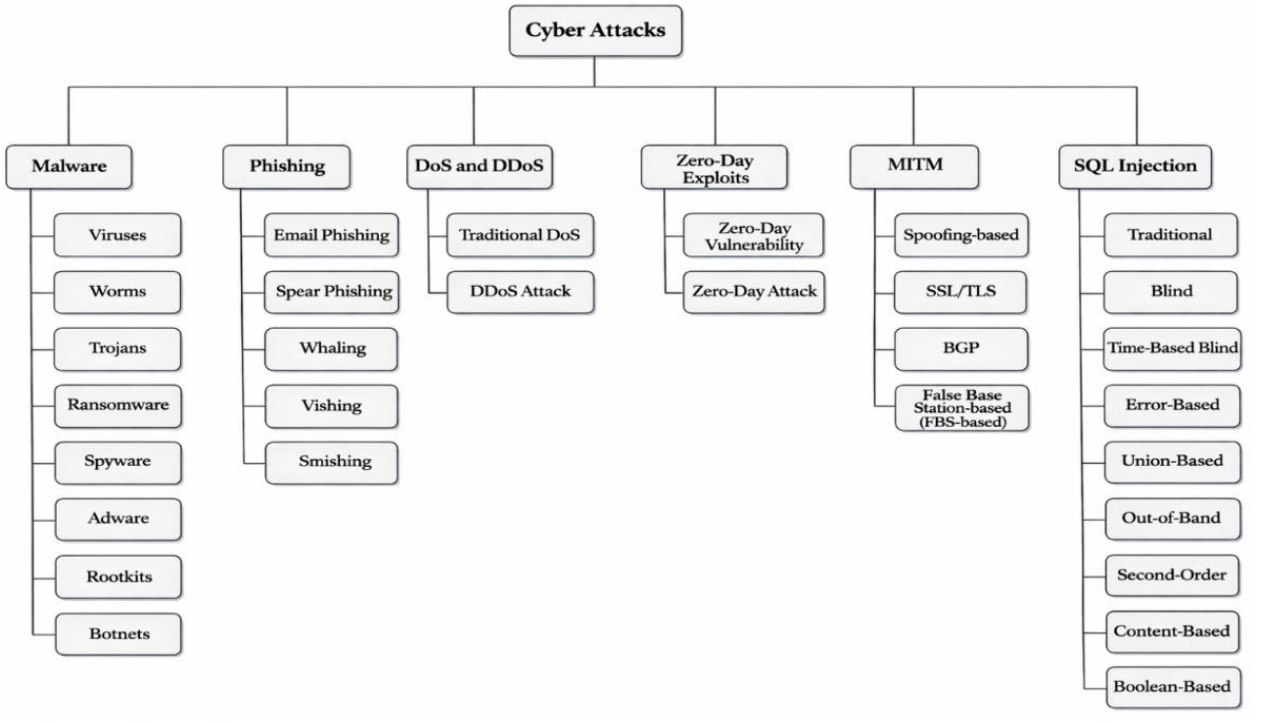

Fig. 3: Attack vectors.[38]

in which a malicious third-party intercept and potentially alters communication between two parties without their knowledge. In this scenario, the attacker positions themselves between the communicating entities, gaining the ability to eavesdrop on or manipulate transmitted data. The term "Man-in-the-Middle" metaphorically describes this interception, where an unauthorized entity intrudes into an otherwise secure exchange.[36]

A typical MITM scenario involves two legitimate endpoints and an attacker who intercepts the communication channel, enabling control over the exchanged messages. These attacks are particularly prevalent in environments such as public Wi-Fi networks, corporate systems, and even encrypted communication channels if encryption keys are compromised.

Effective defenses against MITM attacks include the implementation of secure encryption protocols such as HTTPS, robust authentication mechanisms, and secure key management practices.

Common types of MITM attacks include the following:

1. Spoofing-based MITM: In this attack, the adversary uses techniques such as IP spoofing or ARP spoofing to intercept and manipulate communication between two endpoints without their awareness.
2. SSL/TLS MITM: This attack involves intercepting secure communications by inserting the attacker between the client and server, often by exploiting certificate vulnerabilities or using fraudulent certificates.
3. BGP MITM: In this scenario, the attacker manipulates Border Gateway Protocol (BGP) routing to redirect network traffic through malicious nodes before reaching its intended destination.
4. False Base Station (FBS)-based MITM: This attack forces victims to connect to a rogue base station (e.g., a fake cellular tower), allowing the attacker to intercept and control communication traffic.

2.2.6 SQL injection attacks

SQL injection is a critical security vulnerability in web applications that allows attackers to manipulate database queries by injecting malicious input. The presence of SQL injection vulnerabilities can grant unauthorized access to databases, potentially exposing sensitive user information.[37] Since databases often store confidential data, such as personal and financial information, successful exploitation can lead to severe security breaches and significant consequences. SQL injection is fundamentally a code injection attack in which user-supplied input is improperly incorporated into SQL queries, causing the input to be executed as part of the query. Attackers can exploit this weakness to execute arbitrary SQL commands, retrieve sensitive data, modify database contents, or bypass authentication mechanisms.

Common types of SQL injection attacks include the following:

1. Traditional SQL injection: Direct injection of malicious SQL code into input fields to manipulate database queries.
2. Blind SQL injection: Exploits vulnerabilities where the application does not return explicit error messages, requiring attackers to infer information from application responses.
3. Time-based blind SQL injection: A subtype of blind SQL injection where attackers use time delays to extract information from the database.
4. Error-based SQL injection: Relies on database error messages to gather information about the database structure.
5. Union-based SQL injection: Uses the SQL UNION operator to combine results from multiple queries and extract additional data from the database.

2.3 Security goals

Security goals form the foundation of cybersecurity frameworks and guide organizations in planning, implementing, and evaluating protection mechanisms. These goals are broadly categorized into classical (traditional) and contemporary (emerging or extended) objectives.

Traditional security goals establish the fundamental principles of information security, while contemporary goals address the evolving requirements of modern digital environments. With the increasing complexity of the threat landscape, recent research emphasizes the need to expand beyond traditional objectives to incorporate additional security considerations that enhance resilience, adaptability, and comprehensive protection.[29]

2.3.1 Classic security goals

The fundamental security objectives are commonly defined by the CIA Triad, which provides a foundational framework for evaluating the effectiveness of security systems.

Confidentiality: Ensures that sensitive information is accessible only to authorized individuals or systems. It prevents unauthorized disclosure of data such as personal records, financial information, intellectual property, and confidential organizational assets. Mechanisms such as encryption, authentication protocols, access control policies, and secure communication channels are employed to maintain confidentiality.

Integrity: Ensures that data remains accurate, consistent, and unaltered except by authorized entities. It protects information from unauthorized modification, whether intentional or accidental. Techniques such as digital signatures, hashing, checksums, and version control systems are used to preserve data integrity during storage and transmission.

Availability: Ensures that authorized users can access data, systems, and applications when required. Disruptions may result from hardware or software failures, cyberattacks (such as distributed denial-of-service (DDoS) attacks), or natural disasters. To maintain availability, organizations implement measures such as redundancy, backup systems, disaster recovery planning, and fault-tolerant architectures.

Collectively, these three principles provide a structured approach to securing information assets and are widely referenced in cybersecurity frameworks and standards.[30]

2.3.2 Expanded (contemporary) security goals

The current study extends outside the CIA Triad to resolve unconventional and tenacious threats. The contemporary expanded goals include the following:

Nonrepudiation: This ensures that a person or entity cannot reject the execution of specific tasks, such as sending an SMS or permitting a transaction. The digital signature and cryptographical system are usually utilized to achieve this objective.

Accountability: Accountability requires that entire actions inside a system might be examined to accountable entities. Logging procedures, audit tracks and monitoring schemes allow organizations to sense mismanagement and enforce obedience.

Resilience refers to the ability of schemes to endure, acclimatize to, and improve from cyberattacks or disturbances. Unlike availability, which emphasizes admittance, flexibility emphasizes the endurance of actions throughout and after occurrences. AI-enabled threat recognition, adaptive safety protocols and zero-trust designs are being progressively used to reinforce resilience.[30]

3. Rudiments of AI and ML in cybersecurity

In contemporary times, cybersecurity has become a critical component of information technology due to the increasing reliance of various intellectual and operational activities on digital information, including documentation, reporting, and financial transactions. Consequently, networks must be continuously secured against evolving cyber threats.[39] Security strategies must adapt by learning from dynamic environments and incorporating novel approaches to identify and mitigate vulnerabilities in existing systems.

3.1 Artificial intelligence in cybersecurity

AI Artificial Intelligence (AI) has become a fundamental component of contemporary cybersecurity, enabling more accurate threat detection, response, and mitigation. Its primary focus lies in threat identification, automation, and intelligent decision-making. AI enables security systems to learn, adapt, and respond effectively to increasingly complex attack scenarios.

Unlike traditional security solutions that rely on static rules and signature-based approaches, AI-enabled systems can identify subtle patterns and anomalies within large volumes of data, even in the case of previously unknown or zero-day attacks.[40] By analyzing system traffic and behavioral patterns in real time, AI enhances the detection of malicious activities that might otherwise remain undetected. Furthermore, AI-driven automation reduces the operational burden on security teams by handling repetitive tasks such as log analysis, vulnerability assessment, and patch management, thereby improving the speed and efficiency of incident response.[41] Through predictive analytics, AI enables proactive defense mechanisms by anticipating potential threats and vulnerabilities, allowing organizations to take preventive actions before attacks occur.[42]

Beyond reactive defense, AI also supports proactive strategies such as threat hunting and forensic analysis, thereby strengthening long-term cybersecurity resilience.[43]

3.2 Machine learning standards in cybersecurity

Machine Learning (ML) is a major subfield of Artificial Intelligence (AI) and plays a vital role in cybersecurity by enabling systems to learn from data without explicit programming. ML techniques are typically classified into three primary categories: supervised learning, unsupervised learning, and reinforcement learning, each offering distinct advantages for cybersecurity applications.[44]

Supervised Learning: Supervised learning relies on labeled datasets to train models for predictive tasks. In cybersecurity, it is widely used for malware detection, intrusion detection systems (IDS), and attack classification.[45] Although effective for identifying known threats, its performance depends heavily on the availability and quality of labeled data, which limits its effectiveness against emerging or unknown attacks.

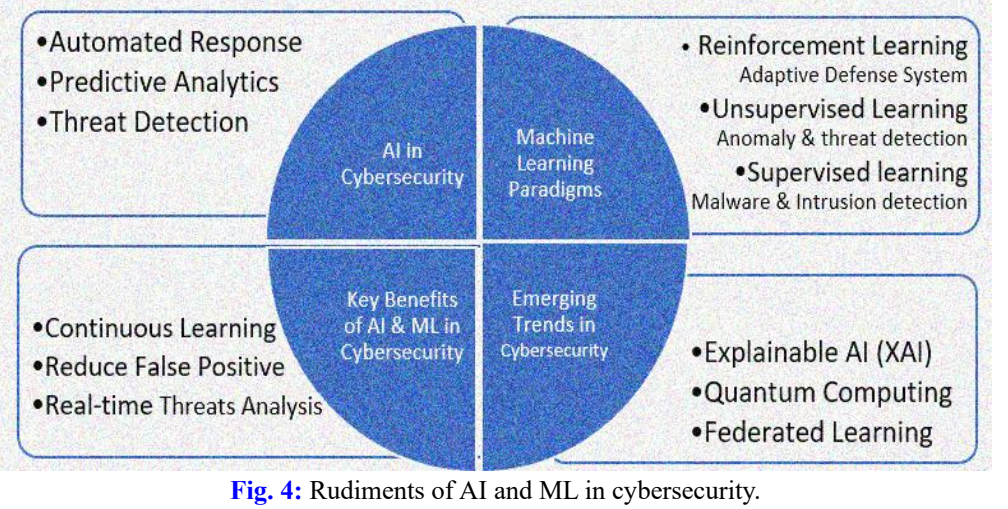

Fig. 4: Rudiments of AI and ML in cybersecurity.

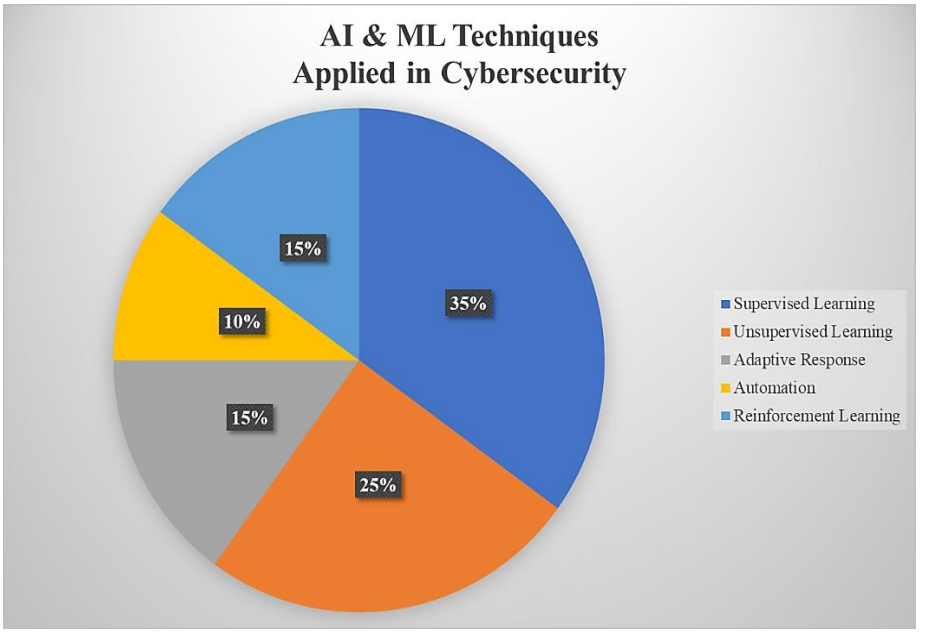

Fig. 5: Share of learning approaches of ML.[7]

Unsupervised Learning: Unsupervised learning focuses on identifying hidden patterns and anomalies in unlabeled data. This approach is particularly effective for detecting zero-day attacks, insider threats, and advanced persistent threats by recognizing deviations from normal behavior.[38] It also supports clustering and analysis of evolving threat patterns.

Reinforcement Learning: Reinforcement learning enables systems to learn optimal defense strategies through interaction with their environment. In cybersecurity, it is applied in adaptive intrusion prevention and automated penetration testing, allowing systems to improve their responses over time based on feedback.[46]

3.3 Synergy amid AI and cybersecurity

The integration of Artificial Intelligence (AI) and Machine Learning (ML) creates a dynamic and adaptive cybersecurity framework capable of addressing real-time threat analysis, automated responses, and continuous learning. This synergy enhances detection accuracy, reduces false positives, and improves threat intelligence, enabling a transition from reactive to proactive security strategies.[47]

4. Systematic literature review approach

The present study adopts a Systematic Literature Review (SLR) approach to examine the role of Artificial Intelligence (AI) and Machine Learning (ML) in cybersecurity. A structured and well-defined search strategy was employed, incorporating relevant keywords, subject domains, and predefined inclusion criteria to identify pertinent scholarly publications.

The literature search primarily focused on open-access and publicly available research articles retrieved through academic search engines and scholarly platforms such as Google Scholar, ResearchGate, and Academia.edu. In addition, references were obtained from publications by well-established academic publishers, including IEEE, Springer, and ACM, as well as from internationally recognized journals indexed in Scopus, the Science Citation Index (SCI), and Web of Science (WoS).

The selection and data collection procedures follow the guidelines proposed by Nachaat Mohamed, ensuring a systematic and transparent review process. To provide a broader perspective, relevant technical reports, web-based resources, and selected industry publications were also included where appropriate.

Furthermore, statistical analyses and graphical representations were incorporated to highlight recent technological developments, market trends, and the economic implications of AI-driven cybersecurity solutions. As the study relies entirely on publicly available information, it aims to provide researchers, practitioners, and general readers with a comprehensive understanding of the intersection of Artificial Intelligence, Machine Learning, and cybersecurity.

5. Hypothetical framework

The theoretical foundation of this study is rooted in the integration of Artificial Intelligence (AI), Machine Learning (ML), and cybersecurity. This integration represents a significant shift from traditional, manual, and reactive security mechanisms toward intelligent, adaptive, and automated defense systems. By leveraging insights from computer science, information theory, behavioral science, and adversarial learning, this framework explains how AI and ML enhance threat detection, response, and prevention in modern cybersecurity environments.

These interdisciplinary approaches contribute to the development of robust security systems capable of addressing complex and evolving cyber threats.[7]

5.1 Role of AI and ML in cybersecurity

Artificial Intelligence (AI) and Machine Learning (ML) form the core foundation of intelligent cybersecurity by enabling automation and data-driven decision-making. Supervised learning approaches support classification-based security tasks such as malware detection, phishing detection, and intrusion detection by utilizing labeled datasets.

Unsupervised learning facilitates anomaly detection by identifying deviations from normal behavior, making it particularly effective for detecting zero-day attacks and insider threats. Reinforcement learning further enhances adaptive defense mechanisms by enabling systems to optimize security decisions through reward-based feedback, thereby supporting continuous improvement in response to evolving threat scenarios.

5.2 Information theory and pattern recognition

Information theory provides essential concepts such as entropy and information gain to quantify uncertainty and detect deviations in security data. Pattern recognition techniques support AI models in identifying meaningful structures within complex datasets, including network traffic and user behavior data. Methods such as classification, clustering, and dimensionality reduction enable the detection of subtle anomalies that may indicate malicious activities.[48,39]

5.3 Behavioral analysis and user profiling

Behavioral analysis in cybersecurity draws on concepts from cognitive science and psychology to model normal user behavior and identify deviations that may indicate potential threats. AI systems continuously learn baseline user behavior and adapt to changes, enabling them to distinguish between benign anomalies and malicious activities. This approach is particularly effective for detecting insider threats and credential misuse.

5.4 Adversarial learning and game theory

Adversarial machine learning focuses on securing AI systems against malicious manipulation. Based on principles from game theory, this approach models attackers and defenders as strategic agents. Concepts such as Nash equilibrium guide the development of robust defense strategies, while techniques such as adversarial training enhance the resilience of models against deceptive or manipulated inputs.

5.5 Automation theory and security orchestration

Automation theory enhances security orchestration by enabling the autonomous management of detection, response, and remediation workflows. Predefined playbooks and decision-tree models streamline responses to security incidents, while AI continuously refines these processes to improve accuracy and response speed.

5.6 Predictive analytics and proactive defense

Predictive analytics leverages time series forecasting, probability theory, and statistical modeling to anticipate potential threats. AI-enabled risk scoring systems prioritize alerts based on likelihood and impact, enabling proactive defense strategies that reduce the success rate of cyberattacks.

5.7 Ethical considerations in AI-driven security

Ethical principles emphasize fairness, accountability, transparency, and privacy in AI-enabled cybersecurity. Explainable Artificial Intelligence (XAI) enhances interpretability, enabling security teams to understand and validate automated decisions while maintaining human oversight.

5.8 AI Bias and its implications for cybersecurity

AI bias can negatively affect threat detection and response accuracy. Bias in training data may lead to an imbalance that causes models to underrepresent emerging or rare threats. If an algorithm is biased, it may overemphasize dominant patterns while failing to detect less frequent but critical anomalies. Bias in threat attribution can ultimately result in inaccurate or ineffective responses.[39]

5.9 Mitigating AI bias in cybersecurity

To mitigate AI bias in cybersecurity, diverse and representative datasets should be utilized to identify and reduce bias during model training. Additionally, Explainable Artificial Intelligence (XAI) combined with human-in-the-loop (HITL) approaches can enhance transparency, accountability, and decision validation.

Privacy concerns remain a significant challenge in federated learning and threat intelligence sharing, particularly in collaborative environments involving multiple organizations and jurisdictions. Although federated learning is designed to enable decentralized model training without direct data sharing, it does not entirely eliminate privacy risks. Advanced attacks, such as model inversion and membership inference, can reconstruct sensitive information from shared models, thereby compromising data confidentiality. Similarly, threat intelligence sharing initiatives may expose proprietary or sensitive operational data if appropriate anonymization and access control mechanisms are not enforced. These risks are especially critical in collaborative defense environments where multiple entities contribute to and utilize shared intelligence. In addition to technical risks, regulatory and compliance requirements further complicate privacy preservation. Data protection regulations such as the GDPR, CCPA, and HIPAA impose strict guidelines on data collection, usage, consent, and cross-border data transfer. Ensuring compliance while enabling effective collaboration requires careful alignment between technical solutions and legal frameworks. To address these challenges, advanced privacy-preserving

techniques have gained prominence.[49,50] Secure multi-party computation enables joint analysis without exposing individual data inputs, while homomorphic encryption allows computation on encrypted data. Differential privacy introduces controlled noise to protect individual data points, and policy-driven data governance ensures accountability and controlled access.[51]

Furthermore, technical safeguards, accountability, transparency, and ethical AI governance play a crucial role in building trust. XAI mechanisms, combined with human oversight and well-defined ethical frameworks, ensure that AI-enabled cybersecurity systems remain fair, accountable, and reliable while effectively balancing security, performance, and privacy.[7]

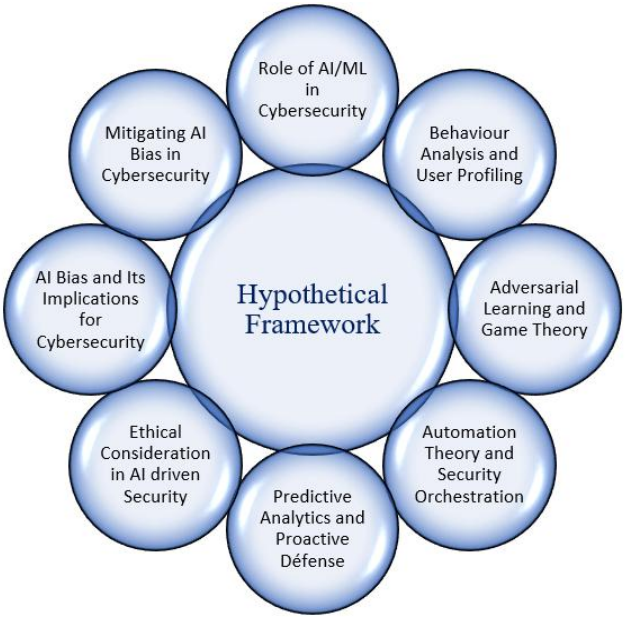

Fig. 6: Hypothetical framework of AI & ML with cybersecurity.

6. Worldwide AI & ML utilization statistics

Artificial Intelligence (AI) and Machine Learning (ML) are increasingly essential to global cybersecurity, with adoption rising significantly as cyber threats become more complex and frequent. According to recent data, approximately 57% of organizations worldwide use AI for anomaly detection, around 50% for malware detection, and about 49% for automating incident response as part of their cybersecurity strategies, highlighting its widespread application in defensive operations.[52]

Market studies indicate that the global AI in cybersecurity market is expanding rapidly, with a compound annual growth rate (CAGR) of approximately 23.6% through 2025. This growth reflects strong demand for intelligent threat detection, predictive analytics, and automated response mechanisms to counter emerging cyber threats.[53]

Another industry report suggests that approximately 74% of enterprises rely on AI-based tools for threat detection and prediction, while 58% of Fortune 500 companies utilize such solutions for proactive security, underscoring their importance in enterprise protection.[54]

In India, the adoption of AI and ML in cybersecurity is particularly notable. A 2025 study conducted by IDC and Fortinet found that approximately 94% of Indian organizations are currently using AI to detect, respond to, and predict cyber threats, indicating a shift from reactive defenses toward proactive, intelligence-driven systems.[55]

Additionally, another report revealed that approximately 97% of Indian organizations have invested in AI/ML technologies, reflecting widespread strategic integration across sectors. Despite this high level of adoption, organizations in India continue to face significant challenges from AI-enabled cyberattacks, with nearly 72% experiencing such threats in the past year. This highlights both the necessity and urgency of implementing advanced security measures.[56]

According to market estimates, India's share of the global AI cybersecurity market is approximately 4.7%, indicating a growing yet still developing presence on the global stage.[57] Overall, AI and ML are transforming cybersecurity worldwide-from threat detection to automated incident response-while in India, they are rapidly becoming essential tools for organizations seeking to strengthen resilience against increasingly sophisticated cyberattacks.

Table 2: AI and M L use in Cybersecurity (Worldwide vs India).

Region	Organizations Using AI/ML in Cybersecurity (%)	AI for Threat Detection (%)	AI for Incident Response (%)
Worldwide	74	57	49
India	94	72	65

Source: Statista; Fortinet–IDC Survey; DSCI Reports; Economic Times (2023–2025).

7. Global adoption of AI & ML in cybersecurity in 2025

Artificial Intelligence (AI) and Machine Learning (ML) have introduced key approaches in modern cybersecurity, emphasizing automation, predictive threat detection, response orchestration, and fraud prevention.

Key Global Statistics: Approximately 55% of organizations worldwide are developing strategies to integrate AI into cybersecurity, reflecting rapid advancement beyond traditional security tools. AI and ML technologies play a significant role in system security and endpoint protection, accounting for over 47% of the market share in security solutions.[58] Furthermore, nearly 88% of cybersecurity vendors have incorporated AI capabilities into their offerings, indicating widespread industry adoption. The prevalence of AI-driven fraud is increasing by approximately 52% annually, particularly targeting the financial services and e-commerce sectors. Additionally, AI-based cybersecurity solutions represent around 72% of new market entrants, highlighting a strong strategic shift toward intelligent security systems.[49]

Regional Adoption Trends: North America leads in AI adoption for cybersecurity, with an approximate adoption rate of 58%. Europe and the Asia–Pacific region follow, with adoption rates of around 42% and 33%, respectively, driven by regulatory initiatives and ongoing digital transformation.[59]

Cybersecurity Domains: The application of AI and ML in cybersecurity can be broadly categorized into four core domains:

1. Threat Detection and Prevention: AI and ML enhance anomaly detection, zero-day threat identification, and response automation, significantly reducing attacker dwell time.
2. Network and Endpoint Security: Machine learning models analyze network traffic, identify patterns, and mitigate attacks in real time, which is critical for distributed enterprise environments.[39]
3. Fraud Detection and Identity Protection: Financial institutions increasingly rely on AI to reduce fraud in digital transactions, with adoption steadily rising.
4. Security Automation and Orchestration: AI enables automated Security Operations Center (SOC) workflows, reducing reliance on manual threat hunting and accelerating response actions.

7.1 Leading countries

The global landscape of AI in cybersecurity from 2021–2025 is evaluated based on overall interest, research productivity, and public discourse, which serve as indicators of technological engagement and strategic focus. This ranking reflects levels of participation and development rather than actual market revenue.[60] The following section presents an overview of the top ten countries and regions where the adoption and application of AI and ML in cybersecurity are most significant.

This analysis highlights global interest in and adoption of AI for cybersecurity, with the United States leading due to its strong technological capabilities and high research output. The United Kingdom and India demonstrate steady engagement and rapid growth in this domain. Countries such as Indonesia and Nigeria reflect increasing awareness and emerging development in cybersecurity markets, indicating a broader shift in global cybersecurity priorities. Meanwhile, Italy, Australia, and Japan have established robust strategies for integrating AI into national security frameworks.[60]

The integration of AI and ML in cybersecurity continues to expand across multiple sectors, including AI innovation ecosystems, technology hubs, cybersecurity infrastructure, startup ecosystems, IoT security, cyber defense, and national governance. Among the top ten countries, three are from Asia and three from Europe, reflecting a balanced global distribution of advancements in AI-driven cybersecurity.

Overall, the data emphasize the critical role of AI in addressing and mitigating evolving digital threats.

7.2 Adoption and necessity of AI in cybersecurity

The development of AI in cybersecurity is not merely a trend; it has become a critical necessity for organizations to stay ahead of emerging threats. The reported statistics on technology adoption rates for AI in cybersecurity highlight both its widespread acceptance and the growing demand for AI-enabled solutions.

Approximately 93% of organizations recognize the impact of AI on threat management, while 69% consider AI essential for incident response. Around 59% support the use of AI for risk assessment, and 64.3% utilize AI to enhance overall cybersecurity capabilities. Additionally, 51% employ AI for threat detection, 62% are actively exploring AI-based cybersecurity solutions, and 51% use AI to address cyberattacks. Further details are presented in Fig. 7.

Table 3: Ranking of the top 10 countries for AI/ML in terms of Cybersecurity adoption.[60]

Rank	Country	AI/ML adoption in Cybersecurity (%)	Key focus area
1	United States	59.33	Focus on cybersecurity advancements
2	United Kingdom	11.94	Growing AI innovation sector
3	India	11.57	Leveraging AI for cybersecurity, emphasizing tech hub
4	Indonesia	4.01	AI's role in enhancing cybersecurity infrastructure
5	Canada	3.45	Active research community and AI startups
6	Germany	2.71	AI in cybersecurity, underpinned by its industrial innovation
7	Nigeria	2.05	AI solutions to address cybersecurity challenges in Africa
8	Italy	1.87	AI technologies in the context of national security
9	Australia	1.68	leveraging AI for national and cyber defense
10	Japan	1.40	AI advancements in cybersecurity and robotics integration

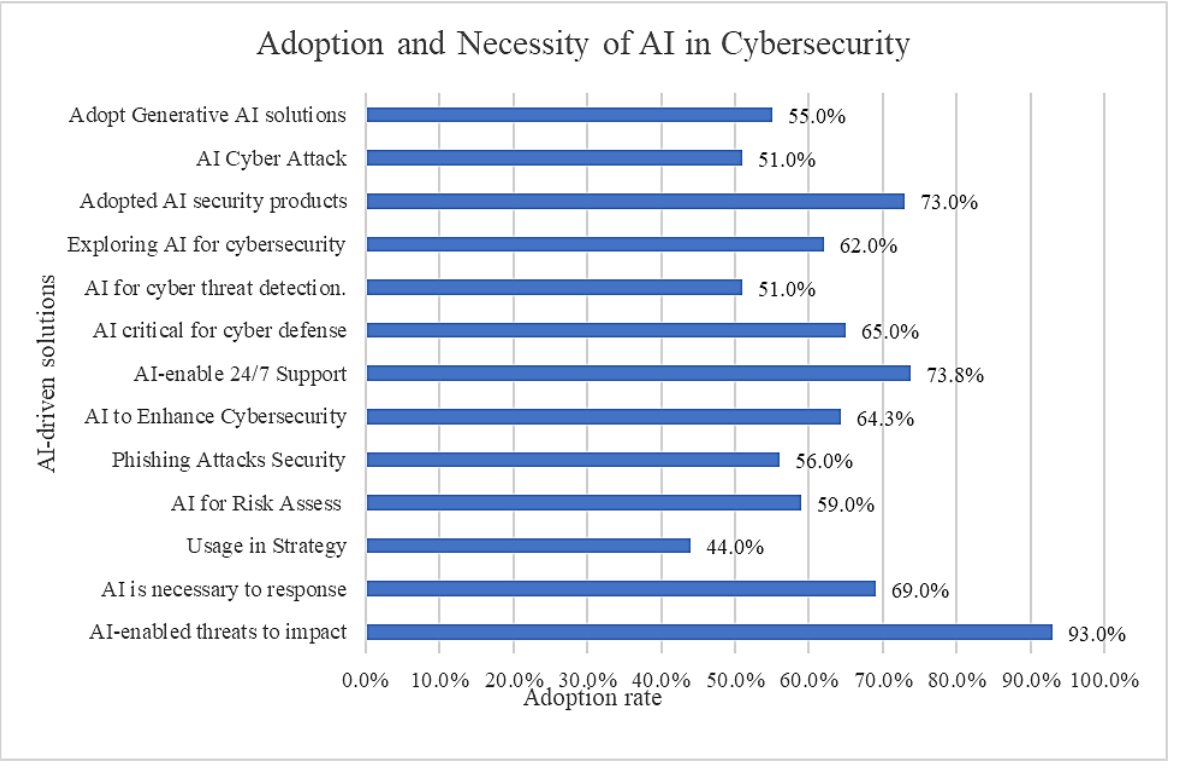

Fig. 7: Adoption and necessity of AI in cybersecurity.

7.3 Performance improvement

Organizations that adopt AI in their cybersecurity strategies report significant performance improvements. The integration of AI and ML enables automated responses, faster threat detection, and a reduction in false positives. By improving operational efficiency and strengthening defensive capabilities, AI enhances security through the automation of routine tasks. A visual analysis of key application areas demonstrates the effectiveness of AI in cybersecurity: triage of Tier-1 threats (67%), detection of zero-day attacks (66%), prediction of future threats (65%), reduction of false positives and noise (65%), and correlation of user behavior with threat indicators (61%).

8. Future prospects

The future prospects of Artificial Intelligence (AI) and Machine Learning (ML) in cybersecurity are illustrated in the projected diagram, which depicts the anticipated growth of the global AI cybersecurity market from 2026 to 2035, highlighting a strong upward trend. The market value is

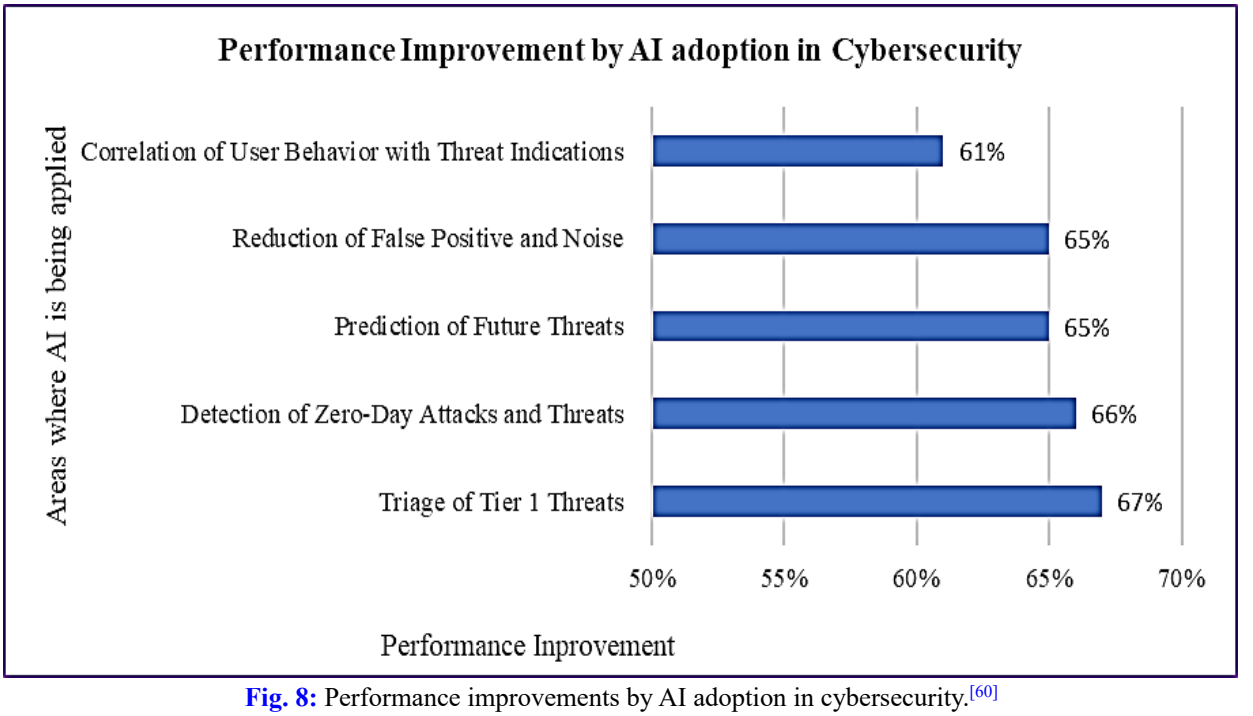

Fig. 8: Performance improvements by AI adoption in cybersecurity.[60]

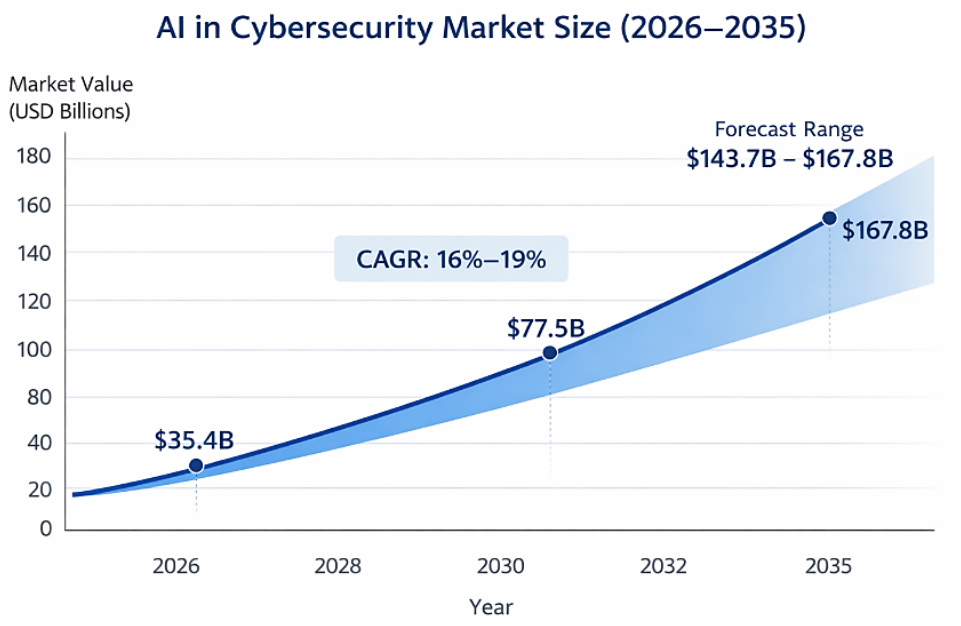

Fig. 9: AI in Cybersecurity (2026–2035).[61,62]

expected to increase from approximately USD 35.4 billion in 2026 to around USD 77.5 billion by 2030, reflecting the rapid adoption of AI-enabled security solutions.[61]

By 2035, the market is projected to reach approximately USD 167.8 billion, with estimates ranging between USD 143.7 billion and USD 167.8 billion. The chart indicates a compound annual growth rate (CAGR) of 16%–19%, driven by the increasing frequency and sophistication of cyber threats, the growing need for automation, and the demand for intelligent threat detection systems.[62]

9. Emerging trends in cybersecurity

The cybersecurity landscape is undergoing rapid transformation as digital infrastructures expand and cyber threats become increasingly sophisticated, automated, and persistent. Traditional security approaches, which primarily rely on predefined rules and signature-based detection, are no longer sufficient to counter advanced attacks such as zero-day exploits, ransomware, and advanced persistent threats (APTs). The adoption of Artificial Intelligence (AI) and Machine Learning (ML) is emerging as a key approach that is redefining how organizations detect, respond to, prevent, and recover from cyber incidents.

Future trends in AI-enabled cybersecurity emphasize not only improvements in detection accuracy but also enhancements in transparency, resilience, scalability, and collaboration. Key developments include Explainable Artificial Intelligence (XAI), federated learning, integration with quantum computing, AI-driven cyber resilience, convergence with IoT security, and sector-specific applications in domains such as banking and healthcare.[7,26]

9.1. Explainable AI (XAI) in cybersecurity

AI-driven automation has become a critical necessity in modern cybersecurity operations. Many Machine Learning (ML) models, particularly deep neural networks, often function as "black boxes," making it difficult for security analysts to understand how specific decisions are made. In cybersecurity contexts-where actions may include blocking users, isolating systems, or escalating alerts-a lack of transparency can lead to uncertainty and operational hesitation.

Explainable Artificial Intelligence (XAI) addresses this challenge by designing models that provide human-interpretable explanations for their outputs. An XAI-enabled system can clearly justify why and how a particular alert was triggered.

The advantages of XAI in cybersecurity include:

- Reduction of false positives and false negatives
- Increased analyst confidence in automated systems
- Faster validation of incident response actions
- Improved auditability and compliance

XAI enables organizations to meet regulatory requirements by ensuring transparency and traceability in decision-making processes. Future cybersecurity architectures and operational models are expected to emphasize not only accuracy but also interpretability, ethical alignment, and transparency.

9.2. Federated learning, privacy preservation and collaborative cyber defenses

Cyber threats are not confined by geographical boundaries; therefore, sharing sensitive data often introduces significant privacy and confidentiality risks. From a security perspective, federated learning (FL) offers an effective solution by enabling decentralized model training without requiring the exchange of raw data.[26]

In a federated learning framework:

1. Each entity trains a local model using its own data.
2. Model parameters (rather than raw data) are shared with a central aggregator.

3. The aggregated global model benefits from collective intelligence while preserving data privacy.

Such decentralized training approaches enable collaborative threat intelligence without exposing proprietary or sensitive data.

Key Applications in Cybersecurity include:

1. Privacy-Preserving Malware Detection: Organizations can collaboratively develop malware detection models while keeping critical datasets local.
2. Collaborative Intrusion Detection Systems (IDS): Financial institutions or security agencies can jointly train anomaly detection models to identify emerging attack patterns.
3. Zero-Day Threat Intelligence Sharing: FL facilitates the detection of emerging threats across organizations while ensuring compliance with regulations such as GDPR and CCPA.

Recent research integrating FL with software-defined networking (SDN) and deep learning frameworks demonstrates enhanced proactive threat detection with minimal computational overhead. As a result, federated learning is expected to play a foundational role in the development of global cyber defense systems

9.3. Integration of AI with quantum computing

The integration of Artificial Intelligence (AI), Machine Learning (ML), and quantum computing represents a transformative shift in cybersecurity capabilities. Quantum computers operate using quantum bits (qubits) and leverage principles such as superposition and entanglement, enabling them to perform complex computations exponentially faster than classical systems for specific tasks.

Conventional encryption schemes such as RSA and Elliptic Curve Cryptography (ECC) rely on the computational difficulty of factoring large numbers and solving discrete logarithmic problems. However, quantum algorithms, particularly Shor's Algorithm, pose a significant threat to these cryptographic frameworks by enabling efficient decryption. This raises critical concerns regarding the need for post-quantum cryptography (PQC). Researchers are actively developing quantum-resistant encryption techniques, including (1) lattice-based cryptography and (2) code-based cryptography. AI can further support the design, testing, and validation of these advanced cryptographic methods.

Quantum computing also enhances AI-driven cybersecurity systems by:

1. Improving large-scale log analysis
2. Detecting anomalies in high-dimensional datasets
3. Enabling real-time pattern recognition across distributed systems
4. Accelerating the training of deep learning models

Quantum-enhanced Security Information and Event Management (SIEM) systems have the potential to analyze complex, multidimensional ransomware patterns in real time. Additionally, quantum-inspired optimization algorithms, such as the Gravitational Search Algorithm (GSA), have been proposed for efficient resource optimization in edge computing environments.

9.4. AI-Driven cyber resilience

Traditional cybersecurity primarily focuses on prevention and mitigation. However, given the inevitability of security breaches, future systems must emphasize cyber resilience-the ability to withstand, recover from, and adapt to cyber incidents.

AI-enabled cyber resilience includes:

1. Self-healing systems
2. Autonomous vulnerability patching
3. Intelligent threat isolation
4. Dynamic security reconfiguration

For example, once a breach is detected, an AI-enabled system can:

1. Identify the compromised component
2. Automatically isolate the affected system
3. Apply necessary security patches
4. Restore services in minimal time

Such autonomous responses reduce the need for human intervention and limit potential damage.

In the context of proactive and predictive defense, AI systems continuously analyze behavioral patterns, threat intelligence feeds, and historical attack data. By anticipating potential vulnerabilities, AI can proactively optimize:

1. Firewall rules
2. Authentication mechanisms
3. Access control policies
4. Network segmentation strategies

This transition from reactive to predictive and adaptive cybersecurity represents the foundation of next-generation digital defense frameworks.

9.5. Convergence of AI and IoT security

The rapid proliferation of Internet of Things (IoT) devices has significantly expanded the attack surface. IoT systems often consist of resource-constrained devices that cannot support traditional security solutions, making them vulnerable to botnet attacks, distributed denial-of-service (DDoS) attacks, and unauthorized access. AI-driven approaches provide scalable and intelligent solutions tailored to IoT environments.

The key capabilities of AI in IoT security include:

1. Behavioral anomaly detection across connected devices
2. Real-time monitoring of device communication patterns
3. Automated identification of compromised nodes
4. Lightweight intrusion detection models

AI techniques continuously analyze IoT traffic patterns, enabling the detection of deviations that may indicate malicious activity.

When integrated with edge computing, AI-enabled IoT security offers several advantages:

1. Faster response times
2. Reduced network congestion
3. Improved scalability
4. Enhanced overall threat mitigation

As IoT devices play a critical role in smart cities, healthcare systems, critical infrastructure, and industrial control systems, AI-driven IoT security has become essential for ensuring robust and resilient protection.

9.6. Sector-specific applications: healthcare and banking

Several critical sectors require exceptionally high levels of cybersecurity due to the sensitive nature of their data. These sectors include healthcare, banking, finance, and government. In the healthcare domain, systems manage electronic health records (EHRs), medical imaging data, and real-time patient monitoring systems, all of which demand robust protection.

AI-enabled cybersecurity enhances healthcare security by:

1. Detecting anomalies in patient data access
2. Securing medical IoT devices
3. Identifying ransomware attacks targeting healthcare institutions
4. Supporting trust and reputation management

AI-based trust evaluation models assess system reliability, safeguard data integrity, and ensure compliance with healthcare regulations.

Similarly, the banking and financial services sector relies heavily on AI for fraud detection and transaction monitoring. AI-driven cybersecurity systems support:

1. Analysis of transaction behavior patterns
2. Detection of unauthorized access attempts
3. Prevention of identity theft
4. Strengthening of authentication mechanisms

Real-time anomaly detection ensures secure digital transactions while maintaining regulatory compliance. With the integration of federated learning and quantum-resistant cryptography, both banking and healthcare sectors can enhance security while preserving privacy and trust.

9.7 Integration into security operations centers (SOCs)

AI-enabled Security Operations Centers (SOCs) play a crucial role in detecting and responding to cyberattacks while strengthening overall security operations. In practice, SOCs integrate AI-driven capabilities to enhance threat detection and response efficiency. AI systems can automate routine tasks such as log analysis and threat prioritization, allowing human analysts to focus on more complex incidents.

From an implementation perspective, modern SOCs combine AI-based Security Information and Event Management (SIEM) systems with correlated alerts from multiple sources. These systems reduce alert fatigue by grouping related events and providing actionable insights.

In a practical case study, a medium-sized organization implemented an AI-enabled SIEM system that reduced false positives by approximately 40% and shortened incident response time by 50%. The system's ability to prioritize high-risk threats significantly improved the overall efficiency and effectiveness of the SOC.

9.8 Autonomous defense

To effectively identify and respond to cyber threats, next-generation AI models are expected to play a critical role in enhancing cyber resilience-the ability to withstand, recover from, and prevent attacks. This paradigm supports the concept of autonomous cyber defense. AI-enabled cybersecurity relies on self-adaptive systems capable of automatically patching vulnerabilities, restoring compromised systems, and strengthening security configurations based on real-time data and threat intelligence. For example, an AI-driven system under attack can detect a breach, isolate affected components, and automatically deploy security patches to eliminate vulnerabilities. These processes are executed with minimal or no human intervention. Such self-healing and adaptive capabilities reduce operational disruption and minimize the impact of cyber incidents, ensuring continuous protection even in dynamic threat environments. In self-healing systems, AI plays a vital role in proactive security frameworks, where advanced models anticipate potential threats and implement appropriate defensive measures. By continuously analyzing threat trends and behavioral patterns, AI systems can predict the likelihood of specific attacks and take preventive actions, such as strengthening authentication mechanisms or optimizing firewall rules. This transition from reactive to proactive and adaptive defense strategies represents a key feature of future cybersecurity models driven by AI and ML.

9.9. Graph neural networks (GNNs) in cybersecurity

Graph Neural Networks (GNNs) have emerged as a powerful AI model in cybersecurity due to their ability to represent and analyze complex relationships among entities such as devices, users, IP addresses, files, and processes. Unlike traditional machine learning methods that treat data as independent samples, GNNs leverage graph structures to capture interdependencies, making them highly effective for detecting sophisticated, multi-stage cyberattacks.

In intrusion detection systems (IDS), GNNs analyze network traffic graphs to identify anomalous communication patterns. For malware detection, they model relationships between system calls and program behaviors, thereby improving classification accuracy. GNNs can also be applied to fraud detection, phishing detection, and advanced persistent threat (APT) identification by uncovering hidden attack paths within large-scale enterprise networks.

Recent studies (2023–2025) highlight the integration of GNNs with threat intelligence graphs and zero-trust architectures to enhance real-time detection and threat response. However, challenges remain in terms of graph construction, adversarial robustness, and model

interpretability.[63]

10. Challenges and limitations

Artificial Intelligence (AI) and Machine Learning (ML) significantly enhance cybersecurity through improved threat detection and response capabilities. However, their implementation presents several challenges, including poor data quality, vulnerability to adversarial attacks, scalability constraints, model inaccuracies, and ethical concerns. Addressing these limitations is essential to ensure the development of trustworthy, transparent, and effective AI- and ML-enabled cybersecurity systems.

10.1 Adversarial attacks

With the increasing adoption of AI and ML in cybersecurity, adversaries are progressively targeting vulnerabilities in these intelligent systems. Adversarial machine learning involves the creation of carefully crafted inputs designed to deceive AI models, posing a significant threat to the reliability of AI-driven defenses. Paradoxically, systems developed to enhance security can themselves become targets of attack.

In adversarial attacks, subtle modifications are introduced into data to trigger incorrect classifications or predictions. For instance, attackers may slightly modify malware code or manipulate network traffic patterns so that malicious activity appears legitimate to an AI system. These minor perturbations are often imperceptible to human observers but sufficient to mislead machine learning models. Similar vulnerabilities have been observed in domains such as image recognition, natural language processing (NLP), and cybersecurity applications.

Although techniques such as adversarial training can improve model robustness, no AI/ML system is completely secure. Continuous updates, monitoring, and the implementation of robust defense strategies are essential to mitigate the evolving risks posed by adversarial threats.

10.2 Data availability and quality

A fundamental challenge in deploying AI and ML in cybersecurity is the need for sufficient high-quality data to train reliable models. Machine learning systems require large datasets to identify patterns, distinguish between normal and malicious behavior, and generate accurate predictions. For tasks such as anomaly detection and malware classification, datasets must include diverse examples of both legitimate activities and cyberattacks. However, obtaining large, well-labeled datasets remains a significant challenge.

Organizations are often reluctant to share cybersecurity data due to privacy concerns, confidentiality agreements, and competitive risks. As a result, publicly available datasets are limited and may not accurately reflect real-world complexity. Even when data are accessible, they are often imbalanced, with relatively few instances of rare but critical threats such as advanced persistent threats (APTs) or zero-day attacks.

Furthermore, cybersecurity data are typically unstructured, incomplete, and noisy, requiring extensive preprocessing. Poor data quality ultimately reduces model accuracy and leads to gaps in threat detection.

10.3 Regulatory compliance and legal constraints

The adoption of AI in cybersecurity is significantly influenced by stringent regulatory frameworks that vary across jurisdictions. Data protection regulations such as the General Data Protection Regulation (GDPR) in the European Union, the California Consumer Privacy Act (CCPA), and cybersecurity standards from the National Institute of Standards and Technology (NIST) and ISO/IEC 27001 impose strict requirements on data collection, storage, and processing.

AI models used in cybersecurity must comply with these regulations by ensuring:

1. Privacy preservation: AI-enabled threat intelligence systems must anonymize sensitive data to prevent unauthorized exposure.
2. Bias mitigation: Compliance with fairness and transparency principles ensures that AI systems do not introduce bias in security decision-making.
3. Auditability and explainability: Regulatory bodies require AI-driven security tools to provide interpretable and justifiable decisions, enabling verification of automated threat detection processes.

Despite these efforts, current AI-enabled cybersecurity solutions often struggle to maintain regulatory compliance, as policies evolve rapidly to keep pace with emerging AI-related risks. Organizations also face challenges in aligning AI-driven security practices with existing and evolving legal frameworks, which can hinder large-scale adoption.

10.4 Scalability and real-time processing

AI-enabled cybersecurity solutions must scale efficiently to handle the increasing variety, volume, and velocity of data generated by modern networks. Large enterprises and cloud environments produce millions of security events per second, requiring AI/ML models to analyze data in real time for rapid threat detection and response. However, scalability remains a major challenge, particularly for traditional algorithms and deep neural networks that demand substantial computational resources.

Real-time processing often depends on high-performance infrastructure, which may be costly and difficult for smaller organizations to implement. Even with cloud and distributed computing, latency can affect analysis and mitigation processes. Furthermore, the evolving threat landscape necessitates adaptive AI/ML systems capable of learning continuously from streaming data without frequent manual retraining.

Scalability also involves coordinating multiple AI systems across cloud platforms, networks, IoT devices, and endpoints. Ensuring seamless integration while maintaining

real-time performance presents a significant technical challenge in modern cybersecurity.

10.5 False positives and false negatives

A key limitation of AI/ML systems in cybersecurity is their propensity to generate false positives (incorrectly flagging benign events as malicious) and false negatives (failing to detect actual threats). Although AI enhances detection capabilities, its accuracy is not perfect, and such errors can significantly impact security operations.

False positives create substantial operational challenges. Excessive incorrect alerts contribute to *alert fatigue*, where security analysts become overwhelmed by notifications from multiple systems. When AI-enabled tools generate a high volume of false alarms, analysts may begin to ignore alerts, potentially overlooking genuine threats. This reduces trust in AI/ML systems and undermines their overall effectiveness.

Conversely, false negatives pose even greater risks. If an AI/ML system fails to detect malware, phishing attacks, or advanced persistent threats, the incident may escalate without intervention, causing significant damage to data, infrastructure, and organizational reputation. In high-risk environments, even a single undetected attack can lead to severe consequences.

Addressing these challenges requires continuous model tuning and regular updates. However, balancing detection sensitivity remains complex, as reducing false negatives often increases false positives. This trade-off necessitates adaptive, real-time learning systems capable of maintaining an optimal balance between accuracy and reliability.

10.6 Ethics and privacy

The deployment of AI in cybersecurity raises significant ethical concerns related to transparency, privacy, and algorithmic bias. AI/ML systems designed for threat detection typically require extensive access to both individual and organizational data, raising important questions about how such data are collected, stored, and utilized.

One of the primary ethical challenges is algorithmic bias. Since AI models learn from historical data, any inherent biases-whether intentional or unintentional-can influence outcomes. In cybersecurity, this may lead to biased profiling, inconsistent monitoring, or disproportionate targeting of specific individuals, groups, or regions. Ensuring fairness and accountability is therefore essential.

Privacy represents another critical concern. Continuous monitoring systems often analyze sensitive data, including user behavior, communication records, and access logs. While such data are necessary for effective threat detection, misuse or inadequate protection can compromise individual privacy rights.

Transparency is equally important. Many advanced AI models operate as "black boxes," limiting their interpretability. The growing emphasis on Explainable Artificial Intelligence (XAI) seeks to address this limitation by enabling more transparent and trustworthy decision-making in cybersecurity operations.

10.7 Real-world execution challenges

Although Artificial Intelligence (AI) and Machine Learning (ML) have demonstrated significant potential in cybersecurity, their real-world deployment faces several challenges, particularly in terms of regulatory compliance and organizational adoption. Addressing these challenges is essential to ensure that AI-enabled security frameworks can be effectively implemented across enterprise environments, government sectors, and critical infrastructures.

10.8 Industry acceptance blockades

Despite achieving high accuracy in detecting cyber threats, the adoption of AI in enterprises remains constrained by integration complexities, trust issues, and cost considerations. The key barriers include:

Legacy System Compatibility: Many organizations still rely on traditional security infrastructures, making it difficult to integrate AI-powered intrusion detection, automated threat analysis, and behavioral analytics into existing Security Information and Event Management (SIEM) systems.

Scalability and Performance: AI-enabled cybersecurity solutions require substantial computational resources and high-quality data to function effectively in real-world environments. Small and medium-sized enterprises (SMEs) often lack the infrastructure needed to deploy such systems at scale.

Trust and Reliability: Organizations may hesitate to adopt AI-driven security solutions due to concerns about false positives, adversarial attacks, and limited transparency in AI-based decision-making. Explainability and human-in-the-loop approaches are critical to bridging this trust gap.

10.9 Workforce and skill gap

The successful deployment of AI in cybersecurity requires highly skilled professionals capable of understanding AI-driven insights, optimizing models, and managing complex threat landscapes. However, the industry faces a significant skill gap, as many cybersecurity professionals lack expertise in AI/ML-based detection and response techniques. Bridging this gap requires structured upskilling programs, specialized AI security certifications, and the integration of AI-focused modules into professional training curricula.

10.9.1 Ethical and adversarial risks

AI systems operating in real-world environments remain vulnerable to adversarial attacks, including model poisoning, data manipulation, and adversarial perturbations. These threats undermine trust in AI reliability. At the same time, ethical safeguards must ensure that automated decisions respect privacy and avoid discriminatory outcomes. Future research should focus on regulatory alignment,

interoperability with legacy systems, improved explainability, and robust defenses against adversarial threats.

10.9.2 False positives and false negatives

AI-enabled systems often face trade-offs between accuracy and reliability. False positives can lead to alert fatigue, while false negatives allow sophisticated attacks to evade detection. Hybrid approaches that combine rule-based methods with AI/ML techniques aim to achieve a better balance; however, optimization remains a complex challenge.

10.9.3 Computational overhead and resource requirements

AI-driven cybersecurity systems require substantial computational resources, continuous model training, and scalable infrastructure. Small and medium-sized enterprises (SMEs) may struggle to meet these requirements. Ongoing research in lightweight deep learning and edge-based AI aims to improve efficiency and scalability.

10.9.4 Explainability and trust issues in ai-based security decisions

Many AI systems operate as "black-box" models, limiting transparency and interpretability. Regulatory frameworks such as GDPR and standards from NIST emphasize the need for explainable and auditable decision-making processes. Enhancing explainability, robustness, and ethical governance is essential for building trustworthy and accountable AI-driven cybersecurity systems.

11. Conclusion

Comparative analyses of global and country-level data clearly demonstrate that Artificial Intelligence (AI) and Machine Learning (ML) have become essential enablers of modern cybersecurity. Global adoption statistics indicate that AI-enabled solutions are no longer experimental but are now operationally embedded across critical security functions. High adoption rates in areas such as AI-based threat detection, network traffic analysis, fraud detection, and defense mechanisms reflect growing confidence in the ability of intelligent systems to manage complex, high-volume cyber threats that exceed human analytical capacity. Country-level analyses further reveal significant disparities and strategic priorities. India has emerged as a global leader in the adoption of AI and ML for cybersecurity, surpassing several developed nations, driven by rapid digital transformation, increasing exposure to cyber threats, and proactive industrial and governmental initiatives. Developed countries such as the United States and the United Kingdom continue to prioritize advanced security automation and innovation, while other nations are progressively expanding adoption to secure critical infrastructure, financial systems, and cloud environments. Future market projections for the period 2026–2035 indicate sustained and robust growth in the AI cybersecurity market, which is expected to exceed USD 140 billion, reflecting a strong compound annual growth rate (CAGR). This trend highlights a global shift toward predictive, automated, and adaptive security systems. Overall, the data confirm that AI and ML are transforming cybersecurity from reactive defense mechanisms into proactive, scalable, and intelligent frameworks capable of addressing evolving digital threats.

CRediT Author Contribution Statement

Pooja Soni: Conceptualization, Literature review, Data curation, Writing - original draft. **Sanjay Gour:** Methodology, Supervision, Writing - review and editing, Validation. All authors have read and agreed to the published version of the manuscript.

Funding Declaration

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study. All information discussed in this review is based on previously published studies cited in the reference list.

Conflict of Interest

There is no conflict of interest.

Artificial Intelligence (AI) Use Disclosure

The authors confirm that no artificial intelligence (AI)-assisted technologies were used in the writing of the manuscript, and no images were generated or manipulated using AI. AI-based tools were used solely for language editing to improve grammar, clarity, and readability, in accordance with journal policy. The authors take full responsibility for the accuracy, originality, and integrity of the work.

Supporting Information

Not applicable

References

[1] S. Samtani, H. Chen, M. Kantarcioglu, B. Thuraisingham, Explainable artificial intelligence for cyber threat intelligence (XAI-CTI), *IEEE Transactions on Dependable and Secure Computing*, 2002, **19**, 2149–2150, doi: 10.1109/TDSC.2022.3168187.

[2] A. Awadallah, K. Eledlebi, J. Zemerly, D. Puthal, E. Damiani, K. Taha, T. Y. Kim, P. D. Yoo, K. K. Choo, M. S. Yim, C. Y. Yeun, Artificial intelligence-based cybersecurity for the metaverse: research challenges and opportunities, *IEEE Communications Surveys & Tutorials*, 2024, **27**, 1008 – 1052, doi:

10.1109/COMST.2024.3442475.

[3] S. Duary, P. Choudhury, S. Mishra, V. Sharma, D. D. Rao, A. P. Aderemi, Cybersecurity threats detection in intelligent networks using predictive analytics approaches, In: 2024 4th International Conference on Innovative Practices in Technology and Management (ICIPTM), IEEE, 2024, 01–05.

[4] D. D. Pandya, S. Gaur, Detection of anomalous value in data mining, Kalpa Publications in Engineering, Proceedings on International Conference on Emerging Trends in Expert Applications & Security, 2018, **2**, 01–06.

[5] S. Lysenko, N. Bobro, K. Korsunova, O. Vasylchyshyn, Y. Tatarchenko, The role of artificial intelligence in cybersecurity: automation of protection and detection of threats, *Economic Affairs*, 2024, **69**, 43–51, doi: 10.46852/0424-2513.1.2024.6.

[6] M. M. Nair, A. Deshmukh, A. K. Tyagi, Artificial intelligence for cyber security: current trends and future challenges, In: Automated secure computing for next-generation systems, Wiley, Hoboken, 2024, 83–114.

[7] N. Mohamed, Artificial intelligence and machine learning in cybersecurity: A deep dive into state-of-the-art techniques and future paradigms, *Knowledge and Information Systems*, 2025, **67**, 6969–7055, doi: 10.1007/s10115-025-02429-y.

[8] M. Malatji, A. Tolah, Artificial intelligence (AI) cybersecurity dimensions: a comprehensive framework for understanding adversarial and offensive AI, AI Ethics, 2024, doi: 10.1007/s43681-024-00427-4.

[9] M. F. A. Sheikh, S. Gaur, H. Desai, S. K. Sharma, A study of performance evaluation of cryptographic algorithm, In: Emerging Trends in Expert Applications and Security (ICETEAS 2018), Advances in Intelligent Systems and Computing Series, 2018, **841**, 379–384.

[10] T. Choithani, A. Chowdhury, S. Patel, P. Patel, D. Patel, M. Shah, A comprehensive study of artificial intelligence and cybersecurity on bitcoin, cryptocurrency and banking system, *Annals of Data Science*, 2024, **11**, 103–135, 10.1007/s40745-022-00433-5.

[11] N. Sharma, N. Jindal, Emerging artificial intelligence applications: metaverse, IoT, cybersecurity, healthcare-an overview, *Multimedia Tools and Applications*, 2024, **83**, 57317–57345, doi: 10.1007/s11042-023-17890-6.

[12] D. K. Sharma, J. Mishra, A. Singh, R. Govil, G. Srivastava, J. C. W. Lin, Explainable artificial intelligence for cybersecurity, *Computers & Electrical Engineering*, 2022, **103**, 108356, doi: 10.1016/j.compeleceng.2022.108356.

[13] R. Das, R. Sandhane, Artificial intelligence in cybersecurity, *Journal of Physics: Conference Series*, 2021, **1964**, 042072, doi: 10.1088/1742-6596/1964/4/042072.

[14] G. Rjoub, J. Bentahar, O. A. Wahab, R. Mizouni, A. Song, R. Cohen, H. Otrok, A. Mourad, A survey on explainable artificial intelligence for cybersecurity, *IEEE Transactions on Network and Service Management*, 2023, **20**, 5115–5140, doi: 10.1109/TNSM.2023.3282740.

[15] K. Michael, R. Abbas, G. Roussos, AI in cybersecurity: the paradox, *IEEE Transactions on Technology and Society*, 2023, **4**, 104–109, doi: 10.1109/TTS.2023.3280109.

[16] Symantec, Internet Security Threat Report (ISTR-19), 2019, 24.

[17] S. Axelsson, The base-rate fallacy and the difficulty of intrusion detection, *ACM Transactions on Information and System Security*, 2000, **3**, 186–205, doi: 10.1145/357830.357849.

[18] R. Sommer, V. Paxson, Outside the closed world: on using machine learning for network intrusion detection, In: 31st IEEE Symposium on Security and Privacy, 2010, 305–316, doi: 10.1109/SP.2010.25.

[19] ENISA, Threat landscape report, European Union Agency for Cybersecurity, 2022, doi: 10.2824/764318.

[20] A. L. Buczak, E. Guven, A survey of data mining and machine learning methods for cybersecurity intrusion detection, *IEEE Communications Surveys & Tutorials*, 2016, **18**, 1153–1176, doi: 10.1109/COMST.2015.2494502.

[21] ENISA, Artificial intelligence and cybersecurity: challenges and opportunities, 2020, doi: 10.2824/808362.

[22] M. Chouhan, S. Gour, A hybrid machine learning framework for proactive threat detection in cybersecurity using intelligent feature, Smart Innovation, Systems and Technologies, Springer Singapore, 2026, 124, 629–640, doi: 10.1007/978-981-95-1353-6_49.

[23] V. Chandola, A. Banerjee, V. Kumar, Anomaly detection: a survey, *ACM Computing Surveys*, 2009, **41**, 1–72, doi: 10.1145/1541880.1541882.

[24] G. Kim, S. Lee, S. Kim, A novel hybrid intrusion detection method integrating anomaly detection with misuse detection, *Expert Systems with Applications*, 2014, **41**, 1690–1700.

[25] S. Gour, V. Khanna, Assessment of efficiency of machine learning algorithms in loan-default prediction, *International Journal of Engineering and Designing Innovation*, 2025, **7**, 01–06, doi: 10.2019/IJEDI/202511001.

[26] S. Gour, P. Soni, Loan default prediction using ensemble machine learning algorithms, *Journal of Smart Sensor and Computing*, 2025, **1**, 25215, doi: 10.64189/ssc.25215.

[27] D. Dave, G. Sawhney, P. Aggarwal, N. Silswal, D. Khut, The new frontier of cybersecurity: emerging threats and innovations, arXiv preprint, 2023, doi: 10.48550/arXiv.2311.02630.

[28] S. T. Erukude, V. C. Marella, S. R. Veluru, AI-driven cybersecurity threats: a survey of emerging risks and

defensive strategies, arXiv preprint, 2026, arXiv:2601.03304.

[29] V. K. Kushi, S. R. Gadag, M. U. Pasha, Sushmitha E., M. Kotari, Emerging threats and innovative solutions in cybersecurity: a comprehensive review, *International Journal of Advances in Computer Science and Technology*, 2024, **13**, 47–54, doi: 10.30534/ijacst/2024/091312024.

[30] A. I. Jony, S. A. Hamim, Navigating the cyber threat landscape: a comprehensive analysis of attacks and security in the digital age, *Journal of Information Technology and Cyber Security*, 2023, **1**, 53–67, doi: 10.30996/jitcs.9715.

[31] S. Razaulla, C. Fachkha, C. Markarian, A. Gawanmeh, W. Mansoor, B. C. M. Fung, C. Assi, The age of ransomware: a survey on the evolution, taxonomy, and research directions, *IEEE Access*, 2023, **11**, 40698–40723, doi: 10.1109/ACCESS.2023.3268535.

[32] A. K. Jain, B. B. Gupta, A survey of phishing attack techniques, defence mechanisms and open research challenges, *Enterprise Information Systems*, 2022, **16**, 527–565, doi: 10.1080/17517575.2021.1896786.

[33] R. Zieni, L. Massari, M. C. Calzarossa, Phishing or not phishing? a survey on the detection of phishing websites, *IEEE Access*, 2023, **11**, 18499–18519, doi: 10.1109/ACCESS.2023.3247135.

[34] M. Gniewkowski, An overview of DoS and DDoS attack detection techniques, In: Zamojski, W., Mazurkiewicz, J., Sugier, J., Walkowiak, T., Kacprzyk, J. (eds) Theory and Applications of Dependable Computer Systems. DepCoS-RELCOMEX 2020. Advances in Intelligent Systems and Computing, vol 1173. Springer, Cham. doi: 10.1007/978-3-030-48256-5_23.

[35] A. Blaise, M. Bouet, V. Conan, S. Secci, Detection of zero-day attacks: an unsupervised port-based approach, *Computer Networks*, 2020, **180**, 107391, doi: 10.1016/j.comnet.2020.107391.

[36] G. N. Nayak, S. G. Samaddar, Different flavours of man-in-the-middle attack, consequences and feasible solutions, In: 2010 3rd International Conference on Computer Science and Information Technology, 2010, 491–495, doi: 10.1109/iccsit.2010.5563900.

[37] M. Nasereddin, A. Al Khamaiseh, M. Qasaimeh, R. Al Qassas, A systematic review of detection and prevention techniques of SQL injection attacks, *Information Security Journal: A Global Perspective*, 2021, **32**, 252–265, doi: 10.1080/19393555.2021.1995537.

[38] T. Victor-Mgbachi, Navigating cybersecurity beyond compliance: understanding your threat landscape and vulnerabilities, *Iconic Research and Engineering Journals*, 2024, **7**, 70–81.

[39] Y. Sabla, S. Gour, Social media networking analytics and growth perspectives, In: ICT: Innovation and Computing (ICTCS 2023), Lecture Notes in Networks and Systems, Springer Nature, 2024, **879**, 273–285, doi: 10.1007/978-981-99-9486-1_22.

[40] S. Hariharan, A. Velicheti, A. S. Anagha, C. Thomas, N. Balakrishnan, Explainable artificial intelligence in cybersecurity: a brief review, In: 2021 4th International Conference on Security and Privacy (ISEA ISAP), IEEE, 2021, 01–12, doi: 10.1109/ISEA-ISAP54304.2021.9689765.

[41] M. Kuzlu, C. Fair, O. Guler, Role of artificial intelligence in the internet of things (IoT) cybersecurity, *Discover Internet of Things*, 2021, **1**, 07–12, doi: 10.1007/s43926-020-00001-4.

[42] S. Kumar, U. Gupta, A. K. Singh, A. K. Singh, Artificial intelligence: Revolutionizing cybersecurity in the digital era, *Journal of Computational Mechanics and Management*, 2023, **2**, 31–42, doi: 10.57159/gadl.jcmm.2.3.23064.

[43] R. Prasad, V. Rohokale, Artificial intelligence and machine learning in cybersecurity, In: Cyber Security: The Lifeline of Information and Communication Technology, Springer, Cham, 2020, 231–247, doi: 10.1007/978-3-030-31703-4_16.

[44] M. Ozkan-Ozay, E. Akin, Ö. Aslan, S. Kosunalp, T. Iliev, I. Stoyanov, I. Beloev, A comprehensive survey: Evaluating the efficiency of artificial intelligence and machine learning techniques on cybersecurity solutions, *IEEE Access*, 2024, **12**, 12229, doi: 10.1109/ACCESS.2024.3355547.

[45] F. Farivar, M. S. Haghighi, A. Jolfaei, M. Alazab, Artificial intelligence for detection, estimation, and compensation of malicious attacks in nonlinear cyber-physical systems and industrial IoT, *IEEE Transactions on Industrial Informatics*, 2019, **16**, 2716–2725, doi: 10.1109/TII.2019.2956474.

[46] A. Massaro, M. Gargaro, G. Dipierro, A. M. Galiano, S. Buonopane, Prototype cross-platform oriented on cybersecurity, virtual connectivity, big data and artificial intelligence control, *IEEE Access*, 2020, **8**, 197939–197954, doi: 10.1109/ACCESS.2020.3034399.

[47] B. Naik, A. Mehta, H. Yagnik, M. Shah, The impacts of artificial intelligence techniques in augmentation of cybersecurity: a comprehensive review, *Complex & Intelligent Systems*, 2022, **8**, 1763–1780, doi: 10.1007/s40747-021-00494-8.

[48] S. Gaur, M. S. Dulawat, A perception of statistical inference in data mining, *International Journal of Computer Science & Communication*, 2012, **1**, 635–658.

[49] S. D. Degadwala, S. Gaur, Privacy preserving system using pseudo-Zernike moment with SURF and affine transformation on RST attacks, *International Journal of Computer Science & Information Security*, 2017, **15**, doi: 10.1007/978-3-319-73712 6_2.

[50] S. D. Degadwala, S. Gaur, A study on privacy preserving system based on progressive VCS and RST attack, In: International Conference on Global Trends in Signal Processing, Information Computing and Communication

(ICSPICC-2016), IEEE, 2016, 38–142, doi: 10.1109/ICGTSPICC.2016.7955285.

[51] S. D. Degadwala, S. Gaur, Two-way privacy preserving system using combined approach: QR code and VCS, In: International Conference on Innovations in Power and Advanced Computing Technologies (I-PACT), IEEE, 2017, 01–05.

[52] Statista, Global AI usage in cybersecurity by use areas, https://www.statista.com/statistics/1610539/global-ai-usage-in-cybersecurity-by-use-areas, Accessed: December 2025.

[53] Enterprise Apps Today, AI use in cybersecurity statistics, https://www.enterpriseappstoday.com/stats/ai-use-in-cyber-security-statistics.html, Accessed: December 2025.

[54] Industry Research, Artificial intelligence-based cybersecurity market report, https://www.industryresearch.biz/market-reports/artificial-intelligence-based-cybersecurity-market-104058, Accessed: December 2025.

[55] Economic Times CISO, AI adoption surges in Indian cybersecurity: 94% of enterprises already using it, https://ciso.economictimes.indiatimes.com/news/cybercrime-fraud/report-ai-adoption-surges-in-indian-cybersecurity-94-of-enterprises-already-using-it/125213801, Accessed: December 2025.

[56] Express Computer, AI adoption in cybersecurity surges across India: 94% of organisations already using it, https://www.expresscomputer.in/news/ai-adoption-in-cybersecurity-surges-across-india-94-of-organisations-already-using-it/129264, Accessed: December 2025.

[57] Grand View Research, AI in cybersecurity market (India), https://www.grandviewresearch.com/horizon/outlook/ai-in-cybersecurity-market/india, Accessed: December 2025.

[58] Electro IQ, AI in cybersecurity statistics, https://electroiq.com/stats/ai-in-cybersecurity-statistics, Accessed: December 2025.

[59] WiFi Talents, AI in the cybersecurity industry statistics, https://wifitalents.com/ai-in-the-cyber-security-industry-statistics, Accessed: December 2025.

[60] All About AI, AI statistics in cybersecurity, https://www.allaboutai.com/resources/ai-statistics/cybersecurity, Accessed: December 2025.

[61] Future Market Insights, Artificial intelligence in cybersecurity market report, https://www.futuremarketinsights.com/reports/artificial-intelligence-in-cybersecurity-market, Accessed: December 2025.

[62] Precedence Research, Artificial intelligence in cybersecurity market, https://www.precedenceresearch.com/artificial-intelligence-in-cybersecurity-market, Accessed: December 2025.

[63] C. Chen, Y. Wu, Q. Dai, H. Y. Zhou, M. Xu, S. Yang, X. Han, Y. Yu, A survey on graph neural networks and graph transformers in computer vision: a task-oriented perspective, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024, **46**, 10297–10318, doi: 10.1109/TPAMI.2024.3445463.

Citation

P. Soni, S. Gour, Artificial intelligence and machine learning in cybersecurity: a review of trends, challenges, and applications, *Journal of Smart Sensors and Computing,* 2026, 2(1), 26202, doi: 10.64189/ssc.26202.