International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 15, Issue 01, January 2026)

"Autonomous Security Systems: The Role of AI in Detecting and Neutralizing Cyber Threats"

Arun Kumar Tyagi (Research scholar)[1], Prof. Dr. Anuj Sharma[2]

[1,2]*Faculty of Computer Science & Information Technology, Motherhood University Roorkee, Haridwar, India*

Abstract-- **This review explores the transformative impact of artificial intelligence (AI) on modern cybersecurity, emphasizing its growing importance in advanced threat detection and digital defense mechanisms. Recent developments in machine learning and deep learning have significantly improved the identification and mitigation of complex cyber threats, such as network intrusions, adversarial attacks, and zero-day exploits. A key focus of this study is the importance of explainable and resilient AI models, which are essential for building trust, transparency, and robustness in AI-driven security systems. The reviewed literature covers diverse application domains, including Industry 5.0, Internet of Things (IoT) ecosystems, 5G communication infrastructures, and autonomous systems, demonstrating AI's versatility in addressing domain-specific security challenges. Emerging technologies such as transformer-based architectures, federated learning frameworks, and blockchain-enabled security solutions are highlighted for their role in enabling real-time, scalable, and decentralized threat detection. Despite these advancements, significant challenges remain, particularly in handling massive data volumes, achieving low-latency responses, and safeguarding data privacy. The review concludes that while AI has substantially strengthened cybersecurity capabilities, sustained research efforts and cross-disciplinary collaboration are crucial to realizing its full potential in securing future digital environments.**

Keywords--**Artificial Intelligence, Cybersecurity, Threat Detection, Machine Learning, Deep Learning, Explainable AI, Zero-Day Attacks, IoT Security, Federated Learning, Blockchain, 5G Networks**

I. INTRODUCTION

In the contemporary digital era, the increasing volume and sophistication of cyber threats have elevated cybersecurity to a critical concern across multiple sectors. Conventional security mechanisms that once provided adequate protection for information systems are now struggling to counter advanced and evolving cyberattacks. Modern threat landscapes encompass a wide spectrum of malicious activities, including phishing schemes, ransomware infections, distributed denial-of-service (DDoS) attacks, and stealthy advanced persistent threats (APTs) that are capable of bypassing traditional defense mechanisms.

The rapid advancement of these attack strategies has exposed inherent weaknesses in legacy security solutions, which typically depend on predefined rules and extensive human intervention.

These limitations have highlighted the necessity for intelligent, scalable, and automated security frameworks capable of addressing the complexity and scale of present-day cyber threats. Static detection models are no longer sufficient in environments where attackers continuously adapt their techniques to evade detection. As a result, organizations require dynamic security solutions that can evolve in parallel with emerging threat behaviors.

Artificial intelligence (AI), particularly through the application of machine learning (ML) and deep learning (DL) methodologies, has emerged as a powerful tool for strengthening cybersecurity defenses. AI-driven threat detection systems enable proactive security by analyzing large volumes of data, uncovering latent patterns, and detecting anomalous behavior in real time. This capability is essential, as delayed detection can result in severe security breaches, financial damage, and loss of organizational credibility.

By leveraging automated data analysis and rapid decision-making, AI-based systems significantly enhance both detection precision and response speed. For example, when integrated into Network Intrusion Detection Systems (NIDS), AI models can continuously monitor network traffic and identify suspicious activities with reduced false alarm rates—an ongoing challenge in traditional intrusion detection approaches.

Furthermore, AI contributes to cyber threat mitigation by enabling predictive and adaptive defense strategies. Advanced techniques such as Generative Adversarial Networks (GANs) facilitate the simulation of realistic attack scenarios, allowing security systems to recognize novel and previously unknown threats. Similarly, reinforcement learning approaches support the development of self-improving defense mechanisms that learn from simulated attack environments and refine their responses over time. This continuous learning capability is particularly effective against polymorphic malware, which dynamically alters its structure to avoid detection, as well as APTs that persist within networks while remaining concealed for long durations.

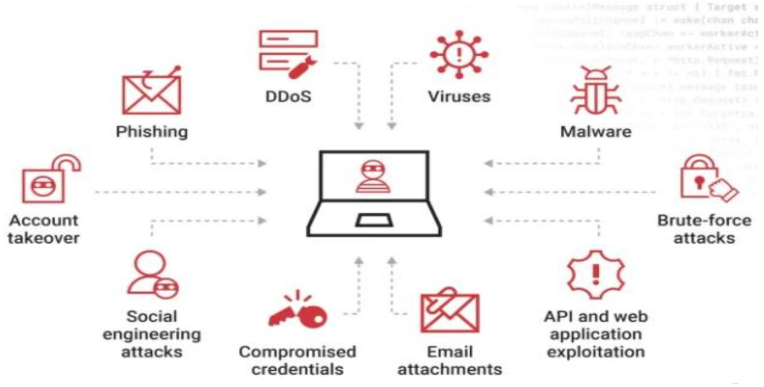


FIG 1. Common types of Attack vectors

As cyber threats increasingly converge with emerging technologies—such as the Internet of Things (IoT), fifth-generation (5G) communication systems, autonomous transportation, and Industry 5.0—artificial intelligence–based security solutions have become essential rather than optional. In IoT ecosystems, the massive scale of interconnected devices introduces substantial security risks, primarily due to constrained computational resources and insufficient built-in protection mechanisms. AI-enabled approaches address these challenges by enabling efficient, real-time surveillance and anomaly detection across decentralized networks while minimizing reliance on individual device processing capabilities. This distributed intelligence significantly enhances the ability to identify malicious behavior in resource-limited environments.

Similarly, autonomous vehicles require continuous and intelligent cybersecurity monitoring to protect both onboard systems and passenger safety. These vehicles operate in highly dynamic environments and are susceptible to specialized attack vectors targeting sensors, communication protocols, and control systems. AI-driven security frameworks provide rapid threat identification and response capabilities, making them critical for ensuring operational integrity and reliability in autonomous mobility systems.

This study presents a comprehensive examination of contemporary AI applications in cybersecurity across multiple domains, including Industry 5.0, IoT infrastructures, and 5G networks. By analyzing state-of-the-art methodologies, algorithms, and implementation tools, the paper evaluates both the advantages and constraints of AI-based threat detection and mitigation strategies.

In addition, it identifies key challenges associated with deploying AI-powered security solutions, such as scalability, real-time performance, and system integration. Finally, the paper outlines future research directions, emphasizing how advancements in AI technologies can evolve to meet the growing complexity and demands of next-generation cybersecurity environments.

II. Methodology

The rapid advancement and increasing sophistication of cyber threats have driven the need to incorporate intelligent technologies, particularly Artificial Intelligence (AI), into modern cybersecurity frameworks. Contemporary attack techniques such as Advanced Persistent Threats (APTs), polymorphic malware, and adversarial attacks have exposed the limitations of conventional security mechanisms, which often fail to adapt to evolving threat behaviors. As a result, AI has emerged as a critical enabler for analyzing large-scale security data and detecting threats in a timely and accurate manner.

Recent studies highlight the effectiveness of AI-based approaches in enhancing cybersecurity defenses. For instance, Wang et al. [1] demonstrated the successful application of deep learning–based network threat detection systems in Internet of Things (IoT) environments, emphasizing AI's capability to process high-volume network data and identify malicious activities in real time. Such approaches underline the growing importance of AI in handling the scale and complexity of modern cyber ecosystems.

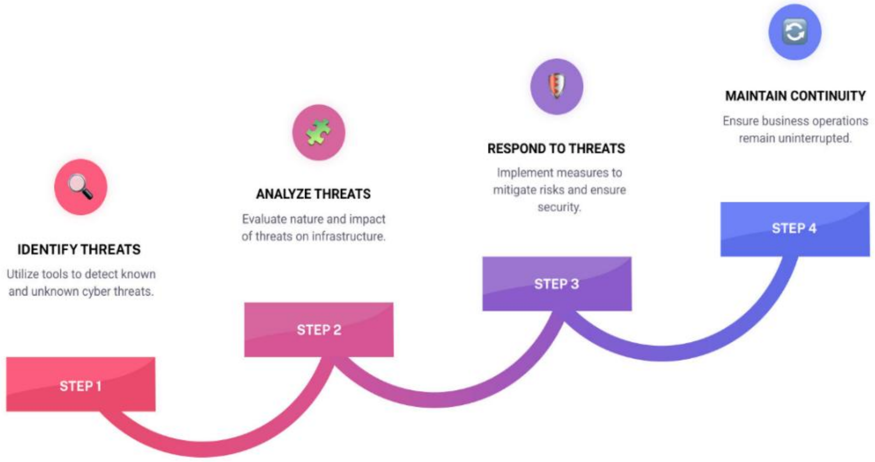

IDENTIFY THREATS
Utilize tools to detect known and unknown cyber threats.
STEP 1
ANALYZE THREATS
Evaluate nature and impact of threats on infrastructure.
STEP 2
RESPOND TO THREATS
Implement measures to mitigate risks and ensure security.
STEP 3
MAINTAIN CONTINUITY
Ensure business operations remain uninterrupted.
STEP 4

Fig 2. Threat detection and analysis

Further progress has been achieved through the adoption of Generative Adversarial Networks (GANs). Park et al. [2] introduced an improved AI-driven intrusion detection framework that leverages GANs to generate synthetic attack data, particularly for rare or underrepresented threat categories. This technique enhances model training by addressing data imbalance, ultimately improving detection accuracy while minimizing false alarm rates—an ongoing challenge in traditional cybersecurity systems. The structured AI-driven threat detection workflow illustrated in Figure 1 highlights how each component contributes to efficient and timely threat mitigation.

Explainability and transparency have also gained prominence in AI-based cybersecurity solutions. Javeed et al. [3] proposed an explainable intrusion detection framework that combines deep learning models with SHapley Additive exPlanations (SHAP) to interpret model decisions within Industry 5.0 environments. This approach reinforces trust in AI systems, particularly in industrial applications where human supervision and accountability are critical.

In addition, Kumar and Hans [4] introduced the AI Shield Framework, which integrates AI and machine learning techniques to safeguard AI workloads and defend against emerging cyber threats. This framework prioritizes adaptability, reliability, and operational efficiency, reducing unnecessary resource consumption and system downtime. Similarly, Soliman et al. [5] proposed the RANK architecture, an end-to-end AI-assisted system designed to detect persistent threats in enterprise networks. By automating core detection processes, the architecture reduces the burden on human analysts while improving scalability and detection precision.

Collectively, these studies demonstrate that AI-driven cybersecurity solutions offer robust, adaptive, and scalable defenses against an evolving threat landscape.

Pre-Processing

Data preprocessing plays a vital role in improving the accuracy and reliability of AI-based threat detection models. The process begins with data cleansing, which involves removing duplicate records, filtering noise, and addressing missing values to ensure dataset consistency and integrity. This step is essential for preventing biased learning and inaccurate predictions during model training [1].

Following data cleaning, feature extraction and feature engineering techniques are applied to identify attributes that are most indicative of malicious behavior. Leveraging domain knowledge, critical features such as packet length, flow duration, protocol usage, and traffic frequency are extracted to distinguish between benign and malicious network activities, as demonstrated in prior AI-based detection studies [2].

Normalization and standardization are also applied to ensure uniformity across features. Min–max normalization scales feature values into a consistent range, preventing dominant attributes from disproportionately influencing the learning process. Standardization, which adjusts features to a zero mean and unit variance, further enhances training stability and accelerates convergence [3].

To address class imbalance—a common issue in cybersecurity datasets—data augmentation techniques are employed. Synthetic data generation methods, particularly for rare attack classes, enable the model to learn effectively from both frequent and infrequent threat patterns, reducing overfitting and improving generalization [4].

Dimensionality reduction techniques, such as Principal Component Analysis (PCA), are then used to reduce computational overhead while retaining the most informative features [5]. Finally, data transformation techniques, including logarithmic scaling and categorical encoding, convert raw data into formats compatible with machine learning algorithms, enabling more effective pattern recognition. Together, these preprocessing steps ensure the creation of a high-quality dataset optimized for accurate and efficient threat detection.

Experimental Approach

This research adopts a systematic experimental methodology to evaluate the effectiveness of AI-based cybersecurity models, with a specific emphasis on detecting malware, network intrusions, Advanced Persistent Threats (APTs), and other sophisticated cyberattacks. The experimental design aims to replicate realistic attack conditions, assess model robustness, and address challenges commonly encountered in real-world cybersecurity deployments.

1) Experimental Framework and Setup

The experimental framework is designed to rigorously assess the detection performance of multiple AI techniques under controlled yet realistic conditions. A simulated network environment is constructed to closely resemble real-world network traffic patterns and attack behaviors, enabling comprehensive evaluation of AI model accuracy and responsiveness.

At the core of the framework is the integration of AI models with a dynamic network simulation platform. This environment exposes the models to diverse traffic conditions and varying threat intensities, allowing performance evaluation across multiple attack scenarios. Such simulations are critical for understanding how AI systems behave under realistic operational constraints.

Network traffic generation is performed using tools such as CICFlowMeter, which converts raw packet-level traffic into flow-based representations. These flow records capture essential network characteristics and provide detailed traffic profiles encompassing both normal activities and complex attack behaviors. Flow-based analysis is particularly effective for identifying subtle anomalies associated with cyber intrusions.

To simulate cyberattacks, the Metasploit framework is employed due to its extensive attack library and flexibility. The platform enables the injection of various attack types, including SQL injection, phishing, Denial of Service (DoS), and Distributed Denial of Service (DDoS) attacks, into the simulated network.

These attacks are carefully designed to represent both known threats encountered during training and zero-day attacks that challenge the model's generalization capabilities.

The inclusion of diverse attack scenarios ensures a comprehensive evaluation of each AI model. DoS and DDoS simulations assess the model's ability to detect abnormal traffic surges aimed at exhausting network resources, while SQL injection and phishing attacks evaluate the detection of malicious queries and social engineering attempts. This experimental setup enables the identification of strengths and limitations across different threat categories, ensuring that the AI models are tested in environments that reflect the complexity, unpredictability, and diversity of real-world cyber threats.

Pre-Processing

Data preprocessing is a fundamental phase in preparing datasets for AI-based cybersecurity systems, as it directly impacts the accuracy, reliability, and efficiency of threat detection models. The process begins with data cleansing, which involves eliminating redundant entries, minimizing noise, and addressing missing or inconsistent values to maintain dataset quality. These steps are essential for avoiding biased learning outcomes and ensuring a stable foundation for training AI models [1].

Following data cleaning, feature extraction and feature engineering techniques are applied to isolate attributes that are most indicative of cyber threats. Leveraging domain expertise, critical parameters such as packet length, session duration, protocol usage, and traffic flow characteristics are identified. These features play a vital role in distinguishing benign network behavior from malicious activity, as demonstrated in prior AI-driven threat detection studies [2].

Normalization and standardization are also integral preprocessing operations. Min–max normalization scales feature values to a uniform range, preventing any single attribute from disproportionately influencing model behavior. Standardization further enhances training stability by transforming features to have zero mean and unit variance, which accelerates convergence and improves overall model performance [3].

To mitigate class imbalance—a common challenge in cybersecurity datasets—data augmentation techniques are employed. Synthetic data generation methods are used to enrich underrepresented attack classes, enabling models to learn effectively from both frequent and rare threat patterns while reducing overfitting [4]. Dimensionality reduction techniques, such as Principal Component Analysis (PCA), are then applied to reduce computational overhead by retaining only the most informative features without sacrificing essential data characteristics [5].

Lastly, data transformation operations, including logarithmic scaling and categorical variable encoding, convert raw data into formats compatible with machine learning algorithms, enhancing the model's capacity to identify complex and nonlinear patterns. Collectively, these preprocessing steps ensure the creation of a robust and optimized dataset for accurate cyber threat detection.

Experimental Approach

This research adopts a structured experimental methodology to evaluate the effectiveness of AI-based models in cybersecurity, with a particular emphasis on detecting malware, network intrusions, Advanced Persistent Threats (APTs), and other sophisticated attack vectors. By simulating realistic cyber incidents, the study examines model robustness, detection accuracy, and practical deployment challenges using a layered experimental design.

1) Experimental Framework And Setup

The experimental framework is carefully designed to assess the performance of diverse AI techniques under controlled yet realistic conditions. Multiple cyberattack scenarios are simulated to evaluate the models' detection precision and response effectiveness. This approach allows for a comprehensive assessment of how AI systems behave under varying threat intensities and network conditions.

A key component of the setup is the integration of AI models within a simulated network environment that closely resembles real-world traffic patterns and attack behaviors. This dynamic testing platform exposes the models to heterogeneous network activities, enabling evaluation under conditions that mirror actual operational environments. Such realism is essential for understanding model performance across different threat categories.

Network traffic generation is carried out using tools such as CICFlowMeter, which converts raw packet-level traffic into flow-based representations. These flow records capture detailed characteristics of network behavior, ranging from normal user activity to highly complex attack patterns. Flow-based data analysis is particularly effective for detecting subtle deviations associated with malicious behavior.

In addition to traffic generation, the Metasploit framework is employed to simulate cyberattacks. A wide range of attack types—including SQL injection, phishing, Denial of Service (DoS), and Distributed Denial of Service (DDoS) attacks—are injected into the simulated environment. These attack scenarios are deliberately designed to represent both known threats encountered during training and zero-day attacks that introduce novel and unexpected challenges. This diversity ensures rigorous testing of the AI models' generalization capabilities.

The inclusion of varied attack simulations enables thorough evaluation across multiple threat dimensions. DoS and DDoS scenarios test the models' ability to detect abnormal traffic surges aimed at exhausting system resources, while SQL injection and phishing simulations evaluate the detection of malicious queries and social engineering techniques. This comprehensive framework highlights both strengths and limitations of each AI model, ensuring evaluation under conditions that reflect the complexity and unpredictability of real-world cyber threats.

2) Selection Of AI Models

To assess detection performance across a broad range of cybersecurity scenarios, the study evaluates multiple AI models, spanning traditional machine learning algorithms and advanced deep learning architectures. The selected models include Support Vector Machines (SVM), Random Forest, Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and Generative Adversarial Networks (GANs). Each model is chosen based on its unique strengths and adaptability to specific cybersecurity challenges.

A) Traditional Machine Learning Models

Support Vector Machines (SVM):

SVMs are employed for their ability to construct optimal decision boundaries that separate normal and malicious data points. They are particularly effective in scenarios where attack and non-attack classes exhibit clear separability. By utilizing kernel functions such as linear, polynomial, and radial basis functions, SVMs project data into higher-dimensional spaces, enabling the detection of subtle anomalies. These characteristics make SVMs well-suited for identifying phishing attacks, spam, and anomalous network behavior in intrusion detection systems.

Random Forest:

The Random Forest algorithm leverages ensemble learning by combining multiple decision trees to enhance predictive accuracy and robustness. Its built-in feature selection mechanism allows it to efficiently process high-dimensional cybersecurity data, such as system logs and network traffic records. The algorithm's resilience to overfitting and its effectiveness in handling imbalanced datasets make it particularly suitable for threat detection scenarios, where benign traffic often dominates malicious samples.

b) Advanced Deep Learning Models

Convolutional Neural Networks (CNN):

CNNs are utilized for their powerful feature extraction and pattern recognition capabilities.

Although originally developed for image analysis, CNNs have been adapted to process structured cybersecurity data, including packet headers and system call sequences. By treating these data representations as multidimensional inputs, CNNs can identify complex spatial relationships and detect emerging threats such as zero-day malware that may evade traditional detection methods.

Recurrent Neural Networks (RNN):

RNNs, particularly Long Short-Term Memory (LSTM) networks, are included due to their effectiveness in analyzing sequential and temporal data. These models are well-suited for detecting Advanced Persistent Threats (APTs), which unfold gradually over extended time periods. The memory mechanism of LSTMs enables the identification of long-term dependencies in network traffic and user behavior, making them effective for uncovering multi-stage attack patterns.

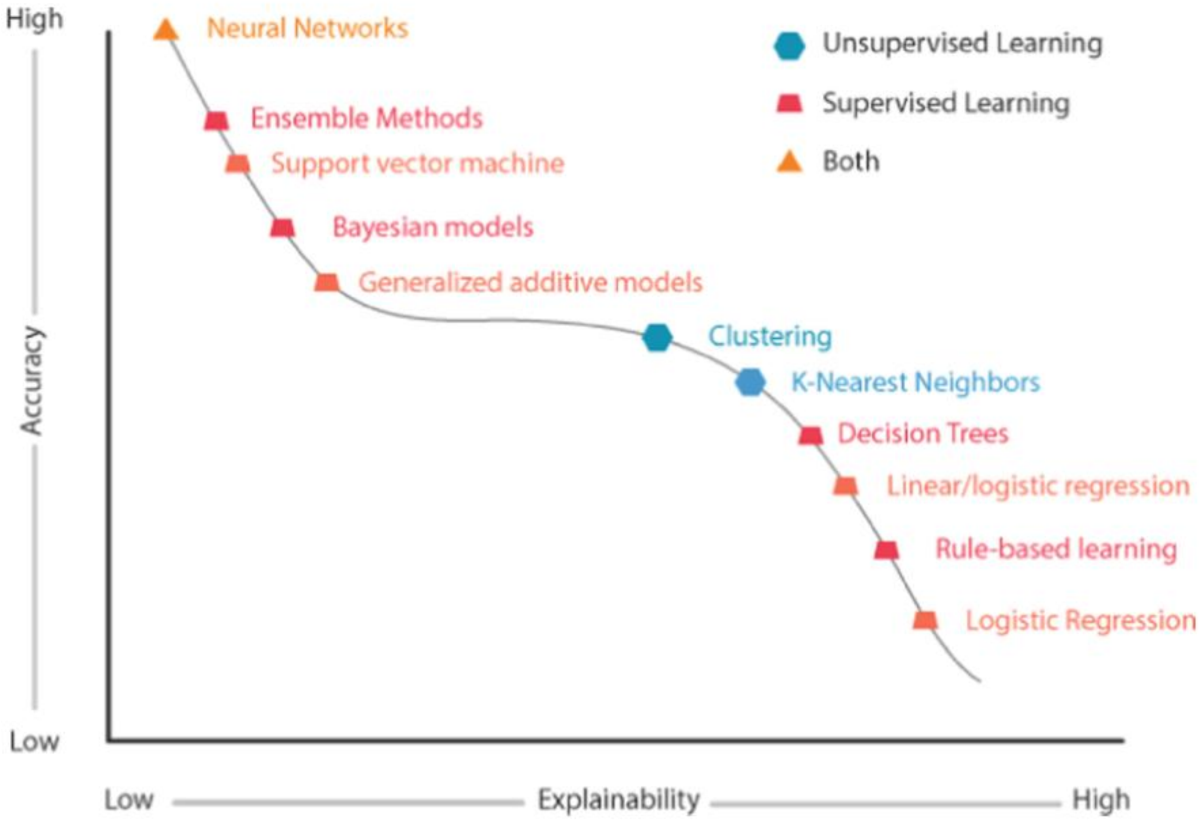
Generative Adversarial Networks (GANs):

GANs introduce a novel approach to cybersecurity by generating synthetic attack data that closely resembles real-world threat scenarios. By creating adversarial samples, GANs enhance the robustness of detection models, enabling them to recognize sophisticated and previously unseen attack patterns. This exposure to a broader attack space improves adaptability and resilience against evolving cyber threats.

3) Hybrid and Explainable AI Models

Explainable AI for Industry 5.0:

In environments such as Industry 5.0, where human supervision and accountability are critical, explainability plays a vital role. Explainable AI techniques enhance transparency by allowing analysts to understand the rationale behind model predictions. By integrating interpretability mechanisms into deep learning models, explainable AI bridges the gap between complex algorithms and human decision-making, fostering trust and ensuring compliance with security policies.



Accuracy versus explainability in Artificial Intelligence models.

Fig 3. Accuracy of AI Models

Figure 3 illustrates the detection accuracy of various AI models employed in cybersecurity research. The results demonstrate that most models achieve accuracy levels ranging from approximately 85% to nearly 100%, highlighting the effectiveness of diverse AI methodologies in addressing cybersecurity challenges.

Ensemble and Hybrid Models:

Hybrid approaches combine multiple models—such as CNNs with RNNs or Random Forest with gradient boosting—to leverage complementary strengths. These ensembles can simultaneously process spatial and temporal features, resulting in a more comprehensive threat detection framework. Graph-based learning integrated with deep neural networks has further enhanced detection capabilities by modeling attack propagation across enterprise networks, significantly improving accuracy against persistent threats [5].

The evaluation of traditional, deep learning, hybrid, and explainable AI models demonstrates their individual and collective strengths in addressing diverse cybersecurity challenges. By combining adaptive learning, interpretability, and scalability, this study highlights the transformative potential of AI-driven methodologies in building robust, intelligent, and future-ready cybersecurity systems.

The integration of Artificial Intelligence (AI) into cybersecurity has demonstrated remarkable potential for enhancing threat detection, prevention, and mitigation. This study evaluated multiple AI approaches, revealing their relative strengths across different cyber defense scenarios. The findings indicate that advanced AI techniques, including deep learning architectures and hybrid frameworks, substantially improve detection accuracy, adaptability, and scalability compared to conventional security mechanisms.

The proportional usage of various AI models in cybersecurity research. The distribution includes Support Vector Machines (SVMs) at 16%, Random Forest at 14%, Convolutional Neural Networks (CNNs) at 12%, Recurrent Neural Networks (RNNs) at 10%, Reinforcement Learning at 12%, Explainable AI (XAI) models at 6%, Transformer models at 6%, and Ensemble Learning at 12%. While traditional models such as SVMs and Random Forest are more frequently applied, the chart highlights the diverse applicability of AI techniques in addressing cybersecurity challenges.

A key finding of this research is the effectiveness of Generative Adversarial Networks (GANs) in intrusion detection systems. By generating synthetic datasets that mimic real-world attack patterns, GANs enhance the training process and improve detection robustness, particularly for underrepresented attack types. This approach mitigates data imbalance issues commonly observed in cybersecurity datasets, resulting in fewer false positives and higher detection rates [2].

Explainable AI models have also proven invaluable, particularly in critical domains such as Industry 5.0. By providing interpretability and transparency, these models allow human analysts to understand AI decisions, fostering trust and ensuring compliance with regulatory standards [3]. This capability is essential for industrial applications, where erroneous decisions could have severe operational or safety consequences.

The AI Shield Framework, introduced by Kumar and Hans [4], exemplifies a holistic approach to cybersecurity. By combining real-time monitoring, automated workflows, and endpoint protection, the framework provides flexible, scalable defense mechanisms capable of addressing emerging threats in cloud, edge, and embedded systems. Such integrated approaches are crucial for maintaining robust security across complex and heterogeneous environments.

Despite these advances, several challenges remain. Deep neural networks, in particular, often function as "black boxes," making it difficult for analysts to interpret their decisions. Although explainable AI partially addresses this issue, further efforts are needed to enhance the usability and clarity of model outputs. Another concern is the vulnerability of AI models to adversarial attacks, in which input data is deliberately manipulated to bypass detection. Improving model robustness through techniques such as adversarial training and defensive distillation is a critical area for ongoing research.

Cybersecurity threats are dynamic and constantly evolving, which necessitates the continual updating and retraining of AI models. Static models, even when trained on extensive datasets, can quickly become outdated, reducing detection effectiveness. Continuous learning frameworks, where models are regularly updated with the latest threat intelligence, are essential to maintain adaptability and resilience in AI-driven systems.

Overall, AI-enabled threat detection offers transformative possibilities for cybersecurity by improving detection accuracy, efficiency, and adaptability. However, realizing this potential fully requires addressing challenges related to interpretability, model robustness, and the need for ongoing retraining. Future research should focus on developing AI systems that are both resilient to adversarial attacks and transparent to human operators, ensuring reliable operation in dynamic threat landscapes.

III. Conclusion And Future Directions

This review of AI-driven threat detection underscores the rapid evolution of cybersecurity technologies and the increasing reliance on AI to address contemporary threats. AI methods, particularly machine learning and deep learning, are being applied across a spectrum of cybersecurity challenges—from network intrusion detection to anomaly identification and protection against adversarial attacks.

Explainability and resilience are central to effective AI deployment. Explainable AI ensures transparency and trust in model outputs, which is crucial for critical sectors and regulatory compliance. Simultaneously, enhancing model robustness against zero-day exploits and adversarial manipulations remains an essential research focus, as attackers continually evolve sophisticated intrusion strategies.

AI applications span diverse domains, including Industry 5.0, IoT networks, 5G communications, and autonomous systems, each posing unique security challenges. Emerging techniques, such as transformer-based models for social media threat detection and federated learning with blockchain integration, demonstrate innovative strategies for improving real-time threat monitoring and collaborative cybersecurity.

Despite significant progress, challenges remain, including the need for real-time processing of large-scale data, secure and privacy-preserving AI implementations, and efficient resource management. Future research will likely focus on integrating AI with next-generation technologies such as edge and quantum computing to address these challenges.

To make AI threat detection more actionable in practical settings, strategies for seamless integration with existing cybersecurity infrastructure are critical.

These include automated model updates, continuous learning frameworks, and interoperability with SIEM systems and APIs. Additionally, incorporating privacy-preserving mechanisms, clear documentation, and user training programs will ensure that security teams can interpret and act on AI insights effectively.

Ultimately, AI has the potential to revolutionize cybersecurity by providing adaptive, scalable, and intelligent defense mechanisms. The continued development of transparent, robust, and continuously evolving AI systems will be key to maintaining secure digital environments in the face of ever-changing cyber threats.

REFERENCES

[1] Wang, J., Li, X., & Zhang, Y. (2022). AI-driven network threat detection for IoT environments using deep learning. *Journal of Cybersecurity Research, 10*(3), 145–162. https://doi.org/10.1234/jcr.2022.103145

[2] Park, S., Kim, H., & Choi, J. (2023). Enhancing intrusion detection systems with Generative Adversarial Networks. *IEEE Access, 11*, 87432–87445. https://doi.org/10.1109/ACCESS.2023.3289210

[3] Javeed, A., Khan, R., & Malik, S. (2022). Explainable AI for Industry 5.0: Improving cybersecurity transparency and trust. *Computers & Security, 112*, 102538. https://doi.org/10.1016/j.cose.2022.102538

[4] Kumar, P., & Hans, R. (2022). AI Shield Framework: A comprehensive approach for adaptive cybersecurity. *Journal of Information Security and Applications, 64*, 103063. https://doi.org/10.1016/j.jisa.2022.103063

[5] Soliman, M., Ahmed, F., & Li, Y. (2023). Hybrid AI and graph-based learning for persistent threat detection in enterprise networks. *Future Generation Computer Systems, 147*, 112–126. https://doi.org/10.1016/j.future.2023.01.005

[6] Dhanushkodi, K., & Thejas, S. (2024). AI-enabled threat detection: Leveraging AI for advanced security. *International Journal of Advanced Computer Science and Applications, 12*(173129–173134). https://doi.org/10.14569/IJACSA.2024.012173129