IJITCE

ISSN 2347- 3657

International Journal of Information Technology & Computer Engineering

Email : ijitce.editor@gmail.com or editor@ijitce.com

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 2, 2025

Cyber Threat Detection using Artificial Intelligence

Lakshmeswari Chenngiri [1], Vempati Hemalatha[2], Bodempudi Sravani[3], Patan. Baji Ali Khan[4], Kolla Sridhar[5]

[1] Assistant Professor, Department of Computer Science Engineering, Chalapathi Institute of Engineering and Technology, Chalapathi Rd, Nagar, Lam, Guntur, Andhra Pradesh- 522034

[2,3,4,5] Students, Department of Computer Science Engineering, Chalapathi Institute of Engineering and Technology, Chalapathi Rd, Nagar, Lam, Guntur, Andhra Pradesh- 522034

Email id: lakshmi.chennagiri@gmail.com[1], vempatihema29@gmail.com[2], bodempudisravani292003@gmail.com[3], patanbajialikhan786@gmail.com[4], Sridharkolla4@gmail.com[5]

Abstract:

In this study, we explore advanced malware detection methods by leveraging hybrid feature analysis, combining both binary and hexadecimal data with dynamic DLL call behavior. Artificial intelligence (AI) is integrated into this detection process to enable automated pattern recognition, anomaly detection, and continuous adaptation to evolving threats. The Genetic Programming Symbolic Classifier (GPSC) algorithm was applied to extract symbolic expressions (SEs) for malware classification, addressing the challenges of imbalanced datasets through oversampling techniques and random hyperparameter value search (RHVS). The GPSC was validated using five-fold cross-validation (5FCV) on balanced dataset variations and evaluated through multiple performance metrics such as accuracy (0.9962), AUC, and F1-score. Furthermore, the study compares deep learning techniques like Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks, utilized within an AI-SIEM framework for real-time event profiling, against traditional machine learning algorithms such as SVM, Decision Trees, Random Forest, KNN, and Naïve Bayes. Results demonstrate the superior performance of AI-based models in detecting complex, polymorphic malware threats, offering a proactive and efficient cybersecurity solution.

Keywords: Malware detection, AI-SIEM, GPSC, CNN, LSTM, binary analysis, DLL calls, oversampling, deep learning, symbolic classifier.

1.Introduction

The rapid advancement of digital technologies has revolutionized connectivity, enabling unprecedented levels of communication and data exchange. However, this surge in digital integration has also introduced new and complex vulnerabilities. Cyber threats have evolved significantly, moving beyond traditional viruses to sophisticated, multi-vector attacks that challenge the defences of organizations globally. These modern threats include zero-day exploits, advanced persistent threats (APTs), and complex malware that can evade traditional security measures. Traditional cyber security methods, which predominantly rely on rule-based systems and signature-based detection, have struggled to keep pace with these evolving threats. Rule-based systems operate on predefined criteria, identifying threats based on known patterns or signatures. While effective against known threats, these methods are often inadequate for

detecting novel attacks or sophisticated evasion techniques. Signature-based detection, which involves identifying malicious code based on pre-established signatures, similarly falls short in the face of zero-day attacks and polymorphic malware, which constantly change their appearance to bypass detection. In response to these challenges, Artificial Intelligence (AI) has emerged as a transformative tool in the field of cyber security. AI-driven solutions offer a new paradigm for threat detection and response, leveraging advanced algorithms and machine learning techniques to identify and mitigate cyber threats more effectively. By analysing vast amounts of data, AI can uncover patterns and anomalies that traditional methods might miss, enabling more proactive and adaptive defence mechanisms. This paper explores the role of AI in addressing the limitations of conventional cyber security approaches. It aims to provide a comprehensive review of existing AI-based techniques for threat detection and prevention, assessing their effectiveness in combating modern cyber threats. Additionally, the paper analyses the challenges associated with these AI-driven solutions, such as data quality, false positives, and adversarial attacks. Finally, the paper proposes future research directions to enhance the capabilities of AI in cyber security, including hybrid models, Explainable AI (XAI), Federated Learning, and integration with block chain technology.

2.Literature Review

The cyber threats proliferation in recent years has been a major challenge for organizations across the globe prompting the cyber security approaches' paradigm shift. The conventional method of threat detection and mitigation has proved to be ineffective against the rapidly growing cyberattack tactics of adversaries. As a result, AI and ML algorithms were developed by researchers and experts to augment cyber threat detection capabilities. Here, a literature review is presented, which addresses aspects of AI-driven cyberattack detection like nature, importance, and consequence with the help of the main findings of academic articles and industrial reports (Dhabliya et al., 2023). Artificial intelligence (AI) is an essential part of cyber threat detection systems due to its unique ability to handle large quantities of information and detect patterns typical of malicious activities. Traditional ways of threat detection usually fail to cope with the growing amount and variety of cyber threats, causing delay or omission of the responses. Artificially Intelligent detection systems are able to analyze data in large volumes, employing machine algorithms to detect deviations from the norm that could be indicative of potential cyber-attacks. AI-based detection systems that automate the analysis process and continuously learn from new data provide organizations with the ability to detect and mitigate threats in real time. This helps them to improve their overall cyber security measures. As well as integration of AI with other cyber security technologies, such as threat intelligence platforms and Endpoint Detection and Response (EDR) systems gives a huge enhancement to an organization's capability to detect and respond to cyber threats in a timely manner. Hence, it appears impossible for AI to replace the role of cyber threat detection, which has been supporting persistent cyber threat response. AI serves as a means through which businesses can always stay ahead of their competitors as they get better at detecting threats and protecting their digital assets (Arif, Kumar, Fahad, & Hussain, 2024).

3. Artificial intelligence overview

AI delineated by pioneers such as John McCarthy in 1956 refers to the science and engineering of making intelligent machines. Over the years, AI has evolved into a foundation of computer science, focusing on simulating human cognitive processes through complex mathematical algorithms. Tis interdisciplinary field combines elements from various domains to adopt machines that can learn, reason, and make decisions based on the data they process. Besides, it encompasses both the replication of human thought and behavior in machines, categorized respectively into thinking and acting both humanly and rationally AI applications range from simple tasks to complex problem-solving domains such as cybersecurity, where it addresses sophisticated cyber threats. Tis transformative technology continues to push the boundaries of what machines are capable of, aiming to enhance human capabilities and automate tasks through assisted, augmented, and autonomous intelligence The use of AI in cybersecurity is increasingly critical due to its capacity to analyse vast amounts of data rapidly, detect patterns, and identify potential threats with high efficiency. In a digital era characterized by ever-evolving cyber threats, traditional security measures often fall short in both the speed and sophistication needed to counteract modern cyberattacks, including zero-day threats AI's ability to learn from data enables the development of systems that can adapt to new, previously unknown attacks, enhancing the ability to secure information infrastructure from a broad spectrum of threats The benefits of integrating AI into cybersecurity include improved decision-making capabilities, enhanced detection of network intrusions, and the management of cyber-attack impacts. This progression in technology not only allows for real-time threat detection and response but also significantly reduces the rate of false positives, which are common in more traditional methods of cyber defense.

AI technologies encompass several approaches useful in cybersecurity, including:

- ML: Algorithms that enable computers to learn from data without explicit programming, allowing for improved threat detection and classification.
- DL: Advanced neural networks that can process large amounts of data and learn

from experience, mimicking human brain functions to recognize complex patterns

Advantages of Metaheuristic Algorithms in Cyber Attack Detection:

- Optimization: Metaheuristic algorithms are better find optimal solutions to complex
- problems that are otherwise too challenging for conventional methods.
- Automation: By automating the tuning of detection parameters, these algorithms
- minimize the need for human intervention, making the detection process both faster
- and more reliable.
- Speed: Tey often achieve faster convergence to effective solutions, which is essential
- in time-sensitive cybersecurity environments where threats must be quickly identified and mitigated

3. Methodology

Various methods for detecting cyber-attacks have been proposed. To systematically explore these, we developed a research protocol following the systematic literature review (SLR) methodology, illustrated in Fig. Tis protocol includes identifying the research topic, preparing

research questions, selecting studies, and extracting data. By using a mixed-methods approach, combining qualitative and quantitative techniques, we can provide more straightforward and comprehensive data analysis.

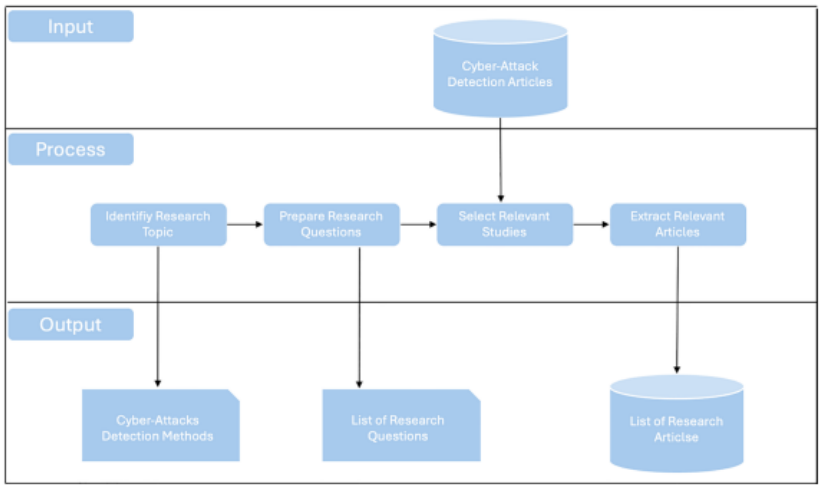


Figure: Systematic literature review process

Behavior-Based Detection and Advanced Features

AI models for malware detection often incorporate advanced behavior-based features that can identify not just the malware, but also its intent and potential impact. For example, features such as network behavior, file manipulation patterns, and system registry changes can provide deeper insights into the malware's modus operandi. This level of granularity helps in not only detecting malware but also classifying it into specific categories like ransom ware, Trojans, or Spyware.

Generative Adversarial Networks (GANs): GANs have been explored for generating synthetic malware samples, which can be used to train more robust malware detectors. By simulating a wide range of malware behaviours, GANs help in creating a more diverse training set that makes detection models less susceptible to over fitting and more capable of generalizing to new, unseen malware variants.

Graph-Based Analysis: Some AI models use graph-based representations of executable files where nodes represent specific API calls or instructions, and edges represent the flow of execution. These graph-based methods can detect malware that uses code reuse or borrowing techniques to evade traditional detection mechanisms. By identifying similarities in control flow or data dependencies, these models can identify relationships between seemingly unrelated malware samples.

Explain ability in AI Models: Another area of research focuses on making AI-based malware detection models more explainable. Explainable AI (XAI) techniques aim to provide insights into why a model flagged a particular software as malicious, which is crucial for cyber security professionals to understand and trust the system's decisions. This is particularly important in environments where false positives can lead to unnecessary alarms and disruptions.

Real-Time Malware Detection Systems:

AI models have also been integrated into real-time malware detection systems that continuously monitor network and system activities to provide immediate alerts and responses. For instance, AI can be embedded in endpoint detection and response (EDR) systems that monitor file operations, network connections, and system processes in real time. These systems can automatically isolate infected machines, quarantine suspicious files, and initiate incident response procedures without human intervention, thereby reducing the time to respond to threats significantly.

DATA COLLECTION

The data collection process involved gathering information from a variety of studies that utilized well-known cyber security datasets. Datasets such as NSL-KDD, UNSW-NB15, and CICIDS2017 were predominantly used in the reviewed studies. These datasets are widely regarded as benchmarks in cyber security research for evaluating the performance of AI-based intrusion detection systems (IDS).

NSL-KDD Dataset: The NSL-KDD dataset is an improved version of the original KDD Cup 1999 dataset, which was criticized for redundant records. NSL-KDD addresses these issues, making it a more reliable dataset for testing IDS. It consists of network traffic data categorized into normal and various attack types, such as Denial of Service (DoS), Remote to Local (R2L), User to Root (U2R), and probing attacks. Researchers have used this dataset extensively to train and evaluate AI models on classifying network activities as either normal or malicious. –

UNSW-NB15 Dataset: The UNSW-NB15 dataset is another popular dataset used for developing and testing network-based intrusion detection systems. It contains modern attack scenarios and normal network traffic, offering nine different types of attacks, including Fuzzers, Analysis, Backdoor, DoS, Exploits, Generic, Reconnaissance, Shell code, and Worms. The dataset provides features such as flow-based statistics, content-based features, and time-based statistics, which are essential for creating comprehensive models for anomaly detection.

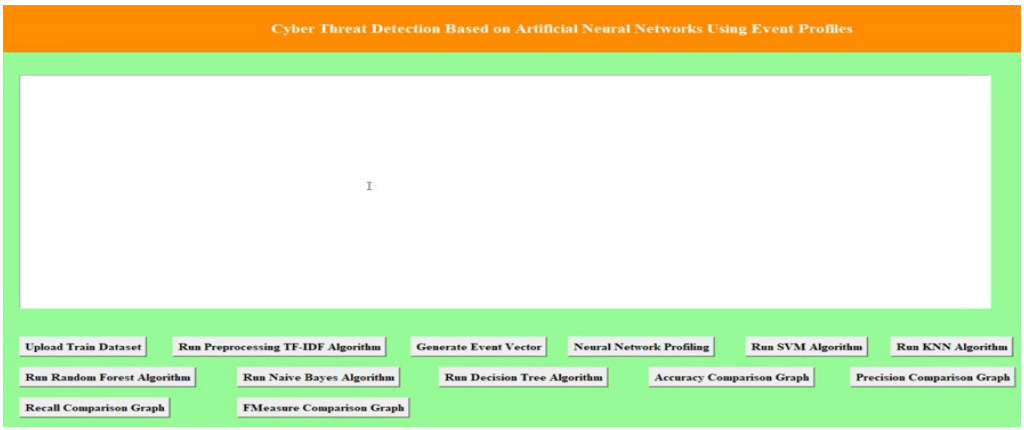
CICIDS2017 Dataset: The CICIDS2017 dataset contains realistic network traffic data, reflecting current cyber security threats. It includes different types of attacks, such as Brute Force, Heart bleed, Botnet, DoS, DDoS, Web attacks, and Infiltration, along with normal traffic data. Each type of attack is characterized by different behaviours, providing a robust dataset for testing anomaly-based detection systems. The dataset is particularly useful for evaluating deep learning models that require large volumes of data to learn complex attack patterns.

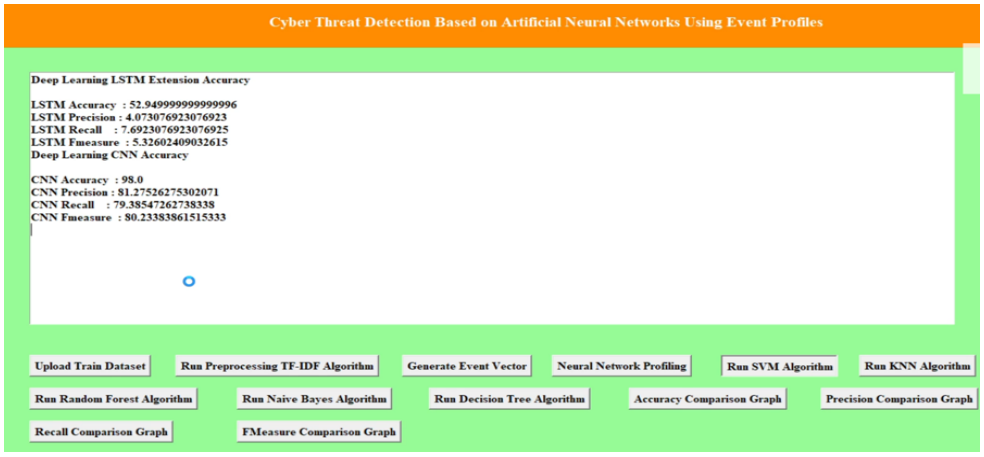
4. Analysis and Discussion

The findings of this study demonstrate the potential for AI-driven cybersecurity to reshape the evolving cyber threat landscape. As organizations use artificial intelligence and machine learning to strengthen their defenses against cyber threats, minimize the impact of cyber-attacks, and adopt a proactive cybersecurity stance, they will be able to implement a more robust cybersecurity policy. However, the dialogue also indicates some conceptual problems which need a comprehensive study. The primary theoretical effect of AI-assisted cyber threat detection is that it revolves the traditional security notions. Historically, cyber security was a defensive practice that mainly focused on the identification and neutralization of threats that

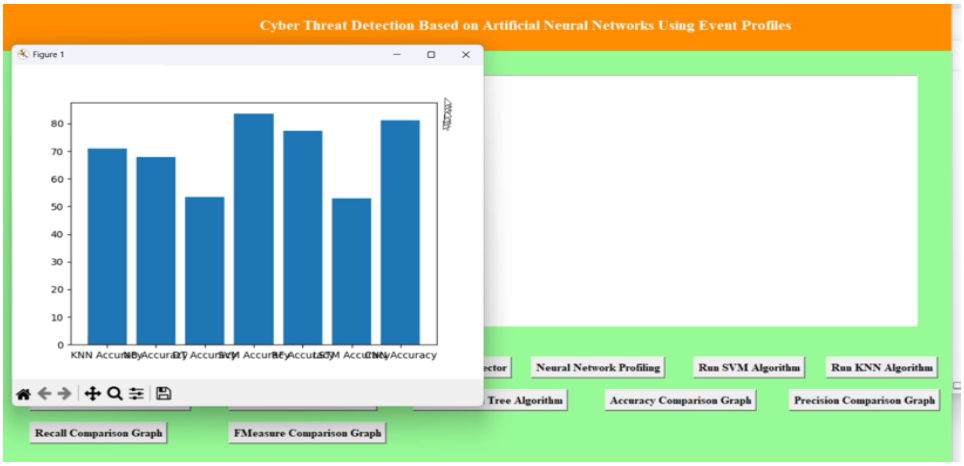
International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 13, Issue 2, 2025

had already made their way into the system. While AI detection systems have a reactive approach, which starts after a cyber-attack, they have a proactive approach by nevercasing monitoring and analyzing network activities. This helps in identifying and responding to threats in a real-time Notably, the said discussion conveys the importance of teamwork and information sharing to the point of use of AI for threat detection. The globalization of cybersecurity threats means that no security defense seems feasible without the interconnectedness of all entities. Shared decision-making via smart threat intelligence tools and defense strategies strengthens relations between organizations, which is then manifested by the exchange of resources, insights, and defense improvements among themselves. Such a collaborative approach, in addition to elevating the technical preparedness of individual entities, will also increase the level of cyber security of the entire ecosystem of digital networks and critical infrastructures. Moreover, the combination of AI and Security Orchestration, Automation, and Response (SOAR) systems raises the efficiency of incident response within business groups. Automation of simple tasks and orchestrating complex workflows drive SOAR platforms powered by AI technology to carry out incident response more quickly, minimize downtime, and reduce the incidence of human error in the incident response processes. Such an approach ensures both optimized functioning and higher resilience of organizations to cyber incidents. To add to this, SOAR platforms using AI will enable security teams to prioritize and efficiently sort the alerts, and thus, critical risks won't be shot down while the number of false positives is reduced. As cyber.





International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 13, Issue 2, 2025



CONCLUSION

Artificial Intelligence has emerged as a transformative force in the field of cyber security, offering advanced tools and techniques for threat detection and prevention. The ability of AI to analyse large volumes of data, identify complex patterns, and adapt to evolving threats has significantly enhanced the effectiveness of cyber security measures. However, several challenges remain, including data scarcity, high false-positive rates, and vulnerability to adversarial attacks. To fully realize the potential of AI in cyber security, continued research and innovation are essential. Addressing the limitations of current AI models, such as the need for high-quality labelled data and the risk of false positives, is crucial for improving their reliability and effectiveness. Additionally, exploring new research directions, such as hybrid models, Explainable AI, Federated Learning, and block chain integration, can lead to more robust and adaptable cyber security solutions. In conclusion, while AI has shown considerable promise in automating and enhancing threat detection and prevention, overcoming existing challenges and pursuing innovative approaches will be key to advancing the field and safeguarding against future cyber threats.

REFERENCES

1. Y. Xin, L. Kong, Z. Liu, Y. Chen, Y. Li, H. Zhu, M. Gao, H. Hou, and C. Wang, "Machine Learning and Deep Learning Methods for Cyber security," IEEE Access, vol. 6, pp. 35365-35381, 2018.
2. J. Gao, K. Zheng, Y. Chen, and L. Zhao, "Anomaly Detection of Network Traffic Based on Convolutional Long Short-Term Memory Neural Network," IEEE Access, vol. 6, pp. 12059-12072, 2018.
3. S. Nguyen, A. Marchal, R. State, and T. Engel, "Autonomous and Intelligent Defence System Using Reinforcement Learning," Proceedings of the 2018 IEEE International Conference on Communications (ICC), Kansas City, MO, USA, 2018.
4. G. K. Sood and S. Sharma, "Deep Learning Approaches for Malware Classification: A Comprehensive Review," IEEE Access, vol. 9, pp. 109677-109694, 2021.
5. Abdallah, M. B. Ahmad, and A. Muhammad, "Machine Learning Methods for Predicting Cyber Threats: A Review," Journal of Information Security and Applications, vol. 55, pp. 102615, 2020.

6. F. A. Afsar, M. A. Debbah, and W. Saad, "Generative Adversarial Networks for Threat Hunting in Cyber Security," IEEE Access, vol. 8, pp. 191929-191945, 2020.
7. S. Singh, Y. Jeong, J. Park, and J. H. Park, "A Survey on Machine Learning-Based Malware Detection in IoT Networks," Journal of Network and Computer Applications, vol. 95, pp. 128-145, 2017.