IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, July 2025
Impact Factor: 7.67

AI based Malware Detection

Prof. Parul Bhanarkar, Mansi Mahamuni, Snehal Jadhav
Professor, School of CSIT (Cybersecurity)
Student, School of CSIT (Cybersecurity)
Symbiosis Skills and Professional University, Kiwale, Pune

Abstract: *Malware remains one of the most prevalent and damaging forms of cyber threats today, compromising millions of devices globally. It can steal sensitive data, degrade system performance, encrypt critical files, and evade detection using advanced techniques. As malware continues to evolve rapidly, traditional detection methods often fall short, especially against previously unseen or zero-day variants. This growing challenge underscores the urgent need for intelligent, adaptive detection mechanisms. In response, this study proposes an advanced, AI-driven approach to malware detection that leverages dynamic deep learning and heuristic methods to identify and classify five prominent malware families: adware, Radware, rootkits, SMS malware, and ransomware. The paper reviews current detection technologies, highlights their limitations, and explores how artificial intelligence, particularly machine learning and deep learning, can enhance malware identification and response. Our analysis aims to guide future research in cybersecurity by advocating for self-learning, autonomous systems capable of handling real-time malware threats. By enhancing the resilience of digital infrastructures, AI-powered solutions can offer robust protection against the ever-growing sophistication of cyber-attacks, ensuring safer computing environments for users and organizations alike.*

Keywords: Malware Detection, Artificial Intelligence, Deep Learning, Heuristic Techniques, Cyber Security, Dynamic Analysis, Malware Classification.

I. INTRODUCTION

The evolution of computing, shaped by pioneers like Babbage and Turing, has led to its integration across all sectors. However, this technological growth has also increased the risk of cyber threats, particularly malware. Since the emergence of *Elk Cloner* in 1982 and *Brain* in 1986, malware has evolved from harmless programs to sophisticated tools for data theft, system disruption, and cybercrime.

High-profile attacks like *Stuxnet* and *CryptoLocker* have shown how damaging malware can be, especially as threats become more evasive and adaptive. In response, traditional security tools like antivirus software and firewalls were developed, but the rapid growth of malware—now with over 450,000 new variants detected daily—has made these tools less effective.

Modern malware uses advanced techniques to bypass detection, affecting various platforms including Windows, Linux, and Android. The rise in interconnected devices and cloud-based infrastructure has further expanded the threat landscape. Deep learning and AI-based methods offer a promising solution for detecting both known and unknown malware by learning patterns and behaviors from data.

This study proposes a dynamic deep learning-based model for autonomous malware detection, aiming to enhance cybersecurity by addressing the limitations of conventional methods and adapting to evolving threats.

Artificial Intelligence and Malware

In this section, we define Artificial Intelligence (AI) and Malware, and provide classifications introduced by various researchers to establish a clear understanding of both domains and their interrelation in cybersecurity.

A. Artificial Intelligence

Artificial Intelligence (AI) refers to the development of computer systems capable of performing tasks that traditionally required human intelligence. It is regarded as a transformative technology across multiple industries, offering efficiency

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, July 2025
Impact Factor: 7.67

gains, automation, and significant cost reductions. Nones et al. describe AI as a rapidly evolving field where machines are enabled to execute tasks that once solely depended on human cognitive abilities.

While AI is sometimes perceived as a replacement for human intelligence, many scholars view it as a tool for intelligence augmentation (IA). From this perspective, AI supports human decision-making by enhancing strategic thinking, problem-solving, and data analysis. This gives AI a strategic edge in modern technological revolutions and makes it a key driver in domains such as security, business intelligence, and big data analytics.

AI includes a broad spectrum of techniques and methodologies. One of the most prominent subsets is Machine Learning (ML), which allows computers to learn from data without explicit programming. ML operates by identifying patterns and adapting its behavior based on experience. It simulates human cognitive functions through logical operations and data-driven learning.

Machine Learning is further supported by technologies like Natural Language Processing (NLP), which enables computers to understand and generate human language. Together, these technologies form the foundation of modern AI applications in malware detection, where systems are designed to recognize malicious behavior patterns and respond autonomously.

The application of AI in cybersecurity—particularly in malware prevention—is increasingly critical. AI-based systems are capable of identifying, classifying, and mitigating malware threats by learning from both labeled and unlabeled data. This ability to adapt and evolve makes AI a powerful tool in combating new and sophisticated malware variants.

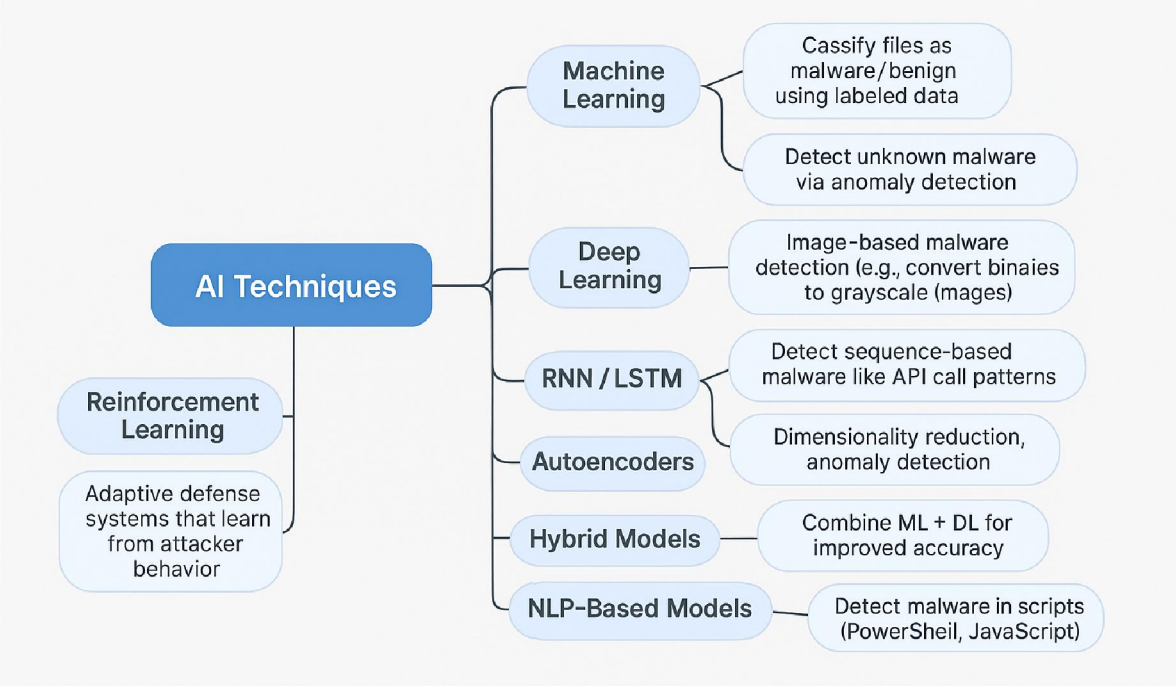

Fig. 1. Types and uses of AI in Malware Detection

B. Malware

Malware, a contraction of "malicious software," refers to any software or code intentionally designed to cause harm to computer systems, networks, or users. It includes a wide range of threats such as computer viruses, worms, trojans, ransomware, spyware, adware, rootkits, keyloggers, and malicious browser extensions (BHOs).

Unlike early software threats that were mostly experimental or non-destructive, modern malware is designed with harmful intent. It aims to disrupt system operations, gain unauthorized access, steal sensitive information, or compromise the availability and integrity of data. Malware can affect not only individual systems but also large-scale networks, mobile devices, and cloud platforms.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, July 2025
Impact Factor: 7.67

The growth of internet connectivity has significantly contributed to the rise in malware activity. From just four hosts in 1969, the number of internet-connected devices surpassed one billion by 2019. This vast digital ecosystem offers attackers numerous points of entry. Malware typically spreads through infected files, malicious websites, email attachments, software downloads, or embedded links. Once executed, a malware program can replicate, install itself silently, and begin executing its malicious payload—often without the user's awareness.

Malware is generally classified based on its purpose and propagation method. For example, trojans disguise themselves as legitimate software, while worms spread independently across networks. Ransomware encrypts user data and demands payment, whereas spyware secretly gathers user information. This classification helps researchers and developers design appropriate detection and prevention strategies.

Figure 2 typically illustrates malware taxonomy, which includes categories based on attack behavior, infection techniques, and target systems. Understanding these classifications is essential in designing robust AI-driven malware detection systems capable of handling both known and emerging threats.

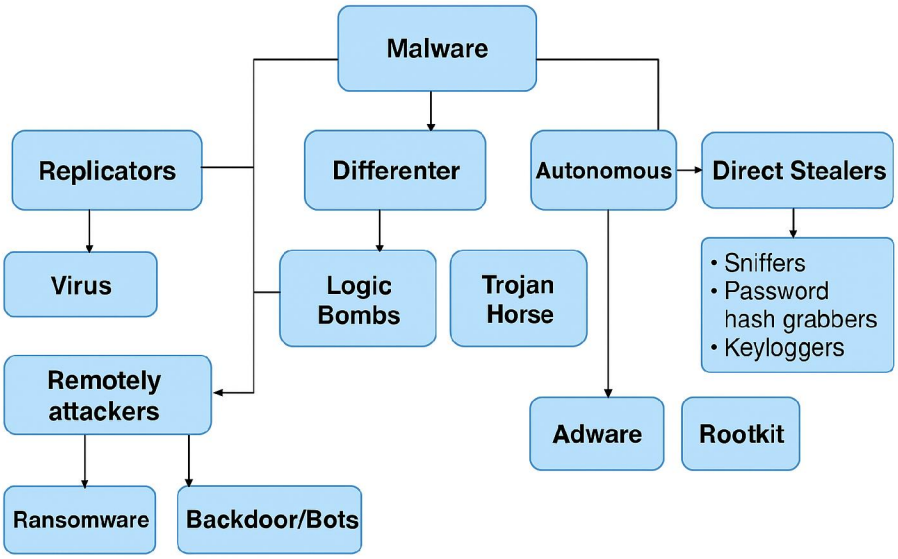

Fig. 2. Malware Classification

Datasets

This section presents two of the most recent and widely adopted malware detection datasets that are suitable for training and evaluating Artificial Intelligence models, particularly in the context of static and dynamic malware analysis. These datasets provide comprehensive and diverse features, making them ideal for modern AI-based malware detection research.

Feature	EMBER 2024	CIC-DGG-2025
Source	Open-source (GitHub, arXiv; FutureComputing4AI/Elastic)	Open-source (Canadian Institute for Cybersecurity – UNB)
Total Samples	3,238,315 (incl. 6,315 evasive)	Large (aggregated from BODMAS, PE-Machine-Learning, etc.)
Sample Types	PE (Win32, Win64, .NET), APK, PDF, ELF	Control-Flow Graphs (CFGs) extracted from Windows, Linux, Android binaries
Labeling	≥ 5 AV detections = malicious; otherwise	Binary labels (malicious/benign) as per

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, July 2025
Impact Factor: 7.67

	benign	source datasets
Feature Types	Byte histograms, PE headers, Rich header, digital signatures, file metadata	Dynamic CFGs, assembly-level graph embeddings
Behavioral Info	Yes (multi-labels: family, packer, vulnerability)	Yes (via dynamic analysis and explainable graph models)
Graph Support	static features only	node/edge details + graph embeddings
Preprocessing	• Time-split (52 wk train / 12 wk test)	
• Normalized feature vectors v3	• Graph extraction via angr	
Train/Test Split	Pre-split (52 weeks train, 12 weeks test)	CSV labels provided; split by user per experiment
Format	CSV + JSON feature vectors	JSON graphs + CSV metadata
Ideal Use Cases	Static detection, large-scale ML experiments, concept-drift research	Dynamic detection, explainable AI, Graph Neural Network training

Table I. Comparison of Latest AI Malware Detection Datasets (till 2025)

II. DATA ANALYSIS AND EXPLORATION

Data analysis and exploration are critical stages in the development of an effective malware detection model. These steps involve examining the dataset to understand its structure, distributions, anomalies, and correlations among features. Exploratory Data Analysis (EDA) helps identify patterns, inconsistencies, and outliers that could influence the modeling strategy. During this phase, visual and statistical methods were applied to assess the balance between malware and benign samples, evaluate feature importance, and ensure that the dataset was suitable for AI-based learning. This process laid the foundation for defining both binary and multi-class classification tasks.

Estimated Market Growth of AI in Cybersecurity for 2025

The global market for artificial intelligence (AI) in cybersecurity is projected to experience substantial growth in 2025, driven by the increasing volume and sophistication of malware attacks. According to various industry reports, the estimated market value ranges between USD 30.79 billion and USD 36.54 billion. Grand View Research estimates the market will reach USD 31.48 billion, while Polaris Market Research projects a similar value of USD 31.38 billion. The Business Research Company provides a slightly lower estimate at USD 30.79 billion, whereas Global Growth Insights offers the most optimistic outlook with a forecast of USD 36.54 billion. These figures underscore the critical importance of AI-powered security solutions in modern cybersecurity frameworks. The increasing reliance on AI for real-time malware detection, behavioral analytics, and automated threat response reflects a broader trend in the cybersecurity landscape—organizations are prioritizing intelligent systems to combat zero-day threats and reduce false positives. This market expansion highlights how AI is not only enhancing threat detection accuracy but also becoming a cornerstone of proactive cybersecurity strategies in both enterprise and governmental infrastructures.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, July 2025
Impact Factor: 7.67

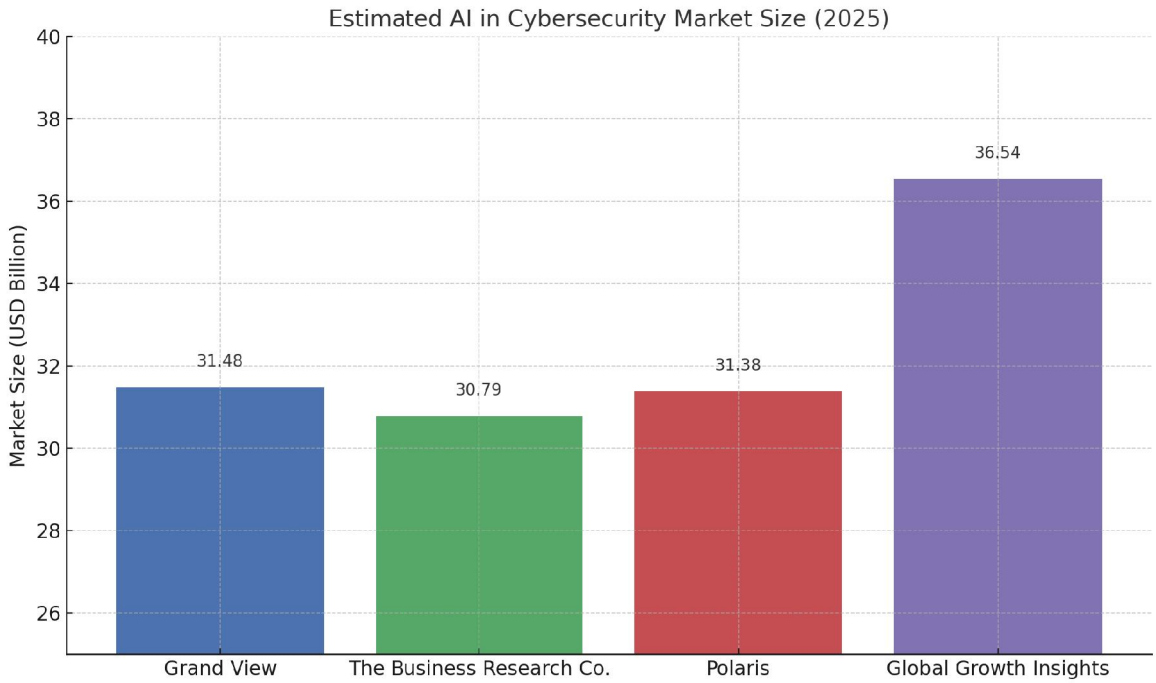

FIG 3 Estimated AI in Cybersecurity Market Size (2025)

Data Preparation

The initial dataset used in this study was collected from multiple sources in unstructured or raw formats. These raw datasets, while rich in content, were not directly usable for machine learning algorithms. Thus, a structured format was developed to support model training and evaluation. As part of the data preparation, labels were consolidated depending on the classification task. For **binary classification**, all malware types were grouped under a single "malware" label, while non-malicious samples were labeled as "benign." For **multi-class classification**, each malware family retained its specific label to enable granular classification. This restructuring facilitated flexible experimentation across different model architectures.

Data Preprocessing

The quality of input data directly affects the performance of AI models. Inaccurate, noisy, or incomplete data can lead to poor generalization and high false-positive rates. Therefore, thorough preprocessing was performed to ensure data quality and consistency before model training. This involved multiple key steps:

- **Data Cleaning**: Removal of duplicates, null values, and irrelevant features.
- **Filtering**: Elimination of noise or inconsistencies that could mislead the learning process.
- **Normalization**: Scaling numerical features to uniform ranges to improve model convergence.
- **Feature Transformation and Encoding**: Transformation of categorical variables and complex structures into machine-readable formats.
- **Feature Extraction**: Identification of relevant indicators that contribute to malware behavior classification.

This preprocessing pipeline was particularly designed to support deep learning architectures, enabling the model to learn complex feature representations from high-dimensional input data.

Step	EMBER 2024 (Static Features)	CIC-DGG-2025 (Graph-Based Features)
Data Cleaning	Remove duplicates, unreadable files	Remove failed or incomplete graph samples

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, July 2025
Impact Factor: 7.67

Labeling	Based on ≥5 AV detections (malicious/benign)	Based on source dataset binary labels
Feature Extraction	Extract byte histograms, PE metadata	Extract control-flow graphs, graph embeddings
Normalization	Apply Min-Max/Z-score to numerical features	Normalize graph-level statistics
Encoding	Encode categorical metadata (e.g., OS, file type)	Embed node/edge attributes
Train/Test Split	52-week train / 12-week test split	User-defined split (e.g., time-based or random)
Output Format	CSV and JSON feature vectors	JSON graphs + CSV metadata

Table II. Essential Preprocessing Steps for EMBER 2024 and CIC-DGG-2025

Malware Detection Using AI

This section presents Artificial Intelligence (AI)-based techniques for malware detection, reviews current detection strategies, highlights their limitations, and discusses how AI can enhance detection capabilities.

A. Malware Detection Techniques

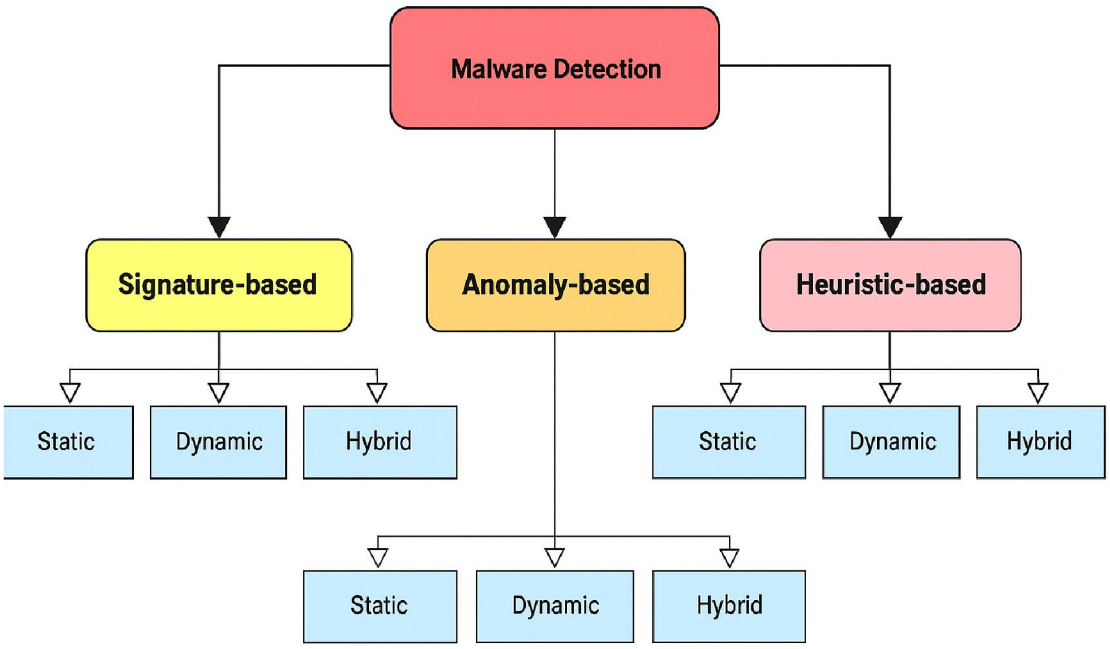

Figure 4 : Malware Detection techniques

Malware detection systems are designed to distinguish between malicious and benign applications by analyzing behaviors and patterns. These techniques are generally categorized into three major types:

- Signature-based Detection
- Anomaly-based Detection
- Heuristic-based Detection

Each method contributes uniquely to identifying threats, and their effectiveness varies depending on the type and novelty of malware.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, July 2025
Impact Factor: 7.67

1) Signature-based Detection

This is the most traditional technique, which involves comparing a file's content or behavior against a database of known malware signatures. If a match is found, the file is flagged as malicious. Signature-based systems are efficient for detecting known threats and are widely used in antivirus programs. However, their major limitation is the inability to detect previously unseen or modified malware (zero-day attacks).

A typical signature-based detection system consists of four core components: a malware database, a scanner, a matching engine, and an alerting mechanism. Additionally, Intrusion Detection Systems (IDS) can maintain statistical models of known traffic and flag traffic that deviates from expected patterns.

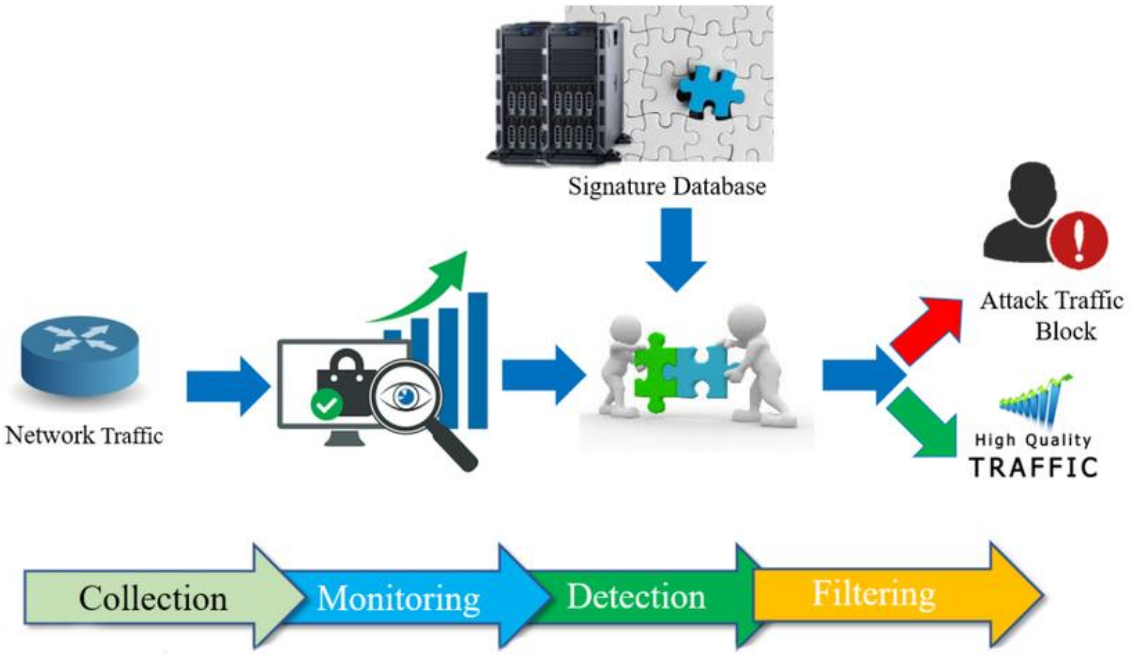

Fig 5 Methodology used in Signature based Intrusion Detection System (IDS) [42]

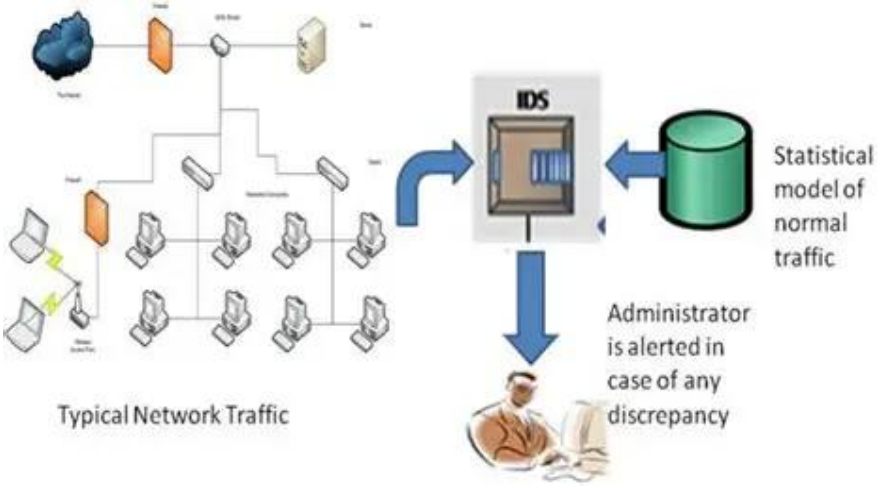

Fig 6 Signature based Intrusion Detection System (IDS) [38]

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, July 2025
Impact Factor: 7.67

2) Anomaly-based Detection

Anomaly-based detection addresses the limitations of signature methods by identifying deviations from normal behavior rather than relying solely on known patterns. These systems analyze system activities or network traffic and flag anomalies that may indicate malicious behavior.

This approach often uses classification algorithms from machine learning to distinguish between normal and suspicious activities. It is particularly effective at detecting unknown and evolving malware but may suffer from higher false-positive rates if not properly trained or tuned.

Anomaly-based Network Intrusion Detection Systems (ANIDS) operate in multiple stages: data collection, preprocessing, behavior modeling, and alerting. If a system's behavior significantly deviates from its learned profile, it is flagged for further analysis.

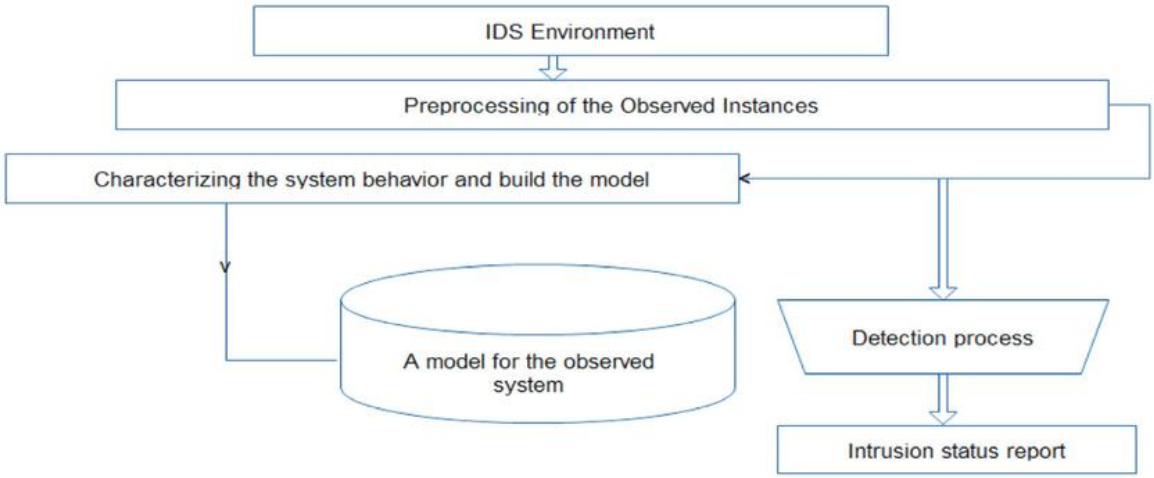

Fig 7 common anomaly based network IDS [43]

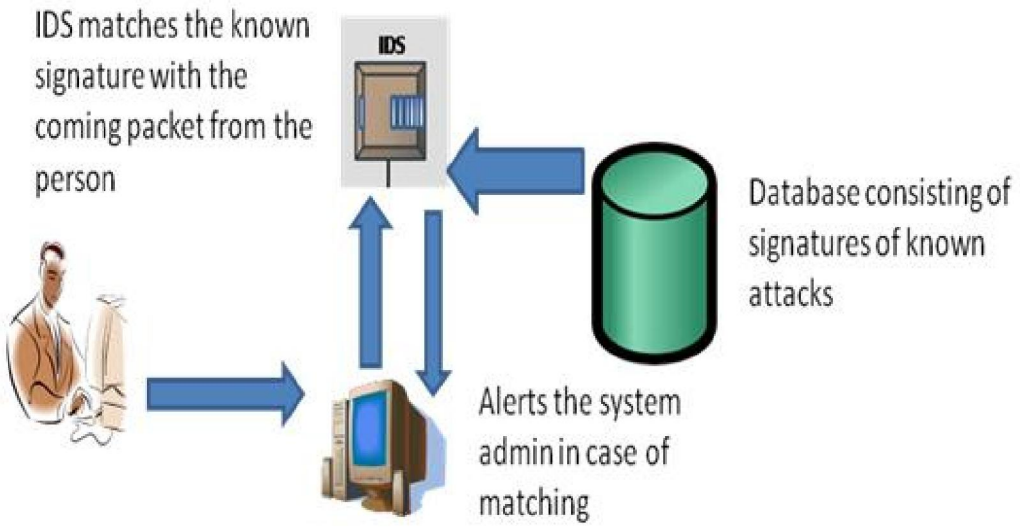

FIG 8 anomaly based IDS [38]

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, July 2025
Impact Factor: 7.67

3) Heuristic-based Detection

Heuristic techniques blend features from both signature and anomaly-based methods, using AI to improve detection accuracy. These systems rely on rule-based logic and probabilistic models to identify suspicious activity, even if it doesn't exactly match known malware patterns.

Machine learning algorithms such as **genetic algorithms** and **neural networks** are often employed in heuristic systems to enhance learning and adaptability. Genetic algorithms, for example, use concepts like inheritance, mutation, and crossover to iteratively improve detection performance.

Heuristic detection offers improved flexibility and predictive capability, especially in dynamic environments where malware evolves quickly. It enables systems to detect novel threats based on behavioral traits rather than known signatures.

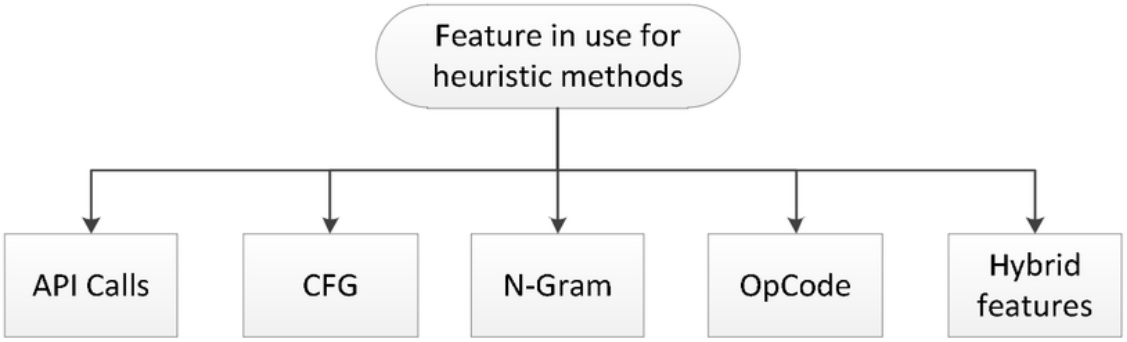

Fig 9 Heuristics Methods Features [46]

Flow of AI-Based Malware Detection

AI-based malware detection systems generally follow a structured flow:

Data Collection: Gathering datasets containing malware and benign samples (e.g., from Kaggle or open-source repositories).

Feature Selection: Techniques like Fisher Score (FS), Chi-Square (CS), Information Gain (IG), Gain Ratio (GR), and Uncertainty Symmetric (US) are used to select the most relevant features.

Classifier Training: Various classifiers are trained and compared based on selected features.

Detection and Evaluation: The trained model is used to detect unknown malware and continuously updated for improved accuracy.

This AI-enhanced pipeline offers a scalable and adaptable approach to malware detection, capable of addressing both known and unknown threats effectively.

III. DISCUSSION

Malware detection approaches have advanced considerably in recent years, with notable improvements in accuracy, adaptability, and responsiveness. Traditional **signature-based techniques** are still relevant, particularly for identifying known malware efficiently. Their effectiveness is largely sustained through continuous updates to signature databases, allowing security tools to respond promptly to familiar threats.

To enhance detection of unknown malware, **generic signature scanning** has been introduced. Unlike conventional methods that rely on exact matches, generic scanning identifies broader code patterns or behaviors, thereby improving the system's ability to detect newly emerging threats that share structural similarities with known malware.

Another key advancement is the use of **heuristic analysis**, which functions in two main modes. The **static approach** assesses code features and structure prior to execution, flagging potentially harmful characteristics. In contrast, **dynamic heuristic analysis** observes how software behaves at runtime, allowing for the identification of suspicious or

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, July 2025
Impact Factor: 7.67

abnormal actions during execution. These complementary techniques enable systems to detect threats based on both code design and behavior, enhancing overall detection coverage.

Integrity verification is also an important element in modern detection systems. This method checks for unauthorized changes by comparing current system or file states with trusted baselines. It adds an extra layer of assurance by confirming the authenticity of critical files and ensuring that no unauthorized modifications have occurred.

The increasing complexity and sophistication of malware have driven the adoption of **Artificial Intelligence (AI)** in detection systems. Techniques such as **machine learning** and **deep learning** allow security systems to learn from large volumes of data, identify complex patterns, and adapt to new and evolving attack methods. AI-powered detection offers benefits like automatic feature selection, real-time analysis, and the ability to detect previously unknown or evasive malware with high accuracy.

Overall, the integration of traditional detection mechanisms with AI-driven technologies results in a more robust and intelligent defense strategy. Continued innovation in this space—particularly in real-time AI implementation and adaptive learning—will be essential for developing next-generation malware detection solutions capable of countering sophisticated cyber threats.

IV. CONCLUSION

Malware poses a severe threat not only to individual computer systems but also to large-scale infrastructures such as data centers, web platforms, and mobile applications. Its impact is particularly critical in sensitive domains like finance and healthcare, where data integrity and privacy are paramount. Safeguarding digital environments against such malicious threats has become a pressing challenge, leading to the continued evolution of **malware detection and prevention systems**.

This study has provided a comprehensive overview of the current landscape of malware detection techniques, with a particular focus on the integration of **Artificial Intelligence (AI)**. AI has emerged as a powerful enabler in building intelligent, adaptive anti-malware systems capable of detecting sophisticated and previously unseen threats with greater accuracy and efficiency.

We began by outlining the foundational concepts of malware and AI, followed by an in-depth discussion on existing detection mechanisms. Limitations of traditional and contemporary techniques were critically examined, emphasizing the need for more scalable, real-time, and intelligent solutions. While many current systems offer improved detection capabilities, they still face challenges such as false positives, model generalization, and adaptability to evolving malware tactics.

Our analysis suggests that AI, particularly through **machine learning and deep learning**, offers a promising direction for enhancing malware detection frameworks. These technologies enable systems to learn from vast datasets, identify complex patterns, and adapt dynamically to new threats.

In conclusion, while notable progress has been made, further research is needed to overcome existing limitations and fully harness the potential of AI in this domain. Continued innovation, real-time adaptability, and hybrid detection models will be key to building next-generation malware detection systems capable of securing future digital ecosystems.

V. FUTURE SCOPE

The future of malware detection will increasingly rely on advanced **AI integration and automation**. As malware grows more sophisticated, techniques such as **deep learning**, **graph-based analysis**, and **real-time behavioral modeling** will be vital. Future research will focus on improving **adversarial robustness**, detecting **zero-day threats**, and building **explainable AI (XAI)** models to enhance system transparency and reliability. Additionally, **cloud-based threat intelligence** and **collaborative detection networks** will enable faster, global response to emerging threats. These advancements will lead to malware detection systems that are more **accurate**, **adaptive**, and **scalable**.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, July 2025
Impact Factor: 7.67

REFERENCES

[1] O. Asaolu, "On the emergence of new computer technologies." Educational Technology Society, vol. 9, pp. 335–343, 01 2006.
[2] Z. Arsic and B. Milovanovic, "Importance of computer technology in realization of cultural and educational tasks of preschool institutions," International Journal of Cognitive Research in Science, Engineering and Education, vol. 4, pp. 9–15, 06 2016.
[3] A. P. Gilakjani, "A detailed analysis over some important issues towards using computer technology into the EFL classrooms," Universal Journal of Educational Research, vol. 2, pp. 146–153, 2014.
[4] H. F. Md Jobair, M. Paul, C. Ryan, S. Hossain, and C. Victor, "Smart connected aircraft: Towards security, privacy, and ethical hacking," International Conference on Security of Information and Networks, 2022.
[5] S. Subramanya and N. Lakshminarasimhan, "Computer viruses," Potentials, IEEE, vol. 20, pp. 16–19, 11 2001.
[6] S. Levy and J. Crandall, "The program with a personality: Analysis of Elk Cloner, the first personal computer virus," 07 2020.
[7] N. Milosevic, "History of malware," 02 2013.
[8] A. P. Namanya, A. Cullen, I. Awan, and J. Pagna Diss, "The world of malware: An overview," 09 2018.
[9] I. Khan, "An introduction to computer viruses: Problems and solutions," Library Hi Tech News, vol. 29, pp. 8–12, 09 2012.
[10] M. Bishop, "An overview of computer viruses in a research environment," USA, Tech. Rep., 1991.
[11] D. B. Patil and M. Joshi, "A study of past, present computer virus performance of selected security tools," Southern Economist, 12 2012.
[12] A. Terekhov. History of the antivirus. [Online]. Available: https://www.hotspotshield.com/blog/history-of-the-antivirus
[13] M. J. Hossain Faruk, H. Shahriar, M. Valero, S. Sneha, S. Ahamed, and M. Rahman, "Towards blockchain-based secure data management for remote patient monitoring," IEEE International Conference on Digital Health (ICDH), 2021.
[14] M. J. Hossain Faruk, "EHR data management: Hyperledger Fabric-based health data storing and sharing," The Fall 2021 Symposium of Student Scholars, 2021.
[15] S. Ryan, R. Mohammad A, H. F. Md Jobair, S. Hossain, and C. Alfredo, "Ride-hailing for autonomous vehicles: Hyperledger Fabric-based secure and decentralize blockchain platform," IEEE International Conference on Big Data, 2021.
[16] D. G. Vigna. (2020) How AI will help in the fight against malware. [Online]. Available: https://techbeacon.com/security/how-ai-will-help-fight-against-malware
[17] H. Hassani, E. Silva, S. Unger, M. Tajmazinani, and S. MacFeely, "Artificial intelligence (AI) or intelligence augmentation (IA): What is the future?" AI, vol. 1, p. 1211, 04 2020.
[18] A. I. Nones, A. Palepu, and M. Wallace. (2019) Artificial intelligence (AI). [Online]. Available: cisse.info/pdf/2019/RR-01-artificial-intelligence.pdf
[19] (2020) Artificial intelligence - reasoning. [Online]. Available: britannica.com/technology/artificial-intelligence/Evolutionary-computing
[20] S. Ahn, S. V. Couture, A. Cuzzocrea, K. Dam, G. M. Grasso, C. K. Leung, K. L. McCormick, and B. H. Wodi, "A fuzzy logic based machine learning tool for supporting big data business analytics in complex artificial intelligence environments," in 2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), 2019, pp. 1–6.
[21] A. Cranage. (2019) Getting smart about artificial intelligence. [Online]. Available: https://sangerinstitute.blog/2019/03/04/getting-smart-about-artificial-intelligence
[22] J. Alzubi, A. Nayyar, and A. Kumar, "Machine learning from theory to algorithms: An overview," Journal of Physics: Conference Series, vol. 1142, p. 012012, 11 2018.
[23] T. Ayodele, Machine Learning Overview, 02 2010.
[24] M. Ahmad, "Malware in computer systems: Problems and solutions," IJID (International Journal on Informatics for Development), vol. 9, p. 1, 04 2020.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 1, July 2025 Impact Factor: 7.67

[25] N. Milosevic, "History of malware," Digital Forensics Magazine, vol. 1, no. 16, pp. 58–66, Aug. 2013.
[26] S. Gupta, "Types of malware and its analysis," International Journal of Scientific Engineering Research, vol. 4, 2013. [Online]. Available: https://www.ijser.org/researchpaper/Types-of-Malware-and-its-Analysis.pdf
[27] Statista. Number of worldwide internet hosts in the domain name system (DNS) from 1993 to 2019. [Online]. Available: https://www.statista.com/statistics/264473/number-of-internet-hosts-in-the-domain-name-system/
[28] S. Poudyal, D. Dasgupta, Z. Akhtar, and K. D. Gupta, "Malware analytics: Review of data mining, machine learning and big data perspectives," 12 2019.
[29] O. Adebayo, M. A., A. Mishra, and O. Osho, "Malware detection, supportive software agents and its classification schemes," International Journal of Network Security Its Applications, vol. 4, pp. 33–49, 11 2012.
[30] A. K. S., "Impact of malware in modern society," Journal of Scientific Research and Development, vol. 2, pp. 593–600, 06 2019.
[31] B. A. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering," Keele University and Durham University Joint Report, Tech. Rep. EBSE 2007-001, 07 2007. [Online]. Available: https://www.bibsonomy.org/bibtex/23f4b30c0fe1435b642467af4cca120ef
[32] C. C. Agbo, Q. H. Mahmoud, and J. M. Eklund, "Blockchain technology in healthcare: A systematic review," Healthcare, vol. 7, no. 2, 2019. [Online]. Available: https://www.mdpi.com/2227-9032/7/2/56
[33] M. O. F. Rokon, R. Islam, A. Darki, E. Papalexakis, and M. Faloutsos, "SourceFinder: Finding malware source-code from publicly available repositories," in RAID, 2020.
[34] N. Sharma and B. Arora, "Data mining and machine learning techniques for malware detection," in *Rising Threats in Expert Applications and Solutions*, Springer Singapore, 2021, pp. 557–567.
[35] S. Sharma, R. Challa, and S. Sahay, *Detection of Advanced Malware by Machine Learning Techniques: Proceedings of SoCTA 2017*, 01 2019, pp. 333–342.
[36] S. Saad, W. Briguglio, and H. Elmiligi, "The curious case of machine learning in malware detection," 2019.
[37] I. Baptista, S. Shiaeles, and N. Kolokotronis, "A novel malware detection system based on machine learning and binary visualization," 05 2019, pp. 1–6.
[38] S. A. Repalle and V. R. Kolluru, "Intrusion detection system using AI and machine learning algorithm," 12 2017.
[39] M. Mohammad, S. Hossain, H. Hisham, H. F. Md Jobair, V. Maria, K. Md Abdullah, A. R. Mohammad, A. Muhaiminul I., C. Alfredo, and W. Fan, "Bayesian hyperparameter optimization for deep neural network-based network intrusion detection," IEEE International Conference on Big Data, 2021.
[40] O. C. Onyedeke, E. Taoufik, M. Okoronkwo, U. Ihedioha, C. H. Ugwuishiwu, and O. B., "Signature based network intrusion detection system using feature selection on Android," International Journal of Advanced Computer Science and Applications, vol. 11, 01 2020, p. 120.
[41] Y. Ye, T. Li, Q. Jiang, Z. Han, and L. Wan, "Intelligent file scoring system for malware detection from the gray list," 01 2009, pp. 1385–1394.
[42] S. Jyoti, A. Bhandari, V. Baggan, M. Snehi, and Ritu, "Diverse methods for signature based intrusion detection schemes adopted," 07 2020.
[43] J. Veeramreddy and K. Prasad, *Anomaly-Based Intrusion Detection System*, 06 2019.
[44] D. Bolzoni and S. Etalle, "Aphrodite: An anomaly-based architecture for false positive reduction," ArXiv, vol. abs/cs/0604026, 2006.
[45] S. Bridges, R. Vaughn, and A. Professor, "Fuzzy data mining and genetic algorithms applied to intrusion detection," 04 2002.
[46] Z. Bazrafshan, H. Hashemi, S. M. Hazrati Fard, and A. Hamzeh, "A survey on heuristic malware detection techniques," 05 2013, pp. 113.
[47] I. Baptista, S. Shiaeles, and N. Kolokotronis, "A novel malware detection system based on machine learning and binary visualization," in *2019 IEEE International Conference on Communications Workshops (ICC Workshops)*, 2019, pp. 1–6.
[48] T. Garfinkel and M. Rosenblum, "A virtual machine introspection based architecture for intrusion detection," NDSS, vol. 3, 05 2003.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, July 2025
Impact Factor: 7.67

[49] S. Li, Q. Zhou, R. Zhou, and Q. Lv, "Intelligent malware detection based on graph convolutional network," The Journal of Supercomputing, 08 2021.
[50] L. Wen and H. Yu, "An Android malware detection system based on machine learning," vol. 1864, 08 2017, p. 020136.