Journal of Informatics and Web Engineering
Vol. 5 No. 1 (February 2026) eISSN: 2821-370X

AI-Driven Malware Analysis and Detection: A Comprehensive Survey of Techniques, Trends and Challenges

Salman Khan[1*], Hasnat Raza[2] and Mansoor Alam[3]
[1,2,3]Facaulty of Computing, Riphah International University, I-14/3 I-14, Islamabad, 46000, Pakistan
**corresponding author: (24730@students.riphah.edu.pk; ORCID: 0009-0005-2926-0804)*

Abstract - Malware represents the most critical threat in cybersecurity, meant to compromise the security for any individual or any organization. These are covert software, designed to perform malicious act like data theft, data alteration, and to interrupt a normal operation of the services. The persistent evolution of malware has called for more sophisticated techniques in its detection and prevention, resulted into direct need of Artificial Intelligence in cybersecurity. Artificial intelligence, using machine learning techniques and rising concepts like neural networks has greatly improved the traditional static and dynamic ways of detecting malware. Advances in AI-driven solutions have made them much more capable than their predecessors of detecting malware and addressing threats in real time. By training machine learning models on vast quantities of data, malicious patterns can easily be detected and identify patterns. With these emerging challenges, AI powers automated real-time analysis and adaptive security posture can effectively mitigate the threat. Large Language Models (LLMs) have revolutionized natural language processing and are increasingly being deployed across a wide range of applications, including text generation, summarization, translation, and detection systems. Recent research related to the methodologies employed in developing detection systems using LLMs, outlines the existing limitations and research gaps, and proposes potential areas for future investigation. The use of AI in malware analysis faces its own challenges with the potential for adversarial attacks and the scale of AI models that can muddy the waters of transparency and trust. Overcoming these challenges will involve the creation of mature, ethical, AI systems and an open dialogue between cybersecurity professionals, sustainable AI development and regulatory compliance all working in concert.

Keywords—Adaptive Security, Adversarial Attacks, Artificial Intelligence, Cybersecurity, Dynamic Analysis, Malware, Machine Learning, Polymorphic Malware, Static Analysis, Large Language Models.

Received: 8 June 2025; Accepted: 12 August 2025; Published: 16 February 2026

1. INTRODUCTION

Malware is one of the most common and devastating threats in the cybersecurity universe. It stands as one of the largest problems at the individual, and corporate levels across worldwide. Malware, in short for malicious software, encompasses a multitude of software archetypes designed to perform unauthorized actions on the infected system [1]. Illustrations of these actions include the theft of information, data corruption, and interfering with service

availability even up to consuming the resources of the system. Malware has made huge Impact on cybersecurity in the form of direct cost - based on the stolen goods, or a ransom. In practice, it can disable essential services, halt business operations, or even subvert data integrity. In terms of a social consequence, the leaked personal information removes trust from the society on digital services [2].

It is imperative to analyze and mitigate the malware that enables best practices and measures to preserve online security thereby maintaining system health. Successful malware management goes beyond just detecting and removing acting malicious software, it also heavily involves taking preventative measures to protect future attacks. Common malware analysis techniques involve static analysis where malware is examined without executing the code and dynamic analysis where malware is executed in a virtual environment and behavior is observed. However, those ways have difficulty matching the pace at which malware techniques are evolving to become more sophisticated like polymorphic malware, that changes its appearance with every new attack, and metamorphic malware, that actually modify its code to continue evading detection [3].

The present study is a comprehensive survey of malware analysis and detection using Artificial Intelligence (AI), especially focusing on Machine Learning (ML) and LLMs. It reviews static and dynamic malware analysis techniques enhanced by AI, explores emerging trends like federated learning, and discusses challenges such as adversarial attacks, model transparency, and resource constraints. The paper highlights AI's transformative potential in cybersecurity while stressing the need for ethical, robust, and interpretable solutions. The literature search for this survey was conducted using various reputable sources and search engines, as summarized in Table 1.

Table 1. Overview of Literature Sources and Publication Counts

Ser	Search Engines/ Sources	Count/ Paper included
1	IEEE	24
2	Springer	10
3	Elsevier	3
4	Cybersecurity	2
5	ACM	4
6	arXiy/ World Scientific/ Appl.Sci	6
7	PLoS ONE/ ICT Express/ World scientific	3
8	Ubiquity/ Symmetry/ Front.Eng.Manag	3
9	Miscellaneous International Journals	17
10	Future Gener. Comp Sys	3
11	Miscellaneous Online Sources	14

In these struggles, a battle has emerged to boost malware analysis and mitigation with the utility of AI. AI is playing an even greater role in the dynamics of malware. It has the capacity to learn from massive sets of data and looking for patterns that can be missed by human analysts. AI in the form of ML models can be trained to detect malicious behaviours and benign behaviours of malware and normal programs respectively. With this training, a myriad of features derived from the software have been taken such as opcode sequences, API calls, and network traffic patterns [4], [5].

Additionally, AI approaches can automate real-time analysis of malware, providing ongoing cybersecurity coverage and rapid adaption to changing malware tactics, without the need for constant manual update of malware statement. This is more important than ever as malware attacks can multiply and evolve faster than ever. AI-driven tools help cybersecurity systems to respond quickly and also proactively identify potential vulnerabilities, making cybersecurity defense to be proactive rather than reactive [6].

Among the various approaches discussed in this study, Figure 1 provides a visual representation of the most common malware analysis techniques, highlighting the methodological frameworks that underpin both static and dynamic examination of malicious software. Figure 1 is a mind map illustrating different types of malwares, including Spyware, Ransomware, Trojans, Bots, and Mobile Malware. Each category branches into specific subtypes or examples, such as Key logger under Spyware and WannaCry under Ransomware. It visually organizes malware classifications and notable variants to show their relationships and effects.

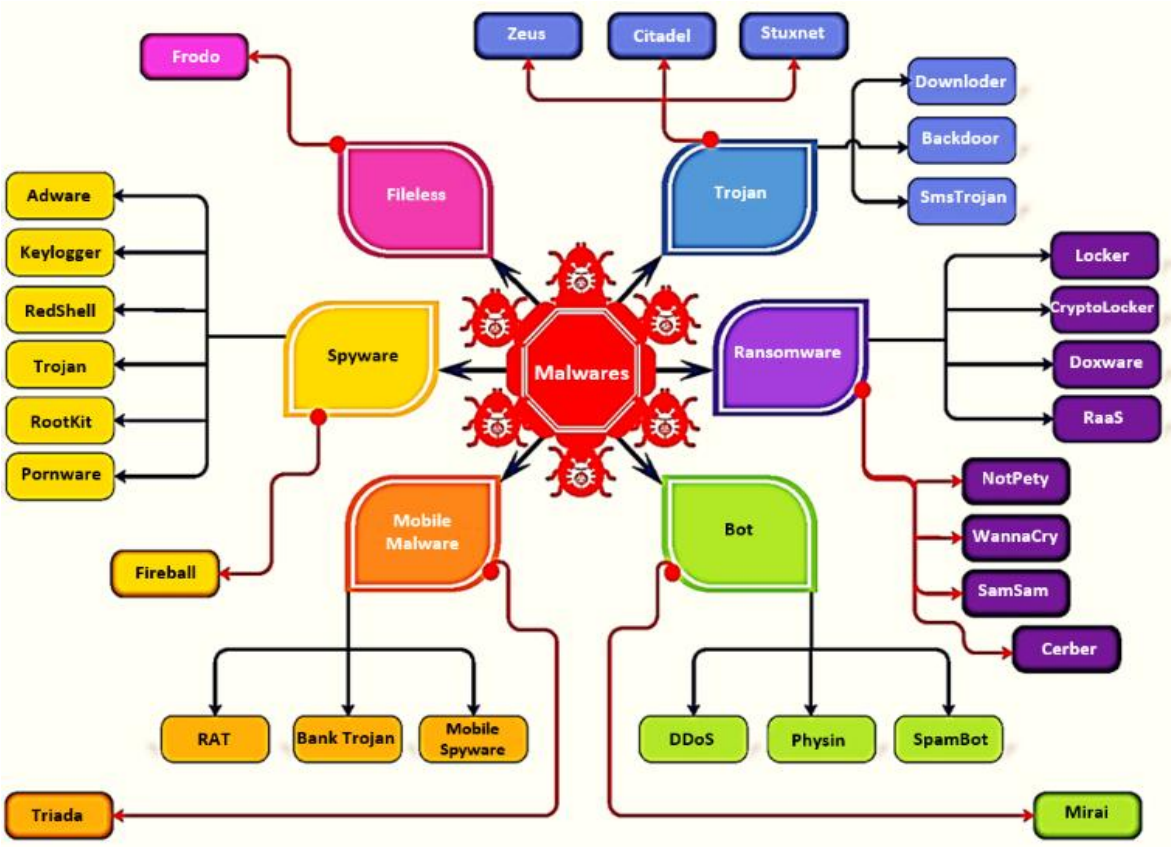

Figure 1. Common Malware Analysis Techniques

The use of AI in malware analysis and mitigation is a significant step forward in the field of cybersecurity integration. AI further helps by improving the fidelity, agility, and predictive capabilities of malware detection systems, arming cybersecurity professionals to get one step ahead of rapidly changing threats in the digital world. The constant advancements in AI are expected not only neutralizing the existing vulnerabilities but also usher in a whole new era of cybersecurity methodologies to come. It is crucial that these AI systems are designed with robustness and ethical considerations in mind as we can only move forward from here and we must make the most out of the efficacy of these AI and ensure they do not lead to new vulnerabilities inadvertently [7].

These advanced protection systems are a beacon of light in the quiet battle against the growing evolution of cyber threats while the technologies of AI itself mature further and further to advance in this fight hackers. AI can process large data sets faster than any human can and can detect problems and anything out of the norm on a scale and speed that otherwise would be impossible with only human analysts. The support for different file types and formats is very important to combat new malware around in the wild that changes and even evolves to get past traditional means of detection. Since they are powered by ML and neural networks, advanced AI systems are able to learn from every interaction adding information about new threats to the knowledge base. This involves more than just identifying the malware in patterns but with a memory of the distinction between false positives and real threats, which increases the accuracy of the cybersecurity.

In addition, the role of AI is not only to detect but also to predict security. Advanced types of predictive analytics with an assist from AI technology can use long-term trends and patterns to project possible security breaches happening in the near future. By preparing defences in advance, rather than reacting to an attack as it occurs, businesses can take a proactive stance to cybersecurity. In one example, if an AI system notices a pattern of break-in attempts on a network of some form, it can launch automated security measures to block similar threats before they are ever even executed, preventing potential attacks from occurring [8].

Furthermore, integration of AI technologies in cybersecurity facilitates the emergence of adaptive security architectures. These systems are designed to be adaptive, learning and more importantly changing with their

environment. In a scalable, AI-driven framework, we can also safely let the algorithms learn and evolve as they grow in a way very similar to how humans develop an immune system that learns from exposure to threats. The capability to dynamically both changes the code and its behavior is especially useful in the context of polymorphic and metamorphic malware, which can morph over sufficiently narrow time window and tailor its appearance and behavior to outrun static security solutions [9]. However, this dependence on AI to fight malware also brings with it several issues that need to be tackled.

The complexity of AI systems also renders them vulnerable to adversaries, who in the worst case can influence the learning process inflows using methods such as poisoning attacks (inserting false data into the system) which distorted the learning and decision process. Furthermore, complexity of AI models may increase its non-transparent and non-interpretable characteristics among decision-making processes that can endanger debugging and trust in AI embedded systems [10].

We need to address them, and that requires the development of "responsible," secure, transparent, and ethical AI systems. This includes verifying the source of data used to train AI models; creating simulated, adversarial environments for vigorous testing of AI that let organizations attack or assess how well an AI will perform; and increasing how much the developers of enthusiastic cybersecurity staff understand behavioural analysis and statistical logic in AI programs. With advancements in these technologies, collaboration between the AI researchers, security experts, and regulatory bodies become mandatory for setting good practices and standards to use AI safely and effectively to curb the malware [11].

In the battle against malware and for better detection, predictive insights, and adaptive security, AI is proving to be an exciting frontier. As the journey of SOAR continues to play out, its continued integration into the cybersecurity strategy will likely upend the way defences are built and operated, providing a new and adaptive answer to the rapidly advancing threat landscape. Consequently, the future of malware defense depends not solely on creating novel AI solutions but on devising a systematic method of improving effective and fair AI technologies.

With the growing sophistication of cyber threats, particularly polymorphic and metamorphic malware, traditional and even early AI-based detection techniques often fall short. To address this, recent research has turned to the use of LLMs—deep learning architectures pre-trained on massive corpora of textual data—to augment malware detection systems. LLMs, such as GPT or BERT-style transformers, are proving valuable in domains beyond natural language processing, especially in understanding complex sequences and contextual patterns found in network and binary data.

LLMs can parse and interpret protocol logs, packet-level sequences, and even obfuscated code, enabling them to function effectively in real-time network monitoring systems. They are particularly well-suited for detecting contextual anomalies—a key weakness in rule-based and signature-based systems. By learning normal traffic behavior patterns and linguistic-like structures in logs and telemetry, LLMs can flag subtle deviations that may signal command-and-control (C2) communications, lateral movement, or exfiltration attempts.

For instance, researchers have shown that transformer-based models trained on packet payloads and flow metadata can identify malware-generated traffic, even when obfuscated using encryption or tunnelling protocols (e.g., DNS tunnelling) [12].

2. MALWARE DETECTION AND ANALYSIS

2.1 Static Malware Analysis Using AI

Static analysis involves examining the malware without executing it. Breaking down the code to recognize patterns and signatures that may be malicious this traditionally involved malware detection based on signatures, i.e., matching the code against a database of specific malware signatures. Unfortunately, modern malware is little bit more sophisticated in their obfuscation techniques to evade detection [13].

A comparative overview of how AI technologies are applied in malware detection, highlighting the distinctive features, methodologies, and outcomes of static and dynamic analysis. Table 2 provides a comparative summary of the application of AI technologies in malware detection, contrasting static and dynamic analysis approaches.

Table 2. Malware Detection through AI: Static and Dynamic Analysis

Analysis Type	Description	AI Technologies Used	Key Studies/Findings	Effectiveness/ Outcomes
Static Analysis	Analyses malware without execution by examining the code to detect suspicious patterns and signatures. Traditionally relies on signature-based detection, which struggles against obfuscation techniques used by modern malware.	ML algorithms, particularly Convolutional Neural Networks (CNNs), which analyze binary code and static features such as byte-level representations.	Used a CNN model trained on raw byte sequences of executable files [13].	Achieved significantly higher detection rates compared to traditional methods. Especially effective in identifying zero-day threats and polymorphic malware, which constantly change their signatures.
Dynamic Analysis	Involves executing malware in a controlled environment to observe its behavior, capturing the actions performed by the malware during runtime. More effective at uncovering the true nature and intent of the malware.	ML models like Recurrent Neural Networks (RNNs) and behavior analysis algorithms that monitor and learn from the operation of the code over time.	Example study not specified in the original text but generally involves using AI to identify deviations in normal operational patterns, indicating potential threats.	Enables real-time detection and mitigation of threats by analysing behavioural patterns, effective against complex and sophisticated malware.

By using AI on top of static analysis, one can apply machine-learning algorithms for analysing the binary code and identifying malicious patterns unknown in the signature bases. For instance, patterns of phishing recorded across similar executables may be observed by CNNs in code embedding from static information from the machine's binaries, such as the byte level, which may be valuable in identifying features of malware [14].

A study by Raff et al. that has been trained using raw byte sequences of executable files; it showed a much greater detection rate than traditional methods! The intelligence-based detection is very successful at tracking zero-day threats and polymorphic malware, which change their signatures with each usage [15].

Static analysis involves the careful examination of an executable's signature without the need to execute the code, aiming to classify the file as malware if the signature appears malicious or as benign if otherwise. This method has the reverse engineering of malware code and involves the detailed processing of extracted features to discern and interpret any malicious activities through a signature-based approach. In this context, a signature refers to a unique identifier4 for a binary file, determined by calculating its cryptographic hash [16].

2.2 AI for Dynamic Malware Analysis

Dynamic Analysis is the process of running the malware in a controlled environment, to see what the malware does. It is highly successful for discovering the true nature of some malware since few malicious behaviours are performed during its runtime. Behavioural analysis uses specialized AI to perform dynamic analysis, and uses fully automated AI features. Models like RNNs learn how the code runs at runtime (how it acts over time) which then helps in the identification of malicious patterns in the code [17].

Dynamic analysis benefits from LLMs' ability to model behavioural logs generated during sandbox executions. Sequences of API calls, file system interactions, or network events are semantically rich and context-sensitive. LLMs help in distinguishing benign from malicious actions not just based on frequency but based on contextual dependencies, much like how word meaning in a sentence depends on its surrounding tokens. Figure 2 describe the

monitoring the execution process, along with data flow dependency which implement ML algorithm for training and testing. So that malware detection has been identified.

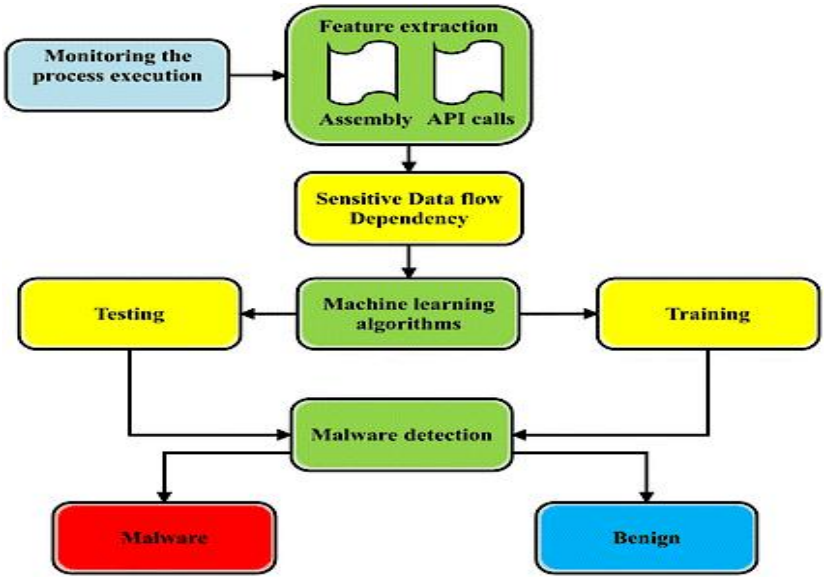

Figure 2. Dynamic Analysis and ML

One of the most significant uses of AI in the field of dynamic analysis is the application of powerful anomaly detection techniques. One common approach is training these systems on a dataset of what typical behavior (User, Process, and System using this same tool (YETI)). The model can then be used to detect deviations as potential malware activity. An interesting story here involves a large cybersecurity company deploying AI powered behavioural-based monitoring to dynamically analyze the network-behavior of installed applications and identify the behavior as coming from an advanced ransomware attack, thereby neutralizing it before large scale damage could ensue [18].

In static analysis, LLMs can be trained on assembly code, bytecode, or PE file strings. When formatted as sequences of tokens, code can be treated like a natural language. Studies have shown that sequence-based models like transformers outperform CNNs and RNNs in learning relationships across large opcode sequences, helping to identify subtle malware signatures [19].

Dynamic analysis benefits from LLMs' ability to model behavioural logs generated during sandbox executions. Sequences of API calls, file system interactions, or network events are semantically rich and context-sensitive. LLMs help in distinguishing benign from malicious actions not just based on frequency but based on contextual dependencies, much like how word meaning in a sentence depends on its surrounding tokens [20].

2.3. AI Methodologies in a Comparative Study in Malware Detection

Such distinct trade-offs draw attention to advantages and disadvantages of different AI approaches to malware detection. For instance, decision tree algorithms offer easily interpretable answers for malware detection, however their simple decision boundaries are no match for the more advanced malware surreptitiously hiding in the background. Others are deep learning techniques like deep neural networks (DNNs), which have better detection performance for stealth or new malware species. But they are computationally expensive, train on huge datasets, and are hard to interpret as they are "black-box" methods [21].

ML methods, such as support vector machines (SVM) and random forests, have been widely adopted in static and dynamic malware analysis. This is particularly useful in malware analysis where SVMs perform well in high-dimensional spaces - where we have a few hundreds or thousands of features extracted from a single piece of malware. The assignment of random forests to a given class of methods also allows for their usefulness in overfitting

and noisy data, as the perform noticeably better than other methods and classifier under highly non-linear and unbalanced data conditions [22].

These methods were analysed comparison-wise in 2019 for different datasets, including the Microsoft Malware Classification Challenge (BIG 2015). Our study showed that the DNN which most of the time were in obviously showing better performance regarding detection accuracy, however, the ensemble methods which are very rare in the top position most of the time the random forests made a good trade-off among high detection accuracy, low computational expense, and direct application [23].

2.4. Emerging Trends in Cybersecurity: From Malware Detection to Federated Learning Applications

This section explores recent advancements in malware detection, federated learning applications, and cybersecurity methodologies across various domains. Static malware analysis using ML on dataset, utilizing string and PE header features can effectively classify malicious detection [24]. Combination of both static and dynamic analysis in the OPEM framework (open-source tool kit for coding) create a machine-learning-based malware detection and neural network model showing improved flexibility and accuracy in identifying threats [25]. Several studies explored federated learning and optimization in energy systems focused on trusted decentralized federated learning, emphasizing privacy and security in distributed environments [26],[27].

Several studies investigated federated learning in the context of power systems. Abnormal power consumption detection system utilizing federated learning, demonstrated improvements in detection efficiency while preserving data privacy through proposed object detection model [28],[29]. Power forecasting spatiotemporal data using federated learning in smart grids, aiming to detect false data injection attacks while ensuring data privacy across different silos [30],[31]. Use of block chain technology in the construction supply chain demonstrates its effectiveness through a case study and threat model [32]. Federated Learning in Network and Security Applications demonstrate various methods for cloud-edge network communications, optimizing latency along with block chain-based decentralized federated learning model, enhancing security and transparency [33], [34]. Poisoning attacks in federated learning, employing normalizing flows mitigate the impact of adversarial inputs [35],[36]. Distributed Control and Cybersecurity Measures tackled the DNS cache poisoning attack, proposing an adaptive caching approach to enhance security in network infrastructure [37].

Cyber Security Breaches Survey 2020, which highlights common threats and the state of cybersecurity in organizations across the UK [38]. Specific case studies and reports discuss incidents like phishing attacks in the construction industry [39], the data breach incident involving Jewson [40], and ransomware attacks targeting Bird Construction [41] as well as Hoffman Construction's health plan data hack [42]. These reports demonstrated the prevalence and impact of cybersecurity threats in critical sectors.

Cyber risk management, focusing on prioritizing threats, identifying vulnerabilities, and applying controls for optimal security posture [43]. Evaluation of ML applications in cyber risk analysis within the construction industry, presented a SWOT (Strength, weakness, opportunities and threats) analysis that highlights the strengths and challenges associated with implementing ML in this domain [44]. NIST released a framework for improving critical infrastructure cybersecurity, offering guidelines for enhancing resilience and preparedness against cyber threats [45].

2.5. AI in Real-Time Malware Mitigation and Response

Another provision of real-time malware mitigation is for the ability to make out the threats promptly when they happen and to neutralize them to limit their resulting damage and slow down their development. AI can work with such large & fast datasets in a way that is unbeatable by humans. ML can detect anomalies that the AI can use to help real-time mitigation.

AI systems, for example, can watch for network traffic patterns and contrast them with its body of historical experiences. Any anomalies, like abnormal outbound traffic or increases in data access, can be immediately highlighted for review. In reality, deep learning models, especially those that use neural networks, are excellent at detecting patterns related to advanced cyber threats such as zero-day exploits when the model has not seen any of the current malware signatures. This was confirmed by a 2020 study by Cisco showing that the models, when implemented in their intrusion prevention systems (IPS), were able to detect and subsequently prevent over 99.8%

of malware encounters from real-time traffic patterns analysis. ML is widely used for general anomaly detection, while deep learning can handle more complex threats with a proven record of high effectiveness in specific cases like Cisco's system.

One of the most impactful uses of LLMs is in stream-based threat detection. Instead of waiting for complete log files, LLMs can process real-time input (such as syslog, Net Flow, or endpoint telemetry) token-by-token. This capability has been explored in real-world implementations such as Cisco's AI-enhanced intrusion prevention system, which used transformer-based models to block over 99.8% of unknown malware by analysing traffic flows in real time. Coupled with SOAR platforms, LLMs can also suggest remediation steps or even generate incident response plans based on learned correlations from historical incidents, improving automated response capabilities.

Unlike conventional ML models (e.g., SVM, Random Forests), LLMs require minimal manual feature engineering. They learn hierarchical representations from raw input, making them ideal for domains where feature crafting is infeasible or limited by expert knowledge.

Moreover, pertaining LLMs on multi-domain datasets (e.g., malware code, cybersecurity forums, phishing emails) enables them to generalize better and detect new attack vectors—such as zero-day exploits or blended threats—that evolve faster than labelled datasets can accommodate. Overview of AI models applied in cybersecurity, outlining their specific implementations and measurable effectiveness, with a focus on how ML and deep learning techniques contribute to anomaly detection and real-time threat mitigation presented in Table 3.

Table 3. AI Models in Cybersecurity: Implementations and Effectiveness

AI Technique	Application in Cybersecurity	Key Implementations	Effectiveness/Results
ML Algorithms	Used for detecting anomalies in real-time, which indicate potential malware activities.	Monitoring network traffic patterns to identify unusual activities like outbound traffic spikes or unauthorized data access.	General application across various systems, no specific study mentioned but widely regarded as effective in anomaly detection.
Deep Learning Models (Neural Networks)	Employed to discern complex patterns in data that are indicative of sophisticated cyber threats, including zero-day exploits.	Cisco's implementation within their IPS in 2020.	Successfully blocked 99.8% of malware encounters by analysing network traffic patterns in real time.

2.6. Automated Systems for Threat Detection and Response Driven by AI

Using AI for security on automated threat detection and response is a revolution in the cybersecurity industry. Using AI technology in these systems helps in a double fold- ability to identify threats more accurately and mobile responding to identified threats automatically. For instance, an AI system, noticing a ransomware attack progress, can also activate some routines to isolate the contaminated network segment and thus stop the malware from reaching the important data centres.

For example, in practice wherever AI-driven automation such as Security Orchestration, Automation, and Response (SOAR) solutions are used. These platforms leverage AI to manage a spectrum of security operations, from the most basic signature updates to the most advanced threat hunting and eradication processes. Now by bringing AI into the equation, SOAR platforms take SIEM capabilities (collect, analyze, and correlate security event data) and response strategies to a whole new playing field, cutting detection to containment time in half, milliseconds in some cases.

For example, a major financial services company used an AI-powered SOAR platform in a live situation, the response time with email-based Phishing Attacks was reduced from several hours to a few minutes, significantly reducing their exposure and potential data loss.

3. EXAMPLES OF AI INTEGRATION INTO CYBERSECURITY DEFENSE MECHANISMS

The impact of AI in cybersecurity is not just restricted to detection and response but also provides a wide repertoire of defense measures that can ensure a stronger defense. AI is key to the development of adaptive security architectures that change based on constant learning of the threat environment. Systems can dynamically change security measures and adapt defences.

AI tracks user behavior, such as keystroke dynamics, mouse movements, and navigation patterns. Even before any malicious activity occurs, these systems are able to identify and detect of departure from the norm, which could mean a compromised account. AI is even being utilized in banking to make better use of behavioural biometrics, which enhance fraud detection at major banks and financial institutions, leading to radical reduction in false positives and an overall improved customer experience.

ML models also used in predictive security, which will predict the future threat scenarios according to the trends and historical data available. This helps to alert the organizations to being proactive in modifying security measures and strategies to resolve the vulnerabilities ahead of time. An international company used AI to anticipate possible attack vectors and then automatically secured certain nodes from which the company receives hacking attempt multiple times.

The integration of AI and ML is poised to transform the SOC by enhancing threat detection capabilities and enabling more effective responses to cyber threats. These advanced technologies fundamentally shift the operational landscape of SOCs, allowing for the faster and more accurate identification of potential security incidents [46].

3.1. Adaptive Security Architectures

A primary driver for the rise in cybersecurity is the integration of AI as a component within Adaptive security architectures. The architecture itself should evolve continually to adapt to ever-changing threats, thus solidifying the resilience and strength of security parameters. Instead of conventional security systems that protect only against static threats, adaptive security architectures take advantage of the learning capabilities of AI to continuously adapt and enhance security in response to ever-changing threats learnt from large datasets. This continuous learning includes studying patterns and behaviours of both normal and malicious activities so new or previously unseen threats can be recognized, and response strategies can be adjusted.

Adaptive security architectures have a significant positive turn out having proactive emphasis. Most of the traditional security systems are based on specific predefined set of rules and signatures that helps in identifying and preventing threats. This is not sufficient against a unique complex malware which can evade these static defences easily. In contrast, when detecting security breaches, highly sophisticated AI-powered adaptive systems can identify minor anomalies and deviations from established norms.

For example, if a network typically exhibits a pattern of data traffic, then any significant deviation from this pattern can be detected and investigated in real-time, even to the point of revealing stealthy attacks before they can become a major threat.

The ability of adaptive security architectures to dynamically security architectures adjustment is really important in the fight against polymorphic and metamorphic malware that change their code to go undetected. Through constantly learning and updating their intelligence-gathering capabilities, adaptive security systems remain vigilant in detection systems to stay on top of these changing threats. This agility turns security from a primarily reactionary tactic into a proactive approach, which creates a smaller window of vulnerability and results in better overall security posture.

One other critical aspect of adaptive security architectures is that they can integrate with many different cyber-security tools and technologies. AI can collect information from various sources, namely network traffic, endpoint events, and even user behavior analytics to craft a more complete threat intelligence framework. This means that in addition to making threat detection more precise and quicker, integration facilitates the process of responding to incidents that are already underway. In the case of an anomaly being detected, system can be programmed to perform predefined actions such as isolation of infected network segments or notifying security team about ongoing attack and ultimately to reduce the impact of an attack.

Adaptive security architecture from application of AI, if designed and programmed properly is truly disruptive. Using continuous learning and the ability to adjust automatically and in real time, these systems represent a very

strong defense against the current evolving threats. They turn you from reactive to proactive security allowing to move quickly and at the same time to be equipped with the intelligence to protect against even the most sophisticated cyber threat. But as AI technology grows even more sophisticated, adaptive security architectures are going to be huge for the future of cybersecurity by making defences more elastic and adaptive against the wide array of threats that continue to plague the overall cybersecurity landscape.

3.2. Behavioural Biometrics

Behavioural biometrics is the AI technology used to analyze and detect behavior patterns that can be used to enhance the security of systems. The approach is to watch and learn these special features of each human interacting with computer systems. These behaviours can range from keystrokes dynamics, mouse movements, touchscreen gestures, typing speed, and even the rhythm of key presses. Through creating a holistic pattern of a typical behavior of a user, which consists of most known interaction fingerprints, the systems can be used to verify authenticator and detect fraud or account takeover.

Behavioural biometrics derive their power from being always-on passive authentication. Behavioural biometrics verifies the identity of the user on a constant basis while the user interacts with the system - vastly different from traditional authentication mechanisms like passwords, or even biometric scans, which occur at isolated points in time [47]. This constant monitoring guarantees that if some unwanted element manages to bypass the initial entry, any behavior that strays from the prospective norm will set off an alarm that will, at the very least, lock down the system if the trespasser continues to exploit the vulnerabilities of the system.

AI algorithms are responsible for the analysis of the extreme amount of data created over user interactions. With the help of ML models over massive datasets help recognize the unique identifies associated with legitimate users. These models can detect nuanced changes and even gracefully roll with the flow, by taking into account small shifts in user behavior over time, thereby improving the discrimination of the system [48].

What makes behavioural biometrics so valuable is that it can identify potential fraud and security breaches before any malicious intent is shown. For example, if an attacker hacks into an account to interact with the system its' behavior compared to the legitimate user behavior is generally different. The AI system has a fast eye on finding these patterns and can take action like locking the account, alerting the real user, notifying security action teams. By adopting this pro-active approach, the window any attacker gets to do nefarious deeds is greatly reduced [49].

In addition to that, because the profile is consistent, behavioural biometrics is very difficult to spoof. While traditional biometrics, such as fingerprints or facial recognition, may be tricked by high-quality replicas or images. By comparison, trying to replicate someone's special behaviours is a whole lot harder. With the episodic nature of the behavioural data, it is possibly even more secure, as frequent and significant resets in the behavioural profile makes it extremely hard for the attackers to imitate the exact same behavior as the legitimate user [50].

Behavioural biometrics is something that makes the life of the user easier by avoiding the need for additional secure measure that are just added on top [51]. Operating systems are able to interact with users in a seamless manner that does not keep prompting the user to authenticate themselves, because the operating system is always checking in the background. This balance between security and usability is especially useful in consumer applications, online banking or corporate networks where the reduction of user friction is critical [52].

To sum up, the behavioural biometrics obviously emerge as something new and innovative for cyber security which in the end also sounds easy to use and to set identity with. These systems use AI to perform pattern analysis and learn from user behavior and as a result, provide a continuous enterprise-wide passive authentication that can determine if there are any anomalies and in real time allow you to take immediate actions in case of a potential breach. With cyber threats getting more and more sophisticated, behavioural biometrics will prove important moving forward in the protection of digital identities, ultimately meaning improved security posture across the ecosystem.

Table 4 summarizes the types, techniques, data sets used, methodology and accuracy utilized in the reviewed articles in the present study.

Table 4. Malware Trends and AI Techniques Review

Ref., Year, Citation	Paper Type	Focus	Technique	Feature	Dataset	Algorithm/ Methodology	Accuracy & Precision (%)
[1], 2021, 35	Review	Evolutionary study of Internet of Things (IoT) malware	ML models	Malware characteristics & behaviours	38,963 IoT malware samples from 36 families, including honeypot-collected samples and commercial interchange samples	An ensemble model for malware classification and lineage analysis	NA
[2], 2015, 162	Review	Network processing in IoT evolution	ML, SDN, NFV	Network processing challenges	NA	Disruptive potential of three aspects of the IoT with respect to network protocols and their processing: the reversal of the client/server architecture, the scavenging of spectral bands, and the federation of Internet gateways	NA
[3], 2019, 120	Analysis	Static analysis of ransomware	Static analysis	Ransomware properties	NA	-	NA
[4], 2018, 229	Review	Overview of malware analysis techniques	Survey/ Review	Various analysis techniques	NA	Survey for malware detection methods like signature-based and heuristic-based	NA
[5], 2020, 174	Analysis	Static malware detection in Android byte-code	Deep learning	Byte-code features	Android byte-code	propose an anti-malware system that uses customized learning models, which are sufficiently deep, and are 'End to End deep learning architectures which detect and attribute the Android malware via opcodes extracted from application bytecode' (Bidirectional long short-term memory (BiLSTMs) neural networks)	NA
[6], 2019, 88	Analysis	Static malware analysis using ML	ML algorithms	String and PE header features	APT1 dataset	Presented implementation of two categories of malware detectors using (a) strings and (b) selected PE header features, respectively. For each category, Author implemented six different ML based classifiers	NA
[7], 2015, 51,253	Review	Deep learning overview	Review of deep learning techniques and applications	Representation learning, supervised learning, convolutional networks, recurrent networks	Various benchmark datasets for image recognition, speech recognition, NLP tasks	Various, including CNN, RNN, Long Short-Term Memory (LSTM)	NA
[8], 2019, 98	Analysis	Static and dynamic malware analysis	ML	Static and dynamic features	39,000 malicious binaries, 10,000 benign files (static); 2,200 malware, 800 benign (dynamic)	The combinations of different features are used for dynamic malware analysis. The different combinations are generated from APIs, Summary Information, DLLs and Registry Keys Changed. Algo Used: Logistic Regression, Decision Tree, Random Forest, Bagging	Accuracy: 99.36 (static), 94.64 (dynamic)

						Classifier, AdaBoost Classifier, Gradient Boosting Classifier	
[9], 2023, 677	Review	Adaptive AI framework for detecting polymorphic and metamorphic malware	Hybrid of dynamic deep learning and heuristic-based analysis	Behavioural pattern extraction, real-time data adaptation	Custom dataset with recent malware samples	Deep learning model combined with heuristics for pattern analysis	NA
[11], 2015, 35	Analysis	Intelligent approaches for static malware analysis	Various (e.g., ML methods)	Intelligent analysis techniques : Mnemonic n-grams, PE header features, API calls, function lengths, strings	992 malicious samples from VX Heaven, 854 benign samples from the System32 folder (Windows7)	paper details some intelligent techniques for malware analysis with all preprocessing steps required to analyze any PE sample like malware classification using mnemonic bi-grams as features Algo Used: Naïve Bayes, IBk, SMO, J48, Random Forest, AdaBoostM1	Accuracy: 96.10 (IBk), 95.78 (SMO)
[12], 2020, 84	Analysis	Android malware detection using network traffic	Two-layer deep learning model with static and network traffic analysis	Network traffic features: Permissions, intents, components, network traffic data	CICAndMal2017: 5,065 benign apps, 429 malware network traffic samples	1) Fully connected neural network for static malware detection 2) Convolutional Auto-Encoder (CAE) for unsupervised feature extraction from network traffic 3) Cascading Convolutional Auto-Encoder and Convolutional Neural Network (CACNN) for supervised malware detection	Accuracy: 99.3 (binary classification), 98.22 (category classification), 71.48 (family classification)
[13], 2018, 117	Analysis	Malware Detection Using ML and Deep Learning	ML & deep learning	Various malware features : Opcode frequency, Windows API calls, system calls	Malicia Project: 11,688 malware samples, 2,819 benign executables	Random Forest, DNNs (DNN-2L, DNN-4L, DNN-7L)	Accuracy: 99.78 Precision:100 Random Forest with Variance Threshold)
[14], 2017, 152	Analysis	Android malware classification using ML	Static analysis and source code analysis using ML	Android features : Permissions, source code (bag-of-words)	M0Droid dataset (200 malicious, 200 benign apps)	SVM with Sequential Minimal Optimization (SMO) - Naive Bayes - C4.5 Decision Trees (J48) - JRIP - AdaBoost - Farthest First clustering - K-means clustering - Expectation Maximization (EM) clustering - Ensemble learning with majority voting using combinations of 3 and 5 algorithms (including SVM, C4.5, Random Forest, JRIP, Logistic Regression)	Accuracy: 95.1 Precision:89
[15], 2022, 68	Analysis	Static malware detection using ML methods	ML	Malware detection features: PE file format features (Subsystem, Size of Optional Header, ID, Sections Min Entropy, etc.)	Ember dataset (1.1M files), unprocessed data from malware security partner of Meraz'18 (malicious and legitimate files)	use of PE file format along with ML statistics to determine whether a particular program is malicious or not. Algo Used: Decision Trees, Random Forest, Gaussian Naïve Bayes, AdaBoost, Gradient Boosting	Accuracy: 99.97 (Random Forest)
[17], 2021, 222	Analysis	Android malware detection with ML classifiers	ML classifiers	Android features	NA		NA
[18], 2018,	Analysis	Quantifying the	Static and dynamic	Known benign and malicious	Dataset from industry: 2 million	Author proposed n-gram and MalConv models are	Accuracy:95.5 (n-gram), 94.1

98		Robustness of ML and Current Anti-Virus	analysis with adversarial modifications	files, adversarial modifications	samples (malicious and benign)	trained on the same corpus. Compare two ML classifiers and four commercial anti-virus products: AV1, AV2, AV3, and AV4	(MalConv), 97.0 (AV1), 81.6 (AV2), 89.2 (AV3), 92.6 (AV4)
[21], 2018, 110	Analysis	Evaluating shallow and deep networks for static PE malware detection	DNNs	PE file features (size, entropy, etc.)	Ember dataset: 1.1M binary files (300K malicious, 300K benign, 300K unlabelled for training, 100K malicious, 100K benign for testing)	Algo used: DNNs, Logistic Regression (LR), Naive Bayes (NB), k-Nearest Neighbour (KNN), Decision Tree (DT), Random Forest (RF), SVM (linear and rbf kernels)	Accuracy,98.9 (DNN) Precision 99.7(DNN)
[22], 2019, 140	Analysis	Robust malware detection with deep learning	Static and dynamic analysis	Various malware features: Opcode sequences, system calls, image processing	Public and private datasets: Ember dataset (PE files), Malimg dataset (images)	a scalable deep learning network architecture for malware detection called ScaleMalNet is proposed with the capability to leverage the application of Big Data techniques to handle vary large number of malware samples	Accuracy: 99.9 (DNN), 97.8 (SVM+LSTM)
[23], 2020, 132	Review	Systematic review of Android malware detection using static analysis	Systematic literature review	Permissions, API calls, intents, hardware components, opcode sequences, program graphs, symbolic execution	98 studies from January 2014 to March 2020	Various static analysis techniques, neural network models, non-neural network models:- Android characteristic-based method - Opcode-based method - Program graph-based method - Symbolic execution-based method	NA
[24], 2013, 100	Analysis	Static malware detection using data mining	Data mining method	Malware features	NA		NA
[25], 2013, 44	Analysis	A Static-Dynamic Approach for Machine-Learning-Based Malware Detection	Hybrid approach combining static and dynamic analysis	Various malware features: Opcode sequences, system calls, operations, raised exceptions	1,000 malware samples (VxHeavens), 1,000 benign samples (collected from computers)	OPEM, an hybrid unknown malware detector which combines the frequency of occurrence of operational codes (statically obtained) with the information of the execution trace of an executable (dynamically obtained). Algorithm Used: KNN, Decision Trees (Random Forest, J48), SVM (RBF Kernel, Polynomial Kernel, Normalized Polynomial Kernel, Pearson VII Kernel), Naive Bayes, Bayesian Networks	Accuracy: 96.60 (SVM with Normalized Polynomial Kernel)
[26], 2022, 15	Specific Technique	Federated Learning	Decentralized Federated Learning	Trust and Security Features	Custom		NA
[27], 2022, 53	Specific Technique	Federated Learning for Recommendations	Trust-Based Federated Learning	Recommendation Features	Custom		NA
[28], 2023, 12	Specific Technique	Abnormal Power Consumption Detection	Improved Federated Learning	Power Consumption Features	Custom		NA
[29], 2023, 7	Specific Technique	Object Detection in Power Operation Sites	Federated Self-Supervised Learning	Object Detection Features	Custom		NA

[30], 2023, 8	Specific Technique	Ultra-Short-Term Power Forecasting	Spatiotemporal Federated Learning	Photovoltaic Forecasting Features	Custom		NA
[31], 2023, 23	Specific Technique	Privacy-Enhancing Cross-Silo Federated Learning for FDIA Detection in Smart Grids	Double-layer encryption scheme, Shamir secret sharing, parallel computing	Smart Grid Features : Local model parameters, training data privacy, FDIA detection in smart grids	Simulation data from a multi-area grid with 64 buses, 58 loads, and 355 measurements	Federated Learning, Double-layer encryption scheme, Shamir secret sharing	Accuracy: 97.5 Precision:97.8
[32], 2020, 19	Research Analysis	Blockchain applied to the construction supply chain: A case study with threat model	Case study analysis	Collaborative systems, information processing, payment actualizations, resource utilization	NA	Blockchain technology	NA
[33], 2023, 10	Specific Technique	Federated Learning in Cloud-Edge Networks	Efficient Federated Learning	Cloud-Edge Communication Features	Custom		NA
[34], 2023, 6	Specific Technique	Blockchain-Based Federated Learning	Blockchain Federated Learning	Blockchain Features	Custom		NA
[35], 2023, 11	Specific Technique	Poisoning Attack Detection	Normalizing Flows	Federated Learning Features	Custom		NA
[36], 2022, 16	Specific Technique	Dual-filtering (DF) schemes for learning systems to prevent adversarial attacks	Dual-Filtering Schemes	Learning System Features: : Input filtering, output filtering, anomaly detection, outlier detection	MNIST, CIFAR-10, ImageNet	Multi-objective Genetic Algorithm (MOGA), Negative Selection Algorithm (NSA), Outlier Detection Methods (OCSVM, IF, VAE)	NA
[37], 2015, 60	Specific Technique	DNS Cache Poisoning Prevention	Adaptive Caching Approach	DNS Security Features	Custom		NA
[38], 2020, N/A	Survey Report	Cybersecurity Breaches	Survey Analysis : Quantitative and qualitative study	Digital footprint, cyber risks, management involvement	Survey data from UK businesses, charities, and educational institutions	Not applicable	NA
[39], 2021, N/A	Article	Phishing Attacks	Case Studies	Construction Industry	NA		NA
[40], 2017, N/A	News Article	Data Breach	Incident Report	Jewson Data Breach	NA		NA
[41], 2021, N/A	News Article	Ransomware Attack	Incident Report	Bird Construction	NA		NA
[42], 2021, N/A	News Article	Data Breach	Incident Report	Hoffman Construction	NA		NA
[43], 2019, N/A	Book	Cyber Risk Management	Risk Management Framework	Various industries	NA		NA
[44], 2022, 18	Review	A Preliminary SWOT Evaluation for the Applications of ML to Cyber Risk	SWOT Evaluation	Strengths, weaknesses, opportunities, threats of ML applications	Review of various sources and previous studies	SWOT analysis framework, various ML techniques for cyber risk analysis	NA

		Analysis in the Construction Industry					
[45], 2018, N/A	Framework	Cybersecurity Framework	Best Practices	Critical Infrastructure	NA		NA
[47], 2019, 23	Analysis	Cybersecurity Threats in Cloud Applications	Deep Learning	Cloud Security Features, Classify Intrusion Attacks in Network Communications	KDD '99, UNSW-NB15, CIC-IDS2017	SABADT (Signature- and Anomaly-Based Attack Detection Technique)	Accuracy: 99.91 (CIC-IDS2017), 98.84 (UNSW-NB15), 99.89 (KDD '99) Precision: 99.89 Highest with KDD
[48], 2023, 19	Review	6G Communications	Index Modulation	6G Communication Features	NA		NA
[49], 2023, 72	Specific Technique	Anomaly Mitigation in Cyber-Physical Systems	Explainable AI (XAI) Framework	Inverter-Based Features	Custom		NA
[50], 2017, 169	Specific Technique	Renewable Energy Management	Virtual Power Plant Management, Imperialist Competitive Algorithm (ICA)	Renewable Energy Features: Thermal load, electricity prices, storage states	Thermal load data from 2004, EEX spot market prices	ICA	NA
[51], 2022, 29	Specific Technique	Frequency Regulation in Power Plants	Grasshopper Optimization Algorithm (GOA) optimized two-stage controller for frequency regulation of grid integrated VPP	Power Regulation Features: Thermal load, electricity prices, storage states	Simulation data from MATLAB	GOA, Firefly Algorithm (FA), Butterfly Optimization Algorithm (BOA), Particle Swarm Optimization (PSO)	NA
[52], 2022, 39	Specific Technique	Edge-Based Byzantine-Robust Federated Learning	Edge-Based Federated Learning	Heterogeneous Data Features	Custom		NA
[53], 2021, N/A	Article	Cybercrime	Industry Analysis	Construction Industry	NA		NA
[54], 2021, 21	Research Analysis	Cybersecurity Risk Assessment in Smart City Infrastructures	Artificial Neural Networks (ANN) for risk assessment	Dynamic network analysis, real-time monitoring, risk classification	Synthetic datasets generated using NS-3 network simulator	Multilayer Perceptron, Backpropagation Algorithm	Accuracy: 97% (ANN classification accuracy)
[55], 2019, N/A	Framework	Information Security Controls	Best Practices	Various industries	NA		NA
[56], 2017, N/A	Regulation	Cybersecurity Requirements	Compliance	Financial Services	NA		NA
[57], 2021, 25	Research Analysis	Cybersecurity in Construction	Text mining, VOSviewer analysis	Cybersecurity risks, digital tools, construction industry	Web of Science (WOS) database	NA	NA

[58], 2023, 17	Journal Article	Cybersecurity Review	Scoping Review	Construction Industry	NA		NA
[59], 2020, 11	Book Chapter	Ransomware Mitigation	Training and Awareness	Various industries	NA		NA
[60], 2021, 34	Review Paper	ML in 3D Printing : ML in 3D Printing: Applications, Potential, and Challenges	Review of various ML techniques in 3D printing	Design for 3D printing, material tuning, process optimization, in-situ monitoring, cloud service, cybersecurity	NA	Various ML techniques including CNN, ANN, SVM, etc.	NA
[66], 2019, 54	Research Analysis	Context Aware Intrusion Detection for Building Automation Systems	Context-aware data structure, anomaly-based behavior analysis	Building Automation Systems: Runtime models, service interactions, functionality patterns	Simulation data from the Smart Building testbed at the University of Arizona	Bayesian Network, RIPPER, Decision Table	NA

4. CURRENT GAPS AND FUTURE DIRECTIONS - RESEARCH PRIORITIES

AI has certainly brought significant benefit to the race of stay-ahead to mitigate against the threats posed by malware, but it has also introduced new vectors of attack for attacker to exploit the weaknesses in the AI systems to turn the tables around [53]. A major challenge is that the learning processes of AI could be compromised by adversaries using poisoning attacks at training time. Poisoning attacks occur when malicious actors insert incorrect pieces of data into the training datasets that will be used to build a ML model. However, if the data is corrupt then it will corrupt the model which impairs the learning process of the model and thus the model can learn incorrectly about classes, and this is bad and this may allow some malware to evade to be detected otherwise. For instance, if an attacker is able to poison the training data which involves benign files and is able to make them look malicious, the AI system might begin tagging correct software as threats, leading you to more false positives and eventually more trust issues in your system [54].

4.1. Transparency and Interpretability

A further major obstacle is the high level of sophistication and non-transparency of AI models, particularly deep learning models. These models are frequently made to be "black box" - they do not reveal how they reached particular decisions or classifications. However, this lack of transparency can make troubleshooting extremely challenging when the system missteps or fails to catch more advanced malware. This is a similar type of roadblock for organizations when it comes to regulatory compliance and auditing - how do you demonstrate that your AI-driven security measures satisfy the standards and protocols that you are bound by making sure that AI models are interpretable and that they can be understood buy human analysts is key for trust and negligence [55].

In addition, deploying AI on malware mitigation necessitates stable and secure infrastructures. As AI system need huge computing power to process and storage heavy amount of data at the moment. The high-performance computing resources are required very much, that is the hindrance for smaller organizations, and with no budget these are too expensive. However, to maintain this ever-evolving security guard, AI models need to be refreshed with new data. And with continued need to train on current data, these models are a serious cost to model maintenance, especially in environments that always have a fast-evolving threat landscape [56].

4.2. Improve Detection Efficiency

A valuable future direction in explainable malware detection is to enhance the design methodologies of malware detectors so that the explanations they generate can assist professionals in more accurately characterizing malware attacks. For example, involve extracting features and employing a decision tree to develop a model capable of determining the maliciousness of applications. In addition, ethical considerations are a key factor in the future

development of AI in cybersecurity. It allows you to identify biases in training data that can result in unfair or even discriminatory outputs, which is important for any organizational that wants to ensure the responsible use of AI. An AI model that learns from biased data that over represent certain kinds of threats or regions may fail to detect or classifies threats from underrepresented areas. Diverse and expansive datasets including a range of use cases and threat types are needed for developing AI systems that are fair & unbiased [57].

A number of future directions have been proposed to overcome these challenges and in order to advance the field. One potential method is the technique known as adversarial training, which involves training a ML model on normal datasets, as well as adversarial examples, which are used to evaluate the model's capability to resist being tampered with and the ability to be improved by hostile examples [58]. Concurrent to this background, XAI, being a nascent field, is finding novel ways to make AI entities more transparent and interpretable [59].

4.3. Mitigate Attacks

In recent research, the primary focus has been on ML attacks, gradient-based attacks, evasion attacks, and poisoning attacks. Evasion attacks involve manipulating malicious input samples during the training phase to circumvent detection by a trained system, and it requires access to the model. Poisoning attacks compromise the integrity of training data by introducing incorrect data since it can mislead the learning process of ML models. This corruption of training data severely undermines the entire training process. AI researchers, cybersecurity experts and regulatory bodies must collaborate to set the standards, and best practices for the use of AI to ensure a secure and ethical manner of AI in the area of malware mitigation. This involves creating data integrity, model transparency, and ethical frameworks to make sure AIs operate not only efficiently but also in a reliable and unbiased way as possible [60].

4.4. Simulating attacks for real-world scenarios

Organizations can use AI to develop attack simulations of adversarial attacks and phishing attempts to prepare against real-world threats in their environment. Organizations can teach cybersecurity personnel to handle real-time attacks properly by running training simulations that minimize downtime and reduce damage [61].

4.5. Challenges and Mitigation Strategies for implementation of LLMs

Despite their transformative potential, LLMs present several significant challenges in cybersecurity applications. First, LLMs are highly data-dependent, requiring extensive, diverse, and high-quality datasets for effective training. In the absence of such datasets, they are prone to inheriting biases and may struggle with generalization. LLMs are susceptible to adversarial attacks; carefully crafted malicious inputs or poisoned training data can lead to misclassification or undesirable behavior, posing security risks. Hence Creation of standardized benchmarks across various detection tasks (e.g., fraud, spam, hate speech) require fair and comprehensive model evaluation [62].

The interpretability of LLMs remains a major concern. Their "black box" nature makes it difficult to understand decision-making processes—an issue that is particularly problematic in cybersecurity, where transparency and explain ability are essential for compliance and the trust of human analysts. Lastly, LLMs are resource-intensive, both in terms of training and inference. This computational burden presents obstacles for deployment in edge computing scenarios, such as IoT devices or real-time threat detection systems.[63-65].

To address these challenges, several mitigation strategies are being explored. One approach is fine-tuning lightweight transformer variants, such as Distil BERT or TinyGPT, which are better suited for resource-constrained environments. Another strategy involves the use of XAI techniques, which help visualize model behavior through attention maps and activation layers, thereby enhancing interpretability. Additionally, federated learning offers a promising solution for training LLMs on decentralized data sources, improving both data privacy and system scalability without the need for centralized storage.

To sum up, AI is emergent revolutionary technology in malware defense, but it brings the umbrella of challenges in data integrity, model transparency, resource expenditure and ethical matters. Meeting these challenges will demand

creative solutions while AI technologies evolve, improving and formalizing standards and practices together. Swimming against the stream by confronting these challenges in a direct manner, AI and ML will finally realize their full potential in cybersecurity and prove their mettle against a dynamic and dangerous adversary [66].

With the rapid evolving of malware IoT in the shape of numerous forms is considerably a big challenge for the designed detection system. Shortage of diversify data and limited scale of obsolete data hampers accuracy and training a model. Survey highlighted the importance of models against sophisticated attacks. Modern malware tactics used obfuscation and evasion tactics against static and dynamic models. These models are expected to be robustness against such adversaries.

Deployment of AI based model in sensitive installation where stakeholder can explain and interrupt the results is critical in explanation. Privacy preserving as a potential approach in federated learning for distributed IoT network, while facilitating a training model. Deep learning methods lead to high accuracy in malware detection but at the cost of high computational cost. Hence mostly it is found unsuitable in case of some IoT applications with limited resources.

Formulation of comprehensive models while integrating insights from behavioural which can be accounted for social engineering and human factors. Technology of block chain can maintain security, detection of logs and temper proof data while maintaining the integrity of malware attributes. Hence ensuring transparency and trustworthiness. Analysis of real time data with edge computing to support IoT devices with limited resources. Hence can provide more scalable solution in real time data analysis. LLMs are likely to play a central role in unified threat management systems, capable of ingesting and reasoning over diverse inputs like email headers, code binaries, user behavior logs, and network traces. Combined with federated and continual learning, LLMs may soon evolve into self-adaptive cybersecurity agents—learning new threats in real-time, just like immune systems adapt to new pathogens [67].

5. CONCLUSION

One of the most vital advancements in preserving the near-future threat landscape is the application of AI to cybersecurity, particularly to malware examination and prevention. The use of AI and AI-based methods are used to support conventional malware analysis for detection purposes such as static and dynamic analysis. AI not only detects but also offers most advanced real-time mitigation and response options which can dynamically change based on new threats. Moving from a reactive security posture to a proactive model then allows IT organizations to anticipate, and proactively contain, potential vulnerabilities, resulting in them shrinking the detection to recovery window.

However, reliance on AI introduces additional risks, like being vulnerable to adversarial manipulation or the fragility or opacity of AI models that hide their inner workings and impede trust and accountability. This can only be done by developing secure, transparent, and ethical AI systems. Going forward, efforts to normalize data integrity, create safe testing environments and increase the general understanding of AI processes by cyber security professionals will all be necessary. AI must also be transparent and developed with ethical considerations as the foundation of providing and applying equal opportunities, discouraging the development and use of AI based on injustice, and ensuring the use of AI is not misused.

More specifically, behavioural biometrics and predictive security are excellent demonstrations of how AI shifted cybersecurity, here allowing for always-on, low friction authentication and even predicting future threats. These capabilities demonstrate how AI can streamline processes and enhance user experiences. This means the next wave of cybersecurity will be characterized by more mature AI solutions and ecosystem-wide efforts to ensure tools are being used ethically and effectively.

In AI introduces a new game changing play as well as an outside of the box manner on how typical cyber security strategies would adapt having a response to threats of scale that is scalable and intelligent. It will be essential for AI scientists, cybersecurity practitioners, and regulatory agencies to collaborate to create these standards and best practices in order to embrace the advantages of AI or eliminate the risks. In addressing these challenges, and harnessing the power of AI, security professionals can better fortify our lines of defense and defend the increasingly complex cyber-landscape that surrounds it.

ACKNOWLEDGMENT

The authors would like to thank the faculty of Riphah University Islamabad for their valuable feedback and insightful discussions during the development of this research and encouragement throughout the study.

FUNDING STATEMENT

The authors received no funding from any party for the research and publication of this article.

AUTHOR CONTRIBUTIONS

Salman Khan: Conceptualization, Data Curation, Methodology, Validation, Writing – Original Draft Preparation;
Hasnat Raza: Project Administration, Writing – Review & Editing;
Mansoor Alam: Project Administration, Supervision, Writing – Review & Editing.

CONFLICT OF INTERESTS

No conflict of interests were disclosed.

ETHICS STATEMENTS

Our publication ethics follow The Committee of Publication Ethics (COPE) guideline. https://publicationethics.org/.

DATA AVAILABILITY

Data availability is not applicable to this paper as no new data were created or analysed in this study.

REFERENCES

[1] H. Wang *et al.*, "An evolutionary study of IoT malware", *IEEE Internet of Things Journal*, vol. 8, no. 20, pp. 15422-15440, 2021, doi: 10.1109/jiot.2021.3063840.

[2] L. Gregorio, "Evolution and disruption in network processing for the Internet of Things", *Ubiquity*, vol. 2015, no. December, pp. 1-14, 2015, doi: 10.1145/2822877.

[3] D. Vidyarthi, and S. Rakshit, "Static malware analysis to identify ransomware properties", *International Journal of Computer Science Issues*, vol. 16, no. 3, pp. 10–17, 2019, doi: 10.5281/zenodo.3252963.

[4] R. Sihwail, K. Omar, and K. Ariffin, "A survey on malware analysis techniques: Static, dynamic, hybrid and memory analysis", *International Journal on Advanced Science, Engineering and Information Technology*, vol. 8, no. 4-2, pp. 1662-1671, 2018, doi: 10.18517/ijaseit.8.4-2.6827.

[5] M. Amin, T. Tanveer, M. Tehseen, M. Khan, F. Khan, and S. Anwar, "Static malware detection and attribution in android byte-code through an end-to-end deep system", *Future Generation Computer Systems*, vol. 102, pp. 112-126, 2020, doi: 10.1016/j.future.2019.07.070.

[6] N. Balram, G. Hsieh, and C. McFall, "Static malware analysis using machine learning algorithms on APT1 dataset with string and PE header features," *2019 International Conference on Computational Science and Computational Intelligence (CSCI)*, Las Vegas, NV, USA, pp. 90-95, 2019, doi: 10.1109/CSCI49370.2019.00022.

[7] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning", *Nature*, vol. 521, no. 7553, pp. 436-444, 2015, doi: 10.1038/nature14539.

[8] M. Ijaz, M. Durad, and M. Ismail, "Static and dynamic malware analysis using machine learning", *2019 16th International Bhurban Conference on Applied Sciences and Technology (IBCAST)*, pp. 687-691, 2019, doi: 10.1109/ibcast.2019.8667136.

[9] A. Djenna, A. Bouridane, S. Rubab, and I. Marou, "Artificial intelligence-based malware detection, analysis, and mitigation", *Symmetry*, vol. 15, no. 3, pp. 677, 2023, doi: 10.3390/sym15030677.

[10] VirusShare. [Online]. Available: https://virusshare.com/, Accessed Nov. 26, 2023.

[11] T. Mithal, K. Shah, and D. Singh, "Case studies on intelligent approaches for static malware analysis", *Emerging Research in Computing, Information, Communication and Applications*, pp. 555-567, 2016, doi: 10.1007/978-981-10-0287-8_52.

[12] J. Feng, L. Shen, Z. Chen, Y. Wang, and H. Li, "A two-layer deep learning method for Android malware detection using network traffic," in *IEEE Access*, vol. 8, pp. 125786-125796, 2020, doi: 10.1109/ACCESS.2020.3008081.

[13] H. Rathore, S. Agarwal, S. Sahay, and M. Sewak, "Malware detection using machine learning and deep learning", In *Big Data Analytics: 6th International Conference, BDA 2018, Warangal, India, December 18–21, 2018, Proceedings*, pp. 402-411, 2018, doi: 10.1007/978-3-030-04780-1_28.

[14] N. Milosevic, A. Dehghantanha, and K. Choo, "Machine learning aided android malware classification*", Computers & Electrical Engineering*, vol. 61, pp. 266-274, 2017, doi: 10.1016/j.compeleceng.2017.02.013.

[15] K. Malik *et al.*, "Static malware detection furthermore, analysis using machine learning methods," *Advances and Applications in Mathematical Sciences*, vol. 21, pp. 4183–4196, 2022.

[16] H. Manthena, S. Shajarian, J. Kimmell, M. Abdelsalam, S. Khorsandroo, and M. Gupta, "Explainable Artificial Intelligence (XAI) for malware analysis: A survey of techniques, applications, and open challenges", *IEEE Access*, vol. 13, pp. 61611-61640, 2025, doi: 10.1109/access.2025.3555926.

[17] P. Agrawal, and B. Trivedi, "Machine learning classifiers for android malware detection", *Advances in Intelligent Systems and Computing*, pp. 311-322, 2020, doi:10.1007/978-981-15-5616-6_22.

[18] W. Fleshman, E. Raff, R. Zak, M. McLean, and C. Nicholas, "Static malware detection & subterfuge: Quantifying the robustness of machine learning and current anti-virus," *2018 13th International Conference on Malicious and Unwanted Software (MALWARE)*, Nantucket, MA, USA, pp. 1-10, 2018, doi: 10.1109/MALWARE.2018.8659360.

[19] E. Raff *et al.*, "Malware detection by eating a whole EXE," *arXiv preprint*, 2018, doi: 10.48550/arXiv.1710.0943.

[20] M. Ijaz, M. H. Durad, and M. Ismail, "Static and dynamic malware analysis using machine learning," *2019 16th International Bhurban Conference on Applied Sciences and Technology (IBCAST)*, Islamabad, Pakistan, pp. 687-691, 2019, doi: 10.1109/IBCAST.2019.8667136.

[21] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, and S. Venkatraman, "Robust intelligent malware detection using deep learning," in *IEEE Access*, vol. 7, pp. 46717-46738, 2019, doi: 10.1109/ACCESS.2019.2906934.

[22] J. Feng, L. Shen, Z. Chen, Y. Wang, and H. Li, "A two-layer deep learning method for Android malware detection using network traffic," in *IEEE Access*, vol. 8, pp. 125786-125796, 2020, doi: 10.1109/ACCESS.2020.3008081.

[23] Y. Pan, X. Ge, C. Fang, and Y. Fan, "A systematic literature review of Android malware detection using static analysis," in *IEEE Access*, vol. 8, pp. 116363-116379, 2020, doi: 10.1109/ACCESS.2020.3002842.

[24] U. Baldangombo, N. Jambaljav, and S. J. Horng, "A static malware detection system using data mining methods," *arXiv preprint,* 2013, doi: 10.48550/arXiv.1308.2831.

[25] I. Santos, J. Devesa, F. Brezo, J. Nieves, and P. Bringas, "OPEM: A static-dynamic approach for machine-learning-based malware detection", *Advances in Intelligent Systems and Computing*, pp. 271-280, 2013, doi: 10.1007/978-3-642-33018-6_28.

[26] A. Gholami, N. Torkzaban, and J. S. Baras, "Trusted decentralized federated learning," *2022 IEEE 19th Annual Consumer Communications & Networking Conference (CCNC)*, Las Vegas, NV, USA, pp. 1-6, 2022, doi: 10.1109/CCNC49033.2022.9700624.

[27] O. Wahab, G. Rjoub, J. Bentahar, and R. Cohen, "Federated against the cold: A trust-based federated learning approach to counter the cold start problem in recommendation systems", *Information Sciences*, vol. 601, pp. 189-206, 2022, doi: 10.1016/j.ins.2022.04.027.

[28] Z.Cai *et al.*, "An improved abnormal power consumption detection system based on federated learning," in *2023 4th International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE 2023)*, Hangzhou, China, pp. 378–382, 2023.

[29] S. Li, J. Hu, X. Chen, Y. Tan, J. Zhang, and P. Li, "An object detection model for electric power operation sites based on federated self-supervised learning," *2023 Panda Forum on Power and Energy (PandaFPE)*, Chengdu, China, 2023, pp. 1706-1710, 2023, doi: 10.1109/PandaFPE57779.2023.10141090.

[30] W. Fu *et al.*, "A spatiotemporal federated learning based distributed photovoltaic ultra-short-term power forecasting method," *2023 IEEE/IAS 59th Industrial and Commercial Power Systems Technical Conference (I&CPS)*, Las Vegas, NV, USA, pp. 1-7, 2023, doi: 10.1109/ICPS57144.2023.10142102.

[31] H. -Y. Tran, J. Hu, X. Yin, and H. R. Pota, "An efficient privacy-enhancing cross-silo federated learning and applications for false data injection attack detection in smart grids," in *IEEE Transactions on Information Forensics and Security*, vol. 18, pp. 2538-2552, 2023, doi: 10.1109/TIFS.2023.3267892.

[32] G. Shemov, B. Soto, and H. Alkhzaimi, "Blockchain applied to the construction supply chain: A case study with threat model", *Frontiers of Engineering Management*, vol. 7, no. 4, pp. 564-577, 2020, doi: 10.1007/s42524-020-0129-x.

[33] J. Duan, J. Duan, X. Wan, and Y. Li, "Efficient federated learning method for cloud-edge network communication", *2023 5th International Conference on Communications, Information System and Computer Engineering (CISCE)*, pp. 118-121, 2023, doi: 10.1109/cisce58541.2023.10142819.

[34] H. Wang, D. Mao, Z. Chen, H. Rao, and Z. Li, "Blockchain-based decentralized federated learning model," *2023 4th International Conference on Information Science, Parallel and Distributed Systems (ISPDS)*, Guangzhou, China, pp. 622-625, 2023, doi: 10.1109/ISPDS58840.2023.10235493.

[35] H. Hu, and L. Yuan, "Poisoning attack in federated learning using normalizing flows," *2023 International Seminar on Computer Science and Engineering Technology (SCSET)*, New York, NY, USA, pp. 310-313, 2023, doi: 10.1109/SCSET58950.2023.00075.

[36] D. Dasgupta, and K. Gupta, "Dual-filtering (DF) schemes for learning systems to prevent adversarial attacks", *Complex & Intelligent Systems*, vol. 9, no. 4, pp. 3717-3738, 2022, doi: 10.1007/s40747-022-00649-1.

[37] H. S. Hmood, Z. Li, H. K. Abdulwahid, and Y. Zhang, "Adaptive caching approach to prevent DNS cache poisoning attack," in *The Computer Journal*, vol. 58, no. 4, pp. 973-985, April 2015, doi: 10.1093/comjnl/bxu023.

[38] J. Emma, *Cyber Security Breaches Survey 2020*, London, UK: Dept. for Digital, Culture, Media & Sport, vol. 2020, pp. 4, 2020.

[39] Infosec, "Phishing Attacks in the Construction Industry." [Online]. Available: https://resources.infosecinstitute.com/topic/phishing-attacks-construction-industry/

[40] P. Kunert, "Shut the front door: Jewson fesses up to data breach," *The Register*. [Online]. Available: https://www.theregister.com/2017/11/14/jewson_suffers_data_breach/

[41] C. Tunney, "Ransomware attack on construction company raises questions about federal contracts," *CBC News*. [Online]. Available: https://www.cbc.ca/news/politics/ransomware-bird-construction-military-1.5434308

[42] R. Korman, "Hoffman construction reports hack of self-insured health plan data," *Engineering News-Record*. [Online]. Available: https://www.enr.com/articles/51232-hoffman-construction-reports-hack-of-self-insured-health-plan-data

[43] C. Christopher, *Cyber Risk Management: Prioritize Threats, Identify Vulnerabilities, And Apply Controls*. New York, NY, USA: Kogan Page Ltd., 2019.

[44] D. Yao and B. Soto, "A preliminary swot evaluation for the applications of ML to cyber risk analysis in the construction industry", *IOP Conference Series: Materials Science and Engineering*, vol. 1218, no. 1, pp. 012017, 2022, doi: 10.1088/1757-899x/1218/1/012017.

[45] NIST, *Framework for Improving Critical Infrastructure Cybersecurity*, ver. 1.1, Gaithersburg, MD, USA: NIST, 2018.

[46] T. R. McIntosh *et al.*, "Inadequacies of Large Language Model benchmarks in the era of Generative Artificial Intelligence," in *IEEE Transactions on Artificial Intelligence*, 2025, doi: 10.1109/TAI.2025.3569516.

[47] S. A. Sokolov, T. B. Iliev, and I. S. Stoyanov, "Analysis of cybersecurity threats in cloud applications using deep learning techniques," *2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, Opatija, Croatia, pp. 441-446, 2019, doi: 10.23919/MIPRO.2019.8756755.

[48] J. Li *et al.*, "Index modulation multiple access for 6G communications: Principles, applications, and challenges," in *IEEE Network*, vol. 37, no. 1, pp. 52-60, January/February 2023, doi: 10.1109/MNET.002.2200433.

[49] A. A. Khan, O. A. Beg, Y. -F. Jin, and S. Ahmed, "An explainable intelligent framework for anomaly mitigation in cyber-physical inverter-based systems," in *IEEE Access*, vol. 11, pp. 65382-65394, 2023, doi: 10.1109/ACCESS.2023.3289887.

[50] M. Kasaei, M. Gandomkar, and J. Nikoukar, "Optimal management of renewable energy sources by virtual power plant", *Renewable Energy*, vol. 114, pp. 1180-1188, 2017, doi: 10.1016/j.renene.2017.08.010.

[51] A. Srivastava, A. Latif, S. Shaoo, D. Das, S. Hussain, and T. Ustun, "Analysis of GOA optimized two-stage controller for frequency regulation of grid integrated virtual power plant", *Energy Reports*, vol. 8, pp. 493-500, 2022, doi: 10.1016/j.egyr.2021.11.117.

[52] F. Zhou, R. Yu, Z. Li, H. Gu, and X. Wang, "FedAegis: Edge-based Byzantine-Robust federated learning for heterogeneous data", *GLOBECOM 2022 - 2022 IEEE Global Communications Conference*, pp. 3005-3010, 2022, doi: 10.1109/globecom48099.2022.10000981.

[53] T. Sawyer, and J. Rubenstone, "Construction Cybercrime is on the Rise," *Engineering News-Record*. [Online]. Available: https://www.enr.com/articles/46832-construction-cybercrime-is-on-the-rise

[54] M. Kalinin, V. Krundyshev, and P. Zegzhda, "Cybersecurity risk assessment in smart city infrastructures", *Machines*, vol. 9, no. 4, pp. 78, 2021, doi: 10.3390/machines9040078.

[55] CIS, *Center for Internet Security Controls*, ver. 7.1, New York, NY, USA: CIS, 2019. [Online]. Available: https://learn.cisecurity.org/20-controls-download

[56] NYCRR, *Part 500 Cybersecurity Requirements for Financial Services Companies*, 2017. [Online]. Available: https://govt.westlaw.com

[57] B. Mantha, and B. Soto, "Cybersecurity in construction: Where do we stand and how do we get better prepared", *Frontiers in Built Environment*, vol. 7, 2021, doi: 10.3389/fbuil.2021.612668.

[58] N. Pargoo, and M. Ilbeigi, "A scoping review for cybersecurity in the construction industry", *Journal of Management in Engineering*, vol. 39, no. 2, 2023, doi: 10.1061/jmenea.meeng-5034.

[59] A. Bello, and A. Maurushat, "Technical and behavioural training and awareness solutions for mitigating ransomware attacks", *Advances in Intelligent Systems and Computing*, pp. 164-176, 2020, doi: 10.1007/978-3-030-51974-2_14.

[60] G. Goh, S. Sing, and W. Yeong, "A review on machine learning in 3S printing: Applications, potential, and challenges", *Artificial Intelligence Review*, vol. 54, no. 1, pp. 63-94, 2020, doi: 10.1007/s10462-020-09876-9.

[61] M. Khayat, E. Barka, M. Adel Serhani, F. Sallabi, K. Shuaib, and H. M. Khater, "Empowering security operation center with artificial intelligence and machine learning-a systematic literature review," in *IEEE Access*, vol. 13, pp. 19162-19197, 2025, doi: 10.1109/ACCESS.2025.3532951.

[62] I. Raji *et al.*, "Closing the AI accountability gap", *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, pp. 33-44, 2020, doi: 10.1145/3351095.3372873.

[63] E. Bender, T. Gebru, A. McMillan-Major, and S. Shmitchell, "On the dangers of stochastic parrots: Can Language Models be too big?", *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, pp. 610-623, 2021, doi: 10.1145/3442188.3445922.

[64] R. Bommasani *et al.*, "On the opportunities and risks of foundation models," *arXiv preprint*, 2021, doi: 10.48550/arXiv.2108.07258.

[65] Z. Lipton, "The mythos of model interpretability", *Communications of the ACM*, vol. 61, no. 10, pp. 36-43, 2018, doi: 10.1145/3233231.

[66] Z. Pan, S. Hariri, and J. Pacheco, "Context aware intrusion detection for building automation systems", *Computers & Security*, vol. 85, pp. 181-201, 2019, doi: 10.1016/j.cose.2019.04.011.

[67] A. Djenna, A. Bouridane, S. Rubab, and I. Marou, "Artificial Intelligence-based malware detection, analysis, and mitigation", *Symmetry*, vol. 15, no. 3, pp. 677, 2023, doi: 10.3390/sym15030677.

BIOGRAPHIES OF AUTHORS

	**Salman Khan** is a research scholar at Riphah International University Islamabad, Pakistan currently pursuing a PhD in Computing. He holds an MS in Information Security and a BS in Software Engineering. His research focuses on cybersecurity, with a particular interest in AI-driven detection systems. His current work explores the integration of Large Language Models (LLMs) in enhancing cybersecurity frameworks. He can be contacted at 24730@students.riphah.edu.pk.
	**Hasnat Raza** is a research scholar at Riphah International University Islamabad, Pakistan currently pursuing a PhD in Computing working at the intersection of cybersecurity and artificial intelligence. He holds MS degree in information security and with a strong focus on developing advanced AI-based security models and LLMs. He can be contacted at hasnatzaidi@gmail.com.
	**Dr. Mansoor Alam** is a faculty member at Riphah International University, Islamabad, Pakistan and as an Adjunct Professor at RMC Multimedia University (MMU), Malaysia with extensive teaching experience. He earned his PhD from France and holds multiple international certifications. He has several impact factor publications to his name and continues to contribute actively to research in computing and cybersecurity. Dr. Alam has also served as an Online Laureate (Facilitator) for the MSIS program jointly offered by Colorado State University, USA, and Saudi Electronic University, Saudi Arabia. Previously, he has held academic and research positions with University Kuala Lumpur, Malaysia Pahang. He can be contacted at m.mansoor@riphah.edu.pk.