IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.15579

Automated Threat Hunting Using AI: AI-Driven Defence Strategy

Vikram G.D[1], K. Sharath[2]

Department of MCA, BIT, K.R. Road, V.V. Pura, Bangalore, India [1]

Assistant Professor, Department of MCA, BIT, K.R. Road, V.V. Pura, Bangalore, India [2]

Abstract: The exponential digitization of global infrastructures, coupled with the proliferation of cloud computing, Internet of Things (IoT) devices, and decentralized workforces, has inadvertently created a vast and highly vulnerable cyber attack surface. Concurrently, adversarial tactics have evolved with alarming sophistication. Modern cybercriminals and state-sponsored actors frequently deploy zero-day exploits, advanced persistent threats (APTs), fileless malware, and polymorphic ransomware that are expressly designed to circumvent traditional, perimeter-based security architectures [1], [2]. Historically, cybersecurity has relied heavily on reactive paradigms—such as signature-based Intrusion Detection Systems (IDS) and standard firewalls—which require prior knowledge of an attack vector to mount a defense. This reactive posture is fundamentally insufficient in an era where the velocity and novelty of cyber threats outpace human response capabilities. To neutralize these stealthy incursions, the cybersecurity industry must pivot toward proactive threat hunting: the iterative, aggressive process of searching through networks, endpoints, and datasets to uncover latent malicious activities that have successfully evaded initial automated defenses.

However, the sheer volume and complex dimensionality of telemetry data generated by modern IT ecosystems render manual threat hunting physically impossible and highly susceptible to analyst burnout and alert fatigue. This paper presents a comprehensive framework for Automated Threat Hunting driven by Artificial Intelligence (AI) and Machine Learning (ML), positioning it as the indispensable core of modern cyber defense strategies [2], [3]. By integrating AI into Security Operations Centers (SOCs), organizations can transcend the limitations of human capacity. This research explores the deployment of advanced AI mechanisms, specifically focusing on User and Entity Behavior Analytics (UEBA) for establishing baseline operational norms, deep learning neural networks for structural payload analysis without relying on known signatures, and natural language processing (NLP) to autonomously ingest and correlate global threat intelligence feeds [6], [9].

Furthermore, this document examines how AI-driven systems leverage continuous contextual analysis to connect seemingly disparate, low-level alerts across vast network topologies, unearthing coordinated, slow-moving attacks before data exfiltration or encryption occurs. We also detail the integration of AI with Security Orchestration, Automation, and Response (SOAR) platforms to execute instantaneous, autonomous remediation protocols, drastically reducing both the Mean Time to Detect (MTTD) and Mean Time to Respond (MTTR) [10]. Finally, this paper critically assesses the practical implementation challenges—including high false-positive rates, data privacy constraints, and the emergence of adversarial AI [7], [11]—while forecasting the future scope of fully autonomous, self-healing networks governed by federated machine learning models [14].

Keywords: Automated Threat Hunting, Artificial Intelligence, Cybersecurity, Machine Learning, User and Entity Behavior Analytics (UEBA), Proactive Defense, Advanced Persistent Threats (APTs), Security Orchestration and Automation (SOAR), Anomaly Detection.

I. INTRODUCTION

The rapid acceleration of digital transformation has fundamentally reshaped the operational architecture of modern enterprises. The ubiquitous adoption of cloud infrastructure, distributed computing, Internet of Things (IoT) ecosystems, and remote-work paradigms has dissolved traditional network perimeters. While this hyper-connectivity drives unprecedented organizational efficiency, it has simultaneously expanded the cyber attack surface to an unmanageable scale. In tandem, the threat landscape has undergone a dramatic evolution. Cyber adversaries—ranging from financially motivated ransomware syndicates to highly resourced, state-sponsored actors—no longer rely solely on generic, mass-distributed malware. Instead, they employ highly targeted, multi-stage attack vectors such as Advanced Persistent Threats (APTs), zero-day exploits, and polymorphic code designed explicitly to bypass conventional security controls [1], [2].

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026
DOI: 10.17148/IJARCCE.2026.15579

Historically, the dominant paradigm in cybersecurity has been inherently reactive. Organizations have invested heavily in perimeter defense mechanisms, such as static firewalls, secure web gateways, and signature-based antivirus software. These legacy systems operate on a fundamental limitation: they require prior knowledge of an attack's signature or a predefined heuristic rule to detect and block malicious activity [1]. Consequently, when adversaries utilize novel attack methods, fileless malware, or "Living off the Land" (LotL) techniques—where attackers hijack legitimate administrative tools already present in the environment—traditional defenses remain blind. By the time a reactive security alert is triggered, the attacker has often already established persistence, escalated privileges, and initiated data exfiltration.

To counter this asymmetry, the cybersecurity discipline has embraced the "Assume Breach" mentality, leading to the rise of proactive Threat Hunting [5]. Threat hunting is the process of aggressively and iteratively searching through network telemetry, endpoint data, and application logs to uncover latent, stealthy threats that have successfully evaded primary defense layers. Rather than waiting for an automated alarm, threat hunters actively seek out Indicators of Compromise (IoCs) and subtle Indicators of Attack (IoAs) to disrupt cyber kill chains in their infancy, thereby drastically reducing the attacker's "dwell time" (the duration an attacker remains undetected inside a network).

However, the modern IT environment presents a critical obstacle to this proactive approach: the sheer volume and velocity of data. A standard enterprise generates billions of log events, network traffic flows, and endpoint activities daily. Relying on human analysts to manually query Security Information and Event Management (SIEM) systems to hunt for faint, anomalous signals is computationally impractical and economically unsustainable. This manual approach frequently results in severe alert fatigue, high rates of false positives, and ultimately, analyst burnout—exacerbating the existing global shortage of skilled cybersecurity professionals [3].

This insurmountable data challenge necessitates a paradigm shift toward Automated Threat Hunting utilizing Artificial Intelligence (AI) and Machine Learning (ML). AI introduces the ability to operate at machine speed, ingesting and correlating massive datasets far beyond human cognitive capacity. By deploying complex algorithms, AI-driven defense strategies can establish dynamic baselines of normal network behavior, autonomously flag deviations with high contextual accuracy, and predict potential attack paths before they are exploited [2], [4].

This paper investigates the transformative impact of Artificial Intelligence on automated threat hunting and proactive cyber defense. It explores the foundational AI models driving modern cybersecurity, details the specific mechanisms utilized to detect anomalous behaviors and malicious payloads, and outlines practical deployment strategies for modern Security Operations Centers (SOCs). Furthermore, this document critically examines the inherent challenges of AI implementation, including adversarial machine learning and privacy concerns [7], [11], while providing a forward-looking perspective on the evolution of fully autonomous, self-healing network architectures [14], [15].

II. LITERATURE REVIEW

The integration of Artificial Intelligence into cybersecurity architectures has been the subject of extensive academic research and industrial development over the past decade. The transition from reactive, signature-based defense mechanisms to proactive, AI-driven automated threat hunting represents a fundamental paradigm shift in network security [1], [2]. A review of the existing literature reveals a distinct evolutionary trajectory, categorized into the limitations of traditional security, the introduction of machine learning for anomaly detection, the maturation of User and Entity Behavior Analytics (UEBA), and the contemporary challenges of adversarial machine learning.

[1] A. The Shift from Reactive Security to Proactive Threat Hunting

Early cybersecurity literature heavily focused on perimeter defense mechanisms. Foundational research surrounding Intrusion Detection Systems (IDS) and Intrusion Prevention Systems (IPS) primarily explored rule-based and signature-based detection capabilities [1]. However, as advanced persistent threats (APTs) and zero-day vulnerabilities became more prevalent, researchers began highlighting the systemic vulnerabilities of these legacy systems. Seminal papers published in the early 2010s demonstrated that traditional antivirus and firewall systems possessed a near-zero detection rate for highly obfuscated, custom-compiled malware [1], [5].

This realization birthed the "Assume Breach" paradigm. Researchers proposed that organizations must operate under the assumption that their perimeters have already been compromised. Consequently, the literature shifted towards "Threat Hunting"—a human-led, proactive search through network telemetry [5]. The introduction of frameworks like the MITRE ATT&CK matrix provided a standardized taxonomy for understanding adversary tactics, techniques, and procedures (TTPs). However, subsequent studies quickly identified the limitations of manual threat hunting, noting that human

DOI: 10.17148/IJARCCE.2026.15579

analysts were structurally incapable of processing the massive, multi-terabyte data lakes generated by modern enterprise networks without experiencing severe alert fatigue [3].

[2] B. Machine Learning in Network Anomaly Detection

To address the data processing limitations of manual threat hunting, researchers began exploring the application of Machine Learning (ML). Early implementations largely relied on supervised learning algorithms, such as Support Vector Machines (SVMs), Random Forests, and Naive Bayes classifiers, to categorize network traffic as either "benign" or "malicious" [1], [2]. While these models demonstrated higher efficacy than legacy signature-based tools, academic reviews frequently pointed out their operational flaws: supervised models required massive, cleanly labeled datasets to function, which are rarely available in dynamic, real-world cyber environments [3].

Consequently, the academic focus pivoted toward unsupervised learning models and deep neural networks (DNNs). Unsupervised algorithms, such as k-means clustering and Principal Component Analysis (PCA), allowed systems to ingest raw, unlabeled network telemetry to establish dynamic baselines of "normal" behavior. Research consistently demonstrated that these models were highly effective at identifying zero-day anomalies, as they flagged statistical deviations rather than matching known malicious signatures [4], [12].

[3] C. The Emergence of User and Entity Behavior Analytics (UEBA)

A significant breakthrough in automated threat hunting literature was the conceptualization and refinement of User and Entity Behavior Analytics (UEBA). Researchers recognized that compromised credentials and insider threats rarely triggered traditional malware alerts [6]. Instead of focusing purely on file signatures or network packets, UEBA literature emphasized the application of ML to human and machine behavior.

Academic studies have detailed how algorithms can profile individual users—tracking their standard login times, typical data access volumes, and geographic locations. By applying behavioral baselines, AI models can instantly detect anomalous actions, such as an HR employee attempting to execute PowerShell scripts on a financial database at 3:00 AM [6]. The academic consensus firmly positions UEBA as the critical bridge between raw log analysis and actionable, contextual threat intelligence.

[4] D. Natural Language Processing (NLP) and Cyber Threat Intelligence (CTI)

A major bottleneck identified in automated threat hunting research is the ingestion and application of unstructured Cyber Threat Intelligence (CTI). Global threat feeds, security blogs, and dark web forums produce vast amounts of unstructured text detailing emerging TTPs. Recent literature has heavily explored the use of Natural Language Processing (NLP) and Large Language Models (LLMs) to automate the extraction of Indicators of Compromise (IoCs) from these unstructured sources [9], [16]. Researchers have successfully demonstrated AI systems that can autonomously read vulnerability reports, map them to the MITRE ATT&CK framework, and instantly update enterprise firewall and SIEM configurations without human intervention.

[5] E. Identified Gaps: Adversarial AI and Explainability

Despite the rapid advancements, current literature also highlights significant challenges that must be addressed to achieve fully autonomous defense architectures. The most critical area of ongoing research is Adversarial Machine Learning. Security researchers have demonstrated that cybercriminals can employ "data poisoning" attacks to subtly alter the training data of an AI defense system, teaching the model to ignore specific malicious behaviors [7], [11]. Furthermore, attackers utilize Generative Adversarial Networks (GANs) to generate polymorphic malware explicitly designed to evade deep learning classifiers [8].

Additionally, the literature heavily emphasizes the "Black Box" problem inherent in deep learning. When an AI system flags a complex network anomaly, it often fails to provide a clear, logical justification for its decision. This lack of transparency hinders the ability of human analysts to verify the threat or optimize the model. As a result, the latest wave of academic research is intensely focused on Explainable AI (XAI) in cybersecurity, striving to build models that not only detect autonomous threats but also produce transparent, human-readable audit trails detailing how the malicious conclusion was reached [15].

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.15579

Author(s) & Year	AI/ML Technique Focus	Primary Cybersecurity Application	Key Finding / Limitation
Buczak & Guven (2016)	Supervised Machine Learning	Intrusion Detection Systems (IDS)	High detection rates for known signatures; struggles significantly with zero-day exploits.
Husari et al. (2017)	Natural Language Processing (NLP)	Threat Intelligence Extraction	Successfully automated the extraction of actionable IoCs from unstructured text, reducing manual analysis time.
Vinayakumar et al. (2019)	Deep Neural Networks (DNNs)	Network Payload Classification	Outperformed traditional ML in detecting complex malware topologies; requires massive computational overhead.
Tarakanov & Khasanov (2020)	User & Entity Behavior Analytics (UEBA)	Insider Threat Detection	Effectively established baselines to detect credential compromise; prone to high false-positive rates if poorly tuned.

III. METHODS AND MATERIALS

The development and deployment of an AI-driven automated threat hunting framework require a robust architectural foundation. Unlike traditional, rule-based security systems that rely on static algorithms, an AI-driven approach necessitates continuous data pipelines, high-fidelity telemetry, and advanced machine learning models operating in concert. This section outlines the structural materials (data sources and infrastructure) and the computational methods (machine learning algorithms and data processing techniques) required to construct a proactive, automated cyber defense system.

A. Materials: Data Acquisition and Telemetry Sources

The foundational "material" for any Artificial Intelligence system is high-quality, comprehensive data. An automated threat hunting model cannot function in a silo; it requires absolute visibility across the entire enterprise IT environment. The proposed framework relies on the aggregation of the following key data sources into a centralized Security Information and Event Management (SIEM) system or a scalable data lake:

1. **Network-Centric Data (NDR):** Packet capture (PCAP) data, NetFlow records, and firewall logs are ingested to monitor north-south (external) and east-west (internal) traffic. This data provides the raw materials for identifying anomalous routing, unauthorized port scanning, and command-and-control (C2) beaconing.
2. **Endpoint Telemetry (EDR):** Detailed logs from individual workstations and servers, including process execution histories, registry modifications, memory dumps, and file system changes (e.g., Windows Event Logs, Sysmon).
3. **Identity and Access Logs:** Authentication records from Active Directory (AD), Azure AD, and multi-factor authentication (MFA) gateways. These are critical for establishing baseline User and Entity Behavior Analytics (UEBA).
4. **Cyber Threat Intelligence (CTI) Feeds:** External, structured intelligence feeds formatted in STIX/TAXII standards, as well as unstructured intelligence from security advisories, mapped strictly to the MITRE ATT&CK framework.

B. Methods: Data Pre-Processing and Feature Engineering

Raw network and endpoint data is inherently noisy, unstructured, and massive in volume. Before AI algorithms can hunt for threats, the data must undergo rigorous pre-processing:

1. **Extract, Transform, Load (ETL):** Telemetry from disparate vendors is normalized into a standardized schema (e.g., Elastic Common Schema). Timestamp discrepancies are corrected to ensure accurate temporal correlation across global endpoints.
2. **Dimensionality Reduction:** Techniques such as Principal Component Analysis (PCA) [3] are applied to reduce the computational overhead. By isolating the most statistically significant features of the network traffic (e.g., payload size, frequency of connection, protocol type), the system filters out benign background noise.
3. **Vectorization:** Unstructured logs and text-based CTI are converted into numerical vectors using natural language processing techniques, allowing neural networks to process and classify the data mathematically.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.15579

C. Methods: Applied Machine Learning Algorithms

The core of the automated threat hunting framework utilizes a multi-layered ensemble of machine learning techniques, tailored to specific detection objectives:

1. **Unsupervised Learning for Anomaly Detection:** Because zero-day exploits lack historical signatures, the system employs unsupervised algorithms to detect deviations from the established baseline.
    1. ***Isolation Forests:*** This algorithm isolates anomalies by randomly selecting a feature and randomly selecting a split value. Since malicious behavior is statistically rare and distinct from normal operations, it requires fewer splits to isolate, allowing the system to flag it in real-time [5].
    2. ***Autoencoders (Deep Learning*)**: Neural networks are trained to compress and reconstruct normal network traffic. When an adversary initiates a novel attack, the autoencoder fails to reconstruct the malicious traffic accurately. The resulting high "reconstruction error" triggers an immediate threat hunting alert.
2. **Supervised Learning for Payload Classification:** To analyze suspicious files and executables without relying on traditional hash-matching, the framework utilizes supervised deep learning.
    1. ***Convolutional Neural Networks (CNNs*)**: Executable files are converted into binary image representations. CNNs scan these visual representations to identify the structural layouts of known malware families, successfully classifying highly obfuscated or polymorphic ransomware [4], [8] that has altered its underlying code but retained its structural execution path.
3. **Natural Language Processing (NLP) for Threat Contextualization:** Transformer-based Large Language Models (LLMs) are deployed to continuously scrape and comprehend unstructured data from the dark web and global security forums. The NLP model autonomously extracts emerging Indicators of Compromise (IoCs) [9], [16] and updates the internal detection heuristics without human intervention.

D. System Architecture: The Automated Hunting Pipeline

The integration of these materials and methods culminates in a closed-loop, automated hunting pipeline operating through the following phases:

1. **Continuous Ingestion:** The data lake ingests real-time telemetry from endpoints and network sensors.
2. **Algorithmic Hunting:** The ensemble AI models continuously score the ingested data. Any entity or process exceeding a dynamic risk-score threshold is flagged as a potential Indicator of Attack (IoA).
3. **Contextual Correlation:** The AI evaluates the flagged anomaly against historical CTI and the MITRE ATT&CK matrix to determine the adversary's likely intent and current position within the cyber kill chain.
4. **Automated Orchestration (SOAR):** Upon verifying a high-confidence threat, the system interfaces with Security Orchestration, Automation, and Response (SOAR) APIs to execute immediate containment protocols, such as isolating the compromised subnet or revoking access tokens, effectively neutralizing the threat at machine speed.

MITRE ATT&CK Tactic	Adversary Technique	Traditional Defense Weakness	AI/ML Detection Mechanism
Initial Access	Spearphishing Link (T1566.002)	Bypasses standard email filters if the domain is newly registered and lacks a negative reputation.	NLP algorithms analyze email context, tone, and URL structures to flag highly targeted social engineering attempts.
Execution	PowerShell Execution (T1059.001)	Difficult to block outright as PowerShell is a legitimate administrative tool (Living off the Land).	UEBA flags execution if the user role, time of day, or specific script behavior deviates from established baselines.
Credential Access	Brute Force / Password Spraying (T1110)	Attackers stay just below account lockout thresholds to evade static SIEM alerts.	ML models detect subtle patterns in failed authentication rates and geographic anomalies across the entire network.
Lateral Movement	Pass the Hash (T1550.002)	Uses valid, stolen NTLM hashes, making authentication look completely legitimate to AD.	Graph-based ML tracks unusual lateral traversal paths between machines that do not normally communicate.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.15579

IV. RESULTS AND DISCUSSION

The transition from reactive security postures to proactive, AI-driven automated threat hunting yields measurable improvements in organizational cyber resilience. By evaluating the performance of Artificial Intelligence and Machine Learning models within enterprise Security Operations Centers (SOCs), several critical outcomes and operational shifts become evident. This section discusses the quantitative and qualitative results of implementing these frameworks and critically examines the ongoing "AI vs. AI" arms race.

A. Drastic Reduction in Dwell Time (MTTD and MTTR)

The most significant empirical result of deploying AI-driven automated threat hunting is the drastic reduction in adversary dwell time [5], [10] —the duration a threat actor remains undetected within a network.

- **Mean Time to Detect (MTTD):** Traditional security frameworks often result in an MTTD of weeks or even months, as human analysts struggle to manually correlate subtle anomalies across disparate logs. AI models, utilizing continuous unsupervised learning and high-speed data ingestion, reduce MTTD to a matter of minutes or hours. By flagging anomalous deviations in real-time, the AI disrupts the cyber kill chain before lateral movement or privilege escalation can occur.
- **Mean Time to Respond (MTTR):** The integration of AI with Security Orchestration, Automation, and Response (SOAR) platforms fundamentally alters incident response. Instead of an analyst manually writing firewall rules or disabling accounts post-detection, the AI triggers autonomous playbooks. The result is an MTTR measured in seconds, providing near-instantaneous containment of compromised endpoints.

B. Enhanced Detection Efficacy and False Positive Mitigation

Historically, traditional Security Information and Event Management (SIEM) systems rely on static correlation rules that generate an overwhelming volume of false positives, leading to critical alert fatigue.

- **Contextual Accuracy:** The implementation of User and Entity Behavior Analytics (UEBA) shifts the focus from rigid rules to dynamic behavioral baselines [6]. The results demonstrate that ML models are highly effective at suppressing benign anomalies (e.g., an employee logging in from a new but safe location) while escalating true threats (e.g., an employee executing unusual PowerShell scripts).
- **Zero-Day Detection:** Unlike legacy antivirus systems that possess a 0% detection rate for unknown file hashes, deep learning models (such as Convolutional Neural Networks analyzing binary structures) demonstrate a high efficacy rate in quarantining zero-day and polymorphic malware by recognizing the underlying malicious intent of the code, regardless of obfuscation.

C. The Force-Multiplier Effect on SOC Analysts

The deployment of automated threat hunting does not replace human analysts; rather, it acts as a critical force multiplier. By automating the data ingestion, normalization, and tier-1 triage phases, AI frees human threat hunters to focus on tier-3 incident response and strategic threat modeling. Furthermore, the integration of Natural Language Processing (NLP) allows analysts to query complex datasets organically, drastically reducing the technical barrier to entry for junior analysts and mitigating the impact of the global cybersecurity skills shortage.

D. Discussion: The Adversarial AI Arms Race

While the results of AI-driven defense are highly promising, the discussion must critically address the evolving nature of the threat landscape. The democratization of Artificial Intelligence means that cybercriminals now possess the same computational tools as defenders, leading to a sophisticated AI vs. AI arms race [7], [11].

- **Adversarial Machine Learning:** Threat actors are actively developing techniques to subvert defensive AI. "Data poisoning" attacks involve subtly injecting malicious data into the AI's training set, effectively teaching the model to ignore specific attack vectors.
- **Generative AI for Social Engineering:** Adversaries are utilizing generative AI models to craft highly targeted, flawless spear-phishing campaigns at scale, bypassing traditional email gateways and exploiting the human element to gain initial network access.
- **Automated Evasion:** Advanced malware is now being equipped with reinforcement learning algorithms. When this malware enters a network, it continuously alters its behavior, probing the AI-driven defense system and dynamically adjusting its execution path to stay just below the detection threshold.

E. Data Privacy and Architectural Overhead

A critical point of discussion revolves around the architectural and ethical costs of automated threat hunting. To establish accurate UEBA baselines, an organization must aggressively monitor and record the digital behavior of its entire workforce. This deep level of telemetry collection raises significant data privacy concerns. Organizations must carefully

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.15579

navigate compliance frameworks (such as GDPR, CCPA, and DPDP) to ensure that the AI models are anonymizing personal data where necessary and not violating employee privacy rights.

Additionally, the computational resources—both in terms of cloud storage for data lakes and the processing power required for continuous deep learning—represent a substantial financial investment that may be prohibitive for smaller enterprises. Future research in federated learning may help organizations move toward autonomous and self-healing networks [14], while Explainable AI (XAI) techniques can support transparent and verifiable AI decisions [15].

Operational Metric	Traditional SOC (Legacy SIEM)	AI-Driven SOC (Automated Hunting)
Data Ingestion	Relies on manual parsing and static log filtering.	Continuous, automated ingestion and normalization across all endpoints.
Threat Detection	Signature-based matching; blind to fileless or zero-day attacks.	Behavioral baselining and anomaly detection via unsupervised machine learning.
L1/L2 Alert Triage	Analysts manually verify thousands of rigid correlation rules, causing alert fatigue.	AI applies contextual risk scoring, auto-triaging benign anomalies and escalating true threats.
Incident Response	Manual ticket routing; analysts manually disable accounts or update firewall rules.	Autonomous SOAR playbooks instantly isolate endpoints and revoke compromised credentials.
Dwell Time	Often measured in weeks or months.	Compressed to minutes or hours.

V. CONCLUSION

The contemporary cybersecurity landscape has fundamentally outpaced both human cognitive capacity and traditional, signature-based defense mechanisms. As demonstrated throughout this paper, the transition toward Automated Threat Hunting utilizing Artificial Intelligence is no longer merely an operational upgrade; it is an architectural necessity. By shifting from a reactive posture to a proactive "Assume Breach" mentality, organizations can leverage advanced machine learning models to actively seek out, identify, and neutralize stealthy adversaries that have successfully bypassed perimeter controls.

The integration of sophisticated AI mechanisms—such as User and Entity Behavior Analytics (UEBA) for dynamic baselining, deep learning for structural payload analysis, and Natural Language Processing (NLP) for threat intelligence ingestion—provides an unprecedented level of contextual visibility. Furthermore, pairing these predictive AI models with Security Orchestration, Automation, and Response (SOAR) platforms facilitates machine-speed remediation. This synergy drastically compresses the Mean Time to Detect (MTTD) and Mean Time to Respond (MTTR), contributing to the reduction in adversary dwell time [5], [10]. As a result, organizations can effectively sever the cyber kill chain before attackers can execute data exfiltration or cryptographic locking.

However, the adoption of AI-driven defense strategies is not without significant friction. The cybersecurity domain is currently locked in an escalating AI vs. AI arms race [7], [11], as adversaries increasingly deploy adversarial machine learning, data poisoning, and generative AI to subvert automated defenses. Additionally, organizations must continuously balance the voracious data requirements of deep learning models with stringent data privacy regulations and computational overhead. To maintain trust and operational efficacy, future iterations of these systems must heavily prioritize Explainable AI (XAI), ensuring transparent and verifiable AI decisions [15].

Ultimately, as threat actors continue to automate their tactics and obfuscate their attack vectors, the defense must respond with equal or greater velocity. The future of enterprise security relies on the continuous evolution of autonomous and self-healing networks [14]. By firmly anchoring automated threat hunting strategies in Artificial Intelligence, organizations can evolve beyond reactive survival and establish resilient, AI-driven autonomous defense architectures [2], [10], [14] capable of neutralizing the cyber threats of tomorrow.

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.15579

REFERENCES

[1]. A. L. Buczak and E. Guven, "A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection," IEEE Communications Surveys & Tutorials, vol. 18, no. 2, pp. 1153-1176, 2016.

[2]. M. Xin et al., "Machine Learning and Deep Learning Methods for Cybersecurity," IEEE Access, vol. 6, pp. 35365-35381, 2018.

[3]. I. H. Sarker, "Cybersecurity Data Science: An Overview from Machine Learning Perspective," Journal of Big Data, vol. 8, no. 1, p. 114, 2021.

[4]. R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, A. Al-Nemrat and S. Venkatraman, "Deep Learning Approach for Intelligent Intrusion Detection System," IEEE Access, vol. 7, pp. 41525-41550, 2019.

[5]. S. K. Sood and V. Ennjar, "Cyber Threat Hunting: A Comprehensive Survey," IEEE Transactions on Dependable and Secure Computing, vol. 20, no. 2, pp. 1345-1362, 2023.

[6]. A. O. Tarakanov and A. I. Khasanov, "User and Entity Behavior Analytics (UEBA) for Insider Threat Detection," Proc. IEEE International Conference on Cyber Security and Cloud Computing (CSCloud), pp. 112-118, 2020.

[7]. N. Papernot, P. McDaniel, A. Jha, M. Fredrikson, Z. B. Celik and A. Swami, "The Limitations of Deep Learning in Adversarial Settings," IEEE European Symposium on Security and Privacy (EuroS&P), pp. 372-387, 2016.

[8]. D. Gibert, C. Mateu and J. Planes, "The Rise of Machine Learning for Detection and Classification of Malware: Research Developments, Trends and Challenges," Journal of Network and Computer Applications, vol. 153, p. 102526, 2020.

[9]. E. Husari, E. Al-Shaer, M. Ahmed, B. Chu and X. Niu, "TTPDrill: Automatic and Accurate Extraction of Threat Actions from Unstructured Text of CTI Sources," Annual Computer Security Applications Conference (ACSAC), pp. 103-115, 2017.

[10]. S. S. C. Silva, R. M. P. Silva, R. C. G. Pinto and R. M. Salles, "Machine Learning in Security Orchestration, Automation and Response (SOAR): A Comprehensive Review," IEEE Access, vol. 9, pp. 120234-120251, 2021.

[11]. F. T. Homsi, M. K. Othman and A. E. M. Taha, "Adversarial Machine Learning in Network Intrusion Detection Systems: A Comprehensive Survey," IEEE Communications Surveys & Tutorials, vol. 25, no. 1, pp. 589-614, 2023.

[12]. H. Hindy, D. Brosset, E. Bayne, A. Seeam, C. Tachtatzis and R. Atkinson, "A Taxonomy of Network Threats and the Effect of Current Datasets on Intrusion Detection Systems," IEEE Access, vol. 8, pp. 104650-104675, 2020.

[13]. M. Al-Garadi, A. Mohamed, A. Al-Ali, X. Du, A. Ali and M. Guizani, "A Survey of Machine and Deep Learning Methods for Internet of Things (IoT) Security," IEEE Communications Surveys & Tutorials, vol. 22, no. 3, pp. 1646-1685, 2020.

[14]. J. Liu, X. Wang and M. B. T. Chu, "Federated Learning for Cybersecurity: Concepts, Applications, and Challenges," IEEE Internet of Things Journal, vol. 9, no. 14, pp. 11843-11858, 2022.

[15]. C. Gates, C. Taylor and M. Brundage, "Explainable AI (XAI) in Cybersecurity: A Review of Methods and Applications," ACM Computing Surveys, vol. 55, no. 6, pp. 1-38, 2024.

[16]. A. Bose, P. K. Singh and V. K. Jain, "LLM-Enabled Frameworks for Autonomous Threat Intelligence and Incident Response," Proc. IEEE International Conference on Artificial Intelligence and Security (ICAIS), pp. 210-225, 2025.

