Digitus : Journal of Computer Science Applications
E-ISSN : 3031-3244
Volume. 3, Issue 2, April 2025
Page No: 78-91

Toward Resilient Networks: AI and Deep Learning Strategies for Intrusion Detection

Lia Marthalia
Universitas Jayabaya, Indonesia
Correspondent: lia.marthalia20@gmail.com

Received : February 27, 2025
Accepted : April 12, 2025
Published : April 30, 2025

Citation: Marthalia, L. (2025). Toward Resilient Networks: AI and Deep Learning Strategies for Intrusion Detection. Digitus : Journal of Computer Science Applications, 3 (2), 78-91.

ABSTRACT: As cyber threats become more sophisticated and pervasive, the demand for advanced Network Intrusion Detection Systems (NIDS) has increased dramatically. This narrative review investigates the application of Artificial Intelligence (AI) and Deep Learning (DL) techniques in enhancing NIDS performance, aiming to address the limitations of conventional rule-based systems. The literature was systematically retrieved from reputable databases such as Scopus and IEEE Xplore using keywords including "Network Intrusion Detection," "Deep Learning," and "Cybersecurity." Inclusion criteria focused on peer-reviewed studies that utilized AI models for intrusion detection, particularly within complex domains like IoT and smart grids. The review identifies CNN, LSTM, and DNN as the dominant AI models employed in modern NIDS, achieving detection accuracies ranging from 88% to 99% across benchmark datasets such as NSL-KDD and CICIDS2017. These models also demonstrate reduced false-positive rates and enhanced detection of zero-day attacks. Despite their promise, challenges remain, including regulatory constraints, computational limitations in edge devices, and difficulties in model interpretability. Systemic organizational factors—such as leadership commitment, IT infrastructure readiness, and cybersecurity culture—further affect successful implementation. This study highlights the potential of AI-based NIDS as a strategic approach to cybersecurity enhancement and proposes solutions including Explainable AI, hybrid model designs, and federated learning. The findings support further research into cross-domain applications, model transparency, and real-time scalability to unlock the full potential of intelligent intrusion detection systems..

Keywords: Network Intrusion Detection, Deep Learning, Cybersecurity, Artificial Intelligence, Anomaly Detection, IoT Security, Explainable AI.

INTRODUCTION

The proliferation of digital technologies and the expansion of interconnected systems have brought about a new era of connectivity and convenience. However, this advancement has also introduced unprecedented security risks, particularly in the realm of network intrusion. Modern computer networks, integral to both public and private infrastructures, face a wide array of

complex and evolving cyber threats. Traditional security mechanisms, although foundational, are increasingly inadequate in addressing these challenges. Network Intrusion Detection Systems (NIDS) have emerged as critical tools in identifying and mitigating unauthorized access or anomalies in network behavior. Recent studies underscore the urgency to enhance NIDS capabilities in response to the rising sophistication of cyber-attacks (Mambwe et al., 2024; Rahman et al., 2025).

The emergence of Artificial Intelligence (AI) and Deep Learning (DL) has transformed the landscape of cybersecurity. These technologies offer adaptive learning mechanisms that significantly outperform rule-based detection models, particularly in identifying complex patterns and anomalies within vast network data. According to Habeeb and Babu (2022), DL models have demonstrated superior capabilities in processing and classifying intricate network traffic, paving the way for more robust and intelligent intrusion detection. Among the popular architectures, Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks are frequently employed due to their proficiency in recognizing temporal and spatial patterns in network data (Popli et al., 2021; Rahman et al., 2025).

Recent global reports have indicated a sharp escalation in cyber threats, particularly targeting Internet of Things (IoT) systems and industrial network infrastructures. These attacks have led to substantial financial and reputational damages across various sectors. Alrayes et al. (2023) and Moura et al. (2023) note that the frequency and severity of such incidents have increased annually, underlining the need for more resilient detection systems. The proliferation of internet-connected devices amplifies the potential attack surface, exacerbating the difficulty in maintaining a secure digital environment. Consequently, there is growing consensus among experts and policymakers regarding the critical need to invest in advanced network protection strategies, particularly those leveraging AI and DL (Dapel et al., 2023).

In parallel with this surge in cyber threats is the exponential growth in network data volume. The high dimensionality and sheer size of this data complicate traditional analysis techniques and pose significant challenges to maintaining detection accuracy. As attackers adopt more covert strategies, including steganographic techniques and polymorphic malware, conventional systems often fail to identify malicious behavior with sufficient precision (Mambwe et al., 2024). Compounding the issue is the persistent problem of false positives, which not only drain system resources but also lead to desensitization of security personnel, ultimately reducing system effectiveness (Rahman et al., 2025).

One of the most formidable challenges in modern NIDS is the detection of zero-day attacks. These attacks exploit previously unknown vulnerabilities, rendering signature-based methods largely ineffective. Adaptive AI models, particularly those incorporating unsupervised and semi-supervised learning, offer promising solutions to this problem by identifying deviations from normal behavior without relying on predefined signatures (Popli et al., 2021; Mambwe et al., 2024). However, these models are not without their limitations, especially regarding transparency and interpretability.

The integration of AI into NIDS also introduces concerns surrounding model explainability and domain specificity. While many studies have achieved high accuracy rates using generic deep learning models, they often fail to contextualize their findings within specific operational environments. Alsulami (2024) and Salem et al. (2024) highlight that most implementations utilize standardized datasets such as KDD'99 and CICIDS2017, which, although useful, do not reflect the variability and complexity of real-world traffic. Furthermore, the reliance on "black box" models hinders practical deployment in mission-critical settings where decision traceability is essential (Neupane et al., 2022; Larriva-Novo et al., 2023).

This literature review seeks to address these gaps by systematically evaluating the application of AI and DL techniques in the development and deployment of advanced NIDS. Specifically, it aims to identify which methodologies are most effective in detecting sophisticated and previously unseen attack vectors. By examining recent empirical studies and theoretical advancements, this review will shed light on the capabilities, limitations, and potential improvements in current NIDS frameworks (Krishnaveni et al., 2024; Vinayakumar et al., 2019; Alabbadi & Bajaber, 2025). Additionally, the study will assess the degree to which existing models can be adapted to various operational environments and threat landscapes.

The scope of this review extends to multiple sectors and geographic contexts, with particular emphasis on underrepresented domains. While significant attention has been given to IT infrastructure, sectors such as healthcare, energy, transportation, and industrial control systems remain relatively underexplored in the context of AI-based intrusion detection. For instance, Rajapaksha et al. (2023) and Zubair et al. (2022) have noted the critical importance of cybersecurity in health informatics and energy management systems, yet research tailored to these contexts is still sparse. Furthermore, systems such as railway signaling and Supervisory Control and Data Acquisition (SCADA) networks require customized detection frameworks due to their unique architecture and performance constraints (Qi & Wang, 2024; Kazmi et al., 2023; Dapel et al., 2023).

By compiling and synthesizing findings across diverse applications, this review aims to provide a comprehensive understanding of how AI and DL can enhance the efficacy and scope of intrusion detection mechanisms. The review also intends to illuminate emerging research trajectories, highlight best practices, and identify critical areas where further investigation is necessary. It serves both academic researchers and practitioners seeking to develop or implement advanced NIDS tailored to specific operational needs and evolving threat scenarios.

Ultimately, this study underscores the need for a paradigm shift in how network security is approached, advocating for the integration of intelligent, context-aware, and explainable AI models. Such models not only promise heightened detection accuracy but also the interpretability and adaptability required for deployment in complex, dynamic environments. In doing so, this review contributes to the broader effort to fortify digital infrastructures against the rising tide of cyber threats in an increasingly connected world.

METHOD

This study adopted a structured and rigorous literature review methodology to explore the use of Artificial Intelligence (AI) and Deep Learning (DL) in Network Intrusion Detection Systems (NIDS). The goal was to collect, assess, and synthesize academic studies that discuss how AI technologies are being leveraged to detect and prevent cyber intrusions in computer networks. Given the interdisciplinary nature of the subject, the literature search and selection process was conducted with a high degree of precision to ensure that only the most relevant and high-quality studies were included.

To begin the process, a comprehensive search was conducted across several academic databases known for hosting peer-reviewed and high-impact publications in the field of computer science and cybersecurity. The databases included Scopus, IEEE Xplore, ACM Digital Library, SpringerLink, and Google Scholar. These sources were selected based on their reputability, breadth of indexed journals, and relevance to the fields of AI, cybersecurity, and computer networks. The search was not limited to a single database to avoid bias and to ensure an inclusive retrieval of interdisciplinary studies.

The search strategy relied on a combination of carefully selected keywords and Boolean operators to refine the results. The primary keywords used were "Network Intrusion Detection," "Deep Learning," "Artificial Intelligence," "Cybersecurity," "Intrusion Detection System (IDS)," and "Anomaly Detection." These terms were chosen because they directly relate to the technologies and systems under investigation. In addition to these primary keywords, keyword combinations such as "AI in Network Security," "Deep Learning for Intrusion Detection," and "Machine Learning Techniques for Cybersecurity" were also applied. These combinations helped retrieve articles that might not use the exact same phrasing but still contribute relevant insights to the research focus.

The initial search yielded several thousand articles, necessitating a clear and systematic filtering process to determine which studies would be included in the final analysis. A two-tiered screening process was employed. The first tier involved a title and abstract review to eliminate studies that were clearly unrelated to the research topic. Articles that were solely theoretical without empirical data or that focused on unrelated domains, such as general AI applications without relevance to security, were excluded at this stage. The second tier involved a full-text review to assess methodological clarity, relevance, and the extent of AI application within the study.

To ensure academic rigor, the selection of articles was guided by strict inclusion and exclusion criteria. The inclusion criteria were as follows: first, studies had to be published in peer-reviewed journals to guarantee a certain standard of scholarly quality and reliability, as emphasized by Rahman et al. (2025) and Alabbadi and Bajaber (2025). Second, the research must focus explicitly on the application of AI, machine learning, or deep learning techniques within the context of intrusion detection systems. Third, studies needed to demonstrate topic alignment, ensuring that they addressed aspects of AI in network security or system defense, as validated by Habeeb and Babu (2022) and Arreche and Abdallah (2025). Lastly, methodological transparency was a key

factor; only studies that clearly articulated their experimental design, dataset, and evaluation metrics were considered, consistent with criteria laid out by Sayegh et al. (2024) and Houda et al. (2022).

On the other hand, exclusion criteria helped eliminate studies that did not meet the necessary standards or relevance. Studies that did not utilize AI-based methods, such as those that focused exclusively on rule-based or signature-based intrusion detection techniques, were not included (Zhou et al., 2021; Choudhury & Azad, 2022). Research that failed to address cybersecurity or network protection adequately, or that had weak alignment with the primary themes of the study, was similarly excluded (Zubair et al., 2022; Rajapaksha et al., 2023). Furthermore, any study lacking a robust methodological framework, including insufficient detail on how results were obtained, unclear objectives, or the use of outdated datasets without justification, was considered ineligible for inclusion (Habeeb & Babu, 2022; Neupane et al., 2022). Another critical consideration was timeliness: to maintain the relevance of this review in the face of fast-evolving technologies, the study primarily included articles published within the last five years.

The final selection included studies that met all inclusion criteria and avoided all exclusionary pitfalls. These articles were subjected to in-depth content analysis to extract information about their research objectives, AI methodologies employed, data sources, evaluation metrics, and reported outcomes. The studies ranged across multiple research designs, including experimental implementations of AI-based NIDS, performance comparisons between traditional and intelligent detection systems, and hybrid models combining different machine learning approaches. Though randomized controlled trials are rare in this domain, most of the included articles were empirical studies that utilized either simulated datasets (e.g., KDD'99, CICIDS2017) or real-world traffic data from organizational network environments. Some studies also explored hybrid methodologies that integrated AI with rule-based or statistical models.

Additionally, the quality of each study was evaluated based on several methodological parameters, including the rigor of experimental setup, comprehensiveness of result reporting, reproducibility of the methods, and the clarity of limitations discussed. Where available, the impact factor of the publishing journal and the citation count of the article were used as supplementary indicators of academic significance. Articles presenting explainable AI frameworks were particularly emphasized, given the increasing demand for transparency and interpretability in cybersecurity applications (Larriva-Novo et al., 2023).

Throughout the review process, the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines were informally adapted to structure the inclusion process and ensure transparency and replicability of the review. Although this is a narrative review rather than a full systematic review, the application of PRISMA principles helped maintain consistency and minimize bias during article selection.

This methodological approach ensured that the review provides a well-rounded and current overview of how AI and DL technologies are shaping the future of intrusion detection systems. It allowed for the synthesis of diverse perspectives and technological approaches, offering insights that are both theoretically sound and practically applicable. By maintaining a consistent framework for inclusion, exclusion, and evaluation, this study contributes to the growing body of knowledge

surrounding the application of AI in cybersecurity and identifies promising directions for future research and innovation.

RESULT AND DISCUSSION

The review of contemporary literature on Network Intrusion Detection Systems (NIDS) integrated with Artificial Intelligence (AI) and Deep Learning (DL) reveals a nuanced understanding of architectural frameworks, algorithmic strategies, contextual applications, and performance outcomes. This section presents the synthesized findings structured into four thematic areas: NIDS architecture, deep learning algorithms, application contexts (IoT, smart grid, healthcare), and evaluation of system effectiveness.

Recent research into intrusion detection architectures demonstrates a growing departure from traditional centralized models toward more flexible, scalable, and intelligent systems. Centralized architectures, although still in use due to their simplicity and ease of deployment in constrained environments, are increasingly being overshadowed by distributed and hybrid approaches (Habeeb & Babu, 2022). Federated models allow for decentralized processing and learning, minimizing data transfer overhead and enhancing privacy. Hybrid models have garnered particular interest by combining the deterministic capabilities of rule-based systems with the adaptability of machine learning algorithms. These architectures can adaptively manage data input and detection workflows across diverse environments, offering a higher degree of resilience and scalability (Aldaej et al., 2023).

Empirical findings show that architectural choice directly affects the system's detection capabilities. Long Short-Term Memory (LSTM)-based architectures consistently demonstrate high accuracy levels, exceeding 92% in tests using standard benchmark datasets like NSL-KDD and CICIDS2017 (Sayegh et al., 2024; Souza et al., 2024). These architectures are especially effective at identifying temporal patterns in network traffic, which are crucial for detecting sophisticated and sequentially staged attacks. Convolutional Neural Networks (CNN), on the other hand, have performed exceptionally well in spatial analysis of network packet data. CNN-based models have achieved accuracy rates as high as 96% under certain testing conditions, highlighting their utility in environments like the Internet of Things (IoT), where high-frequency traffic requires rapid and precise processing (Krishnaveni et al., 2024; Deshmukh & Ravulakollu, 2024).

In examining the algorithmic landscape of deep learning applied to NIDS, several patterns emerge. CNN, LSTM, and Deep Neural Networks (DNN) form the backbone of most high-performing detection systems. CNN models are praised for their ability to detect complex spatial features within packet streams, making them particularly suited to environments characterized by high data variability (Deshmukh & Ravulakollu, 2024). LSTM networks excel in identifying sequential attack behaviors, especially in detecting distributed denial of service (DDoS) patterns, due to their temporal memory structure (Sayegh et al., 2024). DNNs offer a versatile framework capable of integrating advanced interpretability tools such as SHAP (SHapley Additive exPlanations), enhancing the transparency of detection outcomes and aiding in forensic analysis (Houda et al., 2022).

From a performance perspective, CNN and LSTM models consistently outperform other architectures. LSTM-based systems report accuracy rates of 92% or higher, while CNN models have achieved up to 99% accuracy when combined with data augmentation techniques and optimized training protocols (Sayegh et al., 2024). DNNs exhibit a broader range of accuracy, typically between 88% and 92%, depending on architectural complexity and dataset quality (Mansour, 2022). These findings underscore the importance of aligning algorithmic selection with the specific requirements of the implementation environment, especially concerning data characteristics and desired explainability.

Contextual adaptation of AI-based NIDS is increasingly vital, particularly in constrained or highly regulated sectors such as IoT networks, smart grids, and healthcare. In IoT environments, where devices often have limited processing power and memory, system designers have employed techniques such as feature selection and dimensionality reduction to reduce computational overhead while preserving detection accuracy (Sayegh et al., 2024). Lightweight models using simplified CNN or LSTM architectures are optimized to operate effectively on edge devices. Studies indicate that through careful pruning and compression, deep learning models can maintain high detection performance even in resource-constrained settings (Alabbadi & Bajaber, 2025).

Healthcare applications introduce additional layers of complexity. Devices in this domain often interface directly with patients, requiring adherence to strict regulatory frameworks such as HIPAA. This imposes constraints on how data is processed, stored, and transmitted, posing challenges to AI integration. Furthermore, the sensitivity of data demands exceptionally low false positive rates to prevent unnecessary interventions or operational disruptions (AboulEla et al., 2024). Similar constraints exist in the energy sector. Smart grids, due to their interconnected nature, are vulnerable to large-scale cyberattacks with the potential to disrupt critical services. NIDS in this context must handle heterogeneous data inputs from various subsystems while complying with security and interoperability standards. Integrated solutions capable of managing variable data types and communication protocols are still in early stages of development, pointing to an urgent research need (Mounir et al., 2025; Hussain et al., 2020).

Evaluation metrics in AI-based NIDS research provide crucial insights into system reliability and real-world applicability. Accuracy remains the most commonly reported metric, with numerous studies documenting rates between 95% and 99% using datasets such as CICIDS-2017 and NSL-KDD (Alsirhani et al., 2025; Thapa et al., 2020). However, accuracy alone can be misleading in datasets with class imbalance. Hence, complementary metrics such as precision, recall, F1 score, and false positive rate (FPR) are also widely reported. Lower FPRs are particularly important in practical deployments, as high rates of false alerts can overwhelm security teams and degrade trust in the system. Advanced DL models like LSTM and hybrid CNN-LSTM configurations have demonstrated FPRs under 5%, whereas many traditional models exhibit FPRs exceeding 10% (Sharmila & Nagapadma, 2023; Mallampati & Seetha, 2024).

Comparative analysis across studies suggests a global trend toward favoring deep learning approaches over traditional machine learning or rule-based systems. While countries such as the United States, China, and members of the European Union lead in the development and testing of these advanced models, there is growing interest in AI-based NIDS research from emerging economies. However, disparities exist in terms of access to large-scale datasets, computational

infrastructure, and regulatory frameworks. This global perspective emphasizes the need for collaborative frameworks that can support model transferability and generalization across regions with varying technical capacities.

In summary, the results of this narrative review highlight significant advancements in AI-driven intrusion detection across multiple dimensions. Architectural innovations, particularly hybrid and federated models, have enhanced scalability and adaptability. Deep learning algorithms such as CNN and LSTM have proven highly effective in both accuracy and speed. Real-world implementations in sectors such as IoT, smart grids, and healthcare demonstrate the versatility and challenges of AI-based NIDS. Finally, evaluation frameworks confirm the superior performance of these systems in reducing false positives and improving detection precision. While much progress has been made, there remains a strong imperative to address context-specific constraints and ensure equitable access to AI technologies in cybersecurity worldwide.

The rise of artificial intelligence (AI) and deep learning techniques has brought a transformative shift in the development and implementation of Network Intrusion Detection Systems (NIDS). Compared to conventional rule-based systems, AI-based NIDS show considerable advantages in their ability to detect complex and evolving threats with greater accuracy and reduced false positives. As noted by Sayegh et al. (2024) and Habeeb & Babu (2022), AI-powered NIDS can maintain detection accuracies above 92%, while minimizing false positives to below 5%, a substantial improvement over traditional models that often exceed 10% in false alarm rates. These findings are reinforced by studies that demonstrate the superior performance of CNN and LSTM models in identifying sophisticated or previously unknown threats, including zero-day attacks (Mounir et al., 2025; Alabbadi & Bajaber, 2025).

The ability of AI models to generalize from learned traffic patterns and respond to novel attack vectors makes them highly adaptive in dynamic environments. Krishnaveni et al. (2024) emphasized that conventional approaches remain static and rule-dependent, struggling to keep pace with the rapid evolution of cyber threats. This adaptability positions AI-based NIDS as a key component in modern cybersecurity frameworks.

Despite their technical promise, the deployment of AI-driven NIDS systems is deeply influenced by systemic and organizational factors. Leadership commitment and managerial support are crucial enablers, as highlighted by Habeeb & Babu (2022). Organizations willing to invest in AI infrastructure and capacity building are more likely to implement effective and sustainable solutions. Moreover, Shtayat et al. (2023) emphasize that successful implementation depends on an organization's technological infrastructure, including computing power and network resources, which must support the processing requirements of AI algorithms.

Another critical systemic factor is organizational culture. According to Souza et al. (2024), institutions that promote cybersecurity awareness and foster openness to technological innovation tend to experience smoother transitions to AI-based systems. Conversely, organizations with rigid hierarchical structures or resistance to change may encounter difficulties in integrating these solutions. The importance of regulatory compliance also plays a pivotal role, particularly in sectors like healthcare and energy. As stated by AboulEla et al. (2024), AI-based NIDS in these domains must adhere to stringent data protection and operational continuity regulations, such as HIPAA,

GDPR, or industry-specific standards. These constraints necessitate additional layers of auditing, transparency, and accountability within AI systems.

Addressing the limitations of AI-driven NIDS systems has become a focal point in current research. One of the major concerns lies in the interpretability of AI models. Black-box models often provide limited insights into their decision-making processes, making it challenging for security analysts to understand the rationale behind alerts. This issue has motivated the integration of Explainable AI (XAI) methods, including LIME and SHAP, which help elucidate the features contributing to classification outcomes. Yang et al. (2025) and Alabbadi & Bajaber (2025) argue that such methods not only enhance human trust in the system but also facilitate system debugging and improvement.

Furthermore, the availability and quality of training data significantly affect model robustness. Torre et al. (2025) propose the adoption of data augmentation techniques and federated learning to diversify training datasets while preserving data privacy. This approach is particularly valuable in distributed environments, such as IoT networks and cross-sector collaborations, where centralized data collection may be infeasible or legally restricted.

Another promising avenue involves the development of hybrid models that combine AI-based detection with traditional rule-based techniques. As discussed by Mahbooba et al. (2021) and Habeeb & Babu (2022), such hybrid architectures leverage the high detection sensitivity of AI and the contextual rule enforcement of traditional systems. These models can potentially reduce detection latency while enhancing accuracy, thereby improving real-time responsiveness in operational settings.

The feasibility of AI implementation in organizational contexts also demands attention. Otoum et al. (2021) advocate for comprehensive feasibility studies and risk analyses before the adoption of NIDS solutions. These assessments must account for the organization's threat landscape, budget constraints, compliance obligations, and technical maturity. Without this groundwork, even high-performing AI systems may fail to deliver expected security benefits due to misalignment with organizational needs.

In practice, these limitations and recommendations reveal that the journey toward integrating AI in NIDS is not purely technical. It is also strategic, requiring multi-level collaboration across technical, managerial, and regulatory domains. Future research must therefore explore not only the development of novel AI algorithms but also their practical deployment within real-world organizational ecosystems.

Research into cross-sector adaptation, especially in healthcare and smart energy systems, remains relatively limited. Although preliminary studies have demonstrated feasibility, full-scale implementation studies are needed to examine long-term performance, maintenance costs, and regulatory adaptability. Additionally, comparative studies across different national contexts would be valuable in assessing how policy environments, cultural attitudes toward technology, and cybersecurity norms influence the success of AI-based NIDS.

Another underexplored area is the ethical dimension of automated threat detection. As AI takes on more decision-making roles in cybersecurity, concerns over accountability, bias in model training, and potential misuse of detection outputs are likely to grow. These challenges necessitate

interdisciplinary collaborations that include legal scholars, ethicists, and policymakers alongside technologists.

In sum, while the integration of AI and deep learning into NIDS offers a transformative leap in intrusion detection capabilities, its practical realization hinges on a confluence of technical innovation, organizational readiness, and regulatory alignment. The current literature provides a robust foundation, but also reveals critical gaps that must be addressed through empirical research, interdisciplinary dialogue, and policy development.

CONCLUSION

This narrative review has explored the evolution, effectiveness, and limitations of Network Intrusion Detection Systems (NIDS) enhanced by Artificial Intelligence (AI) and Deep Learning (DL) algorithms. The analysis of current literature reveals a consistent pattern: AI-based NIDS outperform traditional rule-based systems in detecting complex and previously unseen cyber-attacks. With detection accuracy often exceeding 92% and lower false-positive rates, especially when using CNN and LSTM models, AI integration significantly enhances threat responsiveness and adaptability. These technologies are particularly impactful in dynamic environments such as IoT, smart grids, and healthcare systems, where conventional methods struggle to maintain performance.

However, the implementation of AI-driven systems is not without obstacles. Organizational factors such as managerial commitment, IT infrastructure, data availability, regulatory compliance, and a culture of security profoundly influence adoption success. Moreover, challenges related to model interpretability, dataset limitations, and system integration persist.

To address these barriers, the integration of Explainable AI (XAI) techniques, adoption of data augmentation and federated learning, development of hybrid detection architectures, and comprehensive feasibility studies are recommended. Policymakers and stakeholders must develop supportive frameworks that ensure secure, scalable, and ethical deployment of AI-based NIDS. Future research should investigate cross-domain adaptability, real-time performance under large-scale network conditions, and model transparency in critical sectors. Ultimately, AI-enhanced NIDS, particularly when using advanced deep learning algorithms, represent a transformative strategy to strengthen cyber resilience across interconnected systems.

REFERENCE

AboulEla, S., Ibrahim, N., Shehmir, S., Yadav, A., & Kashef, R. (2024). Navigating the cyber threat landscape: an in-depth analysis of attack detection within iot ecosystems. *AI, 5(2)*, 704–732. https://doi.org/10.3390/ai5020037

Alabbadi, A. and Bajaber, F. (2025). An intrusion detection system over the iot data streams using explainable artificial intelligence (xai). *Sensors, 25(3)*, 847. https://doi.org/10.3390/s25030847

Alayres, F., et al. (2023). Optimal fuzzy logic enabled intrusion detection for secure iot-cloud environment. *Computers Materials & Continua, 74(3)*, 6737–6753. https://doi.org/10.32604/cmc.2023.032591

Alabbadi, A. and Bajaber, F. (2025). An intrusion detection system over the iot data streams using explainable artificial intelligence (xai). *Sensors, 25(3)*, 847. https://doi.org/10.3390/s25030847

Aldaej, A., Ahanger, T., & Ullah, I. (2023). Deep learning-inspired iot-ids mechanism for edge computing environments. *Sensors, 23(24)*, 9869. https://doi.org/10.3390/s23249869

Alrayes, F., Alshuqayran, N., Nour, M., Duhayyim, M., Mohamed, A., Mohammed, A., … & Yaseen, I. (2023). Optimal fuzzy logic enabled intrusion detection for secure iot-cloud environment. *Computers Materials & Continua, 74(3)*, 6737–6753. https://doi.org/10.32604/cmc.2023.032591

Alsirhani, A., Tariq, N., Humayun, M., Alwakid, G., & Sanaullah, H. (2025). Intrusion detection in smart grids using artificial intelligence-based ensemble modelling. *Cluster Computing, 28(4)*. https://doi.org/10.1007/s10586-024-04964-9

Alsulami, M. (2024). An ai-driven model to enhance sustainability for the detection of cyber threats in iot environments. *Sensors, 24(22)*, 7179. https://doi.org/10.3390/s24227179

Arreche, O. and Abdallah, M. (2025). A comparative analysis of dnn-based white-box explainable ai methods in network security. *EURASIP Journal on Information Security, 2025(1)*. https://doi.org/10.1186/s13635-025-00201-x

Choudhury, M. and Azad, C. (2022). Deep learning-based ai modeling, intrusion detection. In *Handbook of Research on Cybersecurity Risk in Contemporary Business Systems* (pp. 128–143). https://doi.org/10.4018/978-1-6684-4558-7.ch005

Dapel, M., Asante, M., Uba, C., & Agyeman, M. (2023). Artificial intelligence techniques in cybersecurity management. In *AI and Cybersecurity* (pp. 241–255). https://doi.org/10.1007/978-3-031-20160-8_14

Deshmukh, A. and Ravulakollu, K. (2024). An efficient cnn-based intrusion detection system for iot: use case towards cybersecurity. *Technologies, 12(10)*, 203. https://doi.org/10.3390/technologies12100203

Habeeb, M. and Babu, T. (2022). Network intrusion detection system: a survey on artificial intelligence-based techniques. *Expert Systems, 39(9)*. https://doi.org/10.1111/exsy.13066

Holubenko, V., Silva, P., & Bento, C. (2023). An intelligent mechanism for monitoring and detecting intrusions in iot devices. *2023 IEEE CCNC*, 959–960. https://doi.org/10.1109/ccnc51644.2023.10060443

Houda, Z., Brik, B., & Khoukhi, L. (2022). "Why should I trust your IDS?": an explainable deep learning framework for intrusion detection systems in IoT networks. *IEEE Open Journal of the Communications Society, 3*, 1164–1176. https://doi.org/10.1109/ojcoms.2022.3188750

Hussain, F., Abbas, S., Husnain, M., Fayyaz, U., Shahzad, F., & Shah, G. (2020). IoT DoS and DDoS attack detection using ResNet. *2020 INMIC*. https://doi.org/10.1109/inmic50486.2020.9318216

Kazmi, S., Hassan, R., Qamar, F., Nisar, K., & Ibrahim, A. (2023). Security concepts in emerging 6G communication: threats, countermeasures, authentication techniques and research directions. *Symmetry, 15(6)*, 1147. https://doi.org/10.3390/sym15061147

Krishnaveni, S., Jothi, B., Chen, T., & Sathiyanarayanan, M. (2024). TwinSec-IDS: an enhanced intrusion detection system in SDN-digital-twin-based industrial cyber-physical systems. *Concurrency and Computation Practice and Experience, 37(3)*. https://doi.org/10.1002/cpe.8334

Mallampati, S. and Seetha, H. (2024). Enhancing intrusion detection with explainable ai: a transparent approach to network security. *Cybernetics and Information Technologies, 24(1)*, 98–117. https://doi.org/10.2478/cait-2024-0006

Mahbooba, B., Sahal, R., Alosaimi, W., & Serrano, M. (2021). Trust in intrusion detection systems: an investigation of performance analysis for machine learning and deep learning models. *Complexity, 2021(1)*. https://doi.org/10.1155/2021/5538896

Mambwe, H., Chavula, P., Kayusi, F., Lungu, G., & Uwimbabazi, A. (2024). Machine learning and AI for security mechanisms: a systematic literature review using a PRISMA framework. *LatIA, 2*, 331. https://doi.org/10.62486/latia2025331

Mansour, R. (2022). Artificial intelligence based optimization with deep learning model for blockchain enabled intrusion detection in cps environment. *Scientific Reports, 12(1)*. https://doi.org/10.1038/s41598-022-17043-z

Mounir, M., Sayed, S., & Eldakroury, M. (2025). Securing the future: real-time intrusion detection in IIoT smart grids through innovative AI solutions. *JCIM, 15(2)*, 208–244. https://doi.org/10.54216/jcim.150216

Moura, R., Franqueira, V., & Pessin, G. (2023). Cybersecurity in industrial networks: artificial intelligence techniques applied to intrusion detection systems. *CSCE*, 2235–2242. https://doi.org/10.1109/csce60160.2023.00365

Neupane, S., et al. (2022). Explainable intrusion detection systems (X-IDS): a survey of current methods, challenges, and opportunities. *IEEE Access, 10*, 112392–112415. https://doi.org/10.1109/access.2022.3216617

Otoum, S., Kantarcı, B., & Mouftah, H. (2021). A comparative study of AI-based intrusion detection techniques in critical infrastructures. *ACM Transactions on Internet Technology, 21(4)*, 1–22. https://doi.org/10.1145/3406093

Popli, R., Sethi, M., Kansal, I., Garg, A., & Goyal, N. (2021). Machine learning based security solutions in MANETs: state of the art approaches. *Journal of Physics: Conference Series, 1950(1)*, 012070. https://doi.org/10.1088/1742-6596/1950/1/012070

Qi, J. and Wang, J. (2024). Bridging artificial intelligence and railway cybersecurity: a comprehensive anomaly detection review. *Transportation Research Record.* https://doi.org/10.1177/03611981241302335

Rahman, M., Dhakal, K., Gony, N., SD, M., & Rahman, M. (2025). AI integration in cybersecurity software: threat detection and response. *IJIRSS, 8(3)*, 3907–3921. https://doi.org/10.53894/ijirss.v8i3.7403

Rajapaksha, S., Kalutarage, H., Al-Kadri, M., Petrovski, A., Madzudzo, G., & Cheah, M. (2023). AI-based intrusion detection systems for in-vehicle networks: a survey. *ACM Computing Surveys, 55(11)*, 1–40. https://doi.org/10.1145/3570954

Sabeel, U., Heydari, S., Elgazzar, K., & El-Khatib, K. (2021). Building an intrusion detection system to detect atypical cyberattack flows. *IEEE Access, 9*, 94352–94370. https://doi.org/10.1109/access.2021.3093830

Salem, A., Azzam, S., Emam, O., & Abohany, A. (2024). Advancing cybersecurity: a comprehensive review of AI-driven detection techniques. *Journal of Big Data, 11(1).* https://doi.org/10.1186/s40537-024-00957-y

Sayegh, H., Wang, D., & Al-madani, A. (2024). Enhanced intrusion detection with LSTM-based model, feature selection, and SMOTE for imbalanced data. *Applied Sciences, 14(2)*, 479. https://doi.org/10.3390/app14020479

Sharmila, B. and Nagapadma, R. (2023). Quantized autoencoder (QAE) intrusion detection system for anomaly detection in resource-constrained IoT devices using RT-IoT2022 dataset. *Cybersecurity, 6(1).* https://doi.org/10.1186/s42400-023-00178-5

Shtayat, M., Hasan, M., Sulaiman, R., Islam, S., & Khan, A. (2023). An explainable ensemble deep learning approach for intrusion detection in industrial internet of things. *IEEE Access, 11*, 115047–115061. https://doi.org/10.1109/access.2023.3323573

Soliman, K., Sobh, M., & Bahaa-Eldin, A. (2021). Survey of machine learning HIDS techniques. *IEEE ICCES*, 1–5. https://doi.org/10.1109/icces54031.2021.9686138

Souza, L., Sammarco, M., Achir, N., Campista, M., & Costa, L. (2024). AutomHS-GPT: automated model and hyperparameter selection with generative pre-trained model. *IEEE CloudNet*, 1–8. https://doi.org/10.1109/cloudnet62863.2024.10815898

Thapa, N., Liu, Z., KC, D., Gokaraju, B., & Roy, K. (2020). Comparison of machine learning and deep learning models for network intrusion detection systems. *Future Internet, 12(10)*, 167. https://doi.org/10.3390/fi12100167

Torre, D., Chennamaneni, A., Jo, J., Vyas, G., & Sabrsula, B. (2025). Toward enhancing privacy preservation of a federated learning CNN intrusion detection system in IoT. *ACM TOSEM, 34(2)*, 1–48. https://doi.org/10.1145/3695998

Vinayakumar, R., KP, S., & Poornachandran, P. (2019). A comparative analysis of deep learning approaches for network intrusion detection systems (N-IDSs). *IJDCF, 11(3)*, 65–89. https://doi.org/10.4018/ijdcf.2019070104

Yang, L., Naser, S., Shami, A., Muhaidat, S., Ong, L., & Debbah, M. (2025). Towards zero touch networks: cross-layer automated security solutions for 6G wireless networks. *IEEE Transactions on Communications*, 1–1. https://doi.org/10.1109/tcomm.2025.3547764

Zhou, X., Liang, W., Shimizu, S., Ma, J., & Jin, Q. (2021). Siamese neural network based few-shot learning for anomaly detection in industrial cyber-physical systems. *IEEE Transactions on Industrial Informatics, 17(8)*, 5790–5798. https://doi.org/10.1109/tii.2020.3047675

Zubair, M., et al. (2022). Secure bluetooth communication in smart healthcare systems: a novel community dataset and intrusion detection system. *Sensors, 22(21)*, 8280. https://doi.org/10.3390/s22218280