Communication in Physical Sciences, 2024 11(4): 1060-1088

Effectiveness of Machine Learning Models in Intrusion Detection Systems: A Systematic Review

Name: Dahunsi Samuel Adeyemi
Received: 27 July 2024/Accepted: 19 October 2024/Published: 26 October 2024

Abstract *While there are several benefits of machine learning (ML) algorithm for intrusion detection, it has been established that there are other issues like time span and classification of data. Thus, this study conducted a systematic review on the effectiveness of machine learning models in intrusion detection systems. Using the meta-synthesis research design, the study adopts a systematic literature review approach. Different databases (Web of Science, Scopus, Google Scholar, IEEE Xplore, and CINAHL) were consulted and the search techniques required the use of Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA). Data were extracted from the nineteen final selected studies, using the data extraction table. Results showed that the commonly used ML models include Random Forest (RF), Support Vector Machine (SVM), Decision Trees (DT), Naïve Bayes (NB), K-Nearest Neighbors (KNN), Logistic Regression (LR), Gradient Boosting, and AdaBoost. Findings showed that the performance metrics used to measure the effectiveness of ML-enhanced intrusion detection systems include accuracy, precision, recall, F1-score, error margin, false positive rate (FPR), false negative rate (FNR), and area under the ROC curve (AUC). It was demonstrated that ML algorithms perform well in detecting various cyber intrusions. The datasets used for training machine learning models include KDD Cup 99, NSL-KDD, UNSW-NB15, Kyoto, CICIDS2017, and Wireless Sensor Network Dataset (WSN-DS). The challenges associated with the application of ML algorithms for intrusion detection systems include data imbalance, high dimensionality, and feature selection complexities. The study concludes that machine learning models have the capacity to detect various cyber intrusions.*

Keywords: *Machine learning, deep learning, intrusion detection systems, effectiveness, intrusion detection*

Dahunsi Samuel Adeyemi
University of Central Missouri, Missouri, US
Email: dxa26930@ucmo.edu
Orcid id: 0009-0007-5485-8052

1.0 Introduction

Generally, intrusion detection systems can be referred to as efforts to prevent intrusions or attacks that can damage the credibility of security services, such as data confidentiality, integrity, and availability. These systems can be categorized into two, which include Signature-based Intrusion Detection Systems (SIDS) and Anomaly-based Intrusion Detection Systems (AIDS) (Khraisat *et al*., 2019). Essentially, the relevance and usefulness of any system can be entrenched when it is secured. Milenkoski *et al*. (2015) noted that intrusion detection systems are also known as computer intrusion detection systems. The authors noted that intrusion detection systems can be considered secure when they possess the qualities of confidentiality, integrity, and availability of their data and services. Thus, the achievement of a secure system should be paramount to a system developer.

Ordinarily, system attacks are often intentional attempts to compromise the system's security features and architectures (Ahmadjee *et al*., 2022). Therefore, system developers and users need to ensure they detect such system attacks in order to achieve or enjoy credible systems. This accentuates the significance of intrusion

detection systems. Ozkan-Okay *et al*. (2021) described an intrusion detection system as a process of keeping an eye on activities taking place on a computer system or network and examining them for indications of possible dangers, such as threats or infractions of usage guidelines or accepted security procedures. The authors noted further that the main reasons for the use of intrusion detection systems (IDS) include potential event detection, information recording, and transmission of recorded information to security administrators.

Panigrahi & Borah (2018) noted that intrusion detection systems are essential in order to reduce unauthorized access to a system or network resources, which use either anomaly detection or misuse detection. The authors noted further that it is important for IDS to identify potential threats through the use of current data. Training IDSs with current data can help fill the gap with respect to potential attacks or abuse. Kabir *et al*. (2018) observed that IDS identify signs of legitimate users abusing system resources or an attacker breaching a computer system. Since discussion surrounding IDS is integral in this age, it becomes imperative for network administrators and security experts to pay attention to it. Thus, it becomes crucial to understand intrusion detection systems, especially from the perspective of use and effectiveness of machine learning models to detect attacks or intrusions.

Heidari *et* al., (2023), averred that IDS is a significant component of the technology used in cybersecurity. It is used to identify harmful activity by monitoring and analyzing network data from various sources. The authors viewed IDS as not just an effort, but a procedure that uses a variety of methods to identify harmful activities that are directed at a system. Bridges *et al*. (2019) emphasized that intrusion detection systems have three components, which include data collection, conversion to select features, and a decision engine. All these components underpin areas such as manual techniques, algorithms, and commercial products, which are geared toward continual monitoring of computing assets for any sign of compromise or attack. Meanwhile, there are different tools or models that can be used to identify or detect intrusion. It has been established that both machine learning and deep learning can be used in intrusion detection systems (Liu & Lang, 2019; Thapa *et al*., 2020). In this study, the focus will be on machine learning models.

Belavagi and Muniyal (2016) noted that an intrusion detection model is a predictive model used to predict the network data traffic as normal or intrusion, which is built for clustering, classification, and prediction. It was concluded that classification and predictive models for intrusion detection are developed by using machine learning classification algorithms, namely Logistic Regression, Gaussian Naive Bayes, Support Vector Machine and Random Forest. Meanwhile, Azizan *et al*. (2021) noted that the machine learning models that are used vary. However, the commonest are random forest (RF), decision jungle (DJ), and support vector machine (SVM). Das *et al*. (2020) noted that trained models are often used to detect strange traffic, which may or may not depend on attack methods and feature vectors. This indicates the importance of machine learning models for intrusion detection.

Meanwhile, there are different learning models that can be used to detect intrusion in a system. These learning models include machine learning, big data, and deep learning, which are all popular methods for intrusion detection systems (Liu & Lang, 2019). Thus, it is left to cybersecurity experts and other professionals to determine which is best for them. This decision can only be achieved through the indices of effectiveness. While there are several benefits of machine learning algorithms for intrusion detection, it has been established that there are other issues like time span and classification of data (Othman *et al*., 2018). These identified challenges may potentially

limit or reduce the effectiveness of machine learning models for intrusion detection systems. Although prior studies have explored AI-driven defence mechanisms, there is limited systematic evidence on their effectiveness, algorithms employed, and ethical implications, leaving a critical gap in understanding their holistic deployment in cybersecurity. The significance of this study lies in its potential to consolidate fragmented evidence on AI-driven autonomous response systems, providing a comprehensive understanding of their effectiveness, algorithms, and ethical considerations. By doing so, the review offers valuable insights to cybersecurity researchers, practitioners, and policymakers, guiding the development of resilient, adaptive, and ethically aligned defence frameworks for autonomous systems It is against the foregoing that this study seeks to examine the effectiveness of machine learning models in intrusion detection systems. The study's findings will answer the following research questions:

(i) What are the various machine learning models applied in intrusion detection systems?
(ii) What are the performance metrics used to measure the effectiveness of machine learning-based intrusion detection systems?
(iii) How effective are the machine learning algorithms used in detecting various types of cyber intrusions?
(iv) What are the datasets commonly used in training and evaluating machine learning models for intrusion detection?
(v) What are the challenges in the application of machine learning algorithms for intrusion detection systems?

2.0 Methodology

Meta-synthesis research design was adopted to understand the effectiveness of machine learning models in intrusion detection systems. The study adopts this design as it offers a structured way of synthesizing prior empirical evidence in the studied area. This involves the formulation of research questions, searching various databases for relevant literature, screening the literature for its suitability to the study, assessing the qualities of the selected literature, extracting relevant information from the final selected literature, and analyzing the collected information to generate themes based on the research questions (Schut *et al*., 2024). This approach is a systematic review of the literature to understand the prevailing themes in a particular study area, which differs from meta-analysis which does not provide an opportunity for qualitative evidence on the data extracted from the literature (Ahn *et al*., 2018). This study adopts a systematic review of literature, which involves using a structured approach to get the relevant literature for the study. This is to achieve a robust and comprehensive approach to select the final literature. This allows the achievement of search results that would provide relevant information for the study. This search technique concerns the use of the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA). The PRISMA framework is designed so that it ensures structured data collection. Helach *et al*. (2023) established that PRISMA is the most popular and widely adopted framework for systematic reviews of literature. The 27-item PRISMA is divided into identification, screening, eligibility, and inclusion. The identification stage concerns literature search, including the sources and databases consulted. The screening stage concerns the evaluation of the titles and abstracts of literature retrieved. The eligibility phase highlights the inclusion and exclusion criteria. Using this PRISMA framework, the final selected literature for this study is fifteen (15).

To start with, five (5) databases were consulted for relevant literature for the study. These databases include Web of Science, Scopus, Google Scholar, IEEE Xplore, and CINAHL.

All these were considered to have literature on machine learning models in intrusion detection systems. Different search terms were used for this study, which are premised on the central aims and the specific objectives highlighted for the study. Boolean operators of "AND" and "OR" were used for this study owing to the nature of their relevance in extending the understanding of the effectiveness of machine learning models in intrusion detection systems. These search terms include "machine learning models and intrusion detection systems", "machine learning models OR intrusion detection systems", "machine learning and intrusion detection", and "machine learning OR intrusion detection".

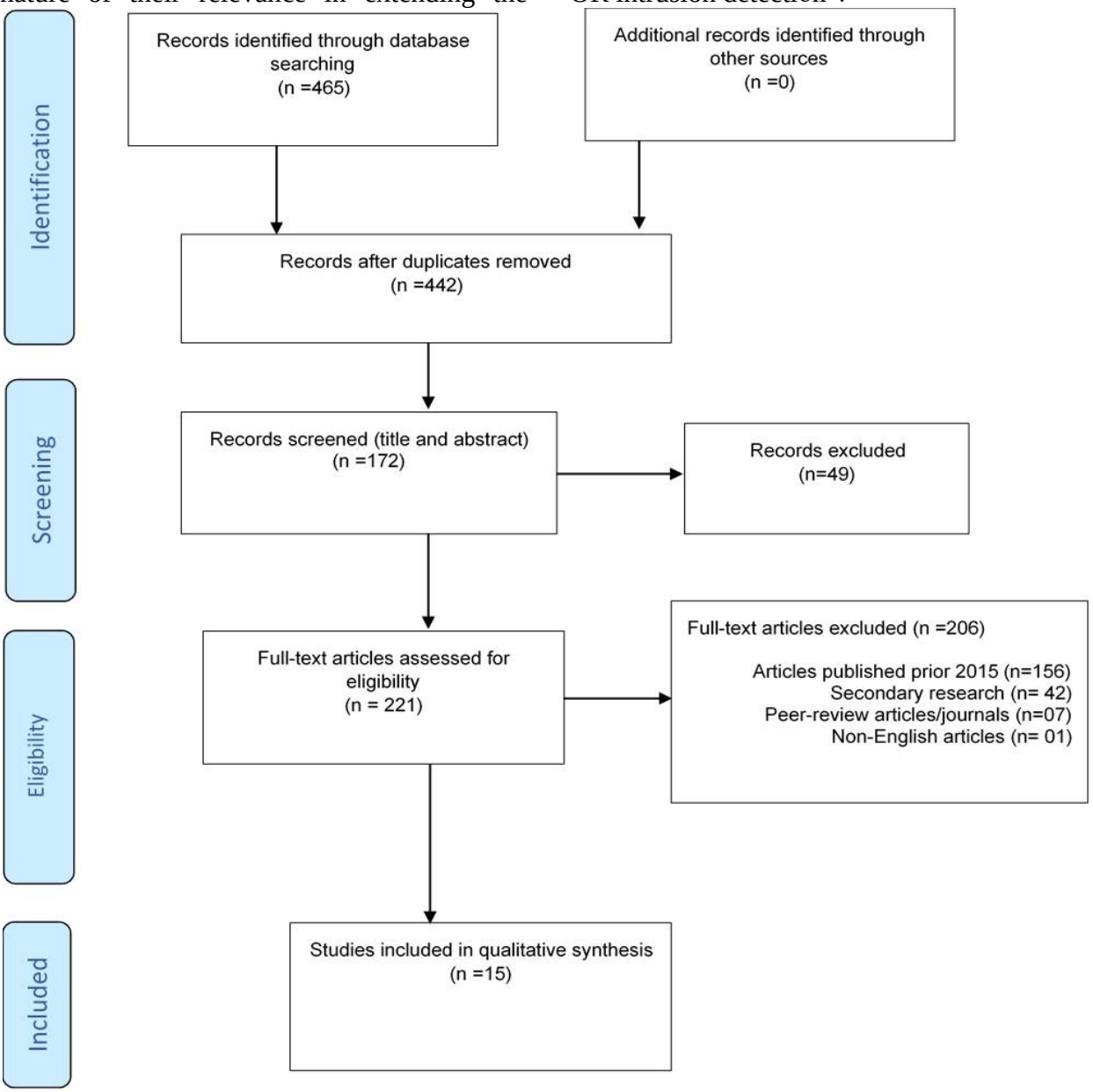

***Fig.* 1: PRISMA diagram flow (Author's self-designed, 2024)**

Having retrieved the relevant literature to the study, scanning was conducted on the methodology sections of the literature to ascertain those studies that were carried out in the focus areas. While this may be considered strenuous, it gives room for an exhaustive assessment of the literature in the studied area.

Meanwhile, the search period was left between 2014 and 2024. This was considered to allow for the retrieval of recent evidence in the area of study.

Table 2: Electronic Search Strategy (Extracts for five databases)

S/N	Search terms	Web of Science	Scopus	Google Scholar	IEE Xplore	CINAHL
				Number of hits		
S1	Machine learning models AND intrusion detection systems	2890	1200	595	752	1502
S2	Machine learning models OR intrusion detection systems	2173	750	123	981	1742
S3	Machine learning AND intrusion detection	23000	21000	17300	27000	5211
S4	Machine learning OR intrusion detection	21000	28000	14000	17000	3590
Databases search limits adopted						
Duplicates removed		89	91	75	82	75
Titles and abstracts checked		56	70	46	28	51
Articles < or = 10 years (2014-2024)		33	19	29	55	24
Secondary research		21	09	12	28	14
Peer-reviewed articles/journals		13	04	05	15	06
English language only		NA	N/A	02	N/A	N/A
Final selected		7	1	2	7	2

Source: Author's Literature Search (2024)

3.0 Results and Discussion

On the various machine learning models applied in intrusion detection systems, the results showed that both traditional machine learning (ML) and deep learning (DL) models are applied to intrusion detection systems (IDS). The results showed that the commonly used ML models include Random Forest (RF), Support Vector Machine (SVM), Decision Trees (DT), Naïve Bayes (NB), K-Nearest Neighbors (KNN), Logistic Regression (LR), Gradient Boosting, and AdaBoost (Abdallah *et al*., 2022; Belavagi & Muniyal, 2016; Mohammed & Hussein, 2022). It emerged in the findings that the deep learning models that are increasingly becoming prominent in intrusion detection systems include Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNN), and Deep Neural Networks (DNN) (Karatas *et al*., 2018; Vinayakumar *et al*., 2019). It emerged in the thematic analysis that hybrid approaches are also used for intrusion detection systems, such as optimized SVM with firefly algorithms and ensemble methods (Shah & Isaac, 2018; Zheng *et al*., 2022).

On the performance metrics used to measure the effectiveness of machine learning-based intrusion detection systems, results showed that ML-based IDS often employ accuracy, precision, recall, F1-score, error margin, false positive rate (FPR), false negative rate (FNR), and area under the ROC curve (AUC) for performance metrics (Almseidin *et al*., 2017;

Elnakid *et al*., 2023; Mahbooba *et al*., 2021). Results showed that accuracy remains the most widely used metric, with RF and SVM models often achieving 98% accuracy on benchmark datasets (Azizan *et al*., 2021; Türk, 2023). However, some of the studies (Almseidin *et al*., 2017; Shah & Isaac, 2018) highlighted the importance of false positive and false negative rates, with emphasis that FPRs undermine IDS reliability. Moreover, the findings showed that F1-scores and recall are significant when datasets are imbalanced, which ensure minority class intrusions are detected effectively (Elnakib *et al*., 2023).

On the effectiveness of machine learning (ML) algorithms in detecting cyber intrusions, the results showed that ML algorithms demonstrate robust performance in detecting various cyber intrusions. Evidence (Mohammed & Hussein, 2022) suggests that Random Forest consistently outperforms other classifiers, which helps in achieving accuracy rates up to 99.96% with low error margins. SVM excels in precision and recall, which makes it effective for detecting both known and novel attacks (Azizan *et al*., 2021). KNN demonstrates superior performance on real-world datasets, especially CICIDS2017, which surpasses other ML classifiers (Alrowaily *et al*., 2019). Results showed that deep learning (DL) models, especially DNNs and hybrid DL frameworks, handle high-dimensional data, which offer scalability and low training times (Karatas *et al*., 2018; Vinayakumar *et al*., 2019). Nonetheless, Alhajjar *et al*. (2021) indicate that adversarial perturbations can degrade performance, which highlight the vulnerability of ML-based IDS.

On the datasets commonly used for training and evaluation, the findings showed that most of the studies (Abdallah *et al*., 2022; Mahbooba *et al*., 2021; Vinayakumar *et al*., 2019) use benchmark datasets such as KDD Cup 99, NSL-KDD, UNSW-NB15, Kyoto, CICIDS2017, and Wireless Sensor Network Dataset (WSN-DS) for training and evaluating ML models. Findings revealed that CICIDS2017 is popular due to its real-world, multi-class, and highly imbalanced network traffic representation (Elnakib *et al*., 2023; Maseer *et al*., 2019). Abdelmoumin *et al*. (2021) revealed that the choice of dataset affects detection accuracy, as imbalanced or different training and testing data often lead to increased false positive rates and poor generalization.

On the challenges in applying ML algorithms to intrusion detection systems, the findings showed that there are several challenges including data imbalance, high dimensionality, and feature selection complexities, which impact detection rates and increase false alarms (Abdallah *et al*., 2022; Abdulmoumin *et al*., 2021). It was shown that model complexity and training costs also limit real-time applicability, especially in IoT environments (Elnakib *et al*., 2023). Results showed that adversarial attacks pose serious threats by exploiting ML vulnerabilities, which cause high misclassification rates (Alhajjar *et al*., 2021). Moreover, some of the studies (Abdallah *et al*., 2022; Zheng *et al*., 2022) highlight that the curse of dimensionality, hyperparameter tuning, and dataset generalization hinder optimal model performance. Ensemble and hybrid models partially reduce these issues, but at the cost of increased computational demands.

Vol 11Implications

The study emphasized the need for policymakers to develop cybersecurity strategies that integrate machine learning (ML) and deep learning (DL) models for intrusion detection systems (IDS). Advanced models like Random Forst (RF), Support Vector Machine (SVM), and deep neural networks (DNN) achieve detection accuracies above 98%, which makes it important to invest in research and infrastructure that support the deployment of such technologies. Furthermore, policymakers should prioritize funding for secure, high-quality, and balanced datasets, such as

CICIDS2017, which influence intrusion detection systems performance. Furthermore, regulatory frameworks must address the vulnerabilities posed by adversarial attacks and ensure that ML-driven IDS meet minimum standards for false positive and false negative rates to maintain public trust. Moreover, national cybersecurity policies should promote collaboration among academia, industry, and government agencies to enhance the development of hybrid ML-DL models, which enables adaptation to evolving cyber threats and enhance overall digital resilience.

The study provides significance for cybersecurity professionals and system administrators. The findings highlight practical pathways to improve intrusion detection. The findings indicate that models like RF, K-Nearest Neighbors (KNN), and hybrid ML-DL frameworks can provide high detection accuracy emphasizes the importance of selecting algorithms aligned with specific network environments and threat profiles. Moreover, practitioners and experts must consider evaluation metrics beyond accuracy, which give due weight to F1-score, precision, and recall, especially in imbalanced datasets where minority class intrusions can go undetected. Since model complexity, data imbalance, and feature selection affect intrusion detection systems performance, there is a need to adopt robust preprocessing techniques and dimensionality reduction strategies. Moreover, the vulnerability of ML models to adversarial perturbations calls for continuous model monitoring, adversarial training, and integrating ensemble methods to harden IDS against evasion attacks. Lastly, experts should adopt scalable solutions with optimized training times, which ensure real-time applicability in dynamic environments like IoT networks.

For the society, the study demonstrated the growing importance of ML-driven intrusion detection systems in safeguarding personal, organizational, and national digital assets. With cyberattacks becoming sophisticated, the deployment of accurate and efficient ML-based intrusion detection systems contributes to the security and reliability of digital services, which enhance public confidence in online systems. The study's findings implied that reduced false positives and improved recall benefit end-users by minimizing service disruptions and ensuring timely detection of malicious activities. Furthermore, as digital transformation accelerates, societies that adopt advanced intrusion detection systems technologies are better positioned to protect sensitive data, support economic growth, and reduce the social and financial impacts of cybercrime. Public awareness campaigns about the benefits and limitations of ML-enhanced cybersecurity tools can foster responsible digital behavior. This also encourages collaboration among citizens, organizations, and government in building a safer cyberspace.

4.0 Conclusion

The systematic review revealed that AI-driven autonomous response systems in cybersecurity have evolved significantly, offering scalable, adaptive, and real-time protection against emerging threats. Regarding the first research question, the findings showed that these systems are characterized by adaptive learning, predictive analytics, and automated incident response, with applications spanning network intrusion prevention, transport systems, and cloud environments.

For the second research question, the review established that supervised and unsupervised learning algorithms dominate the field, with reinforcement learning, neural networks, decision trees, and hybrid approaches (such as ML integrated with fuzzy logic) being the most frequently adopted methodologies. These methods support anomaly detection, simulation-based testing, and proactive risk assessment.

Addressing the third research question, the evidence demonstrated that autonomous systems have high effectiveness in mitigating

cyber threats by improving detection accuracy, minimizing false positives, enhancing situational awareness, and accelerating response times compared to traditional manual approaches.

On the fourth research question, the review identified multiple challenges in deploying these systems. These include adversarial machine learning that manipulates models, data quality and model bias, difficulties in establishing comprehensive regulatory standards, and the limitations of static AI models that require continuous retraining.

Finally, with respect to the fifth research question, the review highlighted several ethical considerations, including transparency, accountability, fairness, privacy, and trust. The findings underscored the importance of explainable AI and human-in-the-loop oversight to ensure responsible and ethical deployment of autonomous response systems in cybersecurity.

The study concludes that AI-driven autonomous response systems represent a transformative shift in cybersecurity, offering adaptability, scalability, and real-time mitigation of threats. Different system architectures and models, such as reinforcement learning, anomaly detection, and hybrid approaches integrating ML, DL, and fuzzy logic, have demonstrated effectiveness in strengthening cyber resilience. Evidence from the reviewed studies shows that these systems reduce detection latency, minimize false alarms, and enhance situational awareness. However, challenges such as adversarial machine learning, data quality limitations, model bias, and regulatory gaps hinder optimal deployment. Also, ethical issues including transparency, accountability, and trust, remain critical as autonomy increases in cyberspace defence. Overall, the review highlights the need for continuous innovation to improve model robustness, integration with human oversight, and the establishment of ethical frameworks that align AI-driven defense systems with broader societal values. Strengthening dataset representativeness, regulatory support, and explainable AI will be essential for achieving resilient and trustworthy autonomous response systems in cybersecurity.

5.0 References

Abdallah, E. E., Eleisah, W., & Otoom, A. F. (2022). Intrusion detection systems using supervised machine learning techniques: A survey. *Procedia Computer Science*, *201*, pp. 205-212.

Abdelmoumin, G., Rawat, D. B., & Rahman, A. (2021). On the performance of machine learning models for anomaly-based intelligent intrusion detection systems for the internet of things. *IEEE Internet of Things Journal*, *9*, 6, pp. 4280-4290.

Ahmadjee, S., Mera-Gómez, C., Bahsoon, R., & Kazman, R. (2022). A study on blockchain architecture design decisions and their security attacks and threats. *ACM Transactions on Software Engineering and Methodology (TOSEM)*, *31*, 2, pp. 1-45.

Ahn, E., & Kang, H. (2018). Introduction to systematic review and meta-analysis. *Korean Journal of Anesthesiology*, *71*, 2, pp. 103-112.

Alhajjar, E., Maxwell, P., & Bastian, N. (2021). Adversarial machine learning in network intrusion detection systems. *Expert Systems with Applications*, *186*, 115782, https://doi.org/10.1016/j.eswa.2021.115782

Almseidin, M., Alzubi, M., Kovacs, S., & Alkasassbeh, M. (2017). Evaluation of machine learning algorithms for intrusion detection system. In *2017 IEEE 15th international symposium on intelligent systems and informatics (SISY)* (pp. 000277-000282). IEEE. doi:10.48550/arXiv.1801.02330

Alrowaily, M., Alenezi, F., & Lu, Z. (2019). Effectiveness of machine learning based intrusion detection systems. In *International Conference on Security, Privacy and Anonymity in Computation,*

Communication and Storage (pp. 277-288). Cham: Springer International Publishing.

Amanoul, S. V., Abdulazeez, A. M., Zeebare, D. Q., & Ahmed, F. Y. (2021). Intrusion Detection Systems Based on Machine Learning Algorithms," *2021 IEEE International Conference on Automatic Control & Intelligent Systems (I2CACIS)*, Shah Alam, Malaysia, 2021, pp. 282-287, doi: 10.1109/I2CACIS52118.2021.9495897. *(I2CACIS)* (pp. 282-287). IEEE.

Azam, Z., Islam, M. M., & Huda, M. N. (2023). Comparative analysis of intrusion detection systems and machine learning-based model analysis through decision tree. *IEEE Access*, *11*, pp. 80348-80391.

Azizan, A. H., Mostafa, S. A., Mustapha, A., Foozy, C. F. M., Wahab, M. H. A., Mohammed, M. A., & Khalaf, B. A. (2021). A machine learning approach for improving the performance of network intrusion detection systems. *Annals of Emerging Technologies in Computing (AETiC)*, *5*, 5, pp. 201-208.

Azizan, A. H., Mostafa, S. A., Mustapha, A., Foozy, C. F. M., Wahab, M. H. A., Mohammed, M. A., & Khalaf, B. A. (2021). A machine learning approach for improving the performance of network intrusion detection systems. *Annals of Emerging Technologies in Computing (AETiC)*, *5*, 5, pp. 201-208.

Belavagi, M. C., & Muniyal, B. (2016). Performance evaluation of supervised machine learning algorithms for intrusion detection. *Procedia Computer Science*, *89*, pp. 117-123.

Bridges, R. A., Glass-Vanderlan, T. R., Iannacone, M. D., Vincent, M. S., & Chen, Q. (2019). A survey of intrusion detection systems leveraging host data. *ACM Computing Surveys (CSUR)*, *52*, 6, pp. 1-35.

Das, S., Ashrafuzzaman, M., Sheldon, F. T., & Shiva, S. (2020). Network intrusion detection using natural language processing and ensemble machine learning. In *2020 IEEE Symposium Series on Computational Intelligence (SSCI)* (pp. 829-835). IEEE.

Elnakib, O., Shaaban, E., Mahmoud, M., & Emara, K. (2023). EIDM: Deep learning model for IoT intrusion detection systems. *Journal of Supercomputing*, *79*, 12, pp. 1-21.

Heidari, A., & Jabraeil Jamali, M. A. (2023). Internet of Things intrusion detection systems: a comprehensive review and future directions. *Cluster Computing*, *26*, 6, pp.3753-3780.

Helach, J., Hoffmann, F., Pieper, D., & Allers, K. (2023). Reporting according to the preferred reporting items for systematic reviews and meta-analyses for abstracts (PRISMA-A) depends on abstract length. *Journal of Clinical Epidemiology*, *154*, pp.167-177.

Karatas, G., Demir, O., & Sahingoz, O. K. (2018). Deep learning in intrusion detection systems. In *2018 International Congress on Big Data, Deep Learning and Fighting Cyber Terrorism* (pp. 113-116). IEEE.

Mohammed, S. Q., & Hussein, M. A. (2022). Performance analysis of different machine learning models for intrusion detection systems. *Journal of Engineering*, *28*, 5, pp. 61-91.

Kabir, E., Hu, J., Wang, H., & Zhuo, G. (2018). A novel statistical technique for intrusion detection systems. *Future Generation Computer Systems*, *79*, pp. 303-318.

Khraisat, A., Gondal, I., Vamplew, P., & Kamruzzaman, J. (2019). Survey of intrusion detection systems: techniques, datasets and challenges. *Cybersecurity*, *2*, 1, pp. 1-22.

Liu, H., & Lang, B. (2019). Machine learning and deep learning methods for intrusion detection systems: A survey. *Applied Sciences*, *9*, 20, pp. 4396-4408.

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), 5538896. https://doi.org/10.1155/2021/5538896.

Maseer, Z. K., Yusof, R., Bahaman, N., Mostafa, S. A., & Foozy, C. F. M. (2021). Benchmarking of machine learning for anomaly-based intrusion detection systems in the CICIDS2017 dataset. *IEEE Access*, *9*, pp. 22351-22370.

Milenkoski, A., Vieira, M., Kounev, S., Avritzer, A., & Payne, B. D. (2015). Evaluating computer intrusion detection systems: A survey of common practices. *ACM Computing Surveys (CSUR)*, *48*, 1, pp. 1-41.

Othman, S. M., Ba-Alwi, F. M., Alsohybe, N. T., & Al-Hashida, A. Y. (2018). Intrusion detection model using machine learning algorithm on Big Data environment. *Journal of big data*, *5*, 1, pp. 1-12.

Ozkan-Okay, M., Samet, R., Aslan, Ö., & Gupta, D. (2021). A comprehensive systematic literature review on intrusion detection systems. *IEEE Access*, *9*, pp. 157727-157760.

Panigrahi, R., & Borah, S. (2018). A detailed analysis of CICIDS2017 dataset for designing Intrusion Detection Systems. *International Journal of Engineering & Technology*, *7*, 3, pp. pp. 479-482.

Patgiri, R., Varshney, U., Akutota, T., & Kunde, R. (2018). An Investigation on Intrusion Detection System Using Machine Learning," *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*, Bangalore, India, 2018, pp. 1684-1691, doi: 10.1109/SSCI.2018.8628676

Schut, M., Adeyemi, I., Kumpf, B., Proud, E., Dror, I., Barrett, C. B., ... & Leeuwis, C. (2024). Innovation portfolio management for the public non-profit research and development sector: What can we learn from the private sector? *Innovation and Development*, 1-19. https://doi.org/10.1080/2157930X.2024.2400779.

Shah, S. A. R., & Issac, B. (2018). Performance comparison of intrusion detection systems and application of machine learning to Snort system. *Future Generation Computer Systems*, *80*, pp. 157-170.

Thapa, N., Liu, Z., Kc, D. B., Gokaraju, B., & Roy, K. (2020). Comparison of machine learning and deep learning models for network intrusion detection systems. *Future Internet*, *12*, 10, pp. 167-187.

Türk, F. (2023). Analysis of intrusion detection systems in UNSW-NB15 and NSL-KDD datasets with machine learning algorithms. *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi*, *12*, 2, pp. 465-477.

Vinayakumar, R., Alazab, M., Soman, K. P., Poornachandran, P., Al-Nemrat, A., & Venkatraman, S. (2019). Deep learning approach for intelligent intrusion detection system. *IEEE Access*, *7*, pp. 41525-41550.

Zhang, C., Jia, D., Wang, L., Wang, W., Liu, F., & Yang, A. (2022). Comparative research on network intrusion detection methods based on machine learning. *Computers & Security*, *121*, 102861, https://doi.org/10.1016/j.cose.2022.102861.

Conflict of interest

All declare there is no conflict of interest

Ethical Consideration

Not applicable

Availability of Data

Data shall be made available upon request.

Funding

None

Authors Contribution

All components of the work were carried out by the author

APPENDIX I
DATA EXTRACTION TOOL
Effectiveness of Machine Learning Models in Intrusion Detection Systems: A Systematic Review

S/N	Research titles and authors	Methodology	Models	Findings
1	Intrusion detection systems using supervised machine learning techniques: A survey. Abdallah *et al*. (2022)	The study employs research design.	The study used datasets, which include KDD'99, NSL-KDD, UNSW-Nb15, and CICIDS2017.	- Supervised machine learning models demonstrate promising classification performance across various datasets. - Random forest consistently achieves high accuracy and low false positive rates, making it one of the best-performing algorithms. - Feature selection is critical for improving classification performance in many cases. For example, it is effective for datasets like NSL-KDD and UNSW-NB15. - Dimensionality reduction techniques, such as PCA and LDA, help address the "curse of dimensionality" and improve model efficiency. - Results showed that machine learning models are highly effective for intrusion detection systems, offering high accuracy, adaptability to different datasets, and the ability to address challenges like data imbalance and high dimensionality.

2	EIDM: Deep learning model for IoT intrusion detection systems. Elnakib *et al*. (2023)	Experimental research design.	The study proposed Enhanced Intrusion Detection Model (EIDM). The models were trained and tested using the CICIDS2017 dataset, and their performance was evaluated based on metrics such as accuracy, precision, recall, and F-score.	- According to the experimental results, the accuracy of the models cannot exceed a certain value as they are affected by the number of classes and the number of samples per class. The more unbalanced the data and the greater number of classes are, the more complex the model becomes to attain high accuracy. However, the model complexity comes at the price of increasing its training and attack identification time costs which are crucial for IDS.
3	Comparative analysis of intrusion detection systems and machine learning-based model analysis through decision tree. Azam *et al*. (2023)	Machine learning (ML) and deep learning (DL) techniques were adopted in IDS systems.	The models used include Network Intrusion Detection Systems (NIDS).	- The decision tree, known for its speed and user-friendliness, is proposed as a model for detecting result anomalies, combining findings from a comparative survey.
4	Performance analysis of different machine learning models for intrusion detection systems. Mohammed and Hussein (2022)	As a case study Information Gain, Pearson and F-test feature selection techniques were used and the obtained results compared to models that use all the features.	Using Python Scikit-Learn library KNN, Support Vector Machine, Naïve Bayes, Decision Tree, Random Forest, Stochastic Gradient Descent, Gradient Boosting and Ada Boosting classifiers were designed.	- One unique outcome is that the Random Forest classifier achieves the best performance with an accuracy of 99.96% and an error margin of 0.038%, which supersedes other classifiers. Using 80% reduction in features and parameters extraction from the packet header rather than the workload, a big performance advantage is achieved, especially in online environments.

5	Deep learning in intrusion detection systems. Karatas *et al*. (2018)	The study adopted experimental research design, using the Intrusion detection systems (IDSs).	The study used of data collection, feature selection/conversion, and decision engine.	- Decision engine directly affects the efficiency of the system and use of machine learning techniques. Also, deep learning emerged as a new approach, which enables the use of Big Data with a low training time and high accuracy rate with its distinctive learning mechanism.
6	A machine learning approach for improving the performance of network intrusion detection systems. Azizan *et al*. (2021)	The study adopted experimental research design. The knowledge discovery in databases (KDD) methodology and intrusion detection evaluation dataset (CICIDS2017) are used in the testing which both are considered as a benchmark in the evaluation of IDS.	The study used machine learning algorithms, which include decision jungle (DJ), random forest (RF) and support vector machine (SVM).	- The average accuracy results of the SVM is 98.18%, RF is 96.76% and DJ is 96.50% in which the highest accuracy is achieved by the SVM. The average precision results of the SVM is 98.74, RF is 97.96 and DJ is 97.82 in which the SVM got a higher average precision compared with the other two algorithms. - The average recall results of the SVM is 95.63, RF is 97.62 and DJ is 95.77 in which the RF achieves the highest average of recall than SVM and DJ. - In overall, the SVM algorithm is found to be the best algorithm that can be used to detect an intrusion in the system.
7	On the performance of machine learning models for anomaly-based intelligent intrusion detection	Experimental research design.	The study used anomaly-based machine learning-enabled intrusion detection systems (AML-IDS) models by tuning hyperparameters and ensemble	- Anomaly-based machine learning-enabled intrusion detection systems (AML-IDSs) show low performance and prediction accuracy while detecting

	systems for the internet of things. Abdelmoumin *et al*. (2021)	learning optimization techniques using the Microsoft Azure ML Studio (AMLS) platform.	intrusions in the Internet of Things (IoT) than that of deep learning-based intrusion detection systems (DL-IDSs). In particular, AML-IDS that employ low complexity models for IoT, such as the principal component machine (PCA) method and the one-class support vector machine (1-SVM) method, are inefficient in detecting intrusions when compared to DL-IDS with the two-class neural network (2-NN) method. - PCA and 1-SVM AML-IDS suffer from low detection rates compared to DL-IDS. The size of the data set and the number of features or variants in the data set may influence how well PCA and 1-SVM AML-IDS perform compared to DL-IDS. We attribute the low performance and prediction accuracy of the AML-IDS model to an imbalanced data set, a low similarity index between the training data and testing data, and the use of a single-learner model. - The intrinsic limitations of the single-learner model have a direct impact on the accuracy of an intelligent IDS. Also, the dissimilarity between testing data and training data leads to an

				increasingly high rate of false positives (FPs) in AML-IDS than DL-IDS, which have low false alarms and high predictability.
8	Comparative research on network intrusion detection methods based on machine learning. Zheng *et al*. (2022)	Experimental research design.	this paper selects the KDD CUP99 and NSL-KDD datasets to conduct comparative experiments on decision tree, Naive Bayes, support vector machines, random forests, XGBoost, convolutional neural networks, and recurrent neural networks.	- The detection accuracy, F1, AUC, and other indicators of these algorithms on different data sets are compared. The experimental results show that the effect of the ensemble learning algorithm is generally better. - The Naive Bayes algorithm has low accuracy in recognizing the learned data, but it has obvious advantages when facing new types of attacks, and the training speed is faster. The deep learning algorithm is not particularly prominent in this experiment, but its optimal results are affected by the structure, hyperparameters, and the number of training iterations, which need further in-depth study.
9	Deep learning approach for intelligent intrusion detection system. Vinayakumar *et al*. (2019)	The study adopted a survey research method. The optimal network parameters and network topologies for DNNs are chosen through the following hyperparameter selection methods with KDDCup 99 dataset. All	One of the approaches used in the study was to classify legitimate and anomalous behavior is to use Machine Learning (ML) techniques. The model used in the study includes KDDCup 99, NSL-KDD, UNSW-NB15, Kyoto, WSN-DS, and CICIDS 2017.	- The study showed that the DNN model learns the abstract and high-dimensional feature representation of the IDS data by passing them into many hidden layers. Through a rigorous experimental testing, it is confirmed that DNNs perform well in comparison with the classical machine learning classifiers.

		the experiments of DNNs are run till 1,000 epochs with the learning rate varying in the range [0.010.5]. The DNN model which performed well on KDDCup 99 is applied on other datasets, such as NSL-KDD, UNSW-NB15, Kyoto, WSN-DS, and CICIDS 2017, to conduct the benchmark.		Finally, we propose a highly scalable and hybrid DNNs framework called scale-hybrid-IDS-AlertNet which can be used in real-time to effectively monitor the network traffic and host-level events to proactively alert possible cyberattacks.
10	Analysis of intrusion detection systems in UNSW-NB15 and NSL-KDD datasets with machine learning algorithms. Türk (2023)	The study adopted the survey research design.	The study conducted a comprehensive attack detection process on the UNSW-NB15 and NSL-KDD datasets using existing machine learning and deep learning algorithms.	- In the UNSW-NB15 dataset, an accuracy of 98.6% and 98.3% was achieved for two-class and multi-class classification, respectively, and 97.8% and 93.4% accuracy were obtained in the NSL-KDD dataset. The results prove that machine learning algorithms are an effective solution for intrusion detection systems.
11	Adversarial machine learning in network intrusion detection systems. Alhajjar *et al*. (2021)	The study adopts survey research design.	The study used NSL-KDD and UNSW-NB15.	- The results show that our adversarial example generation techniques cause high misclassification rates in eleven different machine learning models, along with a voting classifier. Our work highlights the vulnerability of machine learning based network intrusion detection systems (NIDS)

			in the face of adversarial perturbation.
12	Performance comparison of intrusion detection systems and application of machine learning to Snort system. Shah and Isaac (2018)	The study used two open-source intrusion detection systems (IDSs) namely Snort and Suricata for accurately detecting the malicious traffic on computer networks.	- Snort and Suricata were installed on two different but identical computers and the performance was evaluated at 10 Gbps network speed. It was noted that Suricata could process a higher speed of network traffic than Snort with lower packet drop rate but it consumed higher computational resources. Snort had higher detection accuracy and was thus selected for further experiments. It was observed that the Snort triggered a high rate of false positive alarms. To solve this problem a Snort adaptive plug-in was developed. To select the best performing algorithm for Snort adaptive plug-in, an empirical study was carried out with different learning algorithms and Support Vector Machine (SVM) was selected. A hybrid version of SVM and Fuzzy logic produced a better detection accuracy. But the best result was achieved using an optimized SVM with firefly algorithm with FPR (false positive rate) as 8.6% and FNR (false negative rate) as 2.2%.

13	Intrusion detection systems based on machine learning algorithms. Amanoul *et al*. (2021)	The study adopts the survey research design.	The study used machine learning algorithms, which include Bayes Net, Random Forest, and Neural Network. Also, the study used two algorithms from deep learning, which include RNN and LSTM.	- The findings indicate a taxonomy of IDS, which uses the primary dimension of data objects to classify and sum up IDS based on and dependent on deep learning. The findings indicate that this kind of taxonomy is sufficient for researchers in cyber security. The study showed that algorithms from machine learning (Bayes Net, Random Forest, Neural Network) and two algorithms of deep learning (RNN, LSTM) that were tested on KDD cup 99 have accuracy algorithms. The study used a program WEKA to calculate the accuracy.
14	Trust in intrusion detection systems: an investigation of performance analysis for machine learning and deep learning models Mahbooba *et al*. (2021)	Two datasets are used to classify the IDS attack type, including wireless sensor network detection system (WSN-DS) and KDD Cup network intrusion dataset. A detailed comparison of the eight techniques' performance using all features and selected features is made by measuring the accuracy, precision, recall, and F1-score.	The four machine learning techniques are decision tree (DT), K nearest neighbour (KNN), random forest (RF), and naïve Bayes (NB). The four deep learning techniques are LSTM (one and two layers) and GRU (one and two layers).	- Findings show the performance of the machine learning models using the unseen testing KDD dataset. For the RF model, DOS class has achieved the highest accuracy among other models and classes (accuracy of 100%, precision of 100%, recall of 100%, and F-score of 100%). However, NB has the DOS class's worst performances (accuracy of 94.17%, precision of 93.82%, recall of 99.37%, and F-score of 96.52%). For the normal class, RF is the highest performance model (accuracy of 99.98%, precision of

				99.89%, recall of 99.99%, and F-score of 98.94%). - Results show the performance of the machine learning models using the unseen testing KDD dataset. For the RF and DT model, DOS class has achieved the highest accuracy among other models and classes (accuracy of 99.74%, precision of 99.74%, recall of 99.94%, and F-score of 99.84%). However, NB has the DOS class's worst performances (accuracy of 64.87%, precision of 80.27%, recall of 75.10%, and F-score of 77.60%). For the normal class, DT and RF are the highest performance model (accuracy of 99.93%, precision of 99.77%, recall of 99.82%, and F-score of 99.80%). KNN and NB models have achieved the second and third ranks based on accuracy over unseen data by 99.83% and 82.48%, respectively.
15	An investigation on intrusion detection system using machine learning. Patgiri *et al*. (2018)	Experimental research design.	The study used machine learning algorithms, namely, Random Forest and Support Vector Machine (SVM).	- The study demonstrated the comparison between the model's performance before and after feature selection of both Random Forest and SVM.
16	Evaluation of machine learning algorithms for	Experimental research design.	The study used machine learning classifiers based on KDD intrusion dataset. It succeeded to	- The study focus was on false negative and false positive performance metrics in order to

	intrusion detection system. Almseidin *et al*. (2017)		compute several performance metrics in order to evaluate the selected classifiers.	enhance the detection rate of the intrusion detection system. The implemented experiments demonstrated that the decision table classifier achieved the lowest value of false negative while the random forest classifier has achieved the highest average accuracy rate.
17	Performance evaluation of supervised machine learning algorithms for intrusion detection. Belavagi and Muniyal (2016)	The study adopted experimental research design.	In the study, classification and predictive models for intrusion detection are built by using machine learning classification algorithms namely Logistic Regression, Gaussian Naive Bayes, Support Vector Machine and Random Forest. These algorithms are tested with NSL-KDD data set.	- Experimental results shows that Random Forest Classifier out performs the other methods in identifying whether the data traffic is normal or an attack.
18	Benchmarking of machine learning for anomaly-based intrusion detection systems in the CICIDS2017 dataset. Maseer *et al*. (2019)	The study adopted the survey research design.	The study applies 10 popular supervised and unsupervised ML algorithms for identifying effective and efficient ML-AIDS of networks and computers. These supervised ML algorithms include the artificial neural network (ANN), decision tree (DT), k-nearest neighbor (k-NN), naive Bayes (NB), random forest (RF), support vector machine (SVM), and convolutional neural network (CNN) algorithms, whereas the unsupervised ML algorithms	- The ML-AIDS models are tested by using a recent and highly unbalanced multiclass CICIDS2017 dataset that involves real-world network attacks. In general, the k-NN-AIDS, DT-AIDS, and NB-AIDS models obtain the best results and show a greater capability in detecting web attacks compared with other models that demonstrate irregular and inferior results.

			include the expectation-maximization (EM), k-means, and self-organizing maps (SOM) algorithms.	
19	Effectiveness of machine learning based intrusion detection systems. Alrowaily *et al*. (2019)	Experimental research design.	A range of experiments was carried out on seven machine learning algorithms by using the CICIDS2017 intrusion detection dataset. The study used several machine learning algorithms, which include K-Nearest Neighbors (KNN) classifier.	- The experimental results demonstrated that the K-Nearest Neighbors (KNN) classifier outperformed in terms of precision, recall, accuracy, and F1-score as compared to other machine learning classifiers. Nevertheless, all of the used machine learning classifiers except KNN trained their models in a reasonable time.