JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 11(2), JUNE, 2023
ISSN: 2277-0011; Journal homepage: www.atbuftejoste.net

A Comprehensive Evaluation of Machine Learning Algorithms for Network Security Intrusion Detection Systems (IDS)

[1]Ezekiel Ehime AGBON, [2]Abdulkadir Mohammed GIWA, [3]Opeyemi Olajumoke ADEKOGBA, [2]Habib Tunji ADEYEMO, [1]Mohammed NASIR, [1]Abiodun Mogaji BABATUNDE
[1]Department of Electronics and Telecommunication Engineering,
Faculty of Engineering, Ahmadu Bello University, Zaria, Nigeria.
[2]Department of Information and Communication Technology,
Faculty of Ground and Communication Engineering,
Airforce Institute of Technology, Kaduna, Nigeria.
[3]Department of Telecommunication Engineering,
Faculty of Ground and Communication Engineering,
Airforce Institute of Technology, Kaduna, Nigeria.

ABSTRACT
Networks are faced with a wide range of difficulties and threats as a result of the development of Artificial intelligence (AI) technologies and the advent of fifth-generation networks. Concerns about data confidentiality and general network security emerge as more users access the network at once. An Intrusion Detection System (IDS) is frequently used to identify potential threats in order to allay these worries. Network assaults can be divided into two categories: minor and major. Denial-of-Service (DoS) and prob assaults are major types of attacks, whereas User-to-Root (U2R) and Remote-to-Login (R2L) attacks are smaller types. Minor attacks, usually referred to as uncommon attacks, pose a serious risk to host systems and can be particularly difficult to identify. This study investigates several machine learning techniques used in the implementation of these IDS systems and provides an extensive survey on IDS including proposed solution to each of the ML algorithm for IDS.

ARTICLE INFO
Article History
Received: January, 2023
Received in revised form: April, 2023
Accepted: May, 2023
Published online: June, 2023

KEYWORDS
Artificial Intelligence, Denial-of-Service, Intrusion Detection Systems, Machine Learning, Network Security, and Support Vector Machine.

INTRODUCTION

In recent years, the rapid advancement of Artificial Intelligence (AI) technologies and the emergence of fifth-generation (5G) networks have brought about unprecedented challenges and risks for network security (Shobowale *et al*., 2023). As the number of users accessing networks simultaneously continues to rise, concerns regarding data confidentiality and overall network security have become paramount. The need to address these concerns has led to the widespread adoption of Intrusion Detection Systems (IDS) as an essential component of network defense strategies. An IDS plays a crucial role in identifying potential threats within a network. By analyzing network traffic and monitoring system behavior, an IDS can detect and respond to various types of attacks. Network attacks can be broadly classified into two categories: minor attacks and major attacks. Major attacks, such as Denial-of-Service (DoS) and Prob attacks, pose significant risks to network availability and performance (Daniel *et al*., 2023). On the other hand, minor attacks, including User-to-Root (U2R) and Remote-to-Login (R2L) attacks, are characterized by their subtlety and complexity, making them particularly challenging to detect and mitigate. Among these attacks, minor attacks, also known as rare attacks, present a considerable threat to host systems. Their elusive nature and ability to exploit vulnerabilities make them a significant concern for network administrators. Detecting and preventing such attacks requires robust and advanced security measures. In this paper, we present a comprehensive survey on Intrusion Detection Systems (IDS) and specifically focus on the

Corresponding author: A. M. Giwa
engineermohammedagiwa@gmail.com
Department of Information and Communication Technology, Faculty of Ground and Communication Engineering, AFIT, Kaduna.

application of machine learning algorithms in implementing these systems. Machine learning algorithms have shown promising capabilities in enhancing the accuracy and efficiency of IDS by enabling automated detection and response to network threats. By analyzing large volumes of network data and identifying patterns, machine learning algorithms can effectively detect both known and unknown attacks, enhancing the overall security of the network.

The objective of this survey is to provide an overview of the state-of-the-art machine learning algorithms used in IDS implementation. We discuss various techniques and approaches employed in the field, highlighting their strengths, limitations, and potential areas for improvement. Furthermore, we examine the performance of different machine learning algorithms in detecting and mitigating major and minor network attacks. Through this survey, we aim to contribute to the understanding of IDS and provide insights for researchers and practitioners to enhance network security using machine learning-based intrusion detection approaches.

The remainder of this paper is organized as follows: Section 2 provides a comprehensive review of related work in the field of IDS and machine learning algorithms. Section 3 presents the application of machine learning algorithms in implementing the IDS systems. Finally, Section 4 concludes the paper and outlines potential future research directions in this area.

REVIEW OF SIMILAR WORKS

The field of Intrusion Detection Systems (IDS) using machine learning has witnessed significant advancements in recent years. Numerous researchers and practitioners have explored various approaches and algorithms to enhance the detection and mitigation of network attacks. In this section, we provide a comprehensive review of related works in the domain of IDS and machine learning, highlighting the key contributions, methodologies, and findings of existing studies. The review encompasses a wide range of research articles, conference papers, and technical reports that have investigated the application of machine learning techniques in IDS. We examine the different approaches employed by researchers, including supervised learning, unsupervised learning, and hybrid models. Additionally, we analyze the datasets used for training and evaluation purposes, as well as the performance metrics utilized to assess the effectiveness of the proposed IDS systems. Through this review, we aim to identify the current trends, challenges, and gaps in the field of IDS using machine learning. By understanding the existing literature, we can gain insights into the strengths and limitations of different approaches and identify potential areas for further improvement. Moreover, this review serves as a foundation for our study, allowing us to build upon the existing knowledge and contribute novel insights and methodologies to the field.

Munjula and Muniyal, (2017), carried out a study on machine learning algorithms for IDS. The authors noted that network intrusions could be detected through the utilization of an Intrusion Detection System (IDS). Also, machine learning algorithms played a crucial role in predicting whether network behavior corresponded to intrusion or normal activity. In the work, the authors focused on the analysis of prediction using various supervised machine learning algorithms, including Logistic Regression, Gaussian Naive Bayes, Support Vector Machine, and Random Forest. The investigation specifically concentrated on the NSL-KDD dataset, examining the effectiveness of these classification techniques in predicting four distinct types of attacks: Denial of Service, Remote to Local (R2L), Probe, and User to Root (U2R). Furthermore, the authors employed a multi-class classification approach to enhance the accuracy of attack identification. Additionally, the use of K-means in the work of Aung and Min, (2018) highlighted the similarity in assault distribution, and it also showed how accurate the Random Forest algorithm is at spotting intrusions. This paper used the KDD 99 to fully characterize the pattern recognition and machine learning algorithm performance for the four attack categories, including user-to-root (U2R) attacks, which involved unauthorized access to a local super-user, user-to-root (DoS) attacks, which denied legitimate requests to a system, probing attacks, which involved

Corresponding author: A. M. Giwa
engineermohammedagiwa@gmail.com
Department of Information and Communication Technology, Faculty of Ground and Communication Engineering, AFIT, Kaduna.

information gathering, and remote-to-local (R2L) attacks, which involved unauthorized local access from a remote machine. Jayesh *et al.,* (2020) mainly focus and primary goal revolved around the development and presentation of an extensive and all-encompassing system specifically designed for the purpose of detecting intruding attacks. Leveraging the powerful capabilities of Machine Learning, this system was aimed at effectively identifying and discerning unknown attacks by utilizing the knowledge and insights derived from past information obtained from known attacks. To ensure the effectiveness and efficiency of the system, various preprocessing techniques were carefully employed, enabling the data to be properly prepared and optimized for analysis. Moreover, the paper extensively discussed the process of model comparisons, emphasizing the importance of selecting the most suitable models for training and testing the system. This involved evaluating and comparing the performance of different models to determine their effectiveness in accurately detecting and classifying intrusions. Additionally, the paper highlighted the significance of employing robust evaluation techniques to assess the system's overall performance, ensuring its reliability and effectiveness in real-world scenarios. By providing detailed descriptions and explanations of these preprocessing techniques, model comparisons, and evaluation techniques, the paper aimed to equip researchers, practitioners, and security professionals with valuable insights and guidance for the successful implementation and deployment of intrusion detection systems.

In the work of Chauhan *et al.,* (2020), the primary aim of the study was to thoroughly explore and assess the significance of employing machine learning methods for the detection of intrusions. The review extensively covered various machine learning techniques that have been utilized in Intrusion Detection Systems (IDS) to date. The focus was specifically placed on evaluating the effectiveness of these techniques in detecting intrusions within network systems, with a particular emphasis on their predictive accuracy. Furthermore, this literature review critically examined and identified the limitations and drawbacks present in previous studies conducted in this domain. By thoroughly analyzing these limitations, valuable insights were gained into the current state of research in intrusion detection using machine learning methods. This comprehensive analysis served to shed light on the strengths and weaknesses of existing approaches, thereby providing a foundation for further advancements in the field. The ultimate goal of the review was to not only present a comprehensive overview of the state of the art in intrusion detection using machine learning, but also to identify and highlight areas that warrant further exploration and investigation. By identifying the gaps and open research questions in the field, this review aimed to guide future research endeavors, facilitate the development of more robust and accurate intrusion detection systems, and contribute to the advancement of the overall field of network security. Finally, in the work of Alkasassbeh and Almseidin, (2018), the study aimed to highlight the value of the Knowledge Discovery and Data Mining (KDD) dataset in evaluating and testing various machine learning techniques. The researchers recognized the significance of ensuring the creation of a reliable and unbiased experimental dataset, and therefore, dedicated considerable attention to the preprocessing stage of the KDD dataset. Three classifiers, namely J48, MLP, and Bayes Network, were selected for this investigation. Among these classifiers, the results revealed that the J48 classifier outperformed the others in terms of accuracy when it came to identifying and categorizing different types of attacks present in the KDD dataset. Specifically, the J48 classifier demonstrated superior performance in detecting DOS, R2L, U2R, and PROBE attacks, which were the focus of this study.

Based on the conducted review, this paper aimed to provide a comprehensive evaluation of Machine Learning Algorithms in the context of Network Security Intrusion Detection Systems (IDS). The objective was to analyze and assess the effectiveness of various machine learning algorithms in detecting and preventing network security intrusions. By undertaking a thorough examination of existing literature and research in the field, this study aimed to offer valuable insights into the performance,

Corresponding author: A. M. Giwa
engineermohammedagiwa@gmail.com
Department of Information and Communication Technology, Faculty of Ground and Communication Engineering, AFIT, Kaduna.

capabilities, and limitations of different machine learning algorithms employed in IDS. The evaluation encompassed a wide range of algorithms, considering their applicability, accuracy, and efficiency in detecting and mitigating network intrusions.

Through this comprehensive evaluation, the paper sought to contribute to the advancement of IDS technology by providing researchers, practitioners, and security professionals with a comprehensive understanding of the strengths and weaknesses of different machine learning algorithms. The findings of this study aimed to inform the development and implementation of more robust and effective intrusion detection systems, ultimately enhancing the overall security and resilience of network infrastructure

OVERVIEW OF IDS-BASED FUNDAMENTAL CONCEPTS

The objective of this section was to provide an extensive overview of IDS-based fundamental concepts that are based on machine learning. By delving into these concepts, the paper aimed to establish a solid foundation of knowledge for readers in the field of intrusion detection systems. The overview covered a wide range of essential concepts related to IDS and machine learning. It explored the fundamental principles and methodologies behind IDS and how they intertwine with machine learning techniques. This included an examination of key terms, definitions, and concepts relevant to IDS, as well as an exploration of the role of machine learning in enhancing the capabilities of intrusion detection systems. The section went beyond a surface-level understanding and delved into the intricacies of IDS-based fundamental concepts. It provided insights into the various components and processes involved in IDS, such as data collection, preprocessing, feature extraction, and model training. Additionally, it highlighted the significance of machine learning algorithms in analyzing and classifying network data to detect potential intrusions. By presenting this comprehensive overview, the paper aimed to equip readers with a solid understanding of the core principles and concepts underlying IDS and machine learning. This foundational knowledge served as a basis for the subsequent discussions and analyses presented in the paper, enabling readers to grasp the subsequent sections with a deeper level of comprehension and insight.

Concept of IDS System

The concept of Intrusion Detection Systems (IDS) revolves around the use of application software to protect and monitor networks or systems for malicious activities. IDS plays a critical role in maintaining the security of a network by constantly monitoring both hosts and networks for any suspicious behavior or unauthorized access (Shah *et al.,* 2020). One of the primary functions of IDS is to generate alarms or alerts when it detects abnormal or suspicious activities within computer systems. This can include unauthorized attempts to access resources, unusual network traffic patterns, or any other behavior that deviates from the established norms. By promptly identifying and alerting security personnel to potential threats, IDS helps in mitigating the risks associated with malicious activities (Buczak and Guven, 2016)). Figure 1 shows a sample IDS system model. Intrusion Detection Systems are designed to analyze various sources of data, such as network traffic logs, system logs, and event logs, to detect and respond to potential security breaches. IDS employs a range of techniques and algorithms to analyze the data and identify patterns or indicators of intrusion or malicious behavior. These techniques can include rule-based detection, anomaly detection, signature-based detection, and machine learning-based approaches. By continuously monitoring network and host activities, IDS can provide real-time or near-real-time detection of potential security incidents (Unnisa *et al.,* 2022).

Corresponding author: A. M. Giwa
engineermohammedagiwa@gmail.com
Department of Information and Communication Technology, Faculty of Ground and Communication Engineering, AFIT, Kaduna.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 11(2), JUNE, 2023
ISSN: 2277-0011; Journal homepage: www.atbuftejoste.net

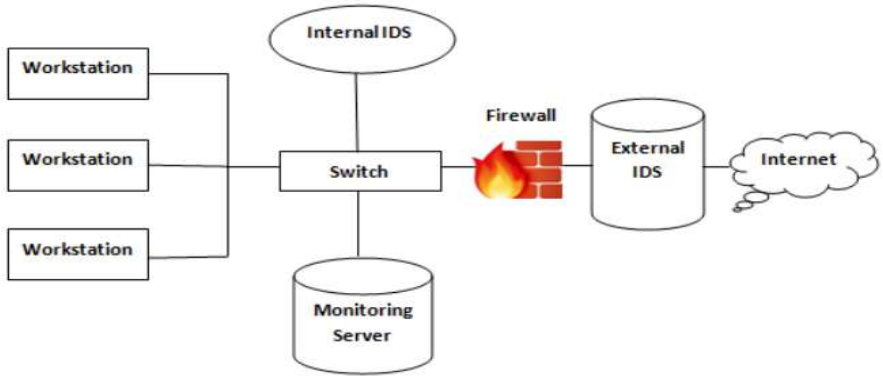

Figure 1: System Model for IDS (Unnisa *et al.,* (2022)

This allows security administrators to take appropriate actions to mitigate the threats and prevent further damage. IDS can also play a vital role in post-incident analysis and forensic investigations by providing detailed logs and information about the detected incidents**.** Conclusively, IDS is an application software that plays a crucial role in network security. It monitors hosts and networks, generates alarms based on abnormal behavior, and detects potential security threats or malicious activities. By providing timely alerts and analysis, IDS help in protecting the integrity, confidentiality, and availability of computer systems and networks.

IDS-Detection-based Methods

IDS detection methods can be broadly classified into two categories: signature-based detection and anomaly detection (See Table 1). These methods are designed to identify and flag potential security threats or malicious activities in a network or system (Unnis *et al.,* 2022).

Signature-Based Detection

Signature-based detection, also known as rule-based detection, relies on predefined signatures or patterns of known attacks. These signatures are created based on the characteristics and behaviors of previously identified malicious activities. The IDS compare incoming network traffic or system events against these signatures to determine if they match any known attack patterns. If a match is found, an alert is generated, indicating the presence of a specific type of attack. Signature-based detection is effective in detecting known attacks and is widely used in IDS systems. Signature-based detection is a fundamental method used in Intrusion Detection Systems (IDS) to identify known attacks and malicious activities. This approach relies on predefined signatures or patterns that correspond to specific attack behaviors. These signatures are created based on extensive analysis and understanding of past attacks and their characteristics (Diaz *et al*., 2022).

To implement signature-based detection, IDS systems employ a signature database that contains a collection of known attack patterns. This database is continuously updated and maintained to keep up with the evolving threat landscape (Shaikh and Gupta, 2022).). When network traffic or system events are analyzed, the IDS compare the observed data against the signatures in the database. If a match is found, indicating that the observed behavior corresponds to a known attack pattern, an alert is generated to notify the system administrator or security personnel. Signature-based detection offers several advantages. First, it is highly effective in detecting attacks for which signatures are available. Since the signatures are based on known attack patterns, this method can quickly and accurately identify well-established threats. Additionally, signature-based detection is computationally efficient, as the matching process involves comparing the observed data with a predefined set of signatures. However, signature-based detection also has limitations. It heavily

Corresponding author: A. M. Giwa
engineermohammedagiwa@gmail.com
Department of Information and Communication Technology, Faculty of Ground and Communication Engineering, AFIT, Kaduna.

relies on the availability of up-to-date and comprehensive signature databases. If a new or modified attack occurs that does not have a matching signature in the database, the IDS may fail to detect it. This makes signature-based detection less effective against novel or zero-day attacks that have not been previously encountered (Shaikh and Gupta, 2022).

Moreover, attackers can employ various evasion techniques to circumvent signature-based detection. By modifying attack patterns or using obfuscation methods, attackers can avoid triggering the existing signatures and thus remain undetected. This poses a significant challenge for signature-based IDS systems and highlights the need for continuous updates and improvements to the signature database. To address the limitations of signature-based detection, hybrid approaches are often employed. These combine signature-based detection with other methods, such as anomaly detection or behavior-based analysis, to enhance the overall effectiveness of intrusion detection. By integrating multiple detection techniques, IDS systems can achieve a higher detection rate, detect novel attacks, and reduce false positives.

Anomaly Detection

On the other hand, anomaly detection, as a complementary method to signature-based detection, takes a different approach to identifying intrusions within a network or system. Instead of relying on predefined attack signatures, anomaly detection algorithms aim to identify deviations from normal or expected behavior. To implement anomaly detection, an IDS establishes a baseline of normal behavior by analyzing historical data or training on a representative dataset. This baseline represents the typical patterns and characteristics of legitimate network traffic or system behavior. The IDS then continuously monitors incoming network traffic or system events and compares them to the established baseline. The comparison involves assessing various attributes and features of the observed data, such as packet size, protocol usage, communication patterns, resource utilization, or system processes. Statistical and machine learning techniques are often employed to analyze these attributes and detect significant deviations from the expected behavior (Martins *et al.,* 2022).

When a deviation or anomaly is detected, it indicates a potential intrusion or suspicious activity. The IDS generate an alert or raises a flag to notify the system administrator or security personnel, who can then investigate the detected anomaly and take appropriate actions. Anomaly detection is particularly effective in detecting novel or previously unseen attacks since it does not rely on predefined attack signatures. By learning from the normal behavior of the network or system, anomaly detection algorithms can identify unusual patterns that may indicate malicious activity. This makes anomaly detection a valuable approach for detecting zero-day attacks, where no prior knowledge of the attack patterns exists (Javaheri *et al.,* 2023). However, implementing effective anomaly detection can be challenging. Establishing a reliable baseline of normal behavior requires accurate and representative training data. It is important to consider the dynamic nature of networks and systems, as the baseline may need to be continuously updated to adapt to evolving conditions and changing patterns of normal behavior. Furthermore, anomaly detection can generate a higher number of false positives compared to signature-based detection. Legitimate but uncommon activities or system changes can trigger an alert, requiring careful analysis and tuning of the detection algorithms to minimize false positives and ensure efficient resource utilization. To overcome these challenges, hybrid approaches that combine signature-based and anomaly detection techniques are often used. By leveraging the strengths of both methods, IDS systems can achieve a more comprehensive and effective intrusion detection capability (Javaheri *et al.,* 2023).

Corresponding author: A. M. Giwa
engineermohammedagiwa@gmail.com
Department of Information and Communication Technology, Faculty of Ground and Communication Engineering, AFIT, Kaduna.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 11(2), JUNE, 2023
ISSN: 2277-0011; Journal homepage: www.atbuftejoste.net

Table 1: Differentiation between Signature-Based Detection and Anomaly Detection

Approach	Signature-based Detection	Anomaly Detection
	Matches incoming data against known attack patterns (signatures)	Identifies deviations from normal behavior patterns
Detection Basis	Predefined signatures of known attacks	Deviations from expected behavior
Training Required	No training required	Requires training on normal behavior
Detection Accuracy	High for known attacks	High for novel or previously unseen attacks
False Positives	Low	Moderate to high
False Negatives	Possible for new or modified attacks	Possible for subtle or stealthy attacks
Adaptability	Limited to known attack signatures	Adapts to evolving attack techniques
Real-time Analysis	Suitable for real-time analysis	Suitable for real-time or offline analysis
Maintenance	Regular updates for new attack signatures	Periodic updates for changes in normal behavior

Both signature-based detection and anomaly detection have their strengths and limitations. Signature-based detection is highly accurate in detecting known attacks, but it may fail to identify new or evolving attack techniques that do not match existing signatures. Anomaly detection, on the other hand, can detect unknown attacks or unusual behaviors, but it may also generate false positives due to legitimate variations in network or system activities. To enhance the effectiveness of IDS, many systems combine both detection methods. This approach, known as hybrid detection, leverages the advantages of both signature-based and anomaly detection techniques. By integrating these methods, IDS systems can achieve a higher detection rate, minimize false positives, and provide comprehensive coverage against a wide range of security threats. In conclusion, IDS detection methods are categorized into signature-based detection and anomaly detection. Signature-based detection relies on predefined attack signatures to identify known attacks, while anomaly detection focuses on detecting deviations from normal behavior. Hybrid approaches that combine both methods are commonly used to enhance the accuracy and effectiveness of IDS systems. By utilizing these methods, IDS plays a crucial role in identifying and mitigating security threats, ensuring the integrity and security of computer networks and systems.

Source of Data-based Method

Under the category of source of data-based methods, there are two distinct approaches: host-based IDS (HIDS) and network-based IDS (NIDS). Each method focuses on different sources of data for detecting intrusions and enhancing network security.

Host-based IDS (HIDS)

Host-based IDS (HIDS) primarily operates at the individual host or endpoint level. It involves monitoring and analyzing the activities and events occurring within a specific system. HIDS collects data from various sources on the host, including log files, system calls, audit trails, and file integrity checks. By examining these sources of data, HIDS can identify suspicious activities or deviations from normal behavior specific to that host. The advantage of HIDS is its ability to provide detailed and granular insights into the activities occurring within a specific host. It can detect attacks or intrusions that may not be

Corresponding author: A. M. Giwa
engineermohammedagiwa@gmail.com
Department of Information and Communication Technology, Faculty of Ground and Communication Engineering, AFIT, Kaduna.

easily observed at the network level. HIDS is particularly useful in protecting individual systems, such as servers or critical endpoints, and can provide valuable information for forensic analysis in case of a security incident (Martins *et al.,* 2022).

Network-based IDS

On the other hand, network-based IDS (NIDS) focus on monitoring and analyzing network traffic flowing through a network segment or an entire network infrastructure. NIDS examines packets and network flows to identify signs of malicious or anomalous behavior. It captures and analyzes network traffic using specialized network sensors or network taps strategically placed at key points within the network. NIDS can detect various network-based attacks, such as port scanning, denial-of-service (DoS) attacks, intrusion attempts, and unauthorized access attempts. By analyzing network traffic patterns, NIDS can detect suspicious activities, anomalies in protocol usage, and patterns indicative of known attack signatures (Gupta *et al.,* 2022).

The main advantage of NIDS is its ability to provide a network-wide view of the security posture and identify threats affecting multiple hosts or network segments. It can detect attacks targeting network infrastructure or widespread attacks that traverse multiple systems. NIDS can be deployed at strategic points within the network, such as at network gateways or switches, to monitor all incoming and outgoing traffic. Both HIDS and NIDS have their strengths and limitations. HIDS excels at providing detailed insights into individual host activities, while NIDS offers a broader perspective on network-wide security. Combining the capabilities of HIDS and NIDS can provide a more comprehensive and effective intrusion detection system, allowing organizations to detect and respond to a wide range of security threats (Amegoces *et al*., 2022).

IDS-Base Machine Learning Algorithm

Machine learning methods play a crucial role in the functioning of Intrusion Detection Systems (IDS) for detecting intrusions. These methods, which are closely related to artificial intelligence, leverage software programs to achieve more precise results without the need for explicit programming. By analyzing historical data, machine learning algorithms can generate new output values, thereby enhancing the detection capabilities of IDS. In this section, we explore a range of machine learning algorithms that are commonly employed in the implementation of IDS. These algorithms encompass diverse methodologies such as Support Vector Machine (SVM), K-nearest Neighbor (KNN), Artificial Neural Networks (ANN), Logistic Regression (LR), Naive Bayes, decision trees, clustering, and hybrid approaches. By utilizing these various algorithms, IDS systems can improve their effectiveness in identifying and classifying intrusions (Chinedu *et al*., 2020; Abubakar *et al*., 2021).

IDS-Based Artificial Neural Network (ANN)

IDS-Based Artificial Neural Network (ANN) is a specialized approach within Intrusion Detection Systems (IDS) that utilizes the power of artificial neural networks to detect and classify intrusions. Artificial neural networks are computational models inspired by the structure and functioning of biological neural networks in the human brain. In the context of IDS, the ANN-based approach involves training a neural network on historical data that contains both normal and malicious network traffic patterns. During the training phase, the neural network learns to recognize and differentiate between these patterns by adjusting its internal parameters and weights (Calugar *et al.,* 2022). Once the ANN is trained, it can be deployed in real-time to analyze incoming network traffic and identify any deviations or anomalies from normal behavior. The ANN-based IDS considers various features and attributes of the network traffic, such as packet headers, payload content, and traffic flow characteristics, to make accurate intrusion detection decisions. The strength of ANN lies in its ability to learn complex patterns and relationships from large volumes of data. It can capture intricate nonlinear dependencies and adapt to changing network environments, thereby improving the detection accuracy and reducing false positives. Additionally, ANNs can handle high-dimensional input data and handle feature extraction automatically, reducing the need for manual

Corresponding author: A. M. Giwa
engineermohammedagiwa@gmail.com
Department of Information and Communication Technology, Faculty of Ground and Communication Engineering, AFIT, Kaduna.

feature engineering (Calugar *et al.,* 2022). However, deploying an ANN-based IDS requires careful considerations such as selecting the appropriate architecture, determining the optimal number of layers and neurons, and fine-tuning the network parameters to achieve optimal performance. Training an ANN can be computationally intensive and may require a substantial amount of labeled training data to ensure effective learning. Generally, IDS-Based Artificial Neural Network offers a powerful and flexible approach for intrusion detection, leveraging the capabilities of artificial neural networks to enhance network security by accurately identifying and classifying intrusions in real-time.

Methods for IDS-Based Artificial Neural Network (ANN)

There are several methods that can be employed for IDS-Based Artificial Neural Network (ANN) in the context of intrusion detection. Some of the commonly used methods include: Feedforward Neural Networks: This is the most basic type of neural network architecture where information flows only in one direction, from the input layer through the hidden layers to the output layer. It is widely used in IDS for pattern recognition and classification tasks (Wu and Hu, 2022). The block diagram for ANN pattern classification AND THE steps for activities is giving in Figures 2 and 3 respectively.

Recurrent Neural Networks (RNN): RNNs are designed to capture sequential dependencies in data by allowing feedback connections, enabling them to handle time-series data effectively. They are suitable for analyzing network traffic and detecting intrusions that exhibit temporal characteristics.

Long Short-Term Memory (LSTM): LSTM is a type of RNN that addresses the vanishing gradient problem and can learn long-term dependencies. It has shown promising results in IDS applications, especially in capturing complex patterns and anomalies in network traffic.

Convolutional Neural Networks (CNN): CNNs are commonly used in image processing, but they can also be applied to network traffic analysis by treating traffic data as multidimensional inputs. They are effective in capturing local patterns and spatial dependencies, making them suitable for detecting attacks based on traffic content.

Hybrid Models: Hybrid models combine multiple neural network architectures or incorporate other machine learning techniques to enhance intrusion detection capabilities. For example, combining an ANN with a decision tree or ensemble methods can improve the overall accuracy and robustness of the IDS.

Transfer Learning: Transfer learning involves leveraging pre-trained neural networks on a related task or dataset and adapting them to the specific IDS domain. This approach can help overcome data scarcity issues and accelerate the training process by utilizing the learned representations from a different but relevant task.

Corresponding author: A. M. Giwa
engineermohammedagiwa@gmail.com
Department of Information and Communication Technology, Faculty of Ground and Communication Engineering, AFIT, Kaduna.

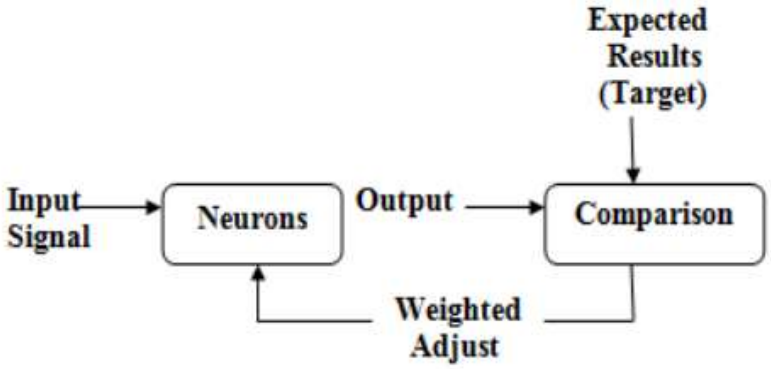

Figure 2: ANN Block diagram for Pattern Classification (Unnisa *et al.,* (2022)

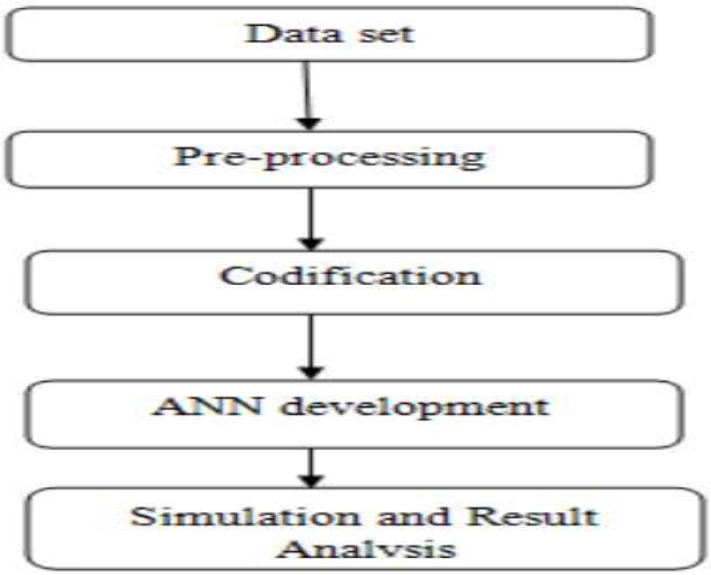

Figure 3: ANN Activity Model (Unnisa *et al.,* (2022)

SVM-based IDS

Intrusion Detection Systems (IDS) using Support Vector Machine (SVM) is a popular approach in the field of network security. SVM is a supervised machine learning algorithm that has been successfully applied to various classification tasks, including intrusion detection. The schematic structure of SVM is shown in Figure 4. SVM is and malicious activities (Zhou *et al.,* 2022).

based on the concept of finding an optimal hyperplane that maximally separates different classes of data points. In the context of IDS, SVM can be trained on labeled datasets containing both normal and attack instances. The goal is to learn a decision boundary that effectively discriminates between normal network traffic

Corresponding author: A. M. Giwa
engineermohammedagiwa@gmail.com
Department of Information and Communication Technology, Faculty of Ground and Communication Engineering, AFIT, Kaduna.

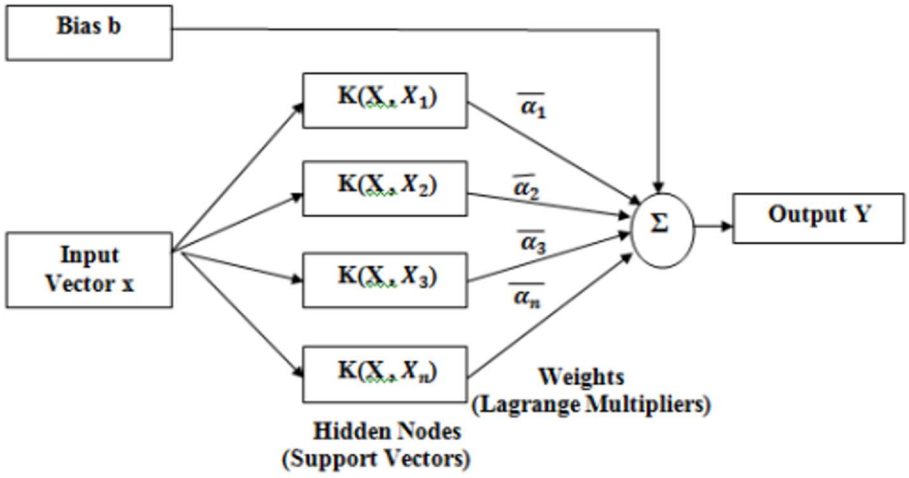

Figure 4: Schematic Diagram of SVM structure (Unnisa *et al.,* (2022)

The process of implementing IDS using SVM typically involves the following steps:

Data Preprocessing: The dataset is prepared by preprocessing the raw network traffic data. This may include removing noise, normalizing the data, and selecting relevant features that are indicative of network intrusions.

Feature Extraction: Relevant features are extracted from the preprocessed data. These features may include packet size, protocol type, source/destination IP addresses, port numbers, and other network-level attributes.

Training Phase: The preprocessed data is divided into training and testing sets. The training set is used to train the SVM classifier by finding the optimal hyperplane that maximizes the margin between different classes. The SVM algorithm learns the boundary based on the labeled instances, classifying them as normal or malicious.

Testing Phase: The trained SVM model is then evaluated on the testing set to assess its performance in detecting intrusions. The model predicts the class label (normal or attack) for each instance in the testing set, and the results are compared with the ground truth labels to calculate various evaluation metrics such as accuracy, precision, recall, and F1 score.

Model Optimization: The SVM model parameters, such as the choice of kernel function, regularization parameter, and kernel width, can be fine-tuned through techniques like cross-validation or grid search to optimize the model's performance (Zhang *et al.,* 2022).

Advantages and Considerations of SVM-based IDS

IDS using SVM offers several advantages, including:

1. SVM is effective in handling high-dimensional feature spaces and can handle complex decision boundaries.
2. SVM has good generalization capabilities, meaning it can perform well on unseen data.
3. SVM is less prone to overfitting, making it robust against noise and outliers in the dataset.
4. SVM can handle both linear and nonlinear classification tasks by utilizing different kernel functions.

However, there are also some considerations when using SVM for IDS:

Corresponding author: A. M. Giwa
engineermohammedagiwa@gmail.com
Department of Information and Communication Technology, Faculty of Ground and Communication Engineering, AFIT, Kaduna.

1. SVM can be computationally expensive, especially when dealing with large datasets.
2. The choice of the appropriate kernel function and its parameters is crucial for achieving good classification performance.
3. SVM may require a significant amount of labeled training data, which can be a challenge in the context of intrusion detection.

Despite these considerations, IDS using SVM has demonstrated promising results in detecting network intrusions. Researchers continue to explore and enhance the performance of SVM-based IDS systems through novel feature selection methods, ensemble techniques, and hybrid approaches that combine SVM with other machine learning algorithms.

KNN-Based IDS

Intrusion Detection Systems (IDS) using the K-Nearest Neighbor (KNN) algorithm is another approach commonly used in network security. KNN is a non-parametric and instance-based machine learning algorithm that classifies data based on their similarity to neighboring instances. The IDS using KNN follows a simple yet effective principle: it classifies an unknown instance by comparing it to its k nearest neighbors in the training dataset (Abdedaime *et al.,* 2022). The class label assigned to the unknown instance is determined by the majority vote of its k nearest neighbors. The KNN classification algorithm is shown in Figure 5.

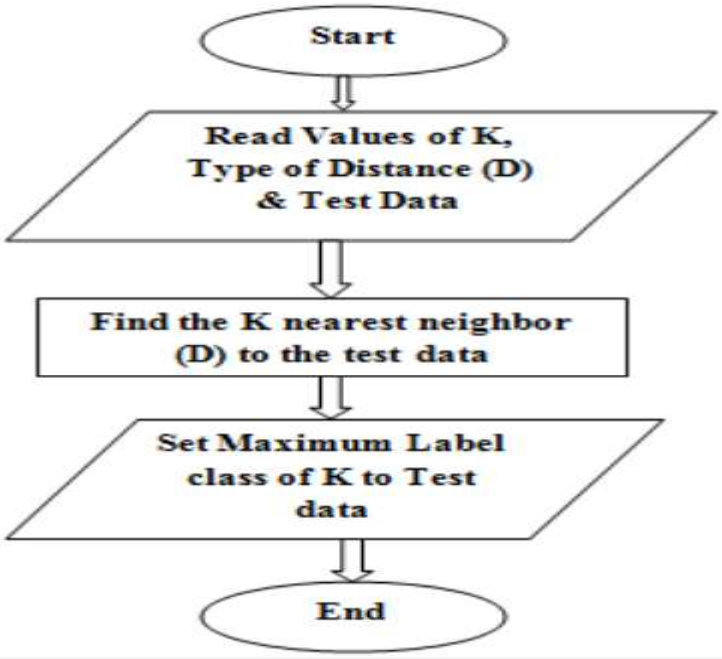

Figure 5: KNN classification Algorithm (Unnisa *et al.,* (2022)

Steps involved in implementing IDS using KNN

The steps involved in implementing IDS using KNN are as follows:

Data Preparation: The raw network traffic data is preprocessed to remove noise, normalize the features, and select relevant attributes for intrusion detection. The dataset is then divided into training and testing sets.

Feature Selection: Relevant features are extracted from the preprocessed data, which may

Corresponding author: A. M. Giwa
engineermohammedagiwa@gmail.com
Department of Information and Communication Technology, Faculty of Ground and Communication Engineering, AFIT, Kaduna.

include attributes such as packet size, source/destination IP addresses, port numbers, protocol type, and other network-level characteristics. Careful selection of features is essential to capture the distinguishing patterns between normal and malicious activities.

Training Phase: During the training phase, the KNN algorithm builds a model by storing the feature vectors and corresponding class labels of the training instances. The value of k, representing the number of nearest neighbors, is determined based on experimentation or domain knowledge.

Testing Phase: The trained KNN model is used to classify unknown instances in the testing dataset. For each test instance, the KNN algorithm calculates the distances to its k nearest neighbors in the training dataset. The class label assigned to the test instance is determined by the majority class among its neighbors.

Evaluation and Performance Metrics: The performance of the IDS using KNN is evaluated by comparing the predicted class labels with the true labels in the testing dataset. Various performance metrics such as accuracy, precision, recall, and F1 score are calculated to assess the effectiveness of the system in detecting intrusions.

Advantage and Disadvantages of KNN-based IDS

IDS using KNN offers some advantages, including (Abdedaime *et al.,* 2022):

1. **Simplicity**: KNN is a straightforward and easy-to-understand algorithm, making it accessible for researchers and practitioners.
2. **Flexibility:** KNN can handle both binary and multi-class classification problems, allowing it to detect various types of network intrusions.
3. **Adaptability:** KNN can adapt to changes in the dataset and detect new types of intrusions without retraining the model.
4. However, there are also some considerations when using KNN for IDS which include the following:
5. **Computational Complexity**: KNN can be computationally expensive, especially with large datasets and high-dimensional feature spaces. Efficient data structures like kd-trees or ball trees are often employed to speed up the nearest neighbor search.
6. **Distance Metric:** The choice of an appropriate distance metric is crucial for accurate classification. Different distance metrics, such as Euclidean distance or cosine similarity, may be more suitable depending on the dataset and the nature of the features.
7. **Determining the Value of k:** The value of k needs to be carefully chosen to balance between underfitting and overfitting. An improper choice of k may lead to biased or inaccurate predictions.

Researchers continue to explore advancements in KNN-based IDS systems, such as incorporating feature selection techniques, ensemble methods, and hybrid approaches to enhance the classification performance. Additionally, efforts are made to optimize the computational efficiency of KNN to handle real-time intrusion detection scenarios.

LR-Based IDS

Intrusion Detection Systems (IDS) using Logistic Regression (LR) is a popular approach in network security for detecting and classifying intrusions. Logistic Regression is a supervised learning algorithm that is commonly used for binary classification problems. It is well-suited for IDS applications where the objective is to determine whether a network activity is normal or represents a potential intrusion. Figure 6 represents the model for RL (Besharah *et al.,* 2019).

Corresponding author: A. M. Giwa
engineermohammedagiwa@gmail.com
Department of Information and Communication Technology, Faculty of Ground and Communication Engineering, AFIT, Kaduna.

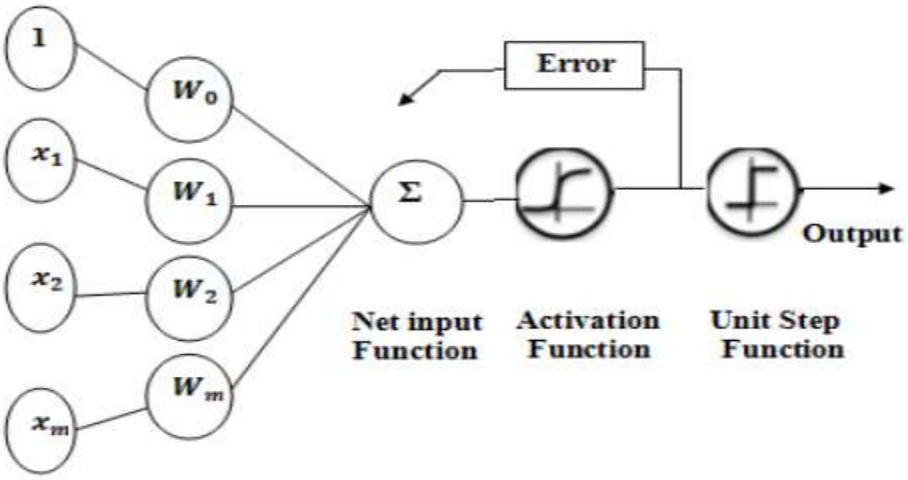

Figure 6: Logistic Regression Model (Unnisa *et al.,* (2022)

The steps for IDS using Logistic Regression include (Besharah *et al.,* 2019):

Data Preparation: The first step in implementing IDS using LR is to prepare the dataset. This involves collecting network traffic data, preprocessing it to remove noise and irrelevant information, and transforming it into a suitable format for LR. The dataset is then divided into training and testing sets.

Feature Extraction: Relevant features are extracted from the preprocessed data to capture the characteristics that distinguish normal network behavior from malicious activities. These features can include attributes such as packet size, source and destination IP addresses, protocol type, port numbers, and timestamps. Feature selection techniques may be applied to select the most informative features for intrusion detection.

Model Training: In the training phase, the LR algorithm learns the parameters of the logistic regression model based on the labeled training data. The objective is to find the optimal set of weights that maximizes the likelihood of the observed data. The LR model estimates the probability of an instance belonging to the intrusion class based on the feature values.

Model Testing: Once the LR model is trained, it can be used to classify unknown instances in the testing dataset. The model calculates the probability of an instance being an intrusion based on its feature values. A decision threshold is then applied to assign a class label (normal or intrusion) to each instance. The threshold can be adjusted to control the trade-off between false positives and false negatives, depending on the desired level of detection accuracy.

Evaluation and Performance Metrics: The performance of the IDS using LR is evaluated using various performance metrics, including accuracy, precision, recall, and F1 score. These metrics provide insights into the effectiveness of the LR model in correctly classifying normal and intrusive network activities. Additionally, techniques like cross-validation can be employed to assess the generalization performance of the model.

Advantages of IDS using Logistic Regression

The following stands for the merits of using LR for IDS (Besharah *et al.,* 2019):

Interpretable Results: LR provides interpretable coefficients that indicate the impact of each feature on the likelihood of an intrusion. This allows security analysts to gain insights into the

Corresponding author: A. M. Giwa
engineermohammedagiwa@gmail.com
Department of Information and Communication Technology, Faculty of Ground and Communication Engineering, AFIT, Kaduna.

important factors contributing to the classification decisions.

Handling Imbalanced Data: Logistic Regression can handle imbalanced datasets, which are common in intrusion detection, where the number of normal instances typically outweighs the number of intrusions.

Scalability: LR is computationally efficient and can handle large-scale datasets, making it suitable for real-time or high-speed network environments.

Considerations and Challenges LR-based IDS

The following stands for the demerits of using LR for IDS:

Feature Selection: Selecting relevant and informative features is crucial for the performance of LR-based IDS. Careful consideration should be given to selecting features that capture the distinguishing patterns of normal and intrusive activities.

Generalization: The LR model needs to generalize well to new and unseen network data. Regularization techniques, such as L1 or L2 regularization, can be applied to prevent overfitting and improve generalization performance.

Handling Multiclass Classification: LR is originally designed for binary classification. To handle multiclass intrusion detection problems, techniques like one-vs-rest or one-vs-one can be employed.

Researchers continue to explore advancements in IDS using Logistic Regression, such as incorporating ensemble methods, feature engineering techniques, and hybrid approaches to improve the detection accuracy and handle complex network environments.

Naïve Bayes and Decision Tress IDS Method

IDS using Naïve Bayes and Decision Trees are two commonly used machine learning algorithms for intrusion detection (Almomani *et al.,* 2021). Figures 7 and 8 represents the Naive Bayes classification steps and Decision Tree Hierarchy respectively.

IDS using Naïve Bayes

Naïve Bayes is a probabilistic classification algorithm based on Bayes' theorem with the assumption of independence between features. It is widely used in various applications, including intrusion detection. Here's how IDS using Naïve Bayes works (Gu and Lu, 2021):

1. **Data Preparation:** Similar to other IDS approaches, the dataset is collected and preprocessed to remove noise, irrelevant information, and normalize the data. The dataset is then divided into training and testing sets.
2. **Feature Extraction:** Relevant features are extracted from the preprocessed data to represent network traffic characteristics. These features can include packet attributes, protocol types, IP addresses, port numbers, and more. Feature selection techniques can be applied to select the most discriminative features for intrusion detection.
3. **Model Training**: In the training phase, the Naïve Bayes algorithm estimates the probability distributions of features for each class (normal or intrusion) based on the training data. It assumes that features are conditionally independent given the class label, which is a simplifying assumption but can still provide effective results for many intrusion detection scenarios.
4. **Model Testing**: Once the Naïve Bayes model is trained, it can be used to classify unknown instances in the testing dataset. The model calculates the probability of an instance belonging to each class (normal or intrusion) based on the observed feature values. The class label with the highest probability is assigned to the instance.
5. **Evaluation and Performance Metrics:** The performance of the IDS using Naïve Bayes is evaluated using various metrics such as accuracy, precision,

Corresponding author: A. M. Giwa
engineermohammedagiwa@gmail.com
Department of Information and Communication Technology, Faculty of Ground and Communication Engineering, AFIT, Kaduna.

recall, and F1 score. These metrics assess the model's ability to correctly classify normal and intrusive instances. Cross-validation can be employed to estimate the model's generalization performance.

Advantages of IDS using Naïve Bayes

1. **Efficiency:** Naïve Bayes is computationally efficient and can handle large-scale datasets, making it suitable for real-time or high-speed network environments.
2. **Handling Missing Values**: Naïve Bayes can handle missing values in the dataset by ignoring the missing values during probability calculations.
3. **Interpretability:** Naïve Bayes provides interpretable results as it estimates the probabilities of an instance belonging to each class based on feature values.
4.

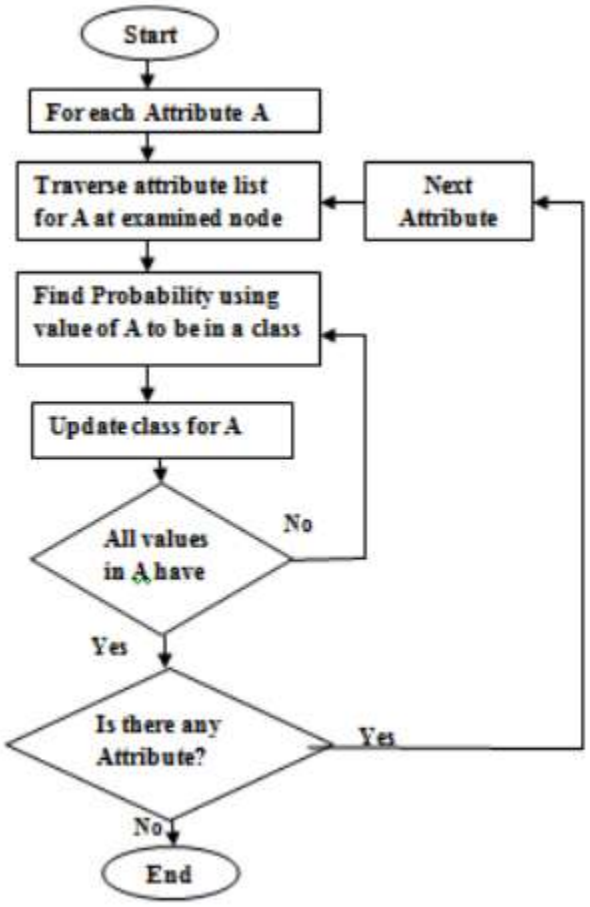

Figure 7: Naive Bayes classification

IDS using Decision Trees

Decision Trees are non-parametric supervised learning algorithms that recursively split the dataset based on the values of features to create a hierarchical decision structure. Here's how IDS using Decision Trees works (Mahbooba *et al.,* 2021):

1. **Data Preparation:** Similar to other IDS approaches, the dataset is collected,

Corresponding author: A. M. Giwa
engineermohammedagiwa@gmail.com
Department of Information and Communication Technology, Faculty of Ground and Communication Engineering, AFIT, Kaduna.

preprocessed, and divided into training and testing sets.

2. **Feature Extraction:** Relevant features are extracted from the preprocessed data to capture the network behavior. These features can include packet attributes, protocol types, source/destination IP addresses, port numbers, and more.
3. **Model Training**: In the training phase, the Decision Tree algorithm builds a tree-like structure by recursively partitioning the dataset based on the values of features. It selects the best splitting criterion (e.g., information gain or Gini index) to create nodes that maximize the separation between normal and intrusive instances.
4. **Model Testing:** Once the Decision Tree model is trained, it can be used to classify unknown instances in the testing dataset. The model follows the decision path based on feature values until it reaches a leaf node, which represents a class label (normal or intrusion).

Evaluation and Performance Metrics: The performance of the IDS using Decision Trees is evaluated using metrics like accuracy, precision, recall, and F1 score. The tree's structure can be visualized to interpret the decision-making process.

Advantages of IDS using Decision Trees

Interpretability: Decision Trees provide interpretable results as the decision rules can be easily understood and visualized. It allows security analysts to gain insights into the features and rules used for intrusion detection (Mahbooba *et al.,* 2021).

1. **Handling Nonlinear Relationships:** Decision Trees can capture nonlinear relationships between features and the target variable, making them suitable for complex intrusion detection scenarios.
2. **Handling Missing Values:** Decision Trees can handle missing values in the dataset by using surrogate splits or imputation techniques.
3. **Feature Importance:** Decision Trees provide a measure of feature importance, indicating which features contribute the most to the classification task.

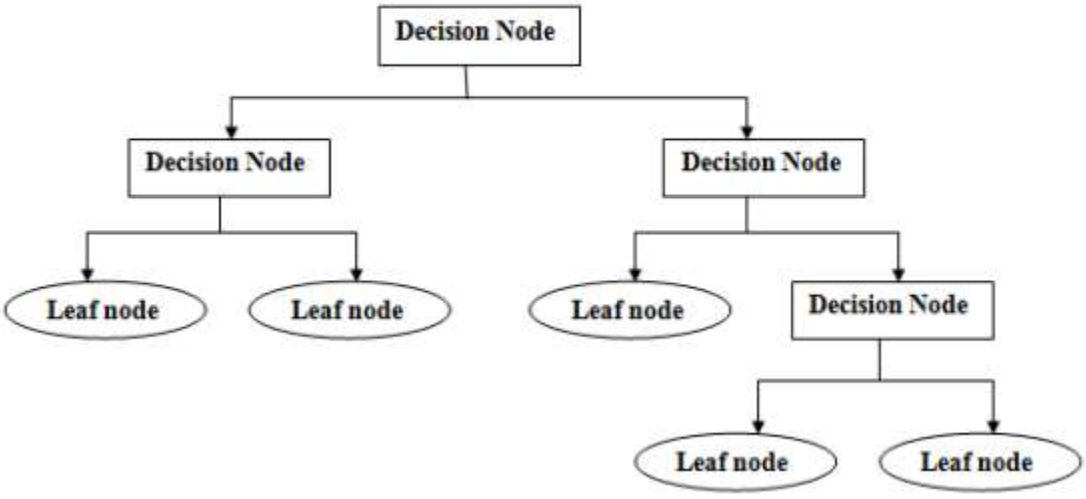

Figure 8: Decision Tree Hierarchy

Corresponding author: A. M. Giwa
engineermohammedagiwa@gmail.com
Department of Information and Communication Technology, Faculty of Ground and Communication Engineering, AFIT, Kaduna.

It is noteworthy that both Naïve Bayes and Decision Trees offer advantages in IDS, and their performance may vary depending on the specific dataset and intrusion detection requirements. Researchers and practitioners continue to explore enhancements and hybrid approaches to improve the accuracy and effectiveness of IDS using these algorithms.

Comparative Analysis of Machine Learning-based IDS

Intrusion Detection Systems (IDS) rely heavily on machine learning methods to detect intrusions. These artificial intelligence-related algorithms use software programs to produce more accurate results without the use of explicit programming. Machine learning algorithms are able to generate new output values by examining old data. The nature of the data must therefore be understood in order to apply these algorithms correctly. The numerous machine learning algorithms examined in this section are those used in IDS implementation. These algorithms cover a wide variety of methodologies, including Support Vector Machine (SVM), K-nearest Neighbor (KNN), Artificial Neural Networks (ANN), Logistic Regression (LR), Naive Bayes, decision trees, clustering, and hybrid approaches. These various algorithms can be used by IDS systems to improve their ability to identify and categorize intrusions. Table 2 shows the comparative analysis of machine learning-based IDS.

Table 2: Comparative Analysis of Machine Learning-based IDS

Algorithm	Advantages	Disadvantages	Similarities	Proposed Solutions to Disadvantages
ANN	- Ability to learn complex patterns	Requires large amounts of labeled data	Nonlinear modeling	Use transfer learning or pre-trained models
SVM	- Effective in high-dimensional spaces	Can be computationally expensive	Binary classification	Employ kernel tricks or use linear SVM
KNN	- Simple and intuitive	Sensitive to irrelevant and noisy features	Instance-based learning	Feature selection or dimensionality reduction
LR	- Interpretability	Limited to linear relationships	Binary/multiclass classification	Use polynomial or interaction terms to capture complexity
Naive Bayes	- Efficient and scalable	Relies on the assumption of feature independence	Probabilistic modeling	Use more advanced probabilistic models
Decision Trees	- Interpretability	Prone to overfitting	Hierarchical decision-making	Use pruning techniques or ensemble methods
K-means	- Simple and fast	Sensitivity to initialization and outliers	Clustering-based approach	Perform multiple runs with different initializations

CONCLUSION

Overall, this survey contributes to the growing body of knowledge on IDS and machine learning algorithms, providing valuable insights for network security practitioners and researchers seeking to enhance the protection of networks against emerging threats in the era of advanced AI technologies and fifth-generation networks.

REFERENCES

Abdedaime, M., Qafas, A., Jerry, M., & Guezzaz, A. (2022, November). A KNN-based intrusion detection model for smart cities security. In *International Conference on Innovative Computing and Communications: Proceedings of*

Corresponding author: A. M. Giwa
engineermohammedagiwa@gmail.com
Department of Information and Communication Technology, Faculty of Ground and Communication Engineering, AFIT, Kaduna.

ICICC 2022, Volume 3 (pp. 265-272). Singapore: Springer Nature Singapore.

Abubakre, O. R., Ubadike, O., Aibinu, A. M., Folorunso, T. A., & Momoh, M. O. (2021). Artificial Neural Network Embedded Optimal Mobile Communication System. *Covenant Journal of Informatics and Communication Technology*.

Alkasassbeh, M., & Almseidin, M. (2018). Machine learning methods for network intrusion detection. *arXiv preprint arXiv:1809.02610*.

Almomani, O., Almaiah, M. A., Alsaaidah, A., Smadi, S., Mohammad, A. H., & Althunibat, A. (2021, July). Machine learning classifiers for network intrusion detection system: comparative study. In *2021 International Conference on Information Technology (ICIT)* (pp. 440-445). IEEE.

Arregoces, P., Vergara, J., Gutiérrez, S. A., & Botero, J. F. (2022, April). Network-based intrusion detection: A one-class classification approach. In *NOMS 2022-2022 IEEE/IFIP Network Operations and Management Symposium* (pp. 1-6). IEEE.

Aung, Y. Y., & Min, M. M. (2018). An analysis of K-means algorithm based network intrusion detection system. *Advances in Science, Technology and Engineering Systems Journal*, *3*(1), 496-501.

Belavagi, M. C., & Muniyal, B. (2017). Multi class machine learning algorithms for intrusion detection-a performance study. In *Security in Computing and Communications: 5th International Symposium, SSCC 2017, Manipal, India, September 13–16, 2017, Proceedings 5* (pp. 170-178). Springer Singapore.

Belavagi, M. C., & Muniyal, B. (2017). Multi class machine learning algorithms for intrusion detection-a performance study. In *Security in Computing and Communications: 5th International Symp*

Besharati, E., Naderan, M., & Namjoo, E. (2019). LR-HIDS: logistic regression host-based intrusion detection system for cloud environments. *Journal of Ambient Intelligence and Humanized Computing*, *10*, 3669-3692.

Buczak, Anna L. and Erhan Guven. "A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection". IEEE Communications Surveys & Tutorials 18.2 (2016): 1153–1176. 10 . 1109 / comst . 2015 . 2494502.

Calugar, A. N., Meng, W., & Zhang, H. (2022, December). Towards Artificial Neural Network Based Intrusion Detection with Enhanced Hyperparameter Tuning. In *GLOBECOM 2022-2022 IEEE Global Communications Conference* (pp. 2627-2632). IEEE.

Chauhan, M., Joon, A., Agrawal, A., Kaushal, S., & Kumari, R. (2021). Intrusion detection system for securing computer networks using machine learning: a literature review. In *Congress on Intelligent Systems: Proceedings of CIS 2020, Volume 1* (pp. 177-189). Springer Singapore.

Chinedu, P. U., Nwankwo, W., Aliu, D., Shaba, S. M., & Momoh, M. O. (2020). Cloud security concerns: assessing the fears of service adoption. *Archive of Science and Technology*, *1*(2), 164-174.

Daniel, A., Shaba, S. M., Momoh, M. O., Chinedu, P. U., & Nwankwo, W. (2021). A computer security system for cloud computing based on encryption technique. *Computer Engineering and Applications*, *10*(1), 41-53.

Díaz-Verdejo, J., Muñoz-Calle, J., Estepa Alonso, A., Estepa Alonso, R., & Madinabeitia, G. (2022). On the detection capabilities of signature-based intrusion detection systems in the context of web attacks. *Applied Sciences*, *12*(2), 852.

Gu, J., & Lu, S. (2021). An effective intrusion detection approach using SVM with naïve Bayes feature

Corresponding author: A. M. Giwa
✉ engineermohammedagiwa@gmail.com
Department of Information and Communication Technology, Faculty of Ground and Communication Engineering, AFIT, Kaduna.

embedding. *Computers & Security*, *103*, 102158.

Gupta, N., Jindal, V., & Bedi, P. (2022). CSE-IDS: Using cost-sensitive deep learning and ensemble algorithms to handle class imbalance in network-based intrusion detection systems. *Computers & Security*, *112*, 102499.

Javaheri, D., Gorgin, S., Lee, J. A., & Masdari, M. (2023). Fuzzy logic-based DDoS attacks and network traffic anomaly detection methods: Classification, overview, and future perspectives. *Information Sciences*.

Mahbooba, B., Timilsina, M., Sahal, R., & Serrano, M. (2021). Explainable artificial intelligence (XAI) to enhance trust management in intrusion detection systems using decision tree model. *Complexity*, *2021*, 1-11.

Martins, I., Resende, J. S., Sousa, P. R., Silva, S., Antunes, L., & Gama, J. (2022). Host-based IDS: A review and open issues of an anomaly detection system in IoT. *Future Generation Computer Systems*, *133*, 95-113.

Mohamed, H., Hefny, H., & Alsawy, A. (2018). Intrusion Detection System Using Machine Learning Approaches. *Egyptian Computer Science Journal*, *42*(3).

Rekha, G., Malik, S., Tyagi, A. K., & Nair, M. M. (2020). Intrusion detection in cyber security: role of machine learning and data mining in cyber security. *Advances in Science, Technology and Engineering Systems Journal*, *5*(3), 72-81.

Shah, Ajay, et al. "Building Multiclass Classification Baselines for Anomaly-based Network Intrusion Detection Systems". IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA) (2020): 759–760. 10. 1109 / DSAA49011.2020.00102.

Shaikh, A., & Gupta, P. (2022). Advanced signature-based intrusion detection system. In *Intelligent Communication Technologies and Virtual Mobile Networks: Proceedings of ICICV 2022* (pp. 305-321). Singapore: Springer Nature Singapore.

Shobowale, K. O., Mukhtar, Z., Yahaya, B., Ibrahim, Y., & Momoh, M. O. (2023). Latest advances on security architecture for 5G technology and services. *International Journal of Software Engineering and Computer Systems*, *9*(1), 27-38.

Unnisa A, N., Yerva, M., & MZ, K. (2022). Review on intrusion detection system (ids) for network security using machine learning algorithms. *International Research Journal on Advanced Science Hub*, *4*(03), 67-74.

Wu, Y., & Hu, X. (2022). An Intrusion Detection Method Based on Fully Connected Recurrent Neural Network. *Scientific Programming*, *2022*.

Zala, J., Panchal, A., Thakkar, A., Prajapati, B., & Puvar, P. (2020). Intrusion Detection System using Machine Learning. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, *42*, 61-71.

Zhang, F., Zhen, P., Jing, D., Tang, X., Chen, H. B., & Yan, J. (2022). SVM based intrusion detection method with nonlinear scaling and feature selection. *IEICE TRANSACTIONS on Information and Systems*, *105*(5), 1024-1038.

Zhou, Y., Xie, L., & Pan, H. (2022). Research on a PSO-H-SVM-Based Intrusion Detection Method for Industrial Robotic Arms. *Applied Sciences*, *12*(6), 2765.

Corresponding author: A. M. Giwa
✉ engineermohammedagiwa@gmail.com
Department of Information and Communication Technology, Faculty of Ground and Communication Engineering, AFIT, Kaduna.