Journal of Harbin Engineering University
ISSN: 1006-7043

Vol 44 No. 9
September 2023

Advancing Intrusion Detection Systems: A Review of Machine Learning and Deep Learning Approaches

Apoorva Jain[1,2], Renu Bagoria[1] and Praveen Arora[3]
[1] Jagannath University, Jaipur
[2] IILM University, Greater Noida
[3] Jagan Institute of Management Studies, Delhi

Abstract: In contemporary computer networks, the utilisation of intrusion detection systems (IDS) plays a vital role in fortifying the security posture against the continuously expanding array of cyber threats. Nevertheless, the effectiveness of traditional rule-based Intrusion Detection Systems (IDS) in detecting and mitigating contemporary threats is hindered by their inherent complexity and sophistication. The application of machine learning (ML) and deep learning (DL) techniques in intrusion detection systems (IDS) has garnered considerable interest as a means to tackle this challenge. The objective of this review paper is to investigate the application of machine learning (ML) and deep learning (DL) methodologies in intrusion detection systems (IDS) with the purpose of improving accuracy, efficiency, and robustness.The paper begins by presenting a comprehensive introduction to conventional Intrusion Detection Systems (IDS) and underscoring their inherent limitations, thereby emphasising the imperative need for more sophisticated Machine Learning (ML) and Deep Learning (DL) methodologies. The text underscores the significance of machine learning and deep learning-based intrusion detection systems (IDS) in addressing the ever-changing landscape of cyber threats. Additionally, it explores prevalent architectures utilised in this domain, including convolutional neural networks (CNNs) and recurrent neural networks (RNNs). Furthermore, the review examines the utilisation of transfer learning and pretraining methods, which have the potential to mitigate the problem of limited data availability and enhance the ability of intrusion detection system models to generalise.
This review paper offers a thorough examination of machine learning (ML) and deep learning (DL) methodologies within the context of intrusion detection systems (IDS). The text underscores the necessity for employing sophisticated methodologies to address the constraints associated with conventional Intrusion Detection Systems (IDS). It delves into multiple facets encompassing architectures, datasets, preprocessing techniques, machine learning (ML) algorithms, and evaluation metrics. The paper additionally provides insights into emerging trends and potential areas for future research in machine learning and deep learning-based intrusion detection systems. It emphasises the significance of addressing crucial challenges in order to facilitate progress in the field of intrusion detection..

Keywords: Intrusion detection systems, Machine learning, Deep learning, Cybersecurity, Network security, IDS architectures, Transfer learning, Anomaly detection.

Introduction

The security of computer systems and networks has emerged as a critical concern in the contemporary interconnected digital environment. In the realm of system security, Intrusion Detection Systems (IDS) hold significant importance as they effectively identify and thwart unauthorised access and malicious activities [1]. This section presents a thorough exposition of Intrusion Detection Systems (IDS), with a particular emphasis on their importance in contemporary cybersecurity [2]. In addition, this paper explores the utilisation of machine learning (ML) and deep learning (DL) methodologies in intrusion detection systems (IDS), alongside an examination of the foundational elements of conventional intrusion detection systems.

An Intrusion Detection System (IDS) functions as a security mechanism that is specifically engineered to monitor and analyse network traffic or system events with the objective of detecting any potentially suspicious or malicious activities in real-time or near real-time. The implementation of Intrusion Detection Systems (IDS) serves to alleviate the potential harm resulting from cyberattacks [3]. In the

realm of intrusion detection systems (IDS), a fundamental categorization can be made, encompassing two primary types: signature-based and anomaly-based. The intrusion detection system (IDS) that is signature-based operates by utilising pre-established patterns or signatures that are associated with recognised attacks. The process involves the comparison of incoming network traffic or system events with a pre-existing database of signatures in order to detect any matches and subsequently generate alerts that indicate potential intrusions. Nevertheless, this methodology demonstrates inherent constraints as it is unable to identify novel or previously unobserved forms of attacks.

On the other hand, anomaly-based intrusion detection systems (IDS) prioritise the detection of deviations from typical system behaviour. The system establishes a standard level of typical activity and generates notifications when observed behaviour deviates significantly from this established norm. The effectiveness of anomaly-based intrusion detection systems (IDS) in detecting unknown or zero-day attacks has been demonstrated [4]. However, it is important to note that this approach may also result in false positives as a consequence of genuine variations in system behaviour.

The increasing intricacy and regularity of cyber threats require the implementation of resilient Intrusion Detection Systems (IDS) in order to safeguard sensitive data and vital infrastructures. In the realm of cybersecurity, Intrusion Detection Systems (IDS) serve a crucial role as a proactive mechanism for identifying and alerting stakeholders to potential security breaches. By promptly detecting and notifying relevant parties, IDS facilitates timely response and the implementation of appropriate measures to mitigate the impact of these breaches. Through vigilant monitoring of network traffic and system events, intrusion detection systems (IDS) are able to effectively identify and detect a wide range of attacks, such as unauthorised access attempts, malware infections, data breaches, and denial-of-service attacks [5]. Furthermore, Intrusion Detection Systems (IDS) assume a crucial function in adhering to industry regulations and standards. Numerous organisations are required to incorporate IDS into their cybersecurity strategies to safeguard customer data, intellectual property, and the uninterrupted operation of their business. Furthermore, Intrusion Detection Systems (IDS) play a crucial role in aiding organisations with incident response, conducting forensic analysis, and identifying vulnerabilities within their systems.

Machine learning (ML) and deep learning (DL) techniques have been recognised as highly effective methods for improving the detection capabilities of intrusion detection systems (IDS). Machine learning algorithms acquire knowledge from past data and generate models capable of identifying patterns and anomalies in network traffic or system behaviour. Deep learning (DL), a subset of machine learning (ML), utilises artificial neural networks that consist of multiple layers. These networks are capable of automatically acquiring hierarchical representations of data, thereby facilitating the extraction of more intricate and abstract features [6]. Machine learning (ML) and deep learning (DL) techniques possess the capability to address the constraints of conventional intrusion detection systems (IDS) through their ability to adapt to changing attack patterns and identify previously unknown attacks. These techniques effectively handle substantial amounts of data, autonomously extract pertinent features, and enhance the precision of detection. In addition, machine learning (ML) and deep learning (DL) based intrusion detection systems (IDS) have the capability to utilise ensemble methods, which involve the amalgamation of multiple models or algorithms, in order to attain improved detection rates and mitigate the occurrence of false positives [7].

Although there has been significant interest in ML and DL-based Intrusion Detection Systems (IDS) in recent years, traditional IDS continue to be relevant in certain situations. Signature-based intrusion detection systems (IDS), for example, remain efficacious in countering well-known attacks due to the regular updates of signature databases that integrate the most recent threat intelligence. The utilisation of anomaly-based intrusion detection systems (IDS) can serve as a valuable supplement to machine learning (ML) and deep learning (DL) techniques, as it offers an extra level of protection against unidentified attacks, despite its vulnerability to false positives. Traditional intrusion detection systems (IDS) also exhibit resource efficiency and

low rates of false positives, rendering them appropriate for environments or systems that have limited resources or specific security needs. Furthermore, it is worth noting that these systems frequently provide interpretability, which enables security analysts to gain a comprehensive understanding of the underlying factors behind alerts and make well-informed judgements.

Machine learning (ML) and deep learning (DL)-based intrusion detection systems (IDS) provide numerous benefits compared to conventional IDS methods. According to the cited source, they have the ability to adapt and acquire knowledge from novel attack patterns, resulting in the efficient detection of previously unidentified attacks [9]. These techniques effectively handle substantial amounts of data and autonomously extract pertinent features, thus diminishing the need for manual rule-based configurations. Furthermore, machine learning (ML) and deep learning (DL)-based intrusion detection systems (IDS) have the potential to enhance their detection capabilities and reduce the occurrence of false positives by utilising ensemble methods. Nevertheless, it is imperative to acknowledge that conventional intrusion detection systems (IDS) continue to maintain significance in certain circumstances. Signature-based Intrusion Detection Systems (IDS), such as the one mentioned, remain highly efficient in the detection of known attacks through the process of comparing network traffic with a comprehensive database of signatures. The utilisation of anomaly-based intrusion detection systems (IDS), despite the inherent risk of generating false positives, can offer a supplementary level of protection against unidentified attacks and serve as a valuable complement to machine learning (ML) and deep learning (DL) methodologies. In addition, conventional intrusion detection systems (IDS) frequently provide interpretability, allowing security analysts to understand the underlying causes of alerts and make well-informed judgements.

In the following sections of this review paper, the utilisation of machine learning (ML) and deep learning (DL) methodologies in the context of intrusion detection systems will be explored . this study will examine several machine learning (ML) algorithms that are frequently utilised in intrusion detection systems (IDS). These algorithms include decision trees, support vector machines (SVMs), and random forests. Additionally, we will also investigate deep learning (DL) architectures, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), commonly employed in IDS. Furthermore, we will examine the obstacles and ongoing research areas in this domain, in conjunction with practical examples and instances where machine learning (ML) and deep learning (DL) based intrusion detection systems (IDS) have exhibited enhanced efficacy. By gaining an understanding of the historical context and importance of Intrusion Detection Systems (IDS), as well as the impact of Machine Learning (ML) and Deep Learning (DL) techniques on improving intrusion detection capabilities, we can obtain a thorough examination of the current advancements in intrusion detection utilising ML and DL. Ultimately, this serves as a valuable contribution to the wider domain of cybersecurity [11].

2 Overview

Intrusion Detection Systems (IDS) refer to security mechanisms that have been specifically developed to oversee network traffic and system events with the purpose of detecting potential intrusions. Traditional intrusion detection systems (IDS) can be categorised into two primary classifications: signature-based and anomaly-based.

- Signature-based intrusion detection systems (IDS) operate by using pre-established patterns or signatures associated with known attacks. These signatures are formulated through a thorough examination of attack behaviors and characteristics. During operation, the IDS compares incoming network traffic or system events to the existing signature database. If a match is detected, indicating a potential unauthorized access, the system generates an alert. Signature-based IDS excel in detecting widely recognized attacks by continuously updating their signatures with the latest threat intelligence [12]. However, their capabilities are limited to identifying attacks only when there are existing signatures available for detection. These systems lack the ability to detect and classify previously unknown or zero-day attacks, making them vulnerable to emerging and unprecedented security risks.

- In contrast, anomaly-based intrusion detection systems (IDS) prioritize the detection of deviations from typical system behavior. A baseline of normal activity is established through the analysis of historical data or pattern extraction from training sets. During operation, the IDS compares observed behavior to the established baseline [13]. If a significant deviation is identified, the system generates an alert. Anomaly-based IDS have proven to be effective in detecting unknown or zero-day attacks due to their ability to discern and identify abnormal patterns. However, these systems may generate false positive results due to genuine fluctuations in system performance, such as system updates or changes in user behavior. Optimizing the anomaly detection threshold is crucial to achieving a delicate balance between detection accuracy and the occurrence of false positives.

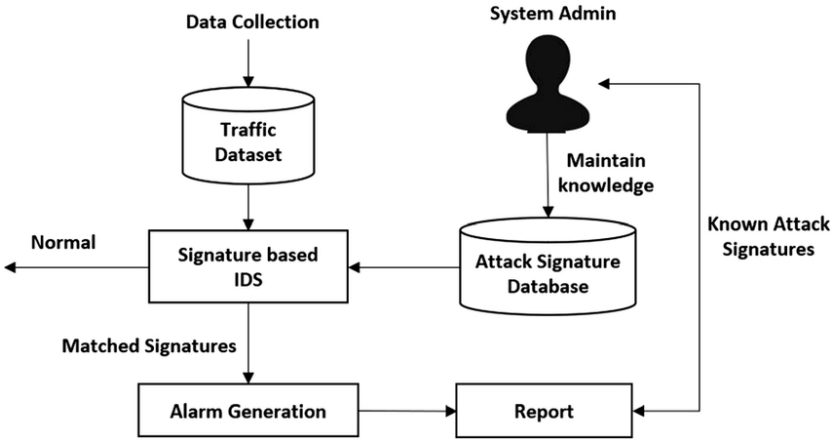

Figure 1 : Signature based Intrusion detection system

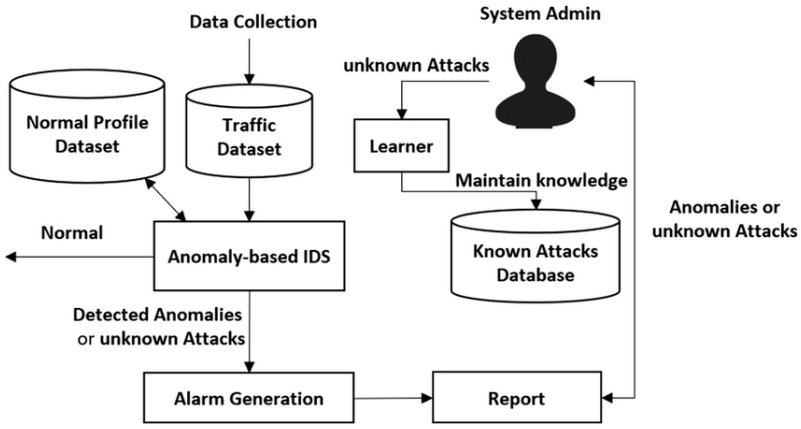

Figure 2 : Anomaly based Intrusion detection system

2.1 Limitations and Challenges of Traditional IDS

The utilisation of machine learning (ML) and deep learning (DL) techniques for intrusion detection has been motivated by the various limitations and challenges encountered by traditional intrusion detection systems (IDS).[14] Firstly, traditional intrusion detection systems (IDS) heavily depend on human experts for the creation and upkeep of signatures or the establishment of rules. The aforementioned procedure is characterised by its requirement of a significant amount of time and effort, as well as its potential inability to effectively adapt to the ever-changing nature of the threat environment. As a result, conventional intrusion detection systems (IDS) may experience delays in identifying novel or evolving attacks. Moreover, conventional intrusion detection systems face challenges in effectively managing the growing quantity and intricacy of network traffic and system events. The extensive magnitude of contemporary networks surpasses the capacity

of manual rule-based configurations, resulting in inadequate coverage or instances of false negatives. Similarly, the ever-changing nature of network traffic and the continuous development of attack techniques pose significant difficulties for conventional Intrusion Detection Systems (IDS) in effectively identifying and distinguishing between legitimate and malicious activities.

3 Machine Learning in Intrusion Detection Systems

Machine learning techniques have emerged as a promising approach to address the limitations of traditional intrusion detection systems (IDS). These techniques leverage their inherent capability to autonomously learn patterns and anomalies from extensive datasets. In the domain of intrusion detection, machine learning (ML) algorithms can be utilised to extract pertinent features from network traffic or system event data. These features are then used to construct models that can effectively classify and identify potential intrusions [15]. Decision trees are widely utilised machine learning algorithms within the field of intrusion detection systems (IDS). These algorithms are responsible for constructing a model that exhibits a tree-like structure, effectively mapping various features to corresponding decisions or classifications. Support Vector Machines (SVMs) are a classification algorithm that employs a hyperplane to accurately distinguish between various classes of data. This characteristic renders SVMs well-suited for the purpose of intrusion detection [15]. Random Forests, a method that integrates multiple decision trees, provide enhanced detection accuracy and robustness in the presence of noise and variability in the dataset [16].

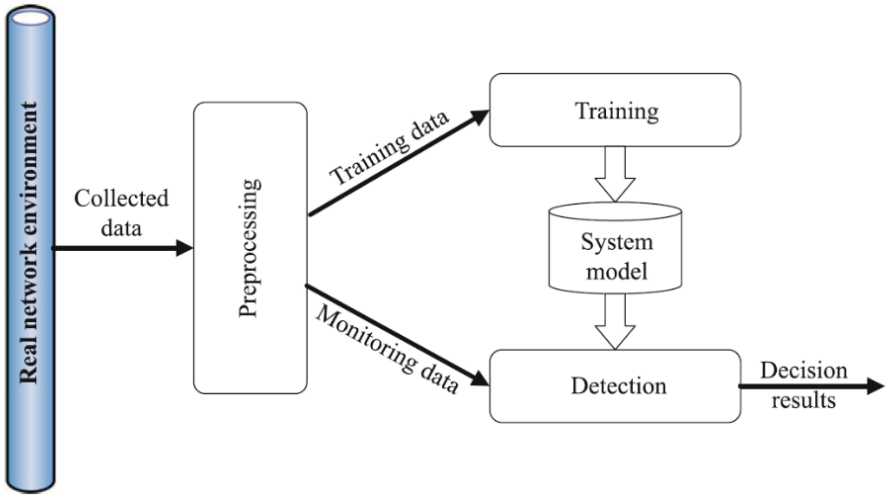

Figure 3: Machine learning based IDS

Machine learning-based intrusion detection systems (IDS) also employ techniques for feature selection and extraction in order to improve the effectiveness of detection. These techniques aim to identify the most informative features for intrusion detection, thereby reducing the dimensionality of the data and enhancing the efficiency and effectiveness of the algorithms. Feature selection techniques, such as Information Gain and Principal Component Analysis (PCA), have the ability to identify a subset of features that possess the greatest discriminatory capability [17]. Feature extraction techniques, such as Autoencoders and Principal Component Analysis (PCA), are employed to convert the original dataset into a reduced-dimensional representation, while preserving crucial information.

In recent times, there has been a growing interest in the application of deep learning (DL) techniques for intrusion detection. This is primarily attributed to the capability of DL models to autonomously acquire intricate data representations. Deep learning models, such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs),

have exhibited remarkable efficacy in tasks related to image and text processing. These models have also been successfully applied in the Intrusion Detection System (IDS) domain, as mentioned in reference [18]. Convolutional neural networks (CNNs) demonstrate exceptional proficiency in extracting spatial features from network traffic data. On the other hand, recurrent neural network (RNN) models are particularly adept at capturing temporal dependencies within sequences of system events. Deep learning-based intrusion detection systems (IDS) possess the capability to acquire hierarchical representations of data, thereby facilitating the extraction of more advanced and abstract features. This, in turn, leads to enhanced accuracy in the detection of intrusions.

One of the noteworthy advantages of machine learning (ML) and deep learning (DL)-based intrusion detection systems (IDS) is their ability to adapt to changing attack patterns. Traditional intrusion detection systems (IDS) frequently encounter difficulties in keeping up with the rapid evolution of novel and emerging threats. This necessitates the manual updating of signatures or rules in order to effectively detect and mitigate these threats. In contrast, machine learning (ML) and deep learning (DL) algorithms possess the ability to acquire knowledge from novel data and adapt their models accordingly. This capability empowers them to identify previously unobserved attacks or modifications of known attacks. The ability to adapt is of utmost importance in the dynamic field of cybersecurity, as malicious actors persistently devise novel strategies to circumvent conventional security measures [19].

Furthermore, machine learning (ML) and deep learning (DL)-based intrusion detection systems (IDS) possess the capability to effectively manage the substantial quantities of data produced by contemporary networks. The exponential growth of network traffic and system events has been attributed to the widespread adoption of connected devices and the increasing complexity of network infrastructures. Machine learning (ML) and deep learning (DL) algorithms possess the ability to effectively handle and analyse extensive datasets, extracting pertinent characteristics, and promptly identifying intrusions either in real-time or nearly real-time [20]. Nevertheless, the implementation of intrusion detection systems (IDS) that rely on machine learning (ML) and deep learning (DL) techniques poses a distinct set of obstacles. One primary obstacle that researchers face is the requirement for annotated training data. Machine learning (ML) and deep learning (DL) algorithms necessitate the use of annotated data in order to acquire knowledge and construct precise models. The acquisition of labelled data for intrusion detection poses challenges due to the limited availability of real-world attack data and the necessity to uphold privacy and legal requirements when utilising sensitive network traffic or system event data. The creation of diverse and representative datasets that encompass a wide range of attack scenarios is of utmost importance in the training of intrusion detection system (IDS) models that are both robust and reliable.

The interpretability and explainability of machine learning (ML) and deep learning (DL)-based intrusion detection systems (IDS) present an additional obstacle. Traditional intrusion detection systems (IDS) typically offer comprehensive logs and explanations of identified incidents, thereby facilitating the comprehension of alerts by security analysts [21]. ML and DL models, specifically deep neural networks, are frequently regarded as opaque systems, posing challenges in comprehending their decision-making mechanisms. Current research is dedicated to the advancement of interpretable machine learning (ML) and deep learning (DL) models and techniques. The objective is to tackle the challenge of enabling security analysts to have confidence in and understand the decisions made by these sophisticated intrusion detection systems (IDS)..

3.1 Dataset and pre processing

The progress of Intrusion Detection System (IDS) research is significantly dependent on the accessibility of datasets that encompass network traffic and encompass labelled instances of both normal and intrusive activities. The aforementioned datasets

are regarded as invaluable assets for the development and assessment of Intrusion Detection System (IDS) models. NSL-KDD and UNSW-NB15 are prominent datasets frequently employed in Intrusion Detection System (IDS) research.

- The NSL-KDD dataset, which originated from the KDD Cup 1999 dataset [22], has garnered significant recognition and acceptance within the field of Intrusion Detection Systems (IDS). The dataset consists of network traffic data obtained from a simulated environment that encompasses a range of attack types. The dataset comprises a total of 41 distinct features, encompassing crucial connection attributes such as protocol type, service, source and destination IP addresses, as well as source and destination ports. The dataset offers a categorised collection of instances, comprising normal instances and instances belonging to four distinct attack types: Denial of Service (DoS), User to Root (U2R), Remote to Local (R2L), and Probe. The NSL-KDD dataset has become widely recognised as a standard for assessing intrusion detection system (IDS) algorithms, primarily because of its comprehensive and varied characteristics.
- The UNSW-NB15 dataset [23] is widely utilised in Intrusion Detection System (IDS) research. It is derived from a real-world setting and encompasses a diverse range of network traffic scenarios. The dataset comprises a total of nine distinct attack types, which encompass DoS (Denial of Service), U2R (User to Root), R2L (Remote to Local), and Reconnaissance, among other categories. The dataset consists of a comprehensive set of 49 features, encompassing attributes related to protocols, flow statistics, and content-based characteristics. The UNSW-NB15 dataset presents notable advancements compared to the KDD Cup 1999 dataset, as it effectively tackles various limitations and integrates a broader range of authentic attack scenarios.

The process of data preprocessing is of utmost importance in the field of Intrusion Detection Systems (IDS) research, as it serves the purpose of preparing the dataset in a manner that facilitates efficient analysis and model training. Various preprocessing techniques are frequently employed on Intrusion Detection System (IDS) datasets, such as normalisation and feature scaling.

Normalisation is a commonly used method for rescaling numerical attributes to a standardised interval, usually ranging from 0 to 1. The purpose of this approach is to ensure equitable contribution of all features to the learning process, thereby mitigating the potential dominance of any single feature resulting from variations in their scales. Normalisation can be accomplished through various techniques, including Min-Max scaling or Z-score normalisation (24). The process of Min-Max scaling involves linearly transforming the feature values in order to fit within a predetermined range. On the other hand, Z-score normalisation is a method that transforms the values of the features to have a mean of 0 and a standard deviation of 1.

An additional significant preprocessing technique is feature scaling, which endeavours to standardise the scales of features. This methodology demonstrates notable advantages when applied to algorithms that exhibit sensitivity to the magnitude of the input features, specifically those that rely on distance calculations. Two commonly used techniques for feature scaling are Standardisation and Robust Scaling (Author, 25). The process of standardisation involves adjusting the feature values in a dataset such that their mean becomes 0 and their standard deviation becomes 1. On the other hand, Robust Scaling is a data preprocessing technique that rescales the features using statistical measures that are resistant to the influence of outliers. This characteristic renders it appropriate for datasets that potentially include outliers.

Preprocessing techniques are employed to ensure that the features undergo suitable transformation and standardisation, thereby enhancing the learning process of Intrusion Detection System (IDS) models. By applying normalisation and scaling techniques to the data, the models are able to accurately identify and analyse patterns and anomalies that exist within the dataset. As a result, this enhances the performance of intrusion detection systems..

3.2 Related work in Machine learning based IDS

The utilisation of machine learning (ML) algorithms has demonstrated significant potential in enhancing the efficacy of intrusion detection systems (IDS). This section presents a comprehensive overview of frequently employed machine learning (ML) algorithms in Intrusion Detection Systems (IDS), encompassing decision trees, support vector machines (SVMs), and random forests.

Decision trees are widely used machine learning algorithms in intrusion detection systems (IDS) because of their straightforwardness and ability to be easily understood and interpreted. The algorithms in question are designed to create a model that takes the form of a tree. In this tree structure, internal nodes correspond to features or attributes, while leaf nodes correspond to decisions or classifications. The process of constructing decision trees involves partitioning the data by evaluating the values of various features, with the aim of generating branches that possess the highest discriminatory power for classification purposes. One notable advantage of decision trees lies in their capacity to effectively capture intricate decision boundaries and discern significant features. Nevertheless, it is important to note that decision trees may exhibit a susceptibility to overfitting in instances where the depth of the tree exceeds a certain threshold. This phenomenon can result in a decline in the overall ability of the tree to generalise and perform well on unseen data. Methods such as pruning and ensemble techniques can help alleviate this constraint.

Support vector machines (SVMs) are extensively employed in intrusion detection systems (IDS) owing to their efficacy in performing binary classification tasks [27]. Support Vector Machines (SVMs) employ a hyperplane to effectively partition the data into distinct classes, leveraging the most discerning features for discrimination purposes. The construction of the hyperplane aims to optimise the margin between the classes, leading to enhanced generalisation performance. Support Vector Machines (SVMs) possess the capability to effectively handle datasets with a high number of dimensions, while also exhibiting robustness in the presence of noise and outliers. These tools are especially valuable in Intrusion Detection Systems (IDS), as their primary objective is to effectively categorise network traffic or system events as either benign or malicious. Nevertheless, Support Vector Machines (SVMs) can pose computational challenges when confronted with extensive datasets, necessitating meticulous hyperparameter selection.

Random forests are a type of ensemble learning technique that integrates multiple decision trees in order to enhance the accuracy and resilience of classification tasks. In a random forest, the construction of each decision tree involves the utilisation of a subset of the training data and a random subset of features. The ultimate categorization is established by consolidating the forecasts made by each individual tree. Random forests are widely recognised for their capacity to effectively manage datasets with a high number of dimensions, effectively handle missing values, and offer reliable estimations of feature importance. Ensemble methods, such as random forests, exhibit reduced susceptibility to overfitting in comparison to individual decision trees, while simultaneously providing enhanced accuracy and robustness.

Furthermore, machine learning-based intrusion detection systems (IDS) make use of feature selection and extraction techniques to enhance the effectiveness of their detection capabilities. Feature selection techniques, such as Information Gain and Principal Component Analysis (PCA) [17], are employed to determine a subset of features that possess the highest discriminatory capability. Feature extraction techniques, such as Autoencoders and Principal Component Analysis (PCA), are employed to convert the initial data into a reduced-dimensional representation, while ensuring the retention of significant information.

Moreover, the utilisation of deep learning (DL) methodologies in intrusion detection systems (IDS) has garnered significant interest owing to their inherent capacity to autonomously acquire intricate data representations. Deep learning models, such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) (18), have exhibited exceptional efficacy in tasks related to image and text processing, and have been successfully ap-

plied in the Intrusion Detection System (IDS) domain. Convolutional neural networks (CNNs) demonstrate exceptional proficiency in extracting spatial features from network traffic data. On the other hand, recurrent neural network (RNN) models are well-suited for capturing temporal dependencies present in sequences of system events. Deep learning-based intrusion detection systems (IDS) possess the capability to autonomously acquire hierarchical representations of data. This ability facilitates the extraction of more sophisticated and abstract features, ultimately resulting in enhanced accuracy for intrusion detection. The advanced machine learning (ML) and deep learning (DL) algorithms exhibit the ability to adapt to changing attack patterns, enabling them to identify previously unencountered attacks or modified versions of known attacks.

ML and DL techniques have been utilised by researchers to analyse diverse datasets with the aim of enhancing the efficacy of Intrusion Detection Systems (IDS). As an illustration, Ferrag et al. (2020) devised an Intrusion Detection System (IDS) named RDTIDS by utilising the CICIDS2017 and BoT-IoT datasets. Their approach demonstrated enhanced accuracy and detection rates in comparison to pre-existing methodologies. The study conducted by Wazirali et al. (2020) aimed to improve the accuracy of intrusion detection systems by addressing the issues of false alarms and detection rates. To achieve this, the researchers utilised the NSL-KDD dataset and implemented the K-nearest neighbour (KNN) algorithm. The study conducted by Krishnaveni et al. (2020) employed the NSL-KDD dataset and the Support Vector Machine (SVM) algorithm in order to construct a robust anomaly detection system. The authors of the study, Chen et al. (2020), devised a highly effective hybrid clustering algorithm for the purpose of intrusion detection. This algorithm was implemented and evaluated using the UCI ML and KDD Cup 99 datasets. In their study, Choras et al. (2020) utilised Artificial Neural Networks (ANN) as a means to identify malware and various types of attacks. This was accomplished by employing the NSL-KDD and CICIDS2017 datasets. The researchers Alqahtani et al. (2020) made a significant contribution to the field of Intrusion Detection Systems (IDS) by designing and implementing a system that utilised a range of machine learning classification algorithms. They conducted an evaluation of the system's performance by employing cybersecurity datasets.

In summary, machine learning (ML) and deep learning (DL) algorithms, including decision trees, support vector machines (SVMs), random forests, convolutional neural networks (CNNs), and recurrent neural networks (RNNs), have demonstrated effective utilisation in the field of intrusion detection system (IDS) research. The aforementioned algorithms exhibit the capacity to enhance the precision and efficacy of intrusion detection systems. Through the utilisation of these sophisticated methodologies, scholars strive to augment the security of computer systems and networks by precisely discerning and alleviating potential intrusions.

3.3 Evaluation Metrics for ML-based IDS

In order to evaluate the performance of machine learning-based intrusion detection systems (IDS), a range of evaluation metrics are employed to quantify the efficacy of the intrusion detection models. These metrics offer valuable insights into the precision, effectiveness, and dependability of the models, enabling comparisons and facilitating informed decision-making. Several commonly employed evaluation metrics in machine learning-based intrusion detection systems (IDS) encompass:

- The metric of accuracy is a direct measure that indicates the proportion of correctly classified instances in relation to the total number of instances [29]. The metric offers a comprehensive assessment of the accuracy of the model and is particularly appropriate in situations where the distribution of classes is equitable. Nevertheless, in situations characterised by imbalanced classes, relying solely on accuracy may prove inadequate due to its potential bias towards the majority class.

- Precision is a metric that quantifies the ratio of correctly identified positive instances to all instances that were predicted as positive. This approach prioritises the accuracy of positive predictions and is particularly advantageous in scenarios where the consequences of false positives are significant. A high level of precision is

indicative of a low occurrence of false positives, thereby minimising the occurrence of unnecessary alarms or alerts.

- Recall, also known as sensitivity or true positive rate, quantifies the ability to correctly identify true positives relative to the total number of positive instances. The aforementioned statement highlights the capacity of the model to identify and respond to unauthorised access or malicious activities, which holds significant importance in scenarios where the consequences of failing to detect such instances are substantial. A high recall value signifies a reduced occurrence of false negatives, thereby guaranteeing an elevated rate of detection.

- The F1 score is a metric that integrates precision and recall, offering a harmonious equilibrium between the two. The harmonic mean of precision and recall is a valuable metric in cases where both false positives and false negatives hold equal significance. The F1 score is a metric that varies between 0 and 1, where a higher value signifies superior performance.

- The false positive rate is a metric that quantifies the ratio of false positive predictions to the total number of instances that are classified as negative. Minimising false positives is of utmost importance in Intrusion Detection Systems (IDS), as it serves to alleviate the workload of security analysts and prevent unwarranted scrutiny of benign activities.

- The Receiver Operating Characteristic (ROC) curve is a graphical representation used in statistical analysis to evaluate the performance of a binary classification model. The receiver operating characteristic (ROC) curve illustrates the relationship between the true positive rate and the false positive rate across different thresholds used for classification. The aforementioned analysis offers a comprehensive perspective on the trade-off between the rate of detection and the rate of false positives, enabling the evaluation and comparison of various models in terms of their performance. The utilisation of the receiver operating characteristic (ROC) curve encompasses the calculation of the area under the curve (AUC), which serves as a widely employed metric for summarising performance. A greater AUC value is indicative of superior performance.

4 Deep Learning in Intrusion Detection Systems

Deep learning (DL) is a specialised domain within the broader field of machine learning (ML), which is concerned with the development of algorithms and models that enable artificial neural networks to acquire knowledge and generate predictions based on intricate and multifaceted datasets. Neural networks are computational models that draw inspiration from the anatomical and functional characteristics of the human brain. These models are composed of interconnected layers of artificial neurons. Deep learning (DL) has had a substantial influence in multiple fields, encompassing computer vision, natural language processing, and, more recently, intrusion detection systems (IDS).

In the domain of deep learning-based intrusion detection systems (IDS), various architectural designs have exhibited considerable potential in the identification and categorization of unauthorised activities. Convolutional neural networks (CNNs) are frequently employed in the analysis and extraction of features from network traffic or system logs. Convolutional neural networks (CNNs) employ convolutional layers to autonomously acquire spatial patterns and hierarchical representations from unprocessed input data. This allows them to efficiently comprehend intricate connections and differentiate between regular and malicious network activities.

Recurrent neural networks (RNNs) are widely utilised as an architecture for intrusion detection systems (IDS). Recurrent Neural Networks (RNNs) are specifically engineered to handle sequential data by integrating feedback connections, which enable the retention of information over multiple time steps. This characteristic renders them well-suited for the examination of temporal interdependencies and identification of intrusions that manifest patterns or behaviours that evolve over time, such as coordinated attacks or reconnaissance operations.

Transfer learning is a methodology employed in deep learning-based Intrusion Detection Systems (IDS) whereby the knowledge acquired from training a neural network on a specific task is transferred and utilised for another task that is closely related. The efficacy of this methodology has been demonstrated in the field of Intrusion Detection Systems (IDS) as a result of the constrained accessibility of annotated intrusion data. The performance of the model can be greatly improved by initially training a deep neural network on a comprehensive dataset, such as publicly accessible non-intrusive network traffic, and subsequently refining the network using a smaller labelled dataset that is specific to the Intrusion Detection System (IDS) domain. The utilisation of pretraining in deep learning-based intrusion detection systems (IDS) enables the network to take advantage of the acquired knowledge about general data patterns. This allows the network to efficiently extract pertinent features from the limited labelled intrusion data. This methodology aids in mitigating the issue of limited data availability and enables the creation of intrusion detection models that are both more precise and resilient.

Alqahtani et al. (2020) conducted a distinct investigation with the objective of constructing a highly effective Intrusion Detection System (IDS) through the utilisation of ensemble feature selection and classification methodologies. The methodology employed by the researchers demonstrated significant improvements in performance, a high level of accuracy, and a low rate of false alarms (FAR). Nevertheless, the authors recognised that the procedure of selecting features and training ensemble models could require significant computational resources and extended processing durations, presenting difficulties in environments with limited resources or when working with extensive datasets.

In their study, Riyaz et al. (2020) introduced an innovative intrusion detection system that was designed specifically to address the unique characteristics and challenges of wireless networks. The researchers were able to achieve an overall detection accuracy of 98.88% by employing the Conditional Random Field and Linear Correlation algorithms. The authors placed significant emphasis on the utilisation of feature selection algorithms and Convolutional Neural Network (CNN) in order to enhance the effectiveness of intrusion detection. However, the provided information did not include any additional details regarding potential limitations.

The research conducted by Sun et al. (2020) focused on the creation of an intrusion detection system (IDS) utilising deep learning techniques. The CICIDS2017 dataset was utilised in this study. The researchers utilised a hybrid network architecture that integrated Convolutional Neural Network (CNN) and Long Short-Term Memory Network (LSTM) components. The system demonstrated a comprehensive accuracy rate of 98.67%. Nevertheless, the researchers issued a warning regarding the potential occurrence of overfitting problems associated with the implementation of this hybrid network.

In the context of InSDN, Elsayed et al. (2021) introduced a novel hybrid deep learning methodology to enhance the performance of Network Intrusion Detection Systems (NIDSs). The researchers employed Convolutional Neural Network (CNN) methodologies, which led to improved performance across all assessment criteria. However, the authors also recognised the potential concerns related to overfitting and model performance. Nevertheless, the provided information did not offer any additional elaboration on these limitations.

Table 1 : Comparison of previous published work

Ref	Objectives	Dataset	Algorithms	Results	Contributions	Limitations
Ferrag et al , 2020	Intrusion Detection System (RDTIDS)	CICIDS2017 dataset BoT-IoT dataset	REP Tree JRip algorithm Forest PA	Highest True Negative Rate (TNR): 98.855% Highest Detection Rate (DR):.	Superiority in terms of accuracy, detection rate, false alarm rate, and time overhead compared to existing schemes	**Average performance for some attack types**
Wazirali et al , 2020	Decrease false alarm rate and enhance detection rate	NSL-KDD	K-nearest neighbor (KNN) with hyperparameter tuning and fivefold cross-validation	Proposed algorithm: 98.87% accuracy	Effective semi-supervised technique for intrusion detection systems (IDSs)	Difficulties in implementing the proposed approach when altering the size of records in the dataset. Unable to identify new kinds of attacks in real-time. Inability to identify new kinds of attacks in real-time. Difficulties in implementing the proposed approach when altering dataset size.
Krishnaveni et al , 2020	Develop an effective anomaly detection system	NSL-KDD	Support Vector Machine (SVM)	IDS accuracy: 96.24%	Effective anomaly detection system for cloud computing	sensitivity to the selection of hyperparameters. SVM requires proper tuning of parameters such as the kernel type, regularization parameter, and gamma value.
Chen et al , 2020	Develop an efficient hybrid clustering algorithm	UCI ML KDD Cup 99	QALO-K (Quantum-Inspired Ant Lion Optimized)	K-means: - AR: 93.44%	Hybrid algorithm combining quantum computing and swarm intelli-	Similar to k-means, the performance of QALO-K may be influenced by the initial values of cluster centers or other

				- DR: 92.83% - FPR: 4.05% - F1: 95.79% ALO-K: - AR: 97.52% - DR: 98.58% - FPR: 6.80% - F1: 98.46% QALO-K: - AR: 98.63% - DR: 98.41% - FPR: 0.44% - F1: 99.14%	gencealgo-rithms with k-means for in-trusion detec-tion Increased ac-curacy and minimized false alarm rate	algorithm-specific parameters.
Choras et al , 2020	Intrusion Detection, the ability to detect malware and other attacks	NSL-KDD, CICIDS2017	Artificial Neural Net-works (ANN)	Multi-class classifica-tion accu-racy: 99.909%	Evaluation of the influence of hyperpa-rameters on classification results	computational complexity and time required for training and tuning the network

Alqahtani et al , 2020	Develop an intrusion detection system (IDS)	Cyber-security datasets	Bayesian Network, Naive Bayes, Decision Tree, Random Decision Forest, Random Tree, Decision Table,Artificial Neural Network	Precision, Recall, F1-score, Accuracy	Utilization of machine learning classification algorithms for intrusion detection	One limitation related to the utilization of machine learning classification algorithms for intrusion detection is the potential for false positives and false negatives.
Alqahtani et al , 2020	Develop an efficient Intrusion Detection System (IDS)		Ensemble feature selection and classification techniques	Strong performance enhancement, high accuracy, low false alarm rate (FAR)	Utilization of ensemble feature selection and classification techniques for IDS in the cloud environment, Efficient classification of network traffic as normal or attack	The process of feature selection and ensemble model training may require more computational resources and longer processing times. This limitation can be especially significant in resource-constrained environments or when dealing with large-scale datasets.
Riyaz et al , 2020	Develop a new intrusion detection system	Wireless networks	Conditional Random Field and Linear Correlation	Overall detection accuracy: 98.88%	Utilization of feature selection algorithms and Convolutional Neural Network (CNN) for effective intrusion detection	None mentioned

Sun et al , 2020	Develop a DL-IDS (deep learning-based intrusion detection system)	CICIDS2017	Convolutional Neural Network (CNN) Long Short-Term Memory Network (LSTM)	Overall accuracy: 98.67%	Utilization of hybrid network (CNN and LSTM) for improved intrusion detection	One limitation related to the utilization of the hybrid network of Convolutional Neural Network (CNN) and Long Short-Term Memory Network (LSTM) for intrusion detection is the potential for overfitting.
Elsayed et al, 2021	Develop a new hybrid DL approach for NIDSs	InSDN	Convolutional Neural Network (CNN)	Higher performance in all evaluation metrics	Utilization of a hybrid DL approach based on CNN for classifying flow traffic into normal or attack classes	Potential overfitting and model performance issues

6. Challenges and Open Research Directions

Although machine learning (ML) and deep learning (DL) techniques have demonstrated potential in the field of intrusion detection systems (IDS), there remain a number of challenges and limitations that necessitate attention and resolution. The aforementioned challenges and open research directions present prospects for future advancements in the field.

- One of the primary obstacles associated with deep learning models pertains to their inherent deficiency in terms of interpretability and explainability. These models are frequently regarded as opaque systems, thereby posing challenges in comprehending and interpreting their decision-making processes. The development of methodologies and techniques that offer elucidation for the decision-making processes of machine learning and deep learning-based intrusion detection systems is of utmost importance. This capability would facilitate the comprehension of intrusion alerts by security analysts and foster confidence in the system.

- The presence of adversarial attacks presents a substantial obstacle to intrusion detection systems (IDS) that rely on machine learning and deep learning techniques. These attacks entail the manipulation of input data in order to deceive the models and circumvent detection. The continuous investigation of creating resilient intrusion detection systems (IDS) models capable of identifying and mitigating adversarial attacks is a persistent area of academic research. Methods such as adversarial training, robust feature representation, and anomaly detection mechanisms have the potential to effectively tackle this challenge and enhance the resilience of Intrusion Detection Systems (IDS) against such attacks.

- Scalability is a crucial requirement for IDS models as they must effectively manage the substantial volume and rapid flow of network traffic in real-time. The optimisation of scalability and efficiency in machine learning and deep learning-based intrusion detection systems (IDS) is of paramount importance in effectively managing network environments of significant scale. The investigation of methodologies aimed at enhancing the efficiency of computational resources and scalability of Intrusion Detection System (IDS) models, while maintaining their accuracy, necessitates thorough re-

search. This encompasses the creation of algorithms and frameworks that are capable of effectively handling and examining vast quantities of data within a reasonable timeframe.

- Generalisation: Intrusion Detection System (IDS) models should possess the capability to effectively generalise their learned knowledge to accurately detect and mitigate novel and previously unseen attack patterns. The performance of models trained on limited and specific datasets may be compromised when confronted with novel attacks. Further investigation is required in order to formulate models that possess the capability to effectively extrapolate to various attack scenarios and adjust to dynamic network environments. The process entails the investigation of various methodologies, including transfer learning, domain adaptation, and continual learning, with the aim of improving the overall generalisation abilities of intrusion detection system (IDS) models.

By acknowledging and tackling these obstacles and investigating the corresponding areas of research, it is possible to achieve progress in enhancing the interpretability, resilience, scalability, and generalizability of intrusion detection systems (IDS) based on machine learning and deep learning (ML/DL). This research endeavour will make a valuable contribution towards the advancement of intrusion detection systems, enhancing their efficacy and dependability. These systems will be capable of accurately identifying and promptly responding to a diverse array of security threats across different network environments.

6.1 Emerging Trends and Future Directions:

The field of intrusion detection system (IDS) research is continuously evolving, and several emerging trends and future directions are shaping the advancement of machine learning (ML) and deep learning (DL)-based IDS. These trends and directions include:

- Deep Learning Architectures: There is a continued exploration of deep learning architectures, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), in the field of IDS. Researchers are investigating how these architectures can be further improved and tailored for IDS applications. Future research may focus on developing novel architectures specifically designed to handle the unique characteristics of network traffic data. This could involve incorporating domain-specific knowledge and designing architectures that can effectively capture and represent complex relationships in network data.

- Transfer Learning and Pretraining: Transfer learning techniques have shown promise in addressing the issue of data scarcity in IDS. By leveraging models pretrained on large-scale datasets, IDS models can benefit from the learned representations of general data patterns. Future research may explore advanced transfer learning methods and investigate the optimal strategies for transferring knowledge from pretrained models to IDS-specific tasks. This includes studying how to effectively fine-tune pretrained models and adapt them to the intrusion detection domain, ultimately improving the generalization capability of IDS models.

- Unsupervised Learning for Anomaly Detection: Unsupervised learning techniques offer the potential to detect unknown or novel attacks without relying on labeled training data. Research in this area may focus on developing robust unsupervised learning algorithms, such as generative models and clustering techniques, for anomaly detection in IDS. These approaches can analyze network traffic patterns and identify abnormal behaviors or deviations from normal patterns. By leveraging unsupervised learning, IDS models can provide early warnings of potential attacks and detect previously unseen intrusion patterns.

These emerging trends and future directions in ML/DL-based IDS research hold great potential for advancing the field. By exploring and harnessing the capabilities of deep learning architectures, transfer learning techniques, and unsupervised learning approaches, researchers can enhance the accuracy, efficiency, and robustness of IDS models. These advancements can contribute to the development of more effective and reliable intrusion detection systems, capable of accurately detecting and mitigating a wide range of security threats in various network environments...

7 Conclusion

In summary, this review paper has presented a thorough examination of intrusion detection systems (IDS) employing machine learning (ML) and deep learning

(DL) methodologies. The review highlights the importance of ML/DL-based Intrusion Detection Systems (IDS) in effectively addressing the challenges presented by the ever-changing landscape of cyber threats. Through the utilisation of sophisticated algorithms and models, intrusion detection systems (IDS) based on machine learning and deep learning (ML/DL) techniques have the potential to greatly enhance the precision, effectiveness, and resilience of identifying unauthorised access attempts. Consequently, this can lead to an overall improvement in the security stance of computer networks.

The limitations of conventional intrusion detection system (IDS) approaches have been brought to attention in the review, emphasising the potential of machine learning and deep learning (ML/DL) techniques to address these limitations. The utilisation of deep learning architectures, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), facilitates the identification of intricate patterns and interconnections within network traffic data. Transfer learning and pretraining methodologies are employed to mitigate the challenge of limited data availability and improve the ability of Intrusion Detection System (IDS) models to generalise.

In addition, the review has identified a number of areas in the field that lack sufficient research and present challenges, necessitating further investigation. The aforementioned factors encompass the interpretability and explainability of intrusion detection systems (IDS) based on machine learning and deep learning techniques, resilience against adversarial attacks, scalability, and generalisation capabilities. Future research endeavours should prioritise the development of methodologies aimed at elucidating the decision-making processes employed by deep learning models. Additionally, efforts should be directed towards bolstering the resilience of Intrusion Detection Systems (IDS) against adversarial attacks, optimising computational resources to ensure scalability, and enhancing the generalisation capability of IDS models in order to effectively detect novel attack patterns.The importance of ML/DL-based Intrusion Detection Systems (IDS) cannot be overstated within the realm of advancing cyber threats. As attacks become more intricate and advanced, conventional rule-based systems may encounter difficulties in maintaining pace. Machine learning and deep learning techniques have the capability to detect and address new or previously unknown attacks in a proactive manner. This is achieved by utilising unsupervised learning methods, anomaly detection, and sophisticated deep learning architectures.

Based on the insights derived from this review, it is advisable for researchers and practitioners to give priority to the resolution of the research gaps that have been identified, as well as to the advancement of intrusion detection systems (IDS) based on machine learning (ML) and deep learning (DL) techniques. The collaboration among academia, industry, and cybersecurity communities plays a vital role in gathering diverse and current datasets, creating IDS models that are flexible and responsive, improving the interpretability of deep learning models, and investigating the possibilities of unsupervised learning for the detection of new or zero-day attacks. In summary, the utilisation of Machine Learning and Deep Learning techniques in Intrusion Detection Systems (IDS) shows great potential in bolstering the security of computer networks. Researchers have the opportunity to enhance the development of intrusion detection systems by focusing on the identified research gaps and continuously investigating emerging trends and techniques. This will ultimately strengthen our defences against evolving cyber threats, leading to more effective and resilient systems.

Compliance with Ethical Standards

All the authors have NO Conflict of Interest.

The review paper does not involve any Human Participant or animal

All the authors are well informed about the manuscript.

Competing Interests

All the authors have common interest in Network & System security.

Research Data Policy and Data Availability Statements

The submitted manuscript is novel, original, unpublished and not under simultaneous consideration by any other

journal.

There was meagre literature available regarding Intrusion Detection system so this review paper would serve as a knowledge base for Information Technology.

References

[1] Halbouni, A., Gunawan, T. S., Habaebi, M. H., Halbouni, M., Kartiwi, M., & Ahmad, R. (2022). Machine learning and deep learning approaches for cybersecuriy: A review. IEEE Access.

[2] Gümüşbaş, D., Yıldırım, T., Genovese, A., & Scotti, F. (2020). A comprehensive survey of databases and deep learning methods for cybersecurity and intrusion detection systems. IEEE Systems Journal, 15(2), 1717-1731.

[3] Zhou, Y., Han, M., Liu, L., He, J. S., & Wang, Y. (2018, April). Deep learning approach for cyberattack detection. In IEEE INFOCOM 2018-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS) (pp. 262-267). IEEE.

[4] Ahmad, R., Alsmadi, I., Alhamdani, W., & Tawalbeh, L. A. (2023). Zero-day attack detection: a systematic literature review. Artificial Intelligence Review, 1-79.

[5] Cañola Garcia, J. F., & Taborda Blandon, G. E. (2022, November). Implementing a Deep Learning Algorithm for Detection of Denial of Service Attacks. In Advances in Computing: 15th Colombian Congress, CCC 2021, Bogotá, Colombia, November 22–26, 2021, Revised Selected Papers (pp. 46-64). Cham: Springer International Publishing.

[6] Kocher, G., & Kumar, G. (2021). Machine learning and deep learning methods for intrusion detection systems: recent developments and challenges. Soft Computing, 25(15), 9731-9763.

[7] Liu, H., & Lang, B. (2019). Machine learning and deep learning methods for intrusion detection s

[8] 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), 4280-4290.

[9] Antunes, M., Oliveira, L., Seguro, A., Veríssimo, J., Salgado, R., & Murteira, T. (2022, March). Benchmarking Deep Learning Methods for Behaviour-Based Network Intrusion Detection. In Informatics (Vol. 9, No. 1, p. 29). MDPI.

[10] Thilakarathne, N. N., Kagita, M. K., Lanka, D. S., & Ahmad, H. (2020). Smart grid: a survey of architectural elements, machine learning and deep learning applications and future directions. arXiv preprint arXiv:2010.08094.

[11] Ahmad, Z., Shahid Khan, A., Wai Shiang, C., Abdullah, J., & Ahmad, F. (2021). Network intrusion detection system: A systematic study of machine learning and deep learning approaches. Transactions on Emerging Telecommunications Technologies, 32(1), e4150.

[12] Lansky, J., Ali, S., Mohammadi, M., Majeed, M. K., Karim, S. H. T., Rashidi, S., ... & Rahmani, A. M. (2021). Deep learning-based intrusion detection systems: a systematic review. IEEE Access, 9, 101574-101599.

[13] Alsoufi, M. A., Razak, S., Siraj, M. M., Nafea, I., Ghaleb, F. A., Saeed, F., & Nasser, M. (2021). Anomaly-based intrusion detection systems in iot using deep learning: A systematic literature review. Applied sciences, 11(18), 8383.

[14] Asharf, J., Moustafa, N., Khurshid, H., Debie, E., Haider, W., & Wahab, A. (2020). A review of intrusion detection systems using machine and deep learning in internet of things: Challenges, solutions and future directions. Electronics, 9(7), p.1177.

[15] Taher, K. A., Jisan, B. M. Y., & Rahman, M. M. (2019, January). Network intrusion detection using supervised machine learning technique with feature selection. In 2019 International conference on robotics, electrical and signal processing techniques (ICREST) (pp. 643-646). IEEE.

[16] Li, X., Chen, W., Zhang, Q., & Wu, L. (2020). Building auto-encoder intrusion detection system based on random forest feature selection. Computers & Security, 95, 101851.

[17] Rawat, S., Srinivasan, A., Ravi, V., & Ghosh, U. (2022). Intrusion detection systems using classical machine learning techniques vs integrated unsupervised feature learning and deep neural network. Internet Technology Letters, 5(1), e232.

[18] Laqtib, S., Yassini, K. E., & Hasnaoui, M. L. (2019, October). A deep learning methods for intrusion detection systems based machine learning in manet. In Proceedings of the 4th International Conference on Smart City Applications (pp. 1-8).

[19] Ferrag, M. A., Friha, O., Maglaras, L., Janicke, H., & Shu, L. (2021). Federated deep learning for cyber security in the internet of things: Concepts, applications, and experimental analysis. IEEE Access, 9, 138509-138542.

[20] Ahmad, Z., Shahid Khan, A., Wai Shiang, C., Abdullah, J., & Ahmad, F. (2021). Network intrusion detection system: A systematic study of machine learning and deep learning approaches. Transactions on Emerging Telecommunications Technologies, 32(1), e4150.

[21] Kim, A., Park, M., & Lee, D. H. (2020). AI-IDS: Application of deep learning to real-time Web intrusion detection. IEEE Access, 8, 70245-70261.

[22] Rawat, S., Srinivasan, A., Ravi, V., & Ghosh, U. (2022). Intrusion detection systems using classical machine learning techniques vs integrated unsupervised feature learning and deep neural network. Internet Technology Letters, 5(1), e232.

[23] Aleesa, A., Younis, M. O. H. A. M. M. E. D., Mohammed, A. A., & Sahar, N. (2021). Deep-intrusion detection system with enhanced UNSW-NB15 dataset based on deep learning techniques. Journal of Engineering Science and Technology, 16(1), 711-727.

[24] Güney, H. (2023). Preprocessing Impact Analysis for Machine Learning-Based Network Intrusion Detection. Sakarya University Journal of Computer and Information Sciences, 6(1), 67-79.

[25] Gupta, K., Sharma, D. K., Gupta, K. D., & Kumar, A. (2022). A tree classifier based network intrusion detection model for Internet of Medical Things. Computers and Electrical Engineering, 102, 108158.

[26] Alzahrani, A. O., & Alenazi, M. J. (2021). Designing a network intrusion detection system based on machine learning for software defined networks. Future Internet, 13(5), 111.

[27] Patel, N. D., Mehtre, B. M., & Wankar, R. (2022, October). Detection of Intrusions using Support Vector Machines and Deep Neural Networks. In 2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO) (pp. 1-5). IEEE

[28] Yao, H., Li, C., & Sun, P. (2020). Using Parametric t-Distributed Stochastic Neighbor Embedding Combined with Hierarchical Neural Network for Network Intrusion Detection. Int. J. Netw. Secur., 22(2), 265-274.

[29] Moualla, S., Khorzom, K., & Jafar, A. (2021). Improving the performance of machine learning-based network intrusion detection systems on the UNSW-NB15 dataset. Computational Intelligence and Neuroscience, 2021, 1-13.

[30] Idrissi, I., Azizi, M., & Moussaoui, O. (2021). Accelerating the update of a DL-based I