Measurement: Sensors 28 (2023) 100827

Contents lists available at ScienceDirect

Measurement: Sensors

journal homepage: www.sciencedirect.com/journal/measurement-sensors

A comprehensive review of AI based intrusion detection system

Sowmya T.[a,*], Mary Anita E.A.[b]

[a] Department of CSE, Christ University Bangalore, CMR Institute of Technology, Bengaluru, India
[b] Department of CSE, Christ University, Bangalore, India

ARTICLE INFO

Keywords:
Intrusion detection system
Machine learning
Deep learning
Artificial intelligence
Ensemble learning

ABSTRACT

In today's digital world, the tremendous amount of data poses a significant challenge to cyber security. The complexity of cyber-attacks makes it difficult to develop efficient tools to detect them. Signature-based intrusion detection has been the common method used for detecting attacks and providing security. However, with the emergence of Artificial Intelligence (AI), particularly Machine Learning, Deep Learning and ensemble learning, promising results have been shown in detecting attacks more efficiently. This review discusses how AI-based mechanisms are being used to detect attacks effectively based on relevant research. To provide a broader view, the study presents taxonomy of the existing literature on Machine Learning (ML), Deep learning (DL), and ensemble learning. The analysis includes 72 research papers and considers factors such as the algorithm and performance metrics used for detection. The study reveals that AI-based intrusion detection methods improve accuracy, but researchers have primarily focused on improving performance for detecting attacks rather than individual attack classification. The main objective of the study is to provide an overview of different AI-based mechanisms in intrusion detection and offer deeper insights for future researchers to better understand the challenges of multi-classification of attacks.

1. Introduction

Cloud computing is a modern technology that enables efficient data storage and processing. Due to the exponential growth of data, cloud computing has become more important than traditional computing in recent years. It is a widely adopted technology because of its scalability, cost-effectiveness, and flexibility [1]. According to Oracle's predictions, there will be 600 times more confidential data in the cloud by 2025. However, security is a major concern in cloud computing due to the distributed and open nature of the cloud architecture. While multi-tenancy is an advantage in cloud computing, it also poses a serious challenge to cloud security (see Fig. 3).

To address vulnerable attacks, many tools and mechanisms have been developed for cloud security, including various intrusion detection systems (IDS). This paper focuses on the study of different intrusion detection mechanisms that can alert network administrators by detecting all known and unknown attacks. Basically, there are network-based, host-based, hypervisor-based, and distributed-based intrusion detection mechanisms. Network-based monitoring monitors the entire traffic in the network, while host-based system detects the traffic only on that particular host [2]. IDS employs signature-based, anomaly-based, and hybrid-based techniques for detection purposes. Signature-based techniques detect intrusion by comparing known patterns or a predefined set of rules, while anomaly-based techniques focus on the current user's activities to recognize interruptions [3]. Anomaly management frameworks have been developed for detecting attacks, but complex rules have to be developed for larger datasets, which can be cost-reducing, time-consuming, and error-prone [4]. However, the firewall mechanism has proven ineffective for multi-cloud environments.

[5] Various ML approaches have been introduced to find anomalies and categorize different types of attacks [6]. ML is a part of artificial intelligence that can learn features and adapt to changing environments [7]. Statistical and ML algorithms have proven to be highly efficient for intrusion detection. In Ref. [8], the author identified the power of various ML algorithms and analysed the effect of ML algorithms for intrusion detection. With the increase in critical data in the cloud, an advanced form of ML algorithms called the DL concept has been introduced [9]. This method follows a layered approach that can recognize audio and speech processing applications.

DL has several advantages, including automatic feature learning, which extracts features automatically and trains the model even for large amounts of data [10]. The performance level of DL is higher than

* Corresponding author.
E-mail addresses: sowmya.t@res.christuniversity.in (T. Sowmya), maryanita.ea@christuniversity.in (E.A. Mary Anita).

https://doi.org/10.1016/j.measen.2023.100827
Received 30 October 2022; Received in revised form 11 May 2023; Accepted 12 June 2023
Available online 21 June 2023

that of ML algorithms [11,12]. studied and identified the importance of various DL algorithms and analysed the performance of each algorithm in the intrusion detection domain. By utilizing an efficient DL algorithm, the main drawback of false-positive and zero-day attacks can also be solved. This paper presents a concise picture of different AI-based intrusion detection approaches in the cloud and network. It reviews and classifies different AI-based IDS and provides a comparison of ML, DL, and ensemble learning-based intrusion detection mechanisms. It also provides a taxonomical view of all the AI-based intrusion detection mechanisms.

The main objective of this paper is to review and classify different AI-based IDS in the cloud and network. The paper provides a fine-grained review of ML, DL, and ensemble models for intrusion detection and proposes a novel classification in the field of AI-based intrusion detection. The paper evaluates the experimental results in tabular form for further use in the research community. This comparative analysis helps researchers compare ML, DL, and ensemble-based intrusion detection in detail. Finally, the paper identifies challenges in several commonly used ML, DL, and ensemble algorithms for the intrusion detection domain and outlines several possible future directions for the research community. The main goal of the study is to provide a comprehensive comparison of ML, DL, and ensemble-based techniques in the intrusion detection domain and insights for future researchers.

The paper is structured as follows: Section 2 provides a review of related work; Section 3 presents a classification and comparison of AI-based intrusion detection systems; Section 4 discusses the challenges faced by these systems; and Chapter 5 concludes the paper.

2. Review of related work

Safeguarding the cloud and network is an indispensable duty performed by Intrusion Detection Systems in Cybersecurity, but at the start IDS relied on humans observing and identifying intrusions which proved ineffective when dealing with advanced types of cyber threats. Cybersecurity experts have realized that with increased use of cloud computing comes a greater need for IDS to identify any possible breaches and ensure safety, and effective ways to identify attacks were proposed by many experts who studied it extensively over time using signature based as well as anomaly based or hybrid methods. While there were several security measures in place, none could identify threats amidst dense cloud and network traffic, but these limitations are addressed by the development of AI-based IDS for quick and effective analysis of attacks. In recent times many research articles have shown different ML algorithms for IDS which further augment the capability of these systems.

In the literature, we presented several research works based on ML, and some of them cover different aspects of DL methods for intrusion detection. One of the old surveys [5] reviewed various ML models for the IDS. Since it is an earlier survey, it doesn't cover the recent developments in the IDS. The use of classic ML models such as support vector machines, decision trees, shallow Nave Bayes networks, and genetic algorithms is reviewed in detail. The survey reviews the intrusion detection frameworks published between 2000 and 2007. In addition to that, the author categorizes the models into three classes: single, hybrid, and ensemble. The division presented in this research work provides a fine-grained overview of different ML-based intrusion detection techniques published between 2000 and 2007.

[13] Provided a study of DL approaches based on intrusion detection approaches and their comparative study. Specifically, the study focuses on various intrusion detection datasets, like network traffic-based and electric network-based datasets. This study analysed the performance of DL models like deep neural networks and RNNs (recurrent neural networks) in two classifications: binary and multiclass classification, under two datasets, namely the CSE–CIC–IDS2018 dataset and the Bot-IoT dataset. In particular, the author compared the performance of DL approaches with ML approaches using performance indicators, namely false alarm, accuracy, and detection rate. Of the recent survey papers [14], this provides the most comprehensive review of ML and DL models employed in IoT security. Further, the threats are classified into cyber and physical, which are popular in the IoT security domain. The performance of various ML and DL models by comparing their advantages and disadvantages is also noted. However, the coverage of ML and DL methods is limited, as few reviews are relevant for intrusion detection in the cloud [15]. analyzes the performance of DL models, namely multilayer perceptron, sparse auto encoder, restricted Boltzmann machine, and MLP with feature embedding's, on NSL-KDD and UNSW-NB15 intrusion datasets. Although the article lacks coverage of DL models like recurrent neural networks on newer intrusion detection datasets,

[16] Presents a taxonomy that classifies DL models into generative and discriminative architectures. However, the authors also note that both CNN and DBN have not been exploited in the field of IDS to detect attacks. Further, the authors compared the performance of DL and shallow learning models in the field of IDS. In the work [17], the author discusses the cyber security technology trends in intrusion detection utilizing ML and DL methods. However, the present paper does not cover all the methods in the intrusion detection domain; the authors concentrate only on the papers published in the past three years. Further, the authors use few benchmark datasets for the model, and the analysis is not uniform. None of the papers covers a deep and insightful analysis of the performance of the model.

3. Classification and comparison of AI based intrusion detection system

Security is really a necessity because of the massive use of data and the internet. Studies have been ongoing in order to develop an automatic detection system for detecting abnormal traffic. With the arrival of cloud computing, more industries are switching to it than ever before. With this trend, Cybersecurity issues will also increase drastically. Therefore, adequate solutions are required to ensure the proper operation of the cloud and network. AI-based intrusion detection provides an attractive solution for detecting and classifying attacks.

The primary goal of IDS is to detect network packets by critically inspecting them and to report them to administrators by generating alarms. After traditional technologies fail, an IDS acts as an adaptable safeguard for network security. Since cloud infrastructure handles enormous traffic and traditional technologies failed to detect and classify the attacks, The IDS can detect binary and multiclass classifications. For binary, the model outputs the label as normal or attack data, and for multiclass, the output will be multivalued, comprising different types of attacks. The core of the detection system is to detect intruders and classify attacks based on the situation. In addition to the detection of intrusions, prevention capabilities can also be provided to stop possible incidents. In Ref. [17], the combined mechanism is called intrusion detection and prevention systems. Based on the method of detection, there are signature-based and anomaly-based detection mechanisms. Signature-based intrusion detection mechanisms identify specific patterns and compare the pattern against the observable events only. Although the method can detect all the known attacks since the patterns are different for detecting novel attacks, the mechanism is not sufficient [18]. IDS techniques can also detect anomalies by detecting any deviation from normal behavior. The mechanism is also called misuse detection. The detection mechanism is purely based on three elements: parameterization, training, and detection. In the parameterization step, the observed behaviors like network connections and hosts are represented. In order to build a classification model that classifies these observed behaviors as normal or abnormal, the training module is utilized. For detection, the classified model constructed in the previous step is used to predict new anomalies.

However, most of the studies are based on the above-mentioned traditional models, and none of the methods are powerful. The problem can be formulated from a ML perspective as well as from a DL

perspective, which is a sub-branch of ML. The output of a ML problem is a label with normal or attack. Even though the previous ML models showed performance when compared with the traditional algorithms, none of them showed excellent performance by showing high accuracy and a low false alarm rate. According to previous studies, the DL method of intrusion detection is advantageous for predicting attacks with high accuracy. In other words, a DL-based intrusion detection system shows better prediction than ML.

We summarized the related studies in terms of classification with respect to ML, DL, and ensemble-based intrusion detection systems. The classification is composed mainly of performance and algorithms. We categorized the related studies in terms of ML, DL, and ensemble-based intrusion detection systems and compared the performance with respect to the performance metrics.

In this section, we propose a classification of AI-based intrusion detection systems. Fig. 1 represents AI-based intrusion detection techniques in detail from a ML, DL, and ensemble learning point of view. The novelty of the classification is to collect all the types of ML, DL, and ensemble learning methodologies that have the capability to detect intruders. However, this classification may help researchers in the field of cyber security detect and classify attacks efficiently (see Fig. 2).

We have surveyed several AI-based IDS methodologies in detail and compiled their performance metrics. Fig. 1 depicts the novel classification of AI-based intrusion detection mechanisms. Finally, a comparative study is performed on the three branches of the AI-based IDS. In this study, we adopted seven models from ML and five models from DL. Since ensemble is a combined methodology of various learning algorithms, we compiled only a few techniques for review purposes. Here we reviewed and compared the performance of popular AI-based algorithms in the domain of intrusion detection using AI techniques. Our detailed evaluation and comparative analysis illustrate the accuracy and other performance metrics of the selected algorithms. This detailed evaluation may become the reference point for future researchers to design and develop an efficient intrusion detection system.

3.1. Review and comparison of ML based intrusion detection system

Cybersecurity uses ML algorithms to make many important predictions to block attacks by dropping the data in order to avoid cybercrimes. This section mainly focuses on the commonly used ML algorithms, which act as a security tool to predict attacks. According to Ref. [19], ML has proven to be the most powerful security tool in detecting attacks, and understanding the system's semantic properties sounds crucial for the development of an anomaly detection system. Understanding the threat model by identifying the attacker's behavior or system environment is useful for constructing the ML-based security tool.

Artificial Intelligence, mainly ML technologies, considers the learning styles to model the algorithm. The property of the input data makes it an ideal choice for the classification of the algorithm. ML offers mainly supervised and unsupervised algorithms for classification based on the training data available. Supervised learning can be taught through teacher-student relationships in which the training dataset trains the target dataset and classifies the labeled dataset. Supervised learning yields accurate performance in the classification tasks in intrusion detection, and it is one of the key models in anomaly detection. Whereas unsupervised learning deals with unlabeled data and allows the user to categorize the data based on similarity. This type of learning

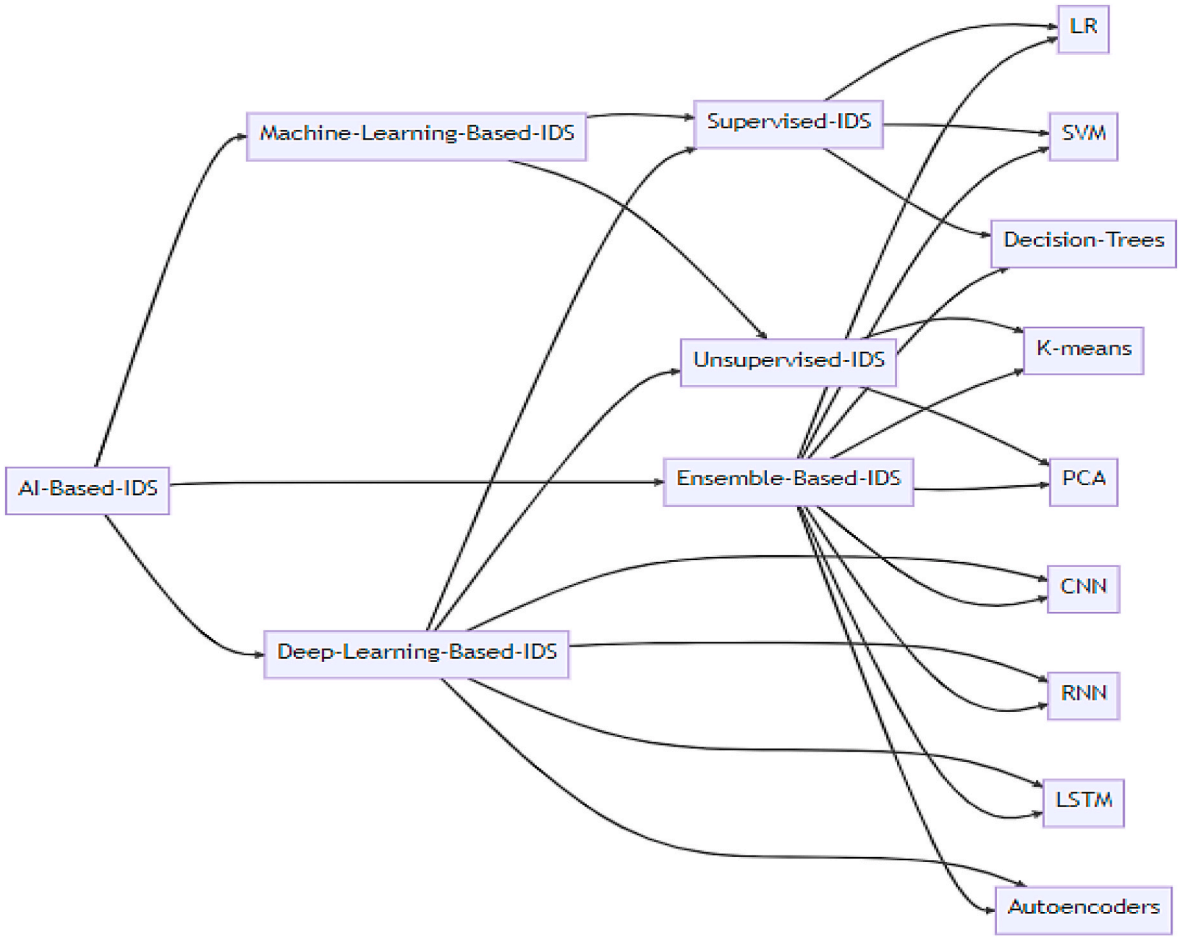

Fig. 1. Classification of AI based IDS.

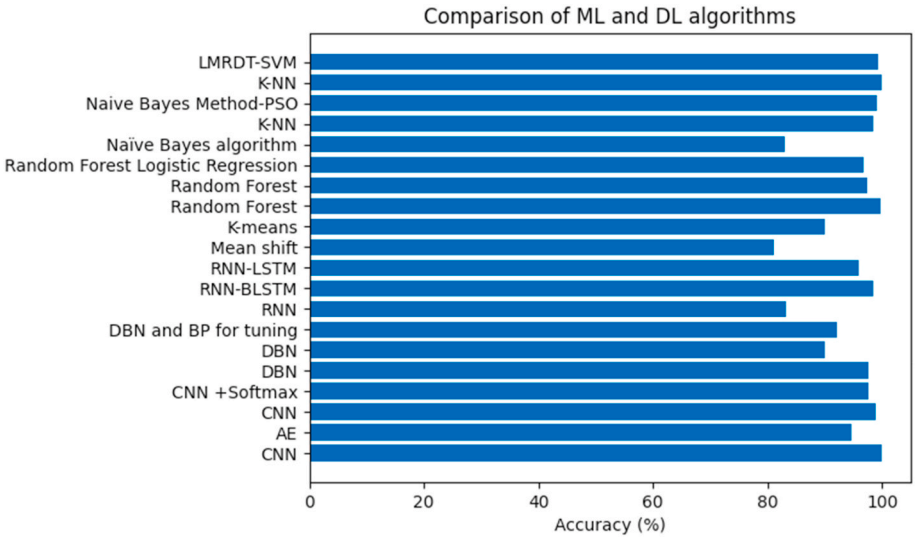

Fig. 2. An analysis of the accuracy performance of ML versus DL algorithms.

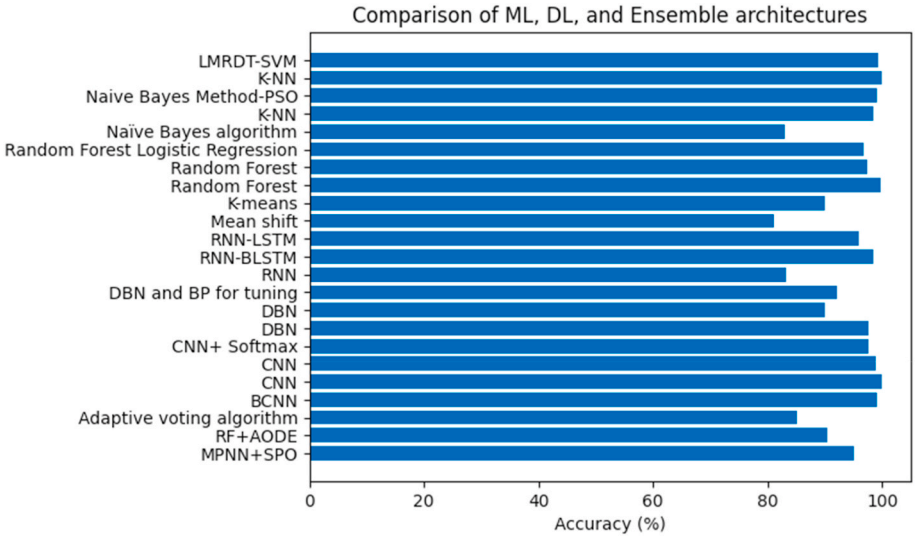

Fig. 3. An analysis of the accuracy performance of ML DL versus Ensemble models.

works as an ideal tool for efficiently analyzing undiscovered patterns. The review covered [20] the importance of supervised and unsupervised ML algorithms and their specific applications. Like all the other approaches, the model has some pitfalls in that it provides less accurate output. Although ML still has so many limitations, for example, the requirement of massive data, ML is used in all industries. Based on the learning style, various taxonomies in intrusion detection have already been introduced by the researchers. The classification consists of training and testing phases.

3.1.1. Supervised ML for intrusion detection

A ML task maps output data based on input-output pair relationships. For empirical analysis, we have considered common machine algorithms for intrusion detection. The main aim of the supervised learning style is to achieve a certain goal based on the training samples. The most common tasks in this learning style are classification and regression mechanisms.

3.1.1.1. Support vector machine. SVM is a classification technique that yields state-of-the-art performance in classification and prediction, and it is one of the key models in ML literature. The key idea of SVM is to construct a high-dimensional feature space from the input vectors in terms of the hyperplane and to segregate them into two classes, namely positive and negative. Positive data represents normal data, and negative data represents abnormal data. The work by Ref. [21] proposes a novel method based on statistically based IDS that performs in two stages. The proposed approach is based on the least squares SVM, in which the entire KDDcup99 dataset is separated into two subgroups. First, the proposed approach selects subsets of samples from the subgroups. Based on the vulnerabilities in the subgroup, the study also considered an optimum allocation scheme. The next step is to extract the samples for intrusion detection using the proposed approach. In Ref. [22], SVM is employed for selecting the features, which in turn provides better accuracy in the classification of attacks.

3.1.1.2. Logistic regression. Logistic regression is one of the supervised algorithms that is used for both binary and multiclass classification tasks. The principle behind logistic regression is to predict the occurrence of an event based on the concept of probability. Logistic regression is not only limited to the context of intrusion detection but also to domains like spam detection [23]. proposed a novel approach for IDS based on the genetic algorithm and logistic regression. The genetic

algorithm with the proposed best feature set selection method is used to reduce the features of the dataset before the classification process starts. The proposed feature selection methodology selects the best feature based on the fitness score, and after that, the logistic regression (LR) algorithm classifies the NSL-KDD dataset into different attacks. The result showed the success of GA-BFSS with the LR algorithm to minimize the number of features, thereby improving logistic regression performance. Another work [24] reported better performance of linear discriminant analysis and LR-based intrusion detection systems for binary and multi-class classification compared to the existing complex supervised ML techniques.

3.1.1.3. KNN algorithm. K-Nearest Neighbor is the simplest supervised ML methodology that can easily learn complicated approaches quickly. KNN is one of the most widely used algorithms for classification and regression tasks. KNN classifies the new data according to similarity and votes the data points based on the Euclidean distance measurement. Suppose the value of K is 5, it will look at the five nearest neighbors and determine how to classify the new input data by measuring with respect to the Euclidean distance [25]. proposed a new model of IDS threat that combines KNN and density peak clustering. The KNN was used for classification, and density peak clustering was used for training purposes. The KDD CUP dataset was used to evaluate the proposed IDS and choose the best 15 features out of 41 features. It was shown that the proposed methodology shows better accuracy of 15% in Probe attacks compared with the existing methodologies. The main drawback of the proposed method is its low detection accuracy [26]. suggested a self-adaptive differential evolution to choose the optimal features in IDS. The proposed approach was validated using the KDD CUP99 dataset, and the entire dataset is subdivided into four smaller sets of data. The proposed SaDE and KNN are applied separately to the four subsets, and the performance shows that the proposed approach may not show better performance if the approach is applied to the whole dataset.

3.1.1.4. Bayesian. The Bayesian classification was developed in the context of supervised learning. An approach relies on the fact that feature likelihood is predicted from the data. To classify, it calculates the probability to determine the likelihood using the Bayes Theorem. For network intrusion, most of the research has been used to detect abnormal traffic. In Ref. [26], we developed a new approach to IDS based on the infinite bounded mixture model and Bayesian. The model was designed to provide security in the IoT. This method uses an infinite bounded mixture model with a Bayesian approach for feature selection, and the Bayesian approach has been mainly employed for feature selection and parameter estimation. Another work [27] proposed a Bayesian Network Model averaging classifier that first subdivides the KDD CUP 99 dataset into subsets of data to build the classifier. Samples from the whole dataset are used for training purposes, and the entire dataset is used for testing. The next step is to perform data discretization on the training and testing data, followed by finding the K-best BN structures based on the training data. The final step is to estimate the conditional probability distribution of each network using Bayesian estimation to generate K-independent BN classifiers that will be applied to the training data. The result of the study showed good accuracy on the NSL-KDD dataset.

3.1.1.5. Random forest algorithm. In this supervised method, the algorithm mainly relies on multiple decision trees, which is one of the key models in ML architecture. Instead of single decision trees, prediction outputs from multiple decision trees are used to predict the final outcome. The approach is based on the ensemble method, which makes Random Forest a good candidate in the field of detecting attacks efficiently in the cloud as well as on a network. For network intrusion detection, random forests act as benchmark models [28]. Provides a survey of various techniques that have been proposed towards the enhancement of intrusion detection. The author also compares various classifiers for intrusion detection on the DARPA dataset. Thus, the random forest is more advantageous in terms of classification than other models. Even though the random forest model shows comparable performance in intrusion detection, the time taken is longer than other classifiers. This makes Random Forest undesirable for real-time intrusion detection applications in the context of real-world applications.

Table 1 compares the performance of various supervised ML (ML) based intrusion detection systems (IDS) in detecting network attacks. The table lists the name of the ML architecture, the dataset used in the study, the corresponding research article, the attacks addressed by the architecture, whether feature reduction or dimensionality reduction was employed, and the performance metrics achieved in the experiments.

The ML-based IDS include SVM, LMRDT-SVM, K-NN, Naive Bayes Method-Particle Swarm Optimization, and Random Forest Logistic Regression. The IDS were tested using different datasets, such as KDD'99, NSL-KDD, KDD CUP-99, UNSWNB-15, and AWID, to evaluate their accuracy in detecting different types of network attacks, including DoS, Probe, U2R, and R2L. The performance metrics range from an accuracy rate of 83% for the Naïve Bayes Algorithm to 99.89% for K-NN. Overall, the results show that the LMRDT-SVM architecture achieved the highest accuracy rate of 99.31%, followed by K-NN and Random Forest Logistic Regression with accuracy rates of 99.89% and 96.78%, respectively.

The table also indicates whether feature reduction or dimensionality reduction was employed in the IDS. The Naive Bayes Method-Particle Swarm Optimization, LMRDT-SVM, and Naïve Bayes Algorithm employed feature reduction, while Random Forest Logistic Regression used dimensionality reduction. The performance metrics reported include detection rate, FAR, error rate, sensitivity, and specificity.

3.1.2. Comparison of supervised ML based IDS

The existing supervised ML methods provide a secured mechanism to detect intruders in the cloud as well as in networking. The feature reduction and dimensionality reduction provide an added advantage in order to detect and classify the attacks. The researchers addressed a secure mechanism to detect active attacks like DoS, probe, U2L, and R2R with good detection rate and accuracy. In Ref. [31], the author addressed DoS, probe, U2L, and R2R attacks using KNN in the KDD dataset with a very good accuracy of 99.89%. All the other approaches used feature reduction and dimensionality reduction, which strengthened the attack prediction mechanism. Even though all the mechanisms have maintained originality by selecting all the relevant features, A comparison of supervised ML-based mechanisms is depicted in Table 1.

3.1.2.1. Unsupervised ML for intrusion detection. The unsupervised learning technique uses clustering based on similarities. The model itself can find the hidden pattern in the unlabeled dataset. The model can find the hidden pattern in the underlying dataset by means of clustering as well as association. Unsupervised learning is real AI that mimics humans' learning skills from past experience.

3.1.2.2. Fuzzy C means clustering algorithm. Fuzzy C means clustering, which is an unsupervised approach that allows assigning data points to one or more clusters. The algorithm assigns membership based on the distance between the cluster centers and data points [38]. The model aims to provide better classification accuracy and stability when tested and trained with the KDD 99 Cup dataset. Fuzzy C means clustering, which is a soft clustering approach in which each data sample is assigned to a cluster based on the likelihood score. The principle behind fuzzy clustering for intrusion detection is to identify and categorize the different types of attacks [39]. The paper shows the comparison between fuzzy C means and K means clustering algorithms for effectively detecting attacks. The paper concludes that fuzzy clustering can be practically used for network intrusion detection. This mechanism

Table 1
Supervised ML based IDS.

ML Architecture	Dataset	Article	Attacks addressed	Feature reduction/Di mensionality reduction	Performance metrics(%)
SVM	KDD'99	Kim et al. [29]	DoS,Probe, U2R, R2L	Yes	Detection rate: DoS-91.6,Probe-35.65, U2R- 12, R2L-22
LMRDT-SVM	NSL-KDD	Huiwen Wang.et al. [30]	No attacks addressed	Yes	Accuracy: 99.31 Detection rate: 99.20
K-NN	KDD	Lin et al. [31]	DoS,Probe, U2R, R2L	No	Accuracy: 99.89
Naive Bayes classifier.	NSL-KDD, UNSW-NB15, and CICIDS2017 datasets.	Monika Vishwakarma. et al. [32]	, DoS,Probe, U2R, R2L	Yes	NSL-KDD:Accuracy: 97 UNSW_NB15: Accuracy:86.9 CIC-IDS2017: Accuracy 98.59
K-NN		Wenchao Li.et al. [33]	No attacks addressed	No	Accuracy- 98.5 FAR-4.63
Naïve Bayes Algorithm	NSL-KDD	Sharmila B.S et al.	DoS,Probe, U2R, R2L	Yes	Accuracy- 83
Random Forest,Logistic Regression		S.Waskle et al.[35]	No attacks addressed	Yes	Accuracy-96.78 Error rate −0.21
Random Forest	UNSWNB-15	Belouch M et al. [36]	No attacks addressed	No	Accuracy-97.49
Random Forest	AWID	Abdulhammed R et al. [37]	802.11 MAC layer attacks	Yes	Sensitivity- 95.53, Specificity-97.75 Accuracy-99.64

detects a maximum of 45.95% and results in an efficient clustering method for detecting attacks.

3.1.2.3. K-means clustering algorithm. K means clustering approach is an unsupervised iterative mechanism which divides the unlabeled dataset into K clusters and assigns a membership for each data sample to a cluster based on the similarities The main idea behind the algorithm is centroid based and allocates each data sample to the closest centroid [40]. presented a hybrid approach to detecting anomalies by combining KNN and Naïve Bayes. Entropy based features selection algorithm is used to select only the important features. The system can detect as well as classify the attacks also. To evaluate the performance and to train and classify the data NSL KDD dataset is used. Another work [41] proposed an intrusion detection approach which is based on KNN and 1R classification in this approach K means clustering is used as a pre classification component and 1R classification is employed to classify the data into multiple classes. KDD Cup 99 dataset.

3.1.3. Comparison of unsupervised ML based IDS

Unsupervised ML-based intrusion detection systems were able to classify different types of attacks, as presented in Table 2. The section reviewed some classification algorithms to improve network security and evaluated their performance with respect to different metrics like accuracy, detection rate, FAR, and many more. Our studies focused on a single classification approach and a multiclassification algorithm in order to improve accuracy. Furthermore, the previews section reviewed the importance of ML algorithms in intrusion detection. From Table 2, we noticed that the best results were shown by the K-means clustering approach, which achieved the best accuracy for classification. However, the detection rate is an important factor in intrusion detection mechanisms and in significant changes. We conclude in Table 3 that in the comparative study of supervised and unsupervised ML, KNN is the best ML algorithm for intrusion detection, providing an accuracy of 99.89%. Although the approach showed better accuracy, it is limited to the selected set of attacks like DoS, Probe, U2R, and R2L.

Table 3
Comparison of ML based IDS based on accuracy.

ML Architecture	Article	Accuracy (%)
LMRDT-SVM	Huiwen Wang.et al. [30]	99.31
K-NN	Lin et a [31]	99.89
Naive Bayes classifier.	Monika Vishwakarma.et al. [32]	98.59
K-NN	Wenchao Li.et al. [33]	98.5
Naïve Bayes algorithm	Sharmila B S et al. [34]	83
Random Forest Logistic Regression	S. Waskle et al. [35]	96.78
Random Forest	Belouch, M et al. [36]	97.49
Random Forest	Abdulhammed, R et al. [37]	99.64
K-Means + RF	K. Samunnisa et al. [42]	92.77

Table 2 presents a comparison of unsupervised ML-based intrusion detection systems (IDS) that do not require labeled data. The table lists the architecture used in the study, the dataset employed, the corresponding research article, the attacks addressed, feature reduction or dimensionality reduction techniques used, and the performance metrics achieved. The unsupervised ML-based IDS include K-means, Fuzzy ARTMAP neural network, and K-means clustering. The datasets used in the studies include KDD'99, KDD Cup'99, and NSL KDD. The attacks addressed in the studies include DoS, Probe, U2R, and R2L attacks. Feature reduction or dimensionality reduction techniques were not used in the K-means clustering approach, while the Fuzzy ARTMAP neural network approach did not address any specific attacks. The performance metrics achieved include accuracy rates of 90% for K-means on KDD'99 and a detection rate of 99.42% and a false acceptance rate of 0.42% for the Fuzzy ARTMAP neural network on KDD Cup'99. Overall, the results indicate that unsupervised ML-based IDS can effectively detect network attacks with high accuracy rates, even without labeled data, making them a promising approach for intrusion detection.

Table 2
UnSupervised ML based IDS.

ML Architecture	Dataset	Article	Attacks addressed	Feature reduction/Di mensionality reduction	Performance metrics(%)
K-Means + RF	NSL-KDD	K. Samunnisa et al. [42]	DoS,Probe, U2R,R2L	Yes	Accuracy-92.77
Fuzzy ARTMAP neural network	KDDCup99	Khazaee et al. b [43]	No attacks adderessed	No	Detection rate-99.42 FAR-0.42
K-means	NSL-KDD	Vipin et al. [44]	DoS,Probe, U2R,R2L	No	82.29

3.1.4. Comparative study of supervised and unsupervised ML based IDS

The above supervised and unsupervised ML schemes are compared based on their performance parameters' accuracy. We observed that the best accuracy rate is observed in the KNN-based supervised learning approach, yielding an accuracy value of 99.89%, followed by Random Forest with an accuracy of 99.64%. The K-means algorithm has a poor detection accuracy of 90% compared to all the other ML-based models for detecting intruders. Early detection of attacks can be predicted accurately with the use of ML techniques. The KNN has been superior when compared to other ML approaches. On the other hand, our findings demonstrate that various ML algorithms perform better with feature reduction and dimensionality reduction. However, the KNN approach [31] worked better without the feature reduction approach.

Feature reduction can improve the accuracy of the algorithms used. Although most of the algorithms improved their accuracy after the feature reduction process and exhibited significant changes, We conclude in Table 3 that in the comparative study of supervised and unsupervised ML, KNN is the best ML algorithm for intrusion detection, providing an accuracy of 99.89%. Although the approach showed better accuracy, it is limited to the selected set of attacks like DoS,Probe, U2R, and R2L.

Table 3 compares the accuracy of various ML (ML) based intrusion detection systems (IDS) in detecting network attacks. The table lists the name of the architecture or algorithm used in the study, the corresponding research article, and the accuracy percentage achieved in the experiments. The ML-based IDS include LMRDT-SVM, K-NN, Naive Bayes Method-Particle Swarm Optimization, Naive Bayes algorithm, Random Forest Logistic Regression, Random Forest, Random Forest and K-means. The accuracy rates range from 83% for Naive Bayes algorithm to 99.89% for K-NN. Overall, the results indicate that K-NN achieved the highest accuracy rate among the listed ML-based IDS.

3.2. Review and comparison of DL based intrusion detection system

DL is a subset of ML and can be regarded as a complex evolution of ML algorithms. DL architectures use layered architectures for modeling complex concepts. The design of DL architectures is inspired by ANN, which in turn is inspired by the human brain. Currently, the DL approach is employed in many fields, such as image processing, speech recognition, and attack detection. DL architectures are different from ML architectures in the context of feature extraction. DL architectures will extract the features automatically, and the algorithm will learn by itself from errors. To make high-quality interpretations, DL architectures need to use a large dataset. This makes DL architectures more advantageous than ML. In addition to that, DL finds more complex relations and mappings between input data and output data [45]. [46] DL has been considered an advanced branch of ML in which multiple layers are connected by neurons, which represent mathematical computation for the learning process.

DL is an improved version of ML that performs feature extraction and classification tasks with multiple consecutive layers without any human intervention. All the multiple layers are interlinked, and the output from one layer is passed as an input for the second layer. Although DL is a subset of ML, the approach can be categorized into supervised and unsupervised. In this section, we discussed different supervised and unsupervised DL models used for analysis. The following DL methodologies for intrusion detection were considered for evaluation purposes:

3.2.1. Supervised DL for intrusion detection

Supervised DL uses a similar concept to supervised ML algorithms, but the difference in methodology is that it uses a computational network. It extracts the features automatically, teaches the learning algorithm by utilizing training data, and can be implemented with unknown data. Supervised learning has proven its effectiveness for intrusion detection by providing a highly accurate detection mechanism. Three supervised DL algorithms for intrusion detection are depicted as follows:

3.2.1.1. Artificial neural network. The structure of an artificial neural network is designed to mimic the human brain's functionality. The network is structured in the form of multiple layers, and the output from one layer is passed as an input for the succeeding layers. The output from the final layer represents the classification task. This makes them useful in domains such as concealed weapon detection, prediction and classification tasks of viruses, etc. Although ANN is a traditional approach in the field of IDS, further improvements have been proposed in order to overcome its disadvantages. From the base approach of ANN, several researchers proposed recurrent neural networks and deep neural networks, which fall under the category of DL approaches. In Ref. [47]'s investigation on artificial neural networks on IDS, the study used ten percent of the KDD CUP 99 dataset, and the feed-forward neural network was trained by the back propagation algorithm to detect and classify the intrusion. Further, the author reports only the false positive rate and false negative rate by comparing the performance of different models.

3.2.1.2. Convolutional neural network. A convolutional neural network is a type of neural network that can learn features automatically and classify attacks. They are made up of neurons, in which mathematical functions are the weighted sum of inputs to obtain the output function as the activation value. CNN is usually composed of several layers; the output from the first layer is passed as input for the next layers. Each layer can learn the features automatically, extract the features, and classify the data. Several studies have shown that the CNN methodology shows promising results for detecting normal and abnormal data [48]. proposed a few-shot DL strategy by combining deep convolutional neural networks with SVM and 1-NN classifiers. For intrusion detection, the author utilized a deep convolutional network, then implemented a support vector machine and 1NN for few-shot intrusion detection. and applied to two datasets, KDD 99 and NSL-KDD, and reported better accuracy of 96.19% and 86.74%, respectively. Finally, in Ref. [49], a convolutional neural network method based on Lenet-5 was proposed to identify the intrusions with a SoftMax regression layer. They observed that the evaluated approach shows promising results on the KDD 99 dataset. The results showed the highest accuracy rate of 97.53%. Furthermore, they observed that for more than 10,000 samples of data, the prediction accuracy rate may hit up to 97.53%.

3.2.1.3. Recurrent neural network. In recent years Recurrent neural networks have become a widely accepted approach due to the eminent computer resources. With the advancement of big data and computing resources DL has been widely adopted by researchers and makes RNN a good candidate for intrusion detection mechanisms Although recurrent neural networks has been introduced long decades ago, but their popularity came recently because of their memory property. Fundamentally speaking, in RNN the output from the previous layer is used as an input for the next layer.RNN can be designed as an IDS classifier that distinguishes normal data and abnormal data. However, with the rapid advancements in computing power, RNN has recently been integrated into the tasks of classification and regression. Although this approach outperforms well compared with other methods, one of the major shortcomings of this methodology was the vanishing gradient problem. This problem happens, and the weights cannot be changed further. Long-term short-memory networks can overcome this issue, and they can learn long-term dependencies easily. Soroush and Jean [57] aimed to improve the accuracy of network intrusion detection by introducing an approach using recurrent neural networks to generate unknown attacks. They conducted experiments on different publicly available datasets and obtained an outstanding detection rate of 16.67%. In Ref. [58], we aim to provide intrusion detection by introducing an approach based on RNN with stochastic gradient descent for classifying intrusions based on the UNSW-NB15 dataset. Through a set of

experiments with different activation functions, the proposed approach outperformed well for the Relu function, with the best accuracy of 98.99%. Another study [59] applied the LSTM-RNN approach to improving accuracy by using a novel multichannel attack detection. Moreover, they evaluated adding data preprocessing, feature abstraction, and multichannel training and detection and discovered that the proposed approach shows a high detection rate. Finally, a voting algorithm is utilized for classifying attack data and normal data. In general, several methods outperformed Bayesian or SVM classifiers.

3.2.1.3.1. Comparison of supervised DL based IDS. In this section, we aim to explore the performance comparison of DL architectures in the context of intrusion detection. Using feature reduction in datasets can maximize performance. The accuracy and other performance metrics of several DL algorithms are illustrated in Table 4. The performance metrics indicate that all the DL models are able to detect and classify the attack accurately. This general comparison of the above models indicates that CNN shows comparable performance that yields desirable accuracy. The purpose of evaluating DL algorithms was to study the ability of each model to improve classification efficiency. We noticed that many studies used the concept of cross-validation to improve accuracy. Additionally, the DL approach goes deeper from one layer to another by integrating neurons. Hence, all these facts contribute to making CNN perform better in attack detection as well as classification by providing an accuracy of 99.9%.

Table 4 compares different supervised DL-based intrusion detection systems (IDS) and their performance on various datasets and attacks. The DL architectures include artificial neural networks (ANN), convolutional neural networks (CNN), and recurrent neural networks (RNN). The datasets used include KDD-99, KDD 99 + NSLKDD, CAIDA DDoS 2007, CICIDS2017, and NSL-KDD. The attacks addressed in these IDS include DoS, U2R, R2L, Probe, Brute force, DDoS, and Botnet. The performance metrics used in the comparison include accuracy and false acceptance rate (FAR). Feature reduction and dimensionality reduction techniques are also applied in some of the models. The table also includes columns for complexity, advantages, and limitations. Overall, the DL-based IDS show promising performance in detecting and preventing various types of attacks, but their complexity and computational cost can be a limitation.

3.2.2. Unsupervised DL for intrusion detection

An unsupervised DL algorithm reduces the manual effort of labeling the dataset by predicting the output by learning itself. Furthermore, irrelevant features in the training data may lead to inaccurate results in supervised learning. Thus, the unsupervised DL approach overcomes the problem by predicting the output by means of finding automatic correlations between input and output. Two unsupervised DL algorithms for intrusion detection have been discussed in detail.

3.2.2.1. Auto encoder. Auto-encoder is an unsupervised ML architecture that has been investigated in this literature and yields state-of-the art performance on classification tasks and dimensionality reduction. Several ML and DL algorithms have been used for network intrusion detection and cloud intrusion detection. Hence, an auto-encoder-based method has been selected for our analysis purposes. The auto-encoder network is composed of three components: the coder, the code, and the decoder. The essence of auto-encoder networks involves using a back propagation algorithm to reduce the reconstruction loss. Reconstruction loss is a measure of how close the output is to the input [60]. presented a DL-based auto-encoder method for intrusion detection using the KDD99 dataset. Their model is stacked with four auto-encoders and achieves a reasonable accuracy above 95%. The proposed approach uses all the features and utilizes only 10% of the KDD99 dataset. Researchers in [, 61, 62] focused on detecting impersonation attacks using a three-layer stacked auto encoder using the AWID dataset. First, they used an artificial neural network as a hidden layer for feature selection. Then the stacked auto encoder is composed of two encoders with a softmax regression layer. They reported a high accuracy of 98.59% in detecting impersonation attacks using a stacked auto encoder with a softmax regression layer. Later in Ref. [62], the authors used the same approach by cascading two encoders with a Kmeans clustering algorithm and evaluating them using the KDD99 dataset. In addition, the approach provides a better accuracy of 94.8% in detecting impersonation attacks. In addition, the [61] approach outperforms compared to Ref. [62] to detect impersonation attacks.

3.2.2.2. Deep belief networks. In this learning mechanism, deep neural networks with a stack of RBMs are connected together. Like the auto-encoder model, the network can complete the dimensionality

Table 4
Supervised DL based IDS.

DL Architecture	Dataset	Article	Attacks addressed	Performance Metrics(%)	Feature reduction/ Dimensionality reduction	Complexity	Advantages	Limitations
ANN	KDD-99	Akashdeep [50]	DoS, U2R, R2L, Probe	DoS-99.93, U2R-96.51, R2L-92.54, Probe-98.7	Yes	Simple	Works well with large datasets	Requires extensive preprocessing and feature engineering
CNN + Softmax	KDD99 + NSLKDD	Lin et al. [51]	Smurf, Neptune, satan, ipsweep, portsweep	Accuracy-97.53	Yes	Moderate	Achieves high accuracy	Computationally expensive
CNN	KDD Dataset	Riyaz et al. [52]	DoS, U2R, R2L, Probe	Accuracy-98.88	Yes	Moderate	Achieves high accuracy	Computationally expensive
CNN	KDD 99 Dataset	Pengju Liu [53]	No attacks addressed	Accuracy-99.953, FAR-0.00022	No	Moderate	High accuracy and low false acceptance rate	Limited to binary classification
RNN-LSTM	CAIDA DDoS 2007	Sanchit Nayyar et al. [54]	Botnet, Bruteforce, DoS, DDos	Accuracy-96	No	Complex	Handles sequential data well	Requires large amounts of training data
RNN-BLSTM	CICIDS2017	S. Sivamohan et al. [55]	Brute force, DoS, DDoS	Accuracy-98.48	Yes	Complex	Handles sequential data well	Requires large amounts of training data
RNN	NSL-KDD	C. Yin et al. [56]	DoS, Probe, U2R, R2L	Training data-99.81, Testing data-83.28	No	Moderate	Achieves high accuracy on training data	Performance drops on testing data

Note: The complexity column indicates the complexity of the model architecture or algorithm, with "simple" indicating a relatively straightforward approach, "moderate" indicating a moderately complex approach, and "complex" indicating a highly complex approach. The advantages and limitations columns provide general benefits and drawbacks of each approach, respectively.

reduction and classification tasks by using the back propagation algorithm [63]. developed a real-time anomaly detection system for real-time vehicle networks. The training mechanism is provided by deep belief networks and then fine-tuned by a stochastic gradient descent layer for binary classification. The experimental results showed higher performance with feed-forward ANN [64]. presented an IDS model that uses DBN as a classifier for classifying network intrusion detections. The NSLKDD dataset, which includes normal and attack classes, was used to evaluate the proposed IDS by employing all the features. It was shown that the proposed DBN methods can produce improved training time and accuracy compared with DBN and DBN-SVM [65]. Implemented a hybrid anomaly detection system with DBN and probabilistic neural networks (PNN). First, DBN is used to learn, and to reduce the features, it is combined with PNN for classification purposes. Furthermore, the addition of the particle swarm optimization algorithm improved the accuracy, detection rate, and false alarm rate by 99.15, 93.25, and 0.615% respectively. From the KDD CUP 99 dataset, 10,000 data points were randomly selected for evaluation purposes.

3.2.2.2.1. Comparison of unsupervised DL based IDS. In Table 5, the use of DBN and AE unsupervised learning algorithms for intrusion detection are compared. The performance of various algorithms is assessed on various datasets.To maximize the accuracy it requires the convergence of feature reduction mechanism. On the other hand, feature reduction helps us to select better features and to eliminate the irrelevant features that lead to accurate results. We found that the Deep Belief network marginally achieved increased performance by providing an accuracy of 97.5% and surpasses AE based intrusion detection mechanism.

Table 5. Shows a comparison of unsupervised DL-based intrusion detection systems (IDS) using different architectures and datasets. The performance metrics used in the evaluation include accuracy, feature reduction or dimensionality reduction, complexity, advantages, and limitations. The architectures include deep belief networks (DBN) and autoencoders (AE), and the datasets used for evaluation are KDD CUP 99 and NSL-KDD. The attacks addressed include DoS, Probe, U2R, and R2L.

The table highlights the advantages and limitations of each architecture. DBN achieved high accuracy in identifying attacks, but it may require large amounts of training data and computational resources. AE is effective in feature reduction and identifying anomalies but may not perform well with complex and diverse datasets. DESC-IDS achieved high accuracy in identifying attacks and is effective in feature reduction but may not work well with complex and diverse datasets. Overall, the table suggests that unsupervised DL-based IDS can be an effective approach to identify network attacks. However, the choice of architecture and dataset should be carefully considered based on the specific needs and limitations of the system.

3.2.2.2.2. Comparative study of supervised and unsupervised DL based IDS. DL can learn the features automatically and compensate for the drawbacks of ML. Another consideration is the huge amount of data we can feed into the architecture for classification purposes. Numerous studies based on CNN have already been addressed; in Ref. [51], they have used CNN and SoftMax regression for classifying the attacks. A reason for this could be that CNN can automatically detect the features without any human intervention. All the models generally performed well on intrusion detection as well as the classification of attacks. RNNs are considered good at extracting time series data as well as textual data efficiently and therefore should perform well on intrusion detection mechanisms, but in our studies, CNN outperformed well on intrusion detection and classification. This suggests that CNN-based models are the more capable models in the benchmark of DL-based intrusion detection. In Table 6, the experiment results show that the accuracies of RNN are lower than those of CNN. It has been observed that about a.40% difference in accuracy rate appeared between CNN and RNN, according to our experiment.

3.2.2.2.3. Comparative study of ML and DL for intrusion detection system. In Table 7, we conducted a different comparison between ML and DL approaches for intrusion detection. No research paper has carried out a comparison between ML and DL in the field of intrusion detection in terms of attack. The major reason for doing this comparative study is to find a suitable model for predicting novel attacks accurately with a lower false alarm rate. We know that the data that spreads across the network and the cloud is vulnerable to attacks. To ensure security, ML and DL algorithms are employed for predicting and classifying attacks. Table 7 provides a complete understanding of selected ML and DL methodologies. The table indicates the accuracy measure of each model in detail. According to Table 7, CNN has the best performance in terms of accuracy. It gained over 99.9% accuracy for different epochs. As the number of epochs increases, accuracy also increases. The author also compared different learning rates, and their model outperformed well for 50 epochs. CNN reached a high level of 99.9% accuracy. Other DL models like AE and RNN also showed good performance in detecting the attacks.

ML algorithms also did their job very well and achieved a high

Table 6
Comparison of DL based IDS base.

DL Architecture	Article	Accuracy
RNN-LSTM	Sanchit Nayyar et al. [54]	96
RNN-BLSTM	S. Sivamohan et al. [55]	98.48
RNN	C. Yin et al. [56]	Training data-99.81 Testing data-83.28
DBN and BP for tuning	D. Kwon et al. [66]	92.1
DESC-IDS	Pengzhou Cheng et al. [67]	96.44
DBN	Z. Alom et al. [45]	97.5
CNN + Softmax	Lin et al. [51]	97.53
CNN	Riyaz et al. [52]	98.88
CNN	Pengju Liu [53]	99.95

Table 5
Unsupervised DL based IDS.

DL Architecture	Dataset	Article	Attacks addressed	Performance Metrics	Feature reduction/ Dimensionality Reduction	Complexity	Advantages	Limitations
DBN	KDD cup 99	Guangzhen et al. [66]	DoS,Probe, U2R,R2L	Accuracy-99.3	No	Medium	High accuracy in identifying attacks	May require large amounts of training data and computational resources.
DESC-IDS	KDD CUP 99	Pengzhou Cheng et al. [67]	No attacks addressed	Accuracy-96.44	Yes	Low	- Effective in feature reduction and identifying anomalies	May not work well with complex and diverse datasets.
DBN	NSL-KDD	Z. Alom et al. [64]	DoS,Probe, U2R,R2L	Accuracy-97.5	No	Medium	High accuracy in identifying attacks	May not perform well with large datasets.
AE	KDD CUP 99	Farahna kian et al.[,60]	DoS,Probe, U2R,R2L	Accuracy-94.71	Yes	Low	Effective in feature reduction and identifying anomalies	May not perform well with complex and diverse datasets.

Table 7
Comparison of ML and DL algorithm.

Learning Architecture	Article	Accuracy (%)
LMRDT-SVM	*Huiwen Wang.*et al. [30]	*99.31*
K-NN	*Lin et a* [31]	*99.89*
Naive Bayes Method	*Monika Vishwakarma .* et al. [32]	*98.59*
K-NN	*Wenchao Li.et al* [33]	*98.5*
Naïve Bayes algorithm	*Sharmila B S* et al. [34]	*83*
Random Forest Logistic Regression	*S. Waskle* et al. [35]	*96.78*
Random Forest	*Belouch, M* et al. [36]	*97.49*
Random Forest	*Abdulhammed, R et al* [37]	*99.64*
K-Means + RF	*K. Samunnisa et al et al.* [42]	*92.77*
Mean shift	*Kumar* et al. [44]	*82.29*
RNN-LSTM	*Sanchit Nayyar* et al. [54]	*96*
RNN-BLSTM	*S. Sivamohan* et al. [55]	*98.48*
RNN	*C. Yin* et al. [56]	*Training data-99.81 Testing data-83.28*
DBN and BP for tuning	*D. Kwon* et al. [66]	*92.1*
DESC-IDS	*Pengzhou Cheng* et al. [67]	*96.44*
DBN	*Z. Alom et al* [45]	*97.5*
CNN + Softmax	*Lin et al* [51]	*97.53*
CNN	*Riyaz* et al. [52]	*98.88*
AE	*Farahna kian* et al.*[,60]*	*94.71*
CNN	*Pengju Liu* [53]	*99.95*

accuracy rate. As we can see in Table 7, KNN also attained an accuracy rate of 99.8. It is interesting to note that the other ML models also showed better accuracy, but KNN was the first to come up with the highest accuracy measure. All the ML and DL models performed well for detecting the attacks as well as classification. Considering the performance metrics in terms of accuracy, detection rate, and FAR, DL models are more capable of detecting the attacks effectively. According to our studies, CNN performs well and can be employed as a base model when building an intrusion detection framework. This study may help future researchers in the field of intrusion detection keep an eye on CNN's approach. The outcome can be considered a fruitful aspect for researchers in detecting attacks. DL proves to be more advantageous in the intrusion detection domain because it has fitting and generalization capabilities. Another aspect is feature engineering, and the capability to handle large data requires DL mechanisms over ML. The results of this comparative study suggest using DL for feature evaluation to detect attacks. Our findings demonstrate that various DL algorithms perform better by selecting the features and reducing the dimensionality of the features. Our results indicate that CNN improves in performance to 100% if feature reduction is applied in advance. For instance, RNN acquired a low accuracy of 83.28 with testing data but can be increased by the feature reduction approach. On the other hand, DL-based IDS obtained a better accuracy of 99.9%. This indicates that CNN can be used as a baseline model for intrusion detection and can be fine-tuned for better performance to deal with intrusion detection. Therefore, expensive ML methods can be replaced with DL methods to yield better attack detection mechanisms. Finally, we found that there is no significant difference between ML and DL for intrusion detection. Both ML and DL performed well in detecting almost all the old attacks and may outdate the existing cloud and network environments. Future researchers can apply this model to the latest set of attacks.

Table 7 compares the accuracy of various ML (ML) and DL (DL) algorithms for intrusion detection systems (IDS). The table lists the name of the algorithm or architecture used in the study, the corresponding research article, and the accuracy percentage achieved in the experiments. The ML-based IDS include LMRDT-SVM, K-NN, Naive Bayes Method-Particle Swarm Optimization, Random Forest Logistic Regression, Random Forest, K-means, and Mean Shift. The DL-based IDS include RNN-LSTM, RNN-BLSTM, RNN, CNN + Softmax, and CNN. The accuracy rates range from 81.2% for Mean Shift to 99.95% for CNN. Overall, the results indicate that the CNN algorithm achieved the highest accuracy rate among all the algorithms tested. Additionally, the DL-based IDS generally achieved higher accuracy rates compared to ML-based IDS.

3.3. Review and comparison of ensemble based intrusion detection system

Ensemble learning is a learning algorithm that combines multiple classifiers to predict the output based on the output of the individual classifiers. Stacking is the most common ensemble technique available to combine the decisions from multiple models and predict the outcome based on that. Each individual model can contribute to the overall performance. Extensive research has been organized for the proposal of ensemble-based intrusion detection to detect and categorize the attacks. Ensemble learning also marks excellent performance by detecting and classifying attacks. It can be used for classification, regression, and feature selection tasks. In the era of the digital world, in order to deal with vast amounts of data, ensemble learning proves to be trustworthy. This learning mechanism deals with large datasets and divides the huge training data into smaller subsets. Each subset of data can be used as an input for the multiple classifiers to predict the combined output. Ensemble learning has become a winning strategy due to its data fusion mechanism. This may lead to improved accuracy in the detection and classification mechanisms. In recent years, the wide use of the ensemble approach has increased and has been deployed in several real-world aspects. According to Ref. [68], there are three major reasons behind the use of this efficient approach: statistical, computational, and representational.

3.3.1. Comparison of.ensemble based intrusion detection system

In this Table 8 a comparative analysis based on the performance is presented between the ensemble techniques of intrusion detection. The BCNN method is being used as it provides a highly efficient solution for the problem of intrusion detection domain. Constituting the results from Table 8 it can be seen that BCNN fetched an accuracy of 99.1 with no feature reduction. If accuracy is a major concern then ensemble based techniques should be chosen as it yields better accuracy for the specified set of attacks.

Table 8 compares the performance of various ensemble-based intrusion detection systems (IDS) using different performance metrics. The table lists the name of the ensemble architecture used in the study, the corresponding research article, the type of attacks addressed, the performance metrics achieved in the experiments, and whether the architecture used feature reduction or dimensionality reduction techniques. The ensemble-based IDS include BCNN, Adaptive Voting Algorithm, RF + AODE, and MPNN + SPO. The performance metrics evaluated include accuracy, false acceptance rate (FAR), and detection rate (DR). Overall, the results indicate that the ensemble-based IDS achieved high accuracy rates, ranging from 85.2% for Adaptive Voting Algorithm to 99.12% for BCNN. The feature reduction or dimensionality reduction techniques used in the architectures vary, with some using these techniques and others not. The main advantage of these ensemble-based IDS is their high accuracy rate, which makes them effective in detecting various types of attacks. However, they may have limitations such as being computationally expensive or not performing well with highly imbalanced data.

3.3.2. Comparative study of ML,DL and ensemble based IDS

In this table, all the AI-based intrusion detection techniques have been compared following the categories of ML, DL, and ensemble techniques. Intrusion detection approaches are reviewed according to the essence of learning methodologies. Ensemble-based methods are made from a combination of ML and DL. Together, all the ML, DL, and ensemble methods obtain better prediction accuracy than traditional methodologies. Even though there is no significant difference between ML, DL, and Ensemble technology, DL shows highly accurate results. Thus, AI-based techniques can be regarded as a potential technique for

Table 8
Comparison of Ensemble based IDS.

Ensemble learning Architecture	Article	Attacks addressed	Performance Metrics	Feature reduction/ Dimensionality Reduction	Complexity	Advantages	Limitations
BCNN	Zhang et al. [69]	No attacks addressed	DR-99.13 FAR- 1.18 Accuracy-99.12	No	High	High accuracy rate	• Computationally expensive
Adaptive voting algorithm	XIANWl et al. [70]	DoS,Probe, R2L and U2R	Accuracy-85.2	Yes	Low	• High accuracy rate • Reduces false positives	• May not perform well with highly imbalanced data
RF + AODE	M A Jabbar et al. [71]	No attacks addressed	Accuracy-90.51 FAR-0.14	No	Moderate	• High accuracy rate • Low false acceptance rate	• May not perform well with highly imbalanced data
MPNN + SPO	Musbau et al. [72]	No attacks addressed	Accuracy-95.02 DR-96.92 FAR-0.01	Yes	High	• High accuracy rate • Effective feature reduction	• Computationally expensive • May not perform well with highly imbalanced data

Note: DR stands for Detection Rate and FAR stands for False Acceptance Rate.

identifying and classifying attacks. All the AI-based methods outperformed well without making any significant difference. In addition, for the latest set of attacks, the AI-based methods should be optimized to improve the detection capability.

Table 9. Compares the accuracy of various ML (ML), DL (DL), and ensemble-based intrusion detection systems (IDS) in detecting network attacks. The table lists the name of the architecture or algorithm used in the study, the corresponding research article, and the accuracy percentage achieved in the experiments. The ML-based IDS include LMRDT-SVM, K-NN, and Naive Bayes Method-Particle Swarm Optimization, while the DL-based IDS include RNN-LSTM, RNN-BLSTM, and CNN. Ensemble-based IDS include BCNN, Adaptive Voting Algorithm, and RF + AODE. The accuracy rates range from 81.2% for Mean Shift to 99.95% for CNN. Overall, the results indicate that DL-based IDS achieved higher accuracy rates compared to ML-based IDS, and that ensemble-based IDS also performed well in detecting network attacks as shown in figure.

Table 9
Comparison of ML,DL and ensemble.

Learning Architecture	Article	Accuracy (%)
LMRDT-SVM	Huiwen Wang.et al. [30]	99.31
K-NN	Lin et a [31]	99.89
Naive Bayes Method	Monika Vishwakarma et al. [32]	98.59
K-NN	Wenchao Li.et al. [33]	98.5
Naïve Bayes algorithm	Sharmila B S et al. [34]	83
Random Forest Logistic Regression	S. Waskle et al. [35]	96.78
Random Forest	Belouch, M et al. [36]	97.49
Random Forest	Abdulhammed, R et al. [37]	99.64
K-Means + RF	*K. Samunnisa* et al. et al. [42]	92.77
K-means	Kumar et al. [44]	82.29
RNN-LSTM	Sanchit Nayyar et al. [54]	96
RNN-BLSTM	S. Sivamohan et al. [55]	98.48
RNN	C. Yin et al. [56]	Training data-99.81 Testing data-83.28
DBN and BP for tuning	D. Kwon et al. [66]	92.1
DEEC-IDS	Pengzhou Cheng et al. [67]	96.44
DBN	Z. Alom et al. [45]	97.5
CNN + Softmax	Lin et al. [51]	97.53
CNN	Riyaz et al. [52]	98.88
CNN	Pengju Liu [53]	99.95
BCNN	Zhang et al. [69]	99.12
Adaptive voting algorithm	XIANWl et al. [70]	85.2
RF + AODE	M A Jabbar et al. [71]	90.51
MPNN + SPO	Musbau et al. [72]	95.02

4. Challenges

AI has played a crucial role in detecting intrusions by providing high accuracy rates in intrusion detection systems. However, it has also presented some challenges, particularly when it comes to effectively detecting multiple attacks. While AI-based models prioritize accuracy, they do not always consider other important performance metrics such as F1 score, FAR, precision, and recall. As a result, there is a need to develop a framework that focuses on these metrics as well. Another issue with AI-based IDS is that most of the publicly available datasets used for attack detection are large and may contain noisy data that can affect performance. To address this, a mechanism can be employed to detect outliers in the data, which can be integrated into the framework to avoid overfitting. Additionally, most frameworks do not consider time complexity and CPU utilization, which can impact system performance. By taking these metrics into account, overall system performance can be improved. It is worth noting that while there is a lot of research on intrusion prevention systems, there is currently no real-world model for this type of system.

This section discuss the challenges that come with using AI-based intrusion detection systems (IDS), and gives some solutions to overcome them.

Challenge 1: Catching multiple attacks: Detecting one attack with an AI-based IDS is accurate, but it becomes tough to detect multiple attacks at the same time.

Solution: To detect multiple attacks effectively, the author suggests designing a framework that focuses not only on accuracy but also considers other metrics like F1 score, false acceptance rate (FAR), precision, and recall. These metrics will optimize the system to catch multiple attacks.

Challenge 2: Poor performance due to noisy data: The publicly available datasets for attack detection are large and may contain noisy data that can hurt the system's performance.

Solution: To improve performance, the author recommends developing a mechanism that detects and removes noisy data. This will help avoid overfitting and improve system performance.

Challenge 3: Ignoring the impact of time complexity and CPU utilization: Most frameworks overlook the effect of time complexity and CPU utilization on system performance.

Solution: To enhance overall performance, the author suggests that the framework should consider time complexity and CPU utilization when designing the system.

Challenge 4: Lack of a real-world model for intrusion prevention systems: Although many studies focus on intrusion prevention systems, there is no real-world model for them yet.

Solution: The author recommends developing a real-world model for

intrusion prevention systems, which will help test and evaluate the framework, leading to improved performance.

In inference, by designing a framework that considers different metrics, removing noisy data, and taking into account time complexity and CPU utilization, AI-based IDS can be optimized to improve performance in detecting and preventing intrusions.

5. Conclusion

In this paper, a comprehensive review of AI-based intrusion detection mechanisms is compared and studied. In this study, the existing AI-based intrusion detection mechanisms are classified, and the mechanisms employed for ensuring security are studied in detail. Compared to traditional security mechanisms, AI-based mechanisms showed the best performance in terms of detection and classification. The survey examines the importance of feature reduction for the performance of intrusion detection. Since feature reduction is employed in all AI-based mechanisms as a mandatory step in order to achieve accurate results, Among the AI-based mechanisms, ML, DL, and ensemble-based mechanisms showed almost similar performances of over 99%. The survey only addressed security attacks such as DoS, Probe, R2L, U2R, and the most common attacks. The AI-based schemes are compared against the accuracy metrics only. The other evaluation metrics, like detection rate and FAR, also have to be compared in future work. Moreover, in the future, we will check the feasibility of AI-based intrusion detection systems for a larger number of the latest attacks based on the existing network and cloud architecture. More importantly, a potential limitation of our study is the limited number of attack datasets. New datasets can also be used to compare performance. In addition, the study of the latest hybrid models for intrusion detection is also recommended in the future. To meet the growing need for security measures, an AI-based intrusion detection mechanism for unknown attacks becomes an open research issue.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

References

[1] Tsehay Admassu Assegie, An optimized KNN model for signature-based malware detection, Int. J. Comput. Eng. Res. Trends 8 (2021) 46–49.

[2] S.G. Kene, D.P. Theng, A review on intrusion detection techniques for cloud computing and security challenges, in: 2nd International Conference on Electronics and Communication Systems (ICECS), 2015, pp. 227–232, https://doi.org/10.1109/ECS.2015.7124898.

[3] S.A. M, P. G, A survey on various intrusion detection system tools and methods in cloud computing, in: 6th International Conference on Computing for Sustainable Global Development (INDIACom), 2019, pp. 439–445.

[4] K. Kulkarni, G. Ahn, H. Hu, Detecting and resolving firewall policy anomalies, IEEE Trans. Dependable Secure Comput. 9 (2012) 318–331.

[5] M. Bhavsingh, M.S. Lakshmi, S.P. Kumar, N. Parashuram, "Improved trial division algorithm by Lagrange" s, Interpol. Funct. 5 (2017) 1227–1231.

[6] Govindraj Chittapur, S. Murali, Basavaraj S. Anami, Copy create video forgery detection techniques using frame correlation difference by referring SVM classifier, Int. J. Comput. Eng. Res. Trends 6 (2019) 4–8.

[7] Teodoro Garcia, Jesus Diaz-Verdejo Pedro, Gabriel Maciá-Fernández, Enrique Vázquez, Anomaly-based network intrusion detection: techniques, systems and challenges, Comput. Secur. 28 (2009) 18–28.

[8] Adel Binbusayyis, Haya Alaskar, Thavavel Vaiyapuri, M. Dinesh, An investigation and comparison of ML approaches for intrusion detection in IoMT network, J. Supercomput. 78 (2022) 17403–17422.

[9] Mesut Uğurlu, Alper Doğru İbrahim, A survey on DL based intrusion detection system, in: 2019 4th International Conference on Computer Science and Engineering (UBMK), 2019, pp. 223–228, https://doi.org/10.1109/UBMK.2019.8907206.

[10] G. Karatas, O. Demir, O.K. Sahingoz, DL in intrusion detection systems, in: 2018 International Congress on Big Data, DL and Fighting Cyber Terrorism (IBIGDELFT), 2018, pp. 113–116, https://doi.org/10.1109/IBIGDELFT.2018.8625278.

[11] K. Kim, M.E. Aminanto, DL in intrusion detection perspective: overview and further challenges, Int. Workshop on Big Data and Inf. Secur. (IWBIS) (2017) 5–10, https://doi.org/10.1109/IWBIS.2017.8275095.

[12] Arwa Aldweesh, Abdelouahid Derhab, Ahmed Z. Emam, DL approaches for anomaly-based intrusion detection systems: a survey, taxonomy, and open issues, Knowl. Base Syst. 189 (2020), 105124.

[13] Ferrag, Mohamed Amine, Leandros Maglaras, Sotiris Moschoyiannis, Helge Janicke, DL for cyber security intrusion detection: approaches, datasets, and comparative study, J. Inf. Secur. Appl. 50 (2020), 102419.

[14] Al Garadi, Mohammed Ali, Amr Mohamed, Abdulla Khalid Al-Ali, Xiaojiang Du, Ihsan Ali, Mohsen Guizani, A survey of machine and DL methods for internet of things (IoT) security, IEEE Commun. Surv. & Tutor. 22 (2020) 1646–1685.

[15] J. Yan, D. Jin, C.W. Lee, P. Liu, A comparative study of off-line DL based network intrusion detection, in: Tenth International Conference on Ubiquitous and Future Networks, ICUFN), 2018, pp. 299–304, https://doi.org/10.1109/ICUFN.2018.8436774.

[16] Elike Hodo, Xavier Bellekens, Andrew Hamilton, Christos Tachtatzis, Robert Atkinson, Shallow and deep networks intrusion detection system: a taxonomy and survey, arXiv preprint arXiv:1701.02145 (2017).1-43.

[17] Zeeshan Ahmad, Adnan Shahid Khan, Wai Shiang Cheah, Johari Abdullah, Farhan Ahmad, Network intrusion detection system: a systematic study of ML and DL approaches, Transactions on Emerging Telecommunications Technologies 32 (2021), e4150.

[18] Garcia Teodoro, Jesus Diaz-Verdejo Pedro, Gabriel Maciá-Fernández, Enrique Vázquez, Anomaly-based network intrusion detection: techniques, systems and challenges, Comput. Secur. 28 (2009) 18–28.

[19] R. Sommer, V. Paxson, Outside the closed world: on using ML for network intrusion detection, in: IEEE Symposium on Security and Privacy, 2010, pp. 305–316, https://doi.org/10.1109/SP.2010.25.

[20] Michael W. Berry, Supervised and Unsupervised Learning for Data Science, Springer International Publishing USA, 2019.

[21] Enamul Kabir, A novel statistical technique for intrusion detection systems" Future Generation, Comput. Syst. 79 (2018) 303–318.

[22] Huaglory Tianfield, Data mining based cyber-attack detection, Syst. simul. technol. 13 (2017).

[23] A. Ponmalar, V. Dhanakoti, An intrusion detection approach using ensemble support vector machine based chaos game optimization algorithm in big data platform, Appl. Soft Comput. 116 (2022), 108295.

[24] Basant Subba, Santosh Biswas, Sushanta Karmakar, Intrusion detection systems using linear discriminant analysis and logistic regression, in: 2015 Annual IEEE India Conference (INDICON), IEEE, 2015, pp. 1–6.

[25] Kai Peng, Victor Leung, Lixin Zheng, Shangguang Wang, Chao Huang, Tao Lin, Intrusion detection system based on decision tree over big data in fog environment, Wireless Commun. Mobile Comput. (2018) 1–10.

[26] Yu Xue, Weiwei Jia, Xuejian Zhao, Wei Pang, An Evolutionary Computation Based Feature Selection Method for Intrusion Detection" Security and Communication Networks, 2018, pp. 1–10.

[27] L. Xiao, Y. Chen, C.K. Chang, Bayesian model averaging of Bayesian network classifiers for intrusion detection, in: 2014 in IEEE 38th International Computer Software and Applications Conference Workshops on 35, 2014, pp. 1302–1310.

[28] Htun hyu Thi, Khaing Kyaw Thet, Anomaly intrusion detection system using random forests and k-nearest neighbor, Int. J. P2P Netw. Trends Technol. (2013) 3.

[29] Dong Seong Kim, Jong Sou Park, Network-based intrusion detection with support vector machines, in: Information Networking: International Conference, ICOIN 2003, Revised Selected Papers, Cheju Island, Korea, 2003, pp. 747–756. February 12-14, 2003.

[30] Huiwen Wang, Jie Gu, Shanshan Wang, An effective intrusion detection framework based on SVM with feature augmentation, Knowl. Base Syst. 136 (2017) 130–139.

[31] Wei-Chao Lin, Ke Shih-Wen, Chih-Fong Tsai, CANN: an intrusion detection system based on combining cluster centers and nearest neighbors, Knowl. Base Syst. 78 (2015) 13–21.

[32] Monika Vishwakarma, Nishtha Kesswani, A new two-phase intrusion detection system with Naïve Bayes ML for data classification and elliptic envelop method for anomaly detection, Decision Analytics Journal 7 (2023), 100233, https://doi.org/10.1016/j.dajour.2023.100233.

[33] Wenchao Li, Ping Yi, Yue Wu, Li Pan, Jianhua Li, New intrusion detection system based on KNN classification algorithm in wireless sensor network, Journal of Electrical and Computer Engineering 1752 (2021) 1–8.

[34] B.S. Sharmila, Nagapadma Rohini, Intrusion detection system using naive bayes algorithm, in: 2019 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE), IEEE, 2019, pp. 1–4.

[35] Subhash Waskle, Lokesh Parashar, Upendra Singh, Intrusion detection system using PCA with random forest approach, Int. Conf. Electron. Sustain. Commun. Syst. (ICESC) (2020) 803–808.

[36] Mustapha Belouch, El Hadaj Salah, Idhammad Mohamed, Performance evaluation of intrusion detection based on ML using Apache Spark, Procedia Comput. Sci. 127 (2018) 1–6.

[37] Razan Abdulhammed, Miad Faezipour, Abdelshakour Abuzneid, Alessa Ali, Effective features selection and ML classifiers for improved wireless intrusion detection, in: 2018 International Symposium on Networks, Computers and Communications (ISNCC), IEEE, 2018, pp. 1–6.

[38] S. Ganapathy, K. Kulothungan, P. Yogesh, A. Kannan, A novel weighted fuzzy C–means clustering based on immune genetic algorithm for intrusion detection, Procedia Eng. 38 (2012) 1750–1757.
[39] Partha Sarathi Bhattacharjee, Md Fujail Abul Kashim, Shahin Ara Begum, A comparison of intrusion detection by K-means and fuzzy C-means clustering algorithm over the NSL-KDD dataset, in: 2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), IEEE, 2017, pp. 1–6.
[40] Hari Om, Aritra Kundu, A hybrid system for reducing the false alarm rate of anomaly intrusion detection system, in: 2012 1st International Conference on Recent Advances in Information Technology (RAIT), IEEE, 2012, pp. 131–136.
[41] Zaiton Muda, Warusia Yassin, Md Nasir Sulaiman, Nur Izura Udzir, Intrusion detection based on k-means clustering and OneR classification, in: 2011 7th International Conference on Information Assurance and Security, IAS), 2011, pp. 192–197.
[42] K. Samunnisa, G. Sunil Vijaya Kumar, K. Madhavi, Intrusion detection system in distributed cloud computing: hybrid clustering and classification methods, Measurement: Sensors 25 (2023), 100612.
[43] Saeed Khazaee, Maryam Sharifi Rad, Using fuzzy c-means algorithm for improving intrusion detection performance, in: 2013 13th Iranian Conference on Fuzzy Systems (IFSC), IEEE, 2013, pp. 1–4.
[44] Vipin Kumar, Himadri Chauhan, Dheeraj Panwar, K-means clustering approach to analyze NSL-KDD intrusion detection dataset, Int. J. Soft Comput. Eng. (IJSCE) 3 (4) (2013) 1–4.
[45] Li Deng, Yu Dong, DL: methods and applications, Foundations and trends® in signal processing 7 (2014) 197–387, 3–4.
[46] Y. Bengio, I. Goodfellow, A. Courville, "DL", MIT press. http://www.deeplearningbook.org, 2016.
[47] G. Poojitha, K.N. Kumar, P.J. Reddy, Intrusion detection using artificial neural network, in: 2010 in Second International Conference on Computing, Communication and Networking Technologies, 2010, pp. 1–7, https://doi.org/10.1109/ICCCNT.2010.5592568.
[48] Md Moin Uddin Chowdhury, Frederick Hammond, Glenn Konowicz, Chunsheng Xin, Hongyi Wu, Li Jiang, A few-shot DL approach for improved intrusion detection, in: 2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON), 2017, pp. 456–462.
[49] Wen-Hui Lin, Hsiao-Chung Lin, Ping Wang, Bao-Hua Wu, Jeng-Ying Tsai, Using convolutional neural networks to network intrusion detection for cyber threats, in: 2018 IEEE International Conference on Applied System Invention (ICASI), IEEE, 2018, pp. 1107–1110.
[50] Ishfaq Manzoor, Neeraj Kumar, A feature reduced intrusion detection system using ANN classifier, Expert Syst. Appl. 88 (2017) 249–257.
[51] Wen-Hui Lin, Hsiao-Chung Lin, Ping Wang, Bao-Hua Wu, Jeng-Ying Tsai, Using convolutional neural networks to network intrusion detection for cyber threats, in: 2018 IEEE International Conference on Applied System Invention (ICASI), IEEE, 2018, pp. 1107–1110.
[52] B. Riyaz, Sannasi Ganapathy, A DL approach for effective intrusion detection in wireless networks using CNN, Soft Comput. 24 (2020) 17265–17278.
[53] Pengju Liu, An intrusion detection system based on convolutional neural network, in: Proceedings of the 2019 11th International Conference on Computer and Automation Engineering, 2019, pp. 62–67.
[54] Sanchit Nayyar, Sneha Arora, Maninder Singh, Recurrent neural network based intrusion detection system, in: 2020 International Conference on Communication and Signal Processing (ICCSP), 2020, pp. 136–140, https://doi.org/10.1109/ICCSP48568.2020.9182099.
[55] S. Sivamohan, S.S. Sridhar, S. Krishnaveni, An effective recurrent neural network (RNN) based intrusion detection via bi-directional long short-term memory, in: 2021 International Conference on Intelligent Technologies (CONIT), IEEE, 2021, pp. 1–5.
[56] Chuanlong Yin, Yuefei Zhu, Jinlong Fei, Xinzheng He, A DL approach for intrusion detection using recurrent neural networks, IEEE Access 5 (2017) 21954–21961.
[57] Soroush M. Sohi, Jean-Pierre Seifert, Ganji Fatemeh, RNNIDS: enhancing network intrusion detection systems through DL, Comput. Secur. 102 (2021), 102151.
[58] M. Al-Zewairi, S. Almajali, A. Awajan, Experimental evaluation of a multi-layer feed-forward artificial neural network classifier for network intrusion detection system, in: 2017 International Conference on New Trends in Computing Sciences (ICTCS), 2017, pp. 167–172, https://doi.org/10.1109/ICTCS.2017.29.
[59] Feng Jiang, Yunsheng Fu, Brij B. Gupta, Yongsheng Liang, Seungmin Rho, Fang Lou, Fanzhi Meng, and Zhihong Tian. DL based multi-channel intelligent attack detection for data security, IEEE trans. Sustain. Comput. 5 (2018) 204–212.
[60] F. Farahnakian, J. Heikkonen, A Deep Auto-Encoder Based Approach for Intrusion Detection System, 2018 20th International Conference on Advanced Communication Technology (ICACT), 2018, pp. 178–183, https://doi.org/10.23919/ICACT.2018.8323688.
[61] Aminanto, Muhamad, Detecting Impersonation Attack in WiFi Networks Using DL Approach, International Workshop on Information Security Applications, 2017, pp. 136–147.
[62] Aminanto, Muhamad Erza, Kwangjo Kim, Improving detection of Wi-Fi impersonation by fully unsupervised DL, in: Information Security Applications: 18th International Conference, vol. 18, WISA 2017, 2017, pp. 212–223.
[63] Min-Joo Kang, Je-Won Kang, Intrusion detection system using deep neural network for in-vehicle network security, PLoS One 11 (2016).
[64] Md Zahangir Alom, VenkataRamesh Bontupalli, M. Tarek, Taha, Intrusion detection using deep belief networks, in: 2015 National Aerospace and Electronics Conference (NAECON), 2015, pp. 339–344.
[65] Guangzhen Zhao, Cuixiao Zhang, Lijuan Zheng, Intrusion detection using deep belief network and probabilistic neural network, in: 2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC), vol. 1, 2017, pp. 639–642.
[66] D. Kwon, K. Natarajan, S.C. Suh, H. Kim, J. Kim, An empirical study on network anomaly detection using convolutional neural networks, in: 2018 IEEE 38th International Conference on Distributed Computing Systems (ICDCS) on, 2018, pp. 1595–1598, https://doi.org/10.1109/ICDCS.2018.00178.
[67] Pengzhou Cheng, Mu Han, Gongshen Liu, DESC-IDS: towards an efficient real-time automotive intrusion detection system based on deep evolving stream clustering, Future Generat. Comput. Syst. 140 (2023) 266–281.
[68] Thomas G. Dietterich, An experimental comparison of three methods for constructing ensembles of decision trees: bagging, boosting and randomization, ML 32 (1998) 1–22.
[69] Jielun Zhang, Fuhao Li, Feng Ye, An ensemble-based network intrusion detection scheme with bayesian DL, in: ICC 2020-2020 IEEE International Conference on Communications (ICC), 2020, pp. 1–6.
[70] Xianwei Gao, Chun Shan, Changzhen Hu, Zequn Niu, Zhen Liu, An adaptive ensemble ML model for intrusion detection, IEEE Access 7 (2019) 82512–82521.
[71] Thomas G. Dietterich, An experimental comparison of three methods for constructing ensembles of decision trees: bagging, boosting, and randomization, ML 40 (2000) 139–157.
[72] Musbau Dogo Abdulrahaman, John K. Alhassan, Ensemble learning approach for the enhancement of performance of intrusion detection system, in: International Conference on Information and Communication Technology and its Applications (ICTA 2018), 2018, pp. 1–8.