Using Machine Learning Techniques for Accurate Attack Detection in Intrusion Detection Systems using Cyber Threat Intelligence Feeds

Ehtsham Irshad[†] and Abdul Basit Siddiqui[††],
ehtsham_irshad@hotmail.com abasit.siddiqui@cust.edu.pk
Department of Computer Science, Capital University of Science & Technology, Islamabad, Pakistan

Abstract
With the advancement of modern technology, cyber-attacks are always rising. Specialized defense systems are needed to protect organizations against these threats. Malicious behavior in the network is discovered using security tools like intrusion detection systems (IDS), firewall, antimalware systems, security information and event management (SIEM). It aids in defending businesses from attacks. Delivering advance threat feeds for precise attack detection in intrusion detection systems is the role of cyber-threat intelligence (CTI) in the study is being presented. In this proposed work CTI feeds are utilized in the detection of assaults accurately in intrusion detection system. The ultimate objective is to identify the attacker behind the attack. Several data sets had been analyzed for attack detection. With the proposed study the ability to identify network attacks has improved by using machine learning algorithms. The proposed model provides 98% accuracy, 97% precision, and 96% recall respectively.

Keywords:
Cyber-threat intelligence (CTI), Security Information and event management (SIEM), Intrusion detection system (IDS), Intrusion prevention system (IPS), Denial of service (DoS), Principal component analysis (PCA), Support vector machine (SVM), Indicators of compromise (IoCs), Network Intrusion Detection System (NIDS), Host based intrusion detection system (HIDS),Random Forest (RF), Decision Tree (DT), Artificial Neural Networks (ANN)

1. Introduction

Cyber-security is now a very important aspect in today's modern world. As networks and systems are increasing very rapidly, protecting networks and data from attacks is a very important aspect of today's research. In recent times, protection from various cyber-attacks is becoming pressing issues. [1-7]. The current setup like hardware and software firewalls, data encryption strategy, and user's authentication are not adequate to meet the challenges of the modern world. Unfortunately, this equipment is not able to secure the computer networks from cyber-attacks [8-10]. The role of artificial intelligence in this field is increasing with the passage of time and it is widely used in every industry [11-14].

Devices such as firewalls and IDS/IPS are used to secure networks. IDS is available in two varieties: signature-based and behavioral/anomaly-based. The common occurrence of false positive alarms, the extended time needed to identify assaults, as well as a failure to identify zero-day attacks, which destroy businesses, are just a few of the issues that are known to plague existing IDSs. Companies lose time in the investigative process due to the flaws in IDS backend engines. Deep packet inspection is carried out by IDS to detect malicious traffic in the network. Every packet that passes through it is examined, and the payload is compared to signature databases. The request is blocked if a match is discovered; otherwise, the network allows it to move on [15–18]. There are two IDS types. A host intrusion system (HIDS) is installed on the host to identify attacks, while a network intrusion system (NIDS) is utilized for network-based activity. The NIDS come in two varieties. One of them is based on signatures. When a request comes to IDS, it checks the request against the signature database because it has a repository of all known attack signatures. The request is denied or rejected if it matches the signature database; else, it is permitted to proceed. The second sort of detection is behavioral, or anomaly based. This kind is employed to identify unidentified attacks, such as zero-day attacks [19-20]. An anomaly is something that is abnormal/exception/outlier. Anomaly detection is a process of detecting these patterns in data that do not have pre-defined normal behavior [21].

Manuscript received April 5, 2024
Manuscript revised April 20, 2024
https://doi.org/10.22937/IJCSNS.2024.24.4.21

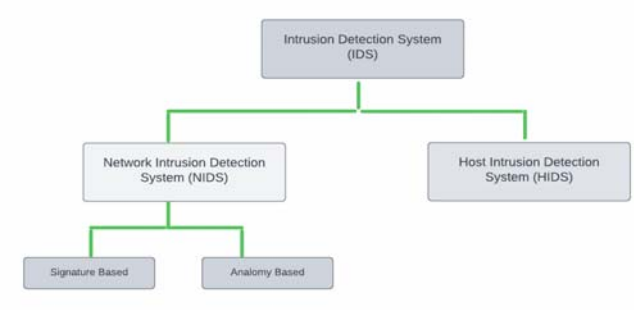

Fig. 1 Types of IDS

From the early 80s, a different mechanism has been devised to detect intrusion in the system. To make the system trustworthy and to fend against attacks, research is being done in this area. The classification of data into several categories in the modern day uses machine learning techniques [22]. Normally, traffic is classified in to two types: normal traffic and attack traffic, but according to some studies, there are really five categories. Attacks are further divided into 4 types: from remote to local, from user to root, probing assault and DoS [23–26].

Machine Learning's Role in Intrusion Detection Systems

While precise defense measures, such as machine learning-based IDS are required. They are being deployed as potential solutions for identifying network attackers [27-30]. To accomplish network intrusion detection (NIDS) employing a combination of algorithms, various ways have been proposed by researchers. There is a need for a method to categorize network assaults because the IDS continues to struggle more accurately with improving detection accuracy despite the significant research efforts [31–34].
Several researchers use ML/DLNN algorithms to evaluate the performance of attack discovery. Machine/deep learning methods can be used in three ways: individually, hybrid and collectively. Several data sets have been used to evaluate machine learning algorithms' performance. The most popular dataset for measuring performance is the NSL KDD Dataset [35–38].

Role of CTI in Intrusion Detection System

The knowledge base that includes context, behavior, actions, and the repercussion of this attack is known as cyber-threat intelligence (CTI). This information will be used to counter attacks in the future. CTI regularly updates information and context about cyber-attacks. It provides multi-source databases that assist cyber defense mechanisms, enabling thorough monitoring, detection, and reaction to online threats [39–41]. Without peering beyond the network of your organization, it can be challenging to predict when, when, and how an attack will occur. Using global CTI feeds will give you information on how an attack is happening and who is behind it. CTI feeds will either provide strategies for anticipating occurrences or potential countermeasures against assaults. These threat feeds will help you develop important defensive security strategies. All-important threat vectors are covered by the CTI feeds, including websites, social media, bot IPs, malicious URLs, phishing URLs, spam, and malicious URLs. It gives businesses a chance to assess the risks and challenges they face in cyberspace and empowers them to decide how best to respond to impending attacks [42–44]. Organizations are concentrating on creating their own knowledge repositories today using data that is available globally. For generating threat feeds, CTI uses internal community and outside sources. Data gathered from corporate security solutions like IDS/IPS, firewalls, and antivirus software, among others, is included in internal feeds. An example of an external source is a threat feed from a public source (an unreliable source), such as an anti-malware domain, or a paid private source from several well-respected and reputable security vendors.

Information sharing between organizations in the appropriate industry is a challenge today. Cyber-threat intelligence plays a key role in giving feeds to security devices such as IDS, SIEM, firewalls etc. These threat feeds are useful for organizations to protect against future attacks [45-46]. A lot of security organizations (AlienVault, threat connect etc.) are providing threat feeds. Organizations are incorporating these threat feeds into their devices [47-50]. These threat feeds are continuously updating. Some of the benefits of CTI feeds are:

- Reducing the risk component.
- Gathering useful information (IoCs).
- Protect your network and prevent data leaks.
- Assessing the security posture.
- Before, during, and after the attack, make use of cyber threat intelligence feeds.
- Transferring CTI feeds.

This research investigation has made the following contributions: -

This paper examines previous strategies for detecting IDS attacks. Machine learning and cyber-threat intelligence feeds play a role in this domain. The second contribution is the use of ML techniques for improved attack detection in IDS. Suggested a framework for accurate attack detection in IDS using different data sets.

The structure of the research paper is as follows. Literature review is included in Part II. The problem statement in section III. The datasets for IDS analysis are described in section IV. Section V outlines the proposed methodology. Part VI presents the results, while Section VII is conclusion and further work.

2. Literature Review

In this proposed study [51] a classifier approach for NIDS by using the tree algorithm is used. The author has proposed a combining tree classifier approach for detecting network attacks. First implemented individual tree algorithms (Random tree, C4.5, NB Tree) on NSL-KDD data to know the accuracy of the individual algorithm for detecting attacks. Then different algorithms are combined to determine the accuracy. In individual classifiers, the random tree gives the best accuracy and the NB tree gives the least accuracy. The author has determined that combining the least and most accurate classifier (Random tree+ NB tree) yields the best accuracy. In this study [52] Intrusion detection system framework is proposed. The author utilized a Bayesian classifier to discover anomalies in the network. NSL-KDD dataset is used as a benchmark. This paper has two phases. The wrapper approach was used for feature scaling. After applying feature scaling 16 feature set was used to obtain results instead of actual 41 features. In this study [53], ml techniques are used to detect security attacks. SVM is utilized in this strategy to enhance the accuracy of attack detection. The NSL-KDD dataset is employed. The 41-feature set is separated into three categories: basic, content, and traffic. In this study [54], the author has proposed a novel approach called the outlier detection to detect intrusion in the network. The NSL-KDD data set was employed to validate the suggested technique. The suggested method takes less execution time and storage to test the dataset as compared to previous machine learning approaches used. The study [55] investigated the viability of merging fuzzy logic with machine learning techniques to detect intrusions. The suggested architecture mined fuzzy association rules using machine learning methods, extracting the best possible rules using a genetic algorithm.

In this research work [56] attack detection method for IDS is proposed. Outliers are removed from the data. When network flow shows abnormal behavior then this concept will help to detect these types of anomalies. The authors of [57] presented a novel concept for intrusion detection systems (IDS). The proposed study proved that if k-means clustering is applied, IDS accuracy improves in detecting attacks. This model performs best when given multiple clusters that correspond to the number of data types in the dataset. When the number of clusters changes, the performance of K- means degrades. In this proposed study [58] it has been elaborated that entropy can detect abnormal network behavior but with a high false rate. SVM model can classify traffic as normal or malicious traffic by learning different features of the network. The goal of this study is to overcome shortcomings of network entropy and support vector machines. So, the authors produced a hybrid solution that incorporates the advantages/ of both techniques. The dataset used in this proposed method is provided by MIT Lincoln laboratory. In this research work [59], authors have used k-means with naive bayes algorithm in IDS. This study shows that the k-means algorithm is not appropriate for anomaly detection because in some cases (especially in passive attacks, observatory attacks, etc.) intrusion behavior is almost the same as normal. If K- means algorithm is used with naive bayes, the detection rate increases with low false alarm. Authors have conducted experiments on Koyot 2006+ dataset. In this study [60] a detailed review of anomaly-based detection in which single, hybrid and ensemble machine learning models are used to evaluate different data sets. This comparison shows that the hybrid and ensemble machine learning approaches give higher accuracy and detection rate. In this study [61] evaluated the performance of J48, MLP, and bayes network classifiers for attack detection in IDS. According to the results, J48 showed the best results for detecting and classifying all attacks in the NSL-KDD dataset. In this study [62], implemented analysis on anomaly detection and presented a comparative review of seven machine learning model performances on Kyoto 2006+ dataset. Radial Bases Function (RDF) performed better under receiver

operating curve (ROC) among seven models. All remaining models outperformed 90% of the time in terms of precision, recall, and accuracy. In the proposed study, presented [63] a hybrid system that uses two detection systems i.e., misuse for signature or already existing types of intrusions and anomaly for new and updated intrusions. KDD Cup dataset used for the training of the system and about 30,000 files from window XP used to perform 2 weeks experiment. Using the NSLKDD dataset, this study [64-65] compared the performance of two supervised machine learning models, ANN and SVM. Four machine learning models are used to create an ensemble model. On two data sets, UNSW NB-15 and UGR'16, random forest, k- closest neighbor (KNN), SVM and logistic regression are applied on emulated and actual network traffic. A review [66] is presented in this research work for detection of attacks in intrusion detection system (IDS). In this proposed study [67], to detect intrusion attacks in a computer network four ML algorithms bayes net, J48, random forest and random tree applied on the KDD Cup dataset to analyze their performance. Random forest and random tree algorithms performed best on test datasets.

This paper investigates ML/DLNN models for intrusion detection systems [68-71]. Seven machine learning models namely Bayesian network, naive bayes classifier, decision tree, random forest, random tree, decision table and artificial neural network are explained, and their performance is tested using KDD cup data in terms of precision, recall, f1-score, and accuracy. Random forest gives the highest performance overall with 94% accuracy. In this paper [72], one-dimensional convolutional neural network-based deep learning method for creating an effective and flexible IDS (1DCNN) is presented. Normal and abnormal network traffic are classified and labelled for supervised learning in the 1D-CNN. Tested this proposed model using the UNSW NB15 IDS dataset to demonstrate the efficacy of approach. To compare performance, random forest (RF) and SVM models based on machine learning, as well as 1D-CNN with varying network parameters and architecture, are utilized. This study's key contribution [73] is the presentation of a HIDS that builds on the well-known consolidated tree construction (CTC) technique to effectively handle class-imbalanced data. At the detector's pre-processing step, a supervised relative random sampling (SRRS) technique was developed to get a balanced sample from a high-class imbalanced dataset. In addition, an advanced multi-class feature reduction approach was devised and built as a filter element to deliver the best standout features from IDS datasets for effective intrusion detection. With the help of cutting-edge IDS, it has been verified by using the NSL-KDD dataset and the CIC-IDS-2017 dataset. This research [74] creates an effective hybrid network-based IDS model (HNIDS) that uses the enhanced genetic algorithm and particle swarm optimization (EGA-PSO) and improved random forest (IRF) methodologies to address the data-imbalance issue. In the first phase, the proposed HNIDS employs hybrid EGA-PSO algorithms to optimize minor data samples and generate a balanced data set to understand the sample properties of small samples more precisely. A PSO approach is used in the suggested HNIDS to enhance the vector. The addition of a multi-objective function to GA improves it by helping to investigate the key features, improve fitness outcomes, reduce dimensions, increase true positive rate (TPR) and decrease

false positive rate (FPR). Using the benchmark datasets NSL-KDD, the performance of the suggested technique is compared with existing approaches. The experimental results show that the proposed HNIDS methodology outperforms different ML algorithms for the NSL-KDD dataset, including SVM, RF, LR, NB, LDA, and CART, with a BCC accuracy of 98.979% and an MCC accuracy of 88.149%. This study [75] provides a detailed analysis of the technologies, procedures, design, and risks posed by compromised internet of things (IoT) devices. Because the number of IoT devices is constantly expanding around the world. By the end of 2020, about 50 billion gadgets will most likely be connected to the Internet. Because of proliferation of IoT devices, the number of IoT-based cyber-attack instances has skyrocketed. So, it is necessary to create new approaches for detecting incidents launched from hacked IoT devices to address this challenge. The best detective control solution against assaults caused by IoT devices in this context uses machine and deep learning techniques. This work [76] described the design, implementation, and testing of a DL-based intrusion detection system based on FFDNNs. The best effective intrusion detection system has yet to be identified. The FFDNN models used in this work were connected to a FEU via IG to reduce input dimension while increasing classifier

accuracy. This study made use of the NSL-KDD dataset. For the binary and multiclass classification problems, the FFDNNs models outperformed SVM, RF, NB, DT, and KNN, both with a full and a FEU-reduced feature space. To detect online intrusions, proposed a unique threat intelligence detection model (TIDM) in this study [77]. The suggested TIDM is designed to handle large amounts of data live and as a result it can identify unknown connections, including zero-day assaults. The TIDM is made up of an optimized filter (Opti Filter), an adaptive and hybrid classifier, and an alert component. The Opti Filter component's key contributions come from its capacity to continually collect data flows and create unlabeled connection vectors. In the second portion of the TIDM, the enhanced growing hierarchical self-organizing map (EGHSOM) and the normal network behavior (NNB) models are integrated to find undiscovered links simultaneously. The suggested TIDM continuously updates the hybrid model in real-time. The Internet of Medical Things (IoMT) [78] is a subset of the Internet of Things (IoT) in which medical devices communicate confidential data. These advancements make it possible for the healthcare industry to communicate with and care for its patients more effectively. However, they have certain drawbacks because to the numerous security and privacy concerns, such as replay attacks, man-in-the-middle attacks, impersonation, privileged insider, remote hijacking, password guessing, DoS assaults, and malware attacks. When one of these attacks targets sensitive data, there is a potential that the attacker may gain access to allowed data or that the data will be altered, rendering it unavailable to authorized users and clients. Researchers in the fields of machine learning and data mining have developed a variety of supervised and unsupervised algorithms to determine the accurate detection of an abnormality. KDD99 or NSL-KDD 99 data sets are the basis for most relevant efforts on IDS [79]. These data sets are regarded as useless for identifying current attack types and have no relevance. In this study, the UNSW-NB15 data set is used as an offline dataset to develop a customized integrated classification-based model for identifying malicious network activity. In comparison to other current decision tree-based models, the suggested integrated classification-based model performs noticeably better at detecting five groups. Also, this study creates its own real-time data set at the NIT Patna CSE lab (RTNITP18), which serves as the suggested intrusion detection model's working example. This RTNITP18 dataset is used as a test set to gauge performance. The proposed model [80] initially performs data preprocessing in two ways: data conversion and data normalization. Also, the suitable selection of features is carried out using the improved fish swarm optimization-based feature selection (IFSO-FS) approach. By incorporating the Levy Flight (LF) concept into the traditional FSO algorithm's search mechanism to get beyond the local optima issue, the IFSO technique was developed.

Sharing threat events and indicators of compromise (IoCs) facilitates critical decision-making about timely and efficient defenses against cyberattacks [81]. Nonetheless, the present threat information sharing solutions make it difficult for threat detection systems (IDS) that use machine learning (ML) methodologies to share information and communicate with one another. As a solution to all of these issues, the platform for orchestrated Information Sharing and Awareness (ORISHA), which facilitates collaboration between threat detection systems and other information awareness components, is offered here. ORISHA is supported by a distributed threat intelligence platform built on an interconnected network of malware information sharing platform instances. This research suggests an efficient deep learning method [82] to increase classification accuracy and shorten training time. AE-IDS (Auto-Encoder Intrusion Detection System), which uses the random forest algorithm, is one such technique. Using feature grouping and feature selection, this technique creates the training set. Upon training, using an auto-encoder, the model can predict the outcomes, drastically cutting down on the amount of time it takes to find results. The experimental findings demonstrate that the suggested approach is preferable in terms of simple training, robust adaptability and high detection accuracy. This proposed study [83] improves IDS detection mechanisms through two processes: a deep neural network (DNN) model with new features for threat detection based on two assumptions related to dealing with zero-day attacks, with low computing power and resources, and a comprehensive detection solution that combines the DNN model and principal component analysis (PCA) to increase security and performance. According to analytical and software results, the suggested detection system, which integrates DNN,

PCA, statistical, and knowledge-based methodologies, surpasses existing IDS. In this study, IMIDS [84] was proposed as an intelligent intrusion detection system (IDS) to protect Internet of Things (IoT) devices. The heart of IMIDS is a lightweight convolutional neural network model that can classify a wide range of cyber threats. This article proposes an attack data generator driven by a conditional generative adversarial network to assist ease the problem of a shortage of training data. IMIDS beats its competitors in the testing, detecting nine different types of cyber-attacks (such as worms, shellcode, and backdoors) with an average F-measure of 97.22%. Furthermore, after being further trained using the information produced by attack data generator, IMIDS detection performance is noticeably enhanced. This work [85] presents the Intrusion Detection Tree ("IntruDTree") machine-learning-based security model, which first considers the ranking of security elements according to their value. This approach reduces computing complexity by reducing feature dimensions, making it beneficial in terms of prediction accuracy for previously unseen test scenarios. Finally, experiments were run on cybersecurity datasets to test the effectiveness of this model and the precision, recall, f score and ROC values were calculated. The MR-IMID is proposed in this article [86] to detect network intrusions using various data categorization jobs. The proposed MR-IMID reliably processes huge data sets utilizing easily available technology. To avoid future disparities, the MR-IMID in this proposed study detects intrusions by anticipating unanticipated test conditions and storing the results in a database. The proposed model outperforms previously published techniques, with detection accuracy of 97.7% and 95.7% throughout the training and validation phases, respectively.

3. Problem Statement

Due to the sophisticated nature of the attacker, the attack's surface is continuously changing. Today organizations are facing a challenging task to safeguard oneself against cyber-attacks. In any network, IDS serves a critical role in protecting against assaults. There is a need to create an automated IDS mechanism that uses machine learning approaches to protect against assaults more correctly and precisely. CTI plays a vital role in providing updated threat feeds to security devices.

4. Dataset for IDS Analysis

In this domain, several data sets are available for exploration. The data sets accessible for experimentation are as follows.

KDD Cup 99 Dataset
It was created in the fifth international conference on knowledge discovery and data mining. Creating a network intrusion detector—a prediction model that can distinguish between incursions and attacks.

NSL-KDD Dataset
Network Security Laboratory-KDD (NSL-KDD), The dataset [22] contains 41 features, the data contains KDDTrain+, KDDTest21+, and KDD Test+ which includes 125973,11,850 and 22,544 records. So, the original matrix of the dataset is of size 125,973x41,11,8500x41 and 22,544x41.Data. The data type is classified into two types of nominal and numerical data type.

Aegean Wi-Fi Intrusion Dataset (AWID) Dataset
It is the most widely used and open IDS dataset. However, AWID is distinguished by character data and an imbalance between attack data and regular data, which may have an impact on how well the intrusion detection system is rated (IDS).

Yahoo Web scope S5 Dataset
This anomaly benchmark consists of annotated anomalous points in real and artificial time-series. The dataset examines the precision with which different anomaly categories, such as outliers and change-points, may be detected.

Numenta Anomaly Benchmark (NAB) Dataset
This dataset is intended to test algorithms for detecting anomalies in streaming web applications. It includes more than 50 annotated real-world and synthetic time series data files, as well as a cutting-edge scoring system built for real-time applications. A scorecard of anomaly detection algorithms, thorough documentation, and entire open-source data and code are all available.

Kyoto 2006+ Dataset
It is based on actual network traffic data collected over a three-year period and classified as normal and attack traffic. In count to the 14 statistical features collected

from the KDD Cup '99 dataset, the Kyoto 2006+ dataset adds 10 more characteristics.

UNSW-NB 15 Dataset
The raw network packets of this dataset were generated by the Australian Centre for Cyber Security (ACCS) to produce a combination of genuine current normal activities and synthetic contemporary attack behaviors. Tcpdump is used to collect 100 GB of raw traffic (e.g., Pcap files). This dataset includes nine attack categories: Fuzzers, Analysis, Backdoors, DoS, Exploits, Generic, Reconnaissance, Shellcode, and Worms. The Argus and Bro-IDS tools, as well as twelve methods, are used to generate a total of 49 characteristics with the class label.

Bot-IoT Dataset
The UNSW Canberra Cyber Range Lab gathered this data by simulating a network environment. The traffic comprises of both regular and botnet traffic. The dataset's source files are available in a variety of formats, including the original Pcap files, created argus files, and csv files. The data were divided into attack categories and subcategories to enable tagging.

ISCX IDS 2012 Dataset
ISCX introduces a systematic method to produce the necessary datasets to meet this goal. A data set created in this regard known as ISCX IDS 2012. The fundamental concept is based on profiles, lower-level network elements, as well as accurate descriptions of intrusions.

CSE-CIC-IDS2018 Data set
We leverage the concept of profiles in this dataset to create datasets that include explicit descriptions of invasions as well as abstract distribution models for apps, protocols, or lower-level network components. It has amassed 16,000,000 occurrences in ten days. This is the latest publicly accessible big data intrusion detection dataset, and it encompasses a wide spectrum of attack strategies.

Data set	Year	No. of Features
KDD- Cup99 [22]	1998	41
NSL-KDD [22]	1999	41
AWID dataset [22]	2015	155
Yahoo Web scope s5 [22]	2015	4 Classes
NAB dataset [22]	2015	58 data streams
Kyoto 2006+ [22]	2006	24
UNSW NB-15 dataset [22]	2015	49
BoT*ıoTdataset* [22]	2019	46
ISCX IDS 2012 [22]	2012	14
CSE-CIC-IDS2018 [22]	2018	81

Table 1 Data sets Summary

5. Methodology

The suggested methodology examines two data sets: NSL KDD and CSE-CIC-IDS2018. These are the two most utilized data sets in IDS analysis of attack detection. There are three stages to the suggested methodology for analyzing the NSL KDD data set. In the first stage, data transformation technologies such as a label encoder are used. The second phase is feature reduction such as PCA, information gain and third phase are using classification techniques such as SVM, RF and DT.

Phase-1
NSL KDD data set consists of both numerical and nominal values, all values are converted to numerical in this phase. Use of a label encoder is made for this reason. It is employed since it is the method that is being used the most in the world. Converting values to a single value has the advantage of generating correct results because machine learning algorithms works good on single type of values.

Phase-2
The next step is the reduction of characteristics after the data have been transformed into numerical form. Several feature reduction methods, including genetic algorithms, linear discriminant analysis (LDA), principal component analysis (PCA), information gain and generalized discriminant analysis (GDA), are employed in the literature. The method of feature reduction that is currently used the most around the globe is PCA. PCA is used here because it is simple to calculate and provide accurate results. Computing systems find it relatively simple to solve problems. By lowering the dimensionality, it helps to improve the performance of machine learning algorithms. The benefit of using PCA is that it reduces data noise to some extent. Approaches such as the genetic algorithm have a high computational cost. Data that has substantial dimensions are difficult to visualize; therefore, PCA simplifies data visualization by decreasing the dimension. The feature set for the proposed study consists of forty-one features. After using PCA, the original forty-one feature set is reduced, and the fourteen best features are chosen. By

lowering the number of data set features, feature reduction approaches have the advantage of speeding up the system and requiring less processing power. The optimal 14-feature set extracted by the PCA is shown in table II.

Sr.#	Feature	Sr.#	Feature
1.	Protocol_type	8.	Srv_count
2.	Service	9.	Duration
3.	Src_bytes	10.	Dst_host_count
4.	Dst_bytes	11.	Wrong_frament
5.	Num_failed_logins	12.	Dst_host_srv_count
6.	Root_shell	13.	urgent
7.	Count	14.	Logged_in

Table II. Optimal Feature Set

Phase-3

The next phase is applying classification algorithm on the data extracted from phase 2 with fourteen features. For classification, the SVM, RF and DT are utilized. These methods produce strong results on various types of data sets.

Figure 2 displays a flow diagram. The NSL-KDD data set serves as the system's input. Using data transformation techniques, the data is reduced to a single numerical value. The features in the data set are then reduced using feature reduction techniques. To distinguish between legitimate and malicious traffic, classification algorithms are used after feature reduction procedures.

Figure 3 displays the proposed methodology for NSL-KDD data set. Three phases make up the proposed methodology. The data preprocessing phase is the first stage. Using data transformation techniques like label encoder, the data set is transformed into numerical values at this phase. Data transformation techniques are used to convert the data set to a single numerical value since machine learning algorithms perform best on single value data sets. The feature reduction is the second stage. During this phase, feature reduction techniques like PCA are used to minimize the feature set. Forty-one characteristics are condensed to fourteen in this phase. Computational power grows when more features are used in the dataset. Hence, feature reduction techniques are utilized to save computational resources. Using machine learning methods for classification is the third phase. Decision tree, random forest, and SVM algorithms are employed in this step to classify data. The training and testing data sets are split 80:20. Machine learning techniques classify the data as either an attack or legitimate/normal traffic.

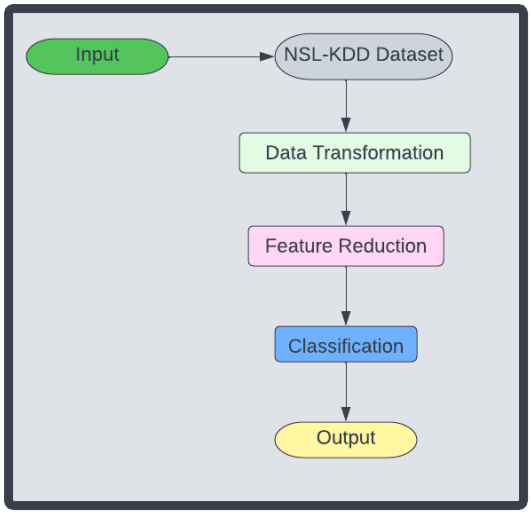

Fig. 2 Data Flow Diagram

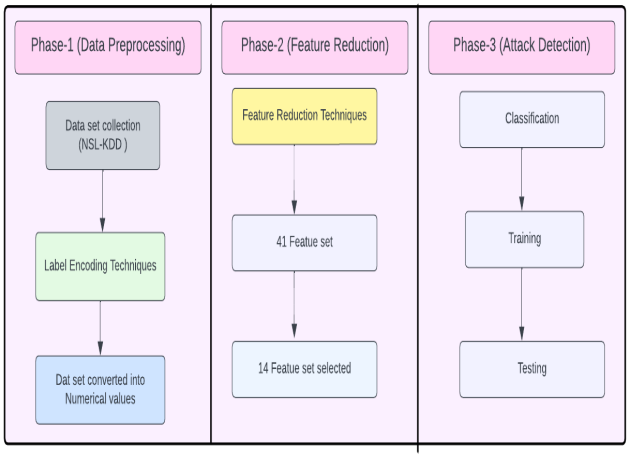

Fig. 3 Proposed Framework

The proposed methodology for analyzing the CSE-CIC-IDS2018 dataset is distributed into three stages. The initial stage is the normalization phase, which employs normalization techniques such as z-score and max normalization. The second phase involves feature reduction techniques like PCA, while the third involves classification methods such as SVM, RF, and DT.

Phase-1

The preliminary step is to standardize the data set.. Because the values in some of the data sets columns are quite high. Normalization techniques are used to balance the values in the data collection. The

advantage of using normalization techniques is that it equalizes all the values of the columns. For this purpose, z-score is used.

Phase-2

In the second phase normalized data set is used for feature reduction, as this data set consists of 81 features which requires more computational power and resources for utilization. For feature reduction, techniques such as PCA are used. 81 feature sets are reduced to 53 feature sets.

Phase-3

The next phase is applying classification algorithm on the data extracted from phase 2 with fourteen features. For classification, the SVM, RF, and DT are employed.

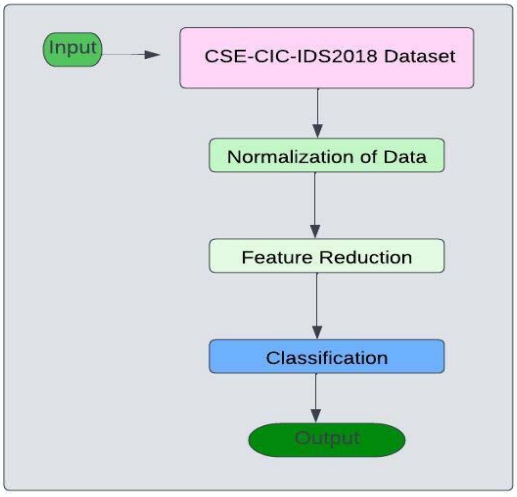

Fig. 4 Data Flow Diagram

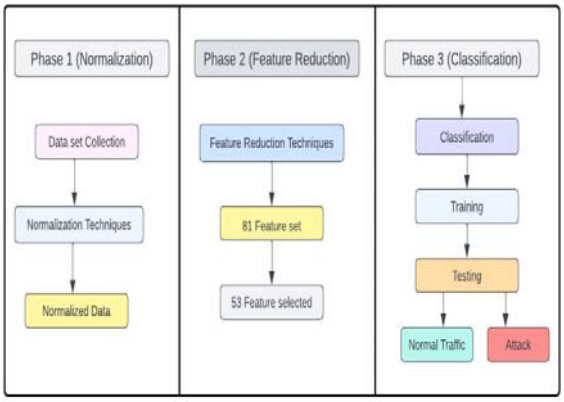

Fig. 5 Proposed Framework

Figure 4 displays a flow diagram. The CSE-CIC-IDS2018 data set serves as the system's input. Using normalization techniques, the entire set of data is normalized. After that, the data set's features are reduced using feature reduction techniques. To distinguish between legitimate and malicious traffic, classification algorithms are used after feature reduction procedures.

Figure 5 displays the proposed methodology for SE-CIC-IDS2018 data set. Three phases make up the suggested methodology. The normalization phase is the first stage. Using normalization techniques like z-score, the data set is normalized in this phase. Normalization is a common approach for preparing data for machine learning. Normalization is the process of converting numeric column values in a dataset to a standard scale while keeping information and not distorting the value ranges. The second stage is feature reduction. During this phase, feature reduction techniques like PCA are used to minimize the feature set. Eighty-one characteristics are condensed to fifty-three in this phase. A data set's computational power grows when more features are used. Hence, feature reduction techniques are utilized to save computational resources. Using machine learning methods for classification is the third phase. Decision tree, random forest, and SVM algorithms are employed in this step to classify data. The training and testing data sets are split 80:20. Machine learning techniques classify the data as either an attack or legitimate/normal traffic.

6. Results

Several performance evaluation metrics, including recall, accuracy, and precision, are employed for experimentation. To learn the specific outcomes, various performance measurements are used. The level of accuracy indicates a model's overall performance. Hence, relying solely on accuracy is a bad idea. Precision identifies the classifier's expected positive results out of all positive findings. Sensitivity is also known as recall. It is preferable to employ precision and recall together rather than separately because they are deficient performance indicators when used alone. Making use of the NSL-KDD data set, the proposed methodology has a 95% accuracy rate, which is higher than that of existing methods.

Using random forest, we achieve accuracy, precision, and recall of 96%, 94%, and 94%, respectively. SVM achieves 94%, 92%, and 92% accuracy, precision, and recall, respectively. The decision tree achieves 92%, 92%, and 91% accuracy, precision, and recall, respectively. The proposed methodology achieves an accuracy of 98% when using the CSE-CIC-IDS2018 data set, which is greater than that of existing methods. Using random forest, we get 98% accuracy, 97% precision, and 96% recall. SVM produces accuracy, precision, and recall of 94%, 95%, and 95%, respectively. The decision tree achieves 93%, 94%, and 94% accuracy, precision, and recall, respectively. Python is used to implement our findings. It uses a Core-I-7 processor and 16 GB of RAM. Table III displays the obtained outcomes. Fig. 6 & 7 shows the comparison of results with NSL-KDD and CSE-CIC-IDS2018 dataset.

Data Set	Classifier	Accuracy	Precision	Recall
NSL KDD	Random Forest	96%	94%	94%
	SVM	94%	92%	92%
	Decision Tree	92%	92%	91%
CSE-CIC-IDS2018	Random Forest	98%	97%	96%
	SVM	94%	95%	95%
	Decision Tree	93%	94%	94%

Table III. Results

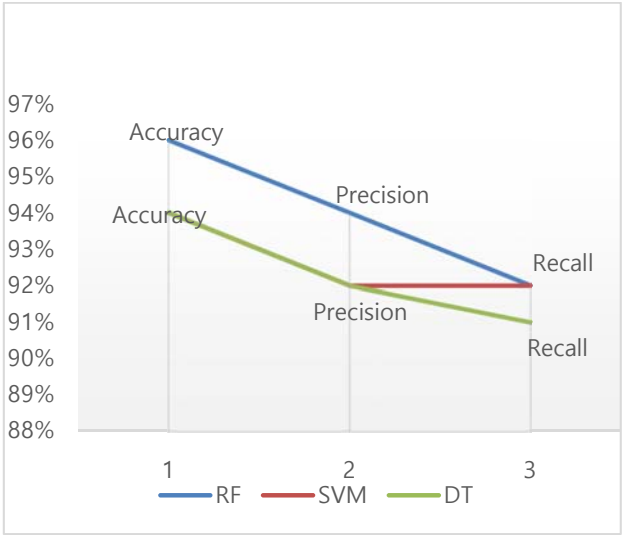

Fig. 6 Comparison of results with NSL-KDD Dataset

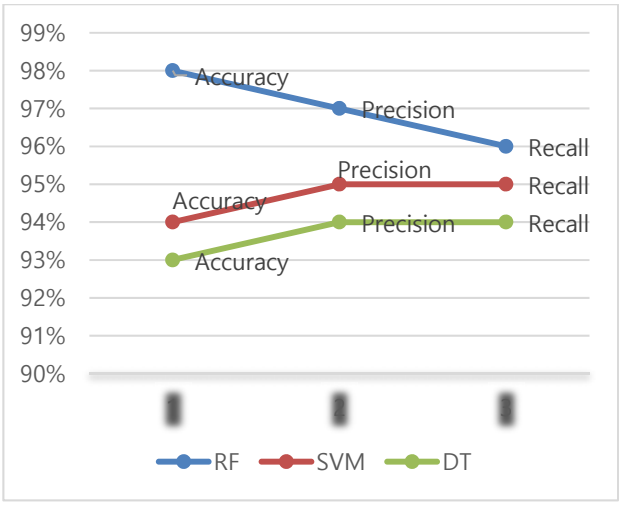

Fig. 7 Comparison of results with CSE-CIC-IDS2018 Dataset

7. Conclusion and Future Work

The cyber-crime rate is rapidly increasing which is a major disadvantage of technology. There are various attacks and ways through which attackers penetrate systems. To protect systems from such attackers, researchers developed numerous solutions based on machine learning algorithms, which are critical in detecting and protecting assets from diverse attacks. Using machine learning approaches, this research study offered a strategy for more precisely detecting attacks in IDS. In the proposed approach two mostly widely used data set NSL-KDD and CSE-CIC-IDS2018 are used for experimentation. With the NSL-KDD data set, this methodology achieves an overall accuracy of 96%, while with the CSE-CIC-IDS2018 data set, it achieves an accuracy of 98%. In IDS, this suggested approach identifies network assaults more correctly and precisely than previous approaches. Deep learning techniques will be employed in the future to improve classification outcomes.

References

[1] Conklin, Art and White, Gregory B, "E-government and cyber security: the role of cyber security exercises", Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS'06), IEEE, vol4, pp79b-79b,
Year 2006.

[2] Leuprecht, Christian and Skillicorn, David B and Tait, Victoria E, "Beyond the Castle Model of cyber-risk and cyber-security", Government Information Quarterly, volume33, pp 250-257, year 2016.

[3] Zwilling, Moti and Klien, Galit and Lesjak, Duan and Wiechetek, and Cetin, Fatih and Basim, Hamdullah Nejat, "Cyber security awareness,

knowledge and behavior: A comparative study", Journal of Computer Information Systems, volume 62, pp 82-97, year 2022.
[4] Rajasekharaiah, KM and Dule, Chhaya S and Sudarshan, E, "Cyber security challenges and its emerging trends on latest technologies", IOP Conference Series: Materials Science and Engineering, volume 981, pp 022062, year 2020.
[5] Tonge, Atul M and Kasture, Suraj S and Chaudhari, Surbhi R, "Cyber security: challenges for society-literature review", IOSR Journal of computer Engineering, volume 2, pp 67-75, 2013.
[6] Von Solms, Rossouw and Van Niekerk, Johan, "From information security to cyber security", computers & security, volume 38, pages 97-102, year 2013.
[7] McNeese, Michael and Cooke, Nancy J and D'Amico, Anita and Endsley, Mica R and Gonzalez, Cleotilde and Roth, Emilie and Salas, Eduardo, "Perspectives on the role of cognition in cyber security", Proceedings of the Human Factors and Ergonomics Society Annual Meeting, volume 56, pages 268-271, year 2012.
[8] Choo, Kim-Kwang Raymond, "The cyber threat landscape: Challenges and future research directions", Computers & security, volume 30, pp719-731, year 2011.
[9] Spence, Aaron and Bangay, Shaun, "Security beyond cybersecurity: side-channel attacks against non-cyber systems and their countermeasures", International Journal of Information Security, volume= 21, pp 437-453, 2022.
[10] Achar, Sandesh," Cloud Computing Security for Multi-Cloud Service Providers: Controls and Techniques in our Modern Threat Landscape", International Journal of Computer and Systems Engineering, volume=16, pages 379-384,2022.
[11] Rowe, Dale C and Lunt, Barry M and Ekstrom, Joseph J, "The role of cyber-security in information technology education", Proceedings of the 2011 conference on Information technology education, pp 113-122, 2011.
[12] Ukwandu, Elochukwu and Ben-Farah, Mohamed Amine and Hindy, Hanan and Bures, Miroslav and Atkinson, Robert and Tachtatzis, Christos and Andonovic, Ivan and Bellekens, Xavier, cyber-security challenges in aviation industry: A review of current and future trends, Information, MDPI, volume 13, pp 146, 2022.
[13] Mahmood, Samreen and Chadhar, Mehmood and Firmin, Selena, "Cybersecurity challenges in blockchain technology: A scoping review", Human Behavior and Emerging Technologies, Hindawi, volume 2022, 2022.
[14] Akpan, Frank and Bendiab, Gueltoum and Shiaeles, Stavros and Karamperidis, Stavros and Michaloliakos, Michalis, "Cybersecurity challenges in the maritime sector" Network, MDPI volume2, pp 123-138, 2022.
[15] Denning, Dorothy E, "An intrusion-detection model", IEEE Transactions on software engineering, pp 222-232, 1987.
[16] Roschke, Sebastian and Cheng, Feng and Meinel, Christoph, "Intrusion detection in the cloud", 2009 eighth IEEE international conference on dependable, autonomic and secure computing, IEEE, pp729-734,2009.
[17] Effendy, David Ahmad and Kusrini, Kusrini and Sudarmawan, Sudarmawan, "Classification of intrusion detection system (IDS) based on computer network.
", 2017 2nd International conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE), IEEE, pp 90-94, 2017.
[18] Uppal, Hussain Ahmad Madni and Javed, Memoona and Arshad, M, "An overview of intrusion detection system (IDS) along with its commonly used techniques and classifications", International Journal of Computer Science and Telecommunications, Citeseer, volume 5, pp 20-24, 2014.
[19] Ashoor, Asmaa Shaker and Gore, Sharad, "Importance of intrusion detection system (IDS)", International Journal of Scientific and Engineering Research, volume 2, pp 1-4,2011.
[20] Liao, Hung-Jen and Lin, Chun-Hung Richard and Lin, Ying-Chih and Tung, Kuang-Yuan, "Intrusion detection system: A comprehensive review", Journal of Network and Computer Applications, volume 36, pp 16-24, 2013.
[21] Wu, Yu-Sung and Foo, Bingrui and Mei, Yongguo and Bagchi, Saurabh, "Collaborative intrusion detection system (CIDS): a framework for accurate and efficient IDS", 19th Annual Computer Security Applications Conference, 2003. Proceedings, IEEE, pp 234-244, 2003.
[22] Khraisat, Ansam and Gondal, Iqbal and Vamplew, Peter and Kamruzzaman, Joarder, "Survey of intrusion detection systems: techniques, datasets and challenges", Cybersecurity, Springer, volume 2, pp 1-22,2019.
[23] Kr. gel, Christopher and Toth, Thomas and Kirda, Engin, "Service specific anomaly detection for network intrusion detection", Proceedings of the 2002 ACM symposium on Applied computing, pp 201-208, 2002.
[24] Hnamte, Vanlalruata and Hussain, Jamal, "An Extensive Survey on Intrusion Detection Systems: Datasets and Challenges for Modern Scenario", 2021 3rd International Conference on Electrical, Control and Instrumentation Engineering (ICECIE), IEEE, pp 1-10, 2021.
[25] Umer, Muhammad Fahad and Sher, Muhammad and Bi, Yaxin, "Flow-based intrusion detection: Techniques and challenges", Computers & Security, volume70, pp 238-254,2017.
[26] Hindy, Hanan and Brosset, David and Bayne, Ethan and Seeam, Amar and Tachtatzis, Christos and Atkinson, Robert and Bellekens, Xavier, "A taxonomy and survey of intrusion detection system design techniques, network threats and datasets", 2018.
[27] Azizjon, Meliboev and Jumabek, Alikhanov and Kim, Wooseong, "2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)}, IEEE, pp 218-224,2020.
[28] Panigrahi, Ranjit and Borah, Samarjeet and Bhoi, Akash Kumar and Ijaz, Muhammad Fazal and Pramanik, Moumita and Kumar, Yogesh and Jhaveri, Rutvij H, "Mathematics, MDPI, volume 9, pp 751, 2021.
[29] Balyan, Amit Kumar and Ahuja, Sachin and Lilhore, Umesh Kumar and Sharma, Sanjeev Kumar and Manoharan, Poongodi and Algarni, Abeer D and Elmannai, Hela and Raahemifar, Kaamran, "A hybrid intrusion detection model using ega-pso and improved random forest method", Sensors, MDPI, volume 22, pp 5986, 2022.
[30] Asharf, Javed and Moustafa, Nour and Khurshid, Hasnat and Debie, Essam and Haider, Waqas and Wahab, Abdul, "A review of intrusion detection systems using machine and deep learning in internet of things: Challenges, solutions and future directions", Electronics, MDPI, volume 9, pp 1177, 2020.
[31] Kasongo, Sydney Mambwe and Sun, Yanxia, "A deep learning method with filter-based feature engineering for wireless intrusion detection system", IEEE access, volume 7, pp 38597-38607, 2019.
[32] Salem, Maher and Al-Tamimi, Abdel-Karim, "A Novel Threat Intelligence Detection Model Using Neural Networks", IEEE Access, volume 10, pp 131229-131245, 2022.
[33] RM, Swarna Priya and Maddikunta, Praveen Kumar Reddy and Parimala, M and Koppu, Srinivas and Gadepalli, Thippa Reddy and Chowdhary, Chiranji Lal and Alazab, Mamoun, "An effective feature engineering for DNN using hybrid PCA-GWO for intrusion detection in IoMT architecture, Computer Communications,
Volume 160, pp 139-149, 2020.
[34] Kumar, Vikash and Sinha, Ditipriya and Das, Ayan Kumar and Pandey, Subhash Chandra and Goswami, Radha Tamal, "An integrated rule-based intrusion detection system: analysis on UNSW-NB15 data set and the real time online dataset", Cluster Computing, Springer, volume 23, pp 1397-1418, 2020.
[35] Alohali, Manal Abdullah and Al-Wesabi, Fahd N and Hilal, Anwer Mustafa and Goel, Shalini and Gupta, Deepak and Khanna, Ashish," Artificial intelligence enabled intrusion detection systems for cognitive cyber-physical systems in industry 4.0 environment", Cognitive Neurodynamic, Springer, volume 16, pp 1045-1057,2022.
[36] Guarascio, Massimo and Cassavia, Nunziato and Pisani, Francesco Sergio and Manco, Giuseppe, "Boosting cyber-threat intelligence via collaborative intrusion detection", Future Generation Computer Systems, volume 135, pp 30-43,2022.
[37] Li, XuKui and Chen, Wei and Zhang, Qianru and Wu, Lifa, "Building auto-encoder intrusion detection system based on random forest feature selection, Computers & Security, volume 95, pp 101851, 2020.

[38] Asif, Muhammad and Abbas, Sagheer and Khan, MA and Fatima, Areej and Khan, Muhammad Adnan and Lee, Sang-Woong, "MapReduce based intelligent model for intrusion detection using machine learning technique", Journal of King Saud University-Computer and Information Sciences, 2021.
[39] T. D. Wagner, K. Mahbub, E. Palomar, and A. E. Abdallah," Cyber threat intelligence sharing: Survey and research directions," Computers & Security, vol. 87, p. 101589, 2019.
[40] T. D. Wagner, E. Palomar, K. Mahbub, and A. E. Abdallah, "A novel trust taxonomy for shared cyber threat intelligence," Security and Communication Networks, vol. 2018, 2018.
[41] V. Mavroeidis and S. Bromander," Cyber threat intelligence model: an evaluation of taxonomies, sharing standards, and ontologies within cyber threat intelligence," in 2017 European Intelligence and Security Informatics Conference (EISIC). IEEE, 2017, pp. 91-98.
[42] M. Conti, T. Dargahi, and A. Dehghantanha," Cyber threat intelligence: challenges and opportunities," in Cyber Threat Intelligence. Springer, 2018, pp. 1-6.
[43] Gartner, "2021 Gartner," https://www.gartner.com, 2021.
[44] R. Brown and R. M. Lee, "The evolution of cyber threat intelligence (cti)": 2019 sans cti survey," SANS Institute: Singapore, 2019.
[45] Tounsi, Wiem and Rais, Helmi, "A survey on technical threat intelligence in the age of sophisticated cyber-attacks", Computers & security, volume 72, pp 212-233,2018.
[46] Ramsdale, Andrew and Shiaeles, Stavros and Kolokotronis, Nicholas, "A comparative analysis of cyber-threat intelligence sources, formats and languages", Electronics, volume 9, pp 824, 2020.
[47] Berndt, Anzel and Ophoff, Jacques, "Exploring the value of a cyber threat intelligence function in an organization", Information Security Education. Information Security in Action: 13th IFIP WG 11.8 World Conference, WISE 13, Maribor, Slovenia, September 21--23, 2020, Proceedings 13, Springer, pp 96-109, 2020.
[48] Zibak, Adam and Simpson, Andrew, "Cyber threat information sharing: Perceived benefits and barriers", Proceedings of the 14th international conference on availability, reliability and security, pp 1-9,2019.
[49] Samtani, Sagar and Abate, Maggie and Benjamin, Victor and Li, Weifeng, "Cybersecurity as an industry: A cyber threat intelligence perspective", The Palgrave Handbook of International Cybercrime and Cyberdeviance, Springer, pp 135-154,2020.
[50] Zibak, Adam and Sauerwein, Clemens and Simpson, Andrew, "A success model for cyber threat intelligence management platforms", Computers & Security, volume 111, pp 102466, 2021.
[51] Kevric, J., Jukic, S. Subasi, A. An effective combining classifier approach using tree algorithms for network intrusion detection. Neural Computing Applications 28, 1051–1058 (2017).
[52] Kabir, Md Reazul, Abdur Rahman Onik, and Tanvir Samad." A network intrusion detection framework based on Bayesian network using wrapper approach." International Journal of Computer Applications 166.4 (2017).
[53] Hagos, Desta Haileselassie, et al." Enhancing security attacks analysis using regularized machine learning techniques." 2017 IEEE 31st International Conference on Advanced Information Networking and Applications (AINA). IEEE, 2017
[54] Divya Goyal, Research Scholar Hardeep Singh, A.P. Dept. CSE at LPU, Jalandhar. Paper on Machine learning Techniques: Outlier Detection and Text summarization, International Journal of Scientific Engineering Research, Volume 5, Issue 3, March-2014 223
[55] IJCSNS International Journal: Intrusion Detection Using Machine learning along Fuzzy Logic and Genetic Algorithms, Y. Dhanalakshmi and Dr.I. Ramesh Babu, Dept of Computer Science Engineering Acharya Nagarjuna University, Guntur, A.P. India.
[56] Chitrakar, Roshan, and Chuanhe Huang." Anomaly based intrusion detection using hybrid learning approach of combining k-medoids clustering and naive bayes classification." 2012 8th International Conference on Wireless Communications, Networking and Mobile Computing. IEEE, 2012
[57] Duque, Solane, and Mohd Nizam bin Omar." Using data mining algorithms for developing a model for intrusion detection system (IDS)." Procedia Computer Science 61 (2015): 46-51.
[58] Agarwal, Basant, and Namita Mittal." Hybrid approach for detection of anomaly network traffic using data mining techniques." Procedia Technology 6 (2012): 996-1003
[59] Muda, Z. Mohamed, Warusia Sulaiman, md nasir Udzir, Nur. (2016). K-Means Clustering and Naive Bayes Classification for Intrusion Detection. Journal of IT in Asia. 4. 13-25. 10.33736/jita.45.2014.
[60] U. S. Musa, M. Chhabra, A. Ali and M. Kaur," Intrusion Detection System using Machine Learning Techniques: A Review," 2020 International Conference on Smart Electronics and Communication (ICOSEC), 2020, pp. 149-155, doi: 10.1109/ICOSEC49089.2020.9215333.
[61] Alkasassbeh and Almseidin. (2018). Machine Learning Methods for Network Intrusions. International Confrernce on Computing, Communication (ICCCNT). Arxiv.
[62] Marzia Z. and Chung-Horng L. (2018). Evaluation of Machine Learning Techniques for Network Intrusion Detection. IEEE. (pp. 1-5)
[63] Dutt t I. et al. (2018). Real Time Hybrid Intrusion Detection System. International Conference on Communication, Devices and Networking (ICCDN). (pp. 885-894). Springer.
[64] Kazi A., Billal M. and Mahbubur R. (2019). Network Intrusion Detect ion using Supervised Machine Learning Technique with feature selection. International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST). (pp. 643-646). IEEE.
[65] Rajagopal S., Poornima P. K. and Kat iganere S. H. (2020). A Stacking Ensemble for Network Intrusion Detect ion using Heterogeneous Datasets. Journal of Security and Communication Networks. Hindawi.
[66] S. Thapa and A.D Mailewa (2020). The Role of Intrusion Detection/Prevention Systems in Modern Computer Networks: A Review. Conference: Midwest Instruction and Computing Symposium (MICS). Wisconsin, USA. Volume: 53. (pp. 1-14).
[67] Chibuzor John Ugochukwu, E. O Bennett.An Intrusion Detection System Using Machine Learning Algorithm Department of Computer Science, International Journal of Computer Science and Mathematical Theory ISSN 2545-5699 Vol. 4 No.1 2018.
[68] Alqahtani H., Sarker I.H., Kalim A., Minhaz Hossain S.M., Ikhlaq S., Hossain S. (2020) Cyber Intrusion Detection Using Machine Learning Classification Techniques. In: Chaubey N., Parikh S., Amin K. (eds) Computing Science, Communication and Security. COMS2 2020. Communications in Computer and Information Science, vol 1235. Springer, Singapore. https://doi.org/10.1007/978-981-15-6648-6 10.
[69] Xin, Y., et al.: Machine learning and deep learning methods for cybersecurity. IEEE Access 6, 35365–35381 (2018).
[70] Ferrag, Maglaras, Moschoyiannis, Janicke (2019). Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study, Journal of Information Security and Applications.
[71] Singh, Geeta and Khare, Neelu, A survey of intrusion detection from the perspective of intrusion datasets and machine learning techniques, International Journal of Computers and Applications, 2021.
[72] Azizjon, Meliboev and Jumabek, Alikhanov and Kim, Wooseong, "1D CNN based network intrusion detection with normalization on imbalanced data", 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), IEEE, pp 218-224,2020.
[73] Panigrahi, Ranjit and Borah, Samarjeet and Bhoi, Akash Kumar and Ijaz, Muhammad Fazal and Pramanik, Moumita and Kumar, Yogesh and Jhaveri, Rutvij H, "A consolidated decision tree-based intrusion detection system for binary and multiclass imbalanced datasets", Mathematics, MDPI, volume 9, pp 751,2021.
[74] Balyan, Amit Kumar and Ahuja, Sachin and Lilhore, Umesh Kumar and Sharma, Sanjeev Kumar and Manoharan, Poongodi and Algarni, Abeer D and Elmannai, Hela and Raahemifar, Kaamran, "A hybrid intrusion detection model using ega-pso and improved random forest method", Sensors, MDPI, volume=22,

Pp 5986,2022.
[75] Asharf, Javed and Moustafa, Nour and Khurshid, Hasnat and Debie, Essam and Haider, Waqas and Wahab, Abdul, "A review of intrusion detection systems using machine and deep learning in internet of things: Challenges, solutions and future directions", Electronics, MDPI, volume 9, pp 1177, 2020.
[76] Kasongo, Sydney Mambwe and Sun, Yanxia, "A deep l earning method with filter-based feature engineering for wireless intrusion detection system", IEEE access, volume 7, pp 38597-38607,2019.
[77] Salem, Maher and Al-Tamimi, Abdel-Karim," A Novel Threat Intelligence Detection Model Using Neural Networks", IEEE Access, volume10, pp 131229-131245, 2022.
[78] RM, Swarna Priya and Maddikunta, Praveen Kumar Reddy and Parimala, M and Koppu, Srinivas and Gadekallu, Thippa Reddy and Chowdhary, Chiranji Lal and Alazab, Mamoun, "An effective feature engineering for DNN using hybrid PCA-GWO for intrusion detection in IoMT architecture", Computer Communications, volume 160, pp 139-149, 2020.
[79] Kumar, Vikash and Sinha, Ditipriya and Das, Ayan Kumar and Pandey, Subhash Chandra and Goswami, Radha Tamal, "An integrated rule-based intrusion detection system: analysis on UNSW-NB15 data set and the real time online dataset", Cluster Computing, Springer, volume 23, pp 1397-1418, 2020.
[80] Alohali, Manal Abdullah and Al-Wesabi, Fahd N and Hilal, Anwer Mustafa and Goel, Shalini and Gupta, Deepak and Khanna, Ashish, "Artificial intelligence enabled intrusion detection systems for cognitive cyber-physical systems in industry 4.0 environment", Cognitive Neurodynamic, volume 16, pp 1045-1057,2022.
[81] Guarascio, Massimo and Cassavia, Nunziato and Pisani, Francesco Sergio and Manco, Giuseppe, "Boosting cyber-threat intelligence via collaborative intrusion detection", Future Generation Computer Systems, volume 135, pp 30-43,2022.
[82] Li, XuKui and Chen, Wei and Zhang, Qianru and Wu, Lifa, "Building auto-encoder intrusion detection system based on random forest feature selection.", Computers & Security, volume 95, pp 101851, 2020.
[83] Al-Fawa'reh, Mohammad and Al-Fayoumi, Mustafa and Nashwan, Shadi and Fraihat, Salam, "Cyber threat intelligence using PCA-DNN model to detect abnormal network behavior", Egyptian Informatics Journal, volume 23, pp 173-185,2022.
[84] Le, Kim-Hung and Nguyen, Minh-Huy and Tran, Trong-Dat and Tran, Ngoc-Duan, "IMIDS: An intelligent intrusion detection system against cyber threats in IoT", Electronics, MDPI, volume 11, pp 524,2022.
[85] Sarker, Iqbal H and Abushark, Yoosef B and Alsolami, Fawaz and Khan, Asif Irshad, "Intrudtree: a machine learning based cyber security intrusion detection model", Symmetry, MDPI, volume 12, pp 754, 2020.
[86] Asif, Muhammad and Abbas, Sagheer and Khan, MA and Fatima, Areej and Khan, Muhammad Adnan and Lee, Sang-Woong, "Journal of King Saud University-Computer and Information Sciences", 2021.