DOI: 10.51153/kjcis.v7i1.198

Accurate Attack Detection in Intrusion Detection System for cyber Threat Intelligence Feeds using Machine Learning Techniques

Ehtsham Irshad[a*], Abdul Basit Siddiqui[a]

[a] Department of Computer Science, Capital University of Science and Technology, Islamabad, Pakistan
ehtsham_irshad@hotmail.com, abasit.siddiqui@cust.edu.pk

*Corresponding Author: Ehtsham Irshad ehtsham_irshad@hotmail.com

Abstract

With the advancement of modern technology, cyber-attacks are continuously rising. Malicious behavior in the network is discovered using security devices like intrusion detection systems (IDS), firewalls, and antimalware systems. To defend organizations, procedures for detecting threats more correctly and precisely must be defined. The proposed study investigates the significance of cyber-threat intelligence (CTI) feeds in accurate IDS detection. The NSL-KDD and CSE-CICIDS-2018 datasets were analyzed in this study.This research makes use of normalization, transformation, and feature selection algorithms. Machine learning (ML) techniques were employed to determine if the traffic was normal or an attack. With the proposed study the ability to identify network attacks has improved using machine learning algorithms. The proposed model provides 98% accuracy, 97% precision, and 96% recall respectively.

Keywords: Cyber-threat intelligence (CTI), Denial of service (DoS), Intrusion Detection System (IDS), Intrusion Prevention System (IPS), Indicators of compromise (IoC), Network Intrusion Detection System (NIDS), Artificial Neural Network (ANN).

1. Introduction

Cyber-security is increasingly an essential component of today's modern world. As networks and systems are increasing very rapidly, protecting data from attacks is an important aspect of today's research. In recent times, protection from various cyber-attacks has become a challenging issue [1-7]. The current system, which comprises firewalls, data encryption techniques, and user authentication procedures, is insufficient to address the threats posed by modern sophisticated attackers. However, these security devices are unable to protect networks against cyber-attacks [8-10]. Artificial intelligence is playing an increasingly important role in this field, and it is now widely used in all industries [11-14].

Firewalls and Intrusion Detection/Prevention Systems (IDS/IPS) are examples of network security devices. The common occurrence of false positive alarms as well as failure to identify zero-day attacks, which destroy businesses are just a few of the issues that are known to plague existing IDS. Companies lose time in the investigative process due to the flaws in IDS backend engines. Deep packet inspection is conducted 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]. IDS are of two types as shown in Figure 1. 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. 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]. Anomaly detection is the process of recognizing patterns in data that do not have predefined usual behavior [21].

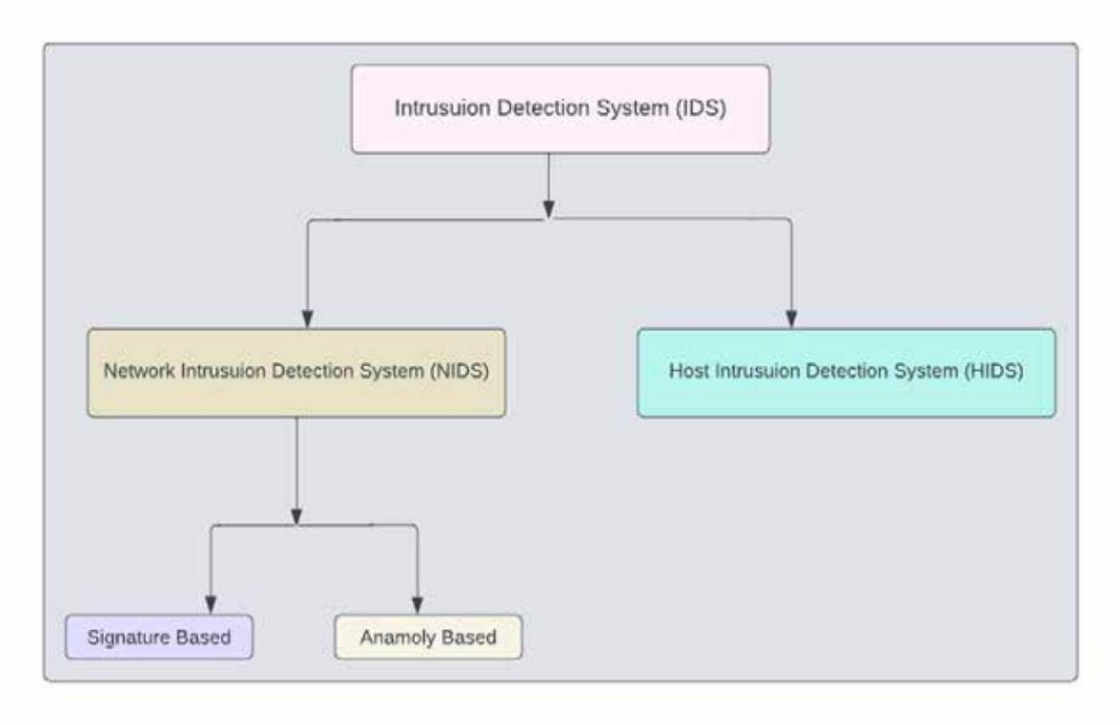

Figure 1. Types of IDS

To classify data into several categories modern-day machine learning techniques are used [22]. Normally traffic is classified into two types: normal and attack traffic. But according to some studies, there are five categories, attacks are further divided into four types: from remote to local, from user to root, probing assault, and DoS [23–26].

1.1 Role of Machine Learning in IDS

To protect against cyber-attacks, precise defense methods such as ML-based intrusion detection systems are required for better protection. They are being used as potential methods for detecting network attackers [27–30]. There is a need to categorize network attacks, despite major research efforts, IDS still struggles to improve detection accuracy [31-34]. ML/deep learning methods can be used in three ways: individually, hybrid, and ensemble-based. The performance of the machine learning technique was evaluated using several datasets. The most popular dataset for measuring performance is the NSL KDD Dataset [35–38].

1.2 Role of CTI in IDS

CTI provides multi-source databases that aid cyber defense mechanisms, allowing for comprehensive monitoring, identification, and response to online threats [39-41]. CTI feeds reveal how an attack occurred and who is behind it. These data feeds will be used to develop significant defensive security methods. The CTI feeds cover all major threat vectors, including websites, social media, bot IP addresses, malicious URLs, phishing URLs, spam, and harmful URLs. It enables organizations to decide how to respond to impending threats [42-44].

CTI generates threat feeds using both internal and external sources. Internal feeds include data from corporate security systems such as IDS/IPS, firewalls, and antivirus software. An external source could be a threat feed from a public (untrustworthy) source, such as an anti-malware domain, or a paid private source from several well-known and reputable security vendors.

Threat feeds are useful for organizations to protect against future attacks. Organizations are incorporating these threat feeds into their devices. Due to the sophisticated nature of the attacker, the attack's surface is continuously changing. There is a need to create an automated IDS mechanism that uses machine learning approaches to protect against assaults more accurately and precisely. CTI plays a vital role in providing updated threat feeds to security devices. [45-50].

This research investigation has made the following contributions:

a. Research has concentrated on finding significant traits and obtaining useful information from datasets. Feature selection strategies are used which help to reduce dimensionality.
b. The study work also contributes to normalizing the dataset. Some values in the dataset can affect the findings. So, normalization is performed on datasets.
c. This investigation uses machine learning techniques for improved attack detection in IDS. This implementation has increased the performance metrics as compared to other techniques.

The following is the structure of the research paper. Part 2 includes a literature review, and section 3 describes the problem statement. Section 4 describes the datasets used for IDS analysis. The proposed methodology is outlined in Section 5. Part 6 contains the results, whereas Section 7 contains the conclusion and future work.

2. Literature Review

According to this study [51], a classifier strategy for NIDS employing the tree algorithm is used. The author proposes a combined tree classifier strategy for identifying network assaults. The author presents an IDS framework [52]. The author used a Bayesian classifier to find abnormalities in the network. The NSL-KDD dataset is used as a standard in this domain. According to this study [53], ML algorithms are utilized to detect security threats. This approach makes use of a support vector machine (SVM) to improve attack detection accuracy. The author presented a novel approach termed outlier detection to detect network intrusion [54]. The NSL-KDD dataset was used to validate the proposed approach.

This study [55] looked at the feasibility of combining fuzzy logic with machine learning approaches to detect intrusions. This research study [56] proposes an attack detection mechanism for IDS. When network flow exhibits anomalous behavior, this idea can help detect problems. The authors [57] introduced a novel paradigm for intrusion detection systems. The suggested study shows that using K-means clustering enhances IDS accuracy in identifying attacks. According to this study [58], entropy can detect anomalous network behavior, although at a high false rate. This study addresses the limitations of network entropy.

In this paper [59], the authors employed K-means and a naïve Bayes algorithm for IDS. When the K-means algorithm is used with naive Bayes, the detection rate increases while producing fewer false alarms. The authors conducted experiments with the Koyot 2006+ dataset. This research [60] provides a comprehensive review of anomaly-based detection, which uses single, hybrid, and ensemble machine learning models to assess distinct datasets. J48 and MLP classifier's performance was evaluated for attack detection in IDS [61]. According to the results, J48 performed the best in detecting and categorizing all assaults in the NSL-KDD dataset. This study [62] carried out anomaly detection analysis and provided a comparative review of seven machine learning model performances on the Kyoto 2006+ dataset.

The authors presented [63] a hybrid system that employs two detection systems: abuse for signature or previously known forms of intrusions and anomaly for new and updated intrusions. Using the NSLKDD dataset, this study [64-65] assessed the performance of two supervised ML models, ANN and SVM. In this proposed study [66], a review is conducted for detecting attacks in IDS. According to this study [67], to detect intrusion threats in a computer network, four ML algorithms are evaluated on the KDD Cup dataset. Random forest and random tree algorithms performed the best on test datasets. This work looks at ML/DLNN models for intrusion detection systems [68–71]. Table 1 presents the outcomes of many previous methodologies.

Table 1. Results of Various Techniques

Author/ Year	Dataset	Technique	Results
A. Alzahrani et al. /2021	NSL-KDD	XGBoost	Precision 92% Recall 89% F1-Score 90%
V. Pai et al. /2021		Random Forest	Accuracy 91% Precision 92% Recall 90% F1-score 92%
A. Halimaa et al. /2019		Support Vector Machine	Accuracy 93%
K. Abu et al/2019	CSE-CICIDS-2018	ANN	Accuracy 91%
M. Fawa'reh et al. 2022		DNN-PCA	Accuracy 96%
J. Kim et. al. / 2019		CNN	Accuracy 95%
V. Kanimozhi et al. /2019		ANN, RF, KNN, SVM, Adaboost, NB	Accuracy 96%, Precision 90% Recall 95% F1- Score 90%
M. Amine et al. / 2019		DNN, RNN, CNN	Accuracy 93%

3. Problem Statement

With the intelligence and diversity of cyber threats increasing, traditional IDS are having problems detecting and mitigating attacks. The sheer volume and complexity of network traffic make it difficult for rule-based IDS to keep up with emerging threats. As a result, there is an urgent need to increase IDS capabilities by employing ML techniques for more precise threat detection. Because of the attack's sophistication and increasing volume, it is difficult to detect them in real time. There

is a need to improve strategies for detecting attacks more precisely to make more accurate decisions concerning the detection of hostile activities.

4. Datasets for IDS Analysis

The KDD Cup 99 dataset was created at the fifth international conference on knowledge discovery and data mining. It was developed at the Network Security Laboratory-KDD (NSL-KDD). It contains forty-one features [22]. The data includes records from KDDTrain+, KDDTest21+, and KDD Test+, totaling 125,973, 11,850, and 22,544. The Aegean Wi-Fi Intrusion Dataset (AWID) is the most widely used and publicly available IDS dataset. AWID is detected by character data as well as an imbalance between attack and regular data.

The Yahoo Web Scope S5 includes labeled anomalous events from both genuine and bogus time series. It tests how well various anomaly types, such as outliers and change points can be identified. The Numenta Anomaly Benchmark (NAB) dataset evaluates approaches for detecting anomalies in streaming web applications. It includes more than 50 annotated real-world and synthetic time series data files. The Kyoto 2006+ dataset is based on actual network traffic data collected over three years and classified as normal or attack traffic.

The UNSW-NB 15 dataset was generated by the Australian Centre for Cyber Security (ACCS) and mixes genuine current normal activities with synthetic contemporary attack behaviors. The UNSW Canberra Cyber Range Lab generated the Bot-IoT dataset by simulating a network. The traffic consists of both standard and botnet traffic. The ISCX IDS 2012 dataset was produced in 2012. The essential notion is built on profiles, lower-level network parts, and precise intrusion descriptions. The CSE-CIC-IDS2018 dataset pioneered the concept of a profile. It has amassed 16,000,000 occurrences in ten days. This is the most recent public big data intrusion detection dataset, and it includes a wide spectrum of attack strategies.

5. Proposed Methodology

The proposed methodology compares two datasets: NSL KDD and CSE-CIC-IDS2018. These are the two most utilized datasets in IDS analysis to detect attacks.

5.1 Methodology for NSL-KDD Dataset

The suggested approach for evaluating the NSL-KDD dataset is divided into three main parts. In the first stage, data transformation techniques are used. The second phase entails decreasing features. The third phase uses classification techniques like support vector machines, random forest, and decision trees to detect risks.

Figure 2 displays the methodology for the NSL-KDD dataset. 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 to a single numerical value since machine learning algorithms perform best on single-value datasets. Feature reduction is the second stage. During this phase, feature reduction techniques like PCA are used to minimize the feature set. Forty-one features are reduced 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. The third phase involves using ML methods to classify data. This step uses a decision tree, random forest, and SVM algorithms to classify data. The training and testing

data sets are split 80:20. Machine learning techniques classify data as assaults or normal traffic.

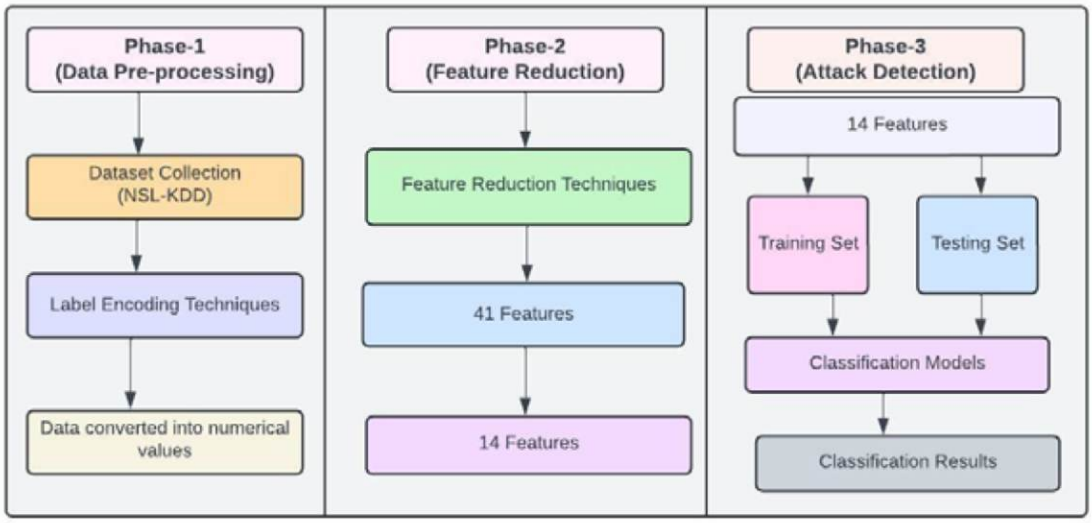

Figure2. Proposed Methodology (NSL-KDD Dataset)

Phase-1

This dataset contains numerical and nominal values. During this phase, all values are converted into numerical representation. This transformation is carried out utilizing a label encoder. It is employed because it is the most often used method. Converting values to a single value has the advantage of delivering correct results, as machine learning algorithms work best with specific types of values.

Phase-2

The subsequent stage is the feature reduction phase. The purpose of feature reduction is to reduce the dataset's forty-one features, which require greater processing resources to compute their values. The literature employs a range of feature reduction techniques, including genetic algorithms, linear discriminant analysis (LDA), principal component analysis (PCA), information gain, and generalized discriminant analysis. PCA is the most popular feature reduction method in the world today. PCA is used here because it is simple to calculate and yields consistent results. Computing systems find it simple to solve problems. Reduced dimensionality enhances the performance of machine learning algorithms. One benefit of using PCA is that it reduces data noise.
Approaches such as the genetic algorithm are computationally complex. Data with many different dimensions is difficult to represent; so, PCA makes visualization of data easier by reducing the dimensions. The proposed study's feature set consists of 41 features. PCA compresses the original forty-one feature set to fourteen main features. A threshold is established, and values more than 0.60 are classified as a feature. In this regard, fourteen feature sets have been selected. By reducing the number of data set features, feature reduction algorithms improve system performance and require less processing power. Table 2 shows the best 14-feature set recovered by the PCA.

Table 2. Optimal Feature Set

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

Phase-3

The following stage is to apply classification algorithms on the data extracted from phase 2 with fourteen features. Classification techniques include SVM, RF, and DT.
Figure 3 depicts a flow diagram. The system accepts the NSL-KDD data set as input. Data transformation techniques are used to convert data into a single numerical value. The features in the data set are then reduced using feature reduction techniques. Following feature reduction methods, classification algorithms are used to distinguish between legitimate and malicious traffic.

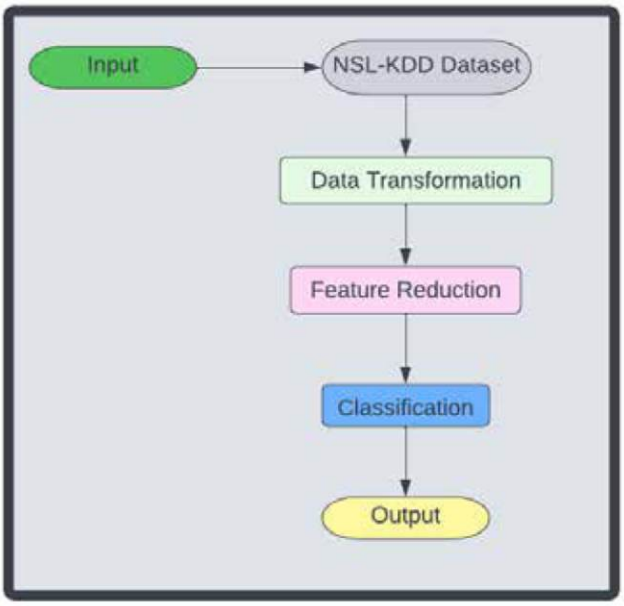

Figure 3. Flow Diagram

5.2 Proposed Methodology for CSE-CIC-IDS2018 Dataset

The proposed technique for analyzing the CSE-CIC-IDS2018 dataset is separated into three stages. The first stage is normalization, which employs methods such as z-score and min-max normalization. The second phase involves feature reduction techniques such as PCA, whereas the third employs classification methods such as SVM, RF, and DT.
Figure 4 depicts the optimal approach for the CIC-IDS2018 dataset. The proposed methodology is made up of three steps. The normalization phase is the first phase. This phase entails normalizing the dataset with techniques like z-score. Normalization is a common method for getting data ready

for machine learning. Normalization is the process of converting numeric column values in a dataset to a standard scale while retaining information and preventing value range distortion. The second step is all about reduction. The feature set is reduced at this step using feature reduction techniques such as PCA.

This phase reduced the number of features from 81 to 53. The processing capability of a data set grows as more features are added. Thus, feature reduction techniques are employed to save computational resources. In the third phase, data is classified using machine learning approaches. In this step, data is classified using decision trees, random forests, and SVM algorithms. The training and testing data sets are split 80:20. Machine learning techniques classify the data as either an attack or normal traffic.

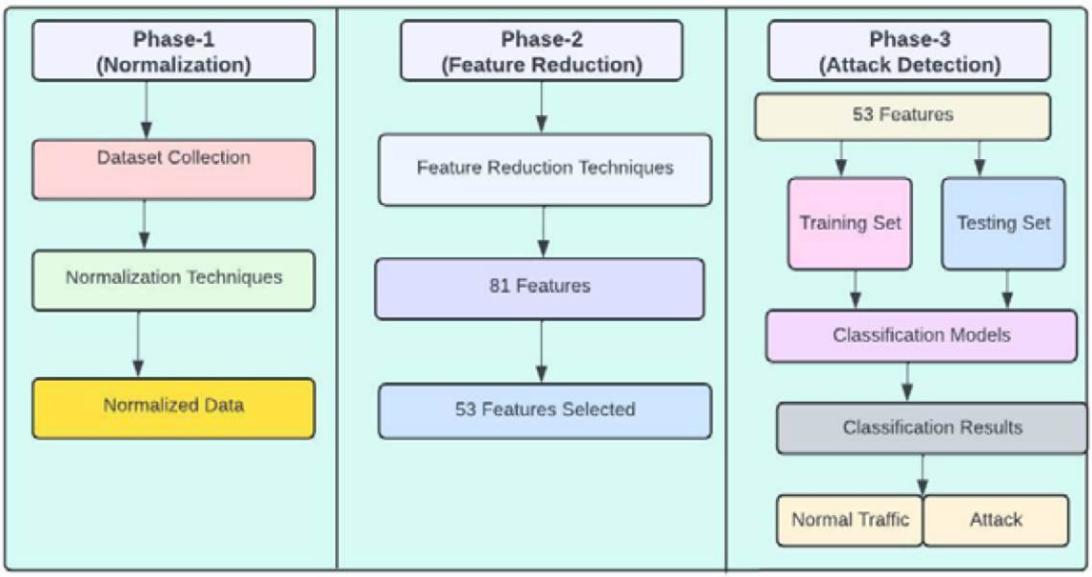

Figure 4. Proposed Methodology (CSE-CIC-IDS2018 Dataset)

Phase-1

The first step is to standardize the data. Because the values in certain columns are high. Normalization procedures are employed to balance the values in the data. The advantage of utilizing normalization procedures is that they equalize all the column values. For this reason, the z-score is used.

Phase-2

In the second phase normalized dataset is used for feature reduction as it consists of eighty-one features that require more computational power and resources for utilization. For feature reduction, a technique like PCA is used. The threshold is set at 0.60. Values higher than this threshold are selected. Eighty feature sets are reduced to fifty-three feature sets.

Phase-3

The following stage is to apply a classification algorithm to the phase 2 data, which contains fourteen features. SVM, RF, and DT are the classification algorithms employed.

Figure 5 depicts a flow diagram. The CSE-CIC-IDS2018 dataset is utilized as the system input. Normalization processes are used to make the data more consistent. The dataset features are then reduced using feature reduction techniques. Following feature reduction methods, classification

algorithms are used to differentiate between legitimate and malicious traffic.

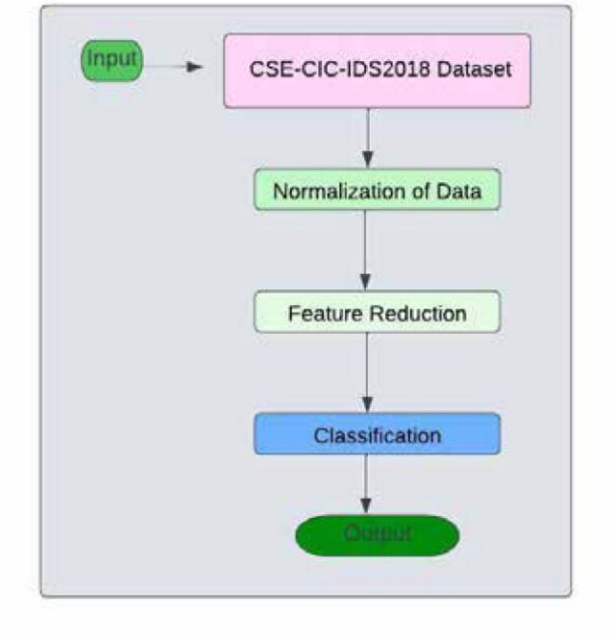

Figure 5. Flow Diagram

6. Results

Several performance evaluation criteria, including Recall, Accuracy, and Precision, are used for experimentation. Accuracy measures a model's overall performance. Relying primarily on accuracy is not an original concept. Precision determines the classifier's expected positive results among all positive discoveries. Sensitivity is sometimes referred to as Recall. Precision and Recall are better employed together than separately because they are ineffective performance measures when used alone. The confusion matrix for the NSL KDD dataset is depicted in Figure 6. This matrix shows both the expected and actual values. The model's anticipated true values compared to predicted false values.

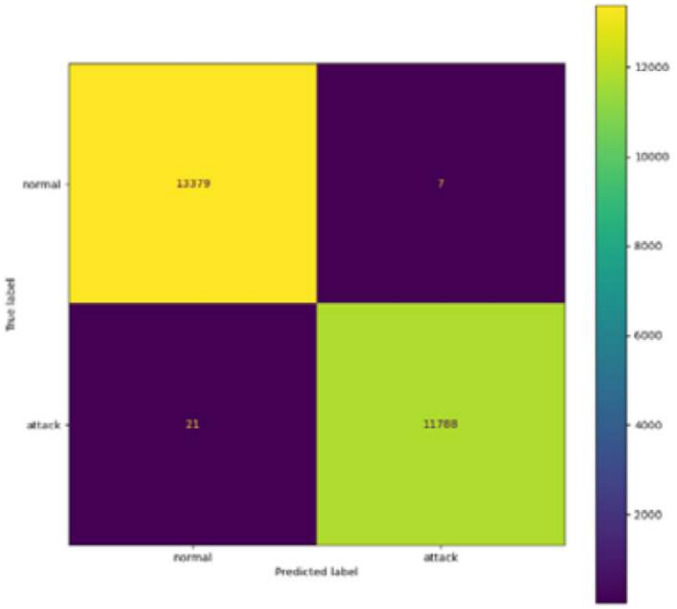

Figure 6. Confusion Matrix (NSL-KDD Dataset)

Using the NSL-KDD dataset, the proposed algorithm achieves 95% accuracy, which outperforms earlier methods. Random forest yields 96% accuracy, 94% precision, and 94% recall. SVM delivers 94% accuracy, 92% precision, and 92% recall rates. The decision tree achieves 92% accuracy, 92% precision, and 91% recall.

The experiment employs cross-validation with a value of k=10. The training and testing datasets have an 80:20 ratio. Using the CSE-CIC-IDS2018 dataset, the proposed methodology surpasses earlier methods, with an accuracy of 98%. Using random forest, we get 98% accuracy, 97% precision, and 96% recall. SVM yields 94% accuracy, 95% precision, and 95% recall, respectively. The decision tree has 93% accuracy, 94% precision, and 94% recall, respectively. The implementation language is Python. Anaconda is used to create an integrated development environment. The implementation testbed is powered by a Core-I-7 CPU with 16 GB of RAM. Figures 7 and 8 compare the results with the datasets NSL-KDD and CSE-CIC-IDS2018.

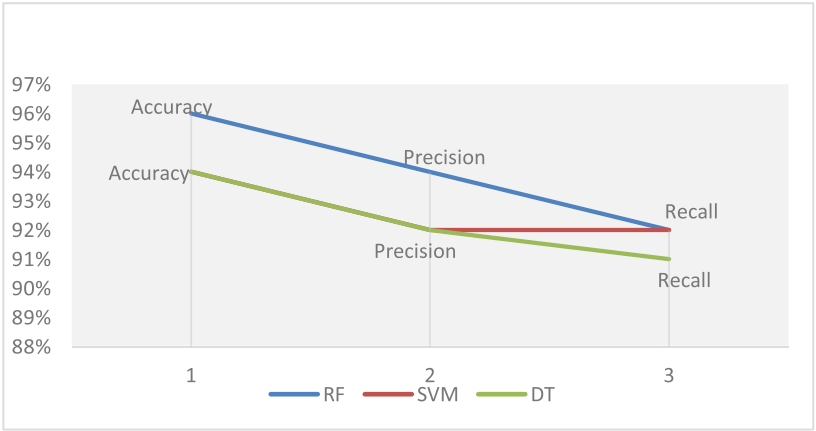

Figure 7. Results (NSL-KDD Dataset)

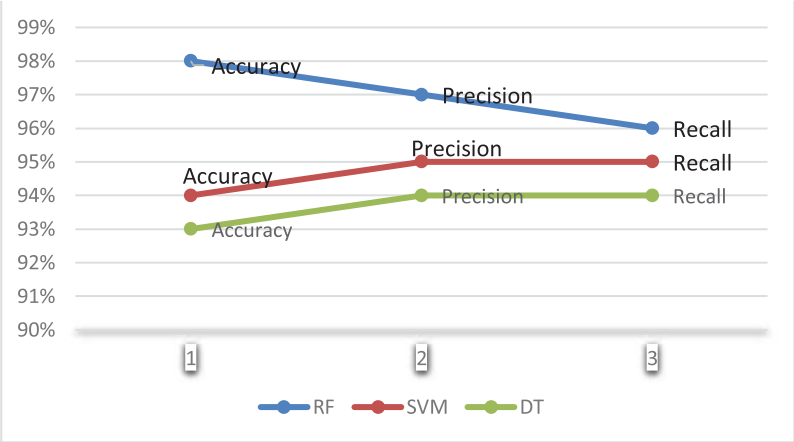

Figure 8. Results (CSE-CIC-IDS2018 Dataset)

7. Conclusion and Future Work

The rate of cybercrime is rapidly increasing, posing a significant disadvantage to technology. There are many attacks and methods by which attackers can breach systems. To secure systems from such attackers, researchers created several solutions based on machine learning algorithms, which are crucial for detecting and safeguarding assets from a variety of threats. Using ML methodologies, this paper proposed a strategy for more precisely detecting attacks in IDS. The suggested approach employs two of the most extensively used datasets for experimentation. This methodology has an overall accuracy of 96% for the NSL-KDD and 98% for the CSE-CIC-IDS2018 dataset. The suggested system identifies network attacks with greater accuracy and precision than previous methods. Deep learning techniques will be utilized in the future to improve categorization accuracy.

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, the 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, volume 33, pp 250-257, the 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, the 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, the 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, the 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, the year 2012.

8. Choo, Kim-Kwang Raymond, "The cyber threat landscape: Challenges and future research directions", Computers & Security, volume 30, pp719-731, the 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 the 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 the intrusion detection system (IDS) based on the computer network. 2017 2nd International Conference 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 the 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. Ashraf, 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 the 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, SwarnaPriya 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, RadhaTamal, "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, MdReazul, Abdur Rahman Onik, and TanvirSamad." A network intrusion detection framework based on Bayesian network using wrapper approach." International Journal of Computer Applications 166.4 (2017).

53. Hagos, DestaHaileselassie, 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. DivyaGoyal, 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. 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 MohdNizam 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, WarusiaSulaiman, md nasirUdzir, 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 Conference 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 Detection 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 Detection 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: fifty-three. (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.