Chapter 39
DDoS Attack Detection Using Ensemble Machine Learning

Adeeba Anis and Md. Shohrab Hossain

1 Introduction

An intentional attempt to block a computer network or website or server from functioning correctly by overflowing it with incoming data is known as a DDoS attack [1]. In simple terms, it impedes genuine users' access to the network or website by acting as a virtual traffic jam that clogs it up [2]. Usually, the attacker forms a "botnet" by coordinating a massive number of computer or devices to concurrently send a massive amount of requests or data to the target network or website. The server can no longer handle the volume of traffic, which results in a major slowdown or server breakdown. This attack is frequently used as a form of extortion, threats, or mischief [3].

DDoS attacks involve a number of drawbacks for both the individuals or organizations who are being attacked. The main objective of a DDoS attack is to disable or disrupt the targeted network or website. As it causes delay and lost sales, this can have serious consequences for businesses. Individuals suffer because they cannot access the websites or Internet services they require. As a result of disruption of services, the business faces financial loss and reputational harm.

Now, intruders have additional opportunity to launch malicious attacks as data is generated from many sources, so it is necessary to defend servers against them [4]. Since the attacker's traffic could be difficult to distinguish from legitimate traffic, DDoS attacks can be challenging to detect and mitigate. Among the techniques and technologies that may be used to identify and lessen, DDoS attacks include network traffic analysis, anomaly detection, rate restriction, blacklisting, and cloud-based DDoS defense.

A. Anis (✉)
Military Institute of Science and Technology, Dhaka, Bangladesh
e-mail: adeebaanis137@gmail.com

Md. Shohrab Hossain
Bangladesh University of Engineering and Technology, Dhaka, Bangladesh

M. Pandit et al. (eds.), *Artificial Intelligence and Sustainable Computing*, Algorithms for Intelligent Systems, https://doi.org/10.1007/978-981-97-0327-2_39

To sum up, preventing an attack using DDoS is essential to maintaining service availability, a positive user experience, minimizing monetary losses, safeguarding reputation, preventing secondary attacks, following the legal requirements, and upholding cybersecurity best practices.

Numerous works have been done on the systems related to intrusion detection. By using DT and NB machine learning techniques on the CAIDA's dataset, Tuan et al. [5] were able to detect DDoS attacks by using SVM, NB, DT, ANN, and K-means. Polat et al. [6] detected DDoS attack on Software-Defined Network (SDN) dataset. On UNSW-NB 15, Azmi et al. [7] used ANN, NB, and DT to detect a DDoS attack. Beulah et al. [8] detected DDoS attack by using ensemble voting technique for SVM and LR on KDD CUP dataset. However, [5–7] didn't use any ensemble machine learning technique while [8] used ensemble voting machine learning for two machine learning algorithms, SVM and LR. Our work *differs* from the previous work in a way that we have used ensemble voting machine learning technique for three machine learning algorithms. We choose three of these four machine learning algorithms—KNN, SVM, DT, and NB—for the ensemble machine learning technique.

The objective of our work is to apply three feature selection techniques (ANOVA, mutual information, and feature importance) on Kaggle dataset to select the most dominant features. We have worked on the top most 10 and 11 features selected by the preceding feature selection techniques.

Our contributions in this paper are (i) deploying traditional machine learning techniques—SVM, KNN, DT, and NB, (ii) implementing ensemble voting technique, and (iii) evaluating the result in relation to accuracy, precision, recall and $F1$-score.

The remaining sections are arranged as follows: The related works are reviewed in Sect. 2. Section 3 contains the explanation of the methodology. The result analysis is presented in Sect. 4. Finally, the paper's conclusion is presented in Sect. 5.

2 Related Work

DDoS attacks are one of the most significant dangers in the field of information technology that must be identified and tackled. The section is related to the detection of DDoS attack using various feature selection techniques used with various machine learning and deep learning algorithms on various datasets.

Several Intrusion Detection Systems have been deployed using machine learning and deep learning algorithms. The capacity of a system to learn from a specific training dataset and the procedure of performing analysis on the given dataset to solve related tasks is known as machine learning [9]. Deep learning is a part of machine learning which is based on Artificial Neural Network (ANN) that uses numerous layers to analyze data and, in some situations, outperforms shallow machine learning algorithms in terms of accuracy. Some of the deep learning algorithms are Deep Brief Networks (DBNs), Deep Neural Networks (DNNs), Recurrent Neural Networks (RNNs), and Convolutional Neural Networks (CNNs) [10]. Again, SVM, KNN, NB, LR, and DT are some machine learning algorithms.

Hailye Tekleselassie [11] proposed a deep learning-based method for detecting DDoS attacks. A combined approach of Convolutional Neural Network (CNN) with stacked autoencoder (SAE) is offered for DDoS detection using the CICIDS2017 dataset. Here, the TP rate, FP rate, precision, recall, F-measure, and accuracy of SMO, Bayes net, and RF have been compared. The study demonstrates that RF performs better in terms of all metrics.

A DNN has been proposed by Bhardwaj et al. in paper [12] to classify the network into normal and DDoS attack. To select features, stacked autoencoder (AE) is used. On two distinct datasets, NSL-KDD and CICIDS2017, the proposed optimized AE with DNN is compared with several state-of-the-art techniques in terms of performance measures, including detection accuracy, precision, recall, and $F1$-score.

Then, many machine learning-based techniques have been considered. In paper [7], Azmi et al. used the data reduction and information gain approach to select features from the UNSW-NB 15 dataset. After selecting the specific features, accuracy, precision, true positive, and false positive have been measured using different algorithms such as ANN, NB, and DT. The accuracy obtained by the selected features comes out to be better than the accuracy obtained by the original dataset.

Feature selection methods are used, in order to get a smaller subset of input features and increase accuracy with a smaller number of features. Araujo et al. [13] have measured the impact of feature selection techniques by XGBoost algorithm on a public dataset to detect DDoS attack classification. Moreover, accuracy, precision, recall, and F1-score metrics have been compared using various algorithms which include mutual information, ANOVA, RFE, XGBoost gain, and ensemble methods and the outcome demonstrates that ANOVA is the most effective method for this dataset.

Polat et al. used a variety of machine learning methods to identify DDoS attacks in Software-Defined Networks (SDN) [6]. At first, filter-based, wrapper-based, and embedded-based feature selection techniques have been applied to select features from the experimental SDN topology dataset. Afterward, accuracy, sensitivity, specificity, precision, and $F1$-score are measured using KNN, SVM, NB, and ANN algorithms. After completing the experiment, it is found that KNN gives the highest accuracy (98.3%) with wrapper feature selection technique.

In paper [5], Tuan et al. evaluated the performance using the accuracy, sensitivity, specificity, false alarm rate (FAR), false positive rate (FPR), AUC, and Matthews Correlation Coefficient (MCC) on the UNBS-NB 15 and KDD99 dataset. In this experiment, supervised machine learning algorithms—DT, SVM, NB, ANN—and unsupervised machine learning algorithms—K-means, X-means—have been used to evaluate the performance.

Beulah et al. [8] proposed an ensemble method by combining SVM and LR to detect DDoS attack on KDD CUP dataset. In this work, voting classifier is used for the ensemble method. The overall performance of the proposed work has a high accuracy and low false positive. The accuracy of the proposed work is 99.2%.

Kumar and Kamatchi proposed an ensemble machine learning technique to detect anomaly-based intrusions [14]. The authors used NSL-KDD dataset to apply algorithms like DT, Bayes classifier, RNN-LSTM, and RF. In this work, ensemble voting

is used to increase the overall performance compared with the existing algorithms. Table 1 offers a summary of the related works including dataset, feature selection technique, classifier, number of features, and metrics.

3 Methodology

To complete the entire process of identifying DDoS attacks on our dataset, this section covers 5 phases. First, data is gathered from a reliable online source. The vital features are then obtained using a variety of feature selection strategies. After that, data is pre-processed to remove any extraneous information, and then traditional and ensemble algorithms are used to detect DDoS attacks along with the performance evaluation shown in Fig. 1.

3.1 Dataset

The dataset used in our research, DDoS classification, is gathered from the trustworthy website Kaggle.com [15]. There are 17,171 rows and 42 columns in the dataset. With the use of this informative dataset, which divided the results into two categories, DDoS and normal, a precise representation of a DDoS attack may be obtained.

3.2 Feature Selection

In order to obtain a subset from the provided dataset with improved accuracy while using fewer features, feature selection technique is used. The features are chosen to use a variety of factors in this feature selection technique [16, 17]. Based on Fig. 2, Analysis of Variance (ANOVA) is a feature selection technique that calculates the F-score for every feature to classify the dataset [18]. In Eq. (1), the ratio of different variance is measured to calculate ANOVA [19].

Let,

- A = variance between groups
- B = variance within groups.

$$F = \frac{A}{B}, \tag{1}$$

$$\text{variance between groups} = \frac{\sum_{i=1}^{k} \sum_{j=1}^{ni} n_i (\overline{Y_i} - \overline{Y})^2}{(n-k)}, \tag{2}$$

Table 1 Related works

References	Year	Dataset	Feature selection technique	Classifier	No. of features	Metrics
Bhardwaj et al. [12]	2020	CICIDS 2017 NSL-KDD	Auto encoder	DNN	23, 40	Accuracy, precision, recall, $F1$-score
Azmi et al. [7]	2021	UNSW-NB 15	Information gain data reduction	ANN, NB, DT	10	Accuracy, precision, true positive rate, false positive rate
Araujo et al. [13]	2021	CICDDoS2019	ANOVA, MI, XGBoost, Gain, RFE	Binary, multiclass	80	accuracy precision recall $F1$-score
Polat et al. [6]	2020	SDN	Filter based wrapper based embedded based	KNN, SVM, NB, ANN	12	Accuracy, sensitivity, specificity, precision, $F1$-score
Hailye Tekleselassie [11]	2021	CICIDS2017	Autoencoder RBM	SMO, Bayes net, RF	21	TP rate, FP rate, precision, recall, F-measure, accuracy
Tuan et al. [5]	2020	UNBS-NB 15 KDD99	Information gain	SVM ANN NB DT K-means	10 9	Accuracy, false alarm rate, sensitivity, specificity, false positive rate AUC
Beulah et al. [8]	2022	KDD CUP	Not mentioned	Ensemble (SVM, LR)	41	Accuracy, true positive rate, false positive rate
Kumar et al. [14]	2020	NSL-KDD	ANOVA	DT, RF, KNN, SVM, DNN, Adaboost	27	Accuracy, precision, recall, F-score, time
This paper	2023	DDoS classification	ANOVA, mutual information, feature importance	KNN, SVM, DT, NB, ensemble (KNN, SVM, DT), ensemble (KNN, SVM, NB), ensemble (SVM, DT, NB), ensemble (KNN, DT, NB)	10, 11	Accuracy, precision, recall, $F1$-score

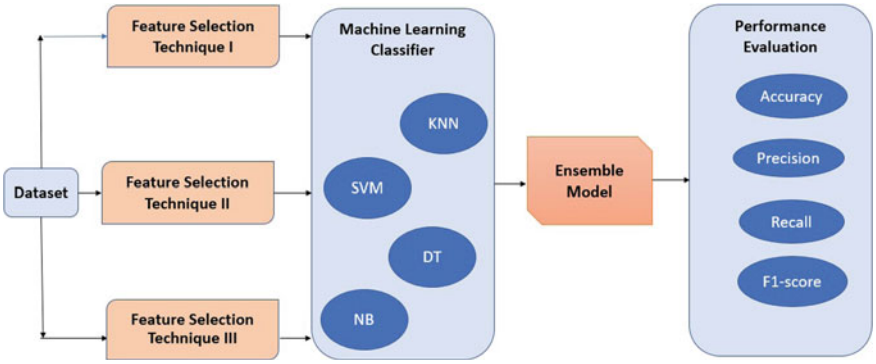

Fig. 1 Diagram of the workflow

Fig. 2 Feature selection using ANOVA

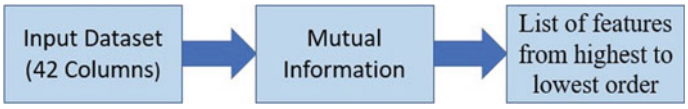

Fig. 3 Feature selection using mutual information

$$\text{variance within groups} = \frac{\sum_i^n n_i (Y_{ij} - \overline{Y_i})^2}{(k-1)}. \tag{3}$$

Another feature selection technique is mutual information that lessens feature redundancy [20] shown in Fig. 3. The mutual information feature selection technique evaluates the relationship between each feature and the target class in order to choose the most effective characteristics for DDoS attack detection [21]. Equation (4) is used to calculate the mutual information.

$$I(M, N) = \sum_{i=1}^{n} \sum_{j=1}^{m} P(M_i, N_i) \log \frac{P(M_i|N_i)}{P(M_i)}. \tag{4}$$

Utilizing the Gini information, the feature importance technique is applied in Fig. 4 to obtain the best features for detecting DDoS attacks. Gini information is calculated using a node's impurity reduction. Again, the impurity is determined by how many samples, out of all the samples, reached the node [22, 23].

Fig. 4 Feature selection using feature importance

3.3 Data Preprocessing

ANOVA, mutual information, and feature importance methodologies have been used to rank features throughout the feature selection process. Afterward, 10, 11, and 12 features have been selected from highest to lowest order from the dataset for each feature selection technique. The remaining features in the dataset have been removed in order for them to be applied in the ensemble model.

3.4 Ensemble Model

In ensemble machine learning technique, various traditional machine learning algorithms are applied first to evaluate the model. After that, multiple models are combined on the basis of simple average or weighted average [24]. Figure 5 shows the algorithms that have been used to create the ensemble models in this paper. In the simple average technique, the mean value is obtained from different machine learning algorithms and then combined to get the ensemble voting value. In weighted average method, as shown in Eqs. (5) and (6), the arithmetic mean value is obtained from different machine learning algorithm given different weight according to their accuracy [25].

$$\hat{y} = \sum_{j=1}^{m} \frac{\hat{y}_j}{m}, \tag{5}$$

$$\hat{y} = \frac{\sum_{j=1}^{m} (w_j \hat{y}_j)}{\sum_{j=1}^{m} w_j}. \tag{6}$$

Moreover, ensemble method works with two different types of voting: hard voting and soft voting [26]. In hard voting approach, the majority of the prediction by the classifiers is calculated; while in soft voting, the probability prediction of each classifier is combined to calculate the result [27].

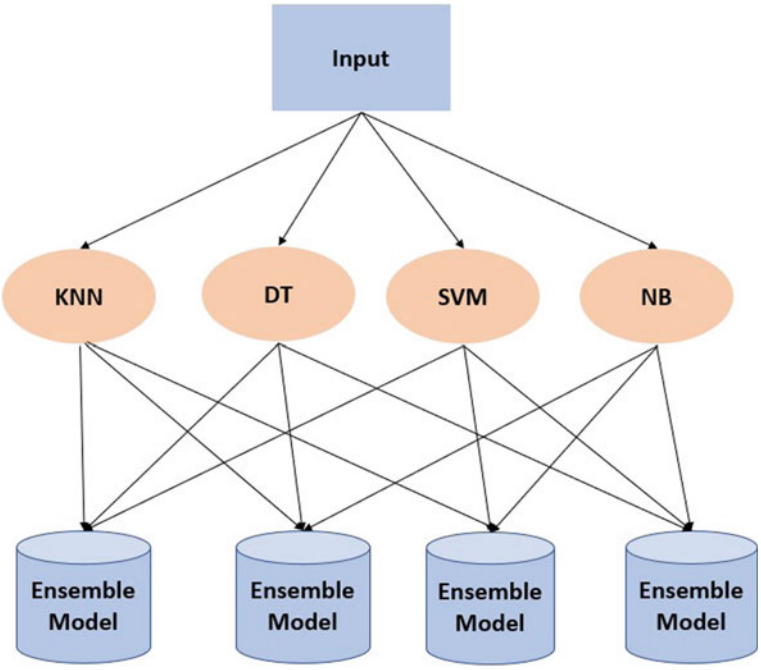

Fig. 5 Ensemble model using voting classifier

4 Result Analysis

Our proposed ensemble model has been compared with the traditional models that are KNN, SVM, DT, and NB. The comparison has been done over the dataset named DDoS classification. To evaluate the proposed ensemble model four performance metrics have been used—precision, recall, $F1$-score, and accuracy. The following provides a description of these metrics.

Accuracy [28] is the ratio of the accurately identified DDoS attack or benign instances to all instances showing in Eq. (7).

$$\text{accuracy} = \frac{\text{TN} + \text{TP}}{\text{TN} + \text{TP} + \text{FN} + \text{FP}}. \tag{7}$$

Precision, as shown in Eq. (8), is the ratio of accurately classified DDoS instances to the total number of incorrect classifications of benign instances as an attack and the DDoS attack instances [29].

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \tag{8}$$

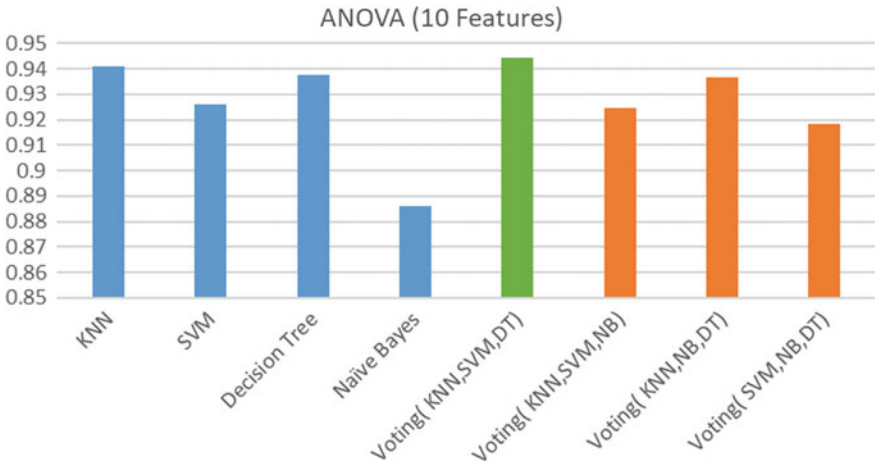

Fig. 6 Accuracy of the models using ANOVA (10 features)

In Eq. (9), recall is measured by the proportion of correctly classified DDoS instances to those that were incorrectly classified as benign instances as well as accurately classified DDoS instances [30].

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \tag{9}$$

Equation (10) shows the application of harmonic mean of recall and precision to calculate $F1$-score [31].

$$F1\,\text{score} = 2 * \frac{P * R}{P + R}. \tag{10}$$

From Fig. 6, it can be seen that Voting (KNN, SVM, DT) has an accuracy score of 0.9245, whereas KNN, a traditional ML algorithm has scored 0.9411 in terms of accuracy. On the other hand, traditional algorithms SVM, DT, and NB have scored 0.9262, 0.9376, and 0.8861 respectively in relation to accuracy.

If precision and recall are taken into consideration, then Voting (KNN, SVM, DT) scores the highest with an $F1$-score of 0.9337, while KNN stands at the second position by having 0.9304 as its $F1$-score followed by DT, Voting (KNN, NB, DT), SVM, Voting (KNN, SVM, NB), Voting (SVM, NB, DT), and NB having 0.9256, 0.9242, 0.9106, 0.9082, 0.9001, and 0.8583 as their $F1$-score which can be seen in Table 2.

From Fig. 7 it can be seen that, the accuracy of Voting (KNN, SVM, DT) stands at the top with 0.9499 in case of 11 features. On the other hand, the lowest accuracy which is 0.8870 is achieved by NB.

Conversely, in terms of precision, Voting (KNN, SVM, NB) stands at the top with 0.9812 and the highest recall is 0.9561 by KNN as given in Table 3. As the recall of

Table 2 Performance of the models using ANOVA (10 features)

Algorithm	Accuracy	Precision	Recall	$F1$-score
KNN	0.9411	0.9702	0.8938	0.9304
SVM	0.9262	0.9764	0.8531	0.9106
DT	0.9376	0.9754	0.8805	0.9256
NB	0.8861	0.9500	0.7828	0.8583
Voting (KNN, SVM, DT)	0.9443	0.9831	0.8890	0.9337
Voting (KNN, SVM, NB)	0.9245	0.9786	0.8473	0.9082
Voting (KNN, NB, DT)	0.9369	0.9816	0.8732	0.9242
Voting (SVM, NB, DT)	0.9185	0.9831	0.8890	0.9001

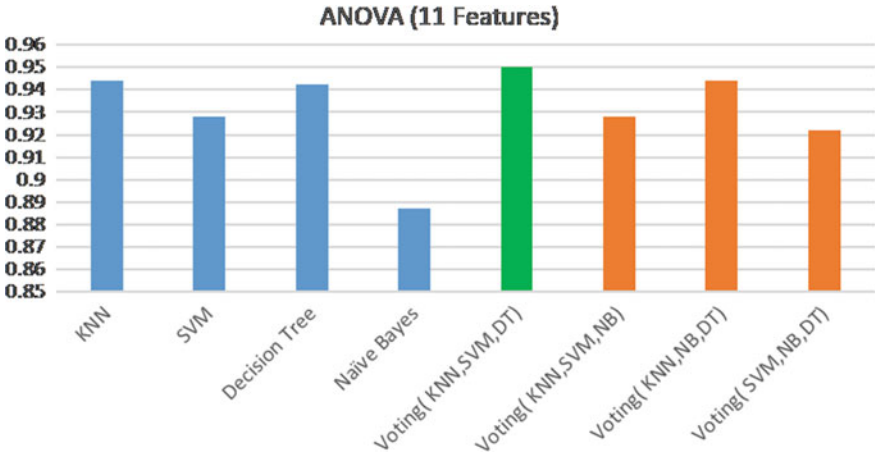

Fig. 7 Accuracy of the models using ANOVA (11 features)

the ensemble model of KNN, SVM, NB is 0.8536 which is the second least value in the recall list, the ensemble model of KNN, SVM, DT ends up at the top with an $F1$-score of 0.9430.

Figure 8 shows 0.9837 to be the highest accuracy which is of Voting (KNN, SVM, DT) followed by KNN's accuracy score of 0.9804. The other ensemble methods have accuracy scores of 0.9718, 0.9720, and 0.9727 by Voting (KNN, NB, DT), Voting (KNN, SVM, NB), and Voting (SVM, NB, DT), respectively.

Table 4 shows Voting (KNN, SVM, DT) tops the $F1$-score list with a score of 0.9814 followed by that of KNN which is 0.9804. On the other hand, NB is at the bottom with 0.8654 as its $F1$-score.

Table 3 Performance of the models using ANOVA (11 features)

Algorithm	Accuracy	Precision	Recall	*F*1-score
KNN	0.9439	0.9197	0.9561	0.9375
SVM	0.9278	0.9782	0.8552	0.9126
DT	0.9425	0.9359	0.9334	0.9346
NB	0.8870	0.9513	0.7838	0.8595
Voting (KNN, SVM, DT)	0.9499	0.9462	0.9397	0.9430
Voting (KNN, SVM, NB)	0.9283	0.9812	0.8536	0.9129
Voting (KNN, NB, DT)	0.9439	0.9465	0.9249	0.9356
Voting (SVM, NB, DT)	0.9222	0.9809	0.8399	0.9049

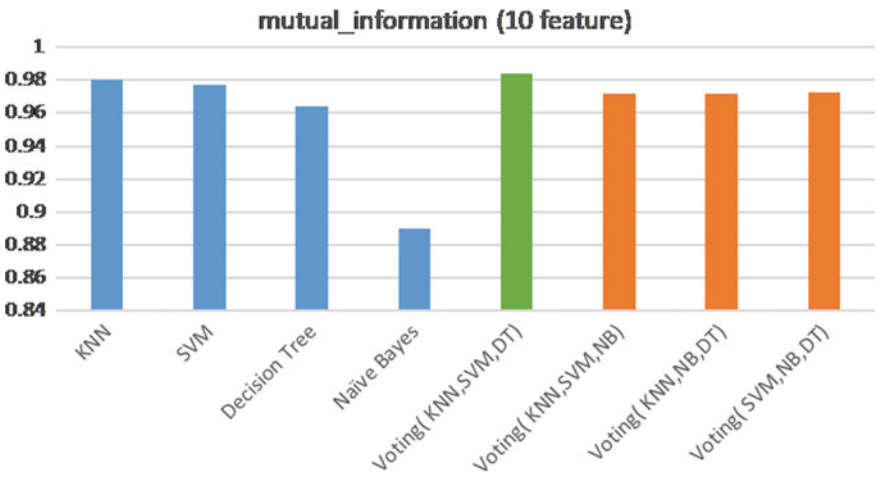

Fig. 8 Accuracy of the models using mutual information (10 features)

Figure 9 reflects that ensemble methods such as Voting (KNN, SVM, DT), Voting (KNN, SVM, NB), Voting (KNN, NB, DT), and Voting (SVM, NB, DT) have scored 0.9856, 0.9830, 0.9786, and 0.9765 in terms of accuracy. Furthermore, most of the traditional algorithms perform well, except NB.

Among the traditional algorithms, KNN is performing the best in case of mutual information (with 11 features) with 98.88% precision and 97.89% recall, showed in Table 5. Other traditional algorithms SVM, DT, and NB have 0.9835; 0.9741, 0.9762; 0.9318 and 0.9541; 0.7918 in precision and recall, respectively.

Figure 10 represents that four models have an accuracy score over 0.98 and among them, Voting (KNN, SVM, DT) has the highest accuracy which is 0.9856. The model

Table 4 Performance of the models using mutual information (10 features)

Algorithm	Accuracy	Precision	Recall	$F1$-score
KNN	0.9804	0.9871	0.9683	0.9776
SVM	0.9769	0.9792	0.9683	0.9737
DT	0.9639	0.9738	0.9434	0.9584
NB	0.8901	0.9399	0.8018	0.8654
Voting (KNN, SVM, DT)	0.9837	0.9856	0.9773	0.9814
Voting (KNN, SVM, NB)	0.9720	0.9836	0.9524	0.9678
Voting (KNN, NB, DT)	0.9718	0.9836	0.9519	0.9675
Voting (SVM, NB, DT)	0.9727	0.9790	0.9588	0.9688

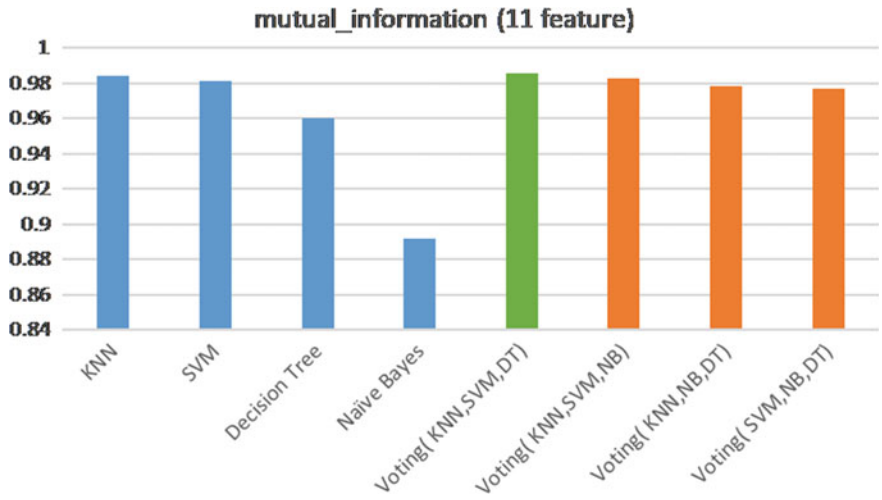

Fig. 9 Accuracy of the models using mutual information (11 features)

is followed by KNN, Voting (KNN, SVM, NB), and SVM with accuracy score of 0.9841, 0.9830, and 0.9814, respectively.

Table 6 shows that, in terms of precision, KNN scored the highest, which is 0.9827. The next highest score for precision is 0.9826 that is for Voting (KNN, SVM, DT). The highest score of $F1$-score is 0.9760 which is for SVM and the lowest $F1$-score is 0.8597 by NB.

Figure 11 reflects that ensemble methods such as Voting (KNN, SVM, DT), Voting (KNN, SVM, NB), Voting (KNN, NB, DT), and Voting (SVM, NB, DT) have scored 0.9886, 0.9704, 0.9676, and 0.9755 in terms of accuracy. In addition, SVM, KNN, and DT score are 0.9814, 0.9760, and 0.9711, respectively.

Table 5 Performance of the models using mutual information (11 features)

Algorithm	Accuracy	Precision	Recall	$F1$-score
KNN	0.9841	0.9888	0.9841	0.9865
SVM	0.9814	0.9835	0.9741	0.9788
DT	0.9599	0.9762	0.9318	0.9535
NB	0.8915	0.9541	0.7918	0.8654
Voting (KNN, SVM, DT)	0.9856	0.9883	0.9789	0.9835
Voting (KNN, SVM, NB)	0.9830	0.9877	0.9736	0.9806
Voting (KNN, NB, DT)	0.9786	0.9875	0.9635	0.9754
Voting (SVM, NB, DT)	0.9765	0.9848	0.9614	0.9730

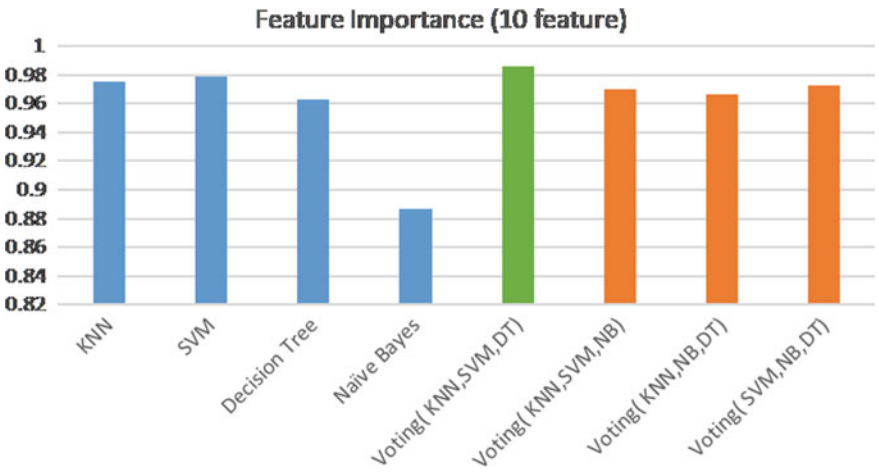

Fig. 10 Accuracy of the models using feature importance (10 features)

Among the traditional algorithms, SVM is performing the best in case of feature importance (with 11 features) with an $F1$-score of 0.9789, showed in Table 7. However, Voting (KNN, SVM, DT) performs far better with an $F1$-score of 0.9886.

5 Conclusion

Both businesses and non-profits of all sizes, DDoS attacks carry a serious risk. Businesses and organizations can lessen their chance of becoming the target of a DDoS attack by adopting precautions for their own safety. We introduced a DDoS

Table 6 Performance of the models using feature importance (10 features)

Algorithm	Accuracy	Precision	Recall	$F1$-score
KNN	0.9751	0.9827	0.9604	0.9714
SVM	0.9788	0.9747	0.9773	0.9760
DT	0.9625	0.9806	0.9334	0.9564
NB	0.8870	0.9495	0.7854	0.8597
Voting (KNN, SVM, DT)	0.9860	0.9826	0.9857	0.9842
Voting (KNN, SVM, NB)	0.9697	0.9788	0.9519	0.9652
Voting (KNN, NB, DT)	0.9669	0.9818	0.9424	0.9617
Voting (SVM, NB, DT)	0.9730	0.9779	0.9604	0.9691

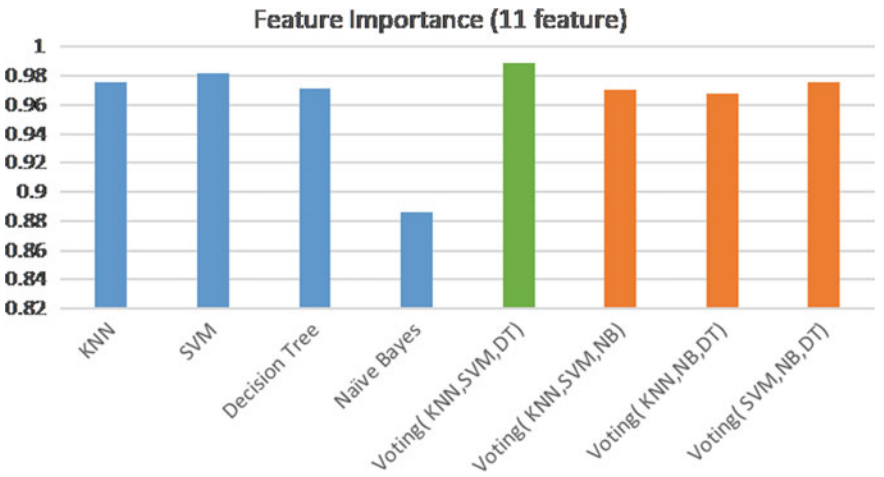

Fig. 11 Accuracy of the models using feature importance (11 features)

attack detection system in this work that uses machine learning methods to identify the attack. We applied ensemble voting technique to detect the attack with more accuracy and fewer features. By analyzing our result, we can conclude that DDoS attack can be detected with more accuracy by feature importance feature selection technique. Moreover, four ensemble models using voting have been implemented and among them the combination of KNN, SVM, and DT performs better than the other algorithms. Finally, as a future work, we are planning to implement our work as real time DDoS attack detection system to reduce the attack in a huge scale.

Table 7 Performance of the models using feature importance (11 features)

Algorithm	Accuracy	Precision	Recall	$F1$-score
KNN	0.9760	0.9833	0.9619	0.9725
SVM	0.9814	0.9753	0.9826	0.9789
DT	0.9711	0.9825	0.9514	0.9667
NB	0.8861	0.9494	0.7833	0.8584
Voting (KNN, SVM, DT)	0.9886	0.9832	0.9910	0.9871
Voting (KNN, SVM, NB)	0.9704	0.9794	0.9530	0.9660
Voting (KNN, NB, DT)	0.9676	0.9824	0.9434	0.9625
Voting (SVM, NB, DT)	0.9755	0.9786	0.9656	0.9721

References

1. Savita TS, Sharma MR (2023) DDoS attack detection using soft voting classifier. J Comput 52(3):66–79
2. Anthi E, Williams L, Javed A, Burnap P (2021) Hardening machine learning denial of service (DoS) defences against adversarial attacks in IoT smart home networks. Comput Secur 108:102352
3. Kumar K, Barver A (2021) A DDoS attack detection using deep learning—a review. IJFMR Int J Multidiscip Res 5(3):1–11
4. Samat NA (2022) Intrusion detection system: challenges in network security and machine learning. Easy Chair Preprint no. 8578
5. Tuan TA, Long HV, Son LH, Kumar R, Priyadarshini I, Son NTK (2020) Performance evaluation of Botnet DDoS attack detection using machine learning. Evol Intell 13:283–294
6. Polat H, Polat O, Cetin A (2020) Detecting DDoS attacks in software-defined networks through feature selection methods and machine learning models. Sustainability 12(3):1035. https://doi.org/10.3390/su12031035
7. Azmi MAH, Foozy CFM, Sukri KAM, Abdullah NA, Hamid IRA, Amnur H (2021) Feature selection approach to detect DDoS attack using machine learning algorithms. JOIV: Int J Inform Visual 5(4):395–401. https://doi.org/10.30630/joiv.5.4.734
8. Beulah M, Pitchai Manickam B (2022) Detection of DDoS attack using ensemble machine learning techniques. In: Soft computing for security applications: proceedings of ICSCS 2021. Springer, pp 889–903
9. Janiesch C, Zschech P, Heinrich K (2021) Machine learning and deep learning. Electron Markets 31(3):685–695
10. Liu H, Lang B (2019) Machine learning and deep learning methods for intrusion detection systems: a survey. Appl Sci 9(20):4396
11. Tekleselassie H (2021) A deep learning approach for DDoS attack detection using supervised learning. In: MATEC web of conferences, vol 348. EDP Sciences, p 01012. https://doi.org/10.1051/matecconf/202134801012
12. Bhardwaj A, Mangat V, Vig R (2020) Hyperband tuned deep neural network with well-posed stacked sparse autoencoder for detection of DDoS attacks in cloud. IEEE Access 8:181916–181929. https://doi.org/10.1109/ACCESS.2020.3028690

13. de Araujo PHHN, Silva A, Junior NF, Cabrini F, Santiago A, Guelfi A, Kofuji S (2021) Impact of feature selection methods on the classification of DDoS attacks using XGBoost. J Commun Inf Syst 36(1):200–214. https://doi.org/10.14209/jcis.2021.22
14. Kumar YV, Kamatchi K (2020) Anomaly based network intrusion detection using ensemble machine learning technique. Int J Res Eng 3:290–297
15. Krishna R. Datasets/Kaggle. https://www.kaggle.com/datasets/ramakrishna0810/ddos-classification. Accessed 10 Jul 2023
16. Kabir MH, Mahmood S, Al Shiam A, Musa Miah AS, Shin J, Molla MKI (2023) Investigating feature selection techniques to enhance the performance of EEG-based motor imagery tasks classification. Mathematics 11(8):1921. https://doi.org/10.3390/math11081921
17. Bagherzadeh F, Mehrani MJ, Basirifard M, Roostaei J (2021) Comparative study on total nitrogen prediction in wastewater treatment plant and effect of various feature selection methods on machine learning algorithms' performance. J Water Process Eng 41:102033. https://doi.org/10.1016/j.jwpe.2021.102033
18. Zaini NAM, Awang MK (2023) Hybrid feature selection algorithm and ensemble stacking for heart disease prediction. Int J Adv Comput Sci Appl 14(2):158–165
19. Azhar M, Ullah S, Ullah K, Shah H, Namoun A, Rahman KU (2023) A three-dimensional real-time gait-based age detection system using machine learning. CMC Comput Mater Contin 75(1):165–182. https://doi.org/10.32604/cmc.2023.034605
20. Ma G, Zhang J, Liu J, Wang L, Yu Y (2023) A multi-parameter fusion method for cuffless continuous blood pressure estimation based on electrocardiogram and photoplethysmogram. Micromachines 14(4):804
21. Hashim MS, Yassin AA. Using Pearson correlation and mutual information (PC-MI) to select features for accurate breast cancer diagnosis based on a soft voting classifier. Iraqi J Electr Electron Eng 43–53 (2023). https://doi.org/10.37917/ijeee.19.2.6
22. Pierzyna M, Saathof R, Basu S (2023) Pi-ML: a dimensional analysis-based machine learning parameterization of optical turbulence in the atmospheric surface layer. arXiv—PHYS—Atmospheric and Oceanic Physics, pp 1–8. arXiv:2304.12177
23. Tikhe SA, Rana DP (2023) Fine-tuned predictive models for forecasting severity level of COVID-19 patient using epidemiological data. In: Frontiers of ICT in healthcare: proceedings of EAIT 2022. Springer, pp 431–442
24. Akhtar MS, Feng T (2022) Comparison of classification model for the detection of cyber-attack using ensemble learning models. EAI Endors Trans Scalable Inf Syst 9(5). https://doi.org/10.4108/eai.1-2-2022.173293
25. Solano ES, Affonso CM (2023) Solar irradiation forecasting using ensemble voting based on machine learning algorithms. Sustainability 15(10):7943. https://doi.org/10.3390/su15107943
26. Atif M, Anwer F, Talib F (2022) An ensemble learning approach for effective prediction of diabetes mellitus using hard voting classifier. Indian J Sci Technol 15(39):1978–1986. https://doi.org/10.17485/IJST/v15i39.1520
27. Karim A, Shahroz M, Mustofa K, Belhaouari SB, Joga SRK (2023) Phishing detection system through hybrid machine learning based on URL. IEEE Access 11:36805–36822. https://doi.org/10.1109/ACCESS.2023.3252366
28. Söğüt E, Erdem OA (2023) A multi-model proposal for classification and detection of DDoS attacks on SCADA systems. Appl Sci 13(10):5993. https://doi.org/10.3390/app13105993
29. Saravanakumar G, Naveen VM, Koushik PH, Sneha C et al (2023) A DDoS attack categorization and prediction method based on machine learning. J Popul Ther Clin Pharmacol 30(9):300–307. https://doi.org/10.47750/jptcp.2023.30.09.030
30. Das S, Venugopal D, Shiva S (2020) A holistic approach for detecting DDoS attacks by using ensemble unsupervised machine learning. In: Advances in information and communication: proceedings of the 2020 future of information and communication conference (FICC), vol 2. Springer, pp 721–738
31. Das S, Mahfouz AM, Venugopal D, Shiva S (2019) DDoS intrusion detection through machine learning ensemble. In: 2019 IEEE 19th international conference on software quality, reliability and security companion (QRS-C). IEEE, pp 471–477. https://doi.org/10.1109/QRS-C.2019.00090