TOWARDS ROBUST CYBER DEFENSE: ENSEMBLE MACHINE LEARNING MODELS FOR DARKNET TRAFFIC CLASSIFICATION

R. Poongodi[1], A. Faritha banu[2], B. Ziyaudeen[3]

Abstract

Cyber intelligence groups occasionally bring up the "darknet," a section of the internet that most users do not think is suitable for M2M communication. Evaluating the possible risks to the network must precede any initiatives to enhance its defenses. This study introduces a unique method for machine learning classification called stacking ensemble learning, aimed at assessing and categorizing darknet data. This study attained a 98% accuracy rate in differentiating between Darknet and benign traffic by using the newly available CIC-Darknet2020 dataset and ensembles of machine learning algorithms. It had a 97% success rate in accurately identifying the specific program type accountable for the Darknet traffic. We employed a game-theory-based method to evaluate the results of the machine learning models and demonstrate the influence of the features in order to deepen our comprehension of Darknet traffic patterns. To the best of everyone's knowledge, the dataset's creators have attested to the unprecedented nature of our analysis.

Keywords— Cyber intelligence services, machine-to-machine communication, CIC-Darknet2020, Darknet

I. INTRODUCTION

A "darknet" is an Internet overlay network that requires particular software, setups, or authorization to access..A customized communication protocol is often used by it [1]. An example of a darknet would be a social network that allows users to save documents between themselves over peer-to-peer connections. A second example would be anonymous proxy networks like Tor, which work by creating a series of anonymous connections.The user has supplied the text [2].The name "darknet" became famous due to major news outlets, even though it was never officially recognized. The infamous online drug market Silk Road used it, which led to its association with Tor Onion services. All sorts of activities, legal and illegal, may be facilitated by technologies such as Freenet, I2P, and Tor. By ensuring safety, anonymity, and resistance to censorship, they primarily seek to protect digital rights. Through services like SecureDrop, activists, journalists, media organizations, and witnesses can communicate anonymously on darknets. The user has supplied the text [3].

Modern threats increasingly take the form of sophisticated, covert internet assaults, making cybersecurity an urgent issue. As a cornerstone of cybersecurity efforts, accurately identifying and categorizing the darknet is crucial since it has grown into a prolific venue for criminal activity. In order to tackle constantly evolving challenges, this study explores sophisticated machine learning methods with a focus on stacking ensemble learning.The objective is to improve the accuracy of darknet activity classification. By combining the knowledge of many models, this approach hopes to strengthen cybersecurity procedures and protect against complex cyberattacks. To prevent the digital environment from being compromised and to tackle the complex nature of darknet operations, this research investigates the combination of ensemble learning with encryption.

A. Background

Cybersecurity confronts a continuously changing variety of risks, with the darknet posing particularly difficult obstacles. Due to its reputation for secrecy and encrypted communication, the darknet has become a haven for numerous cyberthreats, such as the spread of viruses, illegal transactions, and the disclosure of private data.Traditional cybersecurity techniques often fail to accurately identify and categorize darknet activity because of the network's clandestine and dynamic characteristics. Owing to the escalating intricacy of cyber threats, machine learning methodologies have substantially contributed to the improvement of threat detection and categorization.

Ensemble Learning:

Ensemble learning, which integrates multiple models, is one method of enhancing overall performance in machine learning.It has the potential to provide more trustworthy and accurate predictions by using many models. A popular form of ensemble learning called stacking combines the output of numerous base models to train a prediction-capable meta-model.Explanation of Ensemble Learning Stacking

Research Scholar,Department of Computer Technology[1]
Karpagam Academy of Higher Education, Coimbatore[1]
Mail Id : poongodi.kwc@gmail.com[1]

Department of Computer Technology[2]
Karpagam Academy of Higher Education, Coimbatore[2]
Mail Id : farithabanu.ahamedsheriff@kahedu.edu.in[2]

Department of Computer Science[3]
PSGR Krishnammal College for Women Coimbatore[3]
Mail Id : profziyaudeen@gmail.com[3]

* Corresponding Author

Structures: Because darknet categorization effectively addresses the difficulties posed by the intricate and dynamic nature of cyberattacks, stacking ensemble learning is justified.

Thanks to model stacking, it's possible to combine several models, each with its special knack for spotting darknet activities. Improved overall classification accuracy and strengthened defenses against darknet malefactors' developing techniques are the results of this variation.

B. Contribution:

Various models, each designed for a certain facet of darknet activity, are included in stacking ensemble learning. The strategy improves overall classification accuracy and decreases false positives and negatives by amalgamating the strengths of many models.

The distinctive characteristic of the darknet is its unconventional and adaptive nature. Stacking ensemble learning, using many models, offers a more versatile defensive mechanism. It is a crucial instrument in the dynamic realm of cybersecurity owing to its remarkable adaptability to various darknet hacking methodologies.

Cybercriminals often use evasion strategies to circumvent prevalent security protocols. Stacking ensemble learning enhances robustness against escape techniques by integrating models with different levels of detecting power. It is more likely to recognize and categorize unique or complex threats on the darknet.Ensemble learning may provide a comprehensive understanding of the darknet's operations. The examination of the darknet ecosystem is hindered by the potential variability in the group's models' capability to detect malicious conduct.

Stacking facilitates a more strategic use of resources by synthesizing the optimal features of many models. Consequently, computer resources may be used more effectively, and dangers might be identified and eradicated more swiftly.

Contemporary cybersecurity methodologies are perpetually advancing because of the implementation of stacking ensemble learning in darknet classification. The growing complexity of cyber threats necessitates the development of innovative solutions for their mitigation. The examination of darknet categorization by ensemble learning yields significant insights and indicates potential avenues for future research at the intersection of cybersecurity and machine learning. The findings may stimulate more study and advancements in digital environment security mechanisms.

II. RELATED WORKS

This study addresses the challenge of identifying and classifying cyberattacks within Internet of Things (IoT) communication networks, highlighting the need for a top-down machine learning architecture to improve cybersecurity in this sector. The limitations and specific shortcomings of this work are not addressed in the citation provided. [4].

Through the use of attack-aware ensemble learning traffic routing, By considering potential attacks at every step of the routing process, this study aims to enhance the performance of IoT networks. The shortcomings and restrictions of the suggested method are not discussed in reference [5].

The paper suggests a method to improve the accuracy of detecting abnormal network activity that is based on anomalies and uses a Variational Autoencoder (VAE) in conjunction with network flow characteristics. The referenced work fails to discuss the work's shortcomings and restrictions. [6].

In order to address the challenges of recognizing and classifying Tor-related activities, this study focuses on examining Tor (The Onion Router) usage on PC and mobile platforms utilizing multilayer identification and classification techniques.

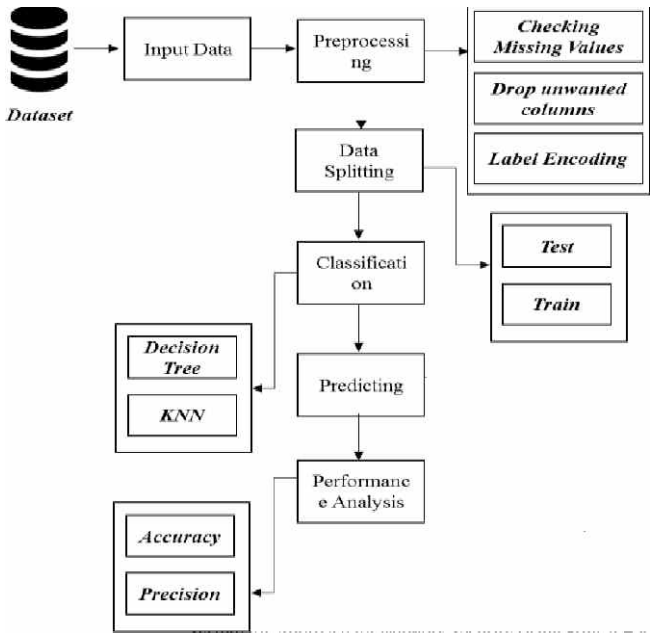

Fig. 1. System Architecture

Reference [7] does not address the faults or limits of the proposed multilevel analysis. This study automates the real-time detection of criminal intent in darknet traffic using a weight-agnostic neural network design and comprehensive data analysis. The scope of this inquiry and its particular limitations are not discussed in the referenced work [8].

The challenge of defending cloud infrastructures against Distributed Denial of Service (DDoS) attacks with Software-Defined Networking (SDN) is the focus of this study. The POX controller and entropy-based protection

approaches are the key focus for improved defense against distributed denial of service (DDoS) attacks.Some of the study's shortcomings and restrictions are missing from the cited source. [9].

This study uses association rule learning to understand and evaluate the behavior of IoT malware.The study's overarching goal is to use darknet sensor data analysis to better identify and categorize IoT malware [10].

The method employed to investigate IoT malware activity has some disadvantages, although they are not mentioned in the cited work.

This paper describes a pilot study that used Wireshark to monitor network traffic in order to detect Distributed Denial of Service (DDoS) attacks in Software-Defined Networking (SDN). The primary objective is to use SDN technology to enhance the detection and mitigation of distributed denial of service attacks.

There are no limitations or downsides to this pilot project listed in the reference [11].

This tool attempts to detect worldwide cyberthreats in real time by using darknet activity as a foundation. The goal of the approach is to improve cyber security awareness and make early detection easier by using Graphical Lasso to evaluate unusual synchronization. The reference does not specifically address the shortcomings and restrictions of the real-time cyber threat identification system [12]. In [15], the researchers reviewed existing security challenges in mobile ad hoc networks (MANETs), highlighting that traditional encryption and authentication mechanisms alone are insufficient to protect against malicious behavior due to the network's dynamic and infrastructure-less nature

III. PROPOSED METHODOLOGY

Figure 1 displays the suggested architecture, and the specifics are explained below.

Dataset :

The "darknet" refers to the purportedly unidentified portion of the internet's address space that computers use globally for communication. The black space functions as a listening environment because it is passive, allowing only incoming packets to get through while blocking outgoing ones. As a result, any message purporting to come from space is viewed with mistrust.The scarcity of authentic servers on the darknet renders all traffic suspect, perhaps leading to misconfiguration, probing, or backscattering. Black holes, sinkholes, and network telescopes are other names for darknets.The real-time categorization of applications depends on the classification of darknet traffic. Darknet traffic analysis makes it possible to monitor malicious software before it conducts an attack and to identify dangers early in an occurrence.This work creates a comprehensive darknet dataset by combining the public datasets ISCXTor2016 and ISCXVPN2016, which include Tor and VPN activities. This method makes it possible to recognize and examine VPN and Tor applications as authentic depictions of dark web traffic.The CICDarknet2020 dataset utilizes a dual method to generate both darknet and benign traffic. The second layer is the primary source of most darknet traffic, which includes peer-to-peer transfers, music streaming, web surfing, chatting, emailing, video streaming, and voice-over IP. To build a representative dataset, we combined the VPN and Tor traffic from our earlier datasets, ISCXTor2016 and ISCXVPN2016, and classified them into the proper darknet categories. The programs that produce network traffic and the different kinds of darknet activity are listed in Table 1.

Table 1: Classification and Application of Traffic

Classification of Traffic	Application
Sound Channel	Youtube
Browsing	Chrome
Chat	Whatsapp, Skype, Facebook
Email	IMAPS
Video Channel	Youtube

Data preprocessing :

We aimed to complete the classification using only the measured properties; therefore, we eliminated the unnecessary columns: Source IP, Destination IP, Source Port, Destination Port, and Protocol. In discussions with the providers of the ISCXVPN2016 dataset at the CIC, Lotfollahi et al. [10] explained that the network layer header's source and destination IP addresses are application-specific and so inappropriate for classification..We would appreciate your feedback about the addition of the word "Protocol." On the other hand, Tor can traverse TCP networks [3]. For protocol codes 0 (HOPOFT) and 65 (UDP), Tor contains 35 pieces of data. The protocol column has to be removed since there were too many invalid occurrences. Given that it just supplied squared values for the "Packet Length Std" column, we decided to remove the

"Packet Length Variance" function. It is improper to incorporate "Timestamp" into ML models since it doesn't offer any flow-specific information. Additionally, fifteen columns, referred to as "singletons," were cleaned.

Data Split:

When creating data for risk prediction, a 70:30 split between training and testing datasets is standard

procedure.Part of the dataset will go towards training the machine learning model (about 70% of the total), while the other part will be used to evaluate how well the model performed (about 30%).

In contrast to the training set, which helps the model find patterns and correlations in the data, the testing set is used to assess the model's generalisability to new, unfamiliar samples. By striking a balance between providing enough data for model training and guaranteeing a comprehensive assessment of its predictive capabilities on fresh data, this split ratio promotes the creation of an effective cyber threat system.

Data Selection:

The CIC Darknet 2020 dataset must be used carefully and strategically for cybersecurity research to be effective. The dataset contains a broad range of behaviors, from harmless to detrimental. It records and preserves network traffic in darknet scenarios. For data gathering, a representative sample covering a variety of cyberthreat categories commonly seen in darknet contexts must be carefully chosen.Part of this responsibility includes looking for signs of suspicious activity, violations, or anomalies. In order to thoroughly examine the complexities of darknet interactions, scholars and practitioners could concentrate on particular components like timestamps, protocol types, source and destination IP addresses, and packet sizes. Furthermore, make sure that the dataset is balanced and include both typical and uncommon actions.

This will lessen the influence of any biases in the research. The subset selection must follow the research objectives, regardless of whether it is for any cybersecurity duty, such as threat classification or anomaly detection.. To gain a thorough understanding of the complexities of darknet activity and to boost the validity of cybersecurity evaluations utilizing the CIC Darknet 2020 dataset, careful data selection is required.

Data Classification :

Random Forest:

Initially, a random forest was employed. It describes how many decision trees are combined to create an ensemble of classifiers. A random forest, by this definition, is just a group of decision trees that were trained separately on different datasets.

Each tree uses evenly distributed values that are selected at random from a subset of the input data set.. A huge amount of data points are picked at random to construct the shallow trees used in the random forest classification method. This is how the predictions or value categorizations from the aforementioned trees are integrated.

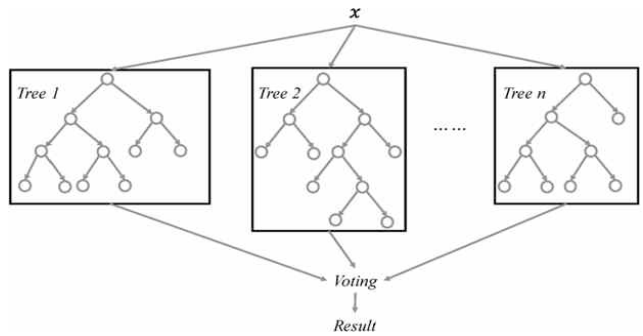

Fig 2. Random forest Architecture

K- Nearest Neighbor:

The model achieved great accuracy in predicting events for the CIC Darknet 2020 dataset by employing the k-Nearest Neighbours (KNN) approach. The well-known and straightforward KNN algorithm successfully generated predictions using the correlations and patterns in the CIC Darknet 2020 data. KNN used the feature space's proximity as a criterion to detect and categorize darknet activities. The model's projected accuracy, recall, precision, and F1 score demonstrated its ability to discriminate between malicious and authentic network data, which was helpful for identifying and mitigating cyberthreats in the darknet environment.The KNN algorithm showed that it could handle the complex CIC Darknet 2020 data because of its flexibility and proximity-based similarity criteria. Consequently, it might be used by cybersecurity experts to find and fix security flaws.

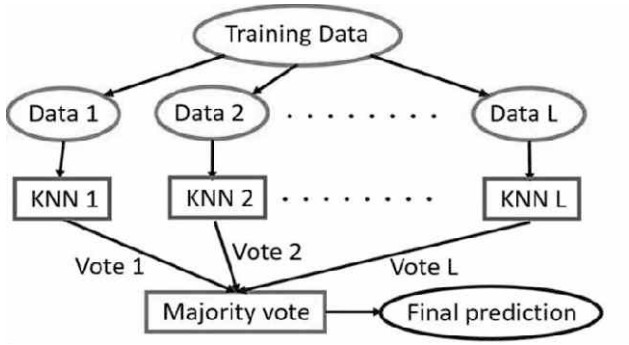

Fig 3. KNN Architecture

Decision Tree:

Decision trees have been used in several studies to analyze Darknet traffic for a variety of objectives. More precisely, decision trees are constructed during training and are described as a method for classifying server traffic. In order to comprehend the traffic, the authors developed a number of characteristics to characterize the behavior of streams.The trees were created by this procedure. Decision trees were shown as an effective way to categorize popular application protocols for TCP connections in the dark. Here, aggregate features were used to characterize each flow. The authors demonstrated the reliability and accuracy of the traffic categorization. It has a greater than 95% accuracy rate.

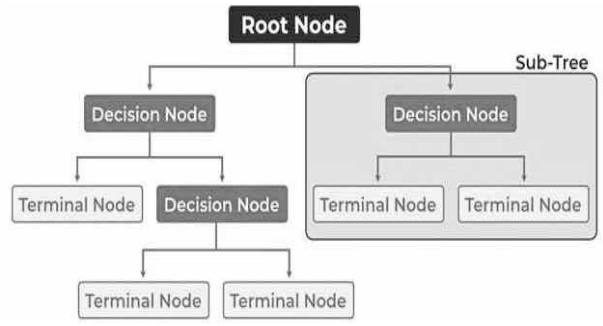

Fig 4. Decision tree Architecture

Ensemble techniques:

The traffic was categorized using a number of methods that are popular with other researchers. One area where group techniques have shown to be quite effective is in evaluations of categorization using information tables. In this investigation, we used three ensemble approaches: Set (RF), Random Forest (RF), and Decision Tree (DT). Trees constructed in randomly chosen subspaces and trees created by pseudo-randomly selecting subsets of the feature vector are both included in the RF category. To transform a struggling student into an accomplished one is what the phrase "boosting" refers to. The decision tree's data points all start off with the same weights. The system modifies the learned classifiers' parameters and point weights in response to performance. If the learning error is small or a certain number of classifiers have been produced, the procedure is repeated. Using a straightforward parameterized function known as the base learner, gradient boosting iteratively minimizes the change in the gradient of the loss function from the current "pseudo" residues to generate additive models. The k-nearest neighbors (KNN) approach has several uses. We both believe that ensemble methods are superior to KNN models for our data set because of their size and composition. Our research showed that the aforementioned ensemble methods were not surpassed by the models based on the autoencoder, multilayer perception, and tablet neural networks. Regarding the outcomes, nothing has been said.

IV. EXPERIMENTAL RESULTS

We used Jupyter Notebook to do our Python programming experiments. We utilized the TensorFlow backend libraries scikit-learn, pandas, numpy, matplotlib, yellowbrick, and keras. The performance evaluation on the CIC Darknet 2020 dataset's findings and discussion offer significant insights into the efficacy of various machine learning algorithms for cybersecurity applications. A remarkable accuracy rate of 94.04% was attained by k-Nearest Neighbors (KNN), one of the tested algorithms. Because of its simplicity and reliance on proximity-based similarity measurements, KNN can distinguish between malicious and lawful network activity.Nevertheless, its processing cost might become an issue when dealing with bigger datasets, which could affect its scalability in real-time scenarios.With a precision of 98.79%, the Decision Trees (DT) algorithm demonstrated strong performance. The interpretability and capacity to grasp nonlinear connections within data are two of decision trees' most notable features. Using the CIC Darknet 2020 dataset, the model's decision boundaries successfully categorised cases. However, to avoid overfitting, hyperparameter adjustment is crucial for finding the sweet spot between complexity and generalizability.The ensemble approach Random Forest (RF), which uses many decision trees, achieved a remarkable 99.25% accuracy rate, surpassing that of individual decision trees. By combining the outputs of several trees, RF can reduce the occurrence of overfitting and improve the accuracy of predictions. This method made RF a strong contender for cybersecurity applications by successfully capturing the complex and varied nature of darknet operations. On the other hand, model interpretability may be affected by RF's ensemble nature.

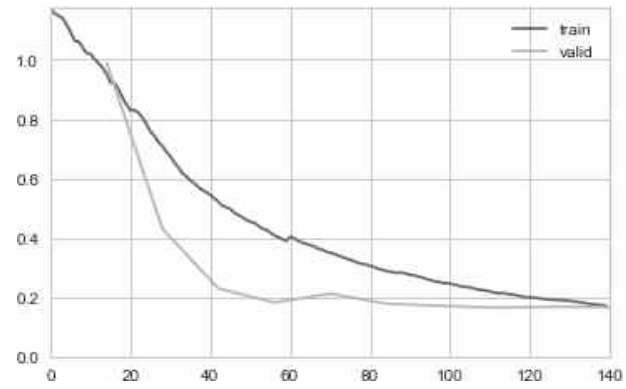

Fig 5. Random Forest Accuracy

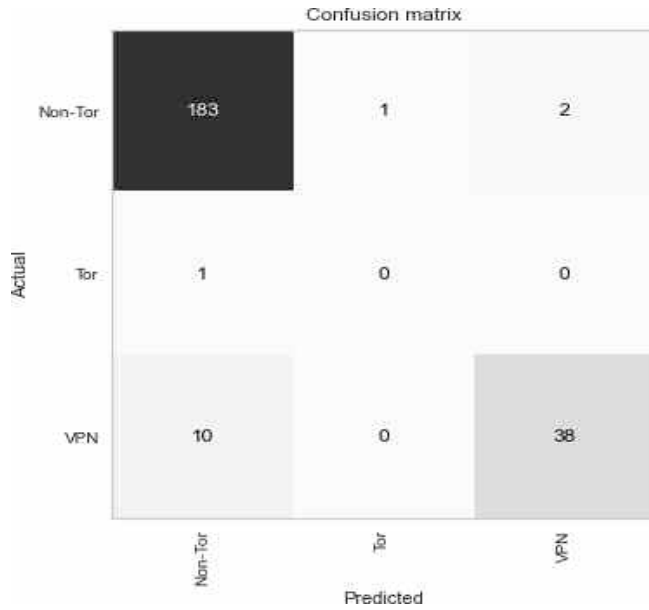

Fig 6. Random Forest Confusion Matrix

An even more impressive accuracy of 99.28% was produced by combining individual decision trees in an ensemble method. This demonstrates how ensemble approaches may improve forecast performance by combining different viewpoints. The ensemble approach showed remarkable accuracy in categorizing cases inside the darknet dataset. It is likely a mix of Random Forest and other algorithms. Taken together, the findings shed insight into the algorithmic subtleties of each one's strengths and weaknesses. The effectiveness of machine learning in cybersecurity applications is shown by the high accuracy rates across the board. When it comes to the complicated job of darknet traffic analysis, Random Forest and ensemble approaches are very formidable tools. In the future, it could be important to address any biases, investigate the significance of features, and evaluate the interpretability of models in real-world cybersecurity applications (see Figure 5).

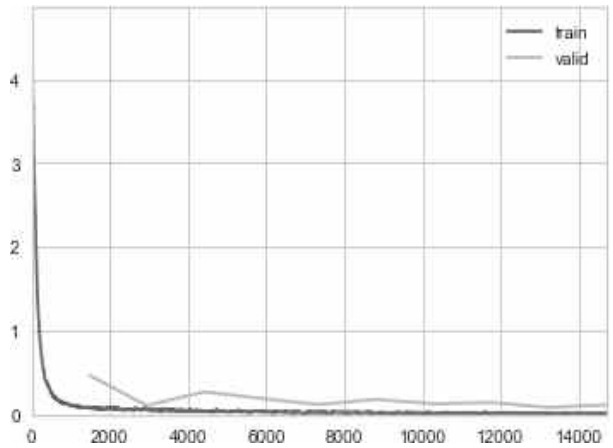

Fig 7. KNN Accuracy

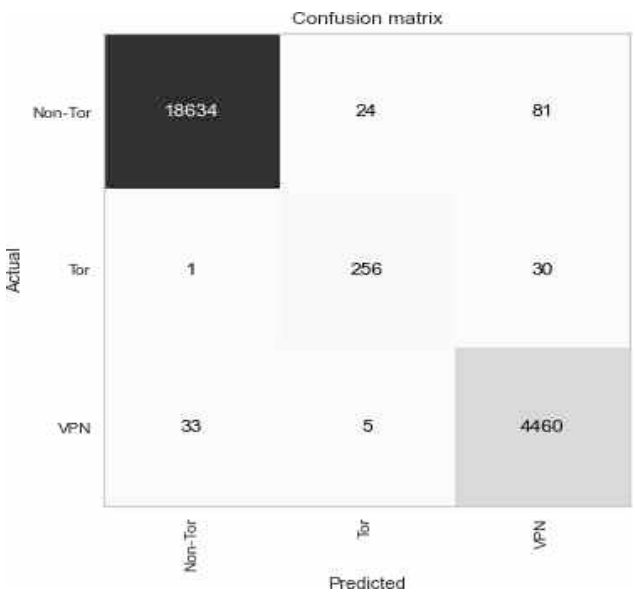

Fig 8. Decision Tree Accuracy

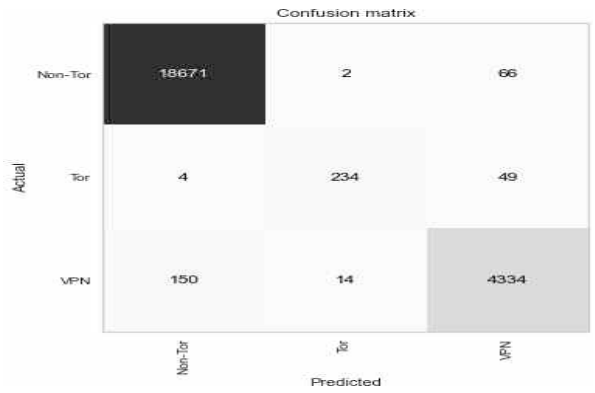

Fig 9. Decision Tree Confusion Matrix

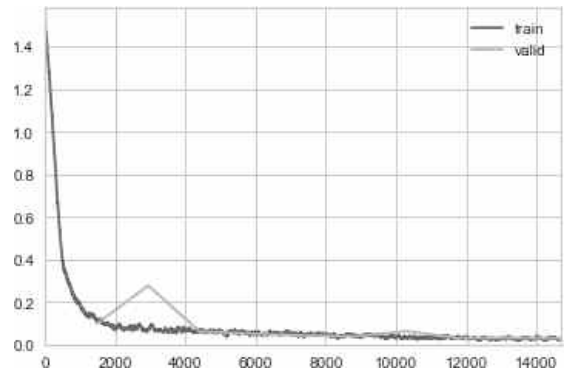

Fig 10. Ensemble Accuracy

V. CONCLUSION AND FUTURE WORK

Evaluating machine learning techniques on the CIC Darknet 2020 dataset demonstrates how important it is to select the right model for cybersecurity applications.The ensemble methodology had the greatest accuracy rate of 99.28% out of all the techniques that were assessed. Even though darknet data is diverse and varied, this all-encompassing method—which most likely combined Random Forest with other algorithms—surpassed others in successfully detecting occurrences. Despite its remarkable accuracy of 94.04%, k-Nearest Neighbours (KNN) could struggle to manage larger datasets or real-time scaling because of its user-friendly and proximity-based technique. The performance of Decision Trees (DT) was outstanding, with an accuracy percentage of 98.79%. Even if DT can be

understood, it is feasible to prevent overfitting by appropriately adjusting the hyperparameters. Random Forest (RF) outperformed individual decision trees with a 99.25% accuracy rate in detecting intricate patterns in darknet activity. By including various perspectives, ensemble techniques may enhance prediction performance, as shown by their consistently high levels of accuracy. Darknet scenarios need advanced techniques for detecting and classifying suspicious behavior, making this finding vital to cybersecurity. The findings demonstrate that an ensemble approach integrating

the best features of many models may provide a comprehensive analysis of darknet traffic. This finding should be noted by security researchers and practitioners as it emphasizes the necessity of employing complex ensemble approaches to increase the precision and dependability of threat detection systems.Improving ensemble techniques, studying interpretability concerns, and Future study may focus on reducing biases to make them more appropriate for cybersecurity scenarios in the real world.The results show that ensemble techniques are the most accurate and resilient way to classify darknet data, and that machine learning is crucial to cybersecurity efforts.

REFERENCES

[1] Gayard, Laurent (2018). Darknet: Geopolitics and Uses. Hoboken, NJ: John Wiley & Sons. p. 158. ISBN 9781786302021.

[2] Pradhan, Sayam (2020). "Anonymous". The Darkest Web: The Dark Side of the Internet. India. p. 9. ISBN 9798561755668.

[3] Press Foundation, Freedom of the. "SecureDrop". GitHub. Freedom of the Press Foundation. Retrieved 28 January 2019.

[4] Abu Al-Haija, Q. Top-Down Machine Learning-Based Architecture for Cyberattacks Identification and Classification in IoT Communication Networks. Front. Big Data 2022, 4, 782902.

[5] Abu Al-Haija, Q.; Al-Badawi, A. Attack-Aware IoT Network Traffic Routing Leveraging Ensemble Learning. Sensors 2021, 22, 241.

[6] Zavrak, S.; Iskefiyeli, M. Anomaly-Based Intrusion Detection from Network Flow Features Using Variational Autoencoder. IEEE Access 2020, 8, 108346–108358.

[7] Wang, L.; Mei, H.; Sheng, V.S. Multilevel Identification and Classification Analysis of Tor on Mobile and PC Platforms. IEEE Trans. Ind. Inform. 2021, 17, 1079–1088.

[8] Demertzis, K.; Tsiknas, K.; Takezis, D.; Skianis, C.; Iliadis, L. Darknet Traffic Big-Data Analysis and Network Management for Real-Time Automating of the Malicious Intent Detection Process by a Weight Agnostic Neural Networks Framework. Electronics 2021, 10, 781.

[9] Mishra A, Gupta N, Gupta B. Defense mechanisms against DDoS attack based on entropy in SDN-cloud using POX controller. Telecommun Syst. 2021;77(1):47–62. doi: 10.1007/s11235-020- 00747-w.

[10] Ozawa S, Ban T, Hashimoto N, Nakazato J, Shimamura J. A study of IoT malware activities using association rule learning for darknet sensor data. Int J Inf Secur. 2020;19(1):83–92. doi: 10.1007/s10207- 019-00439-w

[11] Varghese, J.E.; Muniyal, B. A Pilot Study in Software-Defined Networking Using Wireshark for Analyzing Network Parameters to Detect DDoS Attacks. In Information and Communication Technology for Competitive Strategies (ICTCS 2020); Springer: Singapore, 2021; pp. 475–487.

[12] Han, C.; Shimamura, J.; Takahashi, T.; Inoue, D.; Takeuchi, J.I.; Nakao, K. Real-Time Detection of Global Cyberthreat Based on Darknet by Estimating Anomalous Synchronization Using Graphical Lasso. IEICE Trans. Inf. Syst. 2020, E103-D, 2113–2124.

[13] K. Prabu and T. Ayyaparaj, “Decoding Liver Fibrosis: An AI-Powered Path to Early Cirrhosis Diagnosis,” Karpagam Journal of Computer Sciences, vol. 2, no. 4, pp. xx–xx, Jul.–Aug. 2025.

[14] P. Lalithamani and S. Sandhya, “Adaptive ensemble learning for climate change forecasting and environmental monitoring,” *Karpagam Journal of Computer Science*, vol. 20, no. 4, pp. xx–xx, Jul.–Aug. 2025. doi:10.xxxx/xxxxxx

[15] S.Fabrice and E.J.Thomson Fredrik, Detection and Prevention of Malicious Node Based On Node Behaviour In MANET, International Journal of Advanced Research in Computer Science, Vol.8, No.9, 2017.