ISSN: 2249-7196
IJMRR/April. 2025/ Volume 15/Issue 2s/164-175


# Classification of Darknet Traffic Using Machine Learning

**Petchetti Naga Venkata Sai Kumar**
**P**G scholar, Department of MCA, DNR College, Bhimavaram, Andhra Pradesh.
**Ch.Jeevan Babu**
(Assistant Professor), Master of Computer Applications, DNR college, Bhimavaram, Andhra Pradesh.

***Abstract*** ***Darknet network traffic, characterized by their hidden nature and association with illicit activities,pose significant challenges for network monitoring and security. To address these challenges, thisstudy explores the classification of darknet network traffic using machine learning algorithms. A labeled dataset comprising network traffic captured from darknet environments is collected and preprocessed to extract relevant features. Various machine learning algorithms, including decision trees, random forests, AdaBoost, KNN, Naïve Bayes, Simple CART, Gradient Boosting are evaluated for their effectiveness in classifying darknet traffic. The dataset is divided into training and testing sets, and the selected machine learning models are trained using the training set. The models learn the underlying patterns and relationships between network traffic features followed by relevant classification. The models are then evaluated on the testing set to measure their classification performance in terms of accuracy, precision, recall, and F1-score.The results demonstrate the efficacy of machine learning algorithms in classifying multilayered darknet network traffic. The selected models achieve high accuracy and show promising potential in identifying various types of malicious or illicit activities within the darknet. The study also discusses the interpretability of the models and the insights gained from feature importance analysis, contributing to a better understanding of darknet network behavior. The findings from this research have implications for network security, law enforcement, and cyber security operations, enabling proactive monitoring and detection of darknet activities. The developed classification models can be integrated into real-time traffic analysis systems, aiding in the identification and mitigation of threats originating from the darknet.***

**Keywords:** Darknet Network Traffic, Darknet Traffic Classifier, Machine Learning, Classification, Security, Feature Selection, Ensemble Learning

## I. INTRODUCTION

Darknet refers to a part of the internet that is intentionally hidden and inaccessible through conventional search engines. It is often associated with illicit activities, such as illegal markets, hacking forums, and other forms of cybercrime. Understanding and classifying the network traffic within darknet environments is crucial for detecting and preventing malicious activities.

The research investigates the efficiency of machine learning algorithms in classifying darknet network traffic using a multilayered approach. The multilayered classification approach involves multiple stages of classification, where each stage focuses on a specific aspect or characteristic of the network traffic and compares the performance of each Machine Learning algorithm in classification.

The study aims to evaluate the performance and effectiveness of various machine learning algorithms in accurately classifying different types of darknet network traffic. This may involve the use of supervised learning algorithms, such as decision trees, support vector machines, neural networks, or ensemble methods like random forests or gradient boosting.

The efficiency of the algorithms is assessed based on metrics such as accuracy, precision, recall, and F1 score, which measure the effectiveness of classification in identifying different types of network traffic accurately. The research may also explore the impact of different features, feature selection techniques, or preprocessing methods on the performance of each Machine Learning algorithm and in classifying multilayered Darknet Traffic Classification.

By investigating the efficiency of machine learning algorithms in multilayered classification of darknet network traffic, the research aims to contribute to the development of more robust and accurate methods for detecting and classifying malicious activities

within darknet environments. This can aid in enhancing network security, improving cyber threat intelligence, and facilitating proactive measures to combat cybercrime.

Overall, the research explores the potential of machine learning techniques in addressing the challenges associated with analyzing and classifying complex darknet network traffic, with the goal of advancing the field of network security and facilitating effective countermeasures against cyberthreats.

## 1.1 OBJECTIVES

The objectives of the project revolve around evaluating, designing, optimizing, and validating machine learning algorithms and multilayered classification approaches for darknet network traffic. By achieving these objectives, the project aims to contribute to the development of more efficient and accurate systems for detecting and classifying malicious activities within darknet environments.

The study aims to evaluate the performance and effectiveness of various machine learning algorithms in accurately classifying different types of darknet network traffic. This may involve the use of supervised learning algorithms, such as decision trees, support vector machines, neural networks, or ensemble methods like random forests or gradient boosting. The efficiency of the algorithms is assessed based on metrics such as accuracy, precision, recall, and F1 score, which measure the effectiveness of classification in identifying different types of network traffic accurately. The research may also explore the impact of different features, feature selection techniques, or preprocessing methods on the performance of the algorithms and finds the performance of each Machine Learning algorithm in classifying multilayered Darknet Traffic Classification.

## 1.2 INTRODUCTION

IoT and other communication technologies have dramatically improved our ability to comprehend our environment. Life quality may be improved through the use of IoT technologies, which have the potential to gather and analyze data about the surrounding environment [1]. This circumstance facilitates the development of smart cities by making it easier for things and humans to communicate with each other. There were an estimated 50 billion Internet of Things (IoT) devices by the end of 2020 [2,3].

The IoT is a sophisticated and interconnected system. As a result, it is difficult to meet the security requirements of an IoT system with a large attack surface. The widespread use of IoT has had the unintended consequence of making IoT deployment an interconnected process. There are a number of considerations to keep in mind while deploying an IoT system: security, energy efficiency, analytics approaches, and interoperability with other software applications [3]. The IoT devices, on the other hand, often operate in the absence of a human operator. These devices can therefore be physically accessed by an intruder. Intruders can gain access to private information via eavesdropping on wireless networks used by IoT devices, which are often connected by a communication channel.

As the IoT area continues to evolve, defining a reference architecture that can accommodate both present functionality and future enhancements will be a significant task. As a result, such an architecture must be: scalable, in order to handle a rising number of devices and services without compromising their performance; interoperable, so that devices from different manufacturers may collaborate to accomplish shared objectives; distributive, in order to enable the development of a distributed environment in which data are processed by different entities in a distributed manner after being acquired from various sources; and capable of operating with minimum resources [4].

There is currently no single reference architecture, and building one is proving difficult despite several standardization initiatives. The fundamental issue is the inevitable fragmentation of possible applications, each of which is dependent on a plethora of frequently disparate factors and design standards. This issue must be combined with each supplier's desire to promote its platform for comparable applications [4,5]. Figure 1 depicts

some of the most often seen Internet of Things architectures

Most of us are familiar with the Internet and the World Wide Web (WWW, or web). We regularly access both using web browsers or other networked applications to share information publicly, guided by search engine indexing of the Domain Name System (DNS) over globally bridged Internet Protocol (IP) networks. This publicly accessible and indexed address space is known as the surface web or clearnet. In contrast, the WWW address space which is not indexed by search engines but still publicly accessible is known as the deep web. Private networks within the deep web or networks comprised of unallocated address space are known as darknets and collectively termed the dark web. Figure 1 illustrates the relationship between these layers of the Internet.

The dark web is reached by an overlay network requiring special software, user authorization, or non-standard communication protocols [4]. Many darknets afford users anonymity during communication and thus facilitate many criminal activities, including hacking, media piracy, terrorism, trading illegal goods, human trafficking, and child pornography [2, 20]. Researchers are illuminating darknet traffic with machine learning and deep learning techniques, to better identify and inhibit these criminal activities. This research strives to contribute by promoting accurate classification of traffic features from the well-studied CIC-Darknet2020 [12] dataset. CICDarknet2020 is a collection of traffic features from two darknets, namely The Onion Router (Tor) and a Virtual Private Network (VPN), and also equivalent traffic generated over clearnet sessions using the same applications.

We use classic machine learning techniques, such as Support Vector Machines (SVM) and Random Forest (RF) for classification of the network and application type. We also represent traffic features as grayscale images and apply deep learning architectures as classifiers, including Convolutional Neural Networks (CNN) and Auxiliary-Classifier Generative Adversarial Networks (AC-GAN). To assess the issue of extreme class imbalance within CIC-Darknet2020, we explore the option of dataset augmentation, assessing the Synthetic Minority Oversampling Technique (SMOTE) on the performance of our classifiers. Our results show that RF is the most effective at classifying both traffic type and the underlying application types. Having established baseline classification performance, we consider the confusion of our best classifier in some detail, approaching the problem of darknet traffic detection adversarially.

From the perspective of an attacker we obfuscate the application classes in an attempt to avoid detection. We apply an encoding scheme to transform class features using probability analysis of the CIC-Darknet2020 dataset as a proof-of-concept. We strongly correlate the resulting RF confusion with our obfuscation technique for two attack scenarios, assuming few realistic limitations for traffic modification. We then assess the strength of our obfuscation technique with one defense scenario, by which we demonstrate that we can restore the performance of the RF classifier despite duress. We find that sufficient statistical knowledge of network traffic features can empower either classification or obfuscation tasks.

## II. VLITERATURE SURVEY

### 1. Qingya Yang, Peipei Fu, Junzheng Shi, Bingxu Wang, Zhen Li, Gang Xiong- Multi-Feature Fusion Based Approach for Classifying Encrypted Mobile Application Traffic. (IEEE - 2023)

The paper proposes a novel multi-feature fusion (MFF)- based approach to enhance the accuracy of mobile application traffic classification. The authors extracted packet length sequence, byte sequence, statistical feature, etc. Then, they performed weighted fusions of features based on Relief-F algorithm to achieve the best set of features. Finally, they used machine learning techniques for application classification. In order to obtain the best fusion features a comprehensive analysis on several basic features and perform the weighted fusion of features from different categories by Relief - F algorithm for achieving highly discriminative features. The experiments prove that the effects of fusion features are

better than those of single features in encrypted traffic classification. Then they compared 4 approaches on self-collected dataset. Compared to several other feature extraction methods, MFF achieves an excellent performance with an accuracy of 97.6% for 16 mobile applications and a F1-score of over 99% for VPN- nonVPN.

2. **Hasan Karagol, Oguzhan Erdem, Barkin Akbas, Tuncay Soylu - Darknet Traffic Classification with Machine Learning Algorithms and SMOTE Method. (IEEE - 2022)**

The paper proposes three different Machine Learning (ML) based traffic classification approaches; the binary classification of Darknet and Benign traffic classes (Case1); the quadruple classification of classes Tor, Non-Tor, VPN, Non-VPN (Case2); a traffic classification of eight sub-traffic classes (Case3). Further the proposed paper uses SMOTE method for balancing the sizes of the classes in the traffic dataset and feature selection (FS) algorithm to identify the most effective attributes. For all three cases, classification was performed with six different machine learning algorithms with and without SMOTE and the highest accuracy values were obtained with SMOTE method. The accuracy was 97.22%, 97.16%, 85.99% for cases 1, 2 and 3 respectively.

3. **Zhe Wang, BaiHe Ma, Yong Zeng, XiaoJie Lin, KaiChao Shi, ZiWen Wang – Differential Preserving in XGBoost Model for Encrypted Traffic Classification**

The paper introduced XGBoost model is based on the residual to train the model to fit the real datascene, perform the efficient calculation of the gradient histogram, and realize the Boosting Tree of extreme parallel computing. Classification of the encrypted data into VPN traffic and Non-VPN traffic with the XGBoost model. Then the categorized encrypted traffic is further classified into 14 types of traffic. the row traffic is protected in differential preserving. The proposed model achieves high accurate classification precision. With the differential preserving protected dataset the model does not need to decrypt traffic, and does not diminish traffic privacy. Experiments areperformed on the public dataset ISCX VPN-nonVPN, and the results show that the XGBoost model has an accuracy of 92.4%.

4. **Tianhua Chen, Elans Grabs, Ernests Petersons, Dmitry Efrosinin, Aleksandrs Ipatovs, Janis Kluga - Encapsulated and anonymized network video traffic classification with generative models. (IEEE – 2022)**

Generative machine learning models to classify video streaming category traffic encapsulated by OpenVPN, video streaming category traffic captured in Tor networks, captured video streaming traffic. Furthermore, classification performance metrics are explored in conjunction with differentapplications. Focuses on identifying Tor traffic or VPN traffic from mixed Internet traffic and classifying it according to relevant features, such as applications or traffic categories. Most research works have achieved better classification metrics, but the shortcomings were limited to the detection and classification of traffic under one of the specific scenarios. The experimental results show the classification accuracy is as high as 0.94 for eight categories and 0.97 for three categories. Cybersecurity (CIC) gathered the ISCX VPN-nonVPN traffic dataset and the ISCX Tor-nonTor dataset was used.

5. **Shi-Jie Xu, Guang-Gang Geng, Xiao-Bo Jin, Dong-Jie Liu, Jian Weng - Seeing Traffic Paths: Encrypted Traffic Classification With Path Signature Features. (IEEE – 2022)**

The paper proposes an efficient encrypted traffic classification method named ETC-PS using only packet length information. The traffic path is constructed with a session packet length sequence torepresent the interactions between the client and server. Five path transformations are proposed to exhibit the traffic path structure and obtain different information. Finally, a multiscale path signature feature is extracted as a kind of distinctive feature to train the traditional machine

learning classifier. Experimental results showed that the accuracy and F1 score of ETC-PS achieved stable improvement on six public datasets. Six publicly available datasets with different traffic types of HTTPS/1, HTTPS/2, QUIC, VPN, nonVPN, Tor, and non-Tor are used to conduct closed-world and open-world evaluations to verify the effectiveness of ETC-PS. The experimentalresults demonstrate that ETC-PS is superior to the state-of-the-art methods in terms of accuracy, f1 score, time complexity, and stability

**6. Lazaros Alexios Iliadis, Theodoros Kaifas – Darknet Traffic Classification using Machine Learning Techniques. (IEEE - 2021)**

The paper proposes common machine learning classification algorithms that are employed to identify Darknet traffic. A ROC analysis along with a feature importance analysis for the best classifier was performed, to provide a better visualization of the results. The experiments were conducted in the new dataset CICDarknet2020 and the classifiers were trained to both binary andmulticlass classification. In the first classification task, there were two classes: ”Benign” and ”Darknet”, whereas in the second there were four classes: ”Tor”, ”Non Tor”, ”VPN” and ”Non VPN”. An average prediction accuracy of over 98% was achieved with the implementation of Random Forest algorithm for both classification tasks

**7. Abhishek Gupta– VPN- non VPN Traffic Classification Using Deep Reinforced Naive Bayes and Fuzzy K-means Clustering. (IEEE - 2021)**

The paper introduced Naive Bayes (NB) classifier and fuzzy k-means based deep reinforcement learning (DRL) model, which jointly learns the traffic characterization in VPN-non VPN dataset to classify packets as malicious or legitimate. The DRL process uses an encoder network and a decoding inference network to learn mutually coherent traffic classes in the network segments. The paper proposed a NB approach for traffic classification and characterization using DRL and validates the architecture using the publicly available UNB-CIC VPN-non VPN dataset. The original dataset manufacturers have tested the dataset using KNN and C4.5 algorithms. The accuracy was 84.83%.

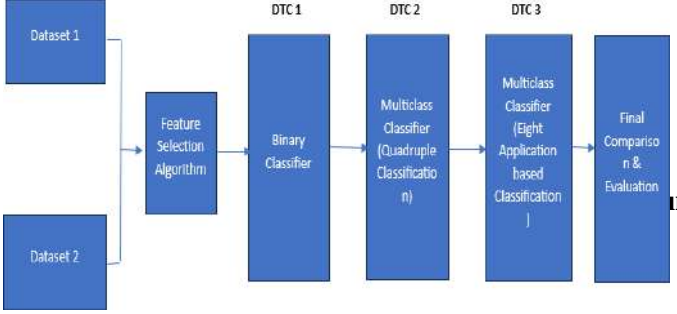
## III. PROPOSED METHOD

### 4.1 BLOCK DIAGRAM OFPROPOSED SYSTEM

**Fig 4.1 Block Diagram of the Proposed System**

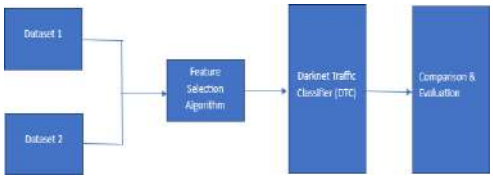
The multilayered classification approach involves multiple stages of classification, where each stage focuses on a specific aspect or characteristic of the network traffic and compares the performance of each Machine Learning algorithm in classification. The proposed system comprises of the following stages:

1. **Input Dataset 1 :** CIC-Darknet2020 dataset consisting of 63 features.
2. **Input Dataset 2 :** NICT Darknet 2022 dataset consisting of 76 features.
3. **Feature Selection Algorithm :** Fisher’s score algorithm calculates the discriminant ratio of each feature based on the fisher’s score in descending order ranking.
4. **Binary Classifier (DTC1) :** Traffic is decomposed into Darknet and Benign classes using Decision tree, Naïve Bayes, KNN, AdaBoost algorithms.
5. **Multiclass Classifier (DTC2) :** Output from DTC1 is again classified into four classes – Tor, Non-Tor, VPN, Non-VPN utilizing Random forests, Decision tree, Naïve Bayes, KNN and Gradient Boosting ML algorithms.




6. **Multiclass Classifier (DTC3) :** Output from DTC2 is again decomposed into eight application classes – Audio stream, Browsing, Chat, Email, P2P, File transfer, Video stream, VOIP using Random forest, Decision tree, Naïve Bayes, KNN and Gradient Boosting ML algorithms.

7. **Final Comparison & Evaluation :** Overall comparison on the basis of Confusion matrix to predict Accuracy, Precision, F1-score, Specificity,




Recall evaluated after DTC1, DTC2 and DTC3 stages along with overall comparison after DTC3 stage.

8. **Analysis of Machine Learning Algorithms :** The final analysis of each Machine Learning algorithm and their efficiency in classification of Darknet network traffic will be analyzed and studied to find out which Machine Learning algorithm performs better in classificationof the proposed system.

**Fig 4.1.2 Detailed Block Diagram of the Proposed Multilayered DTC Classifier System**

**Multilayered Darknet Traffic Classifier (DTC) System :**

- DTC consists of three blocks namely DTC1, DTC2, DTC3.
- DTC1 block performs Binary Classification where the traffic is decomposed into Darknet and Benign classes.
- DTC2 block performs Multiclass classification into four classes - Tor, Non-Tor, VPN, Non-VPN

DTC3 block performs Multiclass classification into eight application classes - Audio stream, Browsing, Chat, Email, P2P, File transfer, Video stream, VOIP

## IV. RESULT

Now-a-days increase usage of networks open new ways for attackers to hack or send malicious request to user system. Darknet also is one of the type attack and all existing Machine Learning algorithms were trained on single Darknet2020 CIC dataset but attackers may find new request variable to deceive existing algorithms. To overcome from this issue author of this paper merging data from two different datasets such as Darknet2020 and ISCX to detect all possible change variables attackers may use.

In propose work both datasets will be input to Fisher algorithm to select high ranking features and then train different algorithms by using different class labels listed below

1) DTC1: First entire dataset will be divided into Darknet and Benign classes where TOR and VPN labels will be consider as Darknet and NONTOR and NONVPN will be consider as Benign to form a binary class labels. Selected features and binary class labels will get trained with Decision Tree, KNN, Naïve Bayes and ADABOOST
2) DTC2 Multi classifier: DTC1 class labels will be divided into 4 different class labels such as TOR, NONTOR, VPN and NONVPN and then trained with Random Forest, Decision Tree, Naïve Bayes, KNN and Gradient Boosting
3) DTC3 Multi classifier: here entire dataset will be split into 8 different traffic categories and trained with 5 different algorithms such as Random Forest, Decision Tree, Naïve Bayes, KNN and Gradient Boosting

Propose algorithm training on two different datasets able to get high accuracy compare to existing single dataset trained algorithms.

We have coded this project using JUPYTER NOTEBOOK and below are the code and output screens with blue color comments

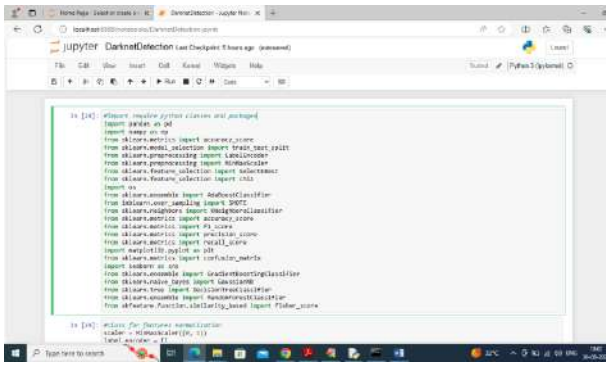


**Fig.5.1 Importing require python classes and packages**

In above screen importing require python classes and packages

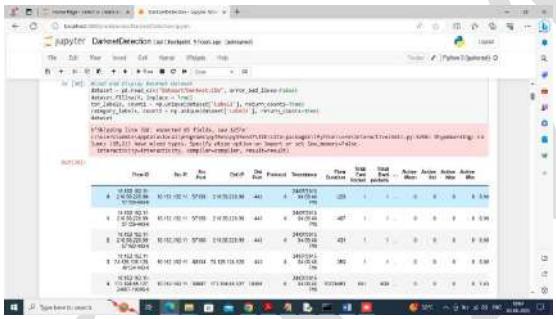


**Fig.5.2 loading and displaying Dartket2020 dataset**

In above screen loading and displaying Dartket2020 dataset to training existing algorithms

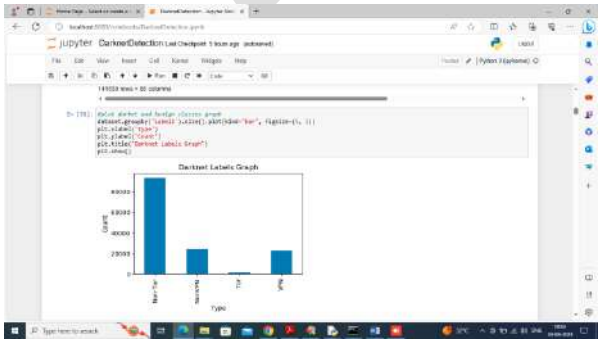


**Fig.5.3 graph of different Darknet**

In above screen plotting graph of different Darknet and benign classes graph where x-axis represents names and y-axis represents count

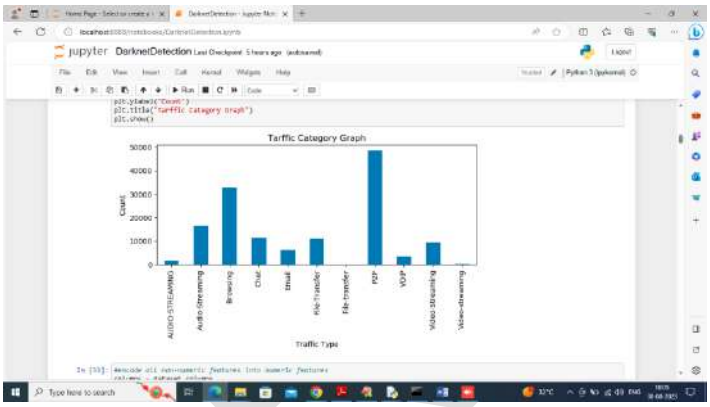


**Fig.5.4. plotting graph of different traffic category**

In above screen plotting graph of different traffic category found in dataset where x-axis represents traffic category and y-axis represents count

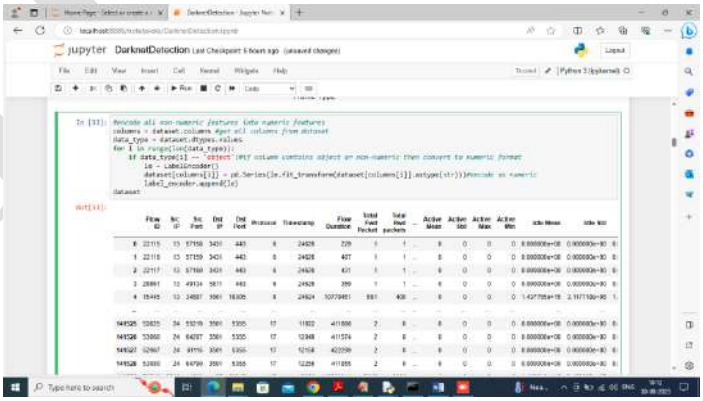


**Fig.5.5 converting all non-numeric columns to numeric columns**

In above screen converting all non-numeric columns to numeric columns

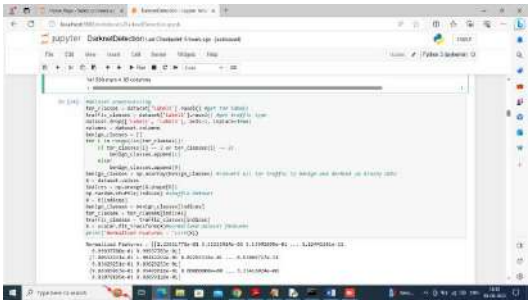


**Fig.5.6 applying pre-processing techniques**

In above screen applying pre-processing techniques such as extracting X and y labels and then shuffling and normalizing dataset values

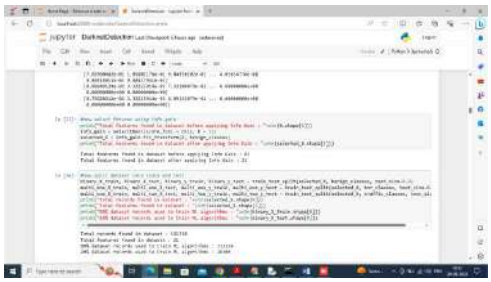


**Fig.5.7. applying existing technique**

In above screen for existing technique applying Information Gain to select relevant features and then selected 21 best features and then splitting dataset into train and test where application using 80% dataset for training and 20% for testing

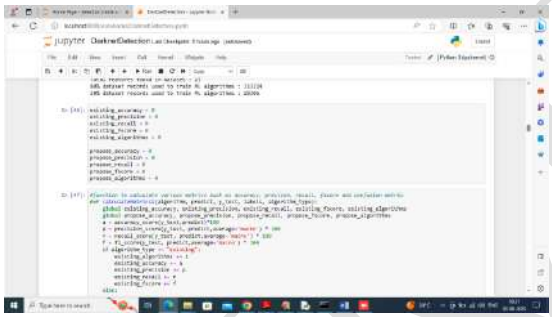


**Fig.5.8 calculation of accuracy**

In above screen defining function to calculate accuracy and other metrics for both existing algorithms and propose algorithms

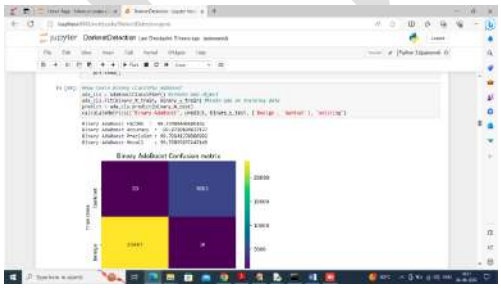


**Fig.5.9 Train ADABOOST**

In above screen training ADABOOST on binary classes using existing single dataset and then ADABOOST got 99.77% accuracy and can see other metrics also and in confusion matrix graph x-axis represents Predicted Labels and y-axis represents true labels and both blue color boxes contains incorrect prediction count which are very few and different color boxes like yellow and other color box contains correct prediction count

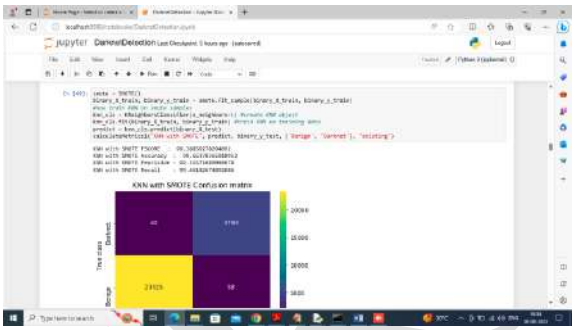


**Fig.5.10 Training existing KNN wit SMOTE**

In above screen training existing KNN wit SMOTE and got accuracy as 99.18% and can see other metrics also

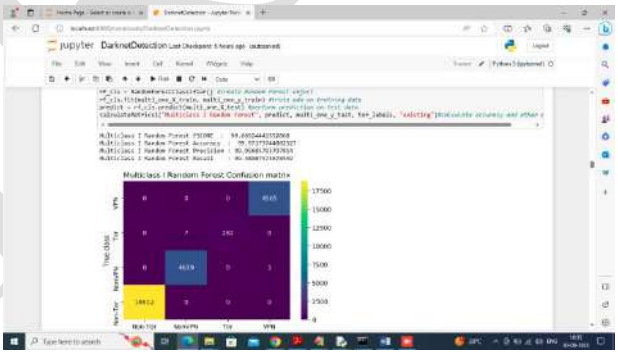


**Fig.5.11 Training existing Random Forest**

In above screen training existing Random Forest on 4 classes as Multi class I dataset and got accuracy as 99.66%

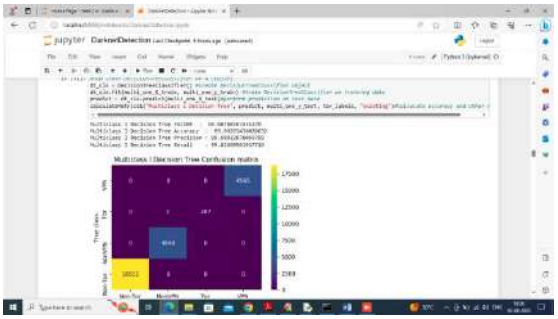


**Fig.5.12 Train decision tree**

In above screen training decision tree and similarly we trained all existing algorithms on single existing

ISSN: 2249-7196
IJMRR/April. 2025/ Volume 15/Issue 2s/164-175


dataset and now in below screen showing propose algorithms by combining two different datasets

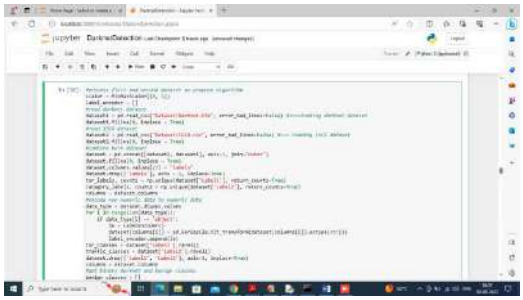


**Fig.5.13 loading both Darknet2020 and ISCX dataset**

In above screen for propose work we are loading both Darknet2020 and ISCX dataset and then applying preprocessing techniques

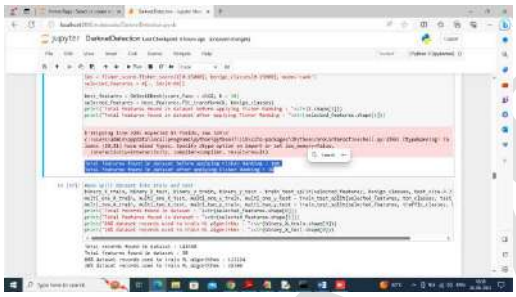


**Fig.5.14 applying FISHER algorithm**

In above screen from propose dataset applying FISHER algorithm to select 30 best features and then splitting into train and test

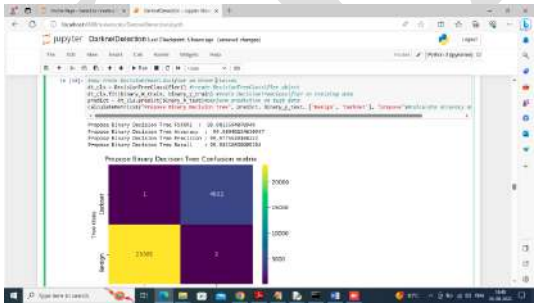


**Fig.5.15 training propose Decision tree on both dataset**

In above screen training propose Decision tree on both dataset features and got accuracy as 99.98% and can see other metrics also

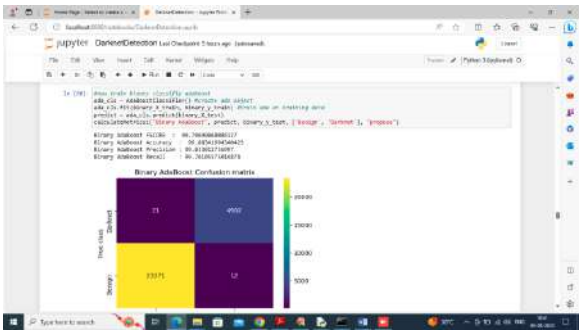


**Fig.5.16 training propose ADABOOST**

In above screen training propose ADABOOST on binary classes and got accuracy as 99.79%

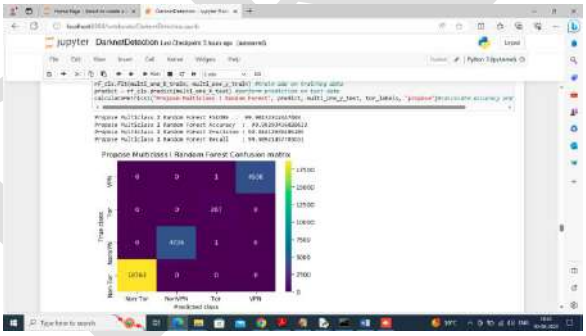


**Fig.5.17 training Random Forest on propose dataset**

In above screen training Random Forest on propose dataset with 4 classes as TOR, NONTOR, VPN and NONVPN and got accuracy as 99.90%

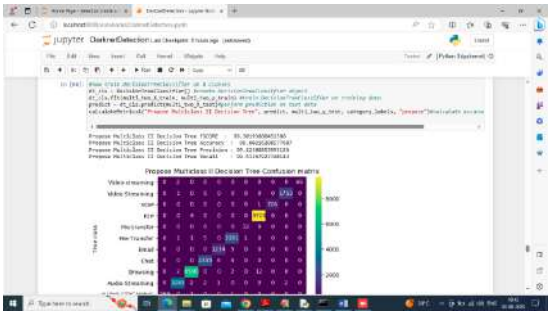


**Fig.5.18 training Decision Tree 8 different classes**

In above screen training Decision Tree 8 different classes as MULTICLASS II and got accuracy as 99.30% and can see other metrics also. Similarly we trained all algorithms from propose and existing LIST. Below is the overall accuracy of all algorithms on existing and propose work

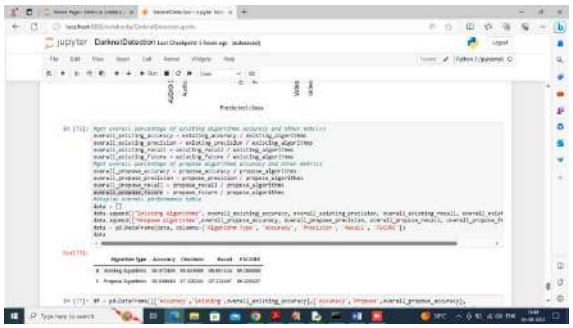


**Fig.5.19 overall accuracy**

In above screen in tabular format we can see overall accuracy and other metrics from both propose and existing algorithms and in both PROPOSE algorithms got high accuracy

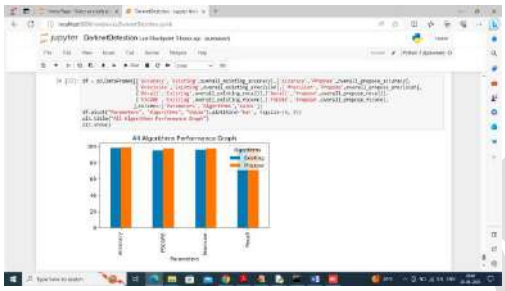


**Fig.5.20 displaying existing and propose performance in graph**

In above screen displaying existing and propose performance in graph format where x-axis represents metrics and blue bar represents existing metrics and orange bar represents propose metrics and from above results we can say combining and selecting features from multiple datasets can increase ML performance

## V. CONCLUSION

To conclude, Machine learning-based classification of Multilayered darknet network traffic offers significant advantages in the field of cybersecurity and network monitoring. The application of machine learning algorithms enables accurate identification and classification of various types of activities within the darknet, including malicious behavior, illegal marketplaces, and potential threats. This approach enhances threat detection, proactive defense, and risk assessment capabilities, empowering organizations to better protect their networks and assets.

By leveraging machine learning algorithms, researchers and cybersecurity professionals have made notable progress in understanding the patterns and characteristics of darknet network traffic. The development of labeled datasets and the exploration of various machine learning models have provided valuable insights into the nature of darknet activities, aiding in cybercrime investigations, threat intelligence, and anomaly detection.

## REFERENCES

1. Q. Yang, P. Fu, J. Shi, B. Wang, Z. Li and G. Xiong, "Multi-Feature Fusion Based Approach for Classifying Encrypted Mobile Application Traffic," 2023 26th International Conference on Computer Supported Cooperative Work in Design (CSCWD), Rio de Janeiro, Brazil, 2023, pp. 1112-1117, doi: 10.1109/CSCWD57460.2023.10152687.

2. H. Karagöl, O. Erdem, B. Akbas and T. Soylu, "Darknet Traffic Classification with Machine Learning Algorithms and SMOTE Method," 2022 7th International Conference on Computer Science and Engineering (UBMK), Diyarbakir, Turkey, 2022, pp. 374-378, doi: 10.1109/UBMK55850.2022.9919462.

3. Z. Wang, B. Ma, Y. Zeng, X. Lin, K. Shi and Z. Wang, "Differential Preserving in XGBoost Model for Encrypted Traffic Classification," 2022 International Conference on Networking and Network Applications (NaNA), Urumqi, China, 2022, pp. 220-225, doi: 10.1109/NaNA56854.2022.00044.

4. T. Chen, E. Grabs, E. Petersons, D. Efrosinin, A. Ipatovs and J. Kluga, "Encapsulated and Anonymized Network Video Traffic Classification With Generative Models," 2022 Workshop on Microwave Theory and Techniques in Wireless Communications (MTTW), Riga, Latvia, 2022, pp. 13-18, doi: 10.1109/MTTW56973.2022.9942564.

5. S. -J. Xu, G. -G. Geng, X. -B. Jin, D. -J. Liu and J. Weng, "Seeing Traffic Paths: Encrypted Traffic Classification With Path Signature Features," in IEEE Transactions on Information Forensics and Security, vol. 17, pp. 2166-2181, 2022, doi:

10.1109/TIFS.2022.3179955.

6. L. A. Iliadis and T. Kaifas, "Darknet Traffic Classification using Machine Learning Techniques," 2021 10th International Conference on Modern Circuits and Systems Technologies (MOCAST), Thessaloniki, Greece, 2021, pp. 1-4, doi: 10.1109/MOCAST52088.2021.9493386.

7. A. Gupta, "VPN-nonVPN Traffic Classification Using Deep Reinforced Naive Bayes and Fuzzy K-means Clustering," 2021 IEEE 41st International Conference on Distributed Computing Systems Workshops (ICDCSW), Washington, DC, USA, 2021, pp. 1-6, doi: 10.1109/ICDCSW53096.2021.00008.

8. W. Cai, G. Gou, M. Jiang, C. Liu, G. Xiong and Z. Li, "MEMG: Mobile Encrypted Traffic Classification With Markov Chains and Graph Neural Network," 2021 IEEE 23rd Int Conf on High Performance Computing & Communications; 7th Int Conf on Data Science & Systems; 19th Int Conf on Smart City; 7th Int Conf on Dependability in Sensor, Cloud & Big Data Systems & Application (HPCC/DSS/SmartCity/DependSys), Haikou, Hainan, China, 2021, pp. 478-486, doi:10.1109/HPCC-DSS-SmartCity-DependSys53884.2021.00087.

W. Maonan, Z. Kangfeng, X. Ning, Y. Yanqing and W. Xiujuan, "CENTIME: A Direct Comprehensive Traffic Features Extraction for Encrypted Traffic Classification," 2021 IEEE 6th International Conference on Computer and Communication Systems (ICCCS), Chengdu, China, 2021, pp. 490-498, doi: 10.1109/ICCCS52626.2021.9449280.

9. S. A. Aswad and E. Sonuç, "Classification of VPN Network Traffic Flow Using Time Related Features on Apache Spark," 2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), Istanbul, Turkey, 2020, pp. 1-8, doi: 10.1109/ISMSIT50672.2020.9254893.

10. G. Baldini, "Analysis of Encrypted Traffic with time-based features and time frequency analysis," 2020 Global Internet of Things Summit (GIoTS), Dublin, Ireland, 2020, pp. 1-5, doi: 10.1109/GIOTS49054.2020.9119528.

11. G. Bovenzi, G. Aceto, D. Ciuonzo, V. Persico and A. Pescapé, "A Big Data-Enabled Hierarchical Framework for Traffic Classification," in IEEE Transactions on Network Science and Engineering, vol. 7, no. 4, pp. 2608-2619, 1 Oct.-Dec. 2020, doi: 10.1109/TNSE.2020.3009832.

12. V. A. Muliukha, L. U. Laboshin, A. A. Lukashin and N. V. Nashivochnikov, "Analysis and Classification of Encrypted Network Traffic Using Machine Learning," 2020 XXIII International Conference on Soft Computing and Measurements (SCM), St. Petersburg, Russia, 2020, pp. 194-197, doi: 10.1109/SCM50615.2020.9198811.

13. G. Aceto, D. Ciuonzo, A. Montieri, V. Persico and A. Pescapé, "Know your Big Data Trade-offs when Classifying Encrypted Mobile Traffic with Deep Learning," 2019 Network Traffic Measurement and Analysis Conference (TMA), Paris, France, 2019, pp. 121-128, doi: 10.23919/TMA.2019.8784565.

14. P. Wang, X. Chen, F. Ye and Z. Sun, "A Survey of Techniques for Mobile Service Encrypted Traffic Classification Using Deep Learning," in IEEE Access, vol. 7, pp. 54024-54033, 2019, doi: 10.1109/ACCESS.2019.2912896