A Novel Approach for Intrusion Detection in 5G Networks using Deep Learning

1st Amit Noel Thokala
SCAI – Software Engineering (MS)
Arizona State University
Tempe, USA
athokal1@asu.edu

2nd Sai Rithvik Vaikuntam
SCAI – Software Engieering (MS)
Arizona State University
Tempe, USA
svaikunt@asu.edu

3rd Harsha Vardhan Yallavula
SCAI – Software Engineering(MS)
Arizona State University
Tempe, USA
hyallavu@asu.edu

4th Hari Priya Gottam
SCAI – Software Engineering (MS)
Arizona State University
Tempe, USA
hgottam@asu.edu

5th Vishwanath Reddy Yasa
SCAI – Software Engineering (MS)
Arizona State University
Tempe, USA
vyasa@asu.edu

6th Yashwanth Reddy Kikkuri
SCAI – Software Engineering (MS)
Arizona State University
Tempe, USA
ykikkuri@asu.edu

***Abstract*—This research paper proposes a novel approach to intrusion detection systems specifically designed for 5G mobile networks. The paper acknowledges the escalating threat landscape of cyber-attacks in 5G networks due to the proliferation of Internet of Things (IoT) devices. It emphasizes the importance of robust IDS to safeguard network integrity, ensure data confidentiality, and mitigate potential security breaches. The research identifies key challenges such as data preprocessing, feature engineering, model selection, and scalability. To address these challenges, the paper proposes a multi-pronged approach that includes data preprocessing techniques, machine learning algorithms, and voting classifier models. The approach is evaluated on the 5G-NIDD dataset, achieving high accuracy in both binary and multi-class classification tasks. The results demonstrate the effectiveness of the proposed approach in detecting various attack types, including HTTPFlood, ICMPFlood, SYNFlood, SYNScan, SlowrateDoS, TCPConnectScan, UDPFlood, and UDPScan. Overall, the paper contributes to the field of cybersecurity by proposing a comprehensive Intrusion Detection Systems solution for 5G mobile networks.**

I. Introduction

With the rapid proliferation of digital technologies and the increasing interconnectivity of devices, the security landscape of modern networks has become increasingly complex and vulnerable to cyber threats. In particular, the advent of 5G networks and the widespread adoption of Internet of Things (IoT) devices have introduced new challenges and heightened the importance of effective intrusion detection and prevention systems. The escalating threat landscape of cyber-attacks underscores the critical need for robust security measures, particularly in the context of 5G networks. These networks, characterized by high-speed connectivity and low-latency communication, serve as the backbone of various critical infrastructures, including e-commerce, healthcare, and communication systems. However, their inherent vulnerabilities, coupled with the evolving nature of cyber threats, pose significant risks to data privacy, network integrity, and service availability.

In response to these challenges, researchers and practitioners have increasingly turned to machine learning and deep learning techniques to develop advanced intrusion detection systems capable of effectively identifying and mitigating cyber threats in 5G networks. By leveraging the power of data-driven analytics and predictive modeling, these systems aim to enhance network security, preemptively detect anomalous behavior, and mitigate the impact of cyber-attacks. In this paper, we present a comprehensive research study focused on the development and evaluation of machine learning-based intrusion detection systems for 5G networks. Our research encompasses various aspects of the IDS lifecycle, including data preprocessing, feature engineering, model selection, and performance evaluation. We propose novel methodologies and algorithms tailored to the unique characteristics of 5G networks, with the aim of achieving high detection accuracy, low false positive rates, and real-time responsiveness.

II. Problem Definition

The problem at hand revolves around the development of robust intrusion detection systems tailored specifically for 5G mobile networks and beyond. With the proliferation of advanced cyber threats and the increasing complexity of network infrastructures, there is a pressing need for sophisticated detection mechanisms capable of identifying and mitigating malicious activities in real-time. The overarching objective is to safeguard network integrity, ensure data confidentiality, and mitigate potential security breaches in the dynamic and evolving landscape of 5G mobile networks.

One of the primary challenges lies in data preprocessing and feature engineering. Given the intricacies of network traffic data, preprocessing and feature engineering are critical steps in extracting relevant information and preparing it for subsequent analysis. Addressing issues such as missing values, outliers, and categorical variables is paramount to ensure the quality and reliability of the data used for intrusion detection.

Another challenge involves selecting and evaluating appropriate models for intrusion detection. With a myriad of machine learning algorithms and voting classifier techniques available, identifying the most suitable models poses a significant challenge. Evaluating the performance of these models in terms of accuracy, precision, recall, and computational efficiency is essential to identify optimal solutions capable of effectively detecting and classifying network intrusions.

Additionally, scalability and adaptability are key considerations in developing intrusion detection systems for 5G mobile networks. As network infrastructures continue to evolve, intrusion detection systems must be scalable and adaptable to accommodate increasing data volumes and changing network dynamics. Solutions should be capable of operating in real-time, providing timely alerts and responses to potential security breaches.

By delineating the problem space and elucidating key challenges, our research aims to advance the state-of-the-art in intrusion detection for 5G mobile networks, paving the way for more resilient and effective cybersecurity measures in the digital era.

III. Related Work

A. An efficient deep learning model for intrusion classification and prediction in 5G and IoT networks

The paper underscores the escalating threat landscape of cyber-attacks, particularly with the pervasive use of the Internet of Things (IoT) in e-commerce, healthcare, and communication systems. As wireless network traffic is predicted to dominate Internet usage, the vulnerability of these networks becomes a critical concern, especially with the continual evolution of technologies like 5G and IoT. The literature review focuses on the burgeoning research in intrusion detection systems (IDS) and the adoption of deep learning techniques for enhanced security. The study introduces a hybrid intrusion detection system designed for low latency and high accuracy, employing an unsupervised pre-training autoencoder for feature extraction and a supervised dense neural network for classification. The evaluation, based on the Aegean Wi-Fi Intrusion Dataset (AWID), demonstrates the models effectiveness, with particular emphasis on its robust performance in detecting specific attack types. Data pre-processing, model architecture, and optimization techniques during the training phase are detailed. The authors compare their model's performance favorably against other machine learning techniques, highlighting its efficiency in feature selection. Future work is suggested to further improve detection accuracy, minimize false positives and negatives, and extend the model's applicability to a broader range of attacks, particularly in mobile and IoT security platforms. [1].

B. Wireless Intrusion and Attack Detection

Paper proposes a new system to detect cyberattacks in 5G networks, specifically focusing on "Wireless Intrusion Detection Systems" (WIDS). Machine learning, particularly deep learning, is highlighted as a potent solution due to its capacity to decipher intricate behaviors, a pivotal attribute in effective network intrusion detection. The literature extensively surveys prior research, encompassing diverse techniques such as Random Forest, Adaboost, Decision Tree, k-Nearest Neighbor, Artificial Neural Networks (ANN), Long Short-Term Memory (LSTM), and Convolutional Neural Networks (CNN). Notably, studies leveraging deep learning, including Autoencoder, DNN, CNN, ANN, and LSTM, exhibit promising outcomes, showcasing their efficacy in achieving precise intrusion detection in the intricate 5G network landscape [2].

C. Research of Machine Learning Algorithms for the Development of Intrusion Detection Systems in 5G Mobile Networks and Beyond

This study delves into the utilization of machine learning and deep learning methodologies for network intrusion detection systems (NIDS) designed for 5G networks [3]. Employing the CIC-IDS-2017 and CSE-CIC-IDS-2018 datasets, the research unfolds across three segments. In the initial phase, the Gradient Boost algorithm demonstrates robust efficacy but grapples with issues related to misclassification in specific attack classes. The second explores unsupervised learning with autoencoder architectures, revealing limitations in achieving sufficient performance compared to traditional machine learning models. In the final phase, a supervised learning approach employing deep neural network (DNN) models attains noteworthy outcomes. However, challenges emerge in accurately distinguishing penetration attacks due to statistical resemblances between benign and penetrating traffic. The research concludes with a comparative analysis, naming the Gradient Boosted Tree model as the most effective intrusion detection in 5G networks.

D. Real-Time Intrusion Detection and Prevention System for 5G and beyond Software-Defined Networks

This paper introduces an innovative Intrusion Detection and Prevention System (IDS/IPS) leveraging a Convolutional Neural Network (CNN) model [4]. The proposed system excels in real-time processing, vital for 5G's high-speed requirements. The CNN model, employed for preprocessing and analysis, demonstrates high accuracy in identifying both known and unknown threat patterns. The architecture ensures scalability, and performance metrics, including accuracy, highlight the system's efficiency. Additionally, the paper introduces the 5G-NIDD dataset, contributing significantly to 5G network security research.

IV. Proposed Approach

A. Data Preprocessing

In any data-driven research or application, the preprocessing stage plays a crucial role in ensuring the quality and reliability of subsequent analyses and models. As part of our research, we conducted extensive data preprocessing, which encompassed a series of tasks aimed at cleaning, transforming, and preparing raw data for analysis and modeling. This included handling missing values, removing outliers, scaling features, encoding

categorical variables, and splitting data into training and testing sets. Additionally, we employed advanced preprocessing techniques such as feature engineering and dimensionality reduction to enhance the performance of our machine learning models. Effective data preprocessing not only improves model accuracy but also facilitates the extraction of meaningful insights from the data, thus underscoring its significance in the research process.

B. Data Visualization

Data visualization serves as a powerful tool in the data analysis process, enabling researchers to gain valuable insights and communicate findings effectively. Through the use of graphs, charts, and other visual representations, complex datasets can be transformed into intuitive visualizations that facilitate comprehension and decision-making. In our research, we employed various data visualization techniques to explore patterns, trends, and relationships within the data. This included creating scatter plots, histograms, and heatmaps to visualize distributions and correlations, as well as employing interactive visualization tools for dynamic exploration. By leveraging data visualization, we not only enhanced our understanding of the underlying data but also communicated our findings in a clear and impactful manner, thereby enriching the research process and facilitating knowledge dissemination.

1) Heatmap for Attack Type vs Attack Tool: : In Fig. 1. The heatmap showcases the relationship between attack types and the corresponding attack tools employed in cyber intrusions. By visually representing this data, we gain insights into the association between different attack types and the specific tools utilized for executing those attacks. This visualization aids in understanding the diversity and sophistication of cyber threats, facilitating targeted defense strategies and resource allocation for mitigating risks.

2) Pie Chart Distribution for Attack Types: : In Fig. 2. Pie chart offers a succinct overview of the distribution of various attack types within a dataset. It provides a clear representation of the proportion of each attack type relative to the total, aiding in understanding the prevalence and significance of different types of cyber threats. This visualization is instrumental in identifying dominant attack vectors, enabling cybersecurity professionals to prioritize threat response and prevention efforts effectively.

3) Bar Chart for Protocol Distribution: : In Fig. 3. Bar chart illustrates the distribution of network protocols observed in cyber traffic. By visualizing the frequency of each protocol, we gain insights into the communication patterns and predominant protocols used in network interactions. This visualization is crucial for network administrators and security analysts to monitor and manage network traffic efficiently, ensuring the integrity, availability, and security of network resources.

4) Scatter Plot for Duration vs Total Bytes: : In Fig. 4. The scatter plot depicts the relationship between the duration of network activities and the total bytes transferred during those interactions. By examining this relationship, we can discern patterns related to the efficiency and intensity of network communications. Understanding how duration correlates with total bytes transferred can inform network optimization strategies and aid in anomaly detection, helping to identify potentially malicious or suspicious activities.

5) Heatmap for Distribution of Each Type of Attack: : In Fig. 5. This heatmap provides a comprehensive overview of the distribution of different attack types across various dimensions, such as time, targets, and attack vectors. By visualizing the frequency and intensity of each attack type, we can identify trends, patterns, and hotspots of malicious activities. This visualization facilitates proactive threat intelligence and incident response, enabling organizations to bolster their defenses and mitigate cybersecurity risks effectively.

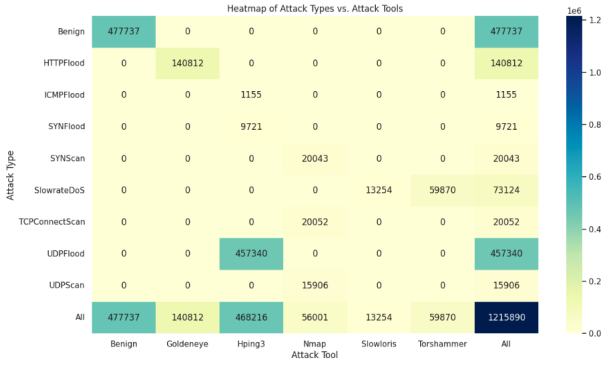

Fig. 1. Heatmap for Attack Type Vs Attack Tool

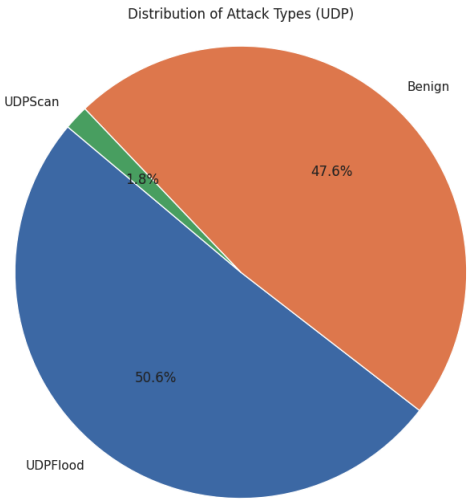

Fig. 2. Piechart for Distribution of Attack types

C. Research of Machine Learning Algorithms for the Development of Intrusion Detection Systems in 5G Mobile Networks and Beyond

In this section, we present the methodology adopted for intrusion detection in 5G networks using machine learning techniques. The proposed approach consists of implementing four distinct models, each leveraging different machine

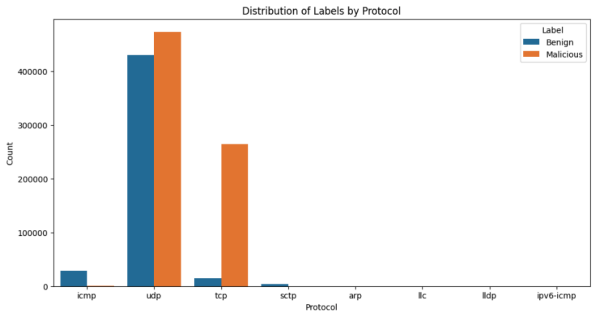

Fig. 3. Protocol Distribution Bar Chart

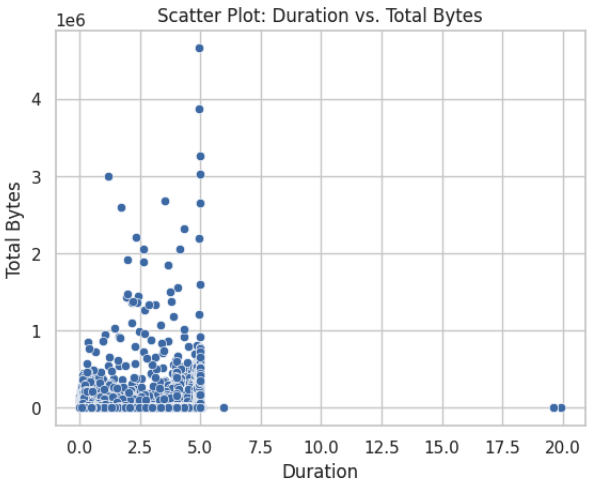

Fig. 4. Scatter Plot Duration Vs Total Bytes

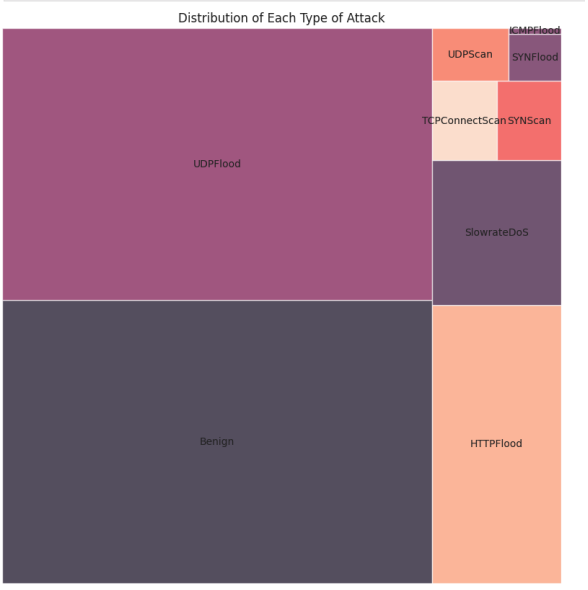

Fig. 5. Distribution of Each Type of Attack

learning algorithms to detect and classify network intrusions effectively.

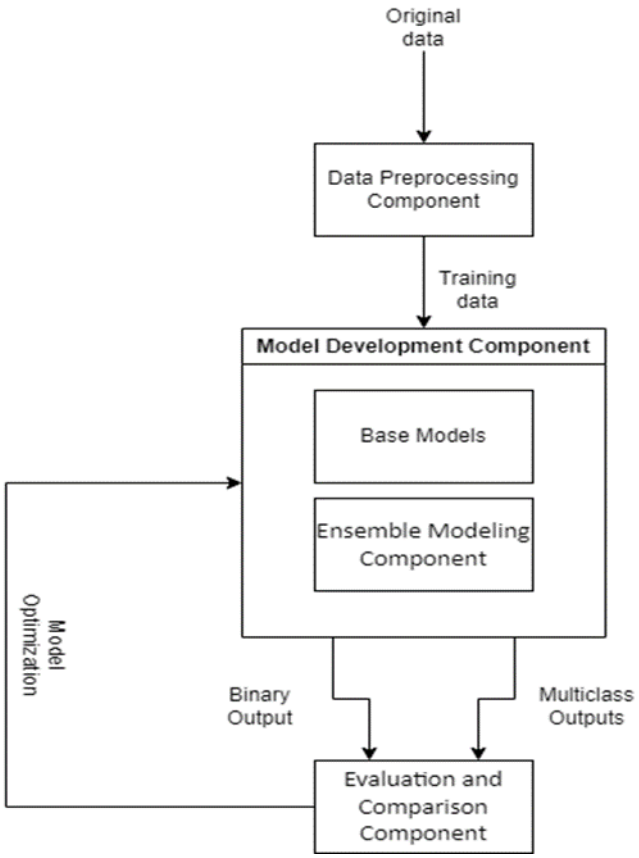

Fig. 6. High level architecture

1) Voting classifier Model: The first model implemented is an Voting classifier model, which aggregates the predictions of multiple base models using both soft and hard voting classifiers. The base models utilized in this technique include Decision Tree, Random Forest, Multi-layer Perceptron, Gradient Boosting Machine (GBM), K-Nearest Neighbors (KNN), and Logistic Regression. We implemented both soft voting and hard voting scheme for the different base models. In the soft voting scheme, the Voting classifier model combines the predicted probabilities from each base model and selects the class with the highest average probability as the final prediction. This approach allows the classifier to take into account the confidence levels of individual models in making predictions. On the other hand, the hard voting scheme combines the class labels predicted by each base model and selects the class with the most votes as the final prediction. This approach operates on the principle of majority voting and can be effective even when individual models have varying levels of accuracy [5].

2) Boosting-LSTM Model: The implemented model comprises two main components: a Boosting classifier for binary classification of network traffic and an LSTM model for predicting the attack type of malicious traffic identified by the Boosting classifier.

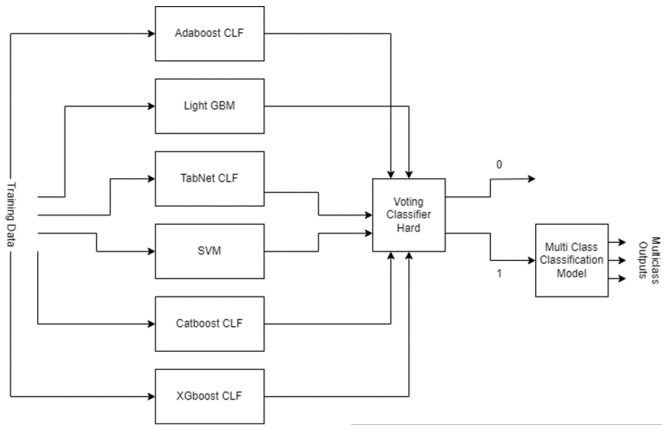

Fig. 7. Voting model architecture

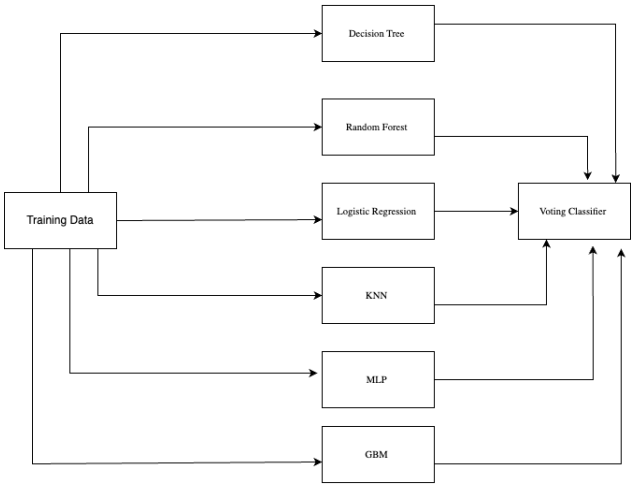

Fig. 8. Boosting model architecture

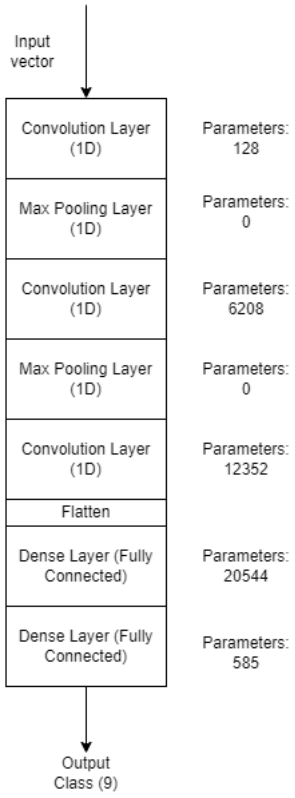

Fig. 9. Multiclass CNN Classifier model architecture

Boosting, a robust ensemble learning technique, amalgamates the predictions of multiple base boosting models to make informed decisions. In our approach, diverse weak learners such as XGBoost, CatBoost, LightGBM, and AdaBoost contribute their predictions to the primary model, thereby enhancing its performance.

This model operates on the premise that benign traffic is legitimate and should be forwarded, while malicious traffic requires further scrutiny. When the Boosting classifier identifies traffic as malicious, the LSTM model steps in to predict the attack type, facilitating the classification of the threat and preventing the forwarding of potentially harmful traffic.

By leveraging the collaborative strengths of Boosting and LSTM, our model effectively addresses the binary classification of network traffic and offers granular insights into the nature of potential attacks, thereby bolstering network security measures.

3) Convolutional Neural Network: Convolutional Neural Network (CNN) architecture was used to develop our third model. CNNs are widely used to perform image classifications and excel at computer vision problems. It is atypical to use them for tabular data despite the latter also being a classification problem. However, for the dataset used (5G-NIDD) network flow data which is sequential, and each pair of rows have some commonalities in their attributes and values. This relation between the neighboring values of the same and adjacent columns allows us to utilize the architecture of CNNs. The first set of results was obtained using without altering the existing relative order of the attributes. In an attempt to maximize spatial correlations, a careful reorganizing of attributes based on domain knowledge is proposed.

4) Recurrent Neural Network: The proposed RNN model is tailored to leverage the sequential nature of network flow data in 5G networks for intrusion detection. In contrast to Convolutional Neural Networks (CNNs), which excel at capturing spatial correlations in image data, RNNs are adept at capturing temporal dependencies in sequential data, making them suitable for analyzing network traffic patterns over time. The RNN model architecture consists of a single Long Short-Term Memory (LSTM) layer followed by a Dense layer with softmax activation for multiclass classification. The LSTM layer enables the model to retain and learn from historical information in the sequential data, allowing it to capture nuanced patterns and correlations.

V. Results

The results of the intrusion detection models evaluated on the 5G-NIDD dataset are summarized in Table I. Each model's performance is assessed in terms of accuracy, precision, and recall for binary classification to identify network traffic as either benign or malicious. Additionally, multi-class classification is performed to classify the attack type associated with malicious traffic. The attack types considered in the multi-class classification are HTTPFlood, ICMPFlood, SYNFlood, SYNScan, SlowrateDoS, TCPConnectScan, UDPFlood, and UDPScan.

Models	Accuracy	Precision	Recall	F1-score
Boosting Classifier (AdaBoost, CatBoost, LightGBM, XGBoost)	0.9996	0.9993	0.9998	0.9996
Hard voting Model (Decision Tree, Random Forest, Logistic Regression, KNN, MLP, GBM)	0.9992	0.9991	0.9990	0.9991
Soft voting Model (Logistic Regression, MLP, GBM)	0.9991	0.9993	0.9984	0.9992
CNN	0.9994	0.9996	0.9993	0.9995
RNN	0.9981	0.9981	0.9981	0.9981

TABLE I
METRICS FOR THE BINARY CLASSIFICATION MODELS.

Attack Types	Boosting LSTM	Hard voting	Soft voting	CNN	RNN
Benign	0.9996	0.9990	0.9992	0.9994	0.9992
HTTPFlood	0.9904	0.9989	0.9987	0.9964	0.9948
ICMPFlood	1.00	0.9916	0.9958	1.0000	0.9916
SYNFlood	0.9971	1.00	1.00	0.9958	0.9963
SYNScan	0.9969	0.9982	0.9992	0.9985	0.9928
SlowrateDoS	0.9972	0.9993	0.9988	0.9873	0.9875
TCPConnectScan	0.9982	0.9936	0.9931	0.9983	0.9987
UDPFlood	1.00	0.9999	0.9997	0.9999	0.9999
UDPScan	0.9987	0.9987	0.9984	0.9972	0.9993

TABLE II
PRECISION FOR THE MULTI-CLASS CLASSIFICATION MODELS.

Attack Types	Boosting LSTM	Hard voting	Soft voting	CNN	RNN
Benign	0.9998	0.9994	0.9989	0.9997	0.9999
HTTPFlood	0.9982	0.9998	0.9999	0.9938	0.9937
ICMPFlood	1.00	1.00	1.00	1.0000	1.00
SYNFlood	1.00	1.00	1.00	0.9984	0.9994
SYNScan	0.9967	0.9989	0.9992	0.9967	0.9976
SlowrateDoS	0.9801	0.9997	0.9997	0.9937	0.9908
TCPConnectScan	0.9984	0.9972	0.9977	0.9961	0.9940
UDPFlood	0.9992	0.9993	0.9994	0.9994	0.9993
UDPScan	0.9969	0.9874	0.9953	0.9984	0.9906

TABLE III
RECALL FOR THE MULTI-CLASS CLASSIFICATION MODELS.

The results of our research, as summarized in Table V, demonstrate the classification performance of various models in a multiclass scenario. Each model was evaluated based on metrics such as accuracy, precision, recall, and F1 score, providing insights into their effectiveness in distinguishing between different classes of attacks.

The Long Short-Term Memory (LSTM) model exhibited a commendable accuracy of 99.73%, with precision, recall, and F1 score values of 99.75%, 99.72%, and 99.73% respectively. This underscores the LSTM model's robustness and proficiency in accurately classifying network traffic instances across multiple attack categories.

In comparison, the Hard Voting classifier achieved a higher accuracy of 99.92%, surpassing the individual LSTM model. While maintaining high precision (99.77%) and recall (99.80%), the Voting classifier model achieved an impressive F1 score of 99.78%. This highlights the efficacy of combining multiple models to enhance overall classification performance.

Similarly, the Soft Voting classifier Model yielded a comparable accuracy of 99.92%, with precision and recall values of 99.81% and 99.89% respectively. Notably, the Voting classifier model achieved the highest F1 score among all models evaluated, reaching 99.85%. This further corroborates the effectiveness of voting classifier methods in improving classification accuracy and robustness.

Furthermore, the Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) models exhibited high accuracies of 99.84% and 99.81% respectively. Both models demonstrated consistent precision, recall, and F1 score values, underscoring their reliability in multiclass classification tasks.

Overall, our results underscore the efficacy of voting classifier methods, particularly soft voting classifier, in enhancing the classification performance in a multiclass scenario. These findings provide valuable insights for the development of robust and accurate classification models for network security applications.

The performance metrics of various machine learning models used for multi-class classification in intrusion detection systems for 5G mobile networks are summarized in Table II, Table III, and Table IV. Precision, recall, F1-score, and accuracy are evaluated across different attack types to assess the effectiveness of each model.

Precision, which measures the proportion of correctly predicted instances among all instances predicted as positive for a particular class, reflects the model's ability to avoid false positives. The Hard Voting Model demonstrates high precision scores across most attack types, indicating its effectiveness in minimizing false positives. Similarly, the Soft Voting Model exhibits high precision scores, suggesting robustness in classifying different attack types with minimal false positives. However, the CNN and RNN models show varying precision scores across different attack types, with generally high precision values for benign traffic but lower precision for certain attack types such as SlowrateDoS and UDPScan.

Recall, also known as sensitivity, measures the proportion of correctly predicted instances of a particular class among all instances of that class in the dataset. It reflects the model's ability to capture all instances of a class without missing any. The Hard Voting Model shows high recall scores across most attack types, indicating its capability to effectively detect instances of various attacks. Similarly, the Soft Voting Model demonstrates high recall scores, suggesting its ability to capture instances of different attack types effectively. However, the CNN and RNN models exhibit varying recall scores across different attack types, with generally high recall values for benign traffic but lower recall for certain attack types like ICMPFlood and SlowrateDoS.

The F1-score, the harmonic mean of precision and recall,

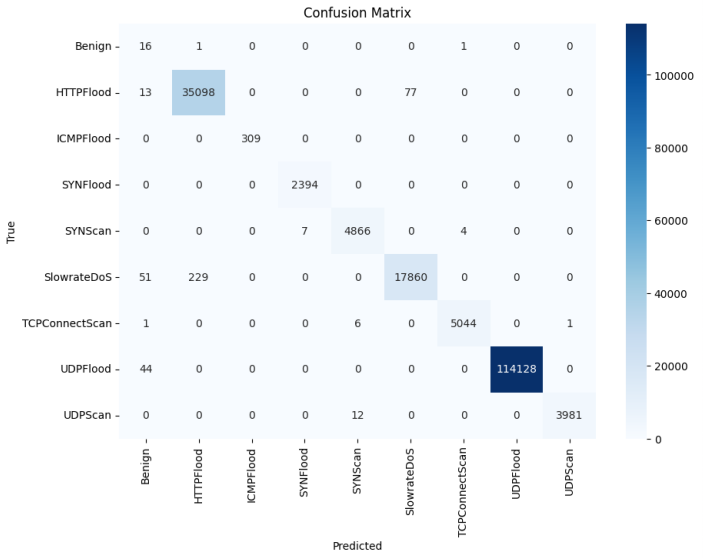

Fig. 10. Confusion Matrix for LSTM MultiClass Classification model

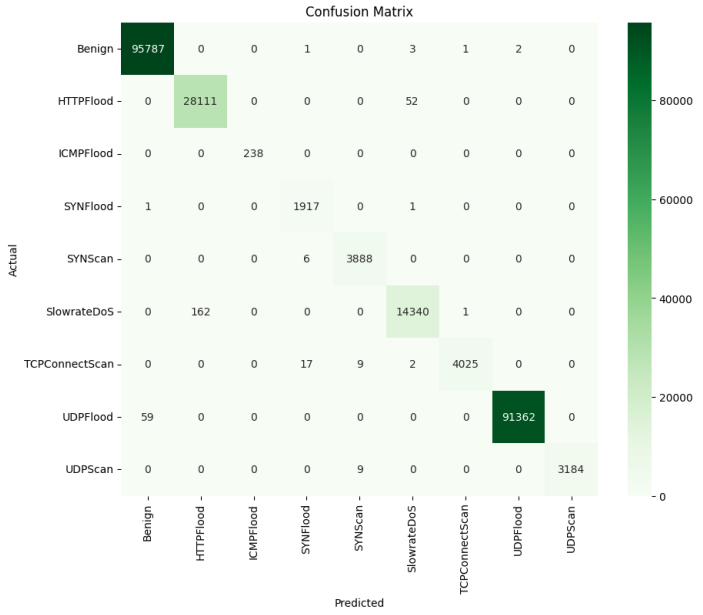

Fig. 11. Confusion Matrix for CNN MultiClass Classification model

Attack Types	Boosting LSTM	Hard voting	Soft voting	CNN	RNN
Benign	0.9996	0.9992	0.9991	0.9995	0.9995
HTTPFlood	0.9943	0.9994	0.9993	0.9951	0.9942
ICMPFlood	1.00	0.9958	0.9979	1.0000	0.9958
SYNFlood	0.9985	1.00	1.00	0.9971	0.9979
SYNScan	0.9968	0.9985	0.9992	0.9976	0.9952
SlowrateDoS	0.9886	0.9995	0.9992	0.9905	0.9892
TCPConnectScan	0.9983	0.9954	0.9954	0.9972	0.9964
UDPFlood	0.9996	0.9996	0.9995	0.9996	0.9996
UDPScan	0.9978	0.9930	0.9968	0.9978	0.9949

TABLE IV
F1-SCORE FOR THE MULTI-CLASS CLASSIFICATION MODELS.

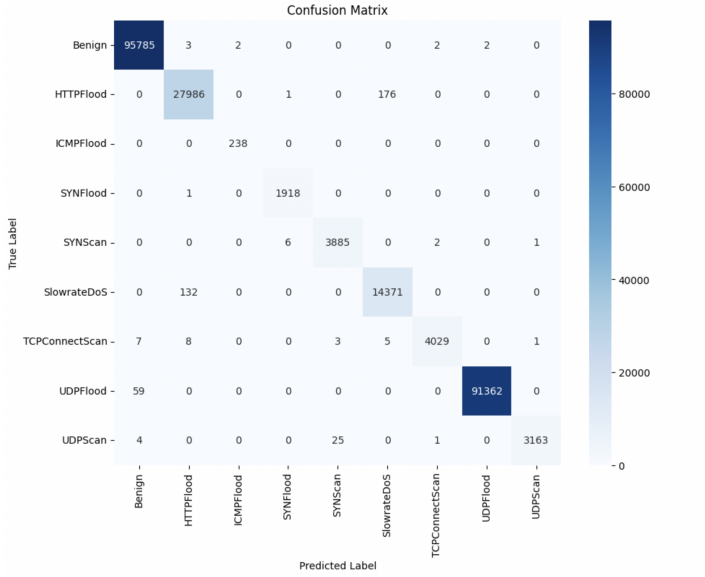

Fig. 12. Confusion Matrix for RNN MultiClass Classification model

Models	Accuracy	Precision	Recall	F1 score
Boosting LSTM	0.9973	0.9975	0.9972	0.9973
Hard voting classifier	0.9992	0.9977	0.998	0.9978
Soft voting classifier	0.9992	0.9981	0.9989	0.9985
CNN	0.9984	0.9984	0.9983	0.9984
RNN	0.9981	0.9981	0.9981	0.9981

TABLE V
ACCURACIES FOR THE MULTI CLASSIFICATION MODELS.

provides a balanced measure of a model's performance. Both the Hard Voting Classifier and the Soft Voting Voting classifier achieve high F1-scores across different attack types, indicating a good balance between precision and recall. The CNN model also achieves a high F1-score, suggesting a good balance between precision and recall for most attack types. Similarly, the RNN model demonstrates a high F1-score, indicating effective performance in capturing instances of various attack types. Accuracy is an important metric for evaluating the overall performance of intrusion detection systems. The CNN model achieves the highest accuracy among all models, indicating its effectiveness in correctly classifying instances across different attack types. The RNN model also demonstrates high accuracy, suggesting robust performance in multi-class classification tasks. Additionally, both the Hard Voting classifier and the Soft Voting classifier exhibit high accuracy, indicating their overall effectiveness in intrusion detection across different attack types.

In our research, we scrutinized the classification performance between slowrate DoS and HTTP flood attacks, both of which target the application layer and employ TCP as the

underlying protocol. Despite the similarities in their modus operandi, the dataset exhibited a notable bias towards HTTP flood attacks, with their sample size roughly doubling that of slowrate DoS instances. This disparity in dataset representation compounded the challenge of accurately distinguishing between the two attack types.

Our analysis encompassed both binary and multiclass classifications, shedding light on the intricacies of model performance. We meticulously examined the confusion matrices of Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN), scrutinizing training and testing times alongside the efficacy of feature selection techniques. Additionally, we compared a gamut of metrics including F1 score, accuracy, recall, and precision to gauge the overall performance of the models.

In both binary and multiclass scenarios, Voting classifier models emerged as frontrunners, demonstrating a remarkable reduction in bias and variance. By amalgamating predictions from diverse individual models, voting classifier methods exhibited superior generalization capabilities, thereby bolstering their performance in discerning between slowrate DoS and HTTP flood attacks. Moreover, the voting classifier approach provided an inherent robustness to overfitting and ensured stability, even in the face of fluctuations in the training data.

The comparative analysis of binary classifications unveiled the prowess of Boosting Long Short-Term Memory (LSTM) networks and other boosting classifiers such as AdaBoost, CatBoost, LightGBM, and XGBoost. Their adeptness with tabular data, reduced reliance on extensive feature engineering, and enhanced interpretability distinguished them from their counterparts, underscoring their suitability for classification tasks in network security.

Our research elucidated the reasons behind the improved results, attributing them to the utilization of more complex models capable of discerning intricate patterns effectively. Furthermore, the inherent ability of LSTM to model sequential data, such as network traffic, by prioritizing recent information played a pivotal role in enhancing classification accuracy. Diverging from the simplistic approach outlined in the original paper, we opted to integrate multiple sophisticated models in our voting classifier approach, thereby circumventing individual model limitations and achieving enhanced performance. .

FUTURE SCOPE

Looking ahead, our research paves the way for future endeavors in real-time data analysis and prediction. The insights gleaned from this study provide a solid foundation for further exploration in the realm of network security, with potential applications spanning various domains requiring robust classification algorithms for threat detection and mitigation

ACKNOWLEDGMENT

We extend our gratitude to Dr. Abdallah Moubayed, our mentor throughout this project conducted under the course SER 517: Software Capstone. His guidance was invaluable in navigating the complexities of our research and achieving our objectives effectively

REFERENCES

[1] Rezvy, S., Luo, Y., Petridis, M., Lasebae, A. and Zebin, T., 2019, March. An efficient deep learning model for intrusion classification and prediction in 5G and IoT networks. In 2019 53rd Annual Conference on information sciences and systems (CISS) (pp. 1-6). IEEE.

[2] Bayana alenazi, Dr. Hala Eldaw Idris, "Wireless Intrusion and Attack Detection for 5G Networks using Deep Learning Techniques

[3] Imanbayev, A.; Tynymbayev, S.; Odarchenko, R.; Gnatyuk, S.; Berdibayev, R.; Baikenov, A.; Kaniyeva, N. Research of Machine Learning Algorithms for the Development of Intrusion Detection Systems in 5G Mobile Networks and Beyond. Sensors 2022, 22, 9957. https://doi.org/10.3390/s22249957

[4] Bocu, R.; Iavich, M.Real-Time Intrusion Detection and Prevention System for 5G and beyond Software-Defined Networks.Symmetry 2023, 15, 110. https://doi.org/10.3390/sym15010110

[5] Wang, J.; Zhou, X.; Hou, Z.; Xu, X. (2022). Homogeneous ensemble models for predicting infection levels and mortality of COVID-19 patients: Evidence from China. Digital Health, 8, 20552076221133692.