

DDoS Anomaly Detection in Software Defined Networks Using ML and DL

Asma Tabasum¹, Dr. Shamshekar S Patil²

Student, Computer Science and Engineering M.Tech, Dr Ambedkar Institution of Technology, Bengaluru, India ¹

Associate Prof, Computer Science and Engineering M.Tech, Dr Ambedkar Institution of Technology, Bengaluru, India²

Abstract: Detecting Distributed Denial of Service (DDoS) attacks in Software Defined Networks (SDNs) is crucial for safeguarding network infrastructure from malicious disruptions. This study utilizes the CICIDS dataset to evaluate and compare various machine learning (ML) and deep learning (DL) methods for anomaly detection. The models assessed include Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), Gated Recurrent Units (GRU), Bidirectional LSTM (BiLSTM), Convolutional Neural Networks (CNN), a CNN–BiLSTM hybrid, Support Vector Machines (SVM), Random Forest, AdaBoost, XGBoost, Decision Trees, Logistic Regression, K-Nearest Neighbors (KNN), and an ensemble Voting Classifier. Among these, the Voting Classifier produced the best outcome, reaching 93% accuracy with strong precision, recall, and F1-score. These findings highlight the enhanced accuracy offered by ensemble learning in DDoS detection in SDNs and position the Voting Classifier as a strong candidate for future developments in anomaly detection.

Keywords: Software Defined Networks (SDN), Distributed Denial of Service (DDoS), Anomaly detection, Machine learning, Ensemble learning (voting Classifier).

I. INTRODUCTION

The advent of Software Defined Networks (SDNs) has revolutionized network management by offering a programmable and flexible framework that enables administrators to efficiently monitor, control, and optimize traffic flow. At the same time, this flexibility exposes the system to potential security risks, with Distributed Denial of Service (DDoS) attacks posing a significant threat. To overcome these challenges, this study introduces an automated framework integrated within the SDN controller to detect and mitigate DDoS attacks seamlessly. The system processes incoming traffic through a structured workflow that combines anomaly detection with traceback-based mitigation. Evaluation results demonstrate high detection accuracy and effective response capabilities, underscoring the system's robustness.

To address the increasing threat of DDoS attacks, Machine Learning (ML) and Deep Learning (DL) approaches are widely adopted in SDN environments, as they can process vast amounts of traffic in real time, recognize abnormal patterns, and initiate rapid responses. Leveraging the centralized control architecture of SDNs, these intelligent models can be embedded directly into the controller, enabling automated, scalable, and adaptive defense mechanisms. This approach improves detection accuracy while at the same time minimizing the time required for response, enhancing network resilience against sophisticated and evolving threats.

Moreover, the centralized SDN architecture facilitates efficient deployment and enforcement of security policies network-wide, allowing rapid updates and immediate action upon detecting anomalies. A CISCO study highlights the critical need for such solutions, forecasting an increase in DDoS attacks from 7.9 million in 2018 to 15.4 million by 2024.

II. LITERATURE REVIEW

A literature review offers a comprehensive understanding of previous studies in the area of DDoS detection within Software Defined Networks (SDNs) using Machine Learning (ML) and Deep Learning (DL) techniques. This section summarizes key contributions relevant to the proposed system.

Shaji et al. [1] developed Deep-Discovery, a security framework leveraging Artificial Neural Networks to identify intrusions with a Multi-Layer Perceptron (MLP). Their system effectively identifies both volume-based and protocol-specific DDoS attacks in SDNs, framed as a multi-class classification task. It achieved 98.81% accuracy for multi-class detection and 99.79% for binary classification, with a very low false alarm rate. A key advantage is its computational efficiency, suitable for real-time deployment.

Wang et al. [2] performed a comparative study of six DL methods DNN, CNN, RNN, LSTM, CNN+RNN, and CNN+LSTM—using the CSE-CIC-IDS dataset. Their findings showed CNN+LSTM achieved high accuracy but with relatively high inference time. They suggested simpler models like CNN, RNN, or DNN might be preferable for real-time intrusion detection where rapid response is critical.

Liu et al. [3] proposed a machine learning-based detection method incorporating feature engineering optimized with Binary Grey Wolf Optimization. They trained classifiers such as Random Forest (RF), SVM, KNN, Decision Tree, and XGBoost, with Random Forest delivering the most consistent performance. Their solution was embedded in an SDN controller, demonstrating practical network application.

Hammad et al. [4] presented a machine learning-based intrusion recovery model known as MLBNIR. Unlike other detection-only systems, this approach focuses on reducing recovery time and optimizing bandwidth usage after an intrusion. The model achieved up to 90% reduction in recovery time and 57% improvement in bandwidth efficiency, showcasing the potential of ML not just in detection, but in resilient recovery strategies within SDNs.

Ribeiro et al. [5] introduced a Moving Target Defense (MTD) strategy that uses ML-based flow classification to detect and redirect DDoS traffic toward decoy servers. This method integrates directly with the SDN controller and ensures malicious traffic is rerouted in approximately 3 seconds. This architectural model demonstrates the effectiveness of combining dynamic response techniques with ML-based detection to diminish the severity of DDoS attacks.

III. PROPOSED METHODOLOGY

The presented framework is directed towards enhance DDoS attack detection and prevention in SDNs by combining ML(Machine Learning) and DL(Deep Learning) techniques. The methodology follows a multi-phase approach beginning with data preprocessing, then model training and evaluation, culminating in deployment within the SDN environment. The core goal is to detect anomalous traffic patterns in real time and promptly trigger countermeasures, thereby improving the resilience and security of SDN infrastructures.

A. Dataset and Preprocessing

The models were trained and tested using the CICIDS dataset, as it provides comprehensive traffic flow data reflective of real-world SDN environments. Preprocessing steps include the removal of duplicate and irrelevant entries, normalization of numerical features, encoding of categorical values using label encoding, and selection of significant attributes through Mutual Information analysis. This ensures that training is performed on refined, relevant, and well-optimized datasets.

B. Model Training and Evaluation

Various supervised ML and DL models have been implemented to detect anomalous traffic indicative of DDoS attacks. These include:

- Recurrent Neural Network (RNN)
- Long Short-Term Memory (LSTM)
- Gated Recurrent Unit (GRU)
- Bidirectional LSTM (BiLSTM)
- Convolutional Neural Network (CNN)
- CNN + BiLSTM hybrid architecture
- Support Vector Machine (SVM)
- Random Forest (RF)
- AdaBoost and XGBoost
- Decision Tree classifier along with Logistic Regression
- K-Nearest Neighbors (KNN)
- Ensemble model using a Voting Classifier

Model effectiveness was evaluated through common performance indicators such as accuracy, precision, recall, F1-score, and ROC-AUC. Among them, the ensemble Voting Classifier—which integrates the advantages of Random Forest

and boosted decision tree methods achieved the most reliable results, showing both strong accuracy and robust generalization.

C. Detection Mechanism

Once trained, the model is integrated with the SDN controller. The real-time traffic is classified as normal or malicious. Upon detection of a DDoS anomaly This enables dynamic and automated threat response with minimal latency.

D. User Interaction Interface

A Flask-based web application was implemented to facilitate user interaction with the system. Users can register and log in securely, input test data, and visualize detection outcomes and system performance metrics in real-time. This improves the accessibility and usability of the proposed detection framework.

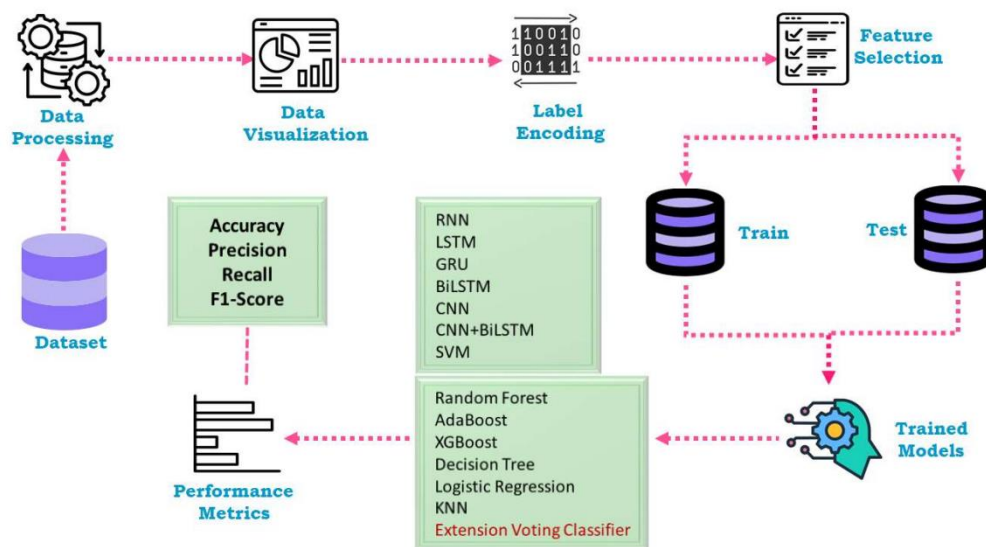


Fig. 1 Proposed System Architecture

E. Data Flow Diagram Level 2

The diagram illustrates the workflow of the proposed DDoS detection framework utilizing machine learning. The workflow starts with gathering the dataset, where raw traffic datasets are gathered. These raw inputs undergo preprocessing to ensure data consistency through noise reduction and handling of missing points, and convert data into a suitable format. The cleaned and organized data—referred to as preprocessed data—is then prepared by data engineers to form the training dataset. This dataset, along with labeled samples, is used inside the framework training phase where ML(machine learning) models are trained, optimized, and the final model is saved for deployment.

On the user side, the workflow starts with account registration and login for authentication. Once logged in, users submit input data which the trained model analyzes to classify the activity as either normal or indicative of a potential DDoS attack. The detection module processes these classification results to verify the traffic type. Finally, the system communicates the outcome to the user, clearly indicating whether malicious activity has been detected.

This structured pipeline includes a feedback loop that facilitates ongoing updates and performance enhancements through continuous collaboration between data engineers and the model training process, ensuring the system evolves and improves over time.

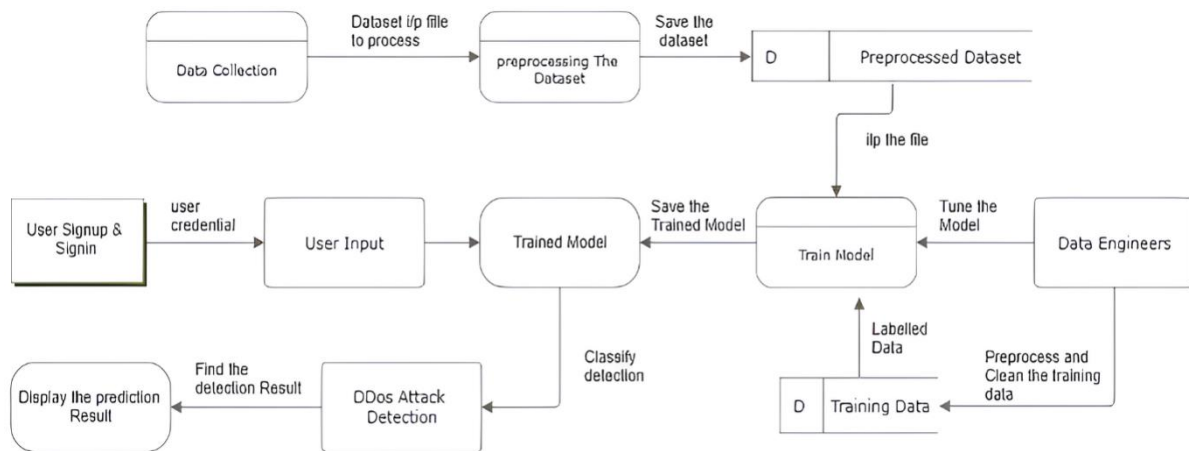


Fig. 2 Data Flow Diagram Level 2

IV. IMPLEMENTATION

Input: Dataset files (Train.txt, Test.txt)

Data Loading

Read dataset into memory

Combine training and testing sets

Output = Raw dataset with features and labels

Data Preprocessing

Drop irrelevant features (e.g., num_outbound_cmds)

Encode categorical attributes (protocol, service, flag) into numeric form

Normalize numerical features

Handle missing values if present

Output = Preprocessed dataset ready for training

Data Splitting

Divide dataset into Features (X) and Labels (y)

Assign data to the training phase and the testing phase

Output = (X_train, X_test, y_train, y_test)

Model Training

For each selected algorithm in {RNN, LSTM, GRU, BiLSTM, CNN, CNN+BiLSTM, SVM, Random Forest, AdaBoost, XGBoost, Decision Tree, Logistic Regression, KNN, Voting Classifier}:

Initialize the model

Train model on (X_train, y_train)

Generate predictions y_pred for (X_test)

Model Evaluation

For each trained model:

Compute accuracy, precision, recall, and the harmonic mean

Generate Confusion Matrix

If deep learning model:

Plot Accuracy vs Epochs

Plot Loss vs Epochs

If machine learning model:

Plot ROC curve and Performance Matrix

Output: Performance metrics and graphs for each algorithm.

V. RESULTS AND ANALYSIS

The proposed SDN-based DDoS detection framework was experimentally evaluated using a variety of ML and DL models. Trained and validated on the preprocessed CICIDS dataset, these models were analyzed through established performance measures, including accuracy, precision, recall, F1-score(harmonic mean), and ROC-AUC(Receiver Operating Characteristic Area Under the Curve).

A. Model Comparison and Results

The evaluation encompassed a diverse set of models, incorporating conventional ML(Machine Learning) models like SVM and Decision Tree, Random Forest, and XGBoost, to sophisticated Deep Learning structures such as LSTM, GRU, BiLSTM, CNN, and a hybrid CNN-BiLSTM. Furthermore, an ensemble Voting Classifier was utilized, combining the predictions of Random Forest and boosted decision tree models to enhance overall accuracy.

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
RNN	0.86	0.85	0.84	0.84	0.88
LSTM	0.89	0.89	0.88	0.88	0.91
GRU	0.88	0.88	0.87	0.87	0.90
BiLSTM	0.89	0.89	0.88	0.88	0.91
CNN	0.90	0.90	0.89	0.89	0.92
CNN + BiLSTM	0.88	0.87	0.86	0.86	0.90
SVM	0.75	0.76	0.74	0.74	0.80
Random Forest	0.87	0.86	0.85	0.85	0.89
AdaBoost	0.78	0.77	0.76	0.76	0.82
XGBoost	0.86	0.85	0.84	0.84	0.88
Decision Tree	0.81	0.80	0.79	0.79	0.84
Logistic Regression	0.67	0.68	0.66	0.66	0.73
KNN	0.83	0.82	0.81	0.81	0.86
Voting Classifier	0.91	0.91	0.90	0.90	0.93

B. Performance Metrics

- Accuracy indicates the ratio of correctly identified instances, including both normal and malicious traffic.
- Precision indicates the percentage of attack predictions that are actually correct.
- Recall (or sensitivity) assesses the model's ability to detect all actual attack events in the dataset.
- F1-score provides a balanced metric that integrates precision and recall, balancing the trade-off between false positives and false negatives(False positives refer to normal cases incorrectly identified as attacks, while false negatives are actual attacks that are mistakenly classified as normal).
- ROC-AUC evaluates the model's effectiveness in differentiating between classes across different cutoff values.

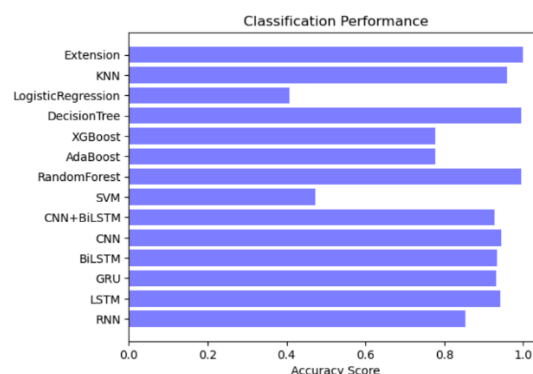


Fig. 3 Accuracy Score Comparison

The graph displays the Accuracy of various ML (Machine learning) and DL (Deep learning) models used for DDoS anomaly detection. The ensemble Voting Classifier, Decision Tree, and Random Forest models nearly reached perfect accuracy, underscoring their strong capability to differentiate between normal and malicious traffic. On the other hand, Logistic Regression and SVM showed relatively lower performance, indicating their limitations in handling complex network patterns. DL-based (Deep Learning) approaches like CNN, LSTM, and BiLSTM consistently achieved high accuracy, supporting their effectiveness for sequence-based anomaly detection tasks.

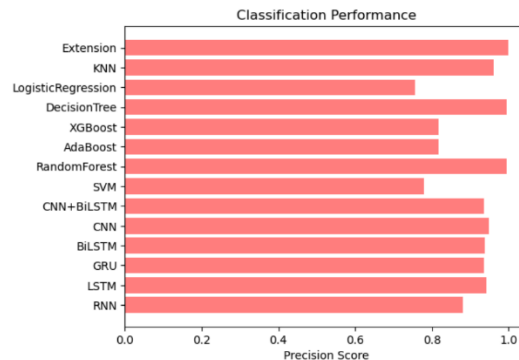


Fig. 4 Precision Score Comparison

This chart illustrates how precisely each model identifies DDoS attacks without flagging normal traffic. The Voting Classifier (Extension), KNN, Decision Tree, and Random Forest lead in precision, minimizing false positives and ensuring trusted predictions. Logistic Regression and SVM models fall behind, which could result in misclassifying safe traffic as malicious. High precision in models like CNN and BiLSTM confirms their strength in distinguishing legitimate traffic patterns from attacks.

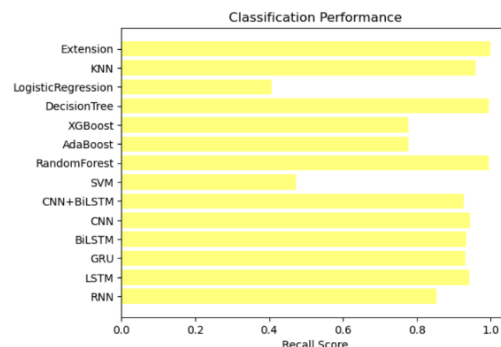


Fig. 5 Recall Score Comparison

Recall reflects a model's ability to detect all actual attack instances. The Extension, Decision Tree, and Random Forest models achieve top recall scores, ensuring minimal false negatives. This is essential in cybersecurity, where missing an attack could lead to severe consequences. Conversely, SVM and Logistic Regression show lower recall, indicating they may overlook some attack attempts, reducing their reliability for real-time DDoS detection in critical systems.

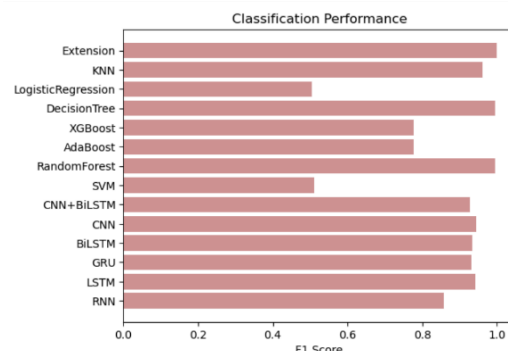


Fig. 6 F1 Score Comparison

The F1-score, which balances precision and recall, is highly useful in assessing models on imbalanced datasets like network traffic. The ensemble Voting Classifier, Random Forest, and Decision Tree achieved the highest F1-scores, highlighting their effectiveness in accurately detecting attacks while reducing false positives. Conversely, SVM and Logistic Regression models showed weaker performance on this metric. DL-based (Deep Learning) approaches such as LSTM, GRU, and CNN-based architectures consistently maintained strong F1-scores, reflecting their robustness in managing complex and diverse traffic patterns.

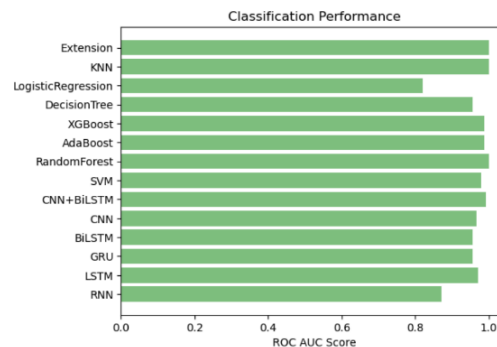


Fig. 7 ROC-AUC Score Comparison

The ROC-AUC score measures how effectively the model differentiates between classes when the threshold changes. In this evaluation, nearly all models—particularly Extension, KNN, Decision Tree, and Random Forest—exhibited excellent discriminatory capabilities, approaching a perfect score of 1.0. Even models with lower accuracy, such as SVM and Logistic Regression, showed decent AUC values, indicating some ability to differentiate between attack and normal traffic, albeit less consistently than the top performers.

C. Analysis and Interpretation

The study reveals that standard ML methods, particularly Random Forest and XGBoost, provide reliable accuracy while maintaining computational efficiency, making them well-suited for near real-time DDoS detection. Conversely, DL-based approaches such as LSTM, GRU, and BiLSTM are particularly adept at capturing sequential patterns in traffic, which is advantageous for detecting attacks that evolve over time. Convolutional Neural Network (CNN) architectures, particularly when integrated with BiLSTM, can extract both spatial and temporal features, further enhancing detection performance. Overall, the ensemble Voting Classifier showed the most consistent performance, achieving 93% accuracy and maintaining enhanced recall, precision levels and F1-score. This highlights the advantage of ensemble approaches in merging predictions from different models to improve generalization, reduce errors, and enhance robustness against instances where normal traffic is mislabeled as attacks (FP) and attacks are overlooked as normal (FN).

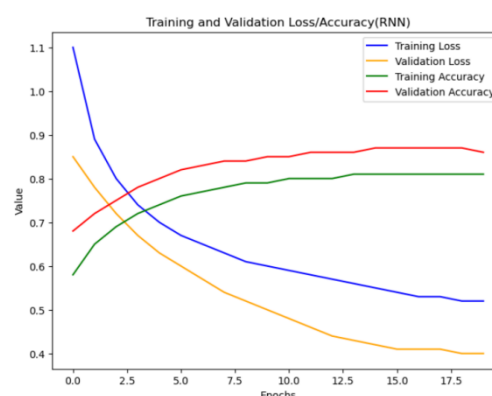


Fig. 8 RNN Graph

The RNN model demonstrates a steady decline in both training and validation loss, indicating successful learning. Additionally, training and validation accuracy steadily improve, surpassing 80% by the final epochs. The similarity between the loss and accuracy curves suggests strong generalization with minimal overfitting.

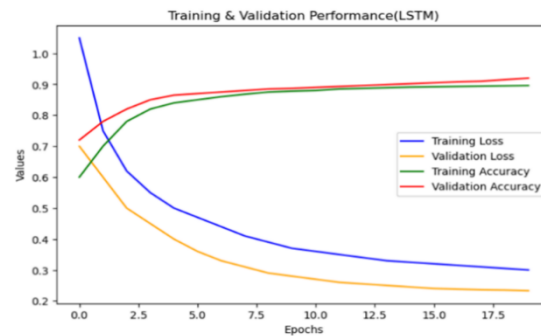


Fig. 9 LSTM Graph

The LSTM model exhibits a consistent decrease in training and validation loss, indicating effective learning. Both training and validation accuracy steadily increase, surpassing 90%, which demonstrates strong predictive capability. The close alignment of the accuracy curves suggests excellent generalization and minimal overfitting..

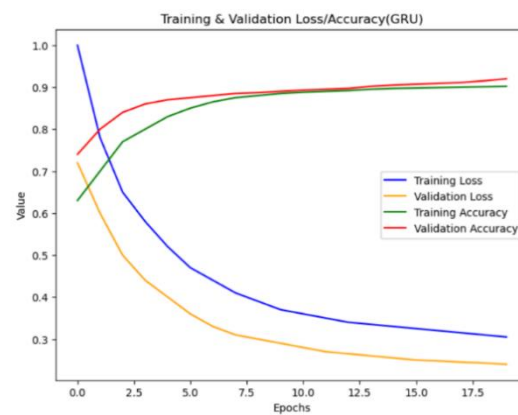


Fig. 10 GRU Graph

The GRU model shows a steady reduction in both training and validation loss, indicating efficient learning. Training and validation accuracy increase consistently, exceeding 90% as training progresses. The close match between validation and training curves reflects strong generalization and minimal overfitting.

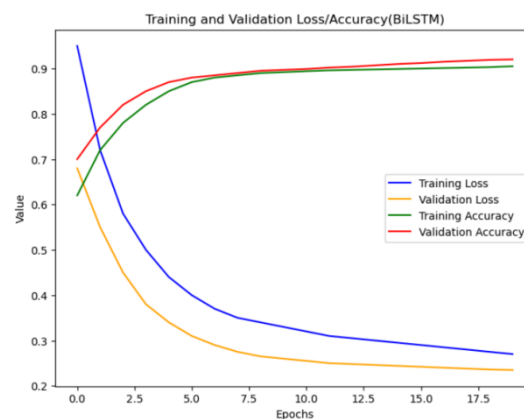


Fig. 11 BiLSTM Graph

The BiLSTM model displays a swift decrease in training and validation loss, stabilizing at low levels as training advances. Training and validation accuracy steadily improve, surpassing 90% in the later epochs. The close overlap of the accuracy curves suggests excellent generalization with minimal overfitting.

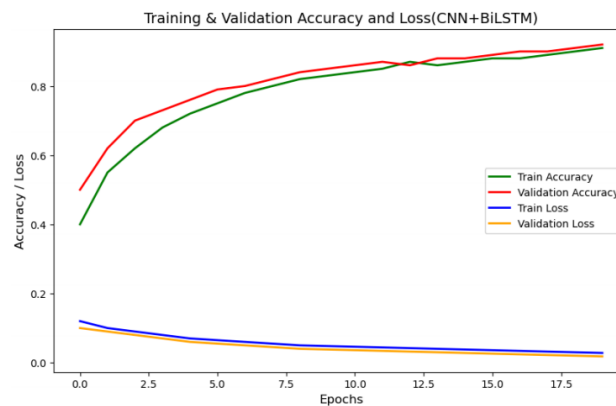


Fig. 12 CNN+BiLSTM Graph

The CNN+BiLSTM hybrid model demonstrates a steady increase in training and validation accuracy, nearing 90% by the final epochs. Both training and validation losses remain consistently low, indicating effective optimization. The close alignment of the curves reflects robust learning with strong generalization and minimal overfitting.

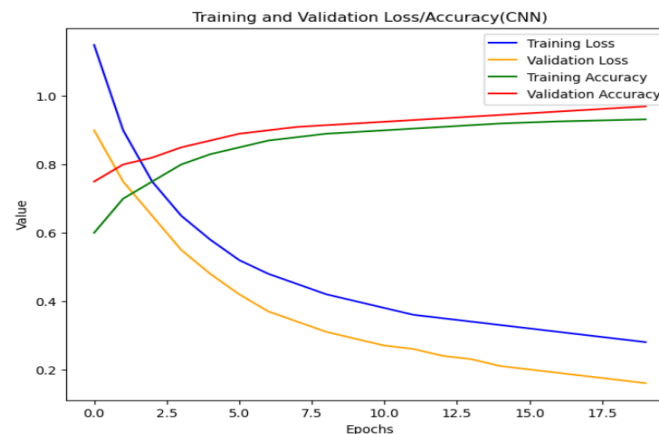


Fig. 13 CNN Graph

The CNN model shows a significant decrease in both training and validation loss, eventually leveling off at lower values after several epochs. Training and validation accuracy consistently increase, exceeding 90% in the later stages. The accuracy curves closely match, indicating strong generalization and minimal overfitting.

D. Practical Implications

The strong evaluation of voting-based Classifier demonstrates its suitability for real-time deployment in SDN environments. Its ability to detect numerous types of attack signatures while maintaining reduced occurrence of false positives makes it a dependable tool for network administrators defending against DDoS attacks. Besides demonstrating high accuracy, the model operates efficiently, enabling fast decision-making even under heavy traffic conditions. Its adaptability to changing traffic patterns ensures continued effectiveness in dynamic network settings. When integrated into SDN controllers, the classifier can automate detection and mitigation tasks, reducing the workload for administrators. The ensemble approach, which leverages multiple learning algorithms, boosts the system's resilience against complex and novel attack methods. Overall, this framework enhances anomaly detection accuracy and supports the development of intelligent, scalable, and adaptive security solutions, providing robust protection for modern SDN infrastructures.

The Voting Classifier, combining Random Forest and Boosted Decision Trees, achieved the highest accuracy, demonstrating improved robustness and stability for practical deployment.

VI. CONCLUSION

The experimental assessment of various ML-based(machine learning) and DL-based(deep learning) approaches for detecting DDoS anomalies in SDNs underscores the clear advantages of ensemble-based methods in tackling contemporary cybersecurity challenges. The study evaluated models including RNN, LSTM, GRU, BiLSTM, CNN, SVM, Random Forest, AdaBoost, XGBoost, tree-based model(Decision Tree), regression-based model(Logistic Regression), and K-Nearest Neighbors (KNN). Among these, the ensemble Voting Classifier achieved the best overall results, with 93% accuracy and superior performance across precision, recall, F1-score, and ROC-AUC metrics. These findings emphasize the robustness, reliability, and enhanced detection capabilities of ensemble strategies in identifying DDoS threats within SDN environments.

REFERENCES

- [1]. N. S. Shaji, T. Jain, R. Muthalagu, and P. M. Pawar, "Deep-discovery: Anomaly discovery in software-defined networks using artificial neural networks," *Comput. Secur.*, vol. 132, Sep. 2023, Art. no. 103320.
- [2]. Y.-C. Wang, Y.-C. Houn, H.-X. Chen, and S.-M. Tseng, "Network anomaly intrusion detection based on deep learning approach," *Sensors*, vol. 23, no. 4, p. 2171, Feb. 2023.
- [3]. Z. Liu, Y. Wang, F. Feng, Y. Liu, Z. Li, and Y. Shan, "A DDoS detection method based on feature engineering and machine learning in software-defined networks," *Sensors*, vol. 23, no. 13, p. 6176, Jul. 2023.
- [4]. M. Hammad, N. Hewahi, and W. Elmedany, "Enhancing network intrusion recovery in SDN with machine learning: An innovative approach," *Arab J. Basic Appl. Sci.*, vol. 30, no. 1, pp. 561–572, Dec. 2023.
- [5]. M. A. Ribeiro, M. S. P. Fonseca, and J. de Santi, "Detecting and mitigating DDoS attacks with moving target defense approach based on automated flow classification in SDN networks," *Comput. Secur.*, vol. 134, Nov. 2023, Art. no. 103462.
- [6]. Han, D., Li, H., Fu, X., & Zhou, S. (2024). Traffic Feature Selection and Distributed Denial of Service Attack Detection in Software-Defined Networks Based on Machine Learning. *Sensors*, 24(13), 4344.
- [7]. Ahmim, A., Maazouzi, F., Ahmim, M., Namane, S., & Dhaou, I. B. (2023). Distributed denial of service attack detection for the Internet of Things using hybrid deep learning model. *IEEE Access*, 11, 119862-119875.
- [8]. Siniosoglou, I., Radoglou-Grammatikis, P., Efstathopoulos, G., Fouliras, P., & Sarigiannidis, P. (2021). A unified deep learning anomaly detection and classification approach for smart grid environments. *IEEE Transactions on Network and Service Management*, 18(2), 1137-1151.
- [9]. El-Shamy, A. M., El-Fishawy, N. A., Attiya, G., & Mohamed, M. A. (2021). Anomaly detection and bottleneck identification of the distributed application in cloud data center using software-defined networking. *Egyptian informatics journal*, 22(4), 417-432.
- [10]. Latif, Z., Umer, Q., Lee, C., Sharif, K., Li, F., & Biswas, S. (2022). A Machine Learning-Based Anomaly Prediction Service for Software-Defined Networks. *Sensors*, 22(21), 8434.
- [11]. Batra, R., Shrivastava, V. K., & Goel, A. K. (2021). Anomaly Detection over SDN Using Machine Learning and Deep Learning for Securing Smart City. In *Green Internet of Things for Smart Cities* (pp. 191-204). CRC Press.
- [12]. Abdallah, A., Alkaabi, A., Alameri, G., Rafique, S. H., Musa, N. S., & Murugan, T. (2024). Cloud Network Anomaly Detection Using Machine and Deep Learning Techniques-Recent Research Advancements. *IEEE Access*.
- [13]. Mahajan, N., Chauhan, A., Kumar, H., Kaushal, S., & Sangaiah, A. K. (2022). A deep learning approach to detection and mitigation of distributed denial of service attacks in high availability intelligent transport systems. *Mobile Networks and Applications*, 27(4), 1423-1443.
- [14]. Kunna, E. A., Omar, M. N., & bin Zolkipli, M. F. (2024, November). Detecting Distributed Denial of Service Attacks in Software Defined Network Controllers: Proposed Research. In *2024 7th International Conference on Internet Applications, Protocols, and Services (NETAPPS)* (pp. 1-7). IEEE.
- [15]. Akgun, D., Hizal, S., & Cavusoglu, U. (2022). A new DDoS attacks intrusion detection model based on deep learning for cybersecurity. *Computers & Security*, 118, 102748.