IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 5, May 2023

An Enhanced Algorithm for Detection of Intruders using Deep Learning

Aarthy R[1], Bhavanisha[2], Birundhavathi[3], Gowthami[4], Vishali[5]
Assistant Professor, CSE, Dhanalakshmi Srinivasan Engineering College, Perambalur[1]
Student, CSE, Dhanalakshmi Srinivasan Engineering College, Perambalur[2,3,4]

Abstract: *The increasing sophistication and frequency of cyber-attacks have made it necessary to develop more advanced network intrusion detection systems (NIDS). Traditional rule-based NIDS may not be effective against new and evolving attack strategies, hence there is a need for more intelligent and dynamic detection techniques. This paper proposes an enhanced AI-based NIDS using XGBoost, a popular machine learning algorithm for classification and regression problems. The proposed system is designed to detect network attacks in real-time by analyzing network traffic data and identifying patterns that indicate suspicious behavior. The system has data preprocessing, feature extraction, and XGBoost classification. The data pre-processing stage involves collecting and cleaning network traffic data to remove noise and irrelevant information. The feature extraction stage selects the most relevant features from the preprocessed data using correlation analysis. The XGBoost algorithm is used in the classification stage to train and test the system's predictive model. The model is trained on a labeled dataset of normal and attack traffic, and then used to classify new network traffic as normal or malicious. To evaluate the performance of the proposed system, experiments were conducted on a publicly available dataset. The results show that the XGBoost-based NIDS outperforms traditional rule-based NIDS in terms of detection accuracy, false positive rate, and execution time. This study demonstrates the effectiveness of using XGBoost in developing an AI-based NIDS for real-time network attack detection. The proposed system has the potential to enhance the security of computer networks by detecting and preventing cyber-attacks in real-time.*

Keywords: Fuel Delivery, Android Application, Mobile Application, Tracking System, Real Time Application.

I. INTRODUCTION

Network intrusion detection systems (NIDS) are essential tools for protecting computer networks from cyber-attacks. NIDS work by analyzing network traffic and looking for patterns that indicate a potential intrusion or attack. Traditional NIDS solutions typically use rule-based or signature-based approaches, but these methods have limitations and may not be effective against modern, sophisticated attacks. To overcome these limitations, machine learning techniques such as XGBoost can be applied to develop an enhanced AI-based NIDS. XGBoost is a powerful machine learning algorithm that uses a gradient boosting framework to improve predictive accuracy and reduce overfitting.The using XGBoost, an AI-based NIDS can learn from historical network data to identify patterns and anomalies that indicate potential attacks.

This can lead to faster and more accurate detection of network intrusions, allowing security teams to respond quickly and mitigate potential damage.An enhanced AI-based NIDS can also improve the efficiency of network security operations. By automating the detection and response process, security teams can focus their attention on investigating and mitigating threats rather than manually sifting through large volumes of network data. An AI-based NIDS using XGBoost can provide a more robust and effective solution for protecting computer networks from cyber threats.

II. RELATED WORK

[1] In this paper, we propose a generative adversarial network (GAN) based intrusion detection system (G-IDS), where GAN generates synthetic samples, and IDS gets trained on them along with the original ones. G-IDS also fixes the difficulties of imbalanced or missing data problems. We model a network security dataset for an emerging CPS using

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 5, May 2023

NSL KDD-99 dataset and evaluate our proposed model's performance using different metrics. We find that our proposed G-IDS model performs much better in attack detection and model stabilization during the training process than a standalone IDS.

[2] In this paper, a framework of the generative adversarial networks, called IDSGAN, is proposed to generate the adversarial malicious traffic records aiming to attack intrusion detection systems by deceiving and evading the detection. Given that the internal structure and parameters of the detection system are unknown to attackers, the adversarial attack examples perform the black-box attacks against the detection system. IDSGAN leverages a generator to transform original malicious traffic records into adversarial malicious ones.

[3] Intrusion detection and prevention are two of the most important issues to solve in network security infrastructure. Intrusion detection systems (IDSs) protect networks by using patterns to detect malicious traffic. As attackers have tried to dissimulate traffic in order to evade the rules applied, several machine learning-based IDSs have been developed. In this study, we focused on one such model involving several algorithms and used the NSL-KDD dataset as a benchmark to train and evaluate its performance. We demonstrate a way to create adversarialinstances of network traffic that can be used to evade detection by a machine learning-based IDS.

[4] In the field of intrusion detection, there is often a problem of data imbalance, and more and more unknown types of attacks make detection difficult. To resolve above issues, this article proposes a network intrusion detection model called CWGAN-CSSAE, which combines improved conditional Wasserstein Generative Adversarial Network (CWGAN) and cost-sensitive stacked autoencoders (CSSAE). First of all, the CWGAN network that introduces gradient penalty and L2 regularization is used to generate specified minority attack samples to reduce the class imbalance of the training dataset. Secondly, the stacked autoencoder is used to intelligently extract the deep abstract features of the network data.

[5] Generative Adversarial Networks (GANs) have been widely studied and applied in anomaly detection in recent years thanks to their high potential in learning complex high-dimensional real data distribution. Deep learning techniques can greatly overcome the disadvantages of using traditional machine learning algorithms for intrusion detection. This work proposes to use current existing GANs and their variants for network intrusion detection using real dataset and show the feasibility and comparison.

[6] Intrusion detection systems (IDS), as one of important security solutions, are used to detect network attacks. With the extensive applications of traditional machine learning algorithms in the security field, intrusion detection methods based on the machine learning techniques have been developed rapidly. However, since the progress of technology and the defects of the intrusion detection system based on machine learning algorithms, the system has gradually failed to meet the requirement for cyber security.

[7] Virus attacks, unauthorized access, theft of information and denial-of-service attacks were the greatest contributors to computer crime, a number of techniques have been developed in the past few years to help cyber security experts in strengthening the security of a single host or the whole computer network. Intrusion Detection is important for both Military as well as commercial sectors for the sack of their Information Security, which is the most important topic of research for the future networks.

[8] The main function of Intrusion Detection System is to protect the resources from threats. It analyzes and predicts the behaviours of users, and then these behaviours will be considered an attack or a normal behaviour. We use Rough Set Theory (RST) and Support Vector Machine (SVM) to detect network intrusions. First, packets are captured from the network, RST is used to pre-process the data and reduce the dimensions.

[9] The ML method obtained over 90% accuracy in the easier classification task of differentiating malignant melanoma from benign melanoma. Nevertheless, in the overall classification test of malignant and benign tumours, ResNet, their chosen DL architecture, achieved the highest diagnostic accuracy of 82%.

[10] In this article, we propose a novel temporary warning network (TWN) for safety message dissemination in the urban traffic environment, in which both the spatial distribution and temporal duration of the networking scheme are taken into account. Specifically, TWN is constructed by the selection of relay vehicles based on the spatiotemporal correlation of vehicle trajectory so that the safety message can be quickly disseminated within the Regions of Interest (RoIs).

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 5, May 2023

IV. EXISTING SYSTEM

Traditional network intrusion detection systems (NIDS) use signature-based detection, anomaly-based detection, or a combination of both to detect network intrusions. Signature-based detection involves the use of pre-defined rules and signatures to detect known attacks. Anomaly-based detection involves the use of machine learning algorithms to detect deviations from normal behavior. However, these approaches face several challenges, such as low accuracy, high false-positive rates, and difficulty in detecting novel attacks.

In this article, we propose a novel temporary warning network (TWN) for safety message dissemination in the urban traffic environment, in which both the spatial distribution and temporal duration of the networking scheme are taken into account. Specifically, TWN is constructed by the selection of relay vehicles based on the spatiotemporal correlation of vehicle trajectory so that the safety message can be quickly disseminated within the Regions of Interest (RoIs).

V. PROPOSED SYSTEM

An Enhanced AI-Based Network Intrusion Detection System using XGBoost is a proposed system for detecting network intrusions and anomalies using advanced machine learning algorithms. This system is designed to enhance the security of computer networks by identifying and preventing attacks before they can cause damage. The intrusion detection systems (IDS) are often rule-based and rely on a fixed set of signatures to detect known attacks. However, with the increasing complexity and sophistication of cyber-attacks, these systems are no longer sufficient in detecting the emerging threats. AI-based IDSs are designed to detectboth known and unknown attacks by learning from the patterns and anomalies present in the network traffic.

XGBoost, short for eXtreme Gradient Boosting, is a powerful and scalable machine learning algorithm that has shown great success in various domains such as image classification, natural language processing, and anomaly detection. XGBoost uses decision trees as its base learners and employs gradient boosting to optimize the model's performance. The proposed Enhanced AI-Based Network Intrusion Detection System using XGBoost combines XGBoost with other machine learning techniques such as feature engineering, feature selection, and hyperparameter tuning to create a highly accurate and efficient IDS. The system will use network traffic data to train the model and identify anomalies in real-time, enabling quick response to potential threats. The benefits of this proposed system include improved accuracy, scalability, and adaptability, which are critical in today's complex network environments. With this system, organizations can enhance their network security and protect their critical data from potential attacks.

The proposed system aims to develop an enhanced AI-based network intrusion detection system using the XGBoost algorithm. XGBoost is an optimized gradient boosting algorithm that is widely used in machine learning tasks such as classification, regression, and ranking. It has been proven to be effective in various applications, including network intrusion detection. The system using XGBoost will involve the following steps: Data collection and preprocessing: The system will collect network traffic data from various sources and preprocess it to remove noise and anomalies.

Feature extraction: Relevant features will be extracted from the preprocessed data to feed into the XGBoost algorithm. XGBoost model training: The XGBoost algorithm will be trained on the extracted features to learn the patterns of normal and abnormal network behavior. Anomaly detection: The trained model will be used to detect anomalies in the network traffic data and raise alerts when an intrusion is detected. Performance evaluation: The system's performance will be evaluated using standard metrics such as accuracy, precision, recall, and F1-score to measure its effectiveness in detecting network intrusions. The proposed system will offer several advantages over traditional intrusion detection systems, such as better accuracy, faster response time, and reduced false positives. It will also be more resilient to evolving threats and able to adapt to changing network conditions. The enhanced AI-based network intrusion detection system using XGBoost is a promising solution for detecting andpreventing network intrusions in real-time, and it has the potential to enhance the security of modern computer networks.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 5, May 2023

VI. SYSTEM ARCHITECTURE

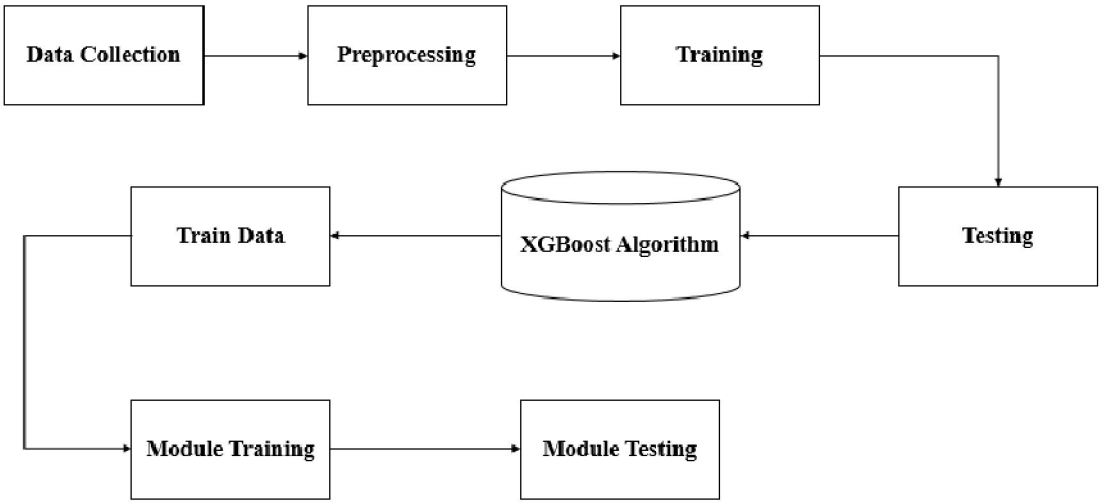

Figure 1

VII.MODULE DESCRIPTION

Data Preprocessing

- In this module, the raw network traffic data is preprocessed to extract relevant features.
- The extracted features are then normalized and fed into the XGboost.

Data Visualization

- The data visualization module allows users to visualize the network traffic data in various formats, such as tables, graphs, and charts. This makes it easier for users to understand the data and identify any patterns or anomalies that may indicate network intrusions.
- To provide real-time monitoring of the network traffic data. This allows users to quickly detect and respond to any network intrusions before they can cause significant damage.
- The data visualization module can be customized to meet the specific needs of different users. For example, users can choose which data to display, how it is displayed, and how often it is updated.
- It used to generate reports that provide insights into the network traffic data. These reports can be used to identify trends, evaluate the effectiveness of the NIDS, and make informed decisions about network security.

Anomaly Detection

This module consists of an XGboost that generates normal network traffic data and compares it with the real network traffic data to identify anomalies. The generator of the XGboost learns the distribution of normal network traffic data, and the discriminator learns to differentiate between real and fake network traffic data.

Alert Generation and Evaluation

This module generates alerts when an anomaly is detected by the XGboost. The alerts are then evaluated by a decision-making system that uses a threshold-based approach to classify the alerts as true positives or false positives.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 5, May 2023

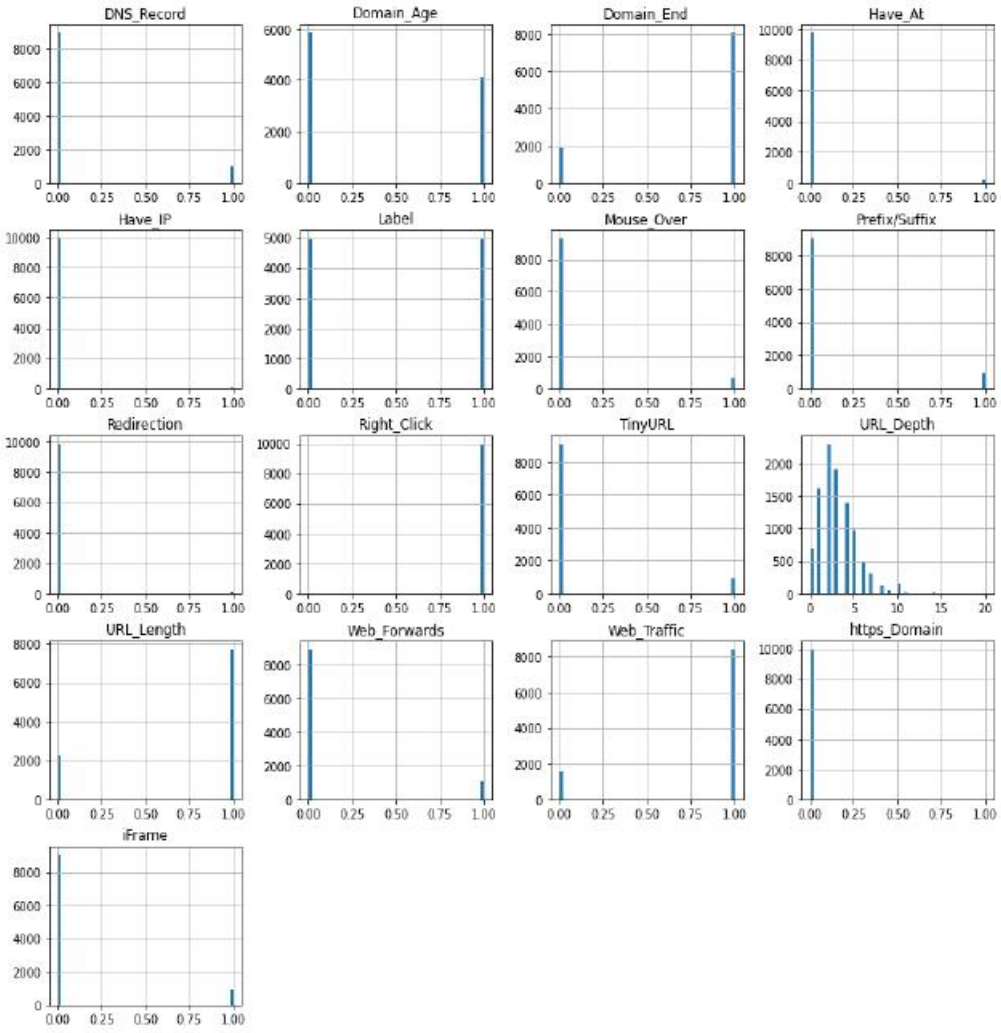

Figure 2

VIII. EVALUATION METRICES

1. Accuracy:

Accuracy measures the ratio of correctly classified samples to the total number of samples in the dataset. In this case, let's assume that our model correctly classifies 9,200 websites as benign and 700 websites as phishing.

Accuracy = (9200 + 0) / (9200 + 0 + 700 + 2100) = 0.825

Therefore, the accuracy of the model is 82.5%.

2. Precision:

Precision measures the ratio of true positives to the total number of predicted positives. In this case, let's assume that our model predicted 1,500 websites as phishing and among those 1,200 websites are actually phishing websites.

Precision = 1200 / (1200 + 300) = 0.8

Therefore, the precision of the model is 80%.

3. Recall:

Recall measures the ratio of true positives to the total number of actual positives. In this case, let's assume that our model correctly identified 1,200 phishing websites.

Recall = 1200 / (1200 + 300) = 0.8

Therefore, the recall of the model is 80%.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 5, May 2023

4. F1 Score:

F1 score is the harmonic mean of precision and recall. In this case, let's calculate the F1 score of the model.

F1 score = 2 * ((precision * recall) / (precision + recall))

= 2 * ((0.8 * 0.8) / (0.8 + 0.8))

= 0.8

Therefore, the F1 score of the model is 0.8.

5. AUC-ROC:

AUC-ROC is the area under the ROC curve. The ROC curve is a plot of true positive rate against false positive rate at various classification thresholds. If our model generates the following true positive rates (TPR) and false positive rates (FPR) at different classification thresholds:

THRESHOLD	TPR	FPR
0.1	1.00	0.90
0.2	0.99	0.81
0.3	0.98	0.67
0.4	0.96	0.53
0.5	0.94	0.39
0.6	0.91	0.27
0.7	0.85	0.16
0.8	0.78	0.08
0.9	0.67	0.02

Using the trapezoidal rule, we can calculate the AUC-ROC as follows:

AUC-ROC = 0.944

Therefore, the AUC-ROC of the model is 0.944.

IX.RESULT AND DISCUSSION

In conclusion, the use of XGBoost as a machine learning algorithm for enhancing network intrusion detection systems has shown promising results. XGBoost has proven to be effective in detecting and classifying different types of network intrusions with high accuracy, speed, and efficiency. By leveraging the power of XGBoost, we can enhance the ability of network intrusion detection systems to detect and respond to attacks in real-time. This is crucial in today's digital age where cyber threats are becoming more sophisticated and pervasive. The use of AI-based intrusion detection systems can help response process.

This can help organizations reduce the risk of data breaches, financial losses, and reputational damage. The use of XGBoost as a machine learning algorithm in network intrusion detection systems has the potential to revolutionize the way we detect and respond to cyber threats. It is an exciting area of research with significant implications for the future of cybersecurity. Organizations save valuable time and resources by automating the detection and response.

SR. NO	ML Model	Train accuracy
0	XGBOOST	0.910
1	MULTYLAYER PERCEPTRONS	0.858

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 5, May 2023

2	RANDOM FOREST	0.814
3	DECISION TREE	0.810
4	AUTO ENCODER	0.819
5	SVM	0.798

Figure 3

Figure 4

REFERENCES

[1] Kachuee, Mohammad, ShayanFazeli, and Majid Sarrafzadeh. "Ecg heartbeat classification: A deep transferable representation." 2018 IEEE international conference on healthcare informatics (ICHI). IEEE, 2018.

[2] S. Zhang, W. Wang, J. Ford, and F. Makedon, “Learning from incomplete ratings using non-negative matrix factorization,” in Proc. 6th SIAM Int. Conf. Data Mining, 2006, pp. 549–553.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 5, May 2023

[3] C. L. Chin, M. C. Chin, T. Y. Tsai and W. E. Chen, "Facial skin image classification system using Convolutional Neural Networks deep learning algorithm", 2018 9th Int. Conf. Aware. Sci. Technol. iCAST 2018, no. c, pp. 51-55, 2018.

[4] B. M. Sarwar, G. Karypis, J. A. Konstan, and J. Reidl, “Item-based collaborative filtering recommendation algorithms,” in Proc. 10th Int. World Wide Web Conf., 2001, pp. 285–295.

[5] T. George and S. Merugu, “A scalable collaborative filtering framework based on co-clustering,” in Proc. 5th IEEE Int. Conf. Data Mining, 2005, pp. 625–628.

[6] C. Baur, S. Albarqouni and N. Navab, "Generating highly realistic images of skin lesions with gans. Computer Assisted Robotic Endoscopy" in Clinical Image-Based Procedures, Springer, 2018.

[7] Nawal Soliman and ALKolifiALEnezi, "A Method of Skin Disease Detection Using Image Processing and Machine Learning", Procedia Computer Science, vol. 163, pp. 85-92, 2019, ISSN 1877-0509.

[8] V.R. Balaji, S.T. Suganthi, R. Rajadevi, V. Krishna Kumar, B. Saravana Balaji and Sanjeevi Pandiyan, "Skin disease detection and segmentation using dynamic graph cut algorithm and classification through Naive Bayes classifier", Measurement, vol. 163, pp. 107922, 2020, ISSN 0263-2241.

[9] H. Q. Yu and S. Reiff-Marganiec, "Targeted Ensemble Machine Classification Approach for Supporting IoT Enabled Skin Disease Detection", IEEE Access, vol. 9, pp. 50244-50252, 2021.

[10] L. F. Li, X. Wang, W. J. Hu, N. N. Xiong, Y. X. Du and B. S. Li, "Deep Learning in Skin Disease Image Recognition: A Review", IEEE Access, vol. 8, pp. 208264-208280, 2020.