An AI-Powered Integrated System for Real-Time Cybersecurity Threat Detection

Bhoomika N B[1], Arya A[2], Nikhil Shastry[3], Ashith[4], Mrs. Thrisha V S[5]
[1,2,3,4]*UG Student, Department of Computer Science and Engineering, Sir M Visvesvaraya Institute of Technology, Bengaluru, Karnataka, India*
[5]*Associate Professor, Department of Computer Science and Engineering, Sir M Visvesvaraya Institute of Technology, Bengaluru, Karnataka, India*

***Abstract*—The expanding size and sophistication of cyber threats have made conventional security measures insufficient. This paper describes an AI-Powered Cybersecurity Threat Detection System based on machine learning algorithms to classify and identify threats like malware, phishing, and brute-force attacks in real time. We enhanced the publically available data by using Generative Adversarial Networks (GANs) to synthesize realistic attack data and the combined dataset was then used to train the machine learning models such as Random Forest and Decision Tree classifiers to provide high detection accuracy. It is executed with Python and deployed via a Flask-based web interface, allowing users to upload network data and obtain instant threat analysis. Experimental results show high precision, recall, and overall accuracy, confirming the effectiveness of the system in early threat detection. The method drastically minimizes manual effort and response time, providing a scalable and automated solution for improving cybersecurity defenses.**

***Index Terms*—Artificial Intelligence, Cybersecurity, Machine Learning, Threat Detection.**

I. INTRODUCTION

In today's digitally connected world, cybersecurity has become a critical concern due to the growing frequency and sophistication of cyber threats. Traditional rule-based security systems often fail to detect novel or evolving attacks, leading to significant data breaches and financial losses. To address these challenges, Artificial Intelligence (AI) and Machine Learning (ML) are increasingly being adopted for their ability to analyze vast datasets and identify patterns indicative of malicious activity. This paper presents an AI-powered threat detection system that leverages supervised learning techniques to detect and classify threats such as malware, phishing, and brute-force attacks. The system offers real-time analysis through a user-friendly web interface, enhancing proactive defense mechanisms.

II. LITERATURE SURVEY

The increasing sophistication of cyberattacks has accelerated the development of intelligent threat detection systems using Artificial Intelligence (AI) and Machine Learning (ML). Traditional signature-based methods, while effective against known threats, fail to adapt to new and evolving attack vectors. Recent research has focused on applying supervised learning algorithms such as Random Forest, Support Vector Machines (SVM), and Decision Trees to classify threats with notable accuracy. For example, Random Forest classifiers have been widely used for their ensemble learning capability, reducing overfitting and improving generalization in malware detection. Additionally, Generative Adversarial Networks (GANs) have emerged as a powerful approach in this domain, particularly for generating synthetic malware samples to augment training datasets. By leveraging the adversarial training process, GANs enhance the robustness of classifiers like Random Forest and SVM, enabling better detection of novel and evolving threats.

It demonstrated the effectiveness of Decision Trees in detecting phishing attempts with minimal computational cost, while other works explored SVM for brute-force attack classification. Additionally, deep learning models such as Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks have shown strong performance in anomaly-based intrusion detection systems (IDS).

However, their complexity and resource-intensive nature often limit real-time deployment.
Datasets such as NSL-KDD, CICIDS2017, and UNSW-NB15 have been commonly used to benchmark model performance, offering diverse scenarios for training and testing. Despite the promising results, many existing systems lack scalability, user accessibility, or real-time capabilities. This project addresses those limitations by implementing a Flask-based application that combines user interaction with efficient ML models, enabling real-time detection of multiple cyber threats in a practical, deployable form.

III. METHODOLOGY

The development of the AI-Powered Cybersecurity Threat Detection System followed a structured methodology, incorporating both data-driven machine learning techniques and practical web application development. The aim was to create a real-time, scalable, and easy-to-use system capable of identifying and classifying various cybersecurity threats such as phishing, brute-force attacks, and malware activity.

1. Data Collection and Preprocessing

The first step involved gathering relevant datasets that contained labeled instances of network traffic, both normal and malicious. Publicly available cybersecurity datasets such as the UNSW-NB15, CICIDS2017, and custom datasets from repositories were analyzed. These datasets include traffic generated from real-world attack scenarios and are rich in features such as IP headers, port numbers, protocols, and payload data.
Once collected, the data underwent a preprocessing stage. This included handling missing values, removing duplicate entries, and converting categorical data into numerical format using techniques like label encoding and one-hot encoding. Normalization was also applied to bring all numerical features to the same scale, ensuring consistent model training.

2. Augmented Dataset Synthesis

Once the data is collected and preprocessed, we used Generative Adversarial Networks (GANs). The GAN framework, comprising a generator and a discriminator was trained to produce realistic attack samples that closely mimic the statistical characteristics of the preprocessed original dataset. The synthetic attack data was combined with the original dataset to create an augmented dataset, enhancing its size and diversity. This augmentation introduced new features, such as varied attack patterns and edge cases not fully represented in the original data, thereby improving the dataset's representativeness

3. Feature Selection

Feature selection was a crucial step to improve model accuracy and reduce computational complexity. We applied correlation analysis and decision tree-based feature importance ranking to identify and retain only the most relevant attributes. This helped in reducing noise in the data and allowed the machine learning models to focus on features that contributed most to classification accuracy.

4. Model Selection and Training

We selected two well-known classification algorithms for this task: Decision Tree and Random Forest. The Decision Tree algorithm was chosen for its simplicity and interpretability, while Random Forest, being an ensemble of multiple decision trees, was selected for its high accuracy and robustness against overfitting. The cleaned and feature-selected dataset was split into training and testing sets, typically in an 80:20 ratio. The models were trained on the training data, and hyperparameter tuning was performed using grid search and cross-validation techniques to optimize model performance. Evaluation metrics such as accuracy, precision, recall, and F1-score were calculated to assess each model's effectiveness.

5. Model Integration and Web Interface Development

Once the models were trained and validated, the next step involved integrating them into a user-friendly application. We developed a web based interface using HTML, CSS, and JavaScript for the frontend, and used Flask (a lightweight Python web framework) for the backend. Flask facilitated the seamless integration of our trained ML models into the application.
Users can upload their network traffic data files in CSV format through the web interface. Upon submission, the data is preprocessed and passed

through the trained model in real time. The system then analyzes the input and displays whether the data contains any malicious activity and classifies the type of threat.

6. Output and Visualization
The final output is displayed in a simple and informative manner on the web interface. Each prediction is accompanied by a confidence score, and users are informed whether the uploaded data is "safe" or poses a "threat," along with the type of threat detected (e.g., malware, phishing, or brute-force). The system provides immediate feedback, making it suitable for real-time monitoring and decision-making.

This methodology ensures that the proposed system is both technically robust and practically usable. It demonstrates how machine learning can be effectively applied to cybersecurity for proactive threat detection, thereby helping organizations and individuals safeguard their digital infrastructure.

IV. RESULTS AND DISCUSSIONS

The AI-Powered Cybersecurity Threat Detection System was evaluated using a variety of real-world datasets, focusing on detecting malware, brute-force attacks, and phishing attempts. The system employs Generative Adversarial Networks (GANs) to generate synthetic data, augmenting the training dataset and improving the model's ability to generalize and detect previously unseen threats. The use of GANs has been particularly beneficial in addressing data imbalance issues, which are common in cybersecurity datasets where certain attack types are underrepresented.

After training the model with this enhanced dataset, the system achieved an impressive 99% accuracy in detecting and classifying various types of cyber threats. The model's high accuracy is attributed to the robust feature extraction process, combined with the power of GANs for data augmentation. Precision, recall, and F1-score metrics were also evaluated, with all of them indicating exceptional performance. The system consistently demonstrated a low false-positive rate, ensuring that legitimate activities were not incorrectly flagged as malicious.

The real-time performance of the system was another key focus. By integrating the trained model into a Flask-based web application, users can upload data and receive threat analysis results almost instantly. This rapid response is essential in modern cybersecurity, where timely intervention can prevent significant damage from attacks.

The results validate the effectiveness of AI-powered threat detection, especially with the innovative use of GANs. While the system shows outstanding performance in controlled environments, future work will focus on further enhancing its adaptability in dynamic network environments, considering emerging threats and evolving attack techniques.

V. ACKNOWLEDGMENT

The authors would like to express our heartfelt gratitude to Mrs. Thrisha V S, our project guide, for her constant support, guidance, and valuable insights throughout the course of this project. Her expertise and encouragement have been instrumental in shaping the direction of our research.

We would also like to extend our sincere thanks to the Department of Computer Science and Engineering at Sir M Visvesvaraya Institute of Technology, Bengaluru, Karnataka, for providing the necessary resources and infrastructure that facilitated the successful completion of this project. The department's support in terms of academic and technical resources has been invaluable.

REFERENCES

[1] S. M. Shinde and V. S. Rajpurohit, "Anomaly Based Intrusion Detection Using Machine Learning," International Journal of Computer Applications, vol. 180, no. 17, pp. 20–24, Mar. 2018.
[2] D. Dua and C. Graff, "UCI Machine Learning Repository," University of California, Irvine, School of Information and Computer Sciences, 2019
[3] M. Ring, S. Wunderlich, D. Scheuring, D. Landes, and A. Hotho, "A survey of network-based intrusion detection data sets," Computers & Security, vol. 86, pp. 147–167, Sep. 2019.

[4] M. Panda and M. R. Patra, "Network Intrusion Detection Using Naive Bayes," International Journal of Computer Science and Network Security (IJCSNS), vol. 7, no. 12, pp. 258–263, Dec. 2007.

[5] S. Revathi and A. Malathi, "A Detailed Analysis on NSL-KDD Dataset Using Various Machine Learning Techniques for Intrusion Detection," International Journal of Engineering Research and Technology (IJERT), vol. 2, no. 12, pp. 1848–1853, Dec. 2013.

[6] A. Buczak and E. Guven, "A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection," IEEE Communications Surveys & Tutorials, vol. 18, no. 2, pp. 1153–1176, Second Quarter 2016.

[7] S. B. Kotsiantis, "Supervised Machine Learning: A Review of Classification Techniques," Informatica, vol. 31, no. 3, pp. 249–268, 2007.

[8] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron C. Courville, and Yoshua Bengio. "Generative adversarial nets. In Advances in Neural Information Processing Systems", 2014.