International Journal of Contemporary Research in Computer Science and Technology (IJCRCST) e-ISSN: 2395-5325
Volume 9, Issue 1 (April '2026)

CYBER ATTACK PREDICTION FROM TRADITIONAL MACHINE LEARNING TO GENERATIVE ARTIFICIAL INTELLIGENCE

Dr. D Banumathy,
Head and Professor,
Department of Computer Science and Engineering,
Paavai Engineering College,
Namakkal,Tamilnadu,India.
Email Id: peccsehod@paavai.edu.in

V Maheskumar,
Assistant Professor,
Department of Computer Science and Engineering,
Paavai Engineering College,
Namakkal, Tamil Nadu, India.
Email Id: mahestamil@gmail.com

G Sanjay,
PG Scholar,
Department of Computer Science and Engineering
Paavai Engineering College,
Namakkal, Tamil Nadu, India.
Email Id: sanjaypmv2002@gmail.com

Abstract: *The rapid expansion of digital infrastructure has significantly increased the complexity and frequency of cyber threats, necessitating advanced approaches for accurate and timely attack prediction. Traditional machine learning techniques have been extensively employed to analyze network traffic and detect anomalies based on historical data, offering reliable performance in identifying known attack patterns; however, their dependence on predefined features and labelled datasets often limits their ability to generalize emerging and sophisticated threats. In contrast, recent advancements in generative artificial intelligence introduce a more adaptive paradigm by enabling models to learn underlying data distributions and generate realistic representations of potential attack scenarios, thereby enhancing the capability to anticipate previously unseen vulnerabilities. This paper examines the evolution from conventional machine learning methods to generative AI-driven frameworks for cyber-attack prediction, emphasizing their comparative strengths and the potential for integrated approaches to improve the robustness, adaptability, and overall effectiveness of modern cybersecurity systems in dynamic environments.*

Keywords: Cyber Attack Prediction, Machine Learning, Generative Artificial Intelligence, Anomaly Detection, Network Security, Deep Learning, Threat Intelligence.

I.INTRODUCTION

The contemporary digital ecosystem has become an essential foundation for modern society, supporting critical domains such as financial systems, industrial infrastructure, communication networks, and national security frameworks. This extensive digital dependence has simultaneously expanded the attack surface, resulting in a rapidly evolving cybersecurity landscape marked by increasing frequency, sophistication, and impact of cyber threats. Unlike earlier periods where attacks were largely isolated and less coordinated, present-day cyber incidents are often orchestrated by highly skilled adversaries, including state-sponsored groups, organized cybercriminal networks, and ideologically driven entities. These actors leverage advanced tools, automation, and strategic planning, challenging the effectiveness of conventional security mechanisms. Historically, cybersecurity strategies have relied on reactive defense models, where systems such as firewalls and signature-based intrusion detection systems operate by identifying known patterns of malicious activity. While effective against previously documented threats, these approaches are inherently limited by their dependence on predefined signatures and historical knowledge. As a result, they struggle to detect emerging threats that do not match existing patterns, creating a critical gap in defense capabilities, particularly in the presence of zero-day vulnerabilities and continuously evolving malware variants.

The limitations of reactive security approaches become more pronounced when addressing advanced cyber threats such as Advanced Persistent Threats (APTs), which are characterized by their stealth, persistence, and ability to evade detection over extended periods. Similarly, polymorphic and metamorphic malware introduce additional complexity by dynamically altering their structure, effectively bypassing traditional

detection mechanisms. This evolving threat landscape highlights the inadequacy of static defense strategies and emphasizes the need for more adaptive and intelligent security solutions capable of handling uncertainty and novelty.

To address these challenges, the cybersecurity domain has gradually shifted towards proactive defense methodologies, where the focus extends beyond detection to prediction and prevention. Traditional machine learning techniques have played a significant role in this transition by enabling systems to learn behavioral patterns from data and identify anomalies indicative of potential attacks. Algorithms such as Support Vector Machines, Random Forests, and clustering methods have demonstrated effectiveness in distinguishing between normal and malicious activities without relying solely on explicit signatures. However, these models are often constrained by their reliance on structured datasets, extensive feature engineering, and limited ability to generalize across diverse and dynamic environments.

The emergence of generative artificial intelligence introduces a transformative direction in cyber-attack prediction by moving beyond pattern recognition to data generation and simulation. Unlike conventional models that focus on classification tasks, generative approaches learn the underlying distribution of data, enabling the creation of synthetic yet realistic representations of potential attack scenarios. This capability opens new avenues for enhancing cybersecurity systems, including the generation of adversarial data for robust model training, mitigation of data scarcity challenges, and simulation of complex multi-stage attack behaviors. By leveraging these capabilities, generative models provide a more flexible and forward-looking framework for understanding and anticipating evolving cyber threats.

II. PROBLEM STATEMENT

Existing cybersecurity systems largely depend on signature-based and rule-driven approaches, which are reactive in nature and ineffective against emerging threats such as zero-day attacks and advanced persistent threats. These systems struggle to detect unknown attack patterns and often generate a high number of false positives, leading to alert fatigue and reduced efficiency in security operations. Although traditional machine learning techniques improve detection through pattern recognition, they are limited by data imbalance, dependency on labeled datasets, and lack of predictive capability. Hence, there is a need for an intelligent framework that can accurately detect and proactively predict cyber-attacks in dynamic environments.

III.OBJECTIVES

The objective of this research is to develop a hybrid cyber-attack prediction system that combines traditional machine learning with generative artificial intelligence to improve detection accuracy and predictive performance. The study aims to analyze existing methods, design and implement a hybrid model, and evaluate its effectiveness using standard performance metrics. Additionally, it focuses on reducing false positives, enhancing detection of novel attacks, and leveraging generative models to create synthetic data for improved learning and prediction.

IV.LITERATURE SURVEY

The field of cyber-attack prediction has witnessed significant advancements with the integration of machine learning and artificial intelligence techniques. Various research works have explored different approaches ranging from traditional algorithms to advanced deep learning and generative models. In this section, five relevant literature studies are analyzed to understand their methodologies, strengths, and limitations, which form the foundation for the proposed hybrid approach.

Iman Sharafaldin et al. (2017) presented a detailed analysis of the CICIDS2017 dataset for intrusion detection system design. The methodology focuses on generating realistic network traffic with diverse attack scenarios using controlled environments and extracting over 80 flow-based features through CICFlowMeter. This dataset enables effective training and evaluation of intrusion detection models by closely simulating real-world conditions. The structured labeling and inclusion of modern attack types make it highly suitable for machine learning applications. However, the dataset still faces limitations such as class imbalance and potential redundancy in features, which may affect model performance. Additionally, it may not fully represent evolving real-time attack patterns. Despite these limitations, it serves as a standard benchmark for evaluating cybersecurity models.

Paul A. Watters and Ross Brown (2018) conducted a comprehensive review of Random Forest algorithms for intrusion detection. Their methodology involves analyzing multiple studies where Random Forest is applied for classification of network traffic based on extracted features. The model's ensemble learning capability improves accuracy and reduces overfitting, making it effective for high-dimensional data. Feature importance ranking further enhances interpretability and aids in feature selection. However, the approach is limited by its static nature, requiring retraining to adapt to new attack patterns. It also lacks predictive capabilities for unseen threats and depends heavily on labeled datasets. Consequently, while effective for known attacks, its adaptability to dynamic environments remains restricted.

M. Usama et al. (2019) explored the application of Generative Adversarial Networks (GANs) for network intrusion detection. The methodology is based on the adversarial training process between a generator and a discriminator, where synthetic attack data is generated to enhance model training. This approach helps in addressing class imbalance and improves detection of rare attack patterns. GANs also enable simulation of potential attack scenarios, contributing to better generalization. However, training GANs is computationally intensive and may suffer from instability issues such as mode collapse. Additionally, evaluating the quality of generated data remains a challenge. Despite these drawbacks, GAN-based approaches offer strong potential for predictive cybersecurity.

Raghavendra Chalapathy and Sanjay Chawla (2019) reviewed deep learning techniques for anomaly detection in cybersecurity. Their methodology includes the use of models such as Autoencoders, Variational Autoencoders, RNNs, and GANs to learn complex data patterns. These models automatically extract features from raw data and detect anomalies using reconstruction errors or learned representations. Deep learning techniques are highly effective in capturing non-linear relationships in network traffic. However, they require large datasets and high computational resources for training. Moreover, the lack of interpretability makes it difficult for analysts to understand model decisions. These limitations highlight the need for combining deep learning with more interpretable methods.

Tianqi Chen and Carlos Guestrin (2016) introduced XG Boost, a scalable and efficient gradient boosting framework widely used in machine learning applications. The methodology focuses on optimized tree boosting with regularization, parallel processing, and efficient handling of missing data. XG Boost provides high accuracy and speed, making it suitable for real-time intrusion detection systems. Its feature importance capability aids in dimensionality reduction and model optimization. However, the model relies on structured input data and requires careful feature engineering. It also lacks the ability to generate new data or predict unseen attack patterns. Thus, while powerful for classification, it needs to be complemented with generative approaches for enhanced prediction.

VI.EXISTING SYSTEM

Existing cyber-attack detection systems have evolved through multiple paradigms, including signature-based, anomaly-based, traditional machine learning, deep learning, and emerging generative AI approaches. Each of these systems contributes unique strengths in detecting cyber threats but also exhibits critical limitations when dealing with dynamic and sophisticated attack patterns. This section analyzes five key existing approaches to understand their methodologies, advantages, and inherent challenges, which highlight the need for more adaptive and predictive frameworks.

Signature-Based Detection Systems

Signature-based systems rely on predefined patterns or known attack signatures to detect malicious activities within network traffic. The methodology involves comparing incoming data against a continuously updated database of signatures using tools such as rule-based intrusion detection systems. This approach ensures high accuracy and low false positives when detecting known threats due to its deterministic nature. It also enables fast processing, making it suitable for real-time network monitoring. However, these systems fail to detect zero-day attacks and polymorphic malware, as no prior signature exists for such threats. The approach is inherently reactive and requires constant manual updates. Additionally, large signature databases can lead to maintenance overhead and alert fatigue, reducing operational efficiency.

Anomaly-Based Detection Systems

Anomaly-based systems detect cyber-attacks by identifying deviations from established normal behaviour patterns. The methodology involves building statistical or machine learning models to define baseline network activity and flag unusual deviations as potential threats. Techniques such as clustering, density estimation, and one-class classification are commonly used. This approach enables detection of previously unseen attacks and provides adaptability to evolving environments. However, accurately defining "normal" behaviour is challenging in dynamic networks. These systems often generate high false positive rates due to legitimate but unusual activities. The models are sensitive to training data quality and may fail under concept drift conditions. As a result, practical deployment is limited by alert overload and reduced reliability.

Traditional Machine Learning-Based Systems

Traditional machine learning systems utilize supervised and semi-supervised algorithms to classify network traffic as normal or malicious. The methodology involves training models such as Support Vector Machines, Random Forests, and XGBoost using labeled datasets with extracted features. These systems improve detection accuracy by learning complex patterns and reduce false positives compared to anomaly-based methods. Feature importance techniques also enhance interpretability and model optimization. However, their performance is highly dependent on the availability and

quality of labeled data. They struggle with class imbalance, where attack instances are rare compared to normal traffic. Additionally, these models are static and require frequent retraining to adapt to evolving threats. The need for manual feature engineering further limits scalability and generalization.

Deep Learning-Based Systems

Deep learning approaches leverage multi-layer neural networks to automatically learn complex representations from network data. The methodology includes models such as Convolutional Neural Networks, Recurrent Neural Networks, and Autoencoders for feature extraction and anomaly detection. These systems can capture non-linear relationships and temporal dependencies, improving detection of sophisticated and multi-stage attacks. They also reduce reliance on manual feature engineering by learning directly from raw data. However, deep learning models require large volumes of training data and high computational resources. Their black-box nature reduces interpretability, making it difficult for analysts to trust predictions. They are also vulnerable to adversarial attacks and performance degradation due to concept drift. These challenges limit their practical adoption in real-time systems.

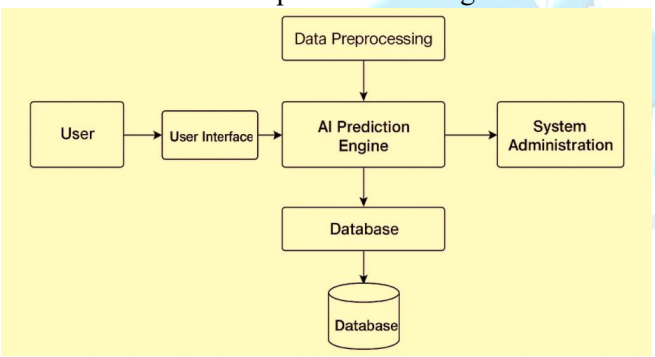


Figure 6.1. Block Diagram

The proposed system presents an advanced hybrid cybersecurity framework that integrates traditional machine learning techniques with Generative Artificial Intelligence (GenAI) to address the critical limitations of existing reactive security mechanisms and enable proactive cyber-attack prediction. As depicted in the proposed system block diagram, the architecture follows a structured and scalable pipeline in which the user interacts with the system through a dedicated user interface, enabling data visualization, monitoring, and control functionalities.

The input data, consisting of network traffic logs, packet captures, and system event records, is first processed through a robust data preprocessing layer designed to handle heterogeneous and high-volume data streams. This layer performs essential operations such as data cleaning, normalization, transformation, and feature extraction, where meaningful attributes like packet size distribution, flow duration, protocol behavior, and temporal patterns are derived using both domain knowledge and automated techniques.

Generative AI-Based Systems

Generative AI systems represent an emerging approach in cybersecurity, focusing on data generation and simulation for improved threat detection. The methodology involves models such as Generative Adversarial Networks and Variational Autoencoders to generate synthetic attack data and simulate potential threat scenarios. These systems help address class imbalance and enhance model robustness by providing diverse training samples. They also enable predictive capabilities by modeling future attack patterns. However, generative models are computationally expensive and difficult to train effectively. Ensuring the quality and realism of generated data remains a major challenge. Additionally, there are concerns regarding misuse of generative models for creating advanced attack strategies. Despite these limitations, they offer strong potential for next-generation cybersecurity solutions.

VII. PROPOSED SYSTEM

The processed data is then forwarded to the core AI prediction engine, which forms the intelligence backbone of the system by combining the strengths of traditional machine learning models, such as XGBoost, with generative models like Variational Autoencoders (VAE) or Generative Adversarial Networks (GANs). The machine learning component performs high-speed classification and baseline anomaly detection by learning discriminative patterns from labeled data, while the generative component captures underlying data distributions and identifies deviations through reconstruction errors or adversarial evaluation, enabling the detection of previously unseen or zero-day attacks.

Furthermore, the generative model contributes to system robustness by creating realistic synthetic attack data, which helps in addressing class imbalance and enhances the generalization capability of the overall model. A fusion mechanism integrates the outputs from both components, combining classification probabilities and anomaly scores to generate a final prediction enriched with confidence levels and contextual insights, thereby reducing false positives and improving decision reliability. The prediction results are then managed by the system administration module, which oversees user authentication, access control, system configuration, and performance monitoring, ensuring secure and efficient system operation.

Simultaneously, all processed data, prediction outcomes, and system logs are stored in a centralized database that supports real-time querying, historical analysis, and continuous model improvement through retraining processes. The entire system operates in a

continuous loop of data ingestion, preprocessing, prediction, and feedback incorporation, where analyst feedback and newly observed patterns are utilized to update and refine the models, allowing the system to adapt dynamically to evolving threat landscapes.

Additionally, the system is designed to support real-time alert generation with prioritized risk levels, enabling security analysts to focus on critical threats without delay. The integration of visualization dashboards enhances interpretability by presenting predictions, anomaly trends, and system insights in an intuitive manner. The architecture also supports modular upgrades, allowing new models, features, or data sources to be incorporated without affecting existing system performance. Furthermore, the system ensures data security through controlled access mechanisms and secure storage practices, maintaining confidentiality and integrity of sensitive information. This extended capability strengthens the overall efficiency, usability, and reliability of the proposed cyber-attack prediction framework.

7.1. Methodology

The methodology of the proposed cyber-attack prediction system is designed to implement a hybrid analytical framework that integrates traditional machine learning with generative artificial intelligence for accurate and proactive threat detection. The approach follows a structured pipeline that includes data collection, preprocessing, model development, hybrid prediction, and evaluation. Each stage is carefully designed to ensure efficient handling of large-scale network data, improved detection of unknown attacks, and continuous system adaptability.

Data Collection and Dataset Selection

The initial step involves collecting network traffic data from reliable sources to train and evaluate the system. Standard datasets such as CICIDS2017 are utilized due to their realistic representation of modern cyber-attacks, including DoS, DDoS, brute force, and web-based attacks. The dataset contains both normal and malicious traffic instances with multiple features representing network behavior. Data may also be collected from real-time network environments to enhance system applicability. The use of diverse and well-labeled datasets ensures that the model learns comprehensive attack patterns and improves generalization.

Data Preprocessing and Feature Engineering

7.2.Key Parameters

S. No	Parameter	Description	Role in System
1	Dataset (CICIDS2017)	Realistic network traffic dataset with modern attack types	Used for training and evaluating ML and GenAI models
2	Flow Duration	Time duration of a network connection	Helps identify abnormal long/short connections
3	Packet Count	Number of packets	Used to detect traffic

The collected raw data is processed to remove noise, handle missing values, and convert it into a structured format suitable for analysis. Data normalization and encoding techniques are applied to ensure consistency across features. Feature engineering is performed to extract meaningful attributes such as packet size, flow duration, protocol type, and traffic patterns. Dimensionality reduction techniques may be used to eliminate redundant features and improve computational efficiency. This stage plays a critical role in enhancing model performance by providing high-quality input data.

Traditional Machine Learning Model Implementation

In this stage, a traditional machine learning model such as XGBoost is implemented for baseline classification of network traffic. The model is trained using labeled data to distinguish between normal and malicious activities. It learns patterns from historical data and provides probability-based predictions. Feature importance analysis is also performed to identify the most significant attributes influencing predictions. This model ensures fast and accurate classification, making it suitable for real-time applications.

Generative AI Model Development

To enhance prediction capability, a generative model such as a Variational Autoencoder (VAE) or Generative Adversarial Network (GAN) is developed. The model learns the underlying distribution of normal network behavior and detects anomalies based on deviations. It can also generate synthetic attack data to address class imbalance and improve model robustness. This approach enables the system to identify previously unseen attack patterns, which are not detectable using traditional methods alone.

Hybrid Prediction and Performance Evaluation

The final stage involves integrating the outputs of both traditional and generative models using a fusion mechanism. The system combines classification results and anomaly scores to generate a final prediction with higher accuracy and reliability. Performance is evaluated using metrics such as accuracy, precision, recall, F1-score, and ROC-AUC. The system is also tested for its ability to detect zero-day attacks and reduce false positives. Continuous feedback and retraining mechanisms are incorporated to improve performance over time.

		transmitted in a flow	spikes and flooding attacks
4	Byte Count	Total data transferred in a flow	Indicates unusual data transfer behavior
5	Protocol Type	Type of protocol (TCP/UDP/ICMP)	Helps classify traffic and detect protocol misuse
6	Source/Destination IP	Network addresses of communication	Used for identifying suspicious communication patterns
7	Feature Dimension	Number of extracted features (80+)	Improves model accuracy through detailed representation
8	ML Model (XGBoost)	Gradient boosting algorithm	Performs fast and accurate classification
9	GenAI Model (VAE/GAN)	Generative model for anomaly detection	Detects unknown attacks and generates synthetic data
10	Anomaly Score	Measure of deviation from normal behavior	Helps identify potential cyber threats
11	Prediction Accuracy	Performance metric of the model	Evaluates effectiveness of the system
12	False Positive Rate	Incorrect attack detection rate	Ensures reliability and reduces alert fatigue

7.3.Flow Diagram

The flow diagram represents the step-by-step working process of the proposed cyber-attack prediction system using a hybrid approach integrated with Generative Adversarial Networks (GAN).

The process begins with the data source, where raw data is obtained from various inputs such as network traffic, system logs, and user-generated data. This data is then passed to the data collection stage, where relevant information is gathered and organized for further processing.

In the next stage, data cleaning is performed to remove noise, missing values, and inconsistencies present in the collected data. This step ensures that the dataset is accurate and suitable for model training. After preprocessing, the cleaned dataset is used in the training phase, where a GAN model is applied to learn data distribution and generate synthetic samples. This helps in improving the dataset quality and addressing issues like class imbalance.

Following this, the system performs classification of cyber-attacks using machine learning techniques. In this stage, the processed data is analyzed to identify whether the activity is normal or malicious and to determine the specific type of cyber-attack. The classification output is then passed to the prediction model, which generates the final prediction based on learned patterns and anomaly detection.

Finally, the detection model utilizes the prediction results to identify and report potential cyber threats. The output is then presented to the user or system administrator for further action. This complete flow ensures a continuous and efficient process of data handling, model training, attack classification, and prediction, enabling accurate and proactive cyber-attack detection.

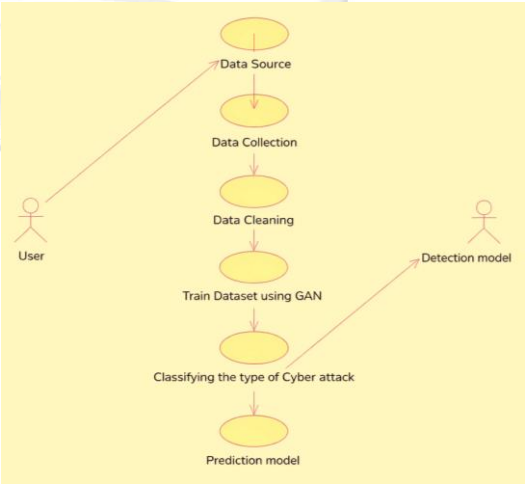


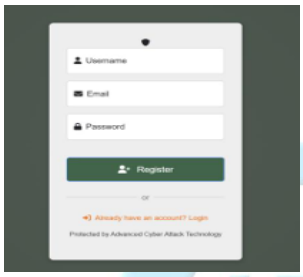
Figure 9.1 Flow Diagram

VIII. RESULT AND DISCUSSION

The performance of the proposed hybrid cyber-attack prediction system was evaluated using a benchmark dataset containing both normal and malicious network traffic instances. The system was tested on its ability to accurately classify known attacks, detect anomalies, and identify previously unseen threat patterns. The integration of traditional machine learning with generative artificial intelligence demonstrated a noticeable improvement in overall detection performance when compared to standalone models.

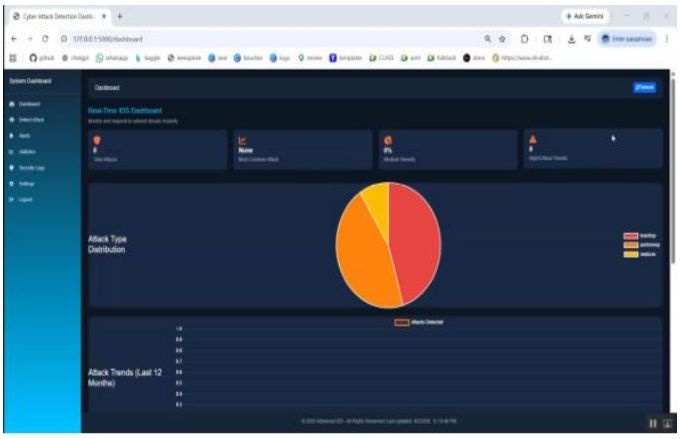
The results indicate that the hybrid model achieved high classification accuracy, with improved precision and recall values, ensuring that most attack instances were correctly identified while minimizing false alarms. The traditional machine learning component provided fast and reliable baseline classification, whereas the generative model enhanced the system's capability to detect anomalies and zero-day attacks by learning underlying data distributions. This combination resulted in a balanced performance, reducing both false positives and false negatives. A key observation from the experimental results is the system's effectiveness in handling class imbalance. The use of generative models to create synthetic attack samples significantly improved the detection rate of rare attack categories, which are typically underrepresented in standard datasets. This contributed to better generalization and robustness of the model across diverse attack scenarios.

Register:

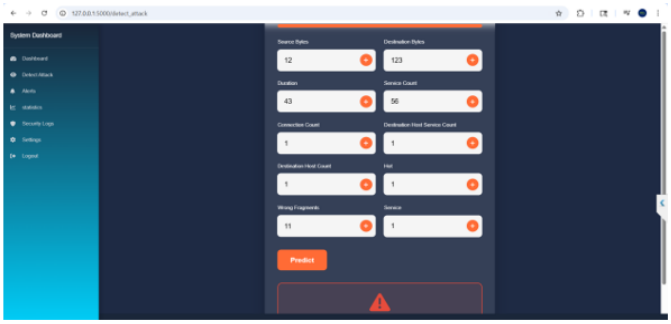


Login:

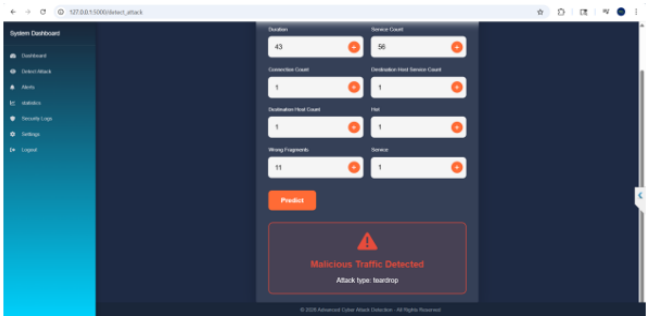
Dashboard:



Prediction:



Detection:



In terms of computational performance, the system demonstrated efficient processing of network data with acceptable latency, making it suitable for near real-time applications. Although the inclusion of generative models increased computational complexity, the overall system maintained a good trade-off between accuracy and processing speed. The modular design also allowed optimization of individual components without affecting the entire pipeline.

The discussion further highlights that the proposed system adapts effectively to evolving cyber threats through continuous learning and feedback mechanisms. By incorporating new data and user feedback into retraining processes, the model improves over time and maintains its relevance in dynamic environments. However, the system's performance is influenced by the quality and diversity of the training dataset, and further improvements can be achieved by incorporating larger real-time datasets and optimizing model parameters.

Advantage Of the Proposed System

The proposed system provides an advanced and efficient solution for cyber-attack prediction by integrating traditional machine learning with generative artificial intelligence. This hybrid approach enhances detection accuracy, enables identification of unknown threats, and supports proactive security measures. The system is designed to handle large-scale network data in real time while maintaining reliability and adaptability, making it suitable for modern dynamic cybersecurity environments.

- ❖ Enables proactive prediction of cyber-attacks instead of only detection
- ❖ Improves accuracy and reliability using hybrid ML + GenAI model
- ❖ Detects zero-day and unknown attacks effectively

- Reduces false positives and false negatives
- Handles class imbalance using synthetic data generation
- Supports real-time data processing and analysis
- Requires less manual feature engineering
- Provides scalable and flexible architecture
- Supports continuous learning and model improvement
- Enhances overall cybersecurity decision-making.

IX.CONCLUSION

This research work presented a comprehensive and advanced hybrid framework for cyber-attack prediction by effectively integrating traditional machine learning techniques with generative artificial intelligence models. The study addressed the critical limitations of existing cybersecurity systems, particularly their reactive nature and inability to detect unknown or zero-day attacks. By leveraging the strengths of models such as XG Boost for fast and accurate classification along with generative models like Variational Autoencoders and Generative Adversarial Networks for anomaly detection and data generation, the proposed system achieved a balanced combination of speed, accuracy, and adaptability. The preprocessing pipeline ensured that high-quality and structured data was provided to the models through effective normalization, feature extraction, and data cleaning techniques. The hybrid prediction mechanism successfully combined classification probabilities and anomaly scores to produce more reliable and context-aware predictions.

The experimental results demonstrated that the proposed system performs effectively across various attack scenarios, including rare and previously unseen threats. The use of synthetic data generation helped in overcoming class imbalance issues, improving the model's ability to generalize across different types of cyber-attacks.

Another important area for future work is the implementation of explainable artificial intelligence techniques, which can provide transparency and interpretability to model predictions. This will help security analysts understand the reasoning behind alerts and improve trust in automated systems. The system can also be extended to include automated response mechanisms, where detected threats can trigger predefined mitigation actions such as blocking suspicious IP addresses or isolating compromised systems, thereby reducing response time and minimizing damage. Furthermore, integration with cloud-based and edge computing environments can enhance flexibility and enable deployment across distributed infrastructures. The screenshots show the step-by-step working of the system from Register and Login to Dashboard, Prediction, and Detection. It clearly explains how the user interacts with the system and how cyber-attacks are identified.

X.REFERENCES

[1]. Sharafaldin, I., Lashkari, A. H., & Ghorbani, A. A. (2018). Toward generating a new intrusion detection dataset and intrusion traffic characterization. ICISSP.

[2]. Watters, P. A., & Brown, R. (2008). Random Forests for Intrusion Detection: A Comprehensive Review. Journal of Information Warfare.

[3]. Usama, M., Qadir, J., Al-Fuqaha, A., & Imran, M. A. (2019). Generative Adversarial Networks for Network Intrusion Detection: A Survey. IEEE Communications Surveys & Tutorials.

[4]. Chalapathy, R., & Chawla, S. (2019). Deep Learning for Anomaly Detection: A Review. arXiv preprint arXiv:1901.03407.

[5]. Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.

[6]. Goodfellow, I., et al. (2014). Generative Adversarial Networks. Advances in Neural Information Processing Systems.

[7]. Kingma, D. P., & Welling, M. (2013). Auto-Encoding Variational Bayes. arXiv preprintarXiv:1312.6114.

[8]. CICFlowMeter. (n.d.). GitHub Repository. Canadian Institute for Cybersecurity.

[9]. Chollet, F., et al. (2015). Keras. GitHub Repository.

[10]. Pedregosa, F., et al. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research.

