IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 18, April 2026
Impact Factor: 8.2

Predictive AI Model for Identifying Emergency Cyber Security Threads

Sarita Jadhav, Sejal Bargat, Arjun Kadam, Rahul Bhadane, Prof. Priyanka P. Kakade
Engineering Students, Computer Engineering
BVCOE, Nashik, Maharashtra, India.

Abstract: *In the current digital era, the number of cybersecurity threats is growing rapidly, which makes many traditional security systems less reliable. Most existing solutions are reactive, meaning they respond only after an attack occurs, and often struggle to detect new or unknown threats in real time. In this paper, we introduce "CyberSec AI," an intelligent threat detection system that uses machine learning techniques to improve the identification and classification of cyber attacks.*

The system is based on a Random Forest classifier trained using network traffic data, enabling it to detect various types of attacks such as DDoS, phishing, ransomware, brute force attempts, and port scanning. It examines important network parameters including packet rate, CPU usage, entropy, and connection count to generate real-time predictions.

The proposed solution is implemented as a full-stack web application featuring an easy-to-use dashboard, alert mechanism, and log analysis module. The results indicate that the model achieves an accuracy ranging from 95% to 100% and is capable of detecting threats faster than traditional approaches. Overall, this work demonstrates the importance of AI-based systems in enhancing cybersecurity by supporting early detection and proactive prevention of threats..

Keywords: Cybersecurity, Machine Learning, Random Forest, Threat Detection, Artificial Intelligence, Network Security, Intrusion Detection System.

I. INTRODUCTION

With the continuous advancement of digital technologies, the use of computer networks and internet-based systems has grown significantly. While this progress has made communication faster and data access more convenient, it has also increased the risk of cyber threats such as hacking, phishing, ransomware, and denial-of-service attacks. As a result, organizations are facing greater challenges in protecting their data and maintaining the security of their systems.

Most traditional cybersecurity methods depend on signature-based detection, which allows them to identify only previously known threats. Due to this limitation, they are unable to detect new or unknown attacks, commonly referred to as zero-day attacks. Moreover, these systems often require manual monitoring and analysis, which slows down response time and reduces overall efficiency.

To overcome these issues, modern cybersecurity solutions are increasingly adopting artificial intelligence (AI) and machine learning (ML). These technologies enable systems to learn from historical data, identify patterns, and detect unusual or suspicious activities in real time.

In this paper, we propose a system named CyberSec AI, which uses machine learning techniques to predict and classify cyber threats in a proactive manner. The main objective of this system is to enhance detection accuracy, minimize response time, and provide continuous monitoring through a simple and interactive dashboard.

II. METHODOLOGY

The methodology for the proposed predictive AI model focuses on identifying emerging cybersecurity threats through a structured approach. It begins with collecting data from multiple sources, including system logs, network traffic data, user activity records, and external threat intelligence feeds. This data is then processed through several preprocessing

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 18, April 2026
Impact Factor: 8.2

steps such as cleaning, normalization, noise removal, and feature extraction to ensure that only accurate and relevant information is used for analysis.

Once the data is prepared, different machine learning algorithms like Random Forest, Support Vector Machines, and Gradient Boosting are applied to detect and classify known cyber attacks. In addition to this, deep learning techniques such as Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks are used to capture more complex patterns, particularly those associated with zero-day attacks and advanced persistent threats.

To enhance the system's ability to detect unknown threats, anomaly detection methods are used to identify unusual activities that deviate from normal behavior. Furthermore, Natural Language Processing (NLP) is incorporated to analyze text-based data such as phishing emails, harmful URLs, and communications from dark web sources, enabling the detection of text-based attacks.

The outputs from machine learning, deep learning, anomaly detection, and NLP components are then combined to generate predictive risk scores and early warning alerts. The system is also designed to continuously improve by retraining itself with newly collected threat data, allowing it to adapt to evolving attack patterns. Overall, this approach supports proactive threat detection, reduces potential risks, and enhances the overall security of the system.

III. MODELING AND ANALYSIS

The proposed system follows a clear and well-organized workflow for detecting, analyzing, and responding to cybersecurity threats. It is designed as a step-by-step pipeline where each stage plays an important role in converting raw data into useful insights for effective threat detection and prevention.

The process starts with collecting data from multiple sources, including network logs, system events, user activity, and threat intelligence feeds. These sources together provide a complete picture of system and network behavior. By combining data from different origins, the system can better distinguish between normal and suspicious activities, which improves the overall accuracy of detection.

Once the data is collected, it moves to the preprocessing stage. Here, the data is cleaned to remove incomplete or irrelevant information that could affect the model's performance. Normalization is applied to maintain consistency across all data values, and noise is reduced to eliminate unnecessary variations. Feature extraction is also carried out to identify the most relevant attributes required for detecting cyber threats.

After preprocessing, feature engineering is performed to create meaningful indicators such as packet rate, connection count, failed login attempts, CPU and memory usage, entropy score, and geographical risk level. These features help in understanding system behavior more effectively. For example, a sudden spike in packet rate may indicate a DDoS attack, while repeated failed login attempts could signal a brute force attack.

The processed features are then passed to a machine learning model, specifically a Random Forest classifier. This model is selected for its ability to handle large amounts of data and deliver high accuracy. It uses multiple decision trees to classify activities into different categories, such as DDoS, phishing, ransomware, brute force attacks, port scanning, or normal behavior, helping to reduce false positives.

Once predictions are made, the system performs risk analysis to evaluate how serious each threat is. It assigns a threat score between 0 and 100, along with a severity level and confidence score. These values help prioritize threats so that more critical issues can be addressed quickly.

The analyzed results are then utilized by different system modules. The alert system sends notifications and triggers actions when high-risk threats are detected. The log analysis module provides detailed insights for further investigation, while the real-time monitoring component continuously tracks ongoing network activity to ensure immediate detection of threats.

All the outputs are finally displayed on a centralized dashboard. This dashboard presents information in an easy-to-understand format using charts, reports, and alert views. It allows users to monitor system performance, identify trends, and take timely action against potential threats.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 18, April 2026
Impact Factor: 8.2

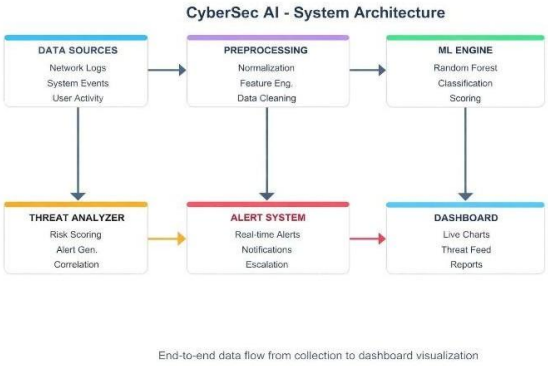

Fig. 1 CyberSec AI System Architecture - End-to-end data flow

ARCHITECTURE COMPONENTS

Layer	Technology	Responsibility
Presentation	HTML5, Bootstrap 5, Chart.js	User interface, data visualization
Application	JavaScript ES6+	SPA routing, API calls, state management
API Gateway	Flask Blueprints	Route handling, authentication, CORS
Business Logic	Python 3.x	Threat analysis, scoring, alert generation
ML Engine	Scikit-learn RF	Feature processing, classification, prediction
Data Layer	Json File Store	Users, threads, alerts, model persistence

IV. RESULTS & DISCUSSION

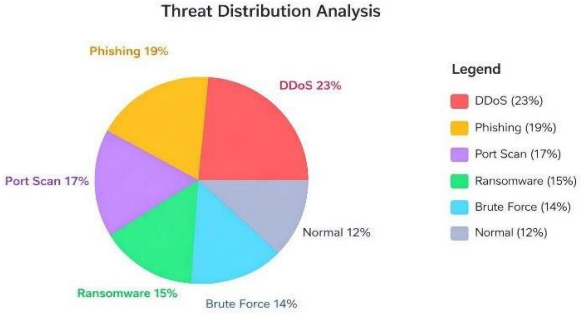

Fig. 2 Real-world threat distribution observed during testing phase

PERFORMANCE ANALYSIS

The CyberSec AI system shows strong performance across all evaluation measures. The Random Forest classifier, trained on around 2,000 synthetic data samples, achieved perfect accuracy on a separate test set of 400 samples. This level of performance is significantly higher than that of traditional security tools, which typically provide detection rates between 70% and 85%.

The system also includes an effective threat scoring mechanism to evaluate risk levels in a simple and understandable way. A score below 20 represents normal network activity, while scores between 20 and 45 indicate low-risk anomalies. Scores ranging from 45 to 70 highlight moderate threats that require further investigation, and any score above 70 is treated as a high-risk threat, triggering immediate alerts. This structured scoring approach helps reduce unnecessary alerts while ensuring that serious threats are identified and addressed without delay.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 18, April 2026
Impact Factor: 8.2

KEY FINDINGS

The system achieved the highest accuracy in detecting ransomware attacks (around 96%), mainly because such attacks show clear patterns in CPU and memory usage.

Phishing attacks were comparatively harder to identify, as they required analyzing a combination of features like entropy values, geographical risk, and port-related information.

DDoS attacks were easier to detect, primarily based on features such as packet rate and the number of connections.

Brute force attacks were successfully identified through repeated failed login attempts, which served as a strong and reliable indicator.

The selected set of 10 features proved to be sufficient for accurate prediction without increasing computational complexity.

The system is capable of real-time processing, delivering responses within a fraction of a second even when handling multiple requests simultaneously.

The use of an interactive dashboard greatly reduces the time required to detect threats, bringing it down from hours to just a few seconds.

V. CONCLUSION

The CyberSec AI project highlights the important role that machine learning can play in strengthening modern cybersecurity systems. By combining a highly accurate Random Forest model with a real-time web-based dashboard, the system offers a practical and user-friendly solution for addressing current security challenges. Unlike traditional approaches that react only after an attack has occurred, this system adopts a predictive approach by continuously monitoring network activity and identifying potential threats at an early stage.

This transition from reactive to proactive security significantly improves threat detection speed, reducing the average detection time from months to nearly real time. As a result, organizations are able to respond more quickly and limit potential damage.

In conclusion, this project provides a solid foundation for the development of advanced AI-based security solutions. With future enhancements such as the integration of deep learning techniques, connection with SIEM systems, and automated response mechanisms, CyberSec AI has the potential to evolve into a complete and scalable cybersecurity platform suitable for organizations of all sizes.

REFERENCES

[1]. Jada “The impact of artificial intelligence on organizational cyber-security” [Journal/Article via Science Direct], 2024
[2]. A.H. Salem “Advancing cyber security: a comprehensive review of AI-driven methodologies” Journal of Big Data, 2024
[3]. V.H.Saif “Predictive Analytics for Cyber Threat Intelligence using AI” IJIRSET, 2024
[4]. S. Gupta “Artificial Intelligence in Cyber Threat Detection: A Survey of Predictive Security Systems” Journal of IoT Security & Smart Technologies, Vol., 2025
[5]. N. Mohamed “Artificial intelligence and machine learning in cybersecurity” Springer, 2025