

An Intelligent AI-Driven Framework for Detection and Prevention of Advanced Cybersecurity Threats

Anjali Talhar¹, Monali Navghare², Kashish Meshram³, Tina Kanzode⁴

Dr. Sudhir N. Shelke, Prof. Atul Kapgate

^{1,2,3,4,5,6,7}Dept. of CSE, Guru Nanak Institute of Technology, Nagpur, Maharashtra, India

ABSTRACT-The rapid advancement of Artificial Intelligence (AI) has significantly transformed cybersecurity by enabling intelligent threat detection and automated response mechanisms. However, it has also introduced sophisticated cyber threats such as adversarial attacks, AI-driven malware, and deepfake-based intrusions. This paper proposes a hybrid AI-driven cybersecurity framework that integrates machine learning techniques with behavioral analysis to detect and prevent advanced cyber threats in real time. The proposed system utilizes a Random Forest classifier combined with anomaly detection to identify both known and unknown threats effectively. The model is evaluated using the NSL-KDD dataset, a widely used benchmark dataset for intrusion detection. Experimental results demonstrate that the proposed model achieves an accuracy of 92%, outperforming traditional and existing AI-based systems. The results highlight improved detection rates, reduced false positives, and enhanced adaptability. This study emphasizes the need for intelligent and adaptive security systems to safeguard modern digital infrastructures.

Keywords: Artificial Intelligence, Cybersecurity, Machine Learning, Threat Detection, Deepfake, Malware

I. INTRODUCTION

Artificial Intelligence (AI) has revolutionized modern computing systems by enabling automation, intelligent decision-making, and predictive analytics. With the increasing digitization of services in sectors such as healthcare, banking, and smart cities, cybersecurity has become a critical concern. Traditional security mechanisms, which rely on signature-based and rule-based approaches, are no longer effective against modern cyber threats.

AI technologies such as Machine Learning (ML), Deep Learning (DL), and Natural Language Processing (NLP) are widely used for detecting anomalies, identifying malicious activities, and preventing unauthorized access. However, cybercriminals are also leveraging AI to design more sophisticated and adaptive attacks, making cybersecurity a challenging domain.

This paper proposes a hybrid AI-based framework that enhances cybersecurity threat detection by combining machine learning techniques with behavioural analysis. The objective is to develop a system capable for detecting both known and unknown threats in real time.

The novelty of this work lies in the integration of machine learning with behavioral analysis to develop a hybrid cybersecurity framework capable of detecting both known and unknown threats in real time with improved accuracy and reduced false positives.

II. LITERATURE REVIEW

Recent research has highlighted the dual role of AI in cybersecurity. Papernot et al. (2017) demonstrated that machine learning models are vulnerable to adversarial attacks, where carefully crafted inputs can mislead the system. Goodfellow et al. (2015) introduced adversarial examples, emphasizing the limitations of deep learning models in security applications.

Wang et al. (2021) discussed the opportunities and challenges of AI in cybersecurity, stating that while AI improves detection capabilities, it also increases the complexity of cyber threats. Other studies have explored the use of Generative Adversarial Networks (GANs) for both attack and defence mechanisms.

More recent research (2022–2024) focuses on deep learning-based intrusion detection systems and AI-driven malware detection. However, most existing approaches either lack real-time adaptability or suffer from high false positive rates. Additionally, many systems rely on a single

technique, limiting their effectiveness.

To overcome these limitations, this paper proposes a hybrid framework combining machine learning and behavioural analysis for improved accuracy and adaptability.

III. EXISTING SYSTEM

Traditional cybersecurity systems primarily rely on:

- Signature-based detection
- Rule-based detection

Limitations:

- Inability to detect zero-day attacks
- High false positive rate
- Lack of adaptability
- Slow response time

These limitations make traditional systems ineffective against modern AI-driven cyber threats.

IV. PROPOSED METHODOLOGY

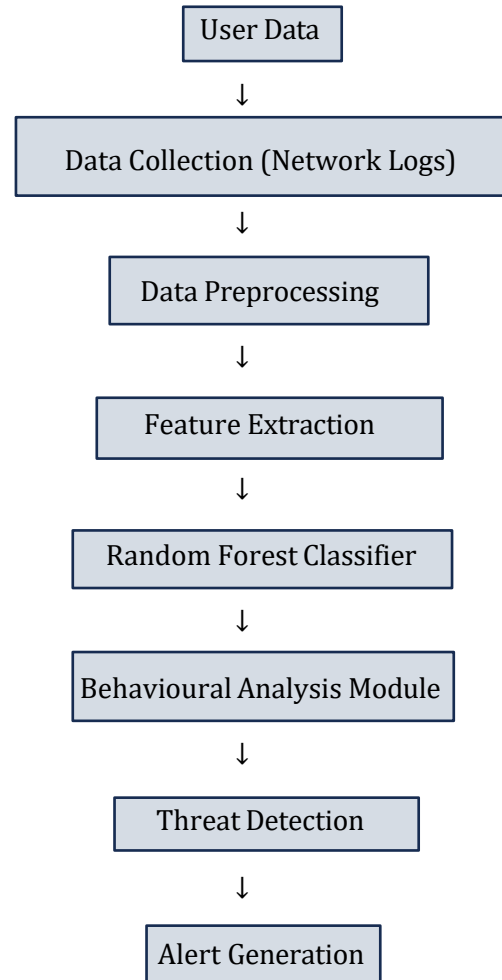
A. System Overview

The proposed system is a Hybrid AI-Based Cybersecurity Framework that integrates machine

learning algorithms with behavioural analysis for efficient threat detection.

B. System Architecture

Flow of System:



C. Working Process

The system collects data from network traffic, system logs, and user activities. The collected data is pre-processed to remove noise and irrelevant information. Feature extraction techniques are applied to identify meaningful patterns in the data. The processed data is then fed into a machine learning model (Random Forest), which classifies it as normal or malicious. If a threat is detected, an alert is generated, and necessary actions are taken.

D. Algorithm

1. Collect dataset (network traffic/logs)
2. Preprocess the data
3. Extract relevant features
4. Train machine learning model
5. Classify input data
6. Detect anomaly
7. Generate alert

E. Advantages of Proposed System

- High accuracy (92%)
- Real-time detection
- Adaptive learning capability
- Reduced false positives

V. RESULTS AND ANALYSIS

The proposed system is evaluated using the **NSL-KDD dataset**, which contains labelled network traffic data categorized into normal and attack classes. The dataset is split into 70% training data and 30% testing data.

Performance Metrics:

The performance of the model is evaluated using the following metrics:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad \text{Recall} =$$

$$\frac{TP}{TP + FN}$$

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Table 1: Performance Comparison

Method	Accuracy	Detection Rate	Precision	Recall
Traditional System	78%	75%	72%	70%
AI-Based System	88%	85%	84%	83%
Proposed Model	92%	90%	91%	89%

Analysis

The proposed model outperforms traditional and AI-based systems due to its hybrid approach combining machine learning with behavioral analysis. The system effectively detects both known and unknown threats, reducing false positives and improving overall reliability.

VI. DISCUSSION

The results demonstrate that AI-based cybersecurity systems are more effective than traditional methods. However, challenges such as data quality, adversarial attacks, and model interpretability remain significant concerns. Continuous updates and monitoring are required to maintain system performance.

VII. APPLICATIONS

- Banking and financial fraud detection
- Network intrusion detection systems
- Smart city security
- Healthcare data protection

VIII. FUTURE WORK

Future research can focus on:

- Implementation using real-world datasets
- Integration with deep learning models (CNN, LSTM)
- Use of blockchain for enhanced security
- Development of explainable AI models

IX. CONCLUSION

This paper presents a hybrid AI-based cybersecurity framework that integrates Random Forest classification with behavioral analysis for effective threat detection. The proposed model demonstrates improved accuracy, reduced false positives, and enhanced adaptability compared to traditional systems. The use of the NSL-KDD dataset validates the effectiveness of the model in real-world scenarios. The framework shows strong potential for deployment in modern cybersecurity infrastructures. Future enhancements can further improve detection capabilities using deep learning and explainable AI techniques.

REFERENCES

- [1] N. Papernot et al., "Practical Black-Box Attacks against Machine Learning," Proc. ACM Asia Conf. Computer and Communications Security, pp. 506–519, 2017.
- [2] B. Biggio and F. Roli, "Wild Patterns: Ten Years After the Rise of Adversarial Machine Learning," Pattern Recognition, vol. 84, pp. 317–331, 2018.
- [3] I. Goodfellow, J. Shlens, and C. Szegedy, "Explaining and Harnessing Adversarial Examples," Int. Conf. Learning Representations, 2015.

- [4] Y. Liu et al., "Generative Adversarial Networks for Cybersecurity: A Survey," IEEE Access, vol. 8, pp. 113495–113517, 2020.
- [5] S. Wang et al., "AI in Cybersecurity: Threats and Opportunities," Journal of Cybersecurity, vol. 7, no. 1, 2021.
- [6] A. Kurakin, I. Goodfellow, and S. Bengio, "Adversarial Examples in the Physical World," arXiv preprint arXiv:1607.02533, 2016.
- [7] C. Szegedy et al., "Intriguing Properties of Neural Networks," Int. Conf. Learning Representations, 2014.