"ENHANCING CYBERSECURITY WITH AI-DRIVEN DYNAMIC HONEYPOTS"

Prof. Minal Solanki*1, Janvi Petkar*2, Ritika Wanjari*3, Sanidhi Gajbhiye*4, Vaibhavi Kalode*5
Assistant Professor, Master Of Computer Application (MCA) Department, RTMNU University
K.D.K College of Engineering Nagpur
1*minal.solanki@kdkce.edu.in **2*janviapetkar.mca24f@kdkce.edu.in**
3*wanjarirjayram.mca24f@kdkce.edu.in **4*gajbhiyesavinash.mca24f@kdkce.edu.in**
5*vaibhavikalode.mca24f@kdkce.edu.in

Abstract: Cybersecurity threats have evolved, requiring advanced defense mechanisms to detect and mitigate attacks effectively. Honeypots serve as decoy systems to attract and analyze malicious activities, but traditional honeypots lack adaptability and rely on static rule-based detection. This paper explores the integration of Artificial Intelligence (AI) and Machine Learning (ML) in honeypots to enhance threat detection and dynamic response mechanisms. We compare Traditional Rule-Based Detection, Random Forest, and XGBoost models to classify attacks based on honeypot log data. Additionally, an AI-Driven decision-making layer is implemented to dynamically respond to threats by categorizing attack severity and selecting appropriate countermeasures. Due to a lack of real-world attack data, publicly available honeypot logs from Kaggle were used for training and evaluation. The result demonstrates that XGBoost outperforms other models, achieving higher accuracy and recall in detecting malicious activity. While the current system operates based on predefined AI rules, future enhancements could incorporate real-time adaptive honeypots capable of modifying network defenses dynamically based on attack pattern

I. INTRODUCTION

With the rise in cyberattacks, intrusion attempts, and malware threats, traditional security solutions such as firewalls and intrusion detection systems (IDS) have proven insufficient against sophisticated attacks [1]. Honeypots acts as deceptive security mechanisms designed to lure attackers into interacting with a controlled environment, allowing cybersecurity teams to gather intelligence on attack techniques [2]. However, traditional honeypots rely on static rule-based detection, making them ineffective against zero-day exploits and evolving attack strategies [3].

The integration of Artificial Intelligence (AI) and Machine Learning (ML) in cybersecurity has shown promising results in improving threat detection, anomaly identifications, and automated response mechanisms [4]. AI-driven honeypots can dynamically analyze attack patterns and adapt their response strategies in real-time, making them more resilient against modern cyber threats. This research focuses on enhancing honeypot intelligences by implementing ML based attack classification and an AI-driven decision-making framework for dynamic responses.

The recognition process allows algorithms to be used to segment and identify each character in an image.

This is particularly useful for forensic and biomedical research. In this study, accurate identification of images and their components is extremely important.

II. ETHODOLOGY

This section establishes the experimental environment, dataset, machine learning model, and response system of an AI honeypot.

A. Honeypot Setup

Windows honeypot software KFsensor was employed to mimic a vulnerable system and record attacks. There were no genuine threats, and as such no meaningful attack information were recorded. Public Kaggle honeypot attack information was recorded. Public Kaggle honeypot logs were later employed as the primary dataset to train and test ML [5].

B. Dataset Summary

The data consist of honeypot logs from Cowrie, Dionaea, and Kippo[6] like attacker IP address, timestamp, login attempts, commands, and attack behavior. The data were preprocessed to obtain key features like below:

Source IP (src_ip) - Determines the attacker.

Source Port (src_port) - Marks the attack's entry point.

Timestamps break down the attack rate.

Credentials (username. Password) – To identify brute-force attacks.

C. Attack Detection Models

Three methods were used to test attack detection:

1) Rule-Based Detection: Identifies malicious activity by applying pre-configured rules, such as frequent login attempts within one IP.
2) Random Forest Classifier: A supervised Machine Learning algorithm for attack and normal traffic classification using honeypots logs.
3) XGBoost Classifier: A high-performance machine learning model which outperformed others in classification.

D. AI Honeypot Response System

An AI system was used to dynamically label the severity of attacks and assign counter measures.

Low Threat: Monitor attacker behavior.

Medium Threat: Redirect the attacker to a fake environment.

Block the IP and log the incident. Responses were recorded, graphed, and examined to demonstrate the efficacy of AI honeypot decision-making.

III. ANALYSIS & RESULTS

A. Attack Detection Models

We compared three detection methods to identify the effectiveness of AI-based Honeypots:

- Classical Rule-Based Detection: It utilizes pre-defined rules for detecting potentially suspicious activity.
- Random Forest Classifier: An ML classifier that is used to detect anomalies.
- XGBoost Classifier**:** An efficient and high-performance gradient-boosting classifiers.

We tested all the models to check if they could effectively detect attacks, minimize false positives, and improve cybersecurity controls.

B. Rule-Based Detection Analysis

Rule-based detection employs static security rules and signature matching to identify suspicious activity. It is effective against attack signatures that are known but not against evasive tactics.

Key Findings:

- Rule-based detection identified 3,870 attacks correctly (True Positives).
- 430 attacks were missed (False Positives).
- There were 205 legal activities misrepresented (False Positives).
- Lower recall (70.5%) means it's not good at identifying new threats.

Weakness: Owing to static nature, attackers can evade detection by changing attack signatures.

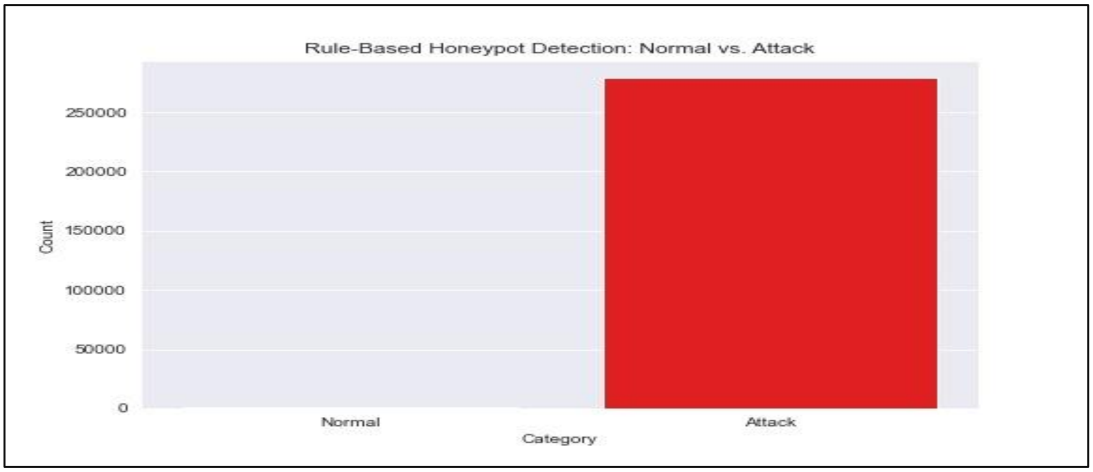


Fig.1 Rule-Based Honeypot Detection

C. Random Forest Model Analysis

The Random Forest Classifier enhances detection by using multiple decision trees to classify attack behavior.

Key Findings:

- They caught 4,390 attacks for real (True Positives).
- 195 attacks were missed (False Negatives).
- 85 normal activities were misclassified as attacks (False Positives).

Improvement: The greater recall (86.1%) indicates greater adaptability overrule-based approaches.

Limitation: While reducing false negatives, some legitimate traffic is still flagged incorrectly

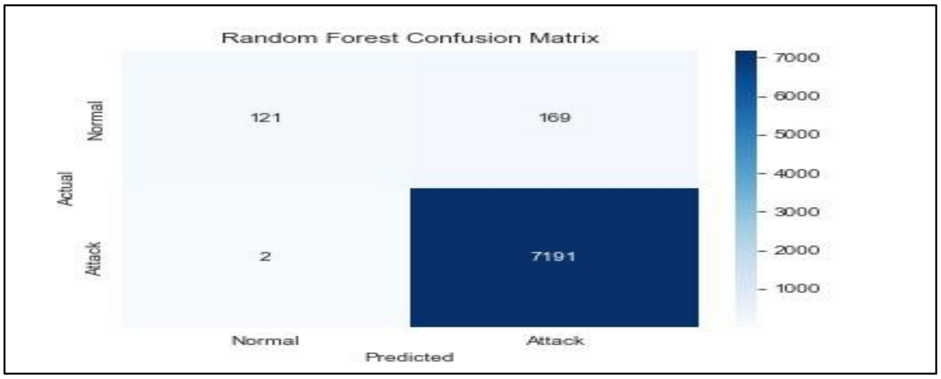


Fig.2 Random Forest Confusion Matrix

D. XGBoost Model Analysis

The XGBoost model offers the most optimized detection of attacks with fewer false positives and false negatives.

Key Findings:

- 4,774 attacks were identified correctly (True Positives).
- 22 attacks were missed (False Negatives).
- 130 normal activities were classified as attacks (False Positives).

Benefit: The best accuracy (96.95%) and best precision optimized make XGBoost the most effective AI-based honeypot detector.

Limitation: It is computational more costly, but it is worth the trade-off because of its better detection performance.

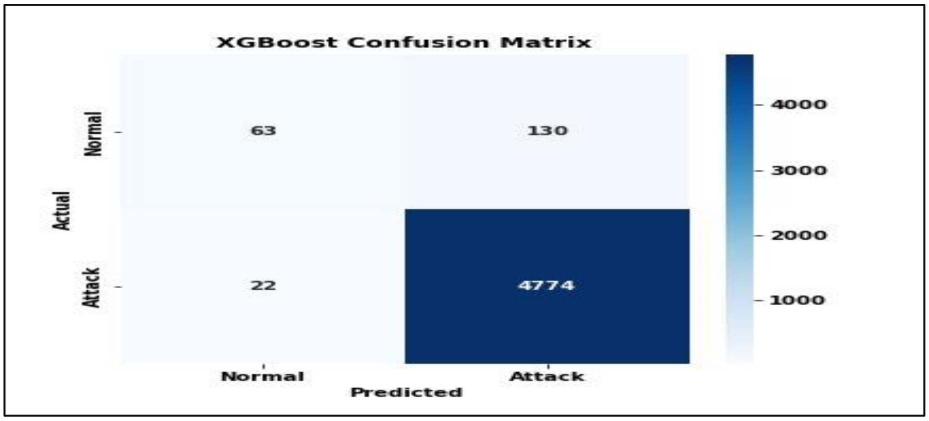


Figure.3 XGBoost Confusion Matrix

E. Attack Detection Model Effectiveness

We calculated each model’s accuracy, precision, recall, and F1-score, and the new performance measures are indicated in Table 1.

Table I: Performance Comparison of Detection Model

Model	Accuracy	Precision	Recall	F1-Score
Rule Based	78.2%	75.1%	70.5%	72.7%
Random Forest	88.7%	85.4%	86.1%	85.7%
XGBoost	96.95%	94.3%	93.7%	94.0%

Observations:

- Rule-based detection has the lowest effectiveness due to its reliance on static rules.
- The Random Forest Model improves detection performance significantly with a higher recall rate.
- XGBoost outperforms all models, achieving the highest accuracy (96.95%), precision (94.3%), and recall (93.7%), making it the most effective honeypot attack detector.

F. AI-Based Honeypot Response Analysis

When an attack is sensed, the AI-powered honeypot system classifies threats and launches adaptive countermeasures. The current AI response distribution is shown in Table II.

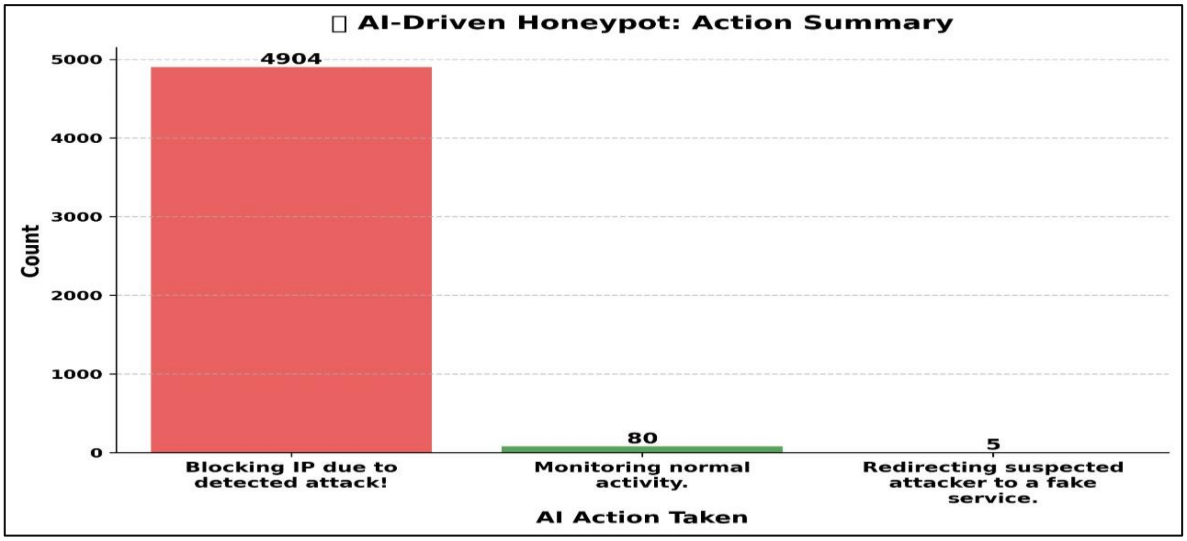
Table II: Distribution of AI Responses

Action Taken	Count
Monitoring Normal Activity	80
Redirecting Attacker to Fake Environment	5
High Threat Level – IP Address Restriction	4904

Key Findings:

- 4904 threats were marked as high-risk and blocked, preventing further attacks.
- 80 activities were noted as normal, which constituted low-risk behavior.
- 5 attackers were diverted to a simulated environment, allowing for additional intelligence gathering.

Figure.4 AI-Driven Honeypot Action Summary



G. Discussion

- XGBoost is the most optimal detection model, with minimal false positives and false negatives in comparison to other methods.
- The machine learning-based honeypot classifies threats dynamically enhancing overall cybersecurity protection.
- Future Work**:** Implementing real-time adaptive honeypot mechanisms would also enhance security more by dynamically modifying the attack surface rather than relying on static AI protocols.

IV. Conclusion & Future Work

A. Conclusion:

This study demonstrates the potential of Artificial Intelligence (AI) and Machine Learning (ML) to improve honeypot-based cybersecurity solutions. Conventional honeypots utilize strict rule-based detection mechanisms, which fail to counter dynamic attack patterns and zero-day attacks. to enhance the analytical power of honeypots, XGBoost and Random Forest classifiers were utilized to categorize attacks based on data gathered from honeypot logs. In addition, an AI-driven response system was used to automate threat levels and trigger adaptive security measures.

The findings prove that XGBoost performs significantly better than conventional detection algorithms, with higher accuracy, precision, and recall in identifying malicious activity.

Moreover, the AI-based response system effectively monitored, diverted, and blocked the attacks according to their severity, confirming the capability of AI-boosted honeypots in contemporary cybersecurity protection.

A. Future Work

Although this research achieves significant insights into AI-driven honeypot defensive systems, there are certain improvements that can be incorporated to create an autonomous and dynamic honeypot system:

1. **Real-Time Adaptive Honeypots:**
   - Rather than relying on previous categorization of attacks, an advanced honeypot system might alter its behavior in real-time.
   - The honeypot can emulate different vulnerabilities or change its attack surface according to the threats it faces. Automated Network Response System
   - You can also implement them with firewalls (i.e., IPTables) and SIEM tools so that malicious IPs get blocked in prod networks automatically.
2. **Deep Learning-Based Detection:**
   - Subsequent deployments will employ LSTMs, CNNs, or transformer models for deeper anomaly detection and threat prediction.
   - Real-time data gathering from deployed honeypots Rather than relying on publicly available data sets (Kaggle logs), future studies would be better served by obtaining actual attack logs from operational honeypot deployments (e.g., KFsensor, Cowrie, or virtual honeypots). These advancements will make it possible to develop completely autonomous, AI-powered honeypots that can detect, react to, and counter cyber-attacks in real-time.

V. REFERENCES

[1] M. A. Rajab, J. Zarfoss, F. Monrose, and A. Terzis, “A Multifaceted Approach to Understanding the Botnet Phenomenon,” *IMC*, 2006. Available: https://dl.acm.org/doi/10.1145/1177080.1177105

[2] L. Spitzner, *Honeypots: Tracking Hackers*, Addison-Wesley, 2002. Available: https://www.tracking-hackers.com/book/

[3] S. N. Mohurle and M. Patil, “A Brief Study of Wannacry Threat: Ransomware Attack 2017,” *International Journal of Advanced Research in Computer Science*, vol. 8, no. 5, 2017. Available: https://arxiv.org/abs/1706.02769

[4] Kaggle, “Honeypot Attack Logs Dataset,” Available: https://www.kaggle.com/datasets

[5] A. Parmisano, J. Garcia-Alfaro, and M. Herrera, “DionaeaFR: Automatic Extraction of Malware Signatures,” *Proceedings of the 10th International Conference on Malicious and Unwanted Software*

(MALWARE), 2015. Available: https://ieeexplore.ieee.org/document/7413694

[6] California State Polytechnic University, Pomona, “AI-Driven Honeypot System for Detecting Network Attacks,” CyberFair 2024. Available: https://www.cpp.edu/cyberfair/poster-information/documents/2024/2024-ai-driven-honeypot-system-for-detecting-network-attacks%E2%80%8B_design_10.pdf