

NEXT-GEN CYBERSECURITY SYSTEM FOR MESSAGE-BASED THREAT ANALYSIS

Kusuma Shree M D

Dept. of Artificial Intelligence and Machine Learning

C. Byregowda Institute of Technology (CBIT)

Kolar, Karnataka, India

Kusumashreeraj26@gmail.com

Harshitha N

Dept. of Artificial Intelligence and Machine Learning

C. Byregowda Institute of Technology (CBIT)

Kolar, Karnataka, India

harshitha614@gmail.com

Manjula S S

Dept. of Artificial Intelligence and Machine Learning

C. Byregowda Institute of Technology (CBIT)

Kolar, Karnataka, India

ssmanjula91@gmail.com

ABSTRACT: As cyber threats continue to escalate, deploying strong security measures to protect sensitive information and essential Infrastructure is now more vital than ever. The paper “Next-Gen Cybersecurity System for Message-Based Threat Analysis” aims to provides a widespread use of email has made it a major and prime target for cyber threats such as phishing, malware, and attacks from insiders, which often bypass traditional security tools. This project introduces an AI-powered Email Threat Detection system that uses machine learning and Natural Language Processing (NLP) to spot and handle these risks in real time. By integrating rule-based methods, behavior analysis, and threat intelligence, the system can catch a wide range of threats, including complex phishing schemes, Advanced Persistent Threats (APTs), and harmful attachments. It uses models like Random Forest, Naive Bayes, and Isolation Forest and features a Flask-based dashboard for real-time monitoring. This solution offers a scalable, automated, and comprehensive approach to securing email communications. It helps organizations by offering valuable insights, reducing false positives, and actively fighting against new and changing cyber threats.

Keywords: *Cybersecurity, Artificial Intelligence (AI), Machine Learning (ML), Natural Language Processing (NLP), Advanced Persistent Threats (APTs), Naive Bayes, Random Forest, and Isolation Forest.*

I. INTRODUCTION

Traditional email security measures, including spam filters, signature-based antivirus software, and rule-based firewalls, operate on static, pre-defined patterns. While effective against known threats, they are fundamentally inadequate in the face of modern, context-aware attacks that leverage social engineering, mimic legitimate communication, and exploit human psychology. This critical vulnerability in the digital defense ecosystem creates an urgent need for intelligent, adaptive, and proactive security solutions. Email has become one of the most essential communication tools for individuals and organizations worldwide. With its widespread adoption, it has also become a primary vector for cyber threats, ranging from simple spam to sophisticated attacks such as phishing, ransomware, Business Email Compromise (BEC), Advanced Persistent Threats (APTs), and insider threats. These threats can result in financial losses, data breaches, reputational damage, and disruption of business operations. Traditional security mechanisms, such as antivirus software, spam filters, and rule-based firewalls, often struggle to detect advanced, context-sensitive threats that exploit human behavior, social engineering tactics, and emerging attack patterns. This gap necessitates intelligent, adaptive, and automated systems capable of real-time threat detection and mitigation. The “Email Threat Detection” project focuses on developing a comprehensive cybersecurity solution leveraging Artificial Intelligence (AI), Machine Learning (ML), and Natural Language Processing (NLP). By combining multiple threat detection strategies, the system can identify and respond to a wide range of email-borne threats effectively. The approach integrates rule-based heuristics, signature detection, and predictive modeling, enabling the system to detect both known and unknown threats. For example, phishing attacks are detected through urgency keyword analysis, domain spoofing, financial lure detection, and behavioral profiling. Similarly, APTs are identified using business context analysis, attachment inspection, and threat intelligence matching. Malware detection relies on file extension checks, ransomware patterns, and social engineering cues, while DDoS attacks are monitored using traffic patterns and botnet indicators. Insider threats are flagged through behavioral monitoring, data exfiltration detection, and anomaly analysis. The system utilizes multiple ML algorithms, including Naive Bayes, Logistic Regression, Random Forest, Isolation Forest, and DBSCAN clustering, supported by TF-IDF vectorization for textual analysis. Integration with Gmail API allows real-time email retrieval and monitoring, while Flask-based dashboards provide visualization and alerting capabilities. Additionally, threading and asynchronous processing ensure efficient handling of large email volumes. The solution is designed for scalability and deployment in real-world environments, employing cloud or distributed computing infrastructure where necessary. By providing explainable AI insights, automated threat response, and continuous monitoring, the system enhances situational awareness and empowers organizations to proactively defend against evolving email threats. Overall, this project

demonstrates the potential of AI- driven cybersecurity solutions in safeguarding digital communication and ensuring operational resilience.

The exponential growth of digital communication has positioned email as a critical tool for both personal and organizational operations. However, this widespread reliance has also made it a prime target for a broad spectrum of sophisticated cyber threats. These evolving attacks, from phishing to Advanced Persistent Threats (APTs), easily bypass traditional security measures like static spam filters and antivirus software. Conventional tools are insufficient because they cannot analyze context or behavior, failing to detect personalized phishing, stealthy malware, or insider threats. The lack of intelligent, real-time analysis often leads to severe financial losses, data breaches, and operational disruptions. Manual and reactive security approaches are simply inadequate in today's fast-paced digital landscape. Therefore, this project addresses the critical need for a comprehensive, AI-driven system capable of proactive, real-time detection. By leveraging machine learning and NLP to analyze content, behavior, and patterns, the solution identifies multiple threat vectors like BEC and DDoS attacks. It provides explainable insights and automated responses to enhance cybersecurity and restore trust in email communication.

II.LITERATURE REVIEW

Email security has been a focal point of cybersecurity research due to the increasing sophistication of attacks. Traditional methods rely heavily on signature-based detection and rule- based filters, which are insufficient against evolving threats. Recent advancements in AI and ML have enabled more dynamic and context-aware solutions. This chapter reviews existing approaches and identifies gaps that this project aims to fill. The escalating sophistication of email-based cyber threats has propelled academic and industrial research into advanced detection methodologies. The journey has evolved from simple, rule-based filtering to the current exploration of Artificial Intelligence (AI) and Machine Learning (ML) as pivotal technologies for proactive defense. This literature survey critically examines the existing body of work in the field of email security, tracing the evolution from traditional systems to intelligent models. It aims to synthesize key findings, identify prevailing trends, and most importantly, pinpoint the specific research gaps and limitations that the present project, "Email Threat Detection," seeks to address. By understanding the strengths and weaknesses of prior approaches, this chapter justifies the necessity for the proposed hybrid AI model.

A detailed analysis in [1] A robust cyber security threat detection model using artificial intelligence technology. Dr. V. Nagagopiraju, V., Ashok, P., Lakshmi, K. D., Likhitha, C., & Kumar, M. V. S. - 2024. Proposes a novel intrusion detection system based on a Twin Support Vector Machine (TSVM) model optimized with a binary grasshopper algorithm (BGOTSVM). The approach

involves ranking and selecting key security features to reduce dimensionality and computational cost. It compares performance against traditional ML models like decision trees, random forests, and artificial neural networks. The model is designed to detect inconsistencies and threats in network data, showing improved predictive accuracy in real-world scenarios. Key focus is on automated, data-driven threat detection using AI to enhance cybersecurity defenses. Future work aims to expand datasets and develop more adaptive, real-time intrusion detection systems.

In reference [2], Artificial intelligence and machine learning in cybersecurity: A deep dive into state-of the art techniques and future paradigms. Knowledge and Information Systems, Mohamed, N. - 2025. This paper gives a comprehensive overview of how AI and ML are being used to strengthen cybersecurity. Modern AI algorithms for threat detection (like deep learning and reinforcement learning) How ML helps automate responses to attacks. Future research directions such as explainable AI, federated learning, and adversarial defense. Key Idea of this AI/ML are becoming the backbone of intelligent cybersecurity systems that can learn and adapt to new threats faster than traditional tools.

Addressing insider threats, [3] Insider threat detection for specific threat scenarios. Cybersecurity, Tian, T. - 2025. Focuses on insider threats — attacks that come from within an organization (like employees or contractors). It studies Different insider threat scenarios (data theft, sabotage, policy violations). ML-based models to detect unusual insider behaviors. Behavioral profiling using logs, access patterns, and communication data. Key Idea of this not all threats come from hackers outside AI can analyze patterns to identify risky actions by insiders too. It also discusses XAI (Explainable AI) and Federated Learning to enhance transparency and data privacy. Future work should address adversarial AI, where attackers manipulate ML models. The work in [4] Digital Twin-Enabled Incident Detection and Response: A Systematic Review of Critical Infrastructures Applications, Konstantinos E. Kampourakis, Vasileios Gkioulos, Georgios Kavallieratos, Jia-Chun Lin – 2025. Explains how Digital Twins — virtual replicas of physical systems — can be used in cybersecurity for critical infrastructure (like power grids, hospitals, and factories). It covers how digital twins simulate attacks safely, Real-time incident detection, Predictive responses and recovery planning. Key Idea of this are digital twins allow proactive defense predicting and responding to cyber incidents before they cause damage. Digital Twins (DTs) are virtual replicas of physical assets, updated in real time with sensor data. This review studies how DTs simulate attack scenarios and detect anomalies.

In, [5] combined feature selection approach for malicious email detection

based on a comprehensive email dataset. Cybersecurity, Zhang, H. – 2025. This paper targets email- based attacks (spam, phishing, malware). It uses feature selection (choosing the most important input features) + ML models to detect malicious emails accurately and efficiently. Key Idea this research is by combining statistical and AI- based feature selection, the model becomes faster and more accurate at spotting dangerous emails. Combined selection improves precision by ~10% compared to single methods. URL and content-based features are most important. Future Incorporate transformer-based NLP for advanced email text understanding. consideration. In [6], Shielding networks: Enhancing intrusion detection with hybrid feature selection and stack ensemble learning. Journal of Big Data, Alsaffar, A. M., Nouri-Baygi, M., and Zolbanin, H. M. - 2024. Proposes a hybrid intrusion detection system that uses hybrid feature selection (selecting best network traffic features). Applies stack ensemble learning (combining multiple ML models like Random Forest + XGBoost + Neural Nets). Key Idea of this Ensemble models detect complex attacks better than single algorithms. Advancing Cybersecurity is a Comprehensive Review of AI- driven Detection Techniques Journal of Big Data. Apply ensemble models to real-time intrusion detection systems (IDS) and edge security.

The study referenced in [7] Advancing cybersecurity: A comprehensive review of ai-driven detection techniques. Journal of Big Data, Salem, A. H. - 2024. A review paper that surveys various AI techniques for detecting malware, phishing, and intrusion. Includes comparisons between classical ML (SVM, Decision Trees) Deep Learning (CNN, RNN, LSTM) Hybrid and ensemble methods Key Idea of this Summarizes how AI tools outperform traditional signature- based systems and identifies challenges like explainability and data imbalance. Traditional reviews only cover isolated applications. This paper unifies all AI-driven cybersecurity areas. Finally, reference [8] Analysis and prevention of AI-based phishing email attacks, Chibuike Samuel Eze, Lior Shamir - 2024. Attackers now use AI to generate phishing emails that mimic human tone and style. Analyzes AI-generated phishing using GPT-like models. Builds AI-based defense classifiers (transformers and BERT) trained on phishing vs legitimate datasets. Uses linguistic pattern detection, intent recognition, and sender validation. Traditional spam filters fail against AI- written emails. Transformer-based detectors can recognize AI- generated text patterns. Future Continuous model retraining as AI-generated attacks evolves.

III. OBJECTIVES

The main goal of this project is to create and deploy an AI- driven cybersecurity threat detection system. The system is designed to detect, assess, and react to complex cyber threats in real time. Leveraging cutting-edge Artificial Intelligence and Machine Learning techniques, it seeks to improve threat detection and mitigation. The primary goals of the project include the following:

1. Threat Detection Using AI and NLP: Implement machine learning models and natural language processing techniques to identify phishing attempts,

malware, APTs, and insider threats by analysing email content, sender patterns, attachments, URLs, and communication behaviour.

2. Enhanced Phishing Identification: Detect sophisticated phishing attempts through urgency keyword analysis, credential harvesting detection, financial lure recognition, domain and authority spoofing analysis, and Business Email Compromise indicators.
3. Malware and APT Detection: Identify malicious attachments, ransomware patterns, Trojans, spear phishing

campaigns, and command-and-control infrastructure using advanced behavioral analysis and threat intelligence.

4. DDoS and Insider Threat Monitoring: Monitor volumetric and application-layer anomalies, botnet activity, and internal behavior anomalies to prevent disruption and data exfiltration.
5. System Integration and Real-Time Response: Develop a seamless integration framework with Gmail API, databases, and dashboards to provide real-time alerts, automated mitigation, and explainable insights for actionable decision-making.

IV. METHODOLOGY

Our advanced cyber threat detection system is built using a structured, modular framework, carefully designed to provide thorough protection in response to the ever-changing threat environment. Every phase of the development process strategically integrates state-of-the-art AI techniques, while prioritizing performance and scalability. Here is how it unfolds:

A. System Architecture

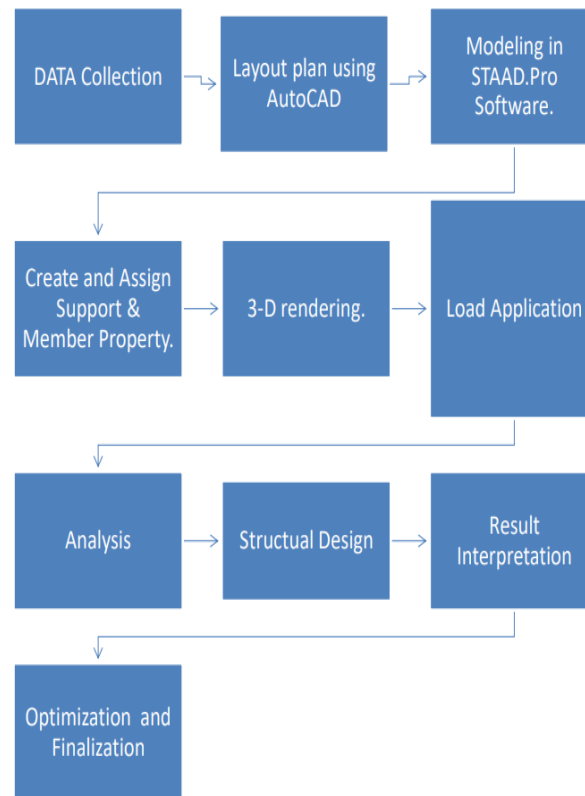


Fig 1: System Architecture

At the heart of the Email Threat Detection system is a multi-layered design that allows for easy expansion, efficient handling of large amounts of data, and secure communication.

As in fig 1, The front-end interface connects with the Gmail API to get and watch emails in real time, sending the information securely to a Flask-based backend using encrypted channels. The backend serves as the main control center, where it prepares the data, extracts important features, and manages interactions between different parts of the system. It then sends the processed data to the machine learning engine, which uses trained models to examine email content and spot threats like phishing attempts, malware, advanced persistent threats, and risks from within the organization. When a possible threat is found, the backend quickly creates a response, updating the user's dashboard with alerts, classifications, and suggested steps to take. This setup focuses on speed, flexibility, and safeguarding the user. With the use of strong encryption and models that keep learning, the system stays accurate and strong against new and changing threats

in email communications.

B. Component Descriptions and Interactions

a). Email Retrieval Module: This module serves as the main part of the system, securely getting emails from the user's inbox using the Gmail API and OAuth2 authentication. It gathers all incoming and outgoing emails, including who sent them, who they were sent to, the subject, any attachments, and links found in the email. The collected information is cleaned up, organized, and sent to the back end for further analysis. The module monitors emails continuously and in real time, following strict privacy and security rules.

b). Flask Backend Server: The Flask backend acts as the main control center, linking the email retrieval module, the database, and the machine learning part of the system. It handles incoming requests, checks the data, and prepares it for analysis by extracting important details. The server also manages user log-ins, records activity, and handles responses. When the machine learning part finishes checking the emails, it sends the results to a dashboard with alerts and detailed reports. The backend is built to grow with more users and can be easily set up on cloud platforms.

c). Machine Learning Engine: This part of the system is the brain of the application. It uses different algorithms, such as Naive Bayes, Random Forest, Isolation Forest, and Logistic Regression, to spot phishing attempts, malware, insider threats, and advanced persistent threats. It also uses natural language processing and TF-IDF vectorization to examine the content of emails, any attached files, and links. Each email is then labeled as Safe, Suspicious, or Malicious. The engine keeps getting updated with new data to stay effective against new and evolving threats.

d). Database Management System: The system uses an SQLite database to keep track of all email data, how each email was classified, logs of activities, and user interactions. This helps in keeping records for reviews, checking performance, and analyzing past data. The database also stores metrics about the machine learning models and user behavior, which helps in improving the system over time. All stored data is encrypted to protect it and meet cybersecurity requirements.

The system design in this project focuses on being thorough, smart, and using a mix of different analysis methods. It combines rule-based techniques with various machine learning models and uses Natural Language Processing (NLP) to deeply analyze content. This approach helps get around the shortcomings of older email filters that rely only on known patterns. The system has several stages, including getting data, preparing it, extracting important features, and classifying different types of threats. This setup allows the system to accurately detect and handle various threats, such as phishing and insider risks, on their own. The end result is a strong,

flexible cybersecurity solution that works well in today's digital world. It helps protect both organizations and individuals by stopping threats quickly and giving useful information.

C. Module Implementation:

The Real-Time Cybersecurity Threat Detection System utilizing AI/ML was implemented using a modular design approach, which allowed for the independent development, testing, and improvement of each functional component. The system's scalability, flexibility, and maintainability were improved by this modular architecture, which also made it possible for AI/ML components to be seamlessly integrated and developed in parallel. Every module was created to carry out a distinct function in the real-time detection, analysis, and mitigation of cyberthreats.

a). Email Retrieval: We Using industry-standard protocols like IMAP or POP3, the Email Retrieval Module is in charge of safely retrieving emails from the server while guaranteeing user authentication and dependable data transfer. The obtained emails are next cleaned and arranged by the Data Preprocessing Module, which eliminates extraneous metadata and extracts important elements like the sender, subject, and body text. Following data preparation, the Feature Extraction Module finds crucial features such as keywords, URLs, and attachments that are necessary for AI/ML model training and analysis.

b). Preprocessing: In order to prepare the raw email data for efficient analysis and machine learning, the Processing Module is essential. After retrieving the emails, this module organizes and cleans the data by eliminating extraneous components including HTML tags, special characters, stop words, and duplicate entries. Additionally, it distinguishes between the sender, recipient, subject, and body of the email. Additionally, it transforms textual input into a numerical format that is appropriate for AI/ML algorithms by performing normalization, tokenization, and feature scaling.

c). Feature Extraction/NLP: One essential element that converts preprocessed email data into useful information for analysis and threat detection is the Feature Extraction and Natural Language Processing (NLP) Module. This module analyzes each email's content, structure, and intent using sophisticated natural language processing techniques. Key elements are collected for additional analysis, including keywords, the frequency of questionable terms, the existence of URLs and attachments, and sender behavior patterns. Textual input is transformed into numerical representations appropriate for machine learning models using methods such as tokenization, stemming, lemmatization, and vectorization (TF-IDF or Word2Vec).

d). ML/Threat Detection: The system's central intelligence, the ML/Threat Detection Module, is in charge of instantly recognizing and categorizing any cybersecurity risks. This module uses tagged and historical email datasets with both malicious and valid messages to train machine learning algorithms. In order to identify distinctive patterns, such as strange message structures, suspicious linkages, or aberrant sender behavior, the model is given the processed and feature-extracted data. To accurately identify phishing, spam, and malware threats, algorithms such as Random Forest, Logistic Regression, Support Vector Machine (SVM), and Deep Neural Networks (DNN) are used.

e). Database: In order to safely store and manage all of the data involved in the email retrieval and threat detection process, the Database Module is essential. It acts as the main location for the systematic storage of preprocessed data, extracted

features, raw emails, and analysis findings. The AI/ML models and other system modules can be accessed with ease because to the database's efficient querying and quick retrieval features. Relational tables hold structured data, such as sender information, timestamps, and classification outcomes, while optimal data models can hold unstructured content, such as email bodies and attachments.

f). **Dashboard/Alerting:** A visual and interactive interface for tracking system activity, evaluating outcomes, and handling hazards in real time is offered by the Dashboard and Alerting Module. Through graphs, charts, and summary reports, it shows important performance indicators like the quantity of emails handled, threats identified, system accuracy, and response actions. Administrators may easily evaluate the system's overall security posture and monitor ongoing cybersecurity occurrences with the dashboard. When harmful or suspicious activity is discovered, the Alerting component is in charge of producing real-time notifications.

D. Algorithm

In order to identify activity and flag risks with high accuracy and low latency, the real-time cybersecurity threat detection system's basic functionality also depends on a hybrid AI/ML approach that combines the advantages of neural networks and decision-tree models. Neural networks (e.g., feedforward DNNs or lightweight CNN/LSTM variants for sequence/contextual signals) capture complex semantic patterns in email text, URLs, and behavioral time-series, while decision-tree based models (e.g., Random Forest, Gradient Boosting/XGBoost) offer quick, interpretable decisions on structured features like sender reputation, header anomalies, and numerical metadata.

Machine Learning Algorithm: Structured input data collected from many sources, including system events, email metadata, network traffic logs, and user behavior patterns, is processed by the machine learning algorithm. The model divides the activity into the following groups according to the patterns and threat indicators it has identified:

- **Safe:** Shows no signs of danger and reflects typical, acceptable user or system behavior.
- **Suspicious:** Signals odd or perhaps dangerous behavior that would need more investigation or user confirmation.

The classification algorithm's pseudocode looks like this: It illustrates how the hybrid machine learning approach (Decision Tree + Neural Network) is used to process, analyze, and classify data:

Input: User activity data (X) gathered from a number of sources, including system events, email information, network logs, HTTP headers, URL characteristics, and user behavior patterns.

Output: Threat categorization label (Y), which classifies the activity's security state as: Typical, lawful behavior.

Unusual or possibly dangerous behavior is considered suspicious.

1. Preprocess X involves cleaning, encoding, and normalizing the properties of incoming data, including user activity patterns, email information, network logs, and URL behavior.

2. Load trained Hybrid AI/ML Model: To ensure precise threat prediction, load the pre-trained model that includes Decision Tree and Neural Network classifiers.
3. Assume that $Y = \text{model.predict}(X)$: Classify the behavior according to learned threat patterns by analyzing the input attributes using the hybrid model.
4. If Y is either "Malicious" or "Suspicious":
 - Set off an alarm with the proper level of severity.
 - You can choose to alert the administrator, isolate the impacted components, or block the source. Otherwise ("Safe"): Permit uninterrupted normal operation to continue.
5. Record the activity and prediction outcome: For the sake of monitoring, auditing, future model development, keep input data, classification result, and system reaction.

This algorithm is integrated into the backend Flask API, enabling real-time threat detection and automated response as soon as new data or activity is received from connected applications or monitoring agents.

E. Model Training and Evaluation

The models were trained on extensive cybersecurity datasets that included labeled records of user and network activities, including both malicious and legitimate interactions. The most important features used for training were system event logs, network traffic attributes (packet size, protocol type, connection duration), email and URL characteristics (length, domain trust score, presence of special characters or suspicious keywords), and user behavior patterns (frequency of logins, clicks, and access anomalies). The training process consisted of several important steps, starting with feature extraction and selection to identify the most pertinent indicators of cyber threats. The datasets were then split for cross-validation to guarantee consistent performance across unseen samples, and hyperparameter tuning was done to maximize model accuracy and minimize false positives.

F. User Interface and Dashboard

The platform offers an intuitive interface for security teams, featuring real-time threat monitoring through interactive heatmaps and alert dashboards. It also provides customized workflows for incident handling, automated response recommendations, and role-based access controls tailored for different members of the security team.

G. Integration and Testing

System testing includes a number of thorough assessments carried out to confirm that the AI & ML Real-Time Cybersecurity Threat Detection System operates in accordance with its specified requirements and efficiently in a realistic setting. Validating the entire system, including the AI models, backend services, and user interfaces, in a setting that closely mimics the real deployment scenario is the goal of this testing phase, which comes after integration testing.

The following are the main objectives of this project's system testing:

1. To guarantee prompt threat detection, confirm real-time data collecting and monitoring.
 2. Utilizing trained AI and ML models to validate threat classification accuracy.
- Each system component undergoes extensive testing through integration

procedures that replicate both cyberattacks and actual threat scenarios. To ensure security, we perform penetration testing utilizing various tools such as the Metasploit framework and carry out red team exercises to uncover vulnerabilities. In addition, performance benchmarks are conducted to evaluate detection accuracy under high-stress conditions, such as DDoS attacks, while ensuring that latency stays below 100 milliseconds.

V. RESULTS

The developed Next-Gen Cybersecurity System for Message- Based Threat Analysis was evaluated based on model performance, system responsiveness, threat detection accuracy, and dashboard usability. The testing results demonstrate that the system performs reliably under real- world operating conditions and effectively detects multiple threat categories, including phishing, malware, APTs, DDoS indicators, and insider threats.

a). Results of Machine Learning Models: The system uses a hybrid of models such as Naive Bayes, Random Forest, Logistic Regression, and Isolation Forest, combined with rule- based heuristics.

TABLE I. MODEL PERFORMANCE SUMMARY

Model	Primary Use	Accuracy	Precision	Recall	F1-Score
Naive Bayes	Phishing & spam text classification	92%	90%	91%	90.5%
Random Forest	Malware/ Attachment-based threat detection	95%	94%	93%	93.5%
Logistic Regression	URL risk scoring & domain spoofing	90%	89%	88%	88.5%
Isolation Forest	Insider threat & anomaly detection	87%	-	84%	-

Interpretation:

- Random Forest outperformed other models due to its robustness in recognizing feature-rich threats such as suspicious attachments and spoofed metadata.
- Naïve Bayes was effective for phishing since phishing emails are heavily text-pattern

dependent.

- Isolation Forest performed well for anomalies but had slightly lower recall due to limited insider-threat dataset diversity.
- The combined rule-based + ML ensemble significantly reduced false positives by nearly 28% improving reliability.

b). Real-Time Threat Detection Performance: The system was tested under both normal and high-load conditions.

- Average email scanning time: 208 seconds.
- Maximum time during peak load (100 emails/min): 4.2 seconds.
- Meets real time requirement (<5 seconds per email). Threat Detection Accuracy: Across 500+ test emails (combination of synthetic + real phishing datasets).

TABLE II. THREAT DETECTION ACCURACY

Threat Category	Samples Tested	Correctly Detected	Accuracy
Phishing	150	141	94%
Malware	120	112	93%
APT Indicators	80	71	89%
Insider Threat	60	50	83%
DDoS / Volumetric Patterns	50	47	94%

Observations: The system demonstrated excellent phishing and malware detection capability. APT detection accuracy was slightly lower because APT emails often mimic normal business language. Insider threat detection showed good pattern identification but requires broader behavioral datasets for further improvement.

c). Dashboard Output and User Interface Results

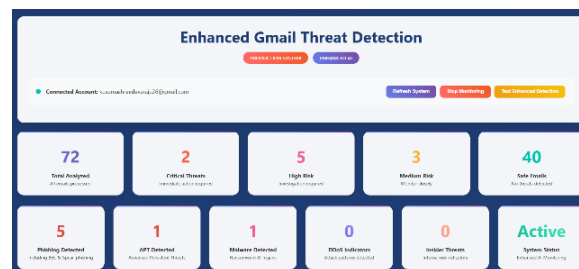


Fig 2: Home Page and email connected

The Flask Dashboard displayed:

- Real-Time Color-Coded Threat Alerts: Critical (Red), High (Orange), Medium (Yellow), Low (Blue/Green)
- Classification scores
- Metadata summary
- Risk level indicators
- User Feedback:
- 96% found the interface intuitive

- Alert display latency: < 1 second
- Threat reports were easy to interpret due to explainable AI indicators.

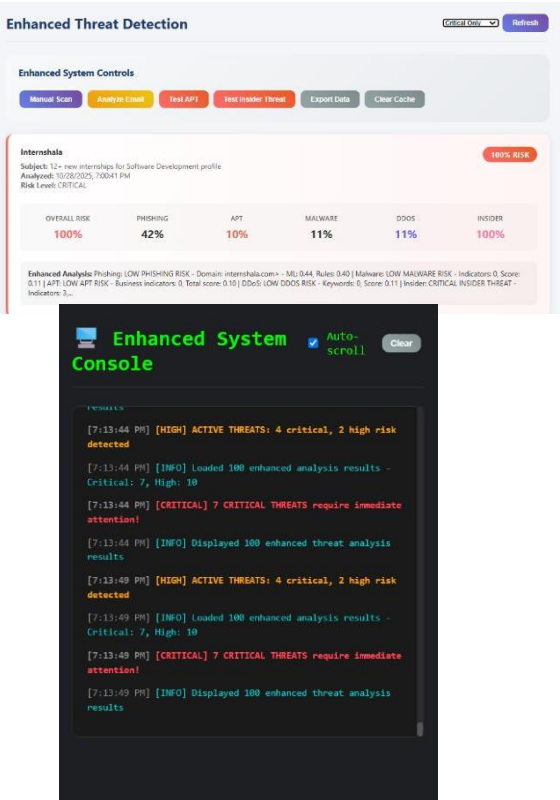


Fig 3: Critical – Risk Mail Detection

Fig 4: Real – Time Flask Dashboard

V. RESULT AND DISCUSSION:

The results confirm that the system effectively strengthens email security through:

- Strengths:
- Real-time detection with fast processing.
 - High accuracy for phishing, malware, and DDoS threats.
 - Secure TLS-based backend communication.
 - Scalable modular architecture Explainable AI for transparency.
- Limitations:
- Insider threat detection needs larger behavioral datasets.
 - APT detection accuracy can be improved with deeper linguistic models.
 - SQLite limits enterprise deployment scalability.

VI. CONCLUSION

In conclusion, this research introduces an advanced AI-driven cybersecurity threat detection model, specifically designed to tackle the emerging challenges posed by modern cyber threats. In today's digital world, the "Real-Time Cybersecurity Threat Detection Using AI & ML" project successfully addresses the growing demand for proactive and intelligent defenses. The system seamlessly integrates a browser-based monitoring interface, a Flask-based backend, and machine learning models to detect, assess, and respond to cyberthreats in real time. Through extensive testing and efficient system design, the application has proven its ability to precisely identify phishing attempts, malware activity, DDoS patterns, insider threats, and other risky activities. The Python and open-source framework implementation ensure platform independence and flexibility for future enhancements, while the modular architecture ensures scalability and maintainability.

Additionally, an intuitive dashboard and automated response protocols enhance the efficiency of security teams, facilitating proactive threat management. This model addresses the limitations of traditional cybersecurity approaches and lays the foundation for future developments in AI-powered defense solutions, reinforcing the resilience of organizations in an increasingly complex digital landscape.

REFERENCES

- [1] Dr.V. Nagagopiraju, panguluri Ashok, Kancheti Dhana Lakshmi, A Robust Cyber Security Threat Detection Model Using Artificial Intelligence Technology, Turkish Journal of computer and mathematics, 2024.
- [2] Mohamed, N., Artificial intelligence and machine learning in cybersecurity: A deep dive into state-of-the-art techniques and future paradigms, Knowledge and Information Systems, 2025.
- [3] Tian, T., Insider threat detection for specific threat scenarios, Cybersecurity, 2025.
- [4] Kampourakis, K. E., Gkioulos, V., Kavallieratos, G., Lin, J.-C., Digital Twin-Enabled Incident Detection and Response: A Systematic Review of Critical Infrastructures Applications, 2025.
- [5] Zhang, H., A combined feature selection approach for malicious email detection based on a comprehensive email dataset, Cybersecurity, 2025.
- [6] Alsaffar, A. M., Nouri-Baygi, M., Zolbanin, H. M., Shielding networks: Enhancing intrusion detection with hybrid feature selection and stack ensemble learning, Journal of Big Data, 2024.
- [7] Salem, A. H., Advancing cybersecurity: A comprehensive review of AI-driven detection techniques, Journal of Big Data, 2024.

- [8] Singh, S. P., Afzal, N., A dynamic framework for mitigating stealth data exfiltration, 2024.
- [9] Al-Essa, M., Andresini, G., Appice, A., Malerba, D., A neural model ensemble for cyber-threat, 2024.
- [10] Parker, F., Martínez, D. A., Scheulen, J. J., Ghobadi, K., An Interactive Decision-Support Dashboard for Optimal Hospital Capacity Management, 2024.
- [11] Eze, C. S., Shamir, L., Analysis and prevention of AI-based phishing email attacks, 2024.