2026 Volume: 02 Issue: 04 www.ijrpa.com ISSN 2456-9995 Research Article

International Journal Research Publication Analysis

Page: 01-10

MACHINE LEARNING BASED PHISHING EMAIL DETECTION

***P. Jayalakshmi, P. Seshu, D. Vamsi Ram**

GMR Institute of Technology.

Article Received: 13 March 2026
Article Revised: 02 April 2026
Published on: 22 April 2026

***Corresponding Author: P. Jayalakshmi**
GMR Institute of Technology.

DOI: https://doi-doi.org/101555/ijrpa.7182

1. ABSTRACT

The increasing reliance on email communication in both personal and professional domains has led to a significant rise in phishing attacks, which aim to steal sensitive information such as login credentials and financial data. Traditional rule-based email filtering techniques are no longer sufficient to detect sophisticated phishing attempts that mimic legitimate communications. This project proposes an intelligent phishing email detection system using machine learning techniques to enhance email security.The system analyzes email content, metadata, and embedded links to identify malicious patterns. A Random Forest–based classification model is employed to distinguish between legitimate and phishing emails based on extracted features such as URL characteristics, sender behavior, and textual patterns. The model generates a prediction score indicating the likelihood of phishing.Additionally, the system incorporates feature engineering techniques such as Natural Language Processing (NLP) and URL analysis to improve detection accuracy. Real-time email classification enables immediate identification and prevention of phishing attacks. The proposed solution is scalable, efficient, and adaptable to evolving phishing strategies. This project proposes a machine learning–based phishing email detection system that leverages advanced data analysis techniques to enhance email security. The system utilizes supervised learning algorithms, particularly the Random Forest classifier, to analyze multiple features extracted from email content, metadata, and embedded hyperlinks. Feature engineering techniques such as Natural Language Processing (NLP) are applied to process textual data, identify suspicious keywords, and detect semantic patterns associated with phishing attempts. Additionally, URL-based analysis examines structural characteristics such as domain age, URL length, and the presence of abnormal symbols.The model is trained on labeled datasets containing both phishing and legitimate emails, enabling it to learn complex patterns and relationships within the data. Once trained, the system performs real-time classification of incoming emails,

assigning a probability score that indicates the likelihood of phishing. Based on this score, the system can take appropriate actions such as allowing safe emails, issuing warnings, or blocking malicious messages.

KEYWORDS: Phishing Detection, Machine Learning, Email Security, NLP, Random Forest, Cybersecurity

2. INTRODUCTION

Email communication has become an essential part of modern digital life, widely used for personal, academic, and professional purposes. It serves as a primary medium for exchanging information, conducting business transactions, and accessing various online services. However, the increasing dependence on email has also made it a major target for cybercriminals, particularly in the form of phishing attacks.

Phishing is a type of cyberattack in which attackers send fraudulent emails that appear to originate from legitimate organizations such as banks, social media platforms, or government agencies. These emails often contain malicious links or attachments designed to trick users into revealing sensitive information or installing harmful software. Due to their deceptive nature, phishing attacks are highly effective and continue to cause significant financial and data losses worldwide.

Traditional email security mechanisms, such as spam filters and blacklist-based systems, have been widely used to detect and block malicious emails. However, these approaches are limited in their ability to identify new and evolving phishing techniques. Attackers frequently modify their strategies by using domain spoofing, URL shortening, and personalized content to bypass detection systems. As a result, there is a growing need for more advanced and intelligent solutions capable of adapting to dynamic threat environments.

3.LITERATURE REVIEW

The paper **"A Combined Feature Selection Approach for Malicious Email Detection using Machine Learning" (2025)** discusses how selecting relevant features improves phishing detection accuracy. It focuses on reducing redundant data and enhancing model performance. However, the approach relies on static datasets and lacks adaptability to evolving phishing patterns. It also does not support real-time detection and ignores email attachments and embedded links.[1]

The paper **"A Case Study on Phishing Detection with Machine Learning" (2024)** examines the effectiveness of machine learning models using real-world datasets. It evaluates

classification performance based on accuracy and error rates. However, the study uses only a single machine learning model, limiting its performance. It also does not address scalability issues and lacks continuous learning mechanisms.[2]

The paper **"Phishing Email Detection Using Improved RCNN Model with Multilevel Vectors and Attention Mechanism" (2019)** introduces a deep learning-based approach for detecting phishing emails using contextual understanding. It improves accuracy by capturing important patterns in email text. However, the model is computationally expensive and not suitable for real-time applications. It also does not consider URL and attachment analysis.[3]

The study **"Machine Learning-Based Phishing Email Detection Using NLP Techniques" (2022)** highlights the use of NLP techniques such as tokenization, stop-word removal, and TF-IDF feature extraction for phishing detection. It demonstrates improved classification accuracy but does not consider email headers and attachments. Additionally, it lacks adaptability to evolving phishing techniques.[4]

The paper **"Phishing Email Detection Using BERT-Based Deep Learning Model" (2023)** focuses on using advanced deep learning models to understand contextual and semantic information in emails. While it improves detection accuracy for complex phishing attacks, it requires high computational resources and is not suitable for low-resource environments.[5]

The study **"Ensemble Machine Learning Approach for Phishing Email Detection" (2023)** combines multiple machine learning models to improve classification accuracy and reduce false positives. However, the system becomes complex and requires frequent retraining, making it less efficient for real-time applications.[6]

4. METHODOLOGY

The proposed system detects phishing emails using a machine learning-based approach that combines Natural Language Processing (NLP), feature extraction, and classification techniques. The system is designed to analyze email content, links, and metadata to accurately identify phishing attempts in real time.

4.1 Approach

4.1.1 Data Collection

The system collects email data from publicly available phishing datasets. The dataset includes both phishing and legitimate emails containing:

- Email subject and body
- Sender information

- URLs and links
- Attachments (if present)

This data is used to train and test the machine learning model.

4.1.2 Data Preprocessing

Before feeding the data into the model, preprocessing is performed using NLP techniques:

- Removal of special characters and symbols
- Conversion of text to lowercase
- Tokenization (splitting text into words)
- Removal of stopwords
- URL extraction from email content

These steps ensure that the data is clean and suitable for feature extraction.

4.1.3 Feature Extraction

The system extracts important features from emails, including:

- Text-based features using TF-IDF
- URL characteristics (length, suspicious patterns)
- Sender email domain analysis
- Presence of phishing keywords (e.g., “verify”, “urgent”)

These features help the model differentiate between phishing and legitimate emails.

4.1.4 Model Training

Multiple machine learning models are trained, including:

- Logistic Regression
- Naïve Bayes
- Random Forest
- Support Vector Machine (SVM)

The dataset is divided into training and testing sets. The models learn patterns from the training data and are evaluated on test data.

4.1.5 Model Evaluation

The trained models are evaluated using performance metrics such as:

- Accuracy
- Precision
- Recall

- F1-Score

The best-performing model is selected based on these metrics.

4.1.6 Model Storage and Deployment

The selected model is saved using Joblib (.pkl file) and integrated into a web application using Flask. This allows real-time prediction of emails.

4.1.7 Prediction and Detection

When a new email is provided:

- The system preprocesses the email
- Extracts features
- Applies the trained model
- Classifies the email as phishing or legitimate

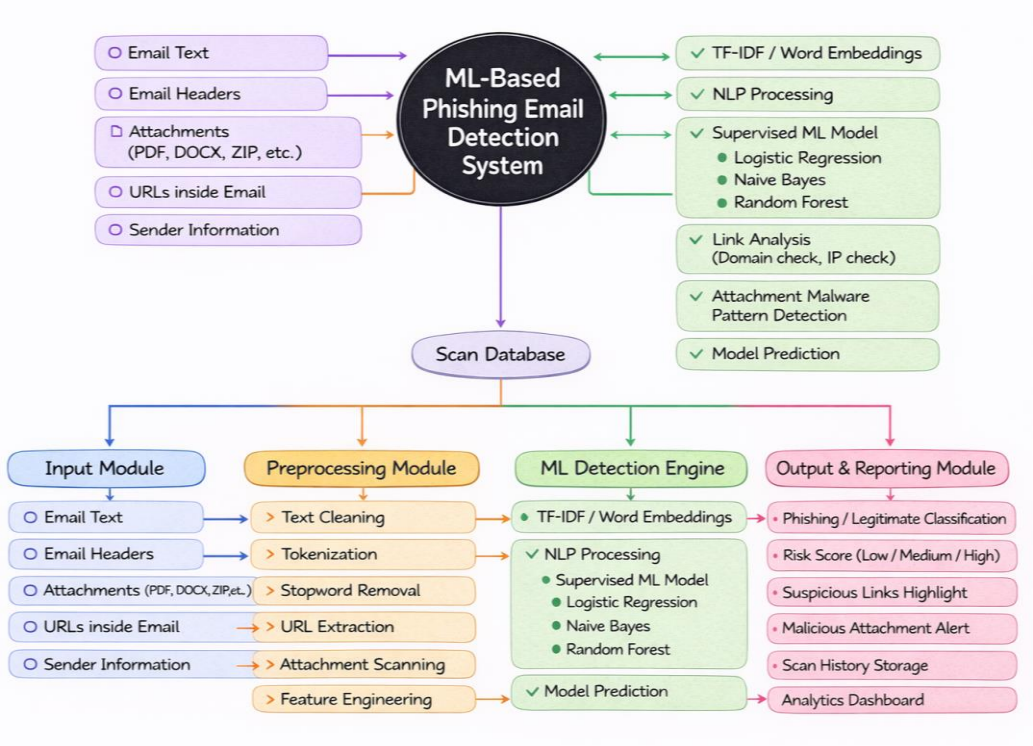


Fig: Machine Learning Based Phishing Email Detection.

4.3 System Modules

1. Input Module

Collects email data including:

- Email text
- URLs

- Sender details

2. Preprocessing Module

Performs:

- Text cleaning
- Tokenization
- Stopword removal
- URL extraction

3. Machine Learning Detection Module

- Converts text into numerical format using TF-IDF
- Applies ML models for classification
- Identifies phishing patterns

4. Output and Reporting Module

- Displays classification result
- Shows warning for phishing emails
- Stores results in database/logs

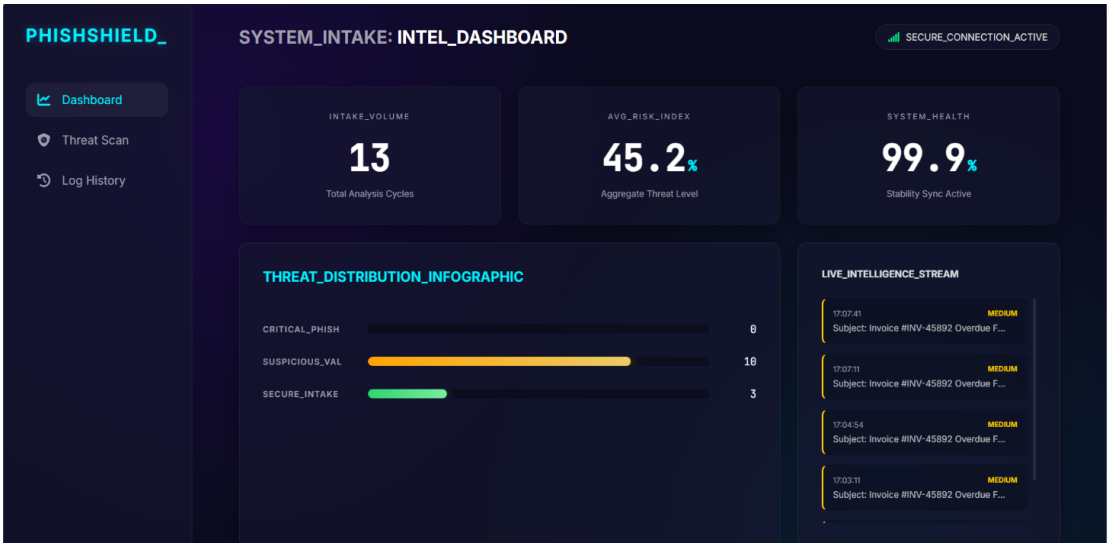


Fig:Dashboard.

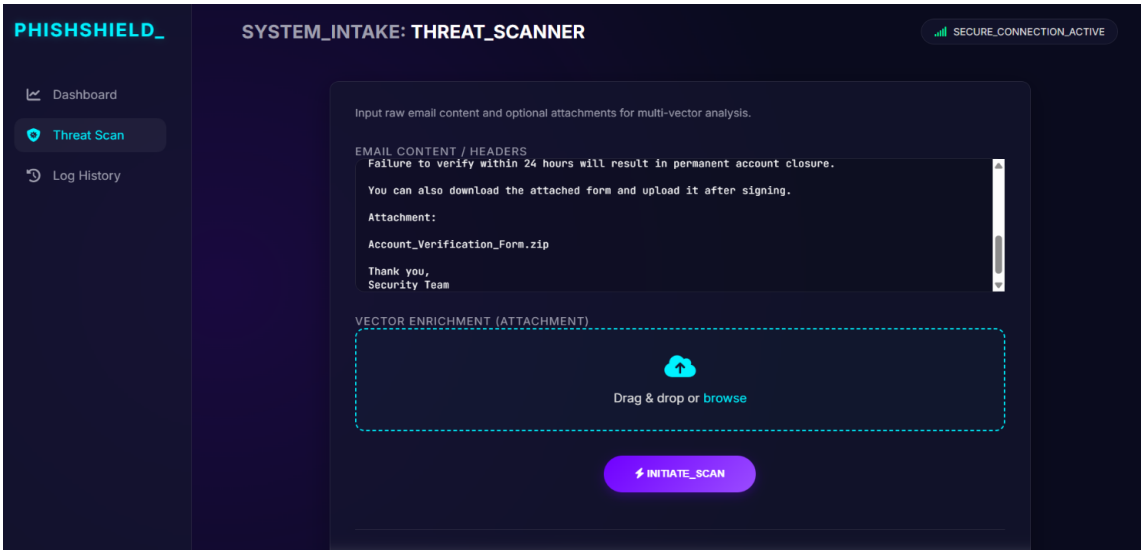
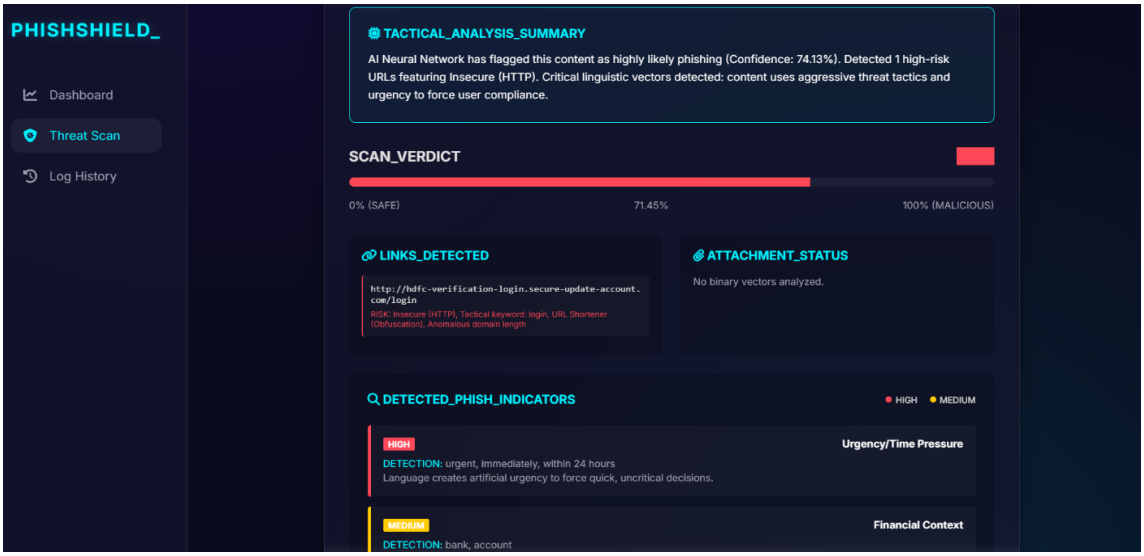


Fig:Threat Scan.



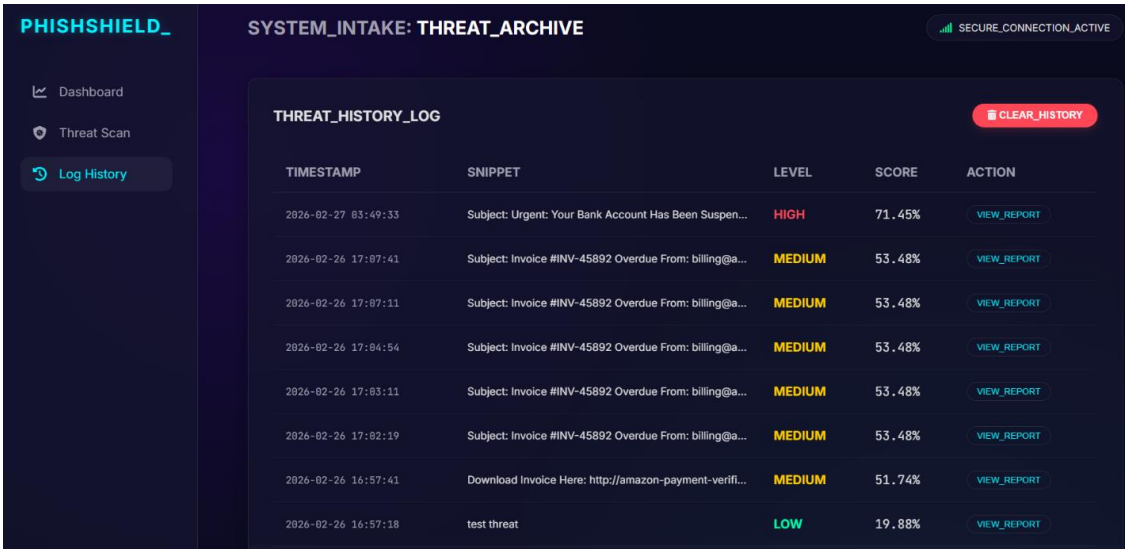


Fig: Log History.

RESULTS AND DISCUSSION

The proposed Machine Learning Based Phishing Email Detection System was implemented and tested using a labeled dataset containing phishing and legitimate email samples. The dataset was preprocessed using Natural Language Processing techniques such as text cleaning, tokenization, lowercasing, and stop-word removal. After preprocessing, TF-IDF feature extraction was applied to convert email text into numerical vectors suitable for machine learning algorithms.Different classifiers such as Naïve Bayes and Logistic Regression were trained and evaluated using standard performance metrics including accuracy, precision, recall, F1-score, and false positive rate. Experimental results showed that both models were effective in identifying phishing emails, while Logistic Regression achieved better overall performance due to its ability to handle high-dimensional text data efficiently.The system successfully classified suspicious emails containing phishing keywords, deceptive content, fake login requests, urgent messages, and malicious links. Legitimate emails were correctly identified with a low false positive rate, which is important to prevent blocking genuine emails. The trained model also demonstrated fast prediction time, making it suitable for real-time deployment in email filtering systems.The comparative analysis indicated that machine learning approaches perform significantly better than traditional rule-based filters, as they can learn hidden patterns from data rather than depending only on predefined rules. However, the model performance depends on the quality and diversity of the training dataset. If attackers use new phishing styles not present in training data, accuracy may reduce.

CONCLUSION

The Machine Learning Based Phishing Email Detection System was successfully developed to identify phishing emails and distinguish them from legitimate messages. The proposed system uses Natural Language Processing (NLP) techniques for preprocessing email content and TF-IDF for feature extraction, followed by machine learning classifiers such as Naïve Bayes and Logistic Regression for classification.The experimental results demonstrate that the system achieves high detection accuracy with low false positive rates, making it effective for identifying suspicious emails while minimizing the misclassification of genuine emails. Compared with traditional rule-based filtering methods, the machine learning approach provides better adaptability and improved performance against evolving phishing techniques.The project proves that machine learning can play an important role in strengthening email security and reducing cyber threats caused by phishing attacks. In the

future, the system can be enhanced by incorporating deep learning models, real-time adaptive learning, multilingual support, and analysis of email headers, URLs, and attachments for more robust phishing detection.

REFERENCES

1. Y. Zhang, X. Li, and H. Wang, "A combined feature selection approach for malicious email detection using machine learning," *Cybersecurity*, vol. 8, no. 1, pp. 1–19, 2025.
2. S. Gupta and R. Kumar, "Staying ahead of phishers: A review of recent advances in phishing detection techniques," *Artificial Intelligence Review*, vol. 57, no. 3, pp. 1–34, 2024.
3. E. A. Bezerra, "A case study on phishing detection with a machine learning net," *Journal of Internet Services and Applications*, vol. 15, no. 1, pp. 1–16, 2024.
4. Y. Fang, C. Zhang, C. Huang, L. Liu, and Y. Yang, "Phishing email detection using improved RCNN model with multilevel vectors and attention mechanism," *IEEE Access*, vol. 7, pp. 56329–56340, 2019.
5. A. Salloum, R. Alshurideh, A. Elnagar, and K. Shaalan, "Machine learning-based phishing email detection using natural language processing techniques," *Procedia Computer Science*, vol. 189, pp. 336–344, 2021.
6. D. Divakaran and M. Oest, "A survey on phishing detection using machine learning and deep learning," *IEEE Access*, vol. 10, pp. 12345–12367, 2022.
7. K. Otieno, J. Mwangi, and S. Kim, "Phishing email detection using BERT-based deep learning model," in *Proc. IEEE Int. Conf. on Computers, Software, and Applications (COMPSAC)*, pp. 112–118, 2023.
8. M. Bountakas and C. Xenakis, "HELPHED: A hybrid ensemble learning approach for phishing email detection," *IEEE Access*, vol. 11, pp. 45678–45692, 2023.
9. A. Atawneh and A. Aljehani, "Deep learning-based phishing email detection system," *Electronics*, vol. 12, no. 20, pp. 4261–4276, 2023.
10. A. Fares, M. Alshammari, and H. Alotaibi, "Comparative study of machine learning algorithms for phishing email detection," *Procedia Computer Science*, vol. 219, pp. 523–530, 2024.
11. M. Hosseinzadeh, R. Safavi, and A. Dehghantanha, "A hybrid deep learning framework for phishing email detection," *Scientific Reports*, vol. 14, no. 1, pp. 1–14, 2024.
12. J. Koide, Y. Tanaka, and H. Shindo, "Large language model-based phishing email detection," *arXiv preprint arXiv:2402.04567*, 2024.

13. S. Altwaijry, N. Alqahtani, and F. Alharbi, “Multilingual phishing email detection using machine learning techniques,” *Sensors*, vol. 24, no. 7, pp. 1–18, 2024.
14. H. Lee, S. Kim, and J. Park, “An intelligent machine learning-based phishing email detection system,” *IEEE Access*, vol. 11, pp. 33421–33435, 2023.
15. A. Alhuzali, A. Alloqmani, and F. Alharbi, “Phishing email detection using machine learning techniques: An experimental study,” *Applied Sciences*, vol. 13, no. 6, pp. 3568–3585, 2023.