International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 11 | Nov - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Phishing E-mail Detector

Authors: **Prof. Vishal Kumar Singh, Vamsee Krishna Gudur, Kavya Neredabilli, Samula Shiv Shankar Reddy, Kolli Likitha**

Department of Computer Science and Engineering, Parul Institute of Engineering and Technology (PIET), Vadodara, Gujarat, India

Abstract

Phishing attacks are a major cybersecurity threat, causing financial losses and data breaches. Conventional spam filters and human detection are not effective against advanced phishing attacks, so AI-based solutions need to be employed. This project creates a Phishing Email Detector based on Python Flask and machine learning to scan the content of emails, identify suspicious patterns, and report potential phishing threats. Through the use of pre- trained data sets and real-time adaptive learning, the system improves detection rates and resists changing phishing techniques. The system under consideration utilizes context- aware detection, behavior link analysis, emotional and tone analysis, and computer vision- based image analysis using AI to detect more efficient fraudulent emails. It derives major email features like headers, body, and included links using machine learning techniques to enhance accuracy. The web application based on Flask provides real-time phishing detection information with a simple interface for advanced email security. The execution takes a systematic methodology, which consists of data input and preprocessing, feature extraction, model training and testing, and deployment. With integration of real-time learning mechanisms and multi-factor email verification, this project is designed to create an intelligent and dynamic anti-phishing system. The primary innovations like context- aware analysis, real-time link behavior observation, and adaptive AI processes contribute to this solution being an efficient and effective counter to contemporary phishing attacks.

I. INTRODUCTION

Phishing remains one of the most damaging forms of cybercrime, leading to major risks in finance, privacy, and reputation, resulting in financial losses, data breaches, and reputational damage for individuals and entities. Conventional spam filters and rule-based detection mechanisms do not detect sophisticated phishing attempts frequently, hence the need to make more powerful and smart security solutions. Our project intends to implement an AI-driven Phishing Email Detector with Python Flask and machine learning (ML) to augment email security by automatically detecting and marking suspicious emails. he proposed system will analyses email content, detect phishing keywords, examine embedded links, and utilize AI-powered image recognition to identify fraudulent logos or altered visuals. Unlike conventional methods, our solution will leverage machine learning models trained on phishing datasets to detect threats with higher accuracy. Additionally, behavioral analysis of links will help identify malicious redirects or phishing websites, further strengthening detection capabilities. One of the major contributions of this project is its real-time adaptive learning capability, through which the system can adapt with new phishing strategies. This guarantees that emerging threats are captured effectively without the need for predefined rules. Our system will also feature multi-factor email authentication and dynamic interface design, providing users with an effective and easy-to-use tool for phishing detection. The end product will be a web application that offers real-time phishing detection information, assisting users in protecting their email communications. Through the use of AI and ML, this project overcomes the shortcomings of conventional spam filters and human detection, providing a more efficient, automated, and smart solution to phishing attacks. In a time of ever-changing cyber threats, this AI-powered solution will be essential in reinforcing email security and defending users from spamming attempts.

II LITERATURE SURVEY

Background & Limitations of Early Methods

Phishing and spam remain prominent cyber threats. Early defences focused on blacklists, heuristics, and rule-based filters that matched sender domains, URL patterns, and fixed indicators [1–3]. These methods worked for known attacks but struggled with zero-day campaigns, obfuscation (URL shortening, homographs), and fast-changing attacker behaviour, creating maintenance burdens and high false negatives [4].

Shift to Classical Machine Learning

With labelled corpora, research moved to supervised ML over email content and metadata. Approaches such as Naïve Bayes, SVM, Logistic Regression, and tree ensembles were widely studied [3,6,11,14]. Feature sets included TF-IDF/n-gram, lexical URL statistics, WHOIS/domain age, and authentication signals (SPF/DKIM/DMARC). Some studies combined header, body, and link features for robust performance [5,12,15]. Despite improved accuracy, handcrafted features remained sensitive to drift and adversarial manipulation [6,14,16].

Deep Learning, NLP, and Representation Learning

Recent work leverages CNN, RNN, LSTM, and hybrid attention-based models to capture semantics beyond surface keywords [7,8,13]. These deep learning approaches improve robustness to obfuscation, while hierarchical attention and hybrid CNN-LSTM models enhance interpretability and detection performance [7,13]. Reviews highlight that deep learning provides stronger generalization compared to conventional ML, though it requires larger datasets and computational resources [8,10].

Data Challenges & Preprocessing

Data quality and preprocessing have been identified as critical challenges in phishing detection [9]. Issues such as imbalanced datasets, noisy labels, and domain distribution shifts limit generalization across organizations and contexts. Comparative studies emphasize the need for standardized benchmarks and consistent evaluation frameworks [10,12].

Behavioural Analysis, Networks & Human Factors

Beyond content, researchers have examined suspicious email behaviours, signature spoofing, and sender network characteristics [4,17]. Simulation studies and awareness training highlight the human element, showing how phishing susceptibility varies across email type, sector, and awareness level [18,19]. These works emphasize that purely technical defences must be complemented with human-centric strategies.

AI-Driven Hybrid & Adaptive Systems

Hybrid systems integrate genetic algorithms, decision trees, and swarm optimization with traditional classifiers for efficiency and adaptability [5,6,15]. More recent studies advocate for adaptive AI approaches that continuously learn from emerging attack strategies [16,20]. Such systems leverage multimodal features and online learning to handle evolving phishing tactics while providing real-time, actionable detection.

Gap & Contribution

Despite progress, challenges remain in zero-day detection, adversarial robustness, and usability for end-users. Current systems often require trade-offs between accuracy, explainability, and deployment cost. Our work proposes an AI-driven phishing detector with Python Flask that (i) fuses text, URL, and image features, (ii) applies behavioural link analysis in a controlled sandbox, and (iii) supports adaptive learning to mitigate drift. The system further integrates reporting, blacklisting, and history features for practical deployment, addressing both technical and human-centric aspects of phishing defence.

III. METHODOLOGY

The Phishing Email Detection System follows a structured AI/ML-driven process for accurate phishing identification. Data is collected from public datasets and email logs, pre- processed, and tokenized to ensure uniformity. Features such as email headers, text content, URLs, and images are analysed with NLP and behavioural link analysis. Supervised models (Random Forest, SVM, LSTM) are trained and optimized to recognize phishing patterns. The trained model is deployed in a Flask-based web app for real-time detection, with adaptive learning to handle evolving threats. Performance is validated using precision, recall, F1-score, and accuracy, while continuous monitoring ensures reliable detection.

Requirement Analysis

The system must detect phishing cues in email content, structure, and links. It incorporates NLP-based tone and urgency detection, image spoof detection, and URL behaviour analysis. The interface is browser-based, offering easy access and interaction.

System Design

The architecture is layered: a data ingestion layer for email preprocessing, a feature extraction module for NLP and AI analysis, a machine learning model for classification, and a detection/alert system for real-time feedback. The UI presents results and allows user validation. A database stores email records, user details, and known phishing patterns.

Development

Frontend uses HTML, CSS, and JavaScript; backend uses Flask for API handling. ML/DL models are trained on labelled datasets, with real-time adaptation for new attacks. Database integration logs detections and supports analysis.

Testing

Testing includes unit, integration, and performance checks, ensuring accuracy and scalability. Security testing protects against adversarial inputs, while user acceptance testing refines usability.

Deployment & Support

The system can be hosted on cloud or on-premises, with Docker for scalability and CI/CD pipelines for updates. Post-deployment, monitoring, retraining, user feedback integration, and security patches ensure continuous improvement.

IV. EXPECTED RESULTS

Phishing E- mail Detector offers a much better solution than conventional phishing detection techniques, providing a more effective, automated, and smart email security solution. Through the combination of machine learning and artificial intelligence, the system gains more accurate detection, less dependent on static rule-based filters that are usually ineffective against sophisticated phishing attacks. Real-time adaptive learning allows the system to adapt with new phishing tactics so that it is always providing consistent security.

The automated detection method avoids the manual screening of emails, conserving time and avoiding human error. Context-aware detection and AI-driven image recognition further assist in the identification of fraudulent requests and manipulated visuals, there by making it more challenging for attackers to impersonate known organizations.
Implemented as a Flask-based web application, the solution is scalable, easy to use, and can be combined with email services to enhance security. Through the use of AI and ML, this phishing detection system offers a multi-layered defense system, safeguarding users against phishing scams and enabling safer email communication.

V. FUTURE WORKS

Adoption of advanced language models

Explore transformer-based architectures such as Distil BERT to replace or complement TF- IDF, enabling the system to better capture subtle patterns in text.

Multilingual and mixed-language capability

Extend preprocessing and modelling to handle multiple languages and code-mixed data, which is increasingly common in user communication.

User-driven active learning

Incorporate a feedback loop where users can correct misclassifications, and those corrections are periodically added into retraining pipelines to improve accuracy with minimal labelling effort.

Model transparency for end-users

Add interpretability features in the interface, highlighting key tokens or phrases that

influenced the prediction, helping users trust the system's output.

Continuous monitoring and reliability

Develop dashboards to track data drift, response latency, and system errors, with automated alerts and fallback mechanisms to older stable models if needed.

Security and robustness

Strengthen authentication with two-factor login, enforce strict password policies, and add safeguards like rate limiting and CAPTCHA to resist attacks.

Privacy and regulatory compliance

Implement tools for automatic removal of personal identifiers, along with mechanisms for data retention control and deletion requests to meet standards such as GDPR and CCPA.

VI. PROJECT SCREENSHOTS

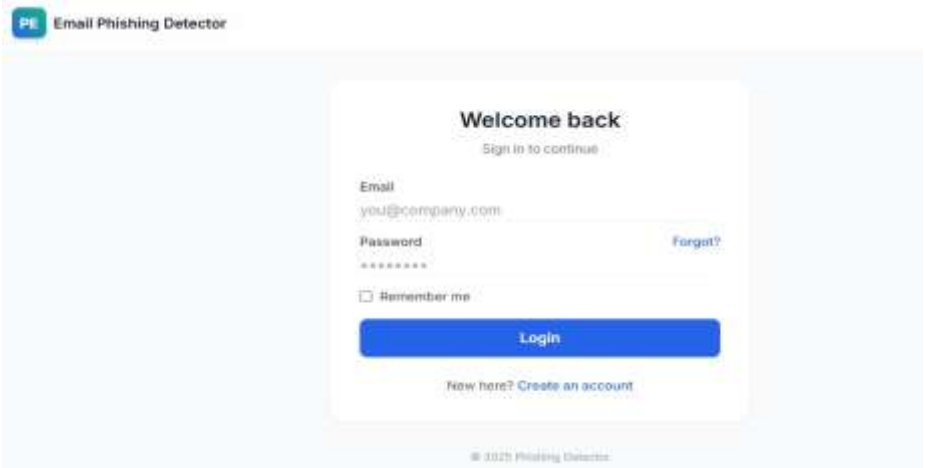

Image 1: Login Page

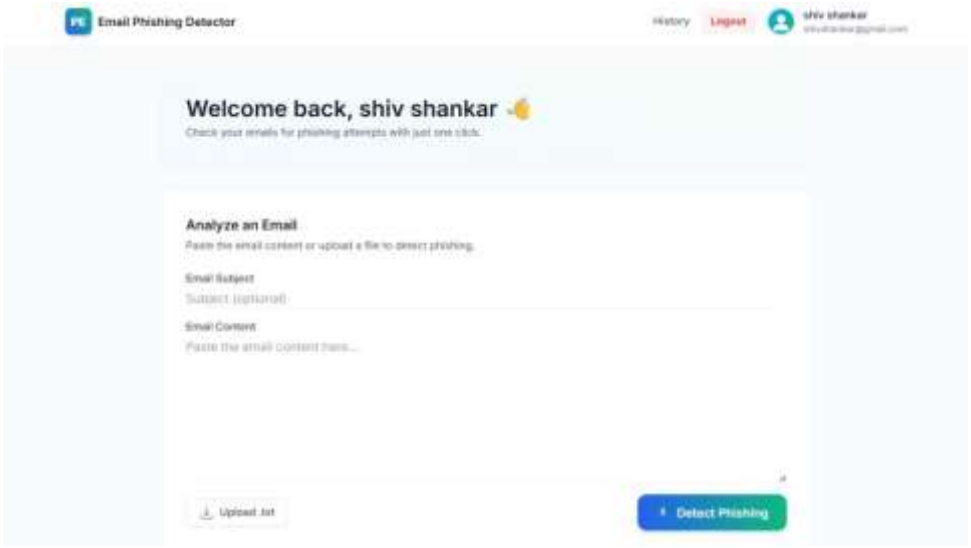

Image 2: User Home Page

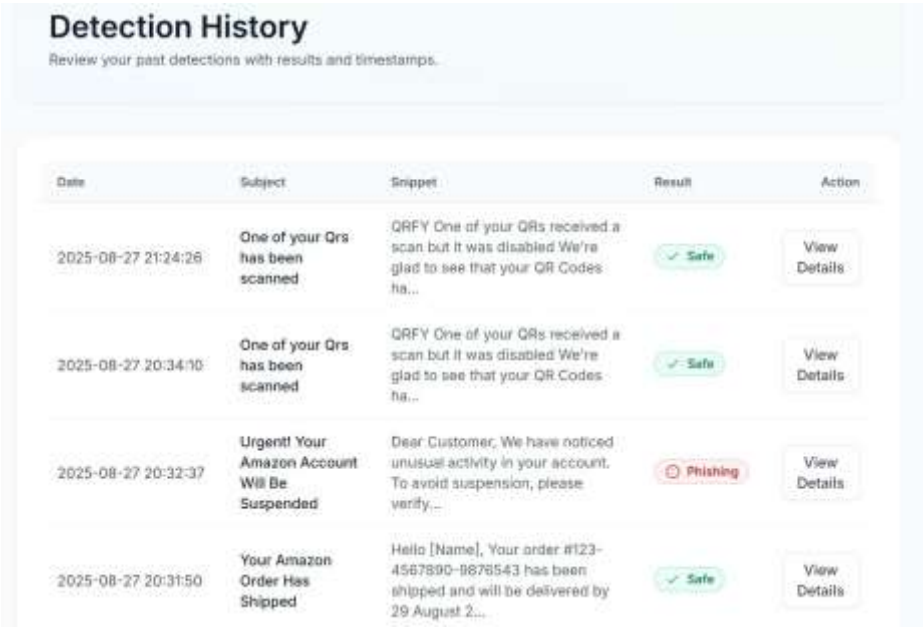

Image 3: Exercise Page

VII. CONCLUSION

This project successfully demonstrates the potential of machine learning for detecting phishing emails through an integrated Flask-based application. By combining traditional NLP methods with supervised models, the system provides a practical solution for identifying suspicious content and safeguarding users against common cyber threats. The addition of features such as history tracking and sender reporting enhances usability and encourages user trust. While the current version establishes a strong foundation, the system can be further expanded with more advanced models, multilingual support, and enhanced security mechanisms. Overall, the work highlights how lightweight AI solutions can be applied to real-world problems and sets the stage for continued development in creating safer digital communication environments.

REFERENCES

1. Improving Email Spam Detection Using Content Based Feature Engineering Approach by Wadi Hijawi and Hossam Faris (2017).

2. Spam Profile Detection in Social Networks Based on Public Features by Ala' M. Al- Zoubi and Jafar Alqatawna (2017).

3. Advanced Spam Email Detection Using Naïve Bayes Classifier: A Machine Learning Approach for Secure Digital Communication by Liansong Wang (2025).

4. Investigating Expert Users' Strategies to Detect Email Signature Spoofing Attacks by Peter Mayer and Damian Poddebniak (2022).

5. A Comprehensive Study on Efficient E-Mail Spam Detection Using Genetic Decision Tree Processing with NLP Features by Safaa S. I. Ismail and Romany F. Mansour (2022).

6. A Study on Email Spam Detection Using an Integrated Approach of Naïve Bayes and Particle Swarm Optimization by Kriti Agarwal and Tarun Kumar (2018).

7. Email Spam Detection Using Bidirectional Long Short Term Memory with Convolutional Neural Network by Sefat E Rahman (2020).

8. Review of Deep Learning Techniques for Phishing Email Detection by Jairo Gutierrez and Akbar Ghobakhlou (2024).

9. Challenges of Data Collection and Preprocessing for Phishing Email Detection by Obianuju N, Nwokonkwo Obi and Anthony I. Otuonye (2024).

10. Email Anti-Phishing: Machine Learning Models and Evaluation Overview by Charles Ikerionwu and Christiana Okoloegbo (2024).

11. Machine Learning for Email Spam Filtering: Review, Approaches and Open Research Problem by Emmanuel Gbenga Dada and Joseph Stephen Bassi (2019).

12. Efficient Email Spam Detection Using Machine Learning Techniques: A Comparative Analysis of Classification Models by Md Abdul Kadir, Sultana Akter and Shivani Rana (2024).

13. Email Spam Detection Using Hierarchical Attention Hybrid Deep Learning Method by Seyhmus Yilmaz and Sultan Zavrak (2023).

14. Email Spam Detection Using Naïve Bayes by Hrithik Vohra and Manoj Kumar (2023).

15. Efficient E-Mail Spam Detection Strategy Using Genetic Decision Tree Processing with NLP Features by Ahmed I. Taloba and Rasha M (2022).

16. A Study of Suspicious E-Mail Detection Techniques by Lokaiah Pullagura and Vinaya Kumari (2024).

17. Characterizing the Networks Sending Enterprise Phishing Emails by Elisa Luo and Grant Ho (2024).

18. Email Phishing Simulations Serve as a Valuable Tool in Fostering a Culture of Cybersecurity Awareness by Kalyankumar Dasari (2025).

19. Phishing Susceptibility Among Healthcare Workers: The Impact of Awareness, Email Type, and Location by Elizabeth N. McElhiney and Darin Challacombe (2025).

20. AI-Powered Phishing Detection: Leveraging Machine Learning for Email Security by Joshua Boluwatife Adelusi and Annie Marcus (2025).