Major Project Presentation

on

AI-Powered Anti-Phishing Email Firewall

in

Partial Fulfillment for BTech Computer Engineering (Regional Language)
Course- Major Project

MOHIT CHARJE	**121B1D050**
AKSHAY RANE	**121B1D038**
OM BABAR	**121B1D046**
KARANSINGH CHAWHAN	**121B1D050**

Under the guidance of

Prof. Rucha Shinde

Department of Computer Engineering (Regional Language)

PCET’s Pimpri Chinchwad College of Engineering

Contents

1. Introduction
2. Problem Definition
3. Motivation
4. Objective
5. Literature Survey
6. Software Requirement Specification
7. Algorithm & Dataset
8. Proposed System
9. Project Plan
10. Software Testing
11. Results
12. Contribution to Sustainable Development Goals
13. Future Scope
14. Conclusion
15. References

Introduction

This project presents an AI-powered phishing detection system that safeguards users from email-based cyberattacks using machine learning and natural language processing. Trained on a large dataset, the deep learning model monitors incoming emails in real time, analyzes their content and URLs, and classifies them as phishing or legitimate, automatically moving threats to the spam folder.

A Flask-based dashboard displays live detection statistics, phishing trends, and allows export of reports. All scan data is stored in a PostgreSQL database for fast access and analysis. Key features include real-time email monitoring, automatic spam handling, and over 90% detection accuracy with real-time dashboard updates.

Problem Definition

The increasing sophistication of phishing attacks poses a major threat to cybersecurity, requiring the development of an intelligent, automated system for detecting and mitigating phishing attempts through machine learning and natural language processing techniques. Common challenges in phishing email detection include:

1. Evolving phishing tactics and social engineering methods
2. Difficulty in analyzing short and deceptive email content
3. High false positives with traditional rule-based spam filters
4. Limited real-time protection for users

Motivation

- **Rising Cyber Threats**: Phishing accounts for over 90% of all cyberattacks, targeting individuals and organizations alike.
- **Inefficiency of Traditional Filters**: Rule-based email filters struggle to adapt to evolving phishing techniques and often produce false positives.
- **Real-Time Protection Needs**: Users require proactive and intelligent systems that can detect threats before any damage occurs.
- **Educational Impact**: Demonstrates the practical use of AI, ML, and NLP in cybersecurity, reinforcing technical learning through real-world application.
- **Affordable Security**: A Raspberry Pi-based solution offers a cost-effective and portable firewall alternative for small businesses and personal use.

Objectives

1) Develop an AI-powered firewall to detect and block phishing emails in real time.
2) Use machine learning and NLP to classify emails as phishing or legitimate based on content and metadata.
3) Build a lightweight, cost-effective system deployable on a Raspberry Pi for personal or small business use.
4) Design a user-friendly dashboard for monitoring threats, viewing logs, and managing email quarantine.
5) Enable real-time alerts and analytics using WebSockets and PostgreSQL integration.
6) Ensure offline processing of emails without relying on cloud APIs, prioritizing data privacy.

Literature Review

Sr No	Title	Year	Methodology	Strengths	Weakness
1	Ai Powered Phishing Detection And Prevention.	2024	The proposed CNN model (OTAMNet) for phishing detection leverages DenseNet architecture for its superior gradient flow and comprehensive feature extraction capabilities. Each layer in OTAMNet receives input from all preceding layers, enabling more diverse feature capture compared to standard CNNs.	High accuracy in phishing detection	Requires continuous model updates
2	Analysis and Prevention of AI Based Phishing Email Attacks	2024	Hybrid framework combining NLP and behavioral analysis to detect AI-generated phishing emails, featuring linguistic feature extraction, transformer-based classification, and adversarial testing against advanced generation models.	Efficient in detecting AI-generated phishing emails	Limited dataset size

Literature Review

Sr No.	Title	Year	Methodology	Strengths	Weakness
3	Next Generation of Phishing Attacks Using AI Powered Browsers	2024	Investigates AI-powered browser-based phishing attacks using a multi-stage experimental approach, combining traffic analysis, browser vulnerability assessment, and simulation testing to identify exploitation patterns, with effectiveness evaluated through controlled experiments and user interaction metrics.	Identifies vulnerabilities in AI-driven browsers	Focuses only on browser-based threats
4	Design of Security System Based on Raspberry-Pi	2019	implements a Raspberry Pi-based security system using a prototype development approach, incorporating sensor integration, embedded software design, and real-time monitoring capabilities, with system performance evaluated through reliability testing, response time measurement, and resource utilization metrics.	Low-cost implementation	Limited processing power

Literature Review

Sr No.	Title	Year	Methodology	Strengths	Weakness
5	Security Surveillance System with Email Notification Using Raspberry Pi	2023	Develops a Raspberry Pi-based surveillance system with email notification capabilities using a practical implementation approach, combining motion detection sensors, camera modules, and automated alerting mechanisms, with performance assessed through detection accuracy, notification reliability, and system latency measurements .	Effective in immediate alerts	Not specific to phishing detection

Sr No	Title	Year	Methodology	Strengths	Weakness
6	Voice Recognition and Document Classification Based Data Analysis for Voice Phishing Detection	2020	The research by Kim et al. (2021) addresses the growing problem of voice phishing (vishing) by proposing a new detection method. Its core mythology is that analyzing the transcribed audio of phone calls using document classification techniques can reveal linguistic patterns indicative of fraudulent activity, offering an innovative way to overcome the limitations of traditional security approaches against this sophisticated social engineering threat.	Innovative approach to detect voice phishing	May require large training datasets
7	Phishing Detection: A Literature Survey	2013	Khonji et al.'s (2013) survey is that a thorough analysis of past phishing detection research is essential to understand the current landscape and pave the way for improved future defenses against this enduring cyber risk.	Extensive historical review	Does not include recent AI-based techniques . .

Software Requirement Specification

1) Functional Requirements

- Fetch and monitor incoming emails from the configured IMAP server.
- Preprocess email content by cleaning text and removing noise (stopwords, special characters).
- Tokenize and pad the email text for model input.
- Predict phishing or legitimate emails using a pre-trained TensorFlow model.
- Detect and verify URLs inside emails using Google Safe Browsing API.
- Identify and flag suspicious emails based on content and URL analysis.
- Move detected phishing emails automatically to the spam folder.
- Support real-time email monitoring with continuous inbox listening.

Software Requirement Specification

1) Functional Requirements

- Maintain and update the last processed email UID to avoid duplicate processing.
- Provide console-based detailed logs for email predictions and actions taken.
- Ensure multi-threaded processing of multiple new emails for better efficiency.
- Design the system to easily extend for future features like malware and attachment detection.

Software Requirement Specification

2) Non-Functional Requirements

- Fast and responsive real-time email scanning and prediction.
- High accuracy in distinguishing phishing and legitimate emails using the trained model.
- Scalable backend capable of handling large volumes of incoming emails efficiently.
- User-friendly and clear console output for easy monitoring and troubleshooting.
- Secure IMAP authentication and safe API communications to protect user credentials and data.
- Reliable and fault-tolerant system with automatic reconnection to the mail server on failures.

Software Requirement Specification

3) Software Requirements

- Python – Programming language
- TensorFlow/Keras – For phishing detection model
- NLTK – Text preprocessing
- imaplib & email – Email access and parsing
- Requests – URL safety check via API
- Pickle – Load tokenizer
- re, os, concurrent.futures – Utilities for text, file, and parallel tasks

Algorithm

Step 1: **Email Retrieval** – Connect to the email server via IMAP and fetch new emails using unique IDs.

Step 2: **Text Preprocessing** – Clean email content using regex and remove stopwords using NLTK.

Step 3: **Tokenization & Padding** – Convert cleaned text into sequences using a Keras Tokenizer and pad them to a fixed length.

Step 4: **Phishing Detection (Deep Learning)** – Load the pre-trained TensorFlow model and predict whether the email is phishing or legitimate.

Step 5: **URL Extraction & Threat Check** – Extract URLs and verify their safety using the Google Safe Browsing API.

Step 6: **Spam Handling** – Identified phishing emails move to spam folder automatically.

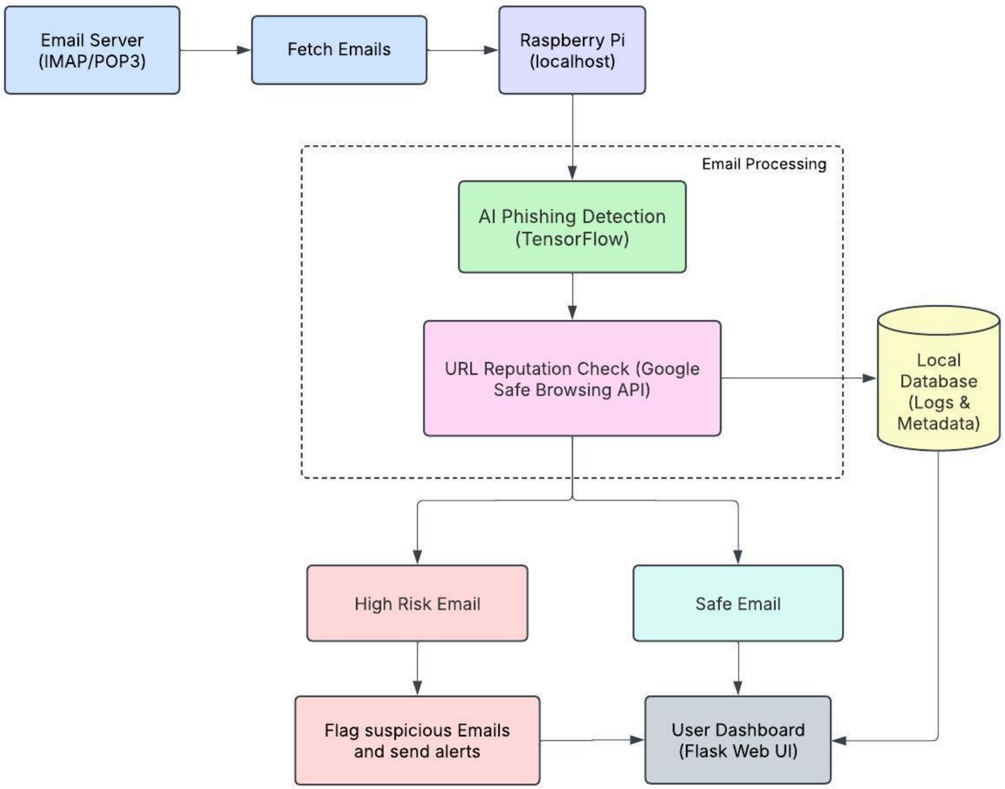
Step 7: **Real-Time Listening** – Continuously monitor the inbox for new emails and repeat the process.

Dataset

Phishing Email Dataset:

- The Phishing Email Dataset repository was used for training and evaluation.
- It contains a large number of real-world phishing and legitimate emails labeled for supervised learning.
- The dataset includes diverse examples of phishing attacks such as malicious links, deceptive subjects, fake sender addresses, and fraudulent content patterns.
- The emails are in raw text format, providing sufficient data for natural language preprocessing and feature extraction.
- We used this dataset to train our TensorFlow-based phishing detection model to accurately classify emails as either phishing or legitimate.
- Cite: Naser Abdullah Alam (2023), “Phishing Email Dataset”, Kaggle, https://www.kaggle.com/datasets/naserabdullahalam/phishing-email-dataset

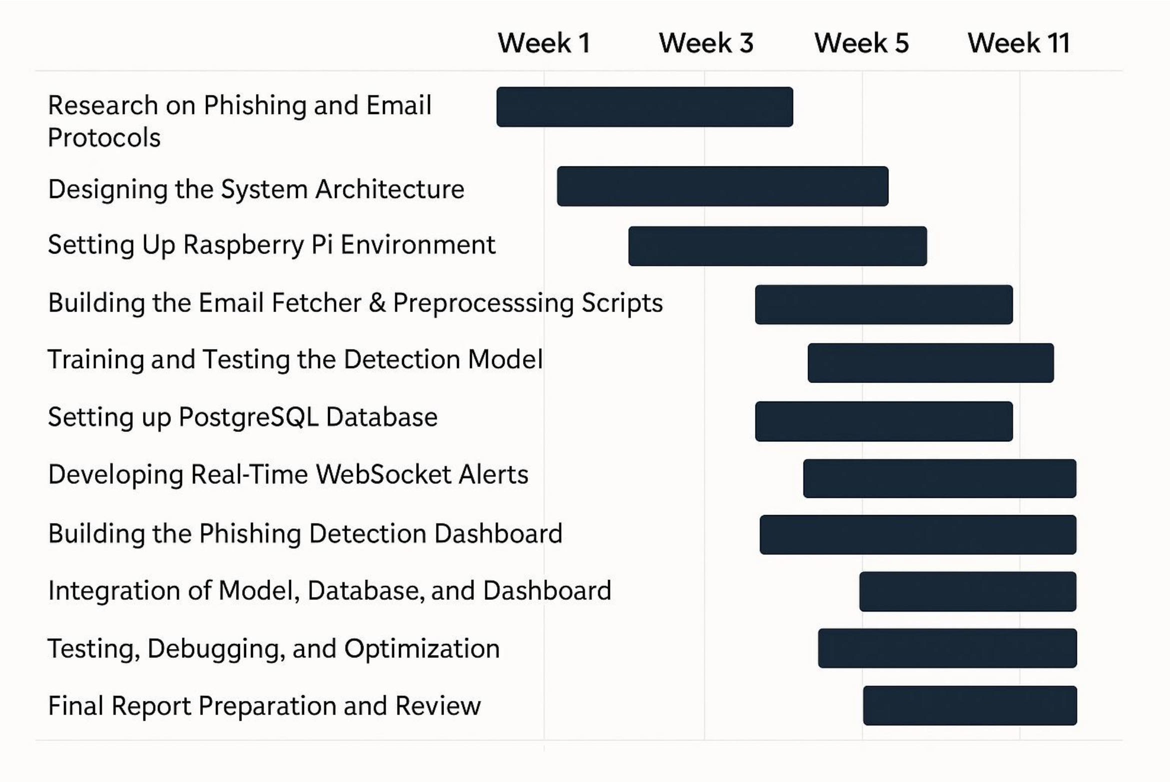
Proposed System



Email Server (IMAP/POP3)
Fetch Emails
Raspberry Pi (localhost)
Email Processing
AI Phishing Detection (TensorFlow)
URL Reputation Check (Google Safe Browsing API)
Local Database (Logs & Metadata)
High Risk Email
Safe Email
Flag suspicious Emails and send alerts
User Dashboard (Flask Web UI)

Fig.1: System Architecture

Project Plan



Timeline Chart

Software Testing

1) Unit Testing

Test ID	Module	Test Case	Input Example	Expected Result
UT01	Email Parser	Valid email header extraction	RFC-2822 formatted email	Correct metadata extraction
UT02	URL Analyzer	Malicious URL detection	hxxps://phishy-site.live/login	Flag as phishing (confidence>0.9)
UT03	NLP Preprocessor	Tokenization accuracy	"Verify your account urgently!"	5 proper tokens generated
UT04	BERT Classifier	Phishing text classification	Phishing email body	P(phishing)>0.85
UT05	Alert Generator	Threat level assignment	Score=0.93	"Critical" alert generated

Table.1: Unit Testing

Software Testing

2) Integration Testing

Test ID	Integrated Modules	Test Scenario	Verification Point
IT01	Parser + NLP + Classifier	End-to-end email analysis	<500ms processing latency
IT02	URL Analyzer + Safe Browsing API	Hybrid URL evaluation	Consistent verdicts across methods
IT03	Model Ensemble	Disagreement resolution	Majority voting applies
IT04	Alert System + Dashboard	Real-time notification delivery	<2s UI update latency

Table.2: Integration Testing

Software Testing

3) Acceptance Testing

Test ID	Dataset	Evaluation Metric	Target Threshold
AV01	Dredze Corpus	Recall (Phishing class)	>0.98
AV02	Enron Emails	Precision (Legitimate class)	>0.96
AV03	Generative Phishing	F1 Score	>0.95
AV04	Multilingual Emails	AUC-ROC	>0.97

Table.3: Acceptance Testing

Software Testing

4) Security Testing

Test ID	Vulnerability Area	Attack Method	Protection Mechanism
ST01	API Endpoints	SQL injection attempts	WAF blocks all .malicious requests
ST02	Model Serving	Adversarial email crafting	Ensemble detects 98% of evasion attempts
ST03	Data Storage	Credential stuffing	MFA blocks unauthorized access
ST04	Network Layer	MITM email interception	TLS 1.3 prevents eavesdropping

Table.4: Security Testing

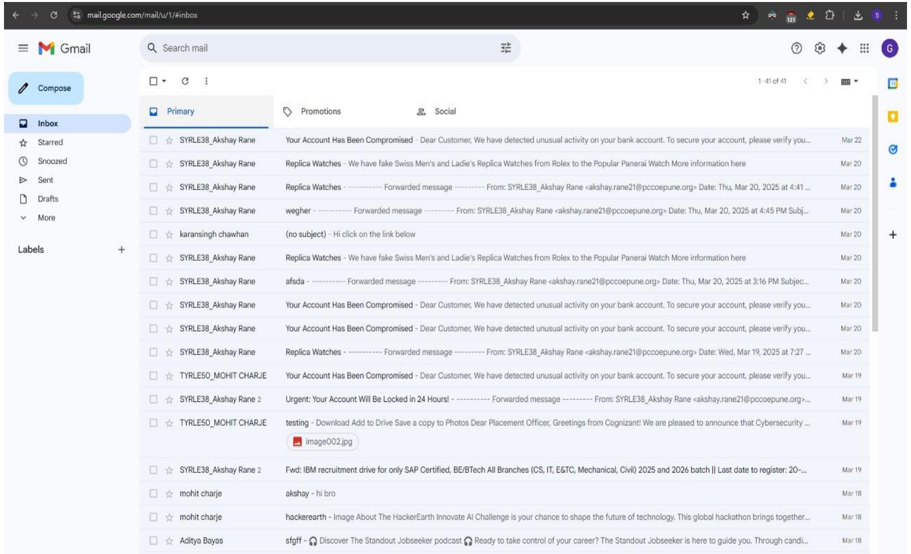
Software Testing

5) Usability Testing

Test ID	User Group	Task	Success Rate
UT01	Security Analysts	Triage 50 phishing alerts	95% correct decisions
UT02	General Users	Identify false positives	80% accuracy
UT03	Administrators	Configure detection thresholds	100% task completion

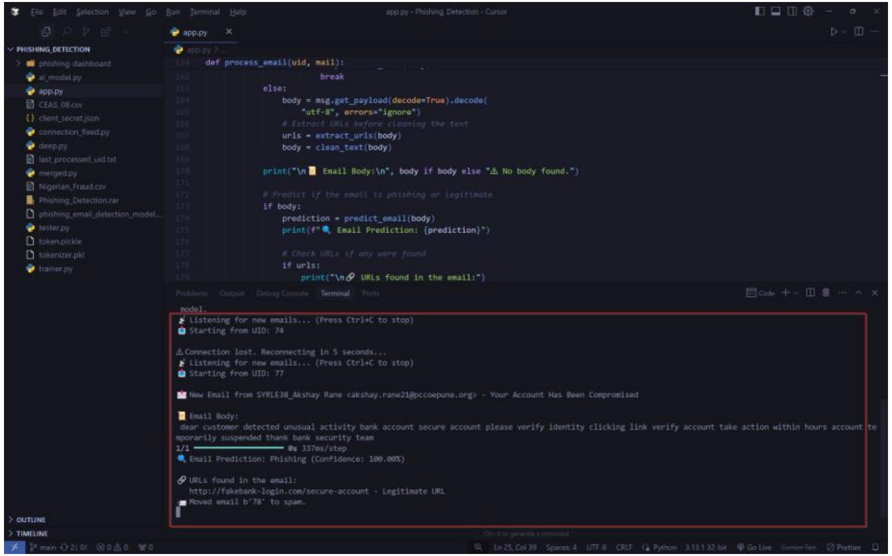
Table.5: Usability Testing

Results



SYRLE38_Akshay Rane	Your Account Has Been Compromised - Dear Customer, We have detected unusual activity on your bank account. To secure your account, please verify you...	Mar 22
SYRLE38_Akshay Rane	Replica Watches - We have fake Swiss Men's and Ladie's Replica Watches from Rolex to the Popular Panerai Watch More information here	Mar 20
SYRLE38_Akshay Rane	Replica Watches - ---------- Forwarded message --------- From: SYRLE38_Akshay Rane <akshay.rane21@pccoepune.org> Date: Thu, Mar 20, 2025 at 4:41 ...	Mar 20
SYRLE38_Akshay Rane	wegher - ---------- Forwarded message --------- From: SYRLE38_Akshay Rane <akshay.rane21@pccoepune.org> Date: Thu, Mar 20, 2025 at 4:45 PM Subj...	Mar 20
karansingh chawhan	(no subject) - Hi click on the link below	Mar 20
SYRLE38_Akshay Rane	Replica Watches - We have fake Swiss Men's and Ladie's Replica Watches from Rolex to the Popular Panerai Watch More information here	Mar 20
SYRLE38_Akshay Rane	afsda - ---------- Forwarded message --------- From: SYRLE38_Akshay Rane <akshay.rane21@pccoepune.org> Date: Thu, Mar 20, 2025 at 3:16 PM Subjec...	Mar 20
SYRLE38_Akshay Rane	Your Account Has Been Compromised - Dear Customer, We have detected unusual activity on your bank account. To secure your account, please verify you...	Mar 20
SYRLE38_Akshay Rane	Your Account Has Been Compromised - Dear Customer, We have detected unusual activity on your bank account. To secure your account, please verify you...	Mar 20
SYRLE38_Akshay Rane	Replica Watches - ---------- Forwarded message --------- From: SYRLE38_Akshay Rane <akshay.rane21@pccoepune.org> Date: Wed, Mar 19, 2025 at 7:27 ...	Mar 20
TYRLE50_MOHIT CHARJE	Your Account Has Been Compromised - Dear Customer, We have detected unusual activity on your bank account. To secure your account, please verify you...	Mar 19
SYRLE38_Akshay Rane 2	Urgent: Your Account Will Be Locked in 24 Hours! - ---------- Forwarded message --------- From: SYRLE38_Akshay Rane <akshay.rane21@pccoepune.org>...	Mar 19
TYRLE50_MOHIT CHARJE	testing - Download Add to Drive Save a copy to Photos Dear Placement Officer, Greetings from Cognizant! We are pleased to announce that Cybersecurity ... image002.jpg	Mar 19
SYRLE38_Akshay Rane 2	Fwd: IBM recruitment drive for only SAP Certified, BE/BTech All Branches (CS, IT, E&TC, Mechanical, Civil) 2025 and 2026 batch || Last date to register: 20-...	Mar 19
mohit charje	akshay - hi bro	Mar 18
mohit charje	hackerearth - Image About The HackerEarth Innovate AI Challenge is your chance to shape the future of technology. This global hackathon brings together...	Mar 18
Aditya Bayas	sfgff - 🎧 Discover The Standout Jobseeker podcast 🎧 Ready to take control of your career? The Standout Jobseeker is here to guide you. Through candi...	Mar 18

Img.1:Client Gmail Inbox

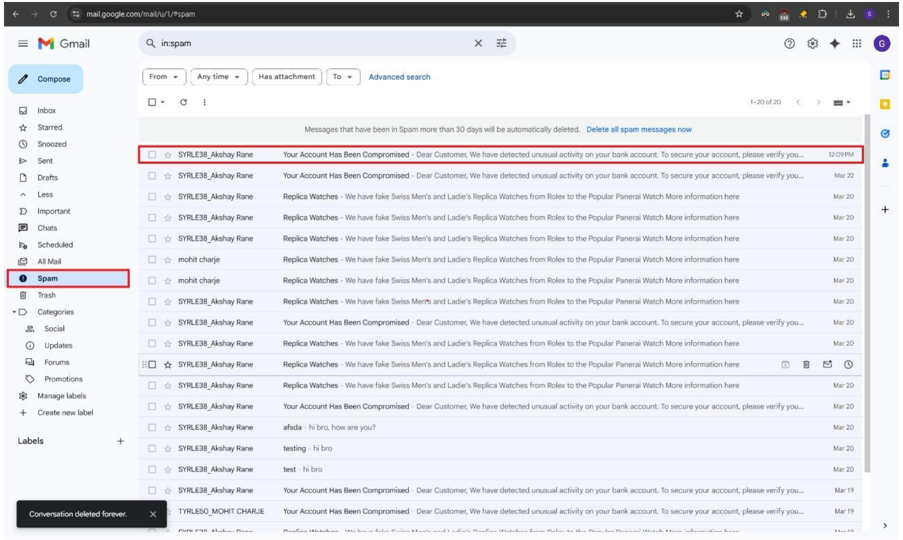


```python
def process_email(uid, mail):
                    break
            else:
                body = msg.get_payload(decode=True).decode(
                    "utf-8", errors="ignore")
                # Extract URLs before cleaning the text
                urls = extract_urls(body)
                body = clean_text(body)

            print("\n📜 Email Body:\n", body if body else "⚠ No body found.")

            # Predict if the email is phishing or legitimate
            if body:
                prediction = predict_email(body)
                print(f"🔍 Email Prediction: {prediction}")

                # Check URLs if any were found
                if urls:
                    print("\n🔗 URLs found in the email:")
```

```
model.
📡 Listening for new emails... (Press Ctrl+C to stop)
📩 Starting from UID: 74

⚠Connection lost. Reconnecting in 5 seconds...
📡 Listening for new emails... (Press Ctrl+C to stop)
📩 Starting from UID: 77

📩 New Email from SYRLE38_Akshay Rane <akshay.rane21@pccoepune.org> - Your Account Has Been Compromised

📜 Email Body:
 dear customer detected unusual activity bank account secure account please verify identity clicking link verify account take action within hours account te
mporarily suspended thank bank security team
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 337ms/step
🔍 Email Prediction: Phishing (Confidence: 100.00%)

🔗 URLs found in the email:
   http://fakebank-login.com/secure-account - Legitimate URL
📨 Moved email b'78' to spam.
```

Img.2: Email Processing App

Results

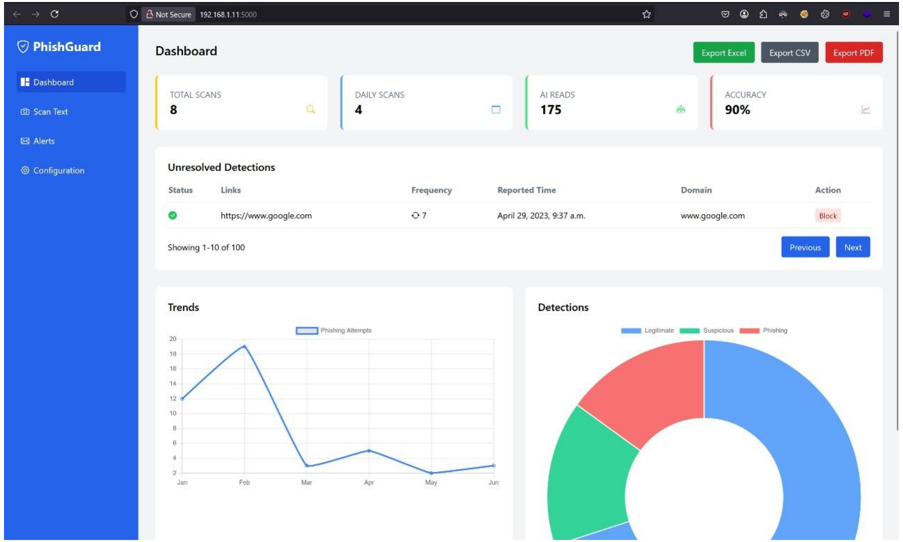


Messages that have been in Spam more than 30 days will be automatically deleted. Delete all spam messages now

SYRLE38_Akshay Rane	Your Account Has Been Compromised - Dear Customer, We have detected unusual activity on your bank account. To secure your account, please verify you...	12:09 PM
SYRLE38_Akshay Rane	Your Account Has Been Compromised - Dear Customer, We have detected unusual activity on your bank account. To secure your account, please verify you...	Mar 22
SYRLE38_Akshay Rane	Replica Watches - We have fake Swiss Men's and Ladie's Replica Watches from Rolex to the Popular Panerai Watch More information here	Mar 20
SYRLE38_Akshay Rane	Replica Watches - We have fake Swiss Men's and Ladie's Replica Watches from Rolex to the Popular Panerai Watch More information here	Mar 20
SYRLE38_Akshay Rane	Replica Watches - We have fake Swiss Men's and Ladie's Replica Watches from Rolex to the Popular Panerai Watch More information here	Mar 20
mohit charje	Replica Watches - We have fake Swiss Men's and Ladie's Replica Watches from Rolex to the Popular Panerai Watch More information here	Mar 20
mohit charje	Replica Watches - We have fake Swiss Men's and Ladie's Replica Watches from Rolex to the Popular Panerai Watch More information here	Mar 20
SYRLE38_Akshay Rane	Replica Watches - We have fake Swiss Men's and Ladie's Replica Watches from Rolex to the Popular Panerai Watch More information here	Mar 20
SYRLE38_Akshay Rane	Your Account Has Been Compromised - Dear Customer, We have detected unusual activity on your bank account. To secure your account, please verify you...	Mar 20
SYRLE38_Akshay Rane	Replica Watches - We have fake Swiss Men's and Ladie's Replica Watches from Rolex to the Popular Panerai Watch More information here	Mar 20
SYRLE38_Akshay Rane	Replica Watches - We have fake Swiss Men's and Ladie's Replica Watches from Rolex to the Popular Panerai Watch More information here	
SYRLE38_Akshay Rane	Replica Watches - We have fake Swiss Men's and Ladie's Replica Watches from Rolex to the Popular Panerai Watch More information here	Mar 20
SYRLE38_Akshay Rane	Your Account Has Been Compromised - Dear Customer, We have detected unusual activity on your bank account. To secure your account, please verify you...	Mar 20
SYRLE38_Akshay Rane	afsda - hi bro, how are you?	Mar 20
SYRLE38_Akshay Rane	testing - hi bro	Mar 20
SYRLE38_Akshay Rane	test - hi bro	Mar 20
SYRLE38_Akshay Rane	Your Account Has Been Compromised - Dear Customer, We have detected unusual activity on your bank account. To secure your account, please verify you...	Mar 19
TYRLE50_MOHIT CHARJE	Your Account Has Been Compromised - Dear Customer, We have detected unusual activity on your bank account. To secure your account, please verify you...	Mar 19

Conversation deleted forever.

Img.3: Phishing email moved to Spam



Dashboard

TOTAL SCANS 8

DAILY SCANS 4

AI READS 175

ACCURACY 90%

Unresolved Detections

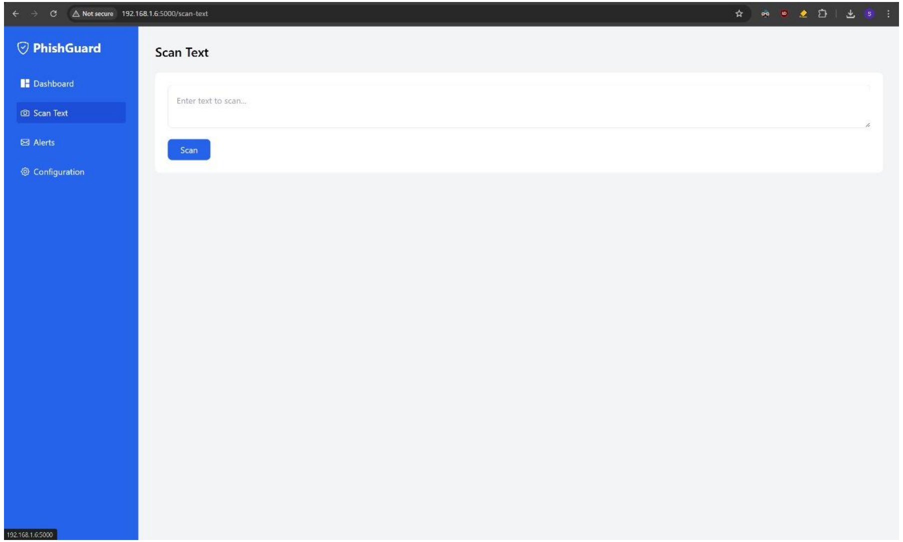
Status	Links	Frequency	Reported Time	Domain	Action
✓	https://www.google.com	7	April 29, 2023, 9:37 a.m.	www.google.com	Block

Showing 1-10 of 100

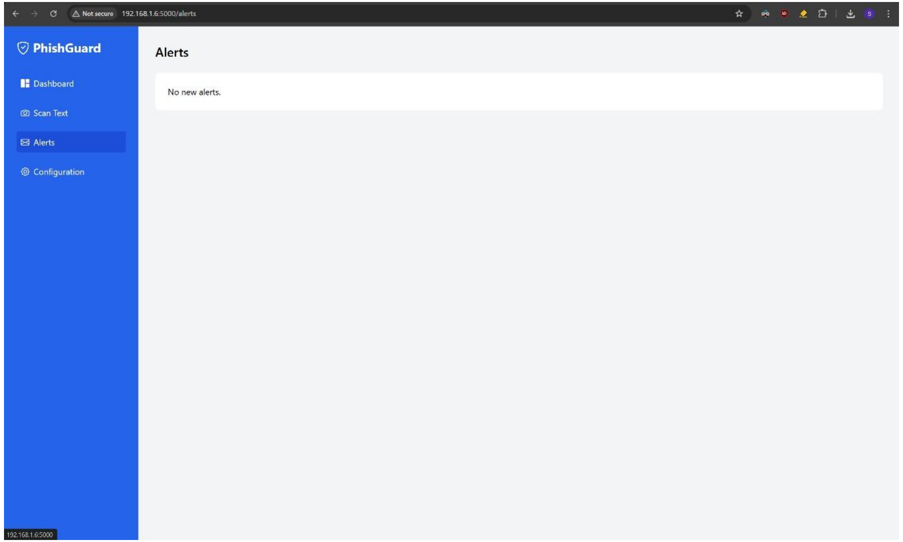
Previous Next

Img.4: Dashboard

Results



Img.6: Scan custom email text content



No new alerts.

Img.7: Receive Alerts on the Dashboard

Contribution to Sustainable Development Goals

- Promotes secure digital infrastructure by protecting users from phishing threats in email communications.
- Supports responsible innovation by using AI ethically to enhance cybersecurity and reduce human error in threat detection.
- Enables digital literacy and awareness by helping users identify phishing attempts and understand online threats.
- Protects personal and financial information, contributing to reduced cybercrime and increased online safety.
- Empowers organizations and individuals with a cost-effective, real-time solution for email threat detection using open-source tools and low-cost hardware like Raspberry Pi.
- Encourages sustainable tech development by leveraging lightweight models and efficient computing, aligning with green computing practices.

Future Goals

1) Enhance AI model with advanced deep learning techniques
2) Expand detection to include SMS, social media, and messaging platforms
3) Add detection of malware and suspicious email attachments
4) Implement adaptive learning for zero-day phishing attacks
5) Optimize deployment for enterprise and cloud environments
6) Enhance dashboard with advanced visualizations and user insights

Conclusion

The project "AI-Powered Anti-Phishing Email Firewall" marks a significant step forward in the field of cybersecurity and email protection. By leveraging machine learning models and real-time threat detection techniques, the system provides a proactive solution to combat phishing attacks. With features such as intelligent email filtering, a real-time monitoring dashboard, and integration with a PostgreSQL database, the project ensures effective identification and management of phishing threats. This solution not only enhances email security for individuals and organizations but also contributes to building a safer digital communication environment. The project sets a strong foundation for future enhancements, such as malware detection and adaptive threat intelligence integration.

References

[1] O. Lamina, W. Ayuba, O. Adebiyi, G. Michael, O.-O. Samuel, and K. Samuel, “Ai Powered Phishing Detection And Prevention,” Path of Science, vol. 10, pp. 4001–4010, Apr. 2024, doi: 10.22178/pos.112-7

[2] C. S. Eze and L. Shamir, “Analysis and Prevention of AI-Based Phishing Email Attacks,” May 2024.

[3] A. Arun and N. Abosata, “Next Generation of Phishing Attacks Using AI Powered Browsers,” arXiv preprint arXiv:2406.12547, Jun. 2024.

[4] M. Al-Rawi, M. Abdulhamid, and S. Sheshai, “Design of Security System Based on Raspberry-PI,” The Scientific Bulletin of Electrical Engineering Faculty, vol. 19, pp. 56– 61, Apr. 2019, doi: 10.1515/sbeef-2019-0022.

[5] N. O. Nwazor and S. I. Orakwue, “Security Surveillance System with Email Notification Using Raspberry Pi,” Iconic Research And Engineering Journals, vol. 6, no. 10, pp. 190– 195, 2023.

[6] J. Kim, G. Hong, and H. Chang, “Voice Recognition and Document Classification-Based Data Analysis for Voice Phishing Detection,” 2021.

[7] M. Khonji, Y. Iraqi, and A. Jones, “Phishing Detection: A Literature Survey,” IEEE Communications Surveys & Tutorials, vol. 15, no. 4, pp. 2091–2121, 2013.

[8] A. C. Bahnsen, I. Torroledo, L. D. Camacho, and S. Villegas, “DeepPhish: Simulating Malicious AI,” arXiv preprint arXiv:1805.07817, May 2018.

[9] O. K. Sahingoz, E. Buber, O. Demir, and B. Diri, “Machine Learning-Based Phishing Detection from URLs,” Expert Syst Appl, vol. 117, pp. 345–357, Mar. 2019, doi: 10.1016/j.eswa.2018.09.051.

THANK YOU !!!!