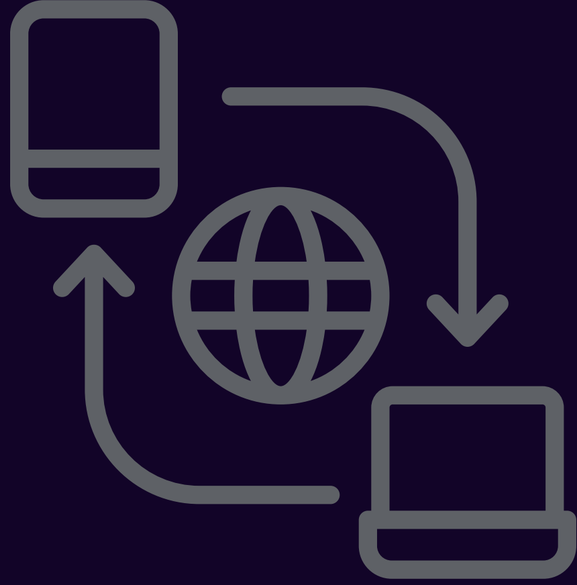


PhishIntel

AI-powered URL and File Scam Detector

TEAM MEMBERS:
Mayur Koregaonkar
Anjali Yadav

Introduction

What is Phishing?

Phishing is a type of cyberattack where hackers trick people into sharing personal info by pretending to be trusted sources, like banks or websites.

They use fake emails, websites, or links to:

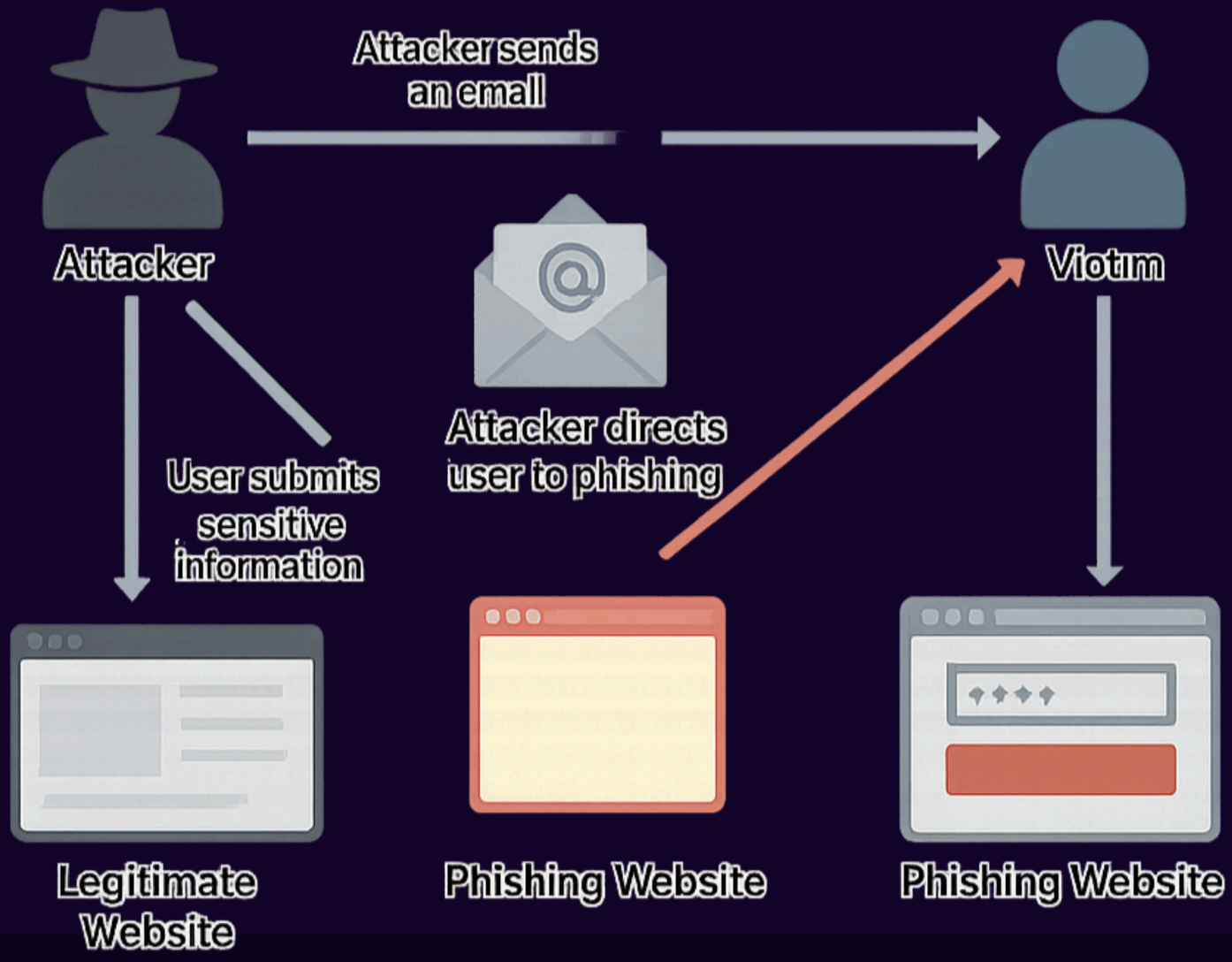
- Steal passwords
- Access bank accounts
- Install malware
- Commit identity theft

Phishing works by fooling people, not by hacking systems, which makes it a serious threat.

Problem Statement

- Phishing is one of the most dangerous and common cybersecurity threats today.
- Users are constantly exposed to scam emails, phishing links, and fake attachments.
- Traditional rule-based detection systems fail when attackers use new tricks (e.g., homograph URLs or scam text variation).

How Does Phishing Work?

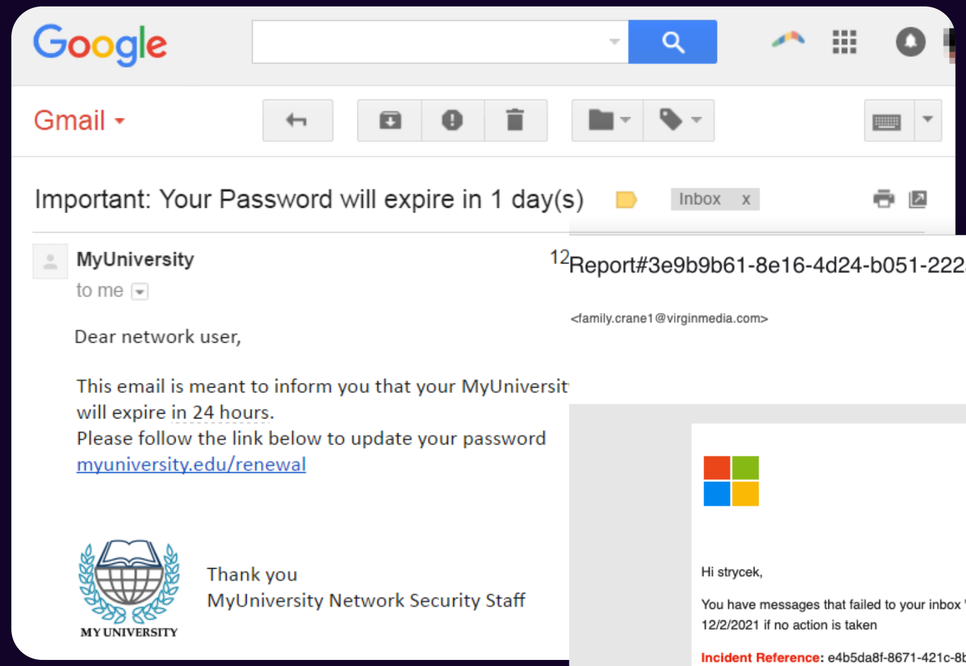


Types of Phishing Attacks

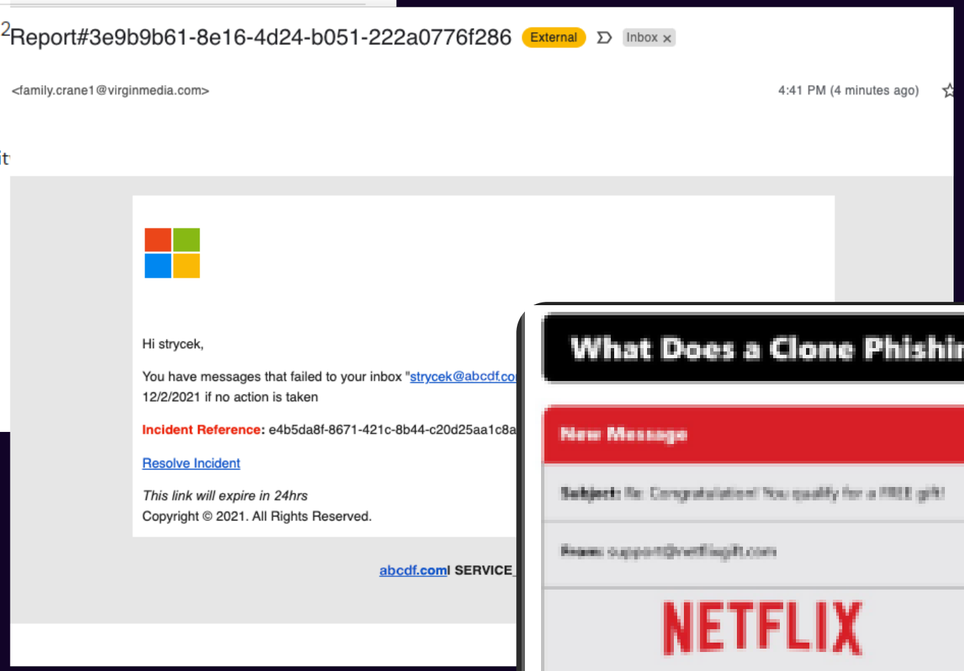
There are various types through which the act of phishing is conducted

- Email Phishing :- attempt to steal sensitive information via an email
- Spear Phishing :- email message are sent to specific people within an organization, usually high privilege
- Clone Phishing :- email that you might have received from an authentic sender but sent from spoofed email id.
- Voice Phishing :- use of fraudulent phone calls to trick people into giving money or revealing personal information
- Smishing (SMS Phishing) :- type of phishing attack that uses text messages (SMS) to trick individuals into revealing personal information or installing malware.
- Angler Phishing (Social Media Phishing)

Emai Phishing

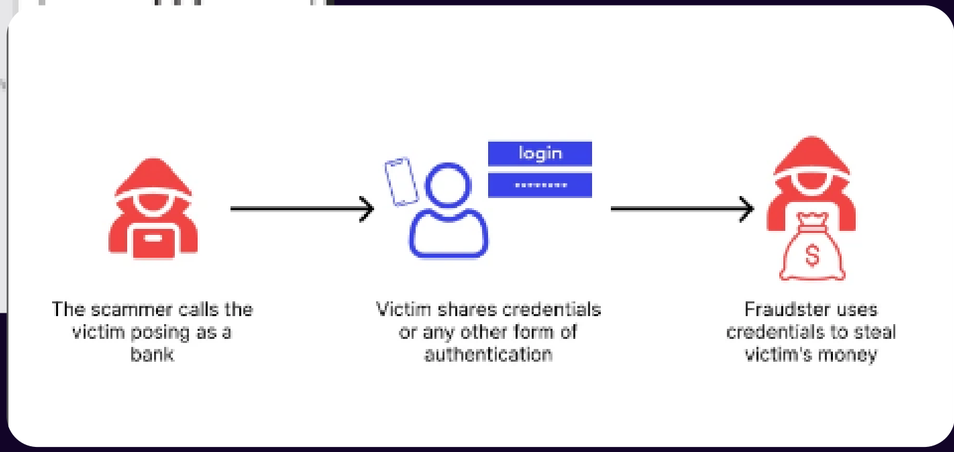


Spear Phishing



Clone Phishing

Voice Phishing



Project Objective

- The main objective of this project is to develop an AI-powered system that can automatically detect whether a given URL is legitimate or a phishing attempt.
- This tool scans URLs and uploaded documents (PDF/TXT).
- Uses two technologies together:
    a. Machine Learning (ML): Trained on phishing datasets.
    b. Google Gemini AI: Analyzes textual and contextual data.
- Goal: Accurate, fast, and intelligent detection.

Features & Workflow

1. User Input
    a. User submits a URL via web form.
    b. Example: http://secure-login.paytm-verification.com
2. Feature Extraction
    a. System analyzes the URL structure:
        i. URL length
        ii. Special characters (@, -)
        iii. IP address vs domain
        iv. Use of HTTPS
3. Machine Learning Prediction
    a. Features sent to a trained ML model (Logistic Regression/SVM)
    b. Predicts based on prior phishing data
4. Output & Explanation
    a. Result shown: “Legit” or “Phishing”
    b. Also displays “Secure” or “Scam” for user clarity

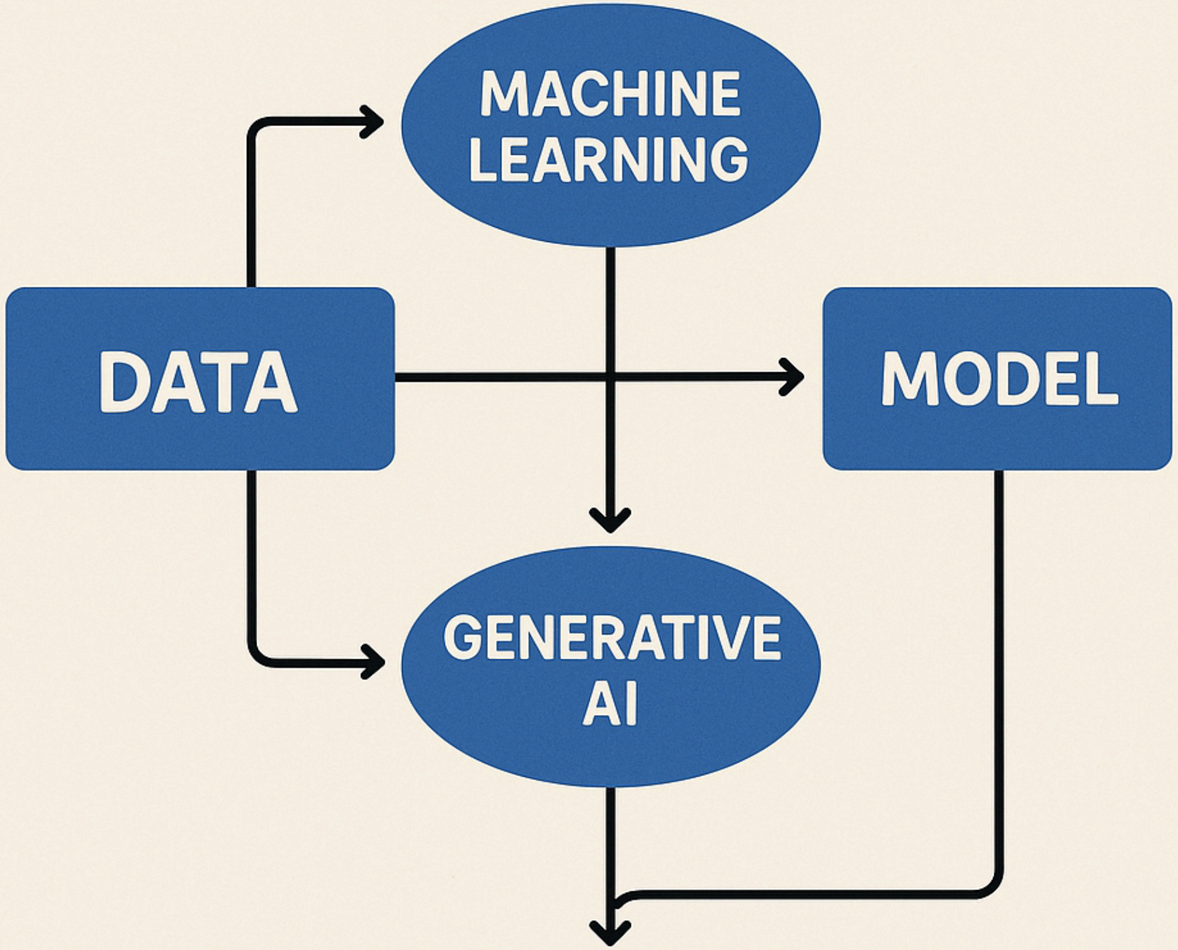
Technologies Used

Tool	Purpose
1. Flask	1. Web backend & API routing
2. HTML/CSS	2. User interface
3. scikit-learn	3. ML-based URL classification
4. TfidfVectorizer	4. Converts URLs to numeric vectors
5. RandomForestClassifier	5. Predicts malicious/safe URLs
6. Gemini AI (google-generativeai)	6. Analyzes scam content in files/URLs
7. PyPDF2	7. Extracts text from PDF files
8. joblib	8. Saves & loads models/vectorizers
9. malicious_phish.csv	9. Dataset for ML training

System Architecture

- Frontend: User inputs URL or uploads file.
- Backend: Flask handles inputs.
- ML Model: Predicts based on trained dataset.
- Gemini AI: Classifies based on natural language reasoning.
- Output: Displays threat result and confidence.

HOW GENERATIVE AI AND ML WORK TOGETHER



Real-World Use Cases

- Email Gateways: Scan incoming email links via Outlook/Gmail APIs.
- Browser Extensions: Warn users of suspicious URLs in real-time.
- Enterprise Firewalls: Block malicious sites on internal networks.
- Training Tools: Simulate phishing to educate employees.
- API for SaaS: Offer URL checking in antivirus, chat apps, CRMs. of body text

🚀 Future Enhancements

✅ File Scanning: Analyze email attachments for malware.
✅ NLP Analysis: Understand context and detect social engineering.
✅ Reputation Scoring: Use VirusTotal, WHOIS, and blacklists.
✅ Live Updates: Auto-train model with real-time threat feeds.
✅ Cloud + Feedback: Add a dashboard & user reporting system.

PREVENTIONS FROM PHISHING ATTACK

- Known what a phishing scam look like
- Don’t click on that link
- Don't give your info to undecured site
- Rotate your Password regulary
- install firewall
- install anti phishing software
- check mail or text on website

Conclusion

- PhishIntel solves a real cybersecurity problem
- Combines traditional ML with modern AI (Gemini)
- Real-time, easy-to-use, and highly effective
- Ready for real-world use and future expansion

Thank you

GitHub Link :- https://github.com/rmayur0323/PhishIntel.git