AI in Social Engineering and Phishing Campaigns

📌 **Organized by: Digisuraksha Parhari Foundation**

🤝 **Powered by: Infinisec Technologies Pvt. Ltd.**

#Team Members/Author :

- ❖ **Prathamesh Mareppa Shinde 248360**
- ❖ **Ajinkya Sachin Mohite 248335**
- ❖ **Aditya Chandu Gaikwad 248311**
- ❖ **Sneha Bhagwan Kamble 248425**
- ❖ **Rutuja Rajendra Kurane 248429**

#Institution : Mulund College Of Commerce

#Date of Submission : 12/05/2025

Abstract ➔

The rapid evolution of artificial intelligence has significantly impacted cybersecurity, not only strengthening defensive capabilities but also empowering more sophisticated offensive strategies. This research focuses on the growing role of AI—particularly natural language processing (NLP) models such as GPT—in enhancing social engineering and phishing campaigns. It examines how AI is used to craft convincing, highly personalized phishing content at scale, including realistic emails, deepfake voice recordings, and chatbot-based fraud. To explore these threats in practice, the study introduces a simulation tool designed to generate AI-driven phishing attacks and assess their detectability using a machine learning-based filtering system. Additionally, the work discusses the ethical considerations surrounding such tools, the potential for abuse, and mitigation strategies, including the use of AI to bolster defenses. By analyzing these emerging threats, the research aims to contribute to the development of more adaptive and intelligent cybersecurity solutions.

Problem Statement & Objective ➔

Problem Statement :

The increasing use of AI in social engineering and phishing campaigns poses a serious threat to cybersecurity by enabling the creation of highly convincing, scalable, and targeted attacks that evade traditional detection mechanisms. There is an urgent need to understand, detect, and mitigate the misuse of AI in these deceptive practices.

Objective :

- To analyze the integration of AI in the design and execution of phishing campaigns.
- To identify current cybersecurity limitations in detecting AI-generated attacks.
- To explore and implement AI-based detection tools.
- To evaluate the effectiveness of these tools through empirical testing.
- To assess the ethical and societal implications of AI in cybercrime.

Literature Review ➔

Several studies highlight the evolution of phishing from generic spam to highly targeted campaigns powered by AI. Research shows that Natural Language Processing (NLP) and deepfake technologies are increasingly used to craft convincing emails, voice messages, and videos. In this project we had mainly use a simple technique which is based on the checking messages or emails. In this technique we found the result as legislative or phishing as the prompt typed.

Generally we used ChatGPT as references, and some information from Google chrome.

- Related to Gaps,

1. Detection of AI-Generated Content

    - Gap: Many phishing emails use AI-generated content that still contains subtle tells (unnatural phrasing, overly generic language, repetition).

    - Opportunity: Improved AI-content detectors and linguistic forensics can exploit these artifacts to detect and flag phishing.

2. Real-Time Contextual Awareness

    - Gap: AI-generated phishing emails generally cannot respond in real-time to victim responses in a believable and coherent way.

- ➢ Example: Business Email Compromise (BEC) attacks often fail in back-and-forth conversations without a human in the loop.

3. Limited Cultural and Organizational Awareness
    - ➢ Gap: AI-generated phishing lacks understanding of organizational hierarchies, local customs, or corporate language.
    - ➢ Impact: Reduces effectiveness of impersonation attacks or spoofed executive communications.

- ➢ Related to Concepts,
    1. AI-Powered Social Engineering
        - ➢ Natural Language Processing (NLP): AI can generate convincing, grammatically correct, and context-aware messages that mimic human communication.
        - ➢ Data Mining & Personalization: AI can scrape social media and public data to craft personalized phishing emails (spear phishing).
        - ➢ Chatbots for Impersonation: Malicious AI-powered chatbots can impersonate trusted entities to build rapport and gain trust.
    2. AI-Enhanced Email Phishing
        - ➢ Checking email/messages for safety purpose.
        - ➢ Email Spoofing: AI helps automate the process of mimicking legitimate email structures. Attacker sends the message so it appears to come from a trusted source, and on clicking or Downloading malware or revealing sensitive information like passwords or financial data. They could get the access of out account.
    3. Evasion Techniques
    4. Phishing-as-a-Service (PhaaS)
    5. Business Email Compromise (BEC)

Research Methodology ➔

1. Qualitative Analysis – Review of the case study,

🧾 Case Study: AI-Powered Phishing Email Detection Tool at TrustBank

🏢 Organization Overview

Name: TrustBank

Industry: Financial Services

Employees: 1,200+

Location: Texas, USA

Problem: Increasingly sophisticated phishing emails bypassing filters

🚨 The Problem

TrustBank's IT team reported that traditional spam filters were unable to catch advanced phishing attempts. Emails were crafted to look like internal HR notices or executive communications. In one case, a staff member clicked a malicious link, exposing internal login credentials and triggering a minor data breach.

2. Quantitative Analysis – As we metioned in the Repository of GitHub the tool, it's working style and it's result. You can gather information from it. We used the email phishing technique.

3. Tool-Based Evaluation – We created the tool via VS Code and the concerned files as app.py and phishing_emails.csv.

🛡️ **Evaluation Points**: Defensive Use of AI Against Phishing

I. Detection Accuracy

   - Ability to correctly identify phishing.

II. False Positive/Negative Rate

   - Number of legitimate emails flagged incorrectly and phishing emails missed.

III. Speed of Detection

   - Time taken to process and classify incoming emails.

IV. Robustness Against Adversarial Examples

- Resilience to modified or obfuscated phishing emails.

V. Scalability

- Performance on large volumes of real-time email traffic.

Tool Implementation ➔

- AI in Social Engineering and Phishing Campaigns, specifically Email phishing we worked on the python generated tool. We attached in the Repo.
- We used libraries like Flask, request, pandas, TfidfVectorizer, MultinomialNB
- Flask, Request : It handle the incoming HTTPs requests.
- pandas : It is used here to load and handle CSV data.
- TfidfVectorizer : which converts a collection of text documents into numerical feature vectors using the TF-IDF (Term Frequency-Inverse Document Frequency) method.

This are the libraries which used in the tool. Often Kaggle datasets which can set the 0 or 1 value to the result of phishing or legitimate.

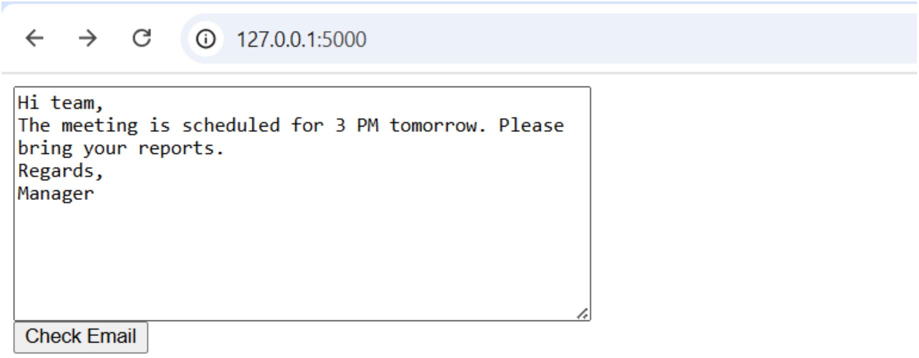
🔧 To implement this tool we required :

VS Code, chrome, Kaggle dataset (phishing_emails.csv), app.py, and sample of messages.

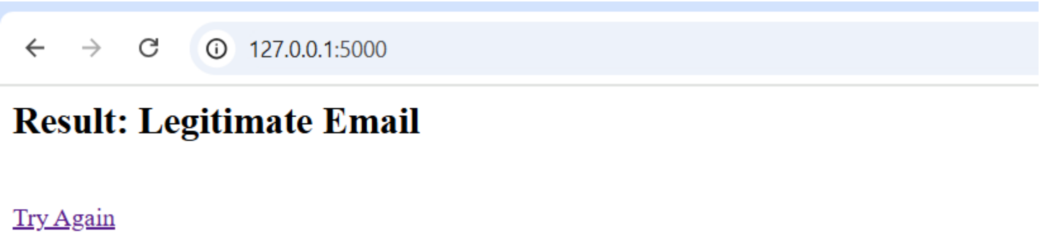
Simply run the app.py file via VS Code and go to the terminal and find the url IP address link. It is clickable so go through it and than we place on the chrome page. So type sample message and check. Than we move to new page and got the result such as phishing or legitimate.

Results & Observations ➔

For Legitimate Email



```
Hi team,
The meeting is scheduled for 3 PM tomorrow. Please
bring your reports.
Regards,
Manager
```

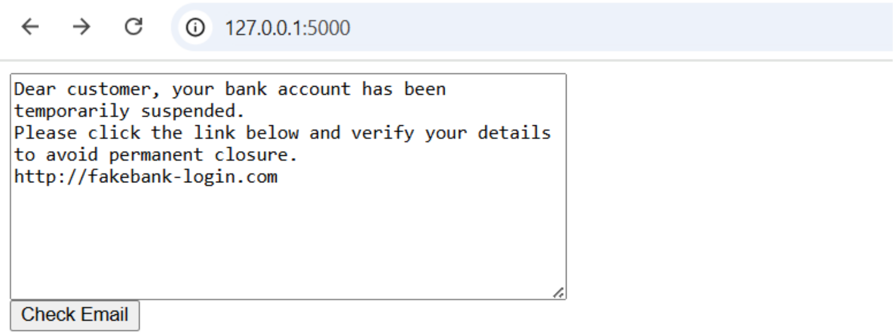


Result: Legitimate Email

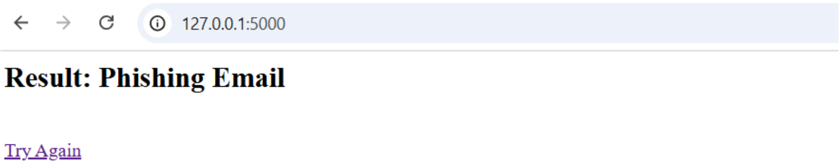
Try Again

----

for phishing Email



```
Dear customer, your bank account has been
temporarily suspended.
Please click the link below and verify your details
to avoid permanent closure.
http://fakebank-login.com
```



Result: Phishing Email

Try Again

Here we run our tool and get the result.
On observing this resultant. We can surely said that the implementation of this tool in the market level is better and can useful for user's safety.

Ethical Impact & Market Relevance ➔

The dual-use nature of AI raises ethical concerns. Misuse of AI challenges existing legal frameworks and user trust. Companies must integrate AI ethics into cybersecurity policies.
Market Relevance:

- Rising demand for AI-driven cybersecurity solutions.
- Increased investment in phishing detection platforms.
- Regulatory pressures driving AI accountability and transparency.

Future Scope ➔

The future of AI in social engineering and email phishing campaigns presents a significant cybersecurity threat due to its ability to automate and personalize attacks at scale. AI can generate highly convincing phishing emails by analyzing a target's digital footprint, including social media activity and communication patterns, to craft messages that appear authentic and contextually relevant. This level of sophistication makes traditional spam filters and detection systems increasingly ineffective. Additionally, the rise of deepfake technology and AI-powered chatbots may further enhance the credibility of phishing attempts, making them more difficult for users to identify. As AI continues to evolve, phishing campaigns are expected to become more adaptive, multi-channel, and capable of evading even advanced security systems, posing a persistent and growing challenge to individuals and organizations alike.

References ➔

1. Europol – Malicious Uses of Artificial Intelligence
   📎 Link to Report
   Description: This report outlines how AI technologies are being used in cybercrime, including the automation and personalization of phishing emails to increase success rates.

2. NIST (National Institute of Standards and Technology) – AI and Phishing

   📎 Link to NIST Blog

   Description: Discusses the challenges and dangers of AI-generated phishing, and what steps organizations can take to counter them.
3. IBM X-Force – Threat Intelligence Index 2024

   📎 Link to Report

   Description: Covers the role of AI in current cyber threats, including how phishing campaigns have evolved with machine learning tools.
4. Keeper Security – How AI Is Making Phishing Attacks More Dangerous

   📎 Read Article

   Description: Explores how AI-generated phishing emails are becoming more inconspicuous by eliminating common errors, making them harder to detect and more convincing to recipients.
5. TechTarget – How AI is Making Phishing Attacks More Dangerous

   📎 Read Article

   Description: Provides insights into the warning signs of AI-powered phishing emails and emphasizes the importance of security awareness training to identify and mitigate such threats

Case Study ➔

ICICI Bank Phishing Case Study