Phishing Website Detection Tool

Project Title

Phishing Website Detection Tool

Submitted By

Soham Pramod Tayade

BSc in Cyber and Digital Science

Dr. D.Y. Patil Arts, Commerce and Science College, Pune

Academic Year: 2024-2025

Problem Statement

Phishing websites attempt to trick users into entering personal information, leading to identity theft and financial fraud.

Objective

To build a tool that detects phishing URLs using rule-based logic and machine learning.

Technologies Used

- Python
- Pandas, Scikit-learn, Regex
- Tkinter for GUI
- Jupyter Notebook for development
- CSV datasets from Kaggle/Mendeley

Methodology

1. Rule-Based Detection: Uses regex and suspicious patterns to flag phishing URLs.
2. Machine Learning: Trained logistic regression model using TF-IDF features to classify URLs.

Implementation & Dataset

Phishing Website Detection Tool

Used a combined dataset from Kaggle and Mendeley. Cleaned data, extracted features, and trained a Logistic Regression model.

Accuracy achieved: 94%.

Also implemented a simple GUI using Tkinter to test URLs.

Result

Rule-based detection flags suspicious patterns.

ML-based model achieves 94% accuracy.

GUI interface allows users to input URLs and check safety.

Conclusion

The tool helps reduce phishing attacks by identifying suspicious URLs before users click on them.

Future Scope

- Add real-time web crawling
- Build a browser extension
- Explore deep learning models

Signature

_____

(Soham Pramod Tayade)

Student, TYBSc Cyber & Digital Science

Date: June 2025