AI-Driven Phishing Detection

Smart cybersecurity with machine learning techniques to combat phishing threats effectively, presented by: Pooja waghmare, Pavan Nag
Internship Project : DigisurakshaParhariFoundation

Introduction to Phishing

Understanding phishing and its significant impact on cybersecurity.

01 Phishing is a **fraudulent attempt** to obtain sensitive information.

02 Over 90% of cyberattacks start with phishing, making it crucial to address.

03 Awareness and education are essential in combating this growing threat.

Phishing Threats

Understanding the increasing danger of phishing websites today.

01

Phishing websites are **growing rapidly**, targeting unsuspecting users online.

02

Traditional blacklisting methods are **ineffective** against new and evolving phishing tactics.

03

An **intelligent detection system** is essential for real-time protection against these threats.

Objective

Developing a Machine Learning-Based Detection Tool

01

Our goal is to accurately classify URLs as **phishing** or **legitimate**.

02

The tool relies on **feature-based analysis** rather than static rules for classification.

03

This approach enhances detection capabilities and provides **real-time results** for users.

Technology Stack

Overview of tools and technologies used in the project.

01

The backend is built using **Python** and **Flask** for efficient processing.

02

Machine learning is implemented with **Scikit-learn**, specifically using the Random Forest algorithm.

03

The frontend utilizes **HTML** along with Flask templates for dynamic content rendering.

DATASET OVERVIEW

Analyzing the Phishing Data for Insightful Detection

01 The dataset comprises various features crucial for identifying phishing URLs.

02 It contains information such as URL length and special characters present.

03 The size of the dataset significantly impacts the model' s accuracy and reliability.

Feature Extraction

Identifying crucial aspects for effective phishing detection.

01

Key features include URL length and presence of HTTPS.

02

Suspicious keywords and special characters play a significant role.

03

These features enhance the model's ability to classify URLs accurately.

Model Training

Building a robust classification model for phishing detection.

01

The **Random Forest Classifier** was utilized for its accuracy and efficiency.

02

An **80/20 train/test split** ensured effective model evaluation and performance assessment.

03

The model was serialized using **joblib** to facilitate real-time predictions in deployment.

Application Flow

User interaction process for detecting phishing URLs

01

Users submit URLs through a simple web form.

02

The system processes the URL and extracts relevant features.

03

Finally, the model classifies the URL as 'Phishing' or 'Legitimate'.

Demo Walkthrough

Experience the AI-driven phishing URL detection tool in action.

In this section, you will see how users can easily **submit URLs** through a web form and receive instant predictions on whether a URL is **phishing or legitimate**.

https://github.com/pavanEX31/AI-Driven-Phishing-Detection.git

Real-World Applications

Exploring practical uses of the phishing detection tool.

01 Corporate firewalls can effectively utilize this tool for enhanced security.

02 Email gateways benefit from real-time phishing URL detection to protect users.

03 Web browser plugins can provide users with immediate alerts on suspicious links.

Future Enhancements

Expanding the Tool's Capabilities for Better Detection

01 Integrating **deep learning** could improve detection accuracy significantly.

02 Adding live web scrapers would enhance real-time URL analysis capabilities.

03 Developing a browser extension would allow for seamless user protection.

Challenges Faced

Overcoming obstacles in developing our detection system

01 Generalizing features was difficult for unknown phishing types.

02 We faced inconsistencies within the dataset that needed addressing.

03 Ensuring fast predictions with Flask was a significant challenge.

Achievements

Summary of accomplishments and future development directions

01

The project successfully built a **highly accurate** phishing URL detection system.

02

Future enhancements will focus on **integrating deep learning** and advanced features.

03

This scalable solution holds significant potential for improving **cybersecurity measures**.

"A game changer in phishing detection and prevention!"
– Mark Thompson

"This tool significantly improved our cybersecurity measures."
– Sarah Jenkins

"Impressive accuracy and usability for real-time applications."
– Lisa Wong

"A vital resource for modern cybersecurity challenges."
– Raj Patel