DELHI POLICE CYBER HACKATHON

TRUSTLINK

TrustLink: Safeguarding Against Deceptive URLs - A Project for Detecting Fake and Spurious URLs

TEAM DETAILS

TEAM LEADER:

Ujjawal Saini (3rd year)

TEAM MEMBERS:

Divyanshu Shukla (2nd year)

Alorika Jain (2nd year)

Dev Bhardwaj (3rd year)

MODEL PROTOTYPE

Project Overview

- TrustLink is an innovative project leveraging machine learning models to detect malicious URLs and enhance online security. By combining static analysis with dynamic classification, our system examines URLs for potential threats, categorizing them into labels such as phishing, malware, benign, or defacement.
- The project incorporates diverse data sources, including curated host lists and a pre-trained text classification model, offering a robust defense against deceptive URLs.
- TrustLink aims to provide users with a reliable solution to identify and avoid malicious links, contributing to a safer online experience.

WORKFLOW

1 — 2 — 3 — 4

Step

User Input

Users input a URL into the TrustLink web application.

Step

Static Analysis

Comparing the entered URL against pre-loaded data from various host lists to identify patterns associated with malicious behavior.

Step

Dynamic Analysis

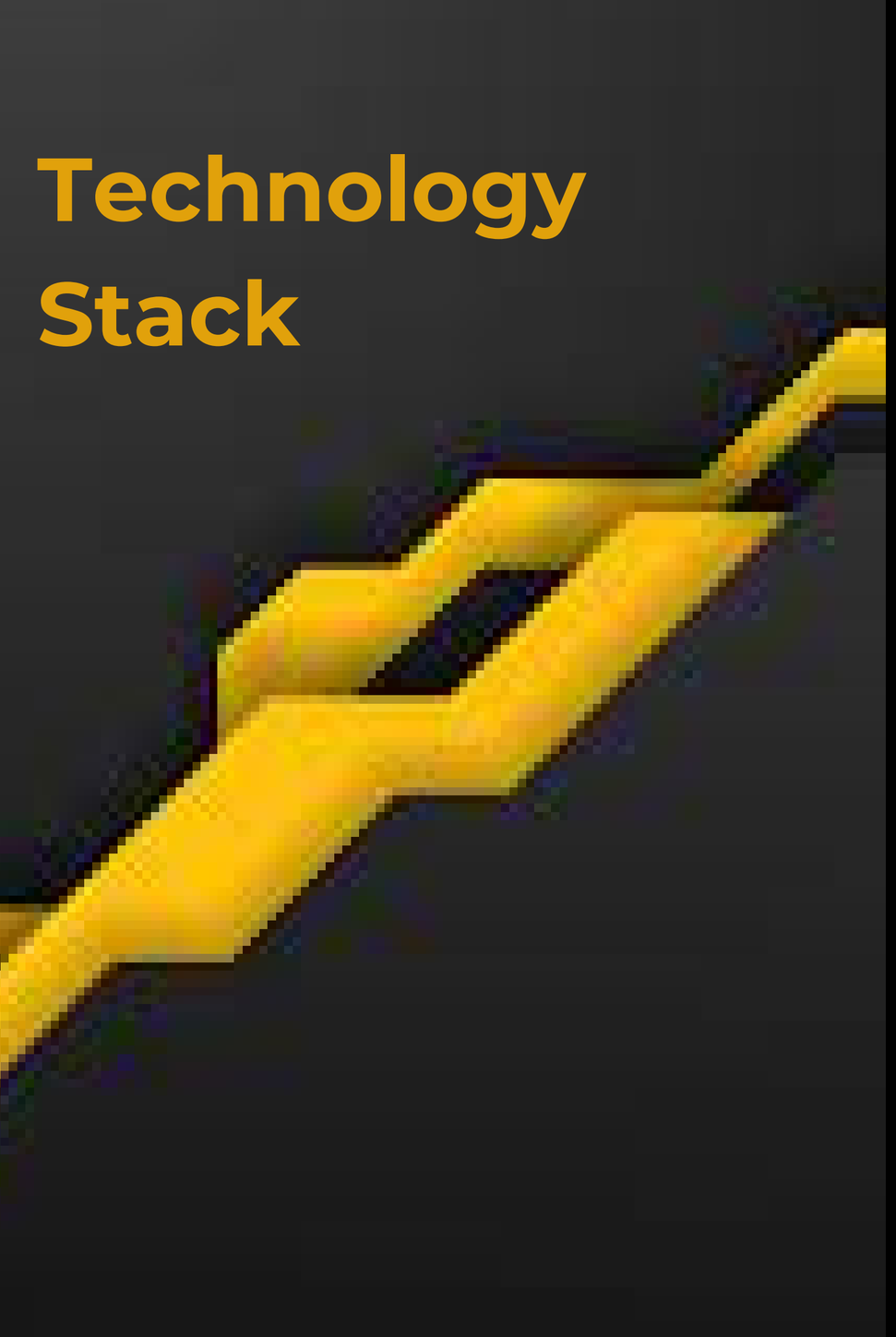
For URLs not found in host lists, the system utilizes a pre-trained text classification model from Hugging Face Transformers for dynamic analysis.

Step

Classification Results

The classification results, including labels such as phishing, malware, benign, or defacement along with corresponding scores, are displayed to the user.

Technology Stack



- Flask for Web Application that allows users to input URLs and receive classification results.
- Transformers Library for ML Model, it provides a pre-trained text classification model for analyzing URLs.
- Python is used as primary programming language for scripting and developing the backend logic of the TrustLink project.
- JSON (JavaScript Object Notation) is used for handling data, particularly for loading and storing host lists and other relevant information.
- A fine-tuned version of microsoft/codebert-base, a pre-trained language model specifically designed for understanding and generating code, has been used.

Data sets collected from reputable websites.

- **Computer Emergency Response Team of the Republic of Turkey**

A curated dataset provided by the national Computer Emergency Response Team (CERT) of the Republic of Turkey, offering insights into local threats.

- **Openphish Phishing Intelligence List**

A dataset dedicated to identifying and combating phishing threats, contributed by Openphish's phishing intelligence efforts.

- **URLhaus by abuse.ch**

A dynamic and constantly updated repository of malicious URLs, actively curated by abuse.ch.

- **Dead Domains List**

Compilation of domains that are inactive or no longer in use, aiding in the identification of potentially abandoned or malicious entities.

Community-Driven Malicious Domains List by Hexxium Creations

A collaborative effort, driven by the cybersecurity community, to compile a comprehensive list of malicious domains.

Data sets collected from reputable websites.

- **Risk Hosts**
  Aggregated data pinpointing hosts associated with elevated cybersecurity risks, enhancing preemptive security measures.

- **Cyber Threat Intelligence Feed by rescure.me**
  A rich source of cyber threat intelligence providing real-time insights into evolving digital threats.

- **Anti-Malware List**
  A comprehensive list of domains actively combating malware, serving as a shield against known malicious entities.

- **List of Malware Domains**
  A curated registry specifically highlighting domains known for hosting and distributing malware.

- **List of Scam URLs by Global Anti-Scam Organization**
  A meticulously curated collection of URLs associated with various scams, contributed by the Global Anti-Scam Organization.

TrustLink Chrome Extension

Enhancement to our project

- Real-Time Threat Detection:
  - TrustLink operates in real-time, analyzing URLs on the fly to detect and warn against potential threats. Experience a heightened sense of security with every click.
- Phishing and Fraud Alerts:
  - With advanced algorithms at its core, TrustLink excels in identifying phishing attempts and fraudulent URLs. Instant alerts ensure you stay one step ahead of cyber scams.

FUTURE ASPECTS

Protection against Typosquatting attacks
Protections against IDN Homograph attacks

- In our future endeavors, we envision enriching our machine learning dataset with additional features, including comprehensive Whois information and the age of the website. By incorporating these insights, we aim to elevate our analysis capabilities, providing a more nuanced and sophisticated approach to predicting the trustworthiness of URLs. This expansion will empower our system to consider not only the current state of a website but also its historical context, enabling us to offer users a more refined and insightful assessment of potential threats.

THANK YOU