chrisheimbuch /
malicious-websites-research-report

Code Issues Pull requests Actions Projects Wiki Security Insights Setting

Flatiron School Data Science Bootcamp Phase 3 Project

0 stars 0 forks 1 watching 1 Branch 0 Tags Activity

Public repository

1 Branch 0 Tags Go to file t Go to file Add file Code

chrisheimbuch documents and additional edits 0c74cd7 · 1 minute ago

documents	documents and additional edits	1 minute ago
source	updates on explainations and styling	2 hours ago
.gitignore	adding start files	4 days ago
README.md	Update README.md	15 hours ago

README

Malicious & Benign Websites - Machine Learning Research Report

This is my Phase 3 project during my time at Flatiron School. It focuses on a detailed exploratory data analysis (EDA) and inferential analysis using Python to examine a dataset of websites labeled as benign or malicious. The dataset includes various features such as DNS, TCP, and IP packets captured from a honeypot that interacted with multiple servers worldwide to collect data.

In the first notebook, I performed the EDA, which can be found in notebook #1 (ending nb1). The second notebook delves into a comprehensive machine learning classification analysis where I trained models to predict whether a website is malicious or benign. The final model achieved a 98.4% accuracy.

This project demonstrates a practical, real-world application that can be deployed by individuals or organizations to safeguard against the ever-evolving threat of malicious web activity, such as spyware, malware, ransomware, and more.

- Documents: includes PDF documents of the inferential EDA, technical machine learning notebook, and slides used in class presentation.
- Source: includes notebook #1 inferential eda source code, notebook #2 machine learning source code, and associated files with the project.

Click here for the Dataset

Releases

No releases published
Create a new release

Packages

No packages published
Publish your first package

Languages

- **Jupyter Notebook** 100.0%