PHISHING WEBSITES IDENTIFIER

A SMART WAY TO IDENTIFY DIFFERENT PHISHING WEBSITES

PROBLEM STATEMENT

- Many online scams nowadays.
- Most dangerous among them are Social Engineering attacks. Phishing is one such attack.
- Phishing stands for a fraudulent process, where an attacker tries to obtain sensitive information from the victim.
- Phishing websites, which are nowadays in a considerable rise, have the same look as legitimate sites. However, their backend is designed to collect sensitive information that is entered by the victim.

OBJECTIVES

- Build a tool to smartly predict if a website is phishing or not.
- Build a predictions service for this tool.
- Build a Machine Learning Pipeline so that the models can be created and trained at runtime.
- Build a UI for this pipeline to enable user to track different pipeline runs.

SOLUTION PROPOSED

- Built the tool which largely uses Machine Learning to make predictions.
- Created an API to enable users to interact with this prediction service.
- Modular Machine Learning Pipeline built.
- Front end for both prediction service and pipelining tool created.
- Dynamically, model configurations can be changed from the front end.
- The tool capable of generating the best model among these Configurations.

RESULTS

- Reduction of False Negative rate to around 1% on testing data.
- Got a recall of approximately 0.97 for the model in production.
- Support for all Machine Learning models present.
- Final Result provided as “Phishing” or “Not Phishing”.
- Capable of providing probabilities instead of hard results.