PROJECT OBJECTIVE

Project Name	Web Phishing Detection
Team ID	PNT2022TMID32768

ABSTRACT

A web service is one of the most important Internet communications software services. Using fraudulent methods to get personal information is becoming increasingly widespread these days. However, it makes our lives easier, it leads to numerous security vulnerabilities to the Internet's private structure. Web phishing is just one of the many security risks that web services face. Phishing assaults are usually detected by experienced users however, security is a primary concern for system users who are unaware of such situations. Phishing is the act of portraying malicious web runners as genuine web runners to obtain sensitive information from the end-user. Phishing is currently regarded as one of the most dangerous threats to web security. Vicious Web sites significantly encourage Internet criminal activity and inhibit the growth of Web services. As a result, there has been a tremendous push to build a comprehensive solution to prevent users from accessing such websites. This report covers machine learning applied science using flask to detect fake URLs by extracting and analyze different features of legitimate and fake URLs. Logistic Regression and algorithm are used to detect fake websites. We perform detailed literature survey and proposed an approach to detect phishing websites by feature extraction and machine learning algorithm.

Problems:
1. Logistic Regression
2. Random Forest Regression/Classification
3. Decision Tree Regression/Classification
4. K-Nearest Neighbors
5. Support Vector Machine

Data pre-processing:
1. Handling the null values and categorical values.
2. Reading the dataset.
3. Identify the dependent and independent values.
4. Split the dataset into train and test sets.

Analysis the dataset through visualization:
1. Univariate analysis
2. Bivariate analysis
3. Multivariate analysis

Applying algorithm:
1. ML Algorithm

Build the web application:
1. Using flask
2. Javascript
3. HTML and CSS