Project Initialization and Planning Phase

Date	26 September 2024
Team ID	LTVIP2024TMID24838
Project Title	Detection of Phishing Websites from URLs
Maximum Marks	3 Marks

Project Proposal (Proposed Solution) template

This project focuses to detect phishing websites from URLs. By utilizing machine learning algorithms and data analytics, the system aims to analyze website URLs in real-time to identify and flag potential phishing attempts, thereby enhancing online security and protecting users from cyber threats.

Project Overview	
Objective	The primary objective of this project is detecting the credibility of a website based on its urls characteristics.
Scope	The project predicts whether a website is a legitimate or a phishing website from its url.
Problem Statement	
Description	When using a website URL, it's crucial to be cautious to avoid falling victim to phishing attacks. Phishing URLs often appear legitimate but can lead to malicious websites designed to steal sensitive information or install malware.
Impact	Detecting phishing websites is crucial for cyber safety as it helps prevent financial losses, identity theft, malware prevention and other malicious activities.
Proposed Solution	
Approach	Machine learning techniques play a crucial role in detecting phishing websites by enabling the identification of patterns and characteristics that distinguish phishing websites from legitimate ones. These techniques can analyze various features of a website, such as its URL, content, and user behavior, to predict whether it is a phishing website or not.
Key Features	-Implementing machine learning model such as Logistic Regression, , Random Forest or other models for detecting the websites.

	- Predicting the credibility of a website using its url.

Resource Requirements

Resource Type	Description	Specification/Allocation
Hardware		
Computing Resources	CPU/GPU specifications, number of cores	e.g., intel i5 / AMD Ryzen 5, 4 cores
Memory	RAM specifications	e.g., 8 GB
Storage	Disk space for data, models, and logs	e.g., 1 TB SSD
Software		
Frameworks	Python frameworks	e.g., Flask
Libraries	Additional libraries	e.g., scikit-learn, pandas, numpy, joblib, regex, tldextract, socket, bs4, whois, favicon, re, google
Development Environment	IDE, version control	e.g., Jupyter Notebook, Google colab
Data		
Data	Source, size, format	e.g., Kaggle dataset, 836 , csv