Project Design Phase-II
Data Flow Diagram & User Stories

Date	10 November 2022
Team ID	PNT2022TMID20907
Project Name	Project - Web Phising Detection
Maximum Marks	4 Marks

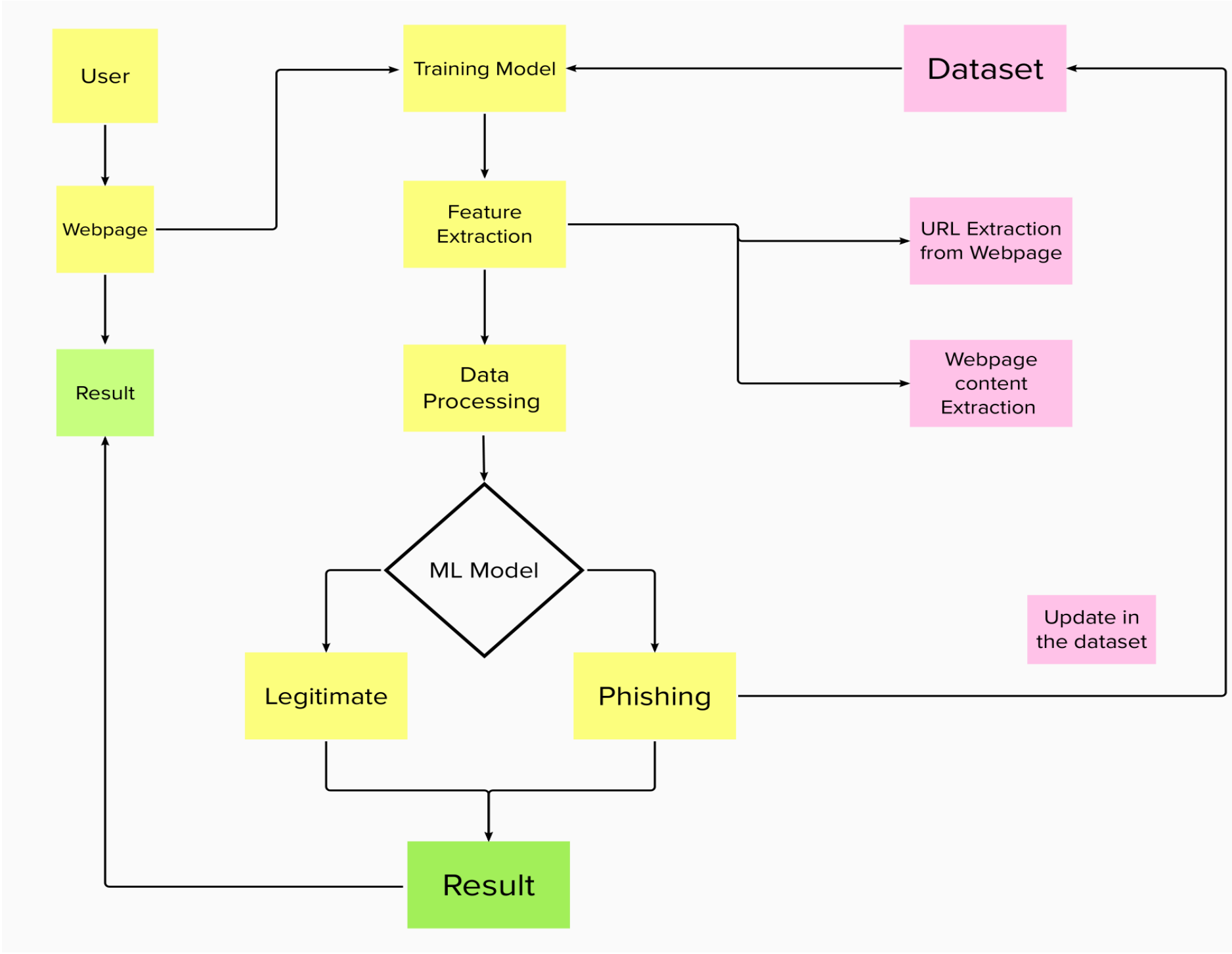
Data Flow Diagrams:

A popular and common method for cybersecurity threats is phishing URLs. Phishing is a type of cybercrime that aims to convince its victims to provide the attacker access to their private and sensitive information. The attacker wants to obtain personal information such user names, passwords, financial account information, information from social networking sites, and addresses.

Then, these confidential login information is frequently exploited for nefarious purposes including fraud, infamy, profit, reputation damage, and many other unlawful acts. This paper provides a thorough analysis of the various systems currently in use for phishing website detection. The technique described here makes use of advanced machine learning to classify webpages as phishing or starting with greater precision and accuracy.

Due to the anonymity offered by the internet and the rapid growth of online transactions, hackers try to trick end users by using techniques like phishing, SQL injection, malware, man-in-the-middle attacks, domain name system tunnelling, ransomware, web trojans, and so on.

Phishing is said to be the most misleading attack among all of these. Usually, the goal is to entice people to divulge sensitive data like system logins or financial information.


User
Training Model
Dataset
Webpage
Feature Extraction
URL Extraction from Webpage
Result
Data Processing
Webpage content Extraction
ML Model
Update in the dataset
Legitimate
Phishing
Result

User Stories

Use the below template to list all the user stories for the product.

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (Mobile user)	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	I can access my account / dashboard	High	Sprint-1
		USN-2	As a user, I will receive confirmation email once I have registered for the application	I can receive confirmation email & click confirm	High	Sprint-1
		USN-3	As a user, I can register for the application through Facebook	I can register & access the dashboard with Facebook Login	Low	Sprint-2
		USN-4	As a user, I can register for the application through Gmail		Medium	Sprint-1
	Login	USN-5	As a user, I can log into the application by entering email & password		High	Sprint-1
	Dashboard					
Customer (Web user)	User Input	USN-1	As a user, I can enter the required URL in the box while awaiting validation.	I can access the website without any problem	High	Sprint-1
Customer Care Executive	Feature Extraction	USN-1	In the event that nothing is discovered during comparison, we can extract features using a heuristic and a visual similarity technique.	As a user I can have comparison between websites for security	High	Sprint-1
Administrator	Prediction	USN-1	The model will use machine learning algorithms like a logistics regression and KNN to forecast the URLs of the websites.	I can accurately forecast the specific algorithms in this way.	High	Sprint-1
	Classifer	USN-2	To create the final product, I will now feed all of the model output to classifier.	I'll use this to identify the appropriate classifier for generating the outcome.	Medium	Sprint-2