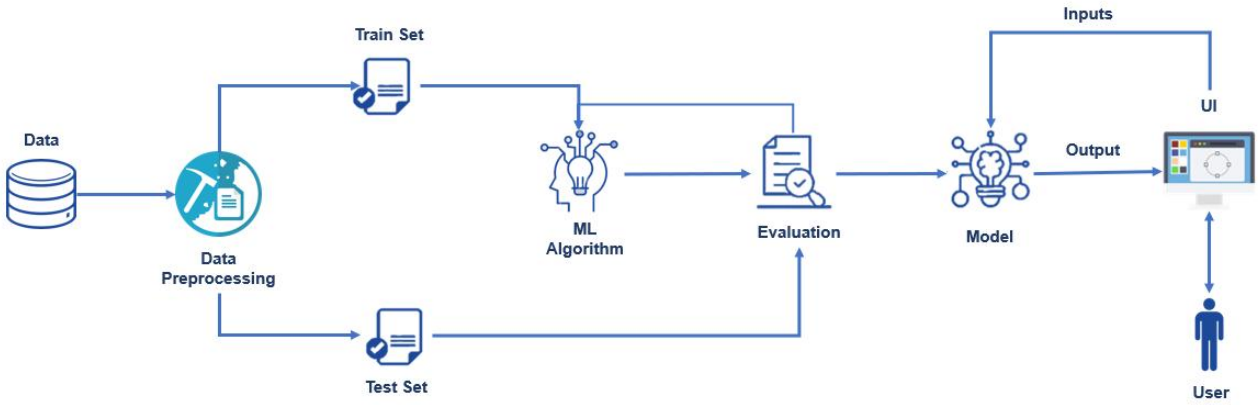
Project Design Phase -1

Solution Architecture

Date	29 September 2022
Team ID	PNT2022TMID49921
Project Name	Project - Web Phishing Detection
Maximum Marks	2 Marks

Architecture:



Google Safe Browsing:

This approach uses the blacklist URLs to discover the phishing attack. A sample URL is taken as input and checked within the blacklist repository. If the URL is present in the black list repository, the URL is termed as suspicious URL, else it is a legitimate website. The main shortcoming of this approach is its inability to detect the phishing URL which aren't present within the blacklist which could increase the false positives rate.

Spoof Guard:

This method scans suspicious websites for phishing symptoms to determine whether the website is legitimate or phishing. Some heuristics include image verification, link verification, URL verification and password field verification. If the total score of the phishing symptoms listed above exceeds the threshold, it is classified as a legitimate phishing website. This method detects zero-day attacks. This method also has a high limit on the number of false positives.

False alert:

This method will use visual phishing detection when the attacker uses the same CSS style to deceive the original website. In this method, CSS style

comparisons are performed on white listed websites with suspicious website styles to detect phishing. Attack

Solution

In a phishing attack, a user is sent a mail or a message that has a misleading URL, using which the attacker can collect important data like the passwords of the banks your money is in. This article gives a short tutorial on how to detect such phishing attempts.

Through phishing attacks, attackers acquire important credentials that can be used for getting access to your bank or other financial accounts. The URLs sent by the attacker look exactly the same as the original applications we use on a daily basis. That is why people often believe these and enter their personal details. A phishing URL can open a Web page that looks similar to the original login page of your bank. Detecting such URLs has become very important of late as such phishing attacks are becoming pretty common. So let’s see how we can check whether a URL is a misleading one or a genuine one using machine learning in Python, as it can help us see the code as well as the outputs. We will be using Jupyter Notebook. You can use Google Colab or Amazon Sagemaker too, if you are more comfortable with those.