Functional Features

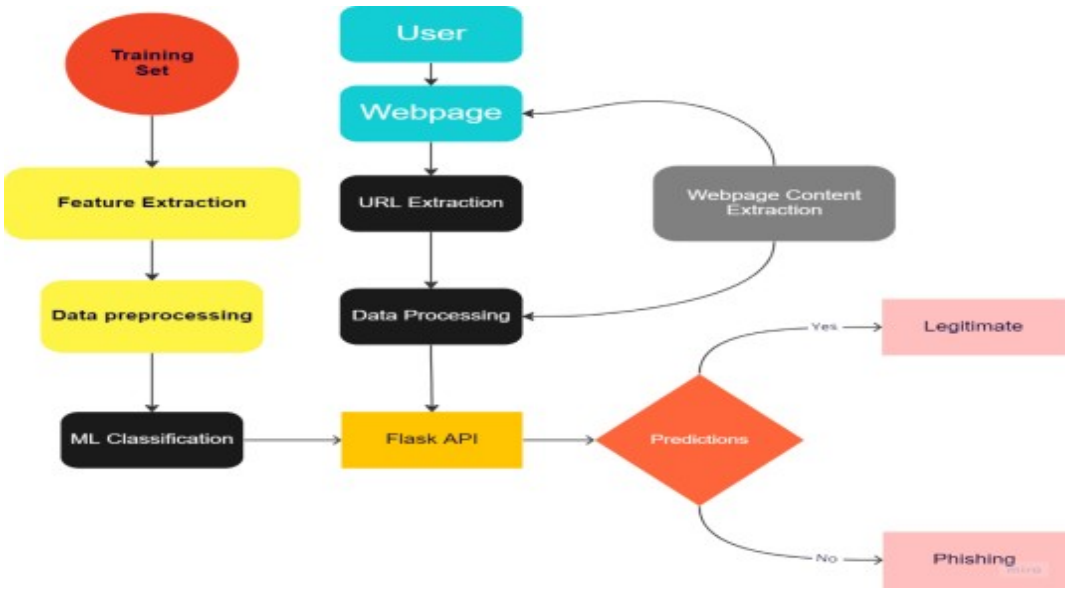
Introduction :

Due of how simple it is to develop a phoney website that closely resembles a legitimate website, phishing is now a top worry for security researchers. Although experts can spot fraudulent websites, not all users can, and as a result, some users fall prey to phishing scams. The attacker's primary goal is to obtain login information for bank accounts. This is the standard technique to identify phishing websites Internet Protocol (IP) addresses that have been blacklisted are added to the antivirus database using the "blacklist" technique. to avoid Attackers on blacklists craftily alter the URL to appear legitimate by obfuscation and many other straightforward ways, such as fast-flux, in which proxies are automatically built to host the website.

Data Analysis:

A method called data preprocessing is used to turn a raw data set into a clean data set. In other words, when data are gathered from various sources, they are gathered in raw form, which makes analysis impossible. It starts with the protocol used to visit the page is where it all starts. The server that is hosting the web page is identified by the fully qualified domain name. It consists of a top-level domain suffix and a registered domain name (second-level domain) (TLD). Due to the requirement that it be registered with a domain name

Registrar, the domain name portion is limited. A domain name plus a subdomain name make up a host name. The subdomain sections are completely in the control of the phisher, who is free to assign any value. The path and file components of the URL are also possible, and they may both be altered at anytime by the phisher. The subdomain name and route are entirely under the phisher's control.



Conclusion:

The purpose of this work was to investigate the feasibility of detecting phishing attacks by separating, using CNNs, the URL and pictures of phishing websites from the URL and photos of authentic websites. Newly produced phishing webpages may be identified using the method we presented, which is based just based on the URL and screenshot of dubious websites. 99.67% classification accuracy is displayed by the suggested model.A conclusion that may be drawn from the data is that integrating URLs characteristics and visual Convolution neural networks are a superior method for detecting similarity since they are more effective than several methods for accomplishing automated feature extraction and categorization distinguishing between authentic and fraudulent websites. Furthermore, it is obvious that growing batch sizes sizes causes the model's accuracy to decrease.We recommend finding a way to automatically identify the shortest URLs and the tiniest screenshot sizes for webpages in order to improve the model in future work and help the suggested method perform as well as possible.September 2020, Volume 12, Number 5 of the International Journal of Computer Networks & Communications (IJCNC)