Project Design Phase-I
Proposed Solution Template

Date	19 September 2022
Team ID	PNT2022TMID03377
Project Name	Project – WEB PHISHING DETECTION
Maximum Marks	2 Marks

Proposed Solution Template:

Project team shall fill the following information in proposed solution template.

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	Internet has dominated the world by dragging half of the world's population exponentially into the cyber world. With the booming of internet transactions, cybercrimes rapidly increased and with anonymity presented by the internet, Hackers attempt to trap the end-users through various forms such as phishing, SQL injection, malware, man-in-the-middle, domain name system tunnelling, ransom ware, web Trojan, and so on. Among all these attacks, phishing reports to be the most deceiving attack. Our main aim of this paper is classification of a phishing website with the aid of various machine learning techniques to achieve maximum accuracy and concise model.
2.	Idea / Solution description	Detection and prevention of phishing websites endure measure continuously a major space for analysis. There are different types of phishing techniques that offer torrential and essential ways that offer attackers to penetrate the data of people and organizations. Uniform resource locator URLs sometimes are also referred to as "Web links" play a vital role in a phishing attack. Uniform resource locator has a vulnerability of redirecting the pages i.e., through the hyperlink; which could redirect to the legitimate website or the phishing site. Different techniques in making phishing sites are emerging day by day. This actually motivated several researchers to put up their concentrate on finding the phishing sites.
3.	Novelty / Uniqueness	Microsoft. Microsoft Security Index Report.
4.	Social Impact / Customer Satisfaction	An exhaustive systematic search was performed on all the indexing databases. The state-of-the-art research related to the web phishing detections was collected. The papers were classified based on the methodologies. Taxonomy was derived by performing a deep scan on the classified papers. The contributions listed in this survey are exhaustive and lists all the state-of-the-art development in this area.
5.	Business Model (Revenue Model)	An exhaustive systematic search was performed on all the indexing databases. The state-of-the-art research related to the web phishing detections was collected. The papers were classified based on the methodologies. Taxonomy was derived by performing a deep scan on the classified papers. The contributions listed in this survey are exhaustive and lists all the state-of-the-art development in the area. A phishing scan starts with spreading bogus e-mail. After receiving an e-mail, ant phishing techniques start working, either by redirecting the phishing mail in the spam folder or by showing a warning when an online user clicks on the link of phishing URL. The lifecycle of phishing attackt in this area.
6.	Scalability of the Solution	The key notable points of our initial work embed: Phishing sites and their domains reveal the features that are different from other sites and domains. (For example, Google; www.google.com and some random phishing website be like; www.google.com).Phishing Uniform Resource Locators and 'domain names' typically have a different length when compared to other websites and domain names the training accuracy and testing accuracy of all the models. The difference between the values of train and test accuracy shows that the models are not over fitting over large dataset