

Project Design Phase-I
Proposed Solution Template

Date	13 October 2022
Team ID	PNT2022TMID01009
Project Name	Web Phishing Detection
Maximum Marks	2 Marks

Proposed Solution Template:

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

SL.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	For an average individual, it is very difficult to differentiate a phishing website from a legitimate website. It is not feasible to identify a fake website just by looking at them because they look similar to the original ones. This makes it hard to trust any websites. Users also share sensitive information like passwords, card details, etc in websites in order to utilize the websites. Presence of these fake makes us feel less secure. These phishing sites also affects the reputation of the original website and its organisation. Users feel frustrated, unsecured, confused, stressed, worried, annoyed, etc. which makes them give up.
2.	Idea / Solution description	The solution to the problems described is a Phishing Website Detection application which analyses the URL of the site and give report about the site. It will detect 30 different features related to the URL of the website like length, IP address, HTTPS token etc. We had tested with three algorithms like Logistic Regression, Decision tree, Random Forest Tree. Among them, Random Forest Tree has given best result with more accuracy. It is an application to secure users from phishing websites.
3.	Novelty / Uniqueness	It uses Random Forest Tree algorithm, which provided higher accuracy rate among others with concept of binary classification on the website's properties like IP address, HTTPS token, etc to identify the originality of the websites.

4.	Social Impact / Customer Satisfaction	Phishing sites created fear among users' mind about losing their sensitive data. It has major impact on user's mental health and their privacy. By identifying the fake certificates, the user can access and provide information without any fear. It makes them feel that their data was not leaked. It provides immense
		customer satisfaction as it is very user-friendly and is easily accessible. It is a place where user can find security of the website.
5.	Business Model (Revenue Model)	The revenue is generated usage of the application. More the user using the application the more the revenue generated. If successful, it can also be developed as a browser extension.
6.	Scalability of the Solution	This application is very scalable as it can identify fake websites and original websites using binary classification ML models.