Project Design Phase-I
Proposed Solution

Date	18 October 2022
Team ID	PNT2022TMID34650
Project Name	Project – Web Phishing Detection
Maximum Marks	2 Marks

Proposed Solution :

S .No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	Phishing attack is a simplest way to obtain sensitive information from the users. Aim of the phishers is to acquire critical information like username, password and bank account details. There are a number of users who purchase products online and make payments through e-banking. There are e-banking websites that ask users to provide sensitive data such as username, password & credit card details, etc., often for malicious reasons. This type of e-banking website is known as a phishing website. Web service is one of the key communications software services for the Internet. Web phishing is one of many security threats to web services on the Internet.
2.	Idea / Solution description	In order to detect and predict e-banking phishing websites, we proposed an intelligent, flexible and effective system that is based on using classification algorithms. Malicious URLs on the Internet can be easily identified by analysing it through Machine Learning (ML) technique. Once a user makes a transaction online when he makes payment through an e-banking website our system will use a data mining algorithm to detect whether the e-banking website is a phishing website or not. Recurrent Neural Network(RNN)—Long Short-Term Memory (LSTM) is one of the ML techniques that presents a solution for the complex real—time problems . LSTM allows RNN to store inputs for a larger period . It is similar to the concept of storage in computer. In addition, each feature will be processed according to the uniform distribution . The combination of RNN and LSTM enables us to extract a lot of information from a minimum set of data. Therefore, it supports a phishing detection system to identify a malicious site in a shorter duration.

3.	Novelty / Uniqueness	To detect malicious URLs and alert the users. We are applying ML techniques in order to analyse the real time URLs and produce effective results.
4.	Social Impact / Customer Satisfaction	Phishing is known as the process in which someone attempts to obtain sensitive information such as usernames, passwords, social security number or financial information and personal information such as birthdates, name and addresses by masking themselves as a trustworthy or familiar entity. Phishing is one of the top cyber-crimes that impact consumers and businesses all around the world. Phishing scam can be done not only by hackers, but by anyone with internet access.
5.	Business Model (Revenue Model)	Monitoring of the potential risk areas and an early detection of the hackers can significantly shorten the risk of hacking. With a high quality, well installed and well maintained web phishing system minimises the attacks.
6.	Scalability of the Solution	The below mentioned category of features are extracted from the URL data: • Address Bar based Features • Domain based Features • HTML & JavaScript based Features