**Project Design phase – 1**

|  |  |
| --- | --- |
| Date | 26 september |
| Team ID | PNT2022TMID02425 |
| Project Name | Project – web phishing Detection using Machine Learning |

|  |  |  |
| --- | --- | --- |
| S.No. | Parameter | Description |
| **1.** | **Problem Statement** | Phishing is a non-ethical method comprising both social engineering and technical tricks to capture user’s information and sensitive credentials like financial credentials. Some of the social engineering techniques use spam mails, pretending as a legitimate company or organization, that are specially designed to forefront users to knock-off websites that manoeuvre recipients to fall into the trap which steal financial credentials like user-ids and passwords. Technical intrigue methods install malicious software onto the systems, to capture the data directly, often using systems to intercept users online account user names and passwords. There are multiple websites who ask user to provide sensitive data such as username, password or credit card details etc. often for malicious reasons. Over the years 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, ransomware, web trojan, and so on. Among all these attacks, phishing reports to be the most deceiving attack. |
| **2.** | **Idea /Solution Description** | The model we propsoed is capable of detecting and predicting e-banking phishing websites effectively, using machine learning classification algorithms.The e-banking phishing website can be detected by considering some of the key factors like the domain identitiy, URL, security & envcryption criteria in the phishing detection rate. Inorder to get rid of phishing attacks one needs to change the passwords continously and use anti-phishing websites to prevent from attacks. And never share you personal / financial details over the internet. |
| **3.** | **Novelty Uniqueness** | Machine learning technology consists of many algorithms which requires past data to make a decision or prediction of future data. Using this technique, algorithm will analyze various backlisted and legitimate URL’s and their features to accurately detect the phishing websites including zero-hour phishing websites.for detection of phishing websites we use a set of features like the domain identity,length of the URL and many more and also to make this software computational free we are using SAAS and further planning to make our site to be available in an extension format with UI/UX designing. |
| **4.** | **Social Impact / Customer Satisfaction** | Phishing websites has a lot of effects interms of property loss,data loss,privacy loss, damage of reputation and disruption of operational activities.some of the companies like jp moragan, sony pictures effected by data breach due to phishing .   * JPMorgan Chase experienced , one of the most significant phishing breaches in history. In 2014, the company announced that the contact information for 76 million households and seven million businesses were compromised in the massive attack. * Using phishing and spearphishing emails, which contained malware, the attackers gained access to Sony’s network and performed months of covert reconnaissance. This attack lead to a cost of $100 million.   Customer satisfaction: By using our web phishing detection website the user can check their websites by copy and paste the phising URL. After knowing the result they can be completely safe from above mentioned impacts. |
| **5.** | **Business Model (Revenue Model)** | As long as phishing websites continue to operate/emerge, many more people and companies will suffer privacy leaks and data breaches or financial loses/intellectual losses. However, the existing phishing detection method do not fully analyze the features of phishing and the performance and efficiency of the models only apply to certain limited datasets and upto a set of 10 features. further need to be improved to be applied to the real web environment. |
| **6.** | **Scalibility of the Solution** | There are vast no. of algorithms that has been in existence which provides wide range of features to identify and analysis web phishing. The concept of phishing has become quite common now such that it can be adapted to any logical approach conducted. This tool we proposed gives high performance and optimization. All sorts of web application and ease of preventing users from scam. |