Ideation Phase
Literature Survey

Date	30 September 2022
Team ID	PNT2022TMID52841
Project Name	Project – Web Phishing Detection
Maximum Marks	

*Detection of Phishing Websites using Machine Learning et al.,*is an interactive and responsive website that will be used to detect whether a website is legitimate or phishing. The website is made using different web designing languages which include HTML, CSS, Javascript and Django. The basic structure of the website is made with the help of HTML. CSS is used to add effects to the website and make it more attractive and user-friendly. The website is created for all users, it must be easy to operate with no user should face any difficulty while making its use.The unstructured data of URLs from Phishtank website, Kaggle website and Alexa website are collected. In pre-processing, feature generation is done where nine features are generated from unstructured data. These features are length of an URL, URL has HTTP, URL has suspicious character, prefix/suffix, number of dots, number of slashes, URL has phishing term, length of subdomain, URL contains IP address. An organized dataset is made in which each detail incorporates the paired (0,1) which is then passed to the various classifiers.The three unique classifiers and analyse their presentation based on exactness two classifiers utilized are Decision Tree and Random Forest algorithm. The exactness of various classifiers and discovered Random Forest as the best classifiers which gives the most extreme precision. If the URL entered by a user is found to be a phishing website, a small pop-up will appear on the screen to warn the user regarding this malicious website.

Web Phishing Detection using Machine Learning et al., the current circumstance is that the population's maturity has been wisecracked, causing them to unknowingly give their private information to hackers. Several banned websites have already been established to seem like that of an actual point of contact through obtaining stoners' private information. Passcode, savings account, and shipping information are just a few examples. Late in 2016, the amount of hacking activities was at an all-time high since the company started monitoring this in 2004. The overall identified phishing attacks in 2016 were 1,609. This represents a 65 percent increase over 2015. Within the final quarter of 2004, there would be scamming attempts each month. Machine Learning was used to find the phishing website. The use of machine literacy to surround the supplied features is the basis of Grounded Malware Monitoring Systems. Features are generated by assembling items in a specific order, such as URLs, sphere names, website features, and website content.

Particle Swarm Optimization Based Feature Weighting for Improving Intelligent Phishing Website Detection et al., is mainly consists of PSO based feature weighting.The website features were weighted with the ideal weights by using PSO to enhance the detection of phishing websites. BPNN, SVM, NB, C4.5, RF, and kNN were trained based on the training dataset of features weighted by PSO in order to identify the phishing websites. The classification accuracies of BPNN, SVM, NB, C4.5, RF and kNN were enhanced after applying the proposed PSO based feature weighting. BPNN, SVM, C4.5, RF, and kNN that improved with the PSO based feature weighting achieved better TPR, TNR, FPR, and FNR compared to other stand alone machine learning models. The machine learning models improved with the proposed PSO based feature weighting were able to successfully detect and classify both phishing and legitimate websites. The PSO based feature weighting outperformed these machine learning models with applying IG, Chi-square, Wrapper, GA based features selection, and GA based features weighting.The machine learning models improved by the proposed PSO based feature weighting can be used as alternative solutions to effectively detect phishing websites in order to contribute to providing more confidence for customers of online commerce and business.The use of the most important features set to represent the website can be utilized to speed up the detection process of the phishing website. A faster and improved version of PSO can be used to speed up the performance of the proposed PSO based feature weighting. The PSO based feature weighting with ensemble learning and fusion approaches can produce promising solutions with a higher detection accuracy of phishing websites.

Detecting Phishing Website using Machine Learning et al., the current situation that is majority of the population has been fooled into giving their personal details to hackers without noticing it. Many blacklisted website has been publish to appear as an original site in order to trap user by asking them to input their personal details. For example, password, bank account, email address and etc. Phishing activity in early 2016 was the highest ever recorded since it began monitoring in 2004.The total number of phishing attacks in 2016 was 1,220,523. This was a 65 percent increase over 2015. In the fourth quarter of 2004, there were 1,609 phishing attacks per month. In the fourth quarter of 2016, there was an average of 92,564 phishing attacks per month, an increase of 5,753% over twelve years. According to the Anti-Phishing Working Group (APWG), there are at least 47, 324 phishing attacks and a top-ten American bank estimates that at least US$300 is lost for every hour that a phishing site remains up. Machine learning is that the science of obtaining computers to act while not being expressly programmed. Machine Learning was implement to develop this proposed system. Machine learning techniques identifies phishing URLs typically assess a URL based on some feature or set of features extracted from it.Thus, before coming to conclusion that this was the major problem, related products were examined and compared view their libation before progressing to the proposed project.

Ideation Phase
Define the Problem Statements

Date	30 September 2022
Team ID	PNT2022TMID52841
Project Name	Project – Web Phishing Detection
Maximum Marks	2 Marks