Ideation Phase

Problem Statements

Date	30 October 2022
Team ID	PNT2022TMID36919
Project Name	Web Phishing Detection

Phishing detection techniques do suffer low detection accuracy and high false alarm especially when novel phishing approaches are introduced. Besides, the most common technique used, blacklist-based method is inefficient in responding to emanating phishing attacks since registering new domain has become easier, no comprehensive blacklist can ensure a perfect up-to-date database. Furthermore, page content inspection has been used by some strategies to overcome the false negative problems and complement the vulnerabilities of the stale lists. Moreover, page content inspection algorithms each have different approach to phishing website detection with varying degrees of accuracy. Therefore, ensemble can be seen to be a better solution as it can combine the similarity in accuracy and different error-detection rate properties in selected algorithms. Therefore, this study will address a couple of research:

1. How to process raw dataset for phishing detection?

2. How to increase detection rate in phishing websites algorithms?

3. How to reduce false negative rate in phishing websites algorithm?

4. What are the best compositions of classifiers that can give a good detection rate of phishing website?

Problem Statement (PS)	I am (Customer)	I'm trying to	But	Because	Which makes me feel
PS-1	Banker	Customer retention	The details are from scammer	Appearance Of hoax emails	Unsafe about my information
PS-2	Student	Finding job in website	Receiving scam mail often	Due to unauthorized link	Developing of individual insecurity