Project Title: Web phishing Detection

Team ID: PNT2022TMID20741

Define CS, fit into CC

1. CUSTOMER SEGMENT(S) — CS

- Used in Web Browsers
- Banking Websites
- Military base systems
- Handheld Applications
- Defense and Air force

6. CUSTOMER CONSTRAINTS — CC

- Cyber Security
- Accuracy
- Ease to Access
- Cyber Awareness

5. AVAILABLE SOLUTIONS — AS

- By using natural language processing in MATLAB can give the result accuracy of 95%
- By applying Bayesian network , Stochastic Gradient Descent, Lazy K Star , Logistic model tree and Multilayer Perception in MATLAB/WEKP can provide an accuracy over 95% to 98%

Explore AS, differentiate

Focus on J&P, tap into BE, understand RC

2. JOBS-TO-BE-DONE / PROBLEMS — J&P

To Train the dataset and test it over multiple test cases and predict the accuracy of the result and to build the model in website and cloud. Adding Anti phishing extension in browsers can make an alert to the users who are in dangerous website.

9. PROBLEM ROOT CAUSE — RC

- We Humans could not able to predict when attack can occur.
- Not only in websites, even in banking sectors and defense systems can't able to predict the attack.
- To solve all these problems this technique / solution has developed.

7. BEHAVIOUR — BE

- Developing the efficient application which can able to prevent from any unauthorized means of activity.
- Any individual can gain knowledge about the issue and this system/model can teach how to get cautious when an attack can occur.

Focus on J&P, tap into BE, understand

Identify strong TR & EM

3. TRIGGERS — TR

- Better Accuracy than other Models
- Feasible UI and UX

10. YOUR SOLUTION — SL

- We use Decision Tree , Random Forest , Gradient Boosting algorithm using Python.
- Training and Testing the models with multiple datasets to overcome the accuracy level from existing algorithms.
- Build the model using python flask and host in web application using IBM cloud.

8. CHANNELS of BEHAVIOUR — CH

8.1 ONLINE

In online we can surf any website by adding the extension of anti phishing so that we can be precautious.

8.2 OFFLINE

This is an online platform but in offline we can create an awareness at every public sectors.

4. EMOTIONS: BEFORE / AFTER — EM

- While training multiple datasets the memory efficiency is more so that it was trained in external SSD with high throughput.
- Time is consumed more on predicting the single dataset.

Extract online & offline CH of BE