Project Title: Web phishing Detection

Purpose / Vision

Team ID: PNT2022TMID23752

Define CS, fit into CC

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

Identify strong TR & EM

Explore AS, differentiate

Focus on J&P, tap into BE, understar

Extract online & offline CH of BE

1. CUSTOMER SEGMENT(S)

Who is your customer?
i.e. working parents of 0-5 y.o. kids

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

6. CUSTOMER CONSTRAINTS

What constraints prevent your customers from taking action or limit their choices of solutions? i.e. spending power, budget, no cash, network connection, available devicor need to get the job done?

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

5. AVAILABLE SOLUTIONS

Which solutions are available to the customers when they face the problem or what have they tried in the past? What pros & cons do these solutions have? i.e. pen and paper is an alternative to digital notetaking

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

2. JOBS-TO-BE-DONE / PROBLEMS

Which jobs-to-be-done (or problems) do you address for your customers? There could be more than one; explore different sides.

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

What is the real reason that this problem exists?
What is the back story behind the need to do this job?
i.e. customers have to do it because of the change in regulations.

- **We Humans could not able to predict when attack to can occur.**
- **Not only in websites, even in banking sectors and defense systems can't able to predict the attack. when an attack can occur.**

7. BEHAVIOUR

What does your customer do to address the problem and get the job done?
i.e. directly related: find the right solar panel installer, calculate usage and benefits;
indirectly associated: customers spend free time on volunteering work (i.e. Greenpeace)

- **Developing the efficient application which can able 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**

• To solve all these problems this technique / solution has developed.

3. TRIGGERS

What triggers customers to act? i.e. seeing their neighbour installing solar panels, reading about a more efficient solution in the news.

- **Better Accuracy than other Models**
- **Feasible UI and UX**

4. EMOTIONS: BEFORE / AFTER

How do customers feel when they face a problem or a job and afterwards?

10. YOUR SOLUTION

If you are working on an existing business, write down your current solution first, fill in the canvas, and check how much it fits reality.

- **We use Decision Tree , Random Forest , Gradient Boosting algorithm using Python.**

If you are working on a new business proposition, then keep it blank until you fill in the canvas and come up with a solution that fits within customer limitations,

- **Training and Testing the models with multiple datasets to overcome the accuracy level from existing algorithms.**

8. CHANNELS of BEHAVIOUR

8.1 ONLINE

What kind of actions do customers take online? Extract online channels from #7

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

8.2 OFFLINE

What kind of actions do customers take offline? Extract offline channels from #7 and use them for customer development.

i.e. lost, insecure > confident, in control - use it in your communication strategy & design.

• Build the model using python flask and host in web application using IBM cloud.

• While training multiple datasets the memory so that it was trained in external

• Time is consumed more on predicting the single dataset.

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

efficiency is more SSD with high throughput.