**Detecting Phishing Websites**

**Project Description:**

**Web Phishing:**

There are a number of users who purchase products online and make payment through e- banking. There are e-banking websites who ask users to provide sensitive data such as username, password & credit card details etc., often for malicious reasons. This type of e-banking website is known as a phishing website. Web service is one of the key communications software services for the Internet. Web phishing is one of many security threats to web services on the Internet.

**Common threats of web phishing:**

* Web phishing aims to steal private information, such as usernames, passwords, and credit card details, by way of impersonating a legitimate entity.
* It will lead to information disclosure and property damage.
* Large organizations may get trapped in different kinds of scams.

**Solution Proposed:**

This Project mainly focuses on applying a machine learning algorithm to detect Phishing websites.

In order to detect and predict e-banking phishing websites, we proposed an intelligent, flexible and effective system that is based on using classification algorithms. We implemented classification algorithms and techniques to extract the phishing datasets criteria to classify their legitimacy. The e-banking phishing website can be detected based on some important characteristics like URL and domain identity, and security and encryption criteria in the final phishing detection rate. Once a user makes a transaction online when he makes payment through an e-banking website our system will use a data mining algorithm to detect whether the e-banking website is a phishing website or not.

**Pre-requisites:**

1. **In order to develop this project we need to install following softwares/packages:**

**Anaconda Navigator :**

Anaconda Navigator is a free and open-source distribution of the Python and R programming languages for data science and machine learning related applications. It can be installed on Windows, Linux, and macOS.Conda is an open-source, cross-platform, package management system. Anaconda comes with so very nice tools like JupyterLab, Jupyter Notebook,

QtConsole, Spyder, Glueviz, Orange, Rstudio, Visual Studio Code. For this project, we will be using Jupyter notebook and Spyder

1. **To build Machine learning models you must require the following packages**

**Sklearn:** Scikit-learn is a library in Python that provides many unsupervised and supervised learning algorithms.

**NumPy:** NumPy is a Python package that stands for 'Numerical Python'. It is the core library for scientific computing, which contains a powerful n-dimensional array object

**Pandas:** pandas is a fast, powerful, flexible, and easy to use open source data analysis and manipulation tool,built on top of the Python programming language[.](https://pandas.pydata.org/)

**Matplotlib:** It provides an object-oriented API for embedding plots into applications using general-purpose GUI toolkits

**Flask:** Web framework used for building Web applications

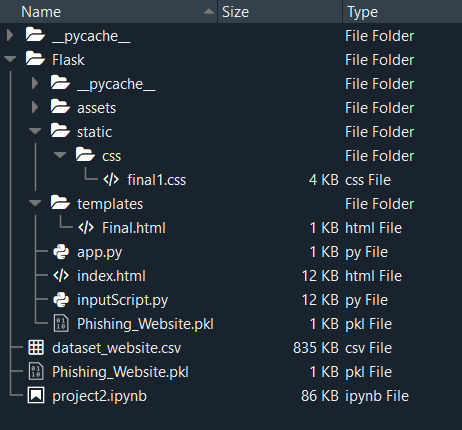
If you are using **anaconda navigator**, follow below steps to download required packages:

* Open anaconda prompt.
* Type “pip install numpy” and click enter.
* Type “pip install pandas” and click enter.
* Type “pip install matplotlib” and click enter.
* Type “pip install scikit-learn” and click enter.
* Type “pip install Flask” and click enter.

**Project folder Structure:**

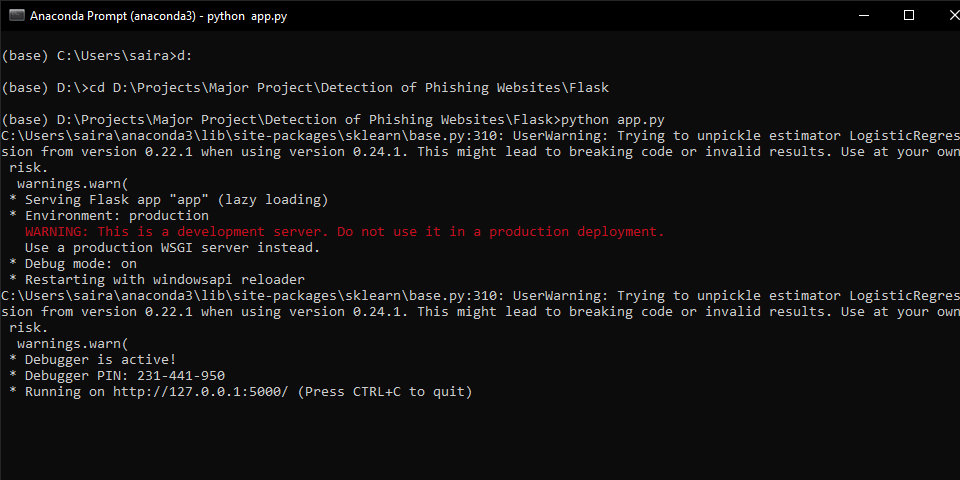
**Let us introduce you to the main project folder downloaded by you in prerequisites.**

* Flask folder contains two sub-folders static and templates. Static folder contains the style sheet used.
* Templates folder has the HTML pages.
* app.py is the python script for server side computing.
* index.html is the home page which should be used for initiating the application.
* inputScript.py has all the parameters of evaluation for a URL.
* Phishing\_Website.pkl is the model file which you have to build.
* dataset\_website.csv is the dataset
* project2.ipynb is the training notebook.

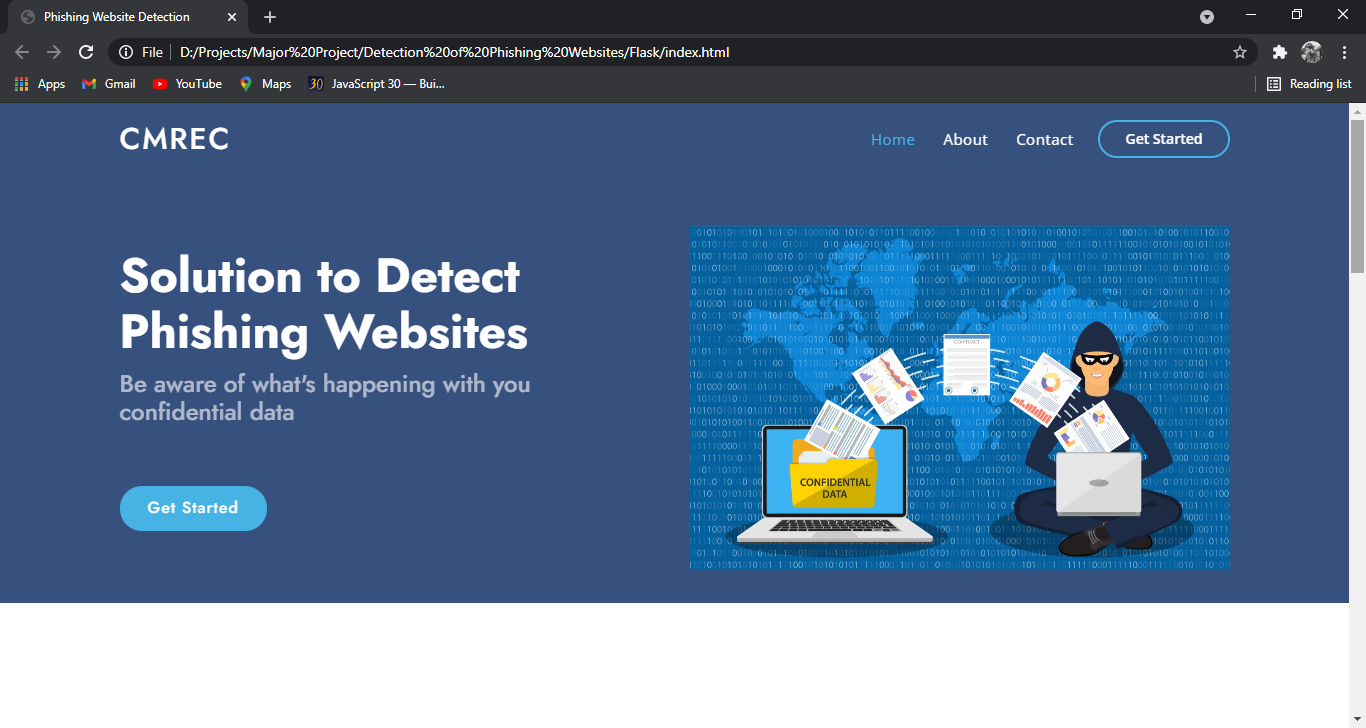


**Execute and test your model**

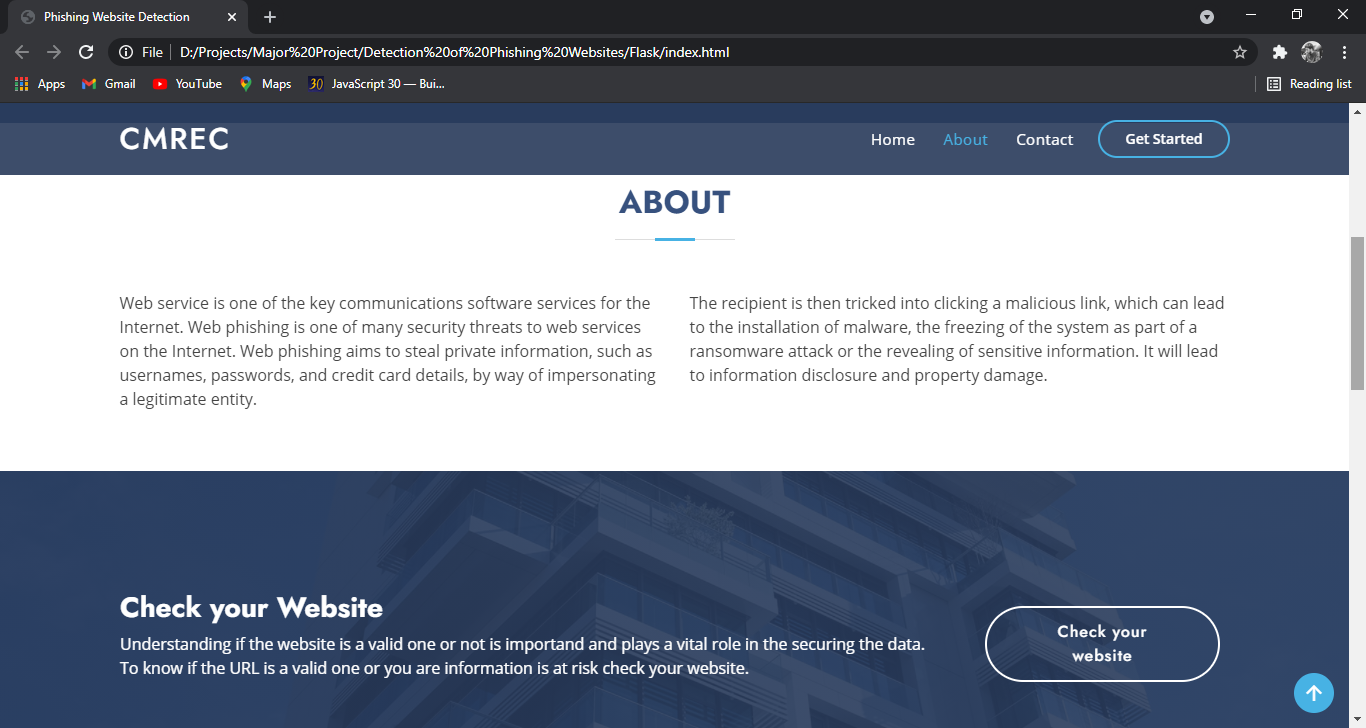
**Step 1: Execute the python code and after the module is running, open index.html page and scroll down to find the buttons to test with.**

****

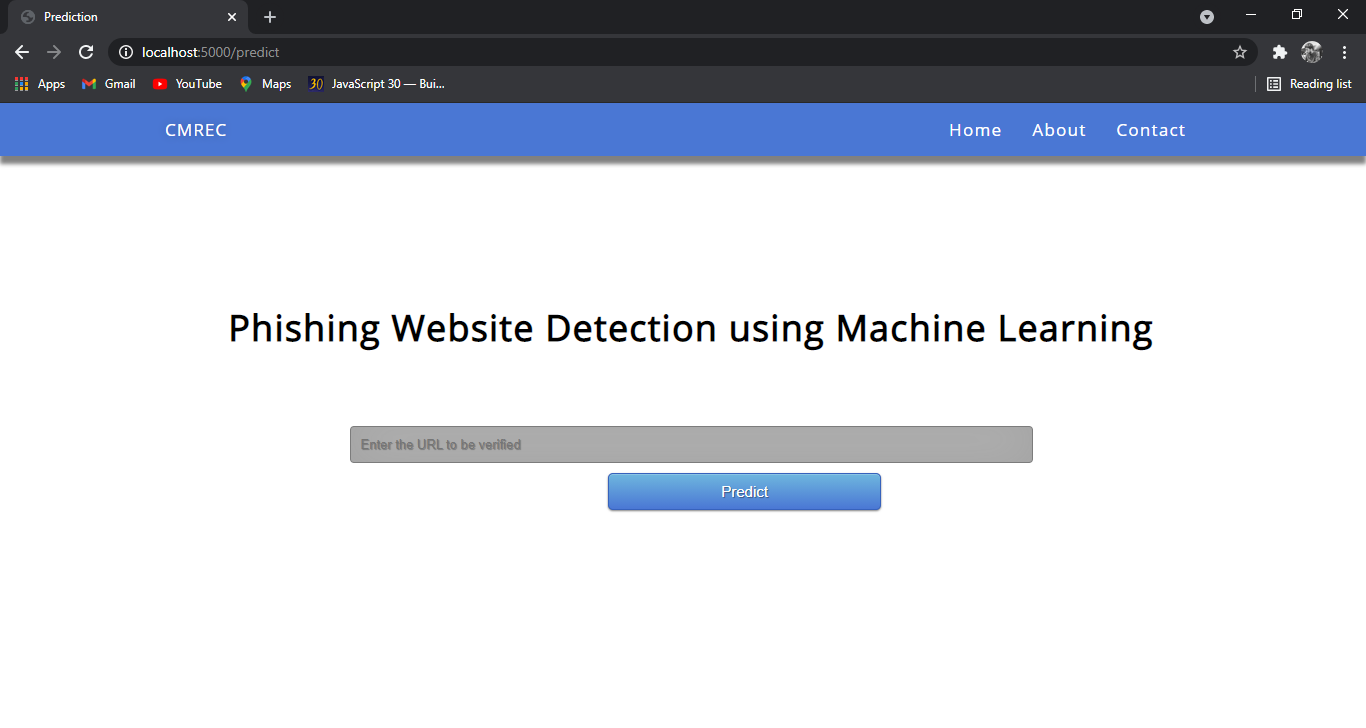
This is the home page of the web application(index.html)



About the project section which gives insights about the project.



When clicked on “Check your website” button, the user will be redirected to the below page where user can specify the URL.



The URL is given, the model analyses and gives the output whether it is a phishing or legitimate website.

