National College of Ireland

Configuration Manual

MSc Research Project
Cybersecurity

Sandhya Rathore
Student ID: x20199805

School of Computing
National College of Ireland

Supervisor: Liam McCabe

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name: Sandhya Rathore

Student ID: X20199805

Programme: MSc in Cybersecurity **Year:** 2021

Module: Research in Computing

Lecturer: Liam McCabe

Submission Due Date: 16 Dec 2021

Project Title: Can a machine learning algorithm, blacklisting & whitelisting approach improve accuracy & efficiency of detecting Phishing Websites?

Word Count: **Page Count:**

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: SandhyaRathore.....

Date: .....15/12/2021.....

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Configuration Manual

Sandhya Rathore
x20199805

1-Introduction

In this manual, GUI is the chrome extension which is used to detect malicious websites by doing feature extraction from the URLs. This Document provides details on setting up the application like Pre-requisite, Configuration steps and resources needed for the application to work. The steps are also shown with the help of screenshots.

1 Pre-requisite

This section discusses about the pre-requisite required to run the program.

2.1 Software Requirement

Software configuration is the bare minimum for a program to function properly. It should be set up as follows:

- Google Chrome
- Python 3.8
- Visual Studio Code/ Jupyter Notebook
- Windows

2 Configuration

This part of the report covers stepwise installation and configuration method for running the application:

3.1 Python installation part

The project utilizes Python3.8. It can be installed from Python Release Python 3.8.3 | Python.org. After installing the Python version, we need to update the Class path in Windows. By Default, the class path should be directed to C:\python38, C:\python38Scripts.

3.2 Visual Studio installation part

For running the Python files, visual studio code will be required. With that Anaconda also need to be installed.

Anaconda installation link: https://docs.anaconda.com/anaconda/install/

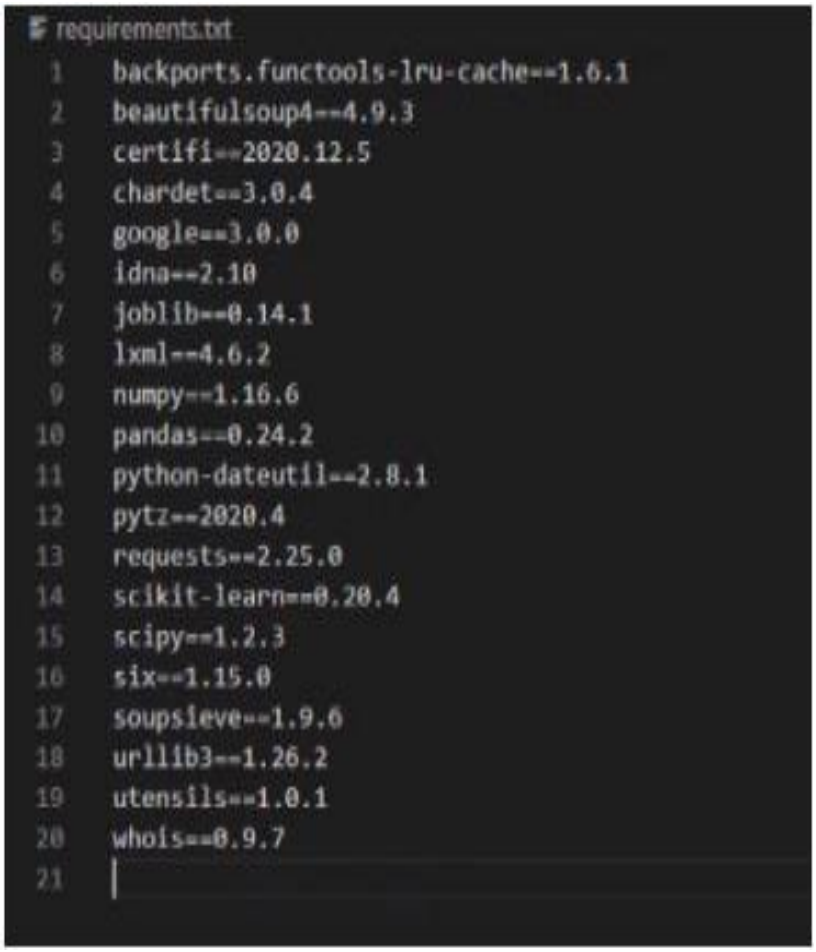
3.3 Jupyter Notebook Installation part

This is optional, you can install it to debug the code properly and also represent the output for the classifiers used in terms of accuracy, precision and also for comparing the results.

3.4 Installing Dependencies

We're ready to install Python dependencies to run the application now that we've set up the above software. The steps for doing so are as follows:

Using pip command, install the packages present in the requirements.txt file.



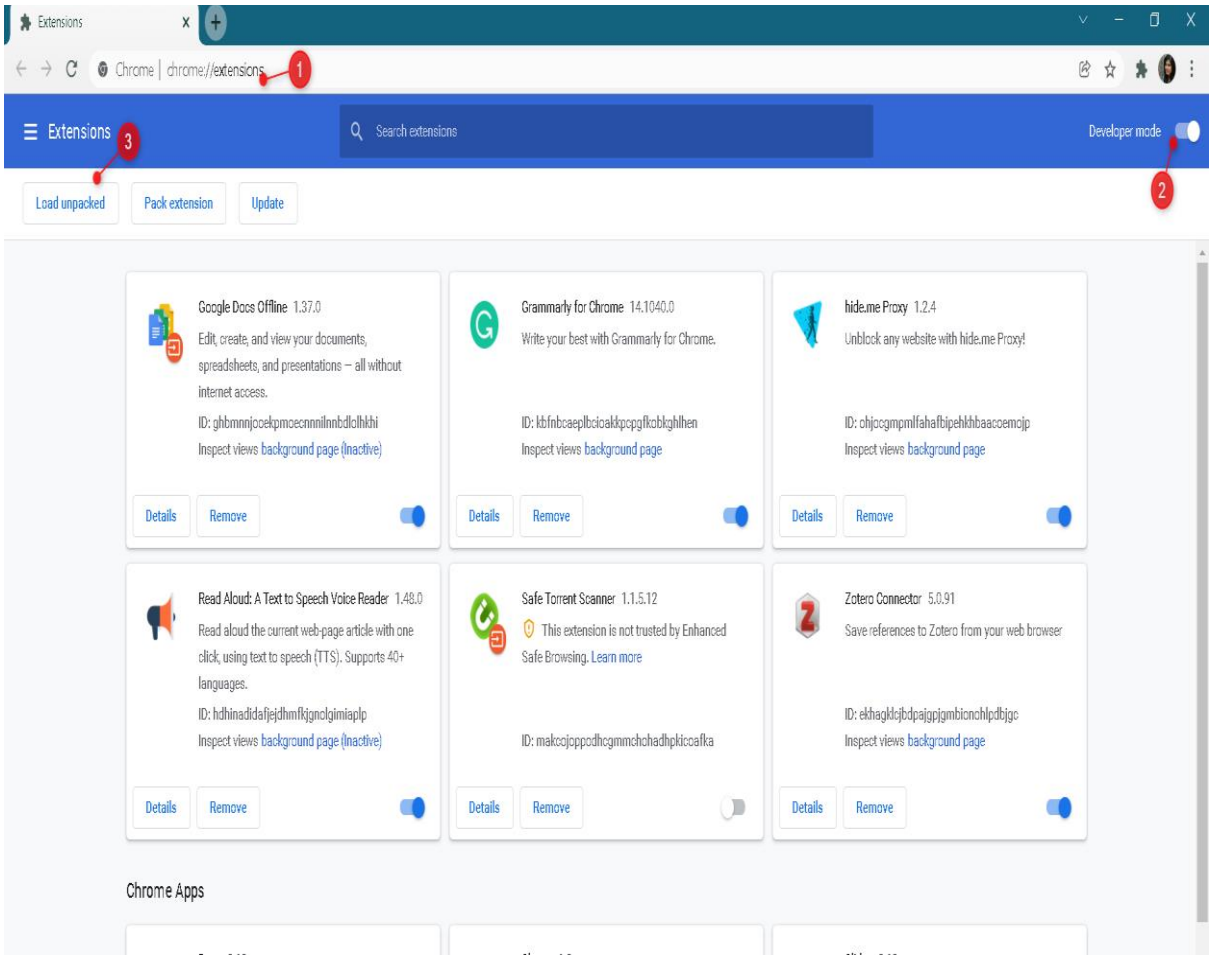
4. Setting up GUI

Below are the steps given to setup the Extension Folder on your Chrome browser which will identify the URLs as Safe or malicious:

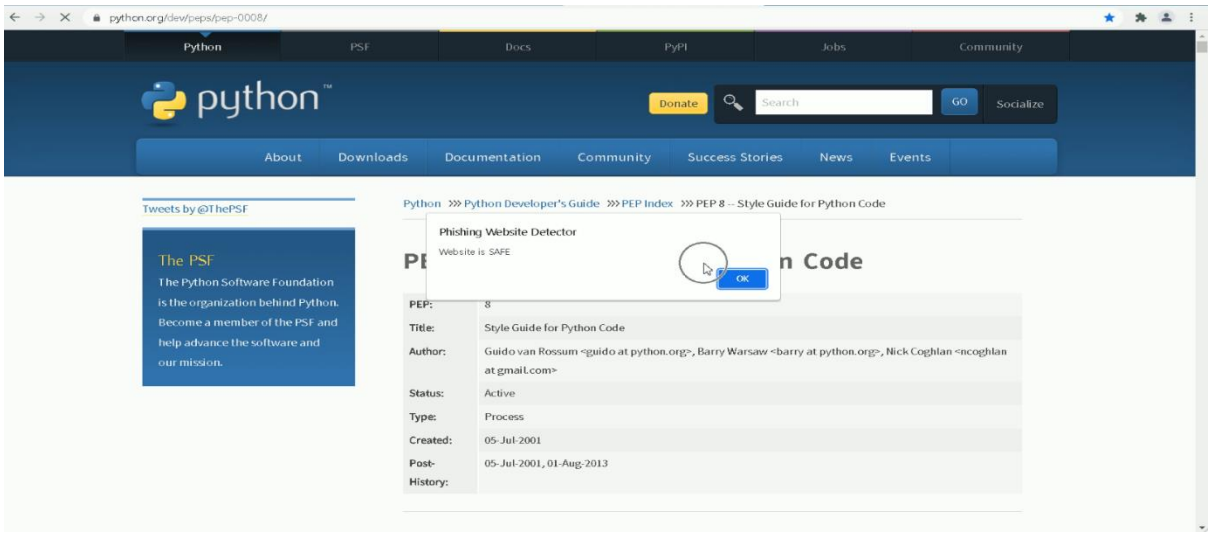
Step 1: Go to chrome extensions

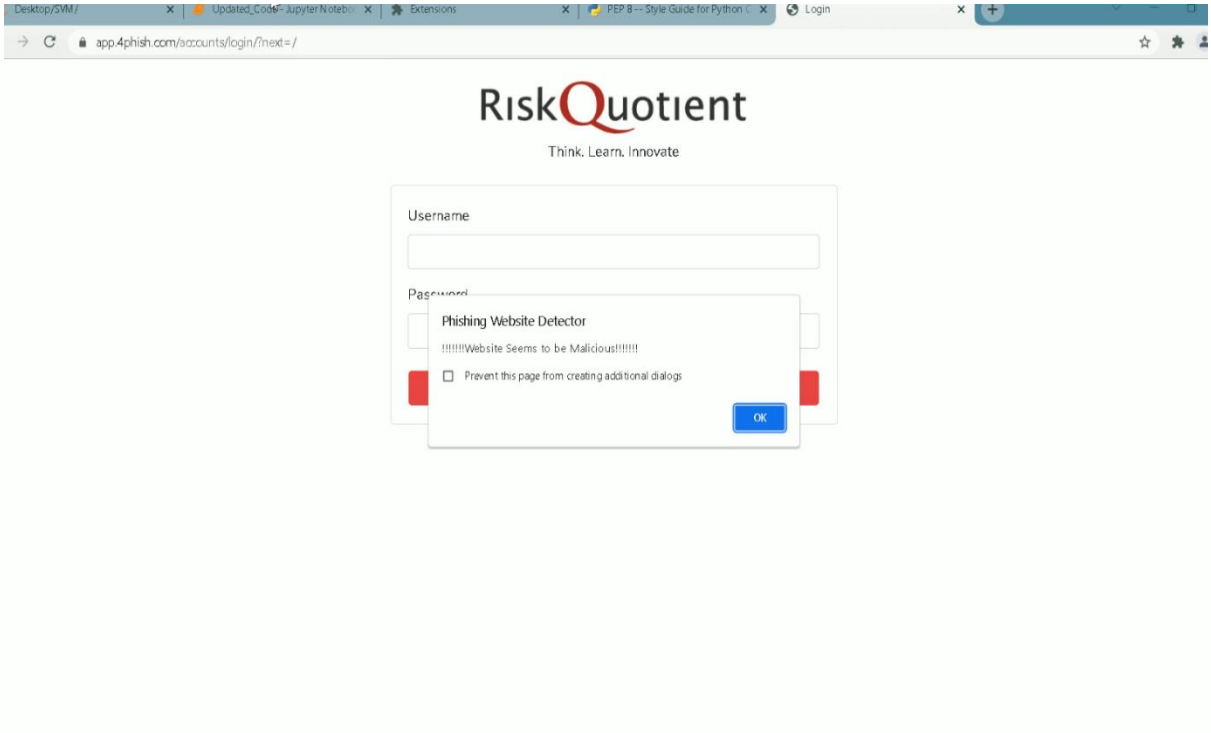
Step 2: Activate the “Developer mode”

Step 3: Click on the Load Unpacked option, browse and select the extension folder.



Step 4: Now, as you visit any URL. The extension will detect the same display if it is a safe or malicious URL.





5. MODEL OUTPUT

We have added all the python dependencies and now we will be able to run the .ipynb file: The. ipynb file will be comparing all the classifiers (Random Forest, Neural Networks, SVM). The below is the output we have received.

```
running random forests...
C:\Users\Sandhya\AppData\Local\Temp/ipykernel_4612/293555000.py:67: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
  model.fit(X_train,y_train)
accuracy = 89.63%
[[1293  162]
 [ 182 1680]]
(2, 2)
TP      FP      FN      TN      Sensitivity     Specificity
1293.0  182.0   162.0   1680.0
        0.89            0.9
1680.0  162.0   182.0   1293.0
        0.9             0.89
0.9071274298056156
runtime = 2.544670343399048 seconds
```

```
running support vector machines...
C:\Users\Sandhya\anaconda3\lib\site-packages\sklearn\utils\validation.py:63: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples, ), for example using ravel().
  return f(*args, **kwargs)
accuracy = 90.05%
[[1254  201]
 [ 129 1733]]
(2, 2)
TP      FP      FN      TN      Sensitivity     Specificity
1254.0  129.0   201.0   1733.0
        0.86            0.93
1733.0  201.0   129.0   1254.0
        0.93            0.86
0.9130663856691253
runtime = 6.011983394622803 seconds
runtime = 38.475990533828735 seconds
```

```
running neural networks...
C:\Users\Sandhya\anaconda3\lib\site-packages\sklearn\utils\validation.py:63: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples, ), for example using ravel().
  return f(*args, **kwargs)
C:\Users\Sandhya\anaconda3\lib\site-packages\sklearn\neural_network\_multilayer_perceptron.py:614: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (200) reached and the optimization hasn't converged yet.
  warnings.warn(
accuracy = 89.03%
[[1246  209]
 [ 155 1707]]
(2, 2)
TP      FP      FN      TN      Sensitivity     Specificity
1246.0  155.0   209.0   1707.0
        0.86            0.92
1707.0  209.0   155.0   1246.0
        0.92            0.86
0.9036527263102171
runtime = 29.918338537216187 seconds
```