A DATA ANALYTICS APPROACH TO THE CYBERCRIME UNDER GROUND ECONOMY

218R1A05B4-P.Pavani

218R1A0552-P.Shravani

ABSTRACT

Despite the rapid escalation of cyber threats, there has still been little research into the foundations of the subject or methodologies that could serve to guide information systems researchers and practitioners who deal with cybersecurity. In addition, little is known about crime-as-a-service (CaaS), a criminal business model that underpins the cybercrime underground. This research gap and the practical cybercrime problems we face have motivated us to investigate the cybercrime underground economy by taking a data analytics approach from a design science perspective. To achieve this goal, we: (1) propose a data analysis framework for analyzing the cybercrime underground; (2) propose CaaS and crimeware definitions; (3) propose an associated classification model, and (4) develop an example application to demonstrate how the proposed framework and classification model could be implemented in practice.

INTRODUCTION

- As the threat posed by massive cyberattacks (e.g., ransomware and distributed denial of service attacks (DDoS)) and cybercrimes has grown, individuals, organizations, and governments have struggled to find ways to defend against them. In 2017, ransomware known as WannaCry was responsible for nearly 45,000 attacks in almost 100 countries.
- Global cyberattacks (such as WannaCry and Petya) are executed by highly organized criminal groups, and organized or national-level crime groups have been behind many recent attacks. Typically, criminal groups buy and sell hacking tools and services on the cybercrime black market, wherein attackers share a range of hacking-related information.

REQUIREMENTS

SOFTWARE REQUIREMENTS

- Operating System : Windows 11.
- Coding Language : Python.
- Front : Python.
- Setup tools and pip : Python.

HARDWARE REQUIREMENTS

- System : Pentium IV
- RAM : 4GB.

EXISTING SYSTEM

- Information Systems (IS) researchers and practitioners are taking an increasing interest in cybercrime, due to the critical issues arising from the rapid increase in cyber threats, few have attempted to put this new interest on a solid foundation or develop suitable methodologies. Previous studies have not analyzed the underground economy behind cybercrime in depth. Furthermore, little is known about CaaS, one of the primary business models behind the cybercrime underground. There is an overall lack of understanding, both in research and practice, of the nature of this underground and the mechanisms underlying it.

PROPOSED SYSTEM

- Propose a data analysis framework for analyzing the cybercrime underground to guide researchers and practitioners
- Define CaaS and crimeware to better reflect their features from both academic research and business practice perspectives
- Use this to build a classification model for CaaS and crimeware
- Build an application to demonstrate how the proposed framework and classification model could be implemented in practice.

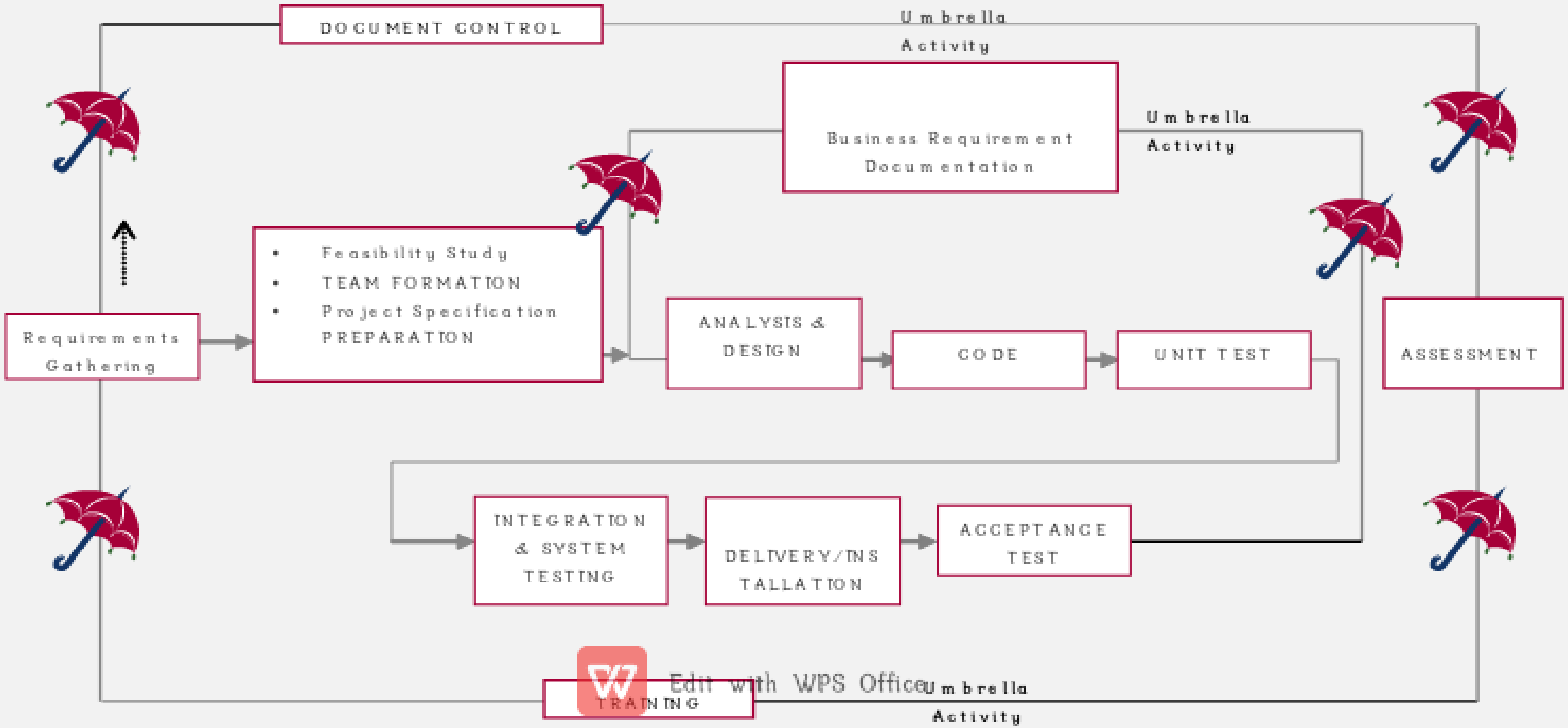
ADVANTAGES

High accuracy

LITERATURE REVIEW

- Crimeware-as-a-service—A survey of commoditized crimeware in the underground marketCrimeware-as-a-service (CaaS) has become a prominent component of the underground economy. CaaS provides a new dimension to cyber crime by making it more organized, automated, and accessible to criminals with limited technical skills. This paper dissects CaaS and explains the essence of the underground economy that has grown around it. The paper also describes the various crimeware services that are provided in the underground market.

SYSTEM ARCHITECTURE



DOCUMENT CONTROL

Umbrella Activity

Business Requirement Documentation

Umbrella Activity

Requirements Gathering

- Feasibility Study
- TEAM FORMATION
- Project Specification PREPARATION

ANALYSIS & DESIGN

CODE

UNIT TEST

ASSESSMENT

INTEGRATION & SYSTEM TESTING

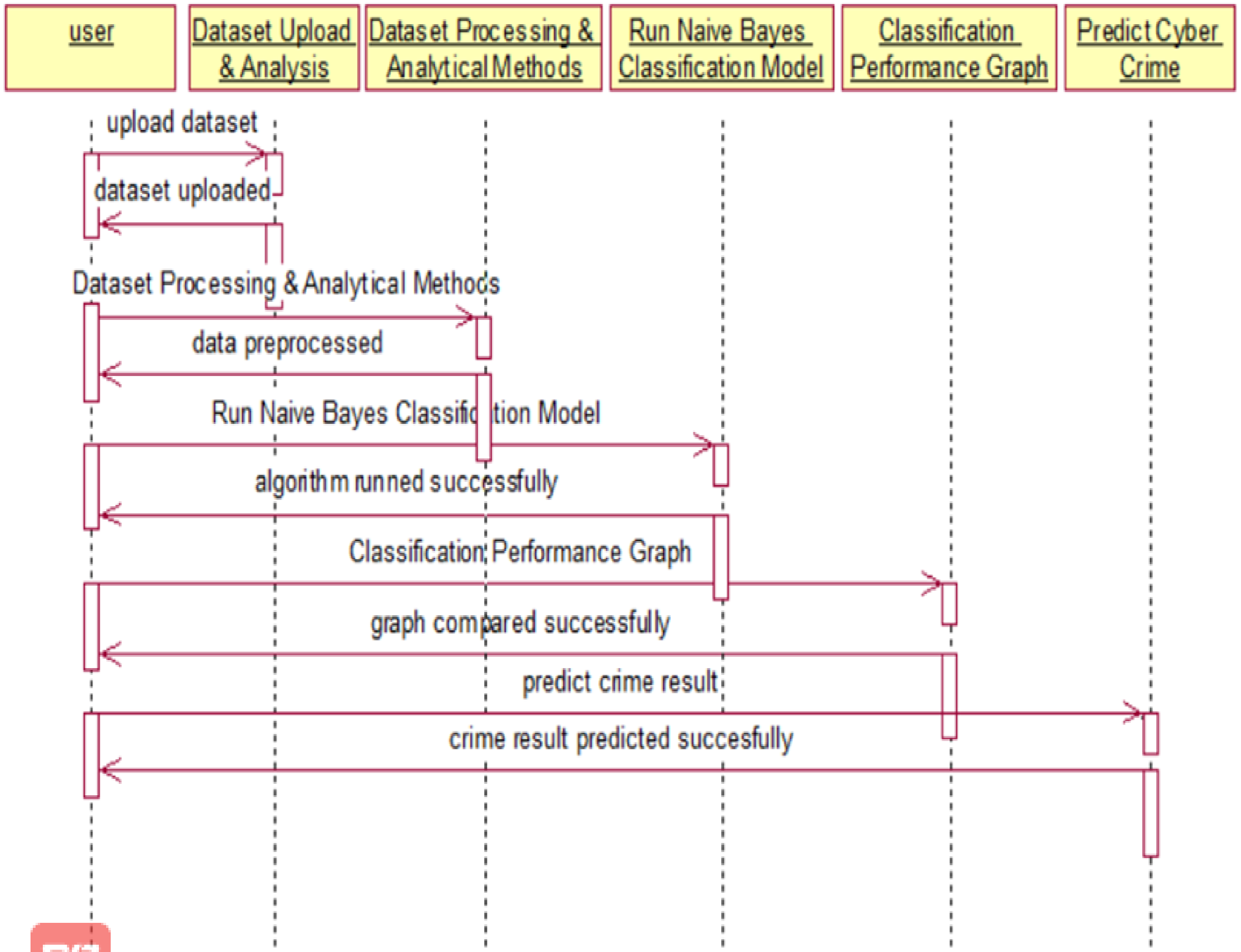
DELIVERY/INSTALLATION

ACCEPTANCE TEST

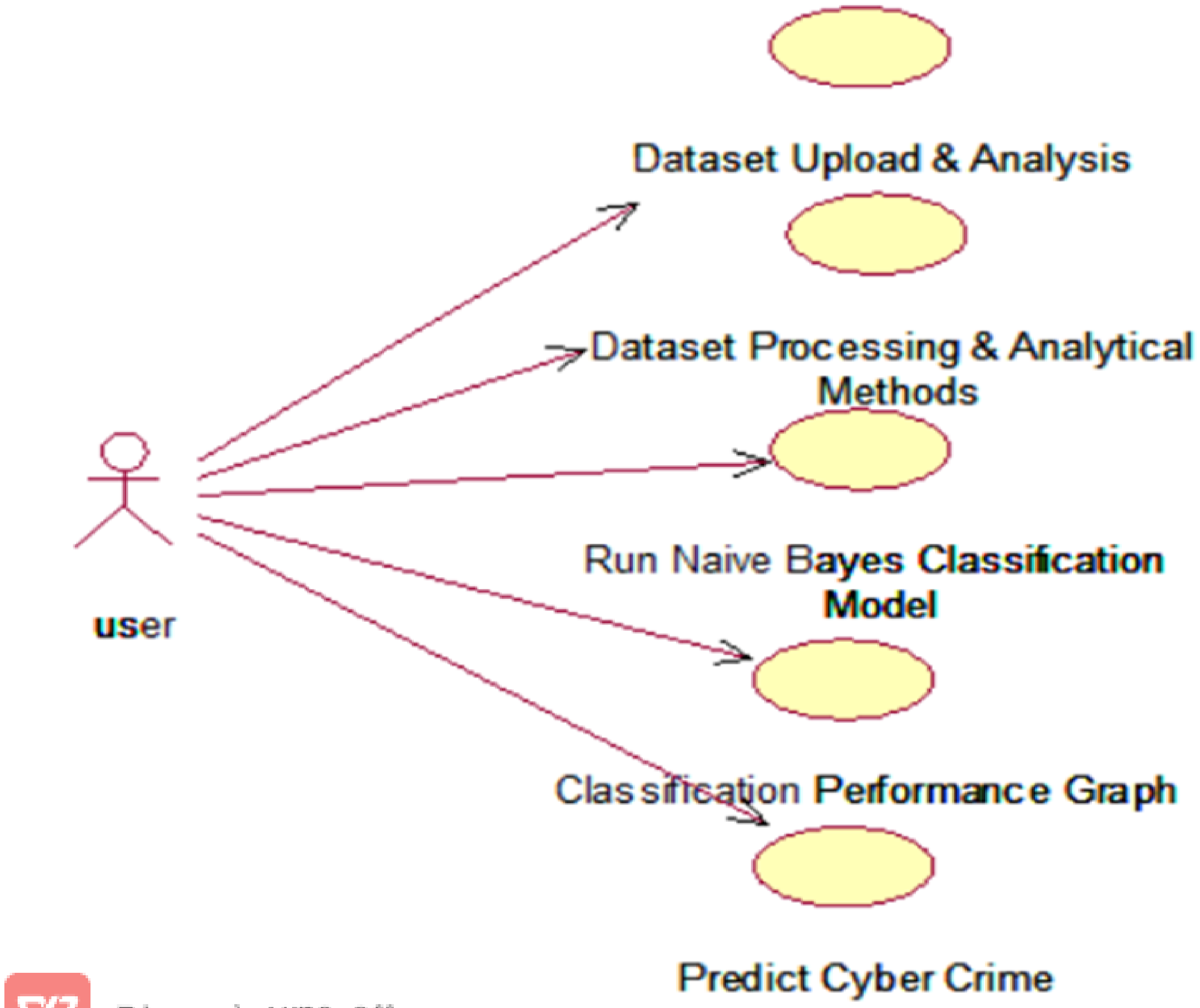
TRAINING

Umbrella Activity

SEQUENCE DIAGRAM


user
Dataset Upload & Analysis
Dataset Processing & Analytical Methods
Run Naive Bayes Classification Model
Classification Performance Graph
Predict Cyber Crime
upload dataset
dataset uploaded
Dataset Processing & Analytical Methods
data preprocessed
Run Naive Bayes Classification Model
algorithm runned successfully
Classification Performance Graph
graph compared successfully
predict crime result
crime result predicted succesfully

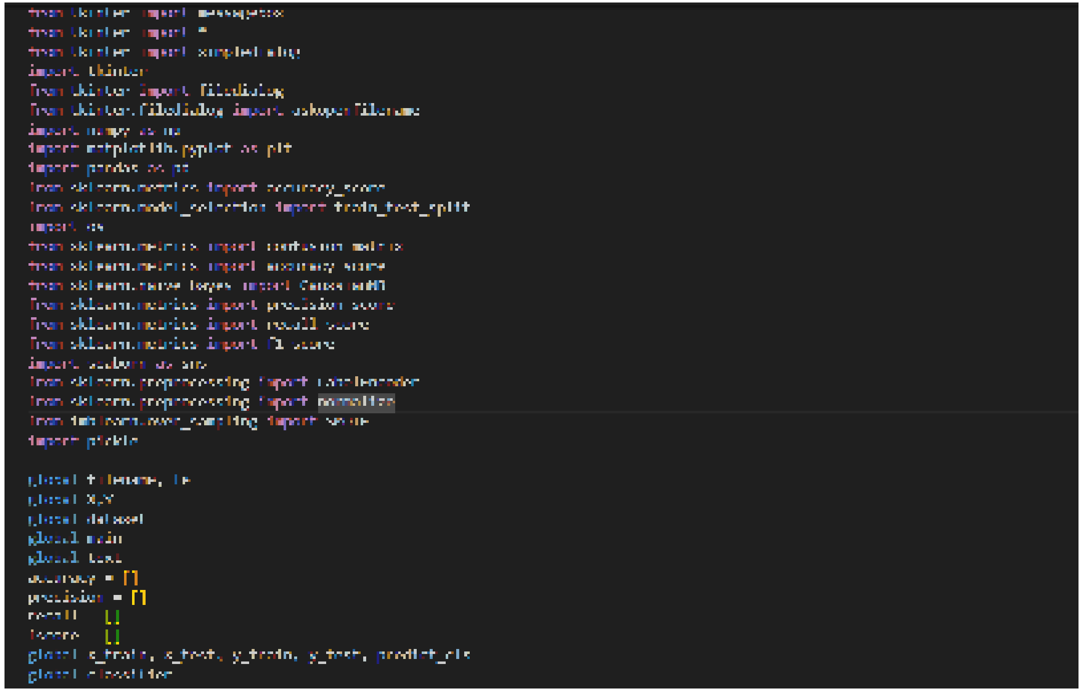
USE CASE DIAGRAM


Dataset Upload & Analysis
Dataset Processing & Analytical Methods
Run Naive Bayes Classification Model
Classification Performance Graph
Predict Cyber Crime
user

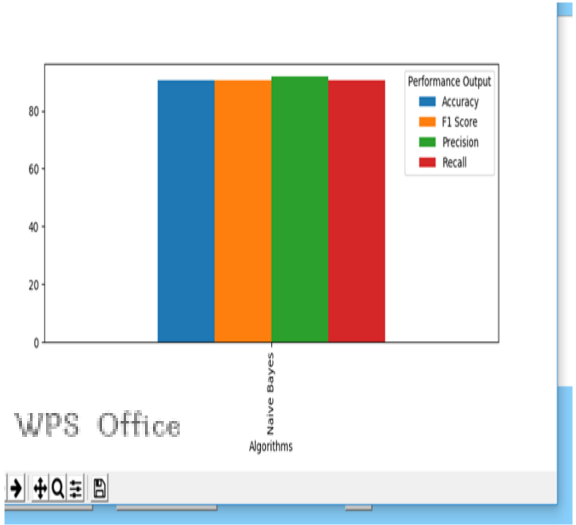
MODULES

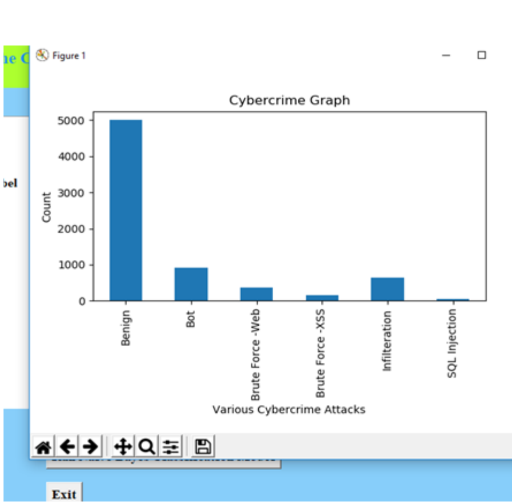
matplotlib	3.1.1
numpy	1.19.2
pandas	0.25.3
pip	23.2.1
scikit-learn	1.0.2
scipy	1.7.3
seaborn	0.10.1
tensorflow	1.14.0
Keras-Applications	1.0.8
Keras-Preprocessing	1.1.2

SOURCE CODE



```
from tkinter import messagebox
from tkinter import *
from tkinter import simpledialog
import tkinter
from tkinter import filedialog
from tkinter.filedialog import askopenfilename
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
from sklearn.metrics import accuracy_score
from sklearn.model_selection import train_test_split
import os
from sklearn.metrics import confusion_matrix
from sklearn.metrics import accuracy_score
from sklearn.naive_bayes import GaussianNB
from sklearn.metrics import precision_score
from sklearn.metrics import recall_score
from sklearn.metrics import f1_score
import seaborn as sns
from sklearn.preprocessing import LabelEncoder
from sklearn.preprocessing import normalize
from imblearn.over_sampling import SMOTE
import pickle

global filename, le
global X,Y
global dataset
global main
global text
accuracy = []
precision = []
recall = []
fscore = []
global x_train, x_test, y_train, y_test, predict_cls
global classifier
```

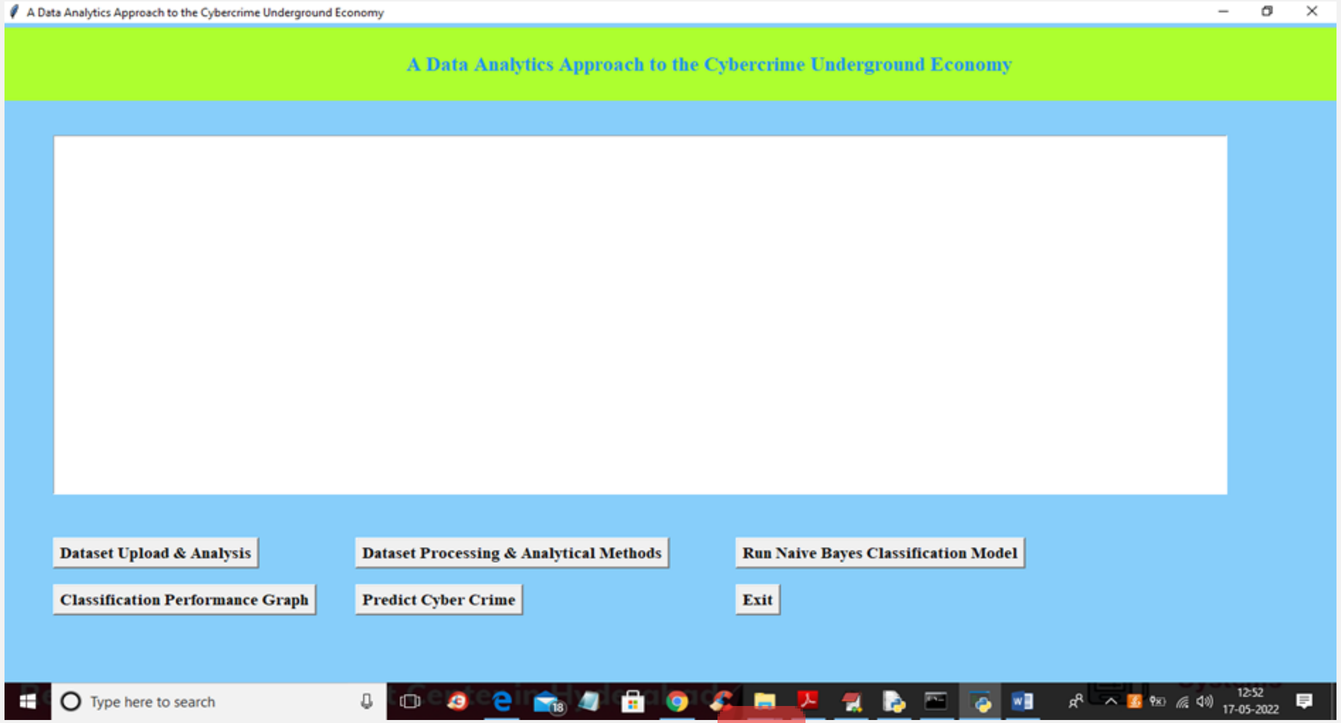


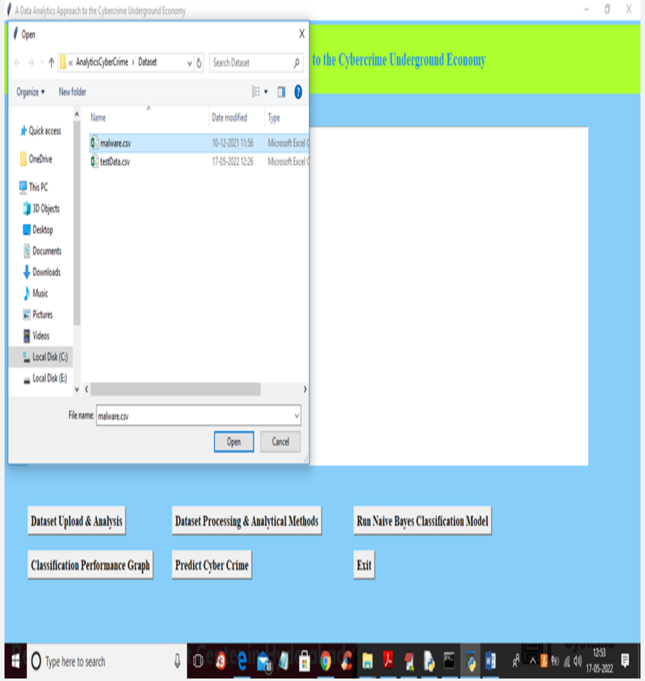


TESTING

Test Case Id	Test Case Name	Test Case Desc.	Test Steps			Test Case Status	Test Priority
			Step	Expected	Actual		
01	Dataset Upload & Analysis	Test whether the Dataset is uploaded or not.	If Dataset is not uploaded	we cannot do further operations	If Dataset uploaded we will do further operations	High	High
02	Dataset Processing & Analytical Methods	Verify the Either Dataset is Preprocess or not into the system	If Dataset may not Preprocess	We cannot do the further operations	Dataset is Preprocessed	High	High
03	Run Naive Bayes Classification Model	Verify the Run Naive Bayes Classification algorithm will run or not	Without training model	we cannot run Naive Bayes Classification algorithm	we can run Naive Bayes Classification algorithm	High	High
04	Classification Performance Graph	Test whether the Graph is displaying or not	Without displaying graph	we cannot do further operations	we can do further operations	High	High
05	Predict Cyber Crime	Verify whether the data is tested or not	Without Predicting result	We cannot get accuracy results	We can get accuracy results	High	High

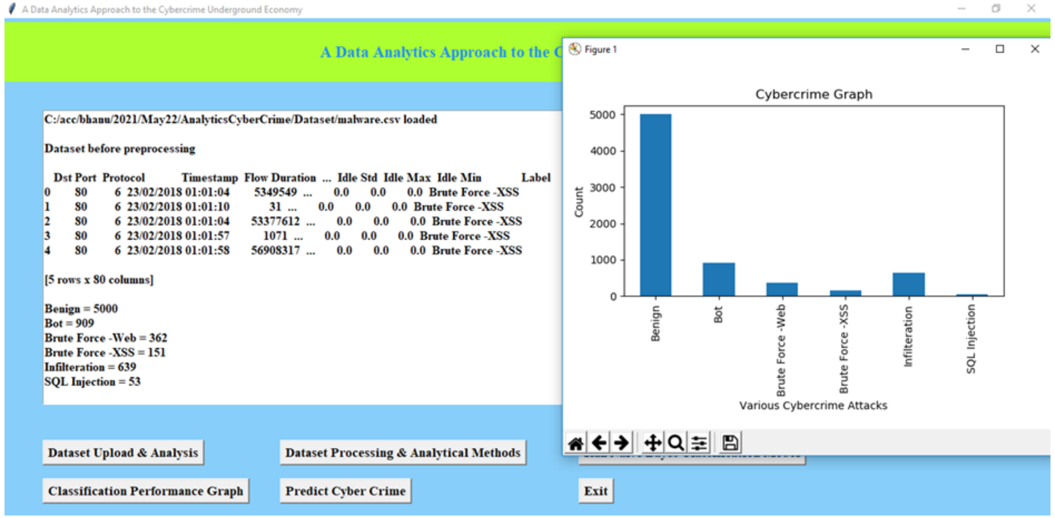
RESULTS


A Data Analytics Approach to the Cybercrime Underground Economy
A Data Analytics Approach to the Cybercrime Underground Economy
Dataset Upload & Analysis
Dataset Processing & Analytical Methods
Run Naive Bayes Classification Model
Classification Performance Graph
Predict Cyber Crime
Exit
Type here to search
12:52
17-05-2022


A Data Analytics Approach to the Cybercrime Underground Economy
Open
« AnalyticsCyberCrime › Dataset
Search Dataset
Organize ▾
New folder
Quick access
OneDrive
This PC
3D Objects
Desktop
Documents
Downloads
Music
Pictures
Videos
Local Disk (C:)
Local Disk (E:)
Open
Cancel
to the Cybercrime Underground Economy
Dataset Upload & Analysis
Dataset Processing & Analytical Methods
Run Naive Bayes Classification Model
Classification Performance Graph
Predict Cyber Crime
Exit
Type here to search

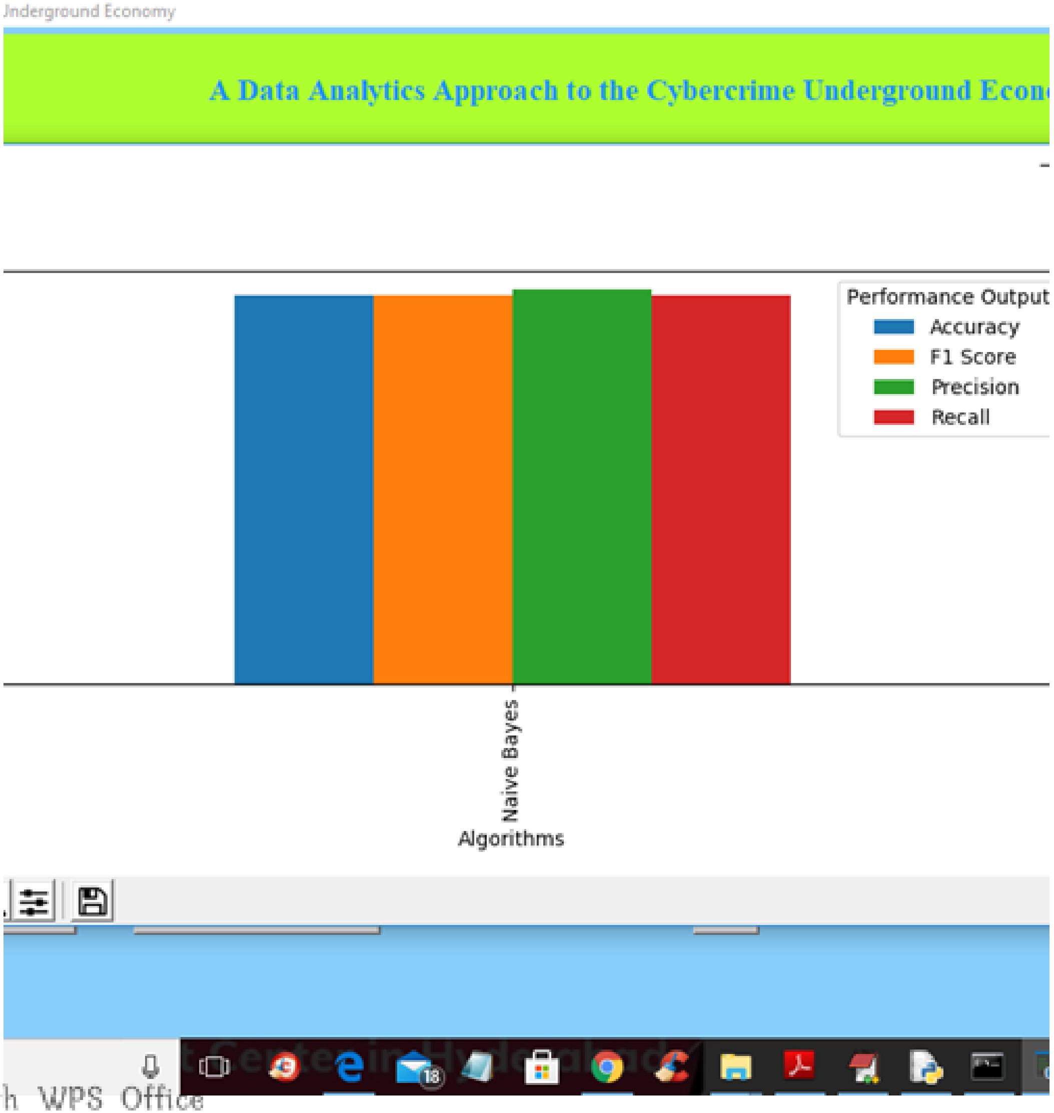
Name	Date modified	Type
malware.csv	10-12-2021 11:56	Microsoft Excel C
testData.csv	17-05-2022 12:26	Microsoft Excel C

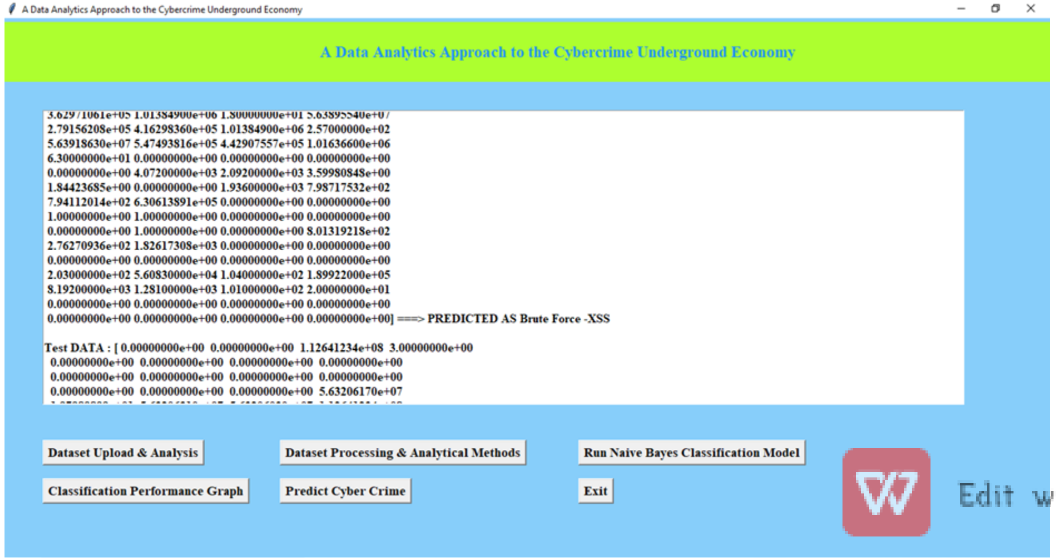
File name: malware.csv



A Data Analytics Approach to the Cybercrime Underground Economy
A Data Analytics Approach to the C
Figure 1
Cybercrime Graph
Count
5000
4000
3000
2000
1000
0
Benign
Bot
Brute Force -Web
Brute Force -XSS
Infilteration
SQL Injection
Various Cybercrime Attacks
Dataset Upload & Analysis
Dataset Processing & Analytical Methods
Classification Performance Graph
Predict Cyber Crime
Exit

```
C:/acc/bhanu/2021/May22/AnalyticsCyberCrime/Dataset/malware.csv loaded

Dataset before preprocessing

  Dst Port  Protocol          Timestamp  Flow Duration  ...  Idle Std  Idle Max  Idle Min           Label
0       80         6  23/02/2018 01:01:04        5349549  ...       0.0       0.0       0.0  Brute Force -XSS
1       80         6  23/02/2018 01:01:10             31  ...       0.0       0.0       0.0  Brute Force -XSS
2       80         6  23/02/2018 01:01:04       53377612  ...       0.0       0.0       0.0  Brute Force -XSS
3       80         6  23/02/2018 01:01:57           1071  ...       0.0       0.0       0.0  Brute Force -XSS
4       80         6  23/02/2018 01:01:58       56908317  ...       0.0       0.0       0.0  Brute Force -XSS

[5 rows x 80 columns]

Benign = 5000
Bot = 909
Brute Force -Web = 362
Brute Force -XSS = 151
Infilteration = 639
SQL Injection = 53
```





A Data Analytics Approach to the Cybercrime Underground Economy
A Data Analytics Approach to the Cybercrime Underground Economy
Dataset Upload & Analysis
Dataset Processing & Analytical Methods
Run Naive Bayes Classification Model
Classification Performance Graph
Predict Cyber Crime
Exit

```
3.62971061e+05 1.01384900e+06 1.80000000e+01 5.63895540e+07
2.79156208e+05 4.16298360e+05 1.01384900e+06 2.57000000e+02
5.63918630e+07 5.47493816e+05 4.42907557e+05 1.01636600e+06
6.30000000e+01 0.00000000e+00 0.00000000e+00 0.00000000e+00
0.00000000e+00 4.07200000e+03 2.09200000e+03 3.59980848e+00
1.84423685e+00 0.00000000e+00 1.93600000e+03 7.98717532e+02
7.94112014e+02 6.30613891e+05 0.00000000e+00 0.00000000e+00
1.00000000e+00 1.00000000e+00 0.00000000e+00 0.00000000e+00
0.00000000e+00 1.00000000e+00 0.00000000e+00 8.01319218e+02
2.76270936e+02 1.82617308e+03 0.00000000e+00 0.00000000e+00
0.00000000e+00 0.00000000e+00 0.00000000e+00 0.00000000e+00
2.03000000e+02 5.60830000e+04 1.04000000e+02 1.89922000e+05
8.19200000e+03 1.28100000e+03 1.01000000e+02 2.00000000e+01
0.00000000e+00 0.00000000e+00 0.00000000e+00 0.00000000e+00
0.00000000e+00 0.00000000e+00 0.00000000e+00 0.00000000e+00] ===> PREDICTED AS Brute Force -XSS

Test DATA : [ 0.00000000e+00  0.00000000e+00  1.12641234e+08  3.00000000e+00
 0.00000000e+00  0.00000000e+00  0.00000000e+00  0.00000000e+00
 0.00000000e+00  0.00000000e+00  0.00000000e+00  0.00000000e+00
 0.00000000e+00  0.00000000e+00  0.00000000e+00  5.63206170e+07
```

CONCLUSION

- This study also has important implications for society. Over the last few years, the world has been facing cyberterrorism and cyberwar threats from nation-sponsored attackers
- As a result, governments should, for example, strengthen their ability to protect their citizens in online virtual environments by enhancing their immediate responses to threats such as cyberespionage and cyberterrorism. This issue therefore has profound implications in terms of the need for a global cyber defense to maintain a cyber-safe environment.

FUTURE SCOPE

Data analytics can be used to identify emerging cyber threats and trends within the underground economy. This includes predicting the types of attacks that might gain traction, the targets they'll go after, and the tools they'll use. Data analytics can be used to profile cybercriminals and their behavior patterns. By analyzing their tactics, techniques, and procedures (TTPs), security professionals can develop better strategies for defense and attribution. As data analytics tools are used to combat cybercrime, ethical considerations regarding privacy, data collection, and the balance between surveillance and civil liberties will become more important. Overall, the future of data analytics in combating the cybercrime underground economy is bright, but it also comes with challenges related to ethics, privacy, and the constant evolution of cyber threats.

REFERENCES

- [1] J. C. Wong and O. Solon. (2017, May 12). Massive ransomware cyber-attack hits nearly 100 countries around the world. [Online]. Available: https://www.theguardian.com/technology/2017/may/12/global-cyber-attack-ransomware-nsa-uk-nhs
- [2] “FACT SHEET: Cybersecurity National Action Plan,” ed: The White House, 2016.
- [3] A. K. Sood and R. J. Enbody, “Crimeware-as-a-service—A survey of commoditized crimeware in the underground market,” Int. J. Crit. Infr. Prot., vol. 6, no. 1, pp. 28–38, 2013.

THANK YOU!