

Configuration Manual

MSc Research Project
Cyber Security

Benetto George
Student ID: x21124485

School of Computing
National College of Ireland

Supervisor: Imran Khan

National College of Ireland
MSc Project Submission Sheet



School of Computing

Student Name: Benetto George

Student ID: X21124485

Programme: MSc Cybersecurity **Year:** 2022 -2023

Module: MSc Research Project

Lecturer: Imran Khan

Submission Due Date: 15 December 2022

Project Title: AdaBoost IDS to detect Zero Day attacks and reduce false positives

 691
Word Count: **Page Count:** ...7.....

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: Benetto George

Date: 15-12-2022

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	<input type="checkbox"/>
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	<input type="checkbox"/>
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.	<input type="checkbox"/>

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

Benetto George
Student ID: x21124485

1 Introduction

This document discusses how the AdaBoost Intrusion Detection System code should be implemented and executed. The project has been coded in Python programming language.

2 System Requirements

Code Editor: Visual Studio Code

Python Version: version3

Operating System Windows 7 or later

Anaconda software has been downloaded and installed on the system. It can be downloaded from the official website given below (“Anaconda | Anaconda Distribution,” n.d.),

(<https://www.anaconda.com/products/individual>)

At the stage given in the image below check both checkboxes

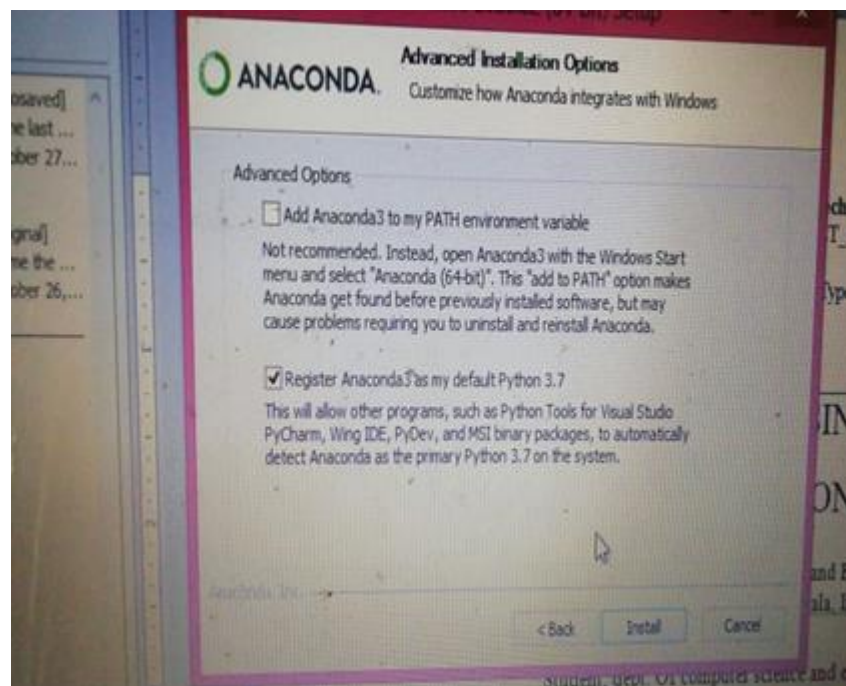


Figure 1: anaconda installation

2.1 Package details

Here in this project, Anaconda software is also used, which is a distribution for Python programming languages for data science. It simplifies package management and deployment. An environment, named “ml_env” is created. It consists of the necessary machine learning python libraries that are custom installed which in turn help to run the whole code. A small number of libraries that are present in the environment are:

- Numpy: Used for array operations
- Sklearn: Used for selection of features and for train-test splitting
- Pandas: Used to read the dataset
- Matplotlib: Is used for data visualization

3 Setting up the environment

- Extract IDS_FULLLCODE.zip folder given at the end of the Thesis report. If you are downloading the artefact from moodle, make sure to download the dataset from Kaggle and paste it in the **Project_Dataset** folder of IDS_FULLLCODE > Machine Learning.
- Extract the ml_env.zip folder from the OneDrive link and paste it into OS(C:) > users > user (it may vary depending on your system) > anaconda3 > envs
- Open anaconda prompt or Windows Command Prompt (cmd) in the project folder
- In the prompt, type command “activate ml_env” (in windows command prompt), or,
- In the anaconda prompt, type “conda activate ml_env”

4 Dataset Source

Dataset used in this research project is taken from an online platform named Kaggle which allows users to access and download various dataset samples. The dataset used has been downloaded and added in the IDS_FULLLCODE folder named Project_Dataset.

5 Code Execution

Anaconda prompt or cmd has been opened. Now, run the following commands,

- Run command: python train.py || To train the model

```

C:\Windows\System32\cmd.e X + v
D:\NCI\Research Methods\Thesis Proj\IDS_FULLCODE\Machine Learning>activate ml_env

(ml_env) D:\NCI\Research Methods\Thesis Proj\IDS_FULLCODE\Machine Learning>python train.py
DoS attacks-SlowLoris      10990
DoS attacks-GoldenEye     41508
Benign                     996077
Name: Label, dtype: int64
Malicious                  52498
Benign                     996077
Name: Label, dtype: int64
Flow Byts/s                4921
dtype: int64
Series([], dtype: float64)
   Dst Port  Protocol  Flow Duration  Tot Fwd Pkts  ...  Idle Std  Idle Max  Idle Min  Label
5000      80         6      12000099         4 ...    0.0    6994310    6994310  Malicious
5001      80         6      11999501         4 ...    0.0    6994337    6994337  Malicious
5002      80         6      12000439         4 ...    0.0    6995703    6995703  Malicious
5003      80         6      11999660         4 ...    0.0    6994306    6994306  Malicious
5004      80         6      12001203         4 ...    0.0    7000221    7000221  Malicious
...      ...      ...      ...      ...      ...      ...      ...      ...      ...
799995     53        17         2529         1 ...    0.0         0         0      Benign
799996     53        17          357         1 ...    0.0         0         0      Benign
799997     53        17          457         1 ...    0.0         0         0      Benign
799998    3389         6      2085036         8 ...    0.0         0         0      Benign
799999     53        17          370         1 ...    0.0         0         0      Benign

[795000 rows x 69 columns]
Malicious    47551
Benign       747449
Name: Label, dtype: int64

```

Figure 2 running train.py

- Run command: python predict.py || To predict whether the packet is Malicious or Benign based on the trained model.
To add the data shown below and in the GUI, there is text file called Test.txt with the values of the 20 features.

```

C:\Windows\System32\cmd.e X + v
(ml_env) D:\NCI\Research Methods\Thesis Proj\IDS_FULLCODE\Machine Learning>
(ml_env) D:\NCI\Research Methods\Thesis Proj\IDS_FULLCODE\Machine Learning>
(ml_env) D:\NCI\Research Methods\Thesis Proj\IDS_FULLCODE\Machine Learning>python predict.py
Enter Dst Port : 0
Enter Protocol : 0
Enter Tot Fwd Pkts : 3
Enter Tot Bwd Pkts : 0
Enter Bwd IAT Min : 0
Enter Bwd IAT Mean : 0
Enter Bwd IAT Max : 0
Enter Idle Max : 56321077
Enter Idle Mean : 56320579
Enter Idle Min : 56320081
Enter Fwd IAT Max : 56321077
Enter Flow IAT Max : 56321077
Enter Flow IAT Std : 704.27835
Enter Fwd IAT Min : 56321077
Enter Fwd IAT Mean : 56320579
Enter Bwd IAT Std : 0
Enter Idle Std : 704.27835
Enter Bwd IAT Tot : 0
Enter Flow Duration : 112641158
Enter Fwd IAT Tot : 112641158
0

*****Result*****
Benign

(ml_env) D:\NCI\Research Methods\Thesis Proj\IDS_FULLCODE\Machine Learning>

```

Figure 3 running predict.py

- Run command: `python gui.py` || To show the Graphical User Interface which takes user input and displays the result.

Intrusion Detection

Home Check

INPUT

Dst Port :	<input type="text" value="80"/>	Fwd IAT Max :	<input type="text" value="53247583"/>
Protocol :	<input type="text" value="6"/>	Flow IAT Max :	<input type="text" value="53247583"/>
Tot Fwd pkts :	<input type="text" value="15"/>	Flow IAT Std :	<input type="text" value="13431392.2"/>
Tot Bwd Pkts :	<input type="text" value="3"/>	Fwd IAT Min :	<input type="text" value="192"/>
Bwd IAT Min :	<input type="text" value="30997119"/>	Fwd IAT Mean :	<input type="text" value="7643258.42"/>
Bwd IAT Mean :	<input type="text" value="52991401"/>	Bwd IAT Std :	<input type="text" value="31104611.9"/>
Bwd IAT Max :	<input type="text" value="74985683"/>	Idle Std :	<input type="text" value="20661603"/>
Idle Max :	<input type="text" value="53247583"/>	Bwd IAT Tot :	<input type="text" value="105982802"/>
Idle Mean :	<input type="text" value="23674282"/>	Flow Duration :	<input type="text" value="107005621"/>
Idle Min :	<input type="text" value="6655865"/>	Fwd IAT Tot :	<input type="text" value="107005618"/>

Figure 4: GUI 1

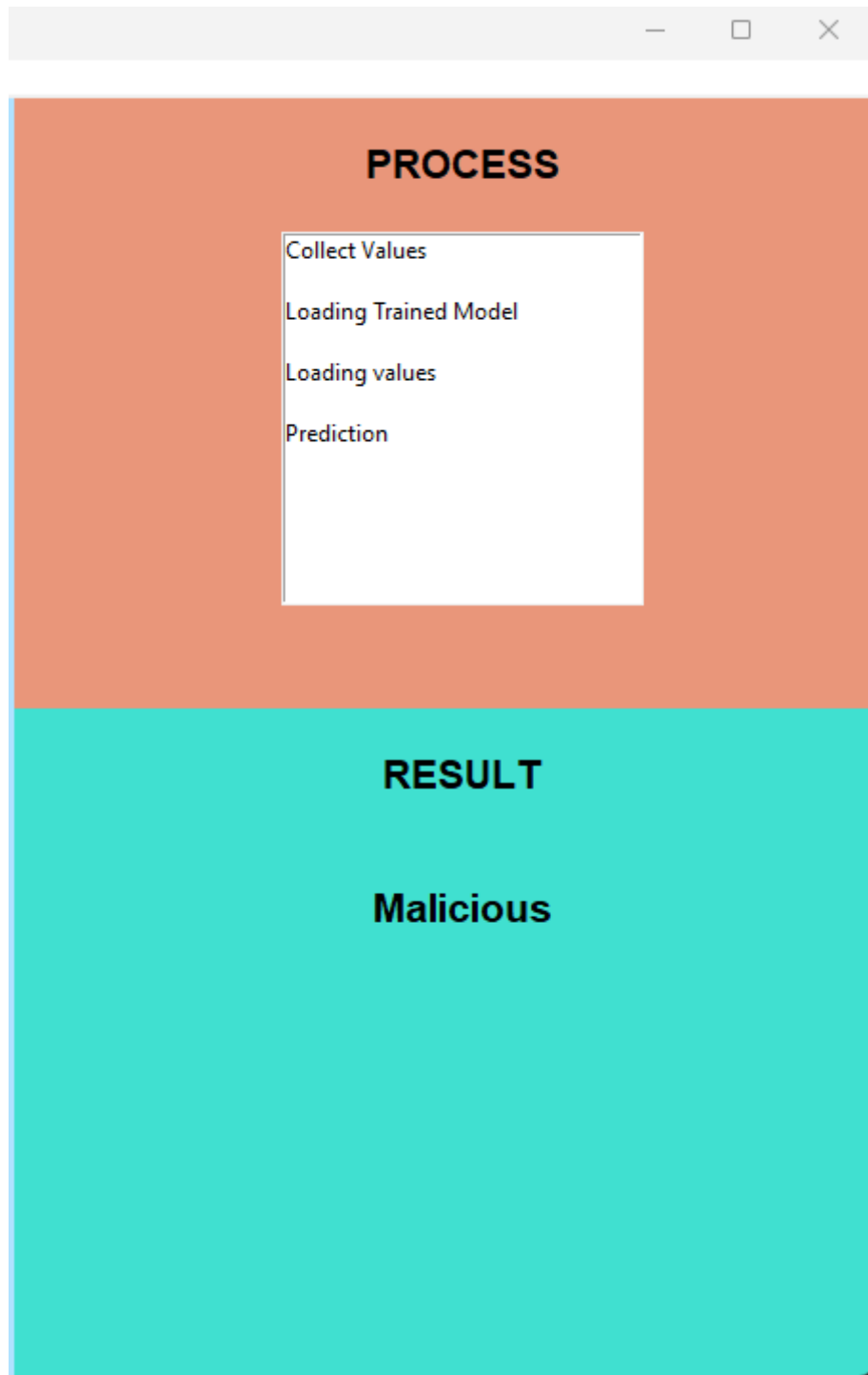


Figure 5: GUI 2

6 References

Anaconda | Anaconda Distribution [WWW Document], n.d. . Anaconda. URL <https://www.anaconda.com/products/distribution> (accessed 12.6.22).