Project Report On

Web Data Mining For Terrorism Analysis

Technology – Python
Modules - Tkinter, Requests, BeautifulSoup4, Pillow, Ttkthemes

Submitted in the partial fulfilment for the **Semester-V Mini-Project** in

MASTER OF COMPUTER APPLICATIONS
(M.C.A.)
[Batch 2019-2021]

Submitted by

Group-22 Team Members

Bamaniya Virat Kanji (191823004) | Solanki Milan Dayaram (191823046n)

Under the guidance of
Prof. Shrey Shah

M.C.A. Head of Department
Dr. Parag Shukla

ACKNOWLEDEGEMENT

The success and final outcome of this project required a lot of guidance and assistance from many people and we extremely privileged to have got this all along the completion of our project. All that we have done is only due to such supervision and assistance and we would not forget to thank them.

We learn lots of thing like Analysis, Designing, Deal with Large number of Datasets, Test-cases, Debugging and lot many.

It is our gratefulness to acknowledge the sincere help and guidance from all our teachers and peers. I would firstly like to thank and gratify to Mini-Project Guide **Prof. Shrey Shah** for their all-time support and sincere guidance, without which project could have never reached to its completion. Their all-time readiness for understanding the problem faced by the students and their ability to solve them with ease is truly appreciable and unforgettable.

Special thanks to head of the department **Dr. Parag Shukla** for their guidance and constant supervision as well as for providing necessary information regarding the Ethics and Technical Concepts, who always inspired us in semester-V Mini-Project.

We would like to thank from core of our heart to Department of MCA for giving the awesome opportunity for doing the Mini-Project, support and guidance from all Teaching staffs of MCA Department which helped us in successfully completing our project work.

Table of Content

Contents

1. PROJECT PROFILE

1.1. Project Description

We use web mining algorithms to mine textual information on web pages and detect their relevancy to terrorism. This system will check web pages whether a webpage is promoting terrorism. Data mining is a technique used to mine out patterns of useful data from large data sets and make the most use of obtained results. Web mining also consists of text mining methodologies that allow us to scan and extract useful content from unstructured data.

1.2. System Requirements

1.2.1. Minimal Hardware Requirements

Processor	Intel(R) Core (TM) i3-4005U CPU @ 1.70 GHz
RAM	4 GB
Hard Disk	250GB
Input Devices	Keyboard, Mouse

1.2.2. Software Requirements

Operating System	Windows XP+ all version
Front END	Python
Modules	Tkinter, Requests, BeautifulSoup4, Pillow, Ttkthemes

2. ABOUT THE TOOLS & TECHNOLOGIES

2.1. Python

- Python can be easy to pick up whether you're a first time programmer or you're experienced with other languages. The community hosts conferences and meetups, collaborates on code, and much more.
- Python is developed under an OSI-approved open source license is administered by the Python Software Foundation.

2.1.1 Python modules

- The Python Package Index (PyPI) hosts thousands of third-party modules for Python. Both Python's standard library and the community-contributed modules allow for endless possibilities.
-

2.1.1.1 Tkinter

- Python has a lot of GUI frameworks, but Tkinter is the only framework that's built into the Python standard library.
- Tkinter has several strengths. It's cross-platform, so the same code works on Windows, macOS, and Linux.
- Visual elements are rendered using native operating system elements, so applications built with Tkinter look like they belong on the platform where they're run.

2.1.1.2 Requests

- Requests allows you to send HTTP/1.1 requests extremely easily. There's no need to manually add query strings to your URLs, or to form-encode your PUT & POST data — but nowadays, just use the json method!
- Requests is one of the most downloaded Python package today, pulling in around 14M Downloads / Week according to GitHub. Requests is a simple, yet elegant HTTP library.

2.1.1.3 BeautifulSoup4

- Beautiful Soup is a Python library for pulling data out of HTML and XML files.
- It works with your favorite parser to provide idiomatic ways of navigating, searching, and modifying the parse tree.
- It commonly saves programmers hours or days of work.

2.1.1.4 Pillow

- The Python Imaging Library adds image processing capabilities to your Python interpreter.
- This library provides extensive file format support, an efficient internal representation, and fairly powerful image processing capabilities.
- The core image library is designed for fast access to data stored in a few basic pixel formats. It should provide a solid foundation for a general image processing tool.

2.1.1.5 Ttkthemes

- A group of themes for the ttk extenstions for Tkinter gathered together by RedFantom and created by various authors.

3. SYSTEM ANALYSIS

3.1. About Existing System

I didn't find any websites or application similar to mine. There can be further enhacements in the project. My application can be made into a website in Django as python libraries are used.

3.2. Feasibility Study

3.2.1. Economical

Economic analysis is the most frequently used method for evaluating the effectiveness of the candidate system. More commonly known as cost/benefit analysis, the procedure is to be determining the benefits and savings that are expected from a candidate and compare them with costs. If benefits outweigh costs, then the decision is made to design and implement the system.

A systems financial benefit must exceed the cost of developing that system. I.e. a new system being developed should be a good investment for the organization. Economic feasibility considers the following

- The cost to conduct a full system investigation.
- The cost of hardware and software for the class of application.
- The benefits in the form of reduced cost or fewer costly errors.
- The cost if nothing changes (i.e. the proposed system is not developed).

The proposed "Web Data Mining For Terrorism Analysis" is economically feasible because the system requires very less time factors.

- The system will provide fast and efficient environment instead of slow and error prone manual system.
- The system will provide service to view various information for proper managerial decision making.

3.2.2. Technical

Technical feasibility centres around the existing computer system (Hardware and Software etc) and to what extend it support the proposed addition. For example, if the current computer is operating at 80 percent capacity - an arbitrary ceiling - then running another application could overload the system or require additional Hardware. This involves financial considerations to accommodate technical enhancements. If the budgets are a serious constraint, then the project is judged not feasible. In this project, all the necessary cautions have been taken care to make it technically feasible. Using a key, the display of text/object is very fast. Also, the tools, operating system and programming language used in this localization process is compatible with the existing one

3.2.3. Operational

People are inherently resistant to change, and computers have been known to facilitate change. An estimate should be made of how strong a reaction the user staff

is likely to have toward the development of a computerized system. Therefore, it is understandable that the introduction of a candidate system requires special efforts to educate and train the staff. The software that is being developed is user friendly and easy to learn. In this way, the developed software is truly efficient and can work on any circumstances, tradition, locales.

3.3. Limitation of Existing System

If terrorists communicate in keywords it will be difficult for the cops to track it.

3.4. Scope of Proposed System

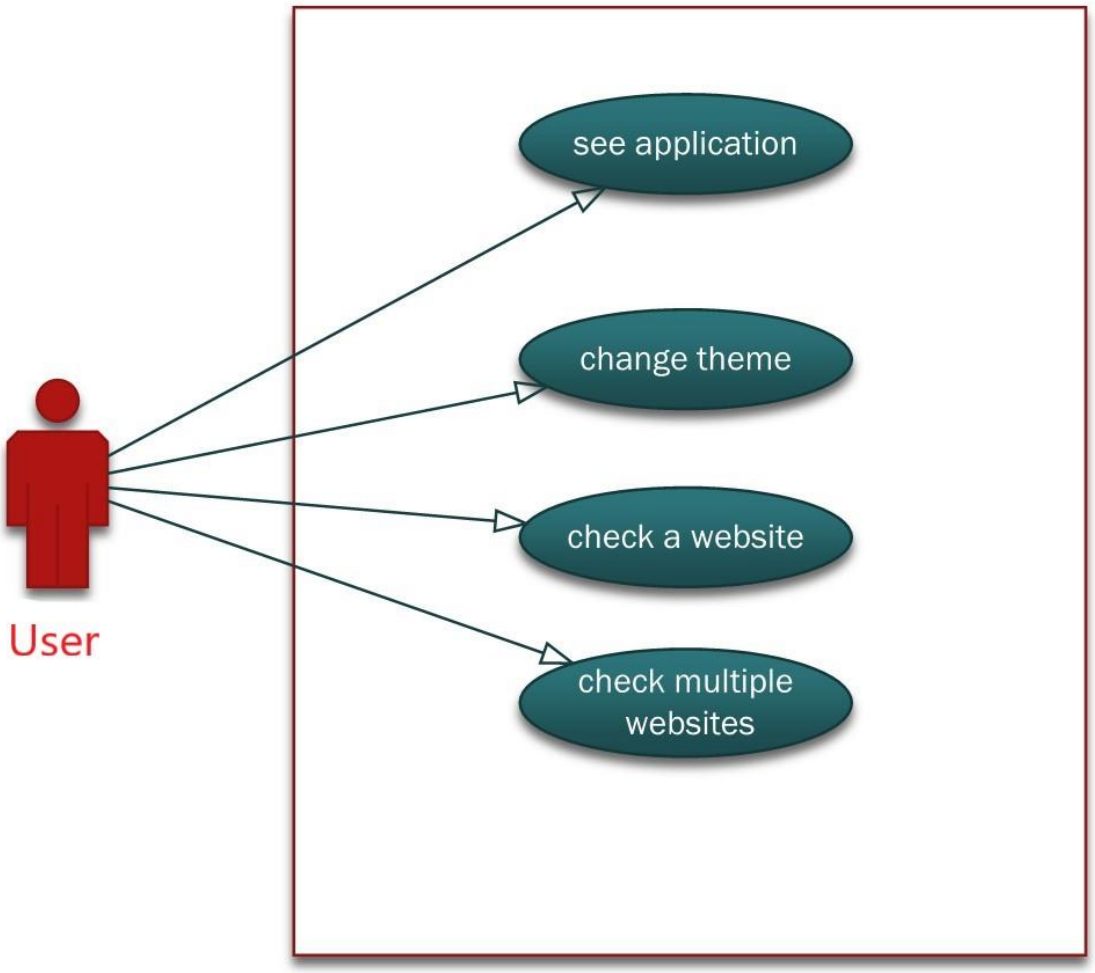
We use web mining algorithms to mine textual information on web pages and detect their relevancy to terrorism.

3.5. Targeted Users

Targeted users are from anywhere in the world. We provide worldwide detection of terrorism.

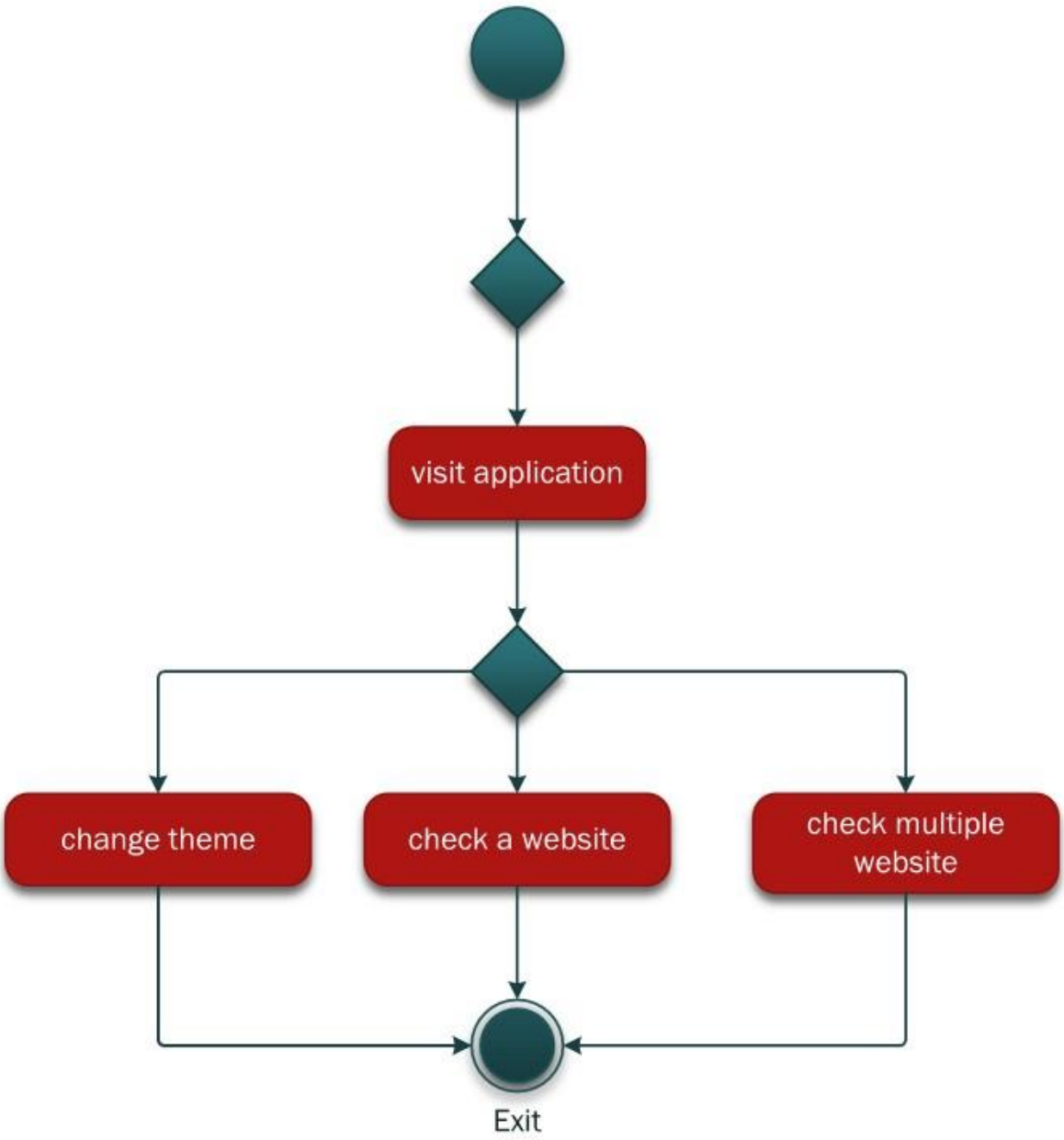
4. SYSTEM DESIGN

4.1. Use Case Diagram



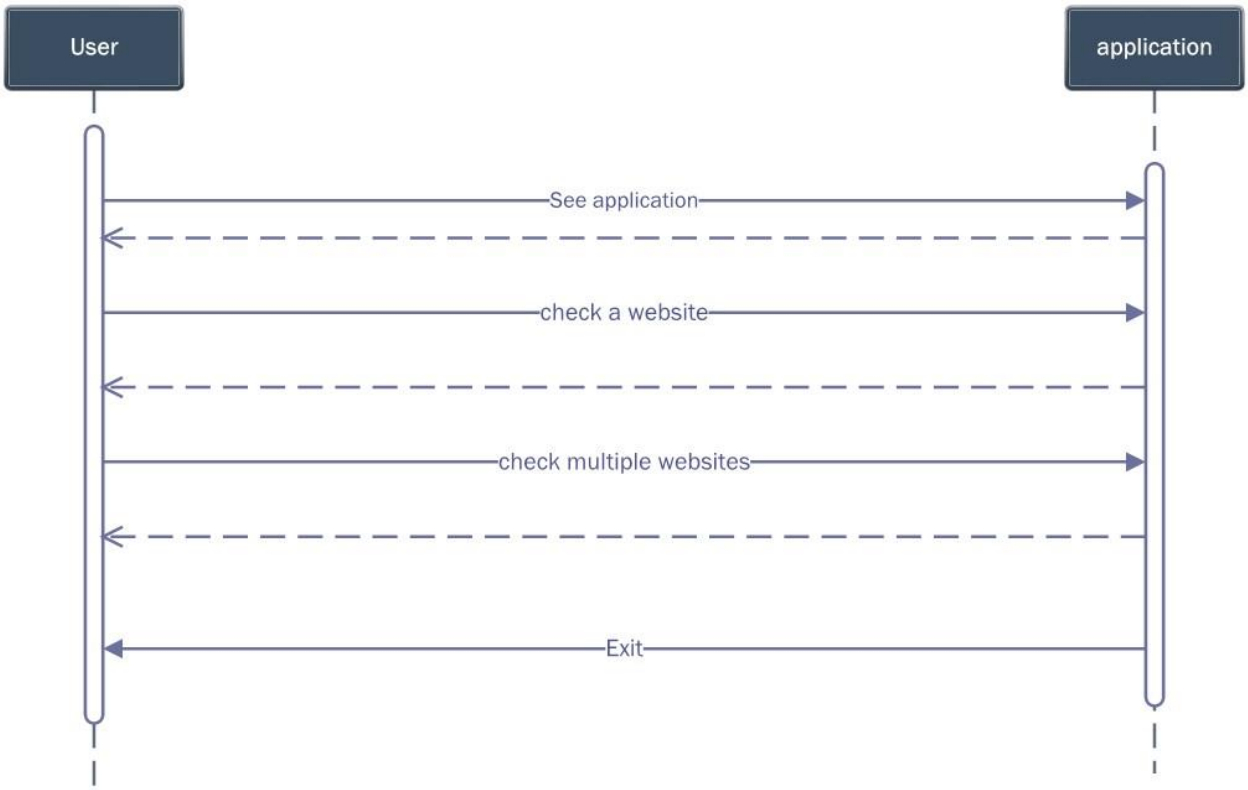
4.2. Activity Diagram

Activity Diagram for User:



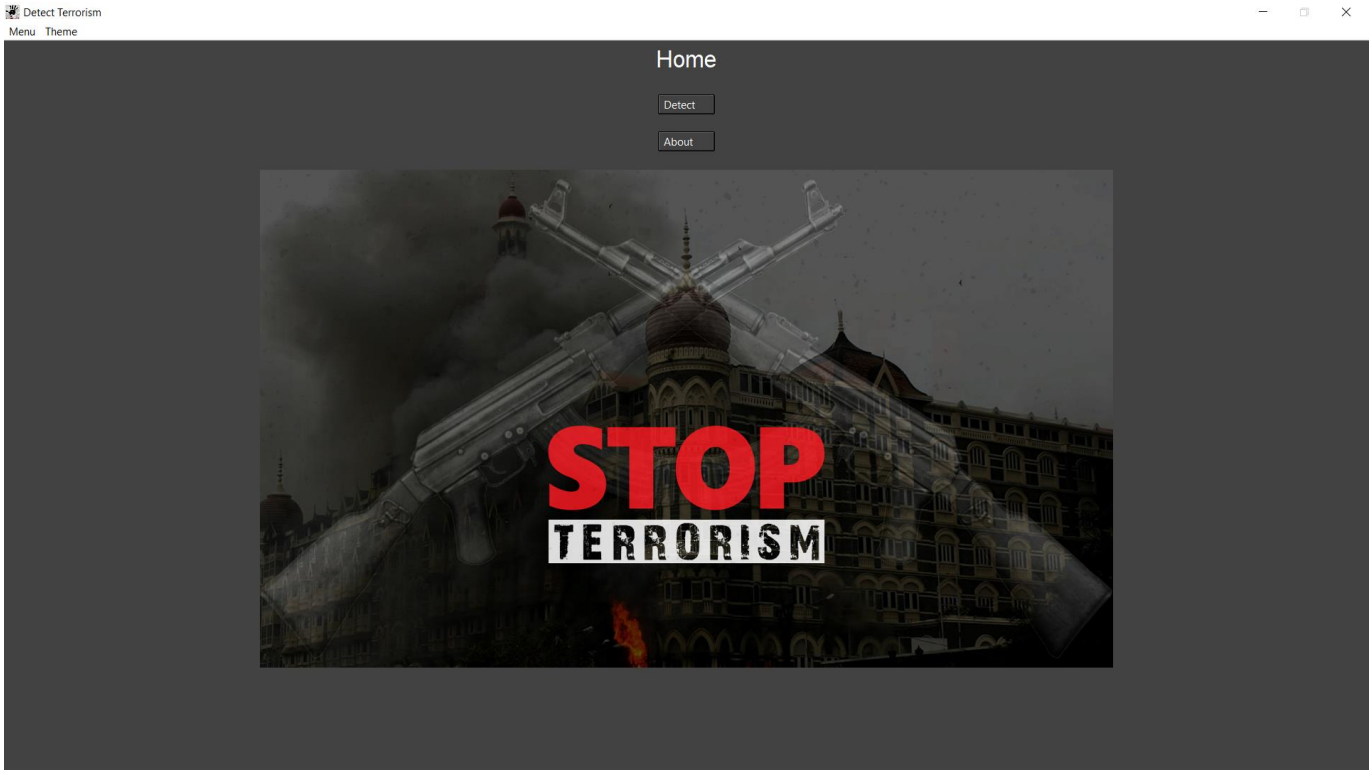
4.3. Sequence Diagram

Sequence Diagram for Donor:

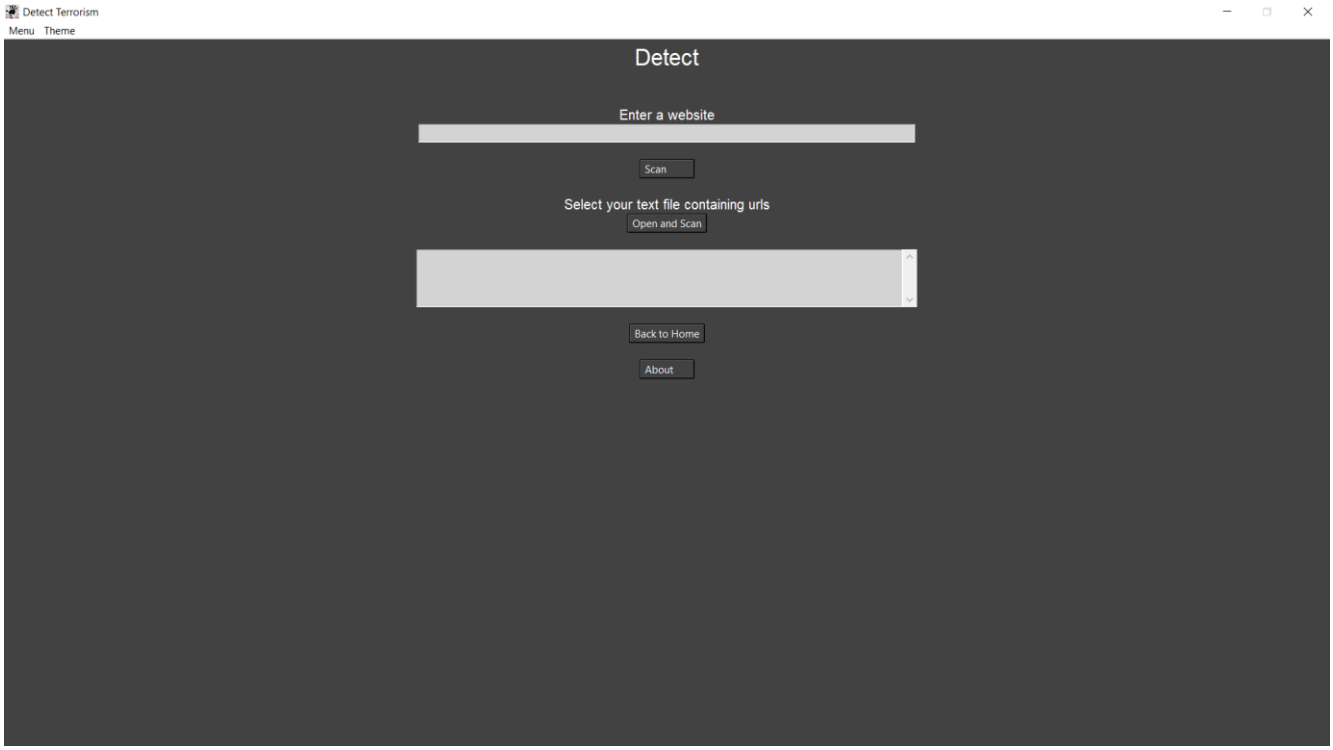


4.4. Screenshots

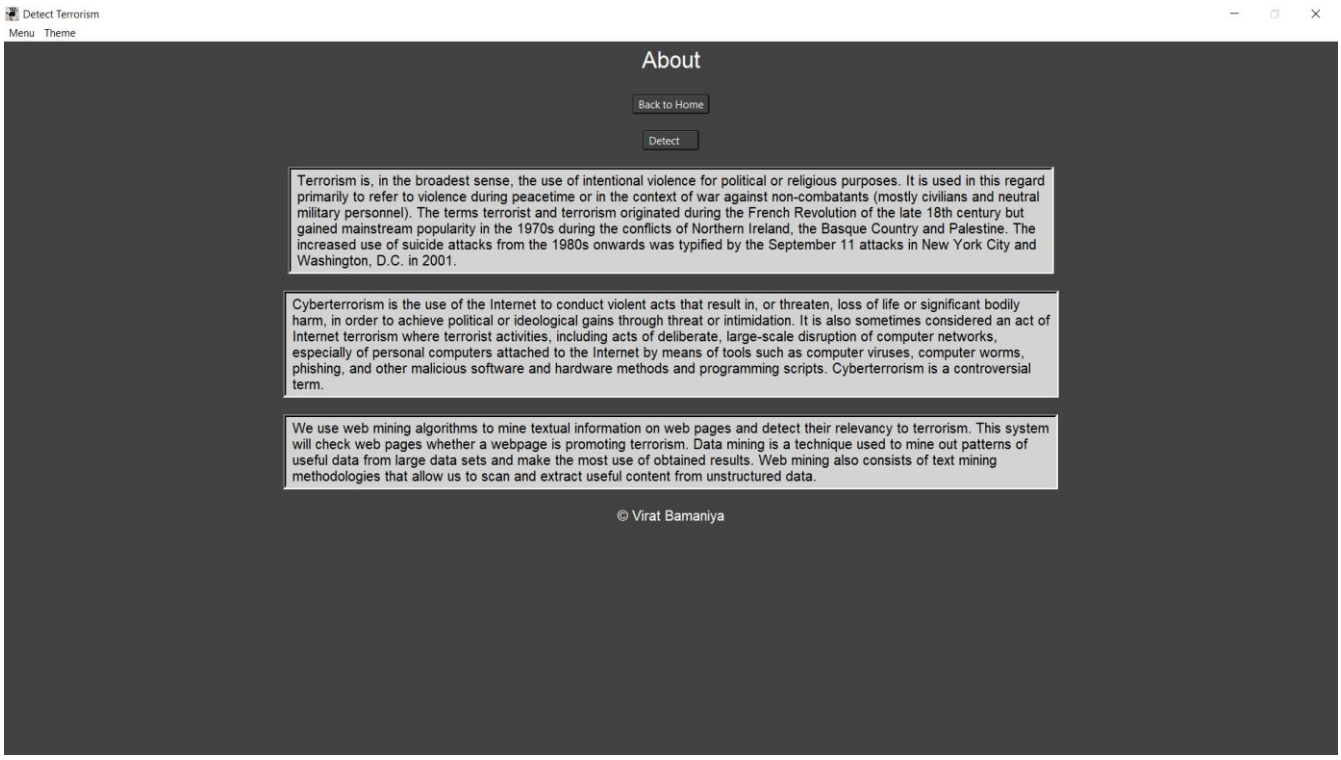
Home page



Detect terrorism page



About Page



WEB REFERENCES:

https://stackoverflow.com/

https://www.geeksforgeeks.org/

https://www.crummy.com/software/BeautifulSoup/bs4/doc/

https://www.javatpoint.com/

https://docs.python.org/3/

https://www.tutorialspoint.com/index.htm

https://www.youtube.com