A Project Report on

Web Data Mining To Detect Online Spread of Terrorism

Submitted in partial fulfillment of the requirement
for the award of the degree of

Bachelor of Technology in Computer Science and Engineering

Under The Supervision of
Ms.Garima Pandey

Submitted By

19SCSE1010306 – Abhishek Yadav

SCHOOL OF COMPUTING SCIENCE AND ENGINEERING

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

GALGOTIAS UNIVERSITY, GREATER NOIDA, INDIA

SCHOOL OF COMPUTING SCIENCE AND ENGINEERING
GALGOTIAS UNIVERSITY, GREATER NOIDA

CANDIDATE'S DECLARATION

I/We hereby certify that the work which is being presented in the project, entitled " Ur Info: A cross-platform Application " in partial fulfillment of the requirements for the award of

BACHELOR OF TECHNOLOGY IN COMPUTER SCIENCE AND ENGINEERING

submitted in the School of Computing Science and Engineering of Galgotias University, Greater Noida, is an original work carried out during the period of JULY-2021 to DECEMBER-2021, under the supervision of **Ms.Garima Pandey**, **Department of Computer Science and Engineering of School of Computing Science and Engineering , Galgotias University, Greater Noida**

. The matter presented in the project has not been submitted by me/us for the award of any other degree of this or any other places.

Abhishek Yadav-19SCSE1010306

This is to certify that the above statement made by the candidates is correct to the best of my knowledge.

Supervisor Name
Ms.Garima Pandey

CERTIFICATE

The Final Thesis/Project/ Dissertation Viva-Voce examination of **Abhishek Yadav, 19SCSE1010306** has been held on 17/11/2021and his/her work is recommended for the award of BACHELOR OF TECHNOLOGY IN COMPUTER SCIENCE AND ENGINEERING.

Signature of Examiner(s)   **Signature of Supervisor(s)**

Signature of Project Coordinator   **Signature of Dean**

Abstract:

In the recent times, terrorism has grown in an exponential manner in certain parts of the world. This enormous growth in terrorist activities has made it important to stop terrorism and prevent its spread before it causes damage to human life or property. With development in technology, internet has become a medium of spreading terrorism through speeches and videos. Terrorist organizations use the medium of the internet to harm and defame individuals and also promote terrorist activities through web pages that force people to join terrorist organizations and commit crimes on the behalf of those organizations. Web mining and data mining are used simultaneously for the purpose of efficient system development. Web mining even consists of many different text mining methods that can be helpful to scan and extract relevant data from unstructured data. Text mining is very helpful in detecting various patterns, keywords, and significant information in unstructured texts. Data mining and web mining systems are used for mining from text widely. Data mining algorithms are used to manage organized data sets and web mining algorithms can be helpful in mining and extracting from unstructured web pages and text data that is available across the web. Websites built in different platforms have varying data structures and that makes it quite difficult to read for a single algorithm.

I.   Introduction

Terrorist organizations are using the internet to spread their propaganda and radicalize youth online and encourage them to commit terrorist activities. To reduce the online foot print of such harmful websites we need to create a system which detects specific keywords in that particular website and if those keywords are found then that website should be blacklisted. Data mining as well as web mining are used together at times for efficient system development [1]. Web mining also consists of text mining methodologies that allow us to scan and extract useful content from unstructured data. Text mining allows us to detect patterns, keywords and relevant information in unstructured texts
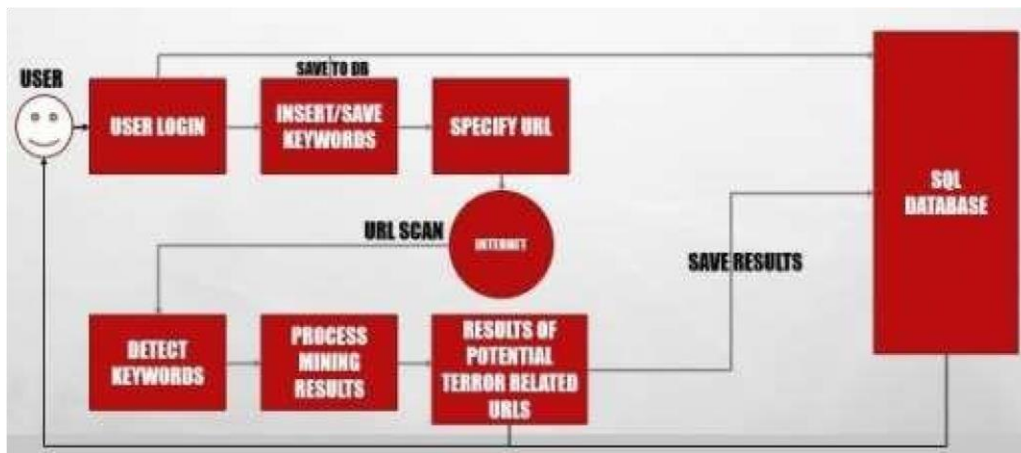
II. Problem Statement

Our system will detect patterns, keywords and relevant information in unstructured texts in web page using web mining as well as data mining . Our system will mine webpage using web mining algorithm to mine textual information on web pages and detect those web pages that are relevant to terrorism. Data mining as well as web mining is used together at times for efficient result.

III. Existing Solutions

Traditionally, there was no such system to keep an eye on various websites or any suspicious words present online. Cops were unable to track the terrorist related website or any person with suspicious information. The ratio of terrorism is high in today's world. There must be a system to track those suspicious word online and bring down the ratio of terrorism.

IV. Model Of Our System

4.1 Hardware and Software Requirements

The Project is loaded in Visual Studio 2010. We used Visual Studio for Design and coding of project. Created and maintained all databases into SQL Server 2008, in that we create tables, write query for store data or record of project.

4.1.1 Hardware Requirement: i3 Processor Based Computer or

higher Memory: 1 GB RAM

Hard Drive: 50 GB Monitor

Internet Connection

4.1.2 Software Requirement: Windows 7 or higher Visual studio

2010. SQL Server 2008.

V. Modules and their Descriptions

The system comprises of 5 modules as follows:

• Login:Here, the admin or the authorized person need to enter the login credentials. • Add KeywordsAfter successful login, admin can add various keyword which specifies terrorism. • Check WebsiteHere, admin can add multiple URL's to scan the website for any suspicious word.

• View/Check all WebsitesAll the URL's which are entered by the admin are listed here. Can check the suspicious keyword.

• Update PasswordSystem allows admin or the authorized person to update their password.

VI. Methodology

We use web mining algorithms to mine textual information on web pages and detect their relevancy to terrorism.[3] Websites created in different platform can be tracked using this application. This system will check web pages whether a webpage is promoting terrorism. This system will classify the web pages into various categories and sort them appropriately. There are two features used in this system that is data mining and web mining. 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[4]. Web mining also consists of text mining methodologies that allow us to scan and extract useful content from unstructured data. This System are used only by the government officials who work for country security. System will help the cops to easily track the susceptible community who are held in terrorism.

Website will have following characteristics:

- Load Balancing:Since the system will be available only the admin logs in the amount of load on server will be limited to time period of admin access.

- Easy Accessibility:Records can be easily accessed and store and other information respectively.

- User Friendly:The Website will be giving a very user-friendly approach for all user. • Efficient and reliable:Maintaining the all secured and database on the server which will be accessible according the user requirement without any maintenance cost will be a very efficient as compared to storing all the customer data on the spreadsheet or in physically in the record books.

- Easy maintenance:Web Data Mining for Terrorism Analysis website is design as easy way. So maintenance is also easy.

VII. **Applications**

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

- Websites created in different platform can be tracked using this application. This system will check web pages whether a webpage is promoting terrorism.

- This system will classify the web pages into various categories and sort them appropriately. There are two features used in this system that is data mining and web mining.

- 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. Data mining as well as web mining are used together at times for efficient system development.

- System will track web pages that are more susceptible to terrorism and will report IP Address to the user who is using the system.

- This System are used only by the government officials who work for country security. System will help the cops to easily track the susceptible community who are held in terrorism.

VI. CONCLUSION AND FUTURE SCOPE

To curb the menace of terrorism and to destroy the online presence of dangerous terrorist organizations like ISIS and other radicalization websites. We need a proper system to detect and terminate websites which are spreading harmful content used to radicalizing youth and helpless people. We analysed the usage of Online Social Networks (OSNs) in the event of a terrorist attack.

We used different metrics like number of tweets, whether users in developing countries tended to tweet, re-tweet or reply, demographics, geo-location and we defined new metrics (reach and impression of the tweet) and presented their models. While the

developing countries are faced by many limitations in using OSNs such as unreliable power and poor Internet connection, still the study finding challenges the traditional media of reporting during disasters like terrorist's attacks. We recommend centres globally to make full use of the OSNs for crisis communication in order to save more lives during such.

VII. REFERENCES

Journal Papers:

[1] Aakash Negandhi, Soham Gawas, Prem Bhatt , Priya Porwal "Detect Online Spread of Terrorism Using Data Mining".IOSR Journal of Engineering Volume 13,17 April 2019. So here they propose an efficient web data mining system to detect such web properties and flag them automatically for human review. Keywords: Anti-Terrorism, Data Mining, Online, Terrorism,World

[2] Avishag Gordon "The spread of terrorism publications: A database analysis",Terrorism and Political Violence journal publishedin Dec 2007.This research note focuses on the spread of terrorism publications from 1988 to 1995 compared to their frequency of appearance from 1996 to 1998. It also identifies the core journals of this research field.