International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 07, July-2022

Detecting Online Spread of Terrorism on Twitter Using Machine Learning

Kokane Saurav Bhairu
Student,
Shree Ramchandra College of Engineering,
Pune

Mane Shidhaji Vijaykumar
Student,
Shree Ramchandra College of Engineering,
Pune

Kshitij Kanojiya
Student,
Shree Ramchandra College of Engineering,
Pune

Chormale Pooja Balaso
Student,
Shree Ramchandra College of Engineering,
Pune

***Abstract*: Detecting terrorist related content on social media is a problem for law enforcement system due to large amount of data. This work is aiming at detecting tweets that are supporting terrorism or sent by terrorist orgs. To do this we use machine learning approach where we make use of two set of features: data dependent features and data independent features. The data dependent features are features that are heavily influenced by the specific dataset while the data independent features are independent of the dataset and can be used on other datasets with similar result. By using this approach, we hope that our method can be used as a baseline to classify violent extremist content from different kind of sources since data dependent features from various domains can be added. In this system, we have developed a front-end system for real-time viewing of the tweets from twitter that are detected using this system. We also have compared the performance of the two different machine learning classifiers, Support Vector Machine (SVM) and Multinomial Logistic Regression and found that the first one works better than the second one.**

Keywords: Social media, Terrorism, twitter, Extremism, machine learning, real-time tweets.

INTRODUCTION:

Nowadays, we have a lot of means of communication. The era called as era of internet.

Internet is highly dynamic means of communication. We can use internet for many reasons like to get any type of information, entertainment, business, microblogging, social media and many all. People also like to share their views on internet. For that they use social media. Nowadays microblogging is the popular ways to share views.

Twitter is most popular microblogging platform. Twitter is most useful way of interaction. Twitter has 380 million regular users and there are 500 million tweets everyday.

But everything has their pros and cons. As we use these platforms for share our views and to tell people what we are doing in our daily lives, some people use them to share their bad intent. Terrorists are one of these people. They use twitter for promote their bad views, to spread fear among people. They also use twitter for recruitment as well as fund raising.

There is our social responsibility to stop this. So, we are building this web app using real time tweets for detecting extremist, radical, elementary and terrorist supporting tweets. And to report and share this information to authorized body to take action.

Akshay Karale made project on Framework for analyzing real time activities on Twitter main features ware countering terrorism and protect human rights as manually reinforcing. Countering violent extremism and radicalization that lead to terrorism. They used database produced by bye Twitter API. Used algorithm SVM and logistic regression.

Mariam Nouth developed a project named Understanding radical mind on twitter. They mostly used magazines to find data about terrorism. They used term frequency inverse frequency bit trigger and word embedding language model for differentiating tweets. Use of feature engineering was significance in model building. They use NLP for tokenizing tweets and understanding sentiment of tweet. They used pro-ISIS Paris2015 attack database from Kaggle databases. There was no real time tweet validation in this project.

Rupali Patil, the procedure of sentiment analysis and its visualization is explained in detail concerning the topic Article 370. Sentiment analysis and opinion mining require detailed knowledge of how twitter and its python client Tweepy works to obtain the results. Python libraries like matplotlib and pandas are also used for simpler analysis and visualization of the tweets acquired. As a whole, from this paper it can be concluded that Pakistan is comparatively more concerned about the impact on its trade and has been somewhat more cynical when it comes to the sentiments of its Twitter users whereas India, on the other hand, is more concerned about increase in terrorism with a slightly positive attitude towards the revocation from the country’s Twitter users.

M. Ashcroft made an attempt to detect jihadist messages from Twitter. They used sentiment analysis to detect if a message supports ISIS or not. They used some keyword to extract tweets from twitter feed. The advantage of this work is it uses three different features such as Time-based features, Sentiment based features and Stylometric features to detect jihadist text. They got almost 90% accuracy using these features. There is a limitation and that is, they didn’t used real time validation of tweets.

Md. Abrar and team of Chittagong university et al. [4] developed a framework for detect real time tweets to detect terrorist activities. They use logistic regression and SVM to analyze tweets. They use tweepy model to create their dataset from real time tweets. They do classification of tweets as they are terrorist supporting or not. They The limitation of this project was, it could not classify terrorist related tweets and terror non supporting tweets.

M. Trupti used Natural Language Processing (NLP) and Machine Learning Technique for sentiment analysis. The advantage of their procedure is they analyzed real-time tweets using Twitter Stream API. The limitation of this study is, they used the word of sentence individually rather than analyzing the sentence individually. So, the semantic meaning was neglected which is present between words. They also failed to analyze if a user's tweet contain sarcasm or they really mean it.

Lee S. compared four text mining methods: Latent Semantic Analysis (LSA), Probabilistic Latent Semantic Analysis (PLSA), Latent Dirichlet Allocation (LDA), and Correlated Topic Model (CTM) using topic model and spam filtering. They concluded that PLSA shows the highest performance and next to LDA, CTM, and LSA in order. One of the limitations of this study is they only considered statistical approach and didn't extend their study to syntactical and morphological approach.

A work on topic discovery based on text mining techniques was presented by Pons P. They proposed a hierarchical clustering algorithm that combines partitioned and agglomerative approaches to produce topic hierarchies. They considered document place, time reference, and textual contents. This resulted in less time complexity while detecting a new topic. The accuracy of their proposed method is not satisfactory.

To limit the reach of cyber-terrorists, several private and governmental organizations are policing online content and utilizing big data technologies to minimize the damage and counter the spread of such information. For example, the UK launched a Counter Terrorism Internet Referral Unit in 2010 aiming to remove unlawful Internet content and it supports the police in investigating terrorist and radicalizing activities online. The Unit reports that among the most frequently referred links were those coming from several OSNs, such as Facebook and Twitter. Similarly, several OSNs are constantly working on detecting and removing users promoting extremist content. In 2018, Twitter announced that over 1:2 million accounts were suspended for terrorist content.

Realizing the danger of violent extremism and radicalization and how it is becoming a major challenge to societies worldwide, many researchers have attempted to study the behavior of pro-extremist users online. Looking at existing literature, we find that a number of existing studies incorporate methods to identify distinguishing properties that can aid in automatic detection of these users. However, many of them depend on performing a keyword-based textual analysis which, if used alone, may have several shortcomings, such as producing a large number of false positives and having a high dependency on the data being studied. In addition, it can be evaded using automated tools to adjust the writing style.

Another angle for analyzing written text is by looking at the psychological properties that can be inferred regarding their authors. This is typically called psycholinguistics, where one examines how the use of the language can be indicative of different psychological states. Examples of such psychological properties include introversion, extroversion, sensitivity, and emotions. One of the tools that automates the process of extracting psychological meaning from text is the Linguistic. Inquiry and Word Count (LIWC) [9] tool. This approach has been used in the literature to study the behavior of different groups and to predict their psychological states, such as predicting depression [10]. More recently, it has also been applied to uncover different psychological properties of extremist groups and understand their intentions behind the recruitment campaigns.

Enghin Omer, in his very interesting research Using machine learning to identify jihadist messages on Twitter analyzed that automatic approach for detection of terrorism. He had used Machine Learning approach to do this. This was a very intelligent project. He Had a very wide Scope for his project. We use a machine learning approach that classifies a tweet as radical or non-radical and our results indicate that an automated approach to aid analysts in
their work with detecting radical content on social media is a promising way forward.

In a study using sentiment analysis of tweets was conducted. They classify a tweet as being negative, neutral or positive. Some of the features are based on the polarity of words. This is determined by using several dictionaries like Dictionary of Affect in Language (DAL) or WordNet which assigns each word a pleasantness score between 1 (negative) and 3 (positive). Other features include counting features (counting the number of positive or negative words) and presence of exclamation marks and capitalized text. The polarity of words features, counting and presence of exclamation marks and capitalized text form what they call senti-features. In their experiments using a SVM classifier and unigram features they get 71.36% accuracy. On the other hand when unigram features are combined with senti-features the result increases to 75.39% showing the contribution of the senti-features for tweets sentiment classification.

Twitter provides a list of most popular topics people tweet about known as trending topics in real time but it is often hard to understand what these trending topics are about. In Twitter trending topics are classified into 18 different categories like sports, politics, technology etc. A bag-of-words approach for text classification is used. For each topic, a document is made from trend definition and varying number of tweets.

With the development of machine learning, it has gradually been applied to the analysis of extremist content and sentiments. Ferrara et al. [5] applied machine learning techniques on social media text to detect the interaction of extremist users. The proposed system has experimented on a set of more than 20,000 tweets generated from extremist accounts, which were later suspended by Twitter. The main emphasis was on three tasks, namely: (i) detection of extremist users, (ii) identifying users having with extremist content, and (iii) predicting users' response to extremists' postings. The experiments are conducted in two dimensions, i.e. time-independent and real-time prediction tasks. An accuracy of about 93% is achieved with respect to extremist detection.

With the same purpose, a machine learning-based technique is proposed by [7] for classifying of extremist affiliations. The Naïve Bayes algorithm is applied with the classical feature set. The system is based on the classification of user reviews into positive and negative classes with less focus on identifying, which sentiment class (positive or negative) is associated with extremist communication. In contrast to Ferrara et al. [5] work, which mainly emphasizes on the classification extremist's affiliations on skewed data; their method applies NB algorithm on balanced data giving more robust results. However, the overall dependencies in the sentence are not considered. This issue can be handled by applying deep learning models based on word embedding features. Hybrid approach for developing sentiment-based applications have received considerable attention of researchers in different domains, such as business, health-care and politics. In such approaches, different features of supervised, unsupervised and semi supervised techniques are adopted.

In the context of extremist affiliation classification, Zeng et al. worked on the Chinese text segmentation issue in terrorism domain using a suffix tree and mutual information. The core module uses mutual information and the suffix tree for manipulating data in terrorism domain. The technique has applicability for processing huge amount of Chinese textual data.

Analyzing militant conversation from online conversation, Prentice et al. investigated the intents and content generated during the militant's conversation on social media with respect to Gaza violence in 2008/2009. Over 50 online text conversations were analyzed by applying both qualitative and quantitative techniques. Their proposed system includes a manual coding approach to detect the presence of a persuasive metaphor and semantics of the underlying text.

The aforementioned studies on detection and classification of social media-based extremist affiliations have used different approaches, such as supervised machine learning, an unsupervised technique like lexicon-based and clustering-based, and hybrid models. However, there is a need to investigate the applicability of state-of-the-art sentiment-based deep learning models for classifying extremist affiliations using social media content.

Lisa K. studied to classify tweets as being multiplier of jihadism. They used machine learning to build a classifier that can analyse twitter to find multiplier of jihadism. They used adaboost classifier to train a model. They analyse both Arabic and English tweets. They obtain 84% accuracy for Arabic tweets and 98% accuracy to English tweets. They didn't test their model in real time environment.

V.Wisdom studied well on twitter data analysis. They used python for their analysis of twitter data. They used python library called tweepy for analyzing tweets. The advantage of this study is they describe numerous analysis on tweets such as term frequencies bigrams terms , most used hashtags, most used mentions, the limitation of the study on basic analysis of twitter data. No advance techniques like ML,NLP, sentimental analysis were not used.

Kathy L. classified twitter trending topic into different categories as the list of trending topics provided by twitter is often hard to understand what these trending topics are about. They classified the twitter trending topics into 18 different categories like politics, Technology, sport etc. They used a bag of words approach using TF-IDF (Term Frequency-Inverse Document Frequency) which provides how important a word is in document. The best accuracy is obtained from using Naïve Bayes Multinomial classifier (63.36%). It performed better than Naïve Bayes(43.31%) and SVM (61.76%).

Kocharekar and Jadhav treated the detection of terrorist activities using sentiment analysis in a distributed system. For this end, they proposed a web API based architecture for collecting data from social networks, using Apache Hadoop, to store large amounts of data in real time and a processing language based on artificial intelligence, for the extraction of meaning from text and identifying the emotions of publications on these social networks.

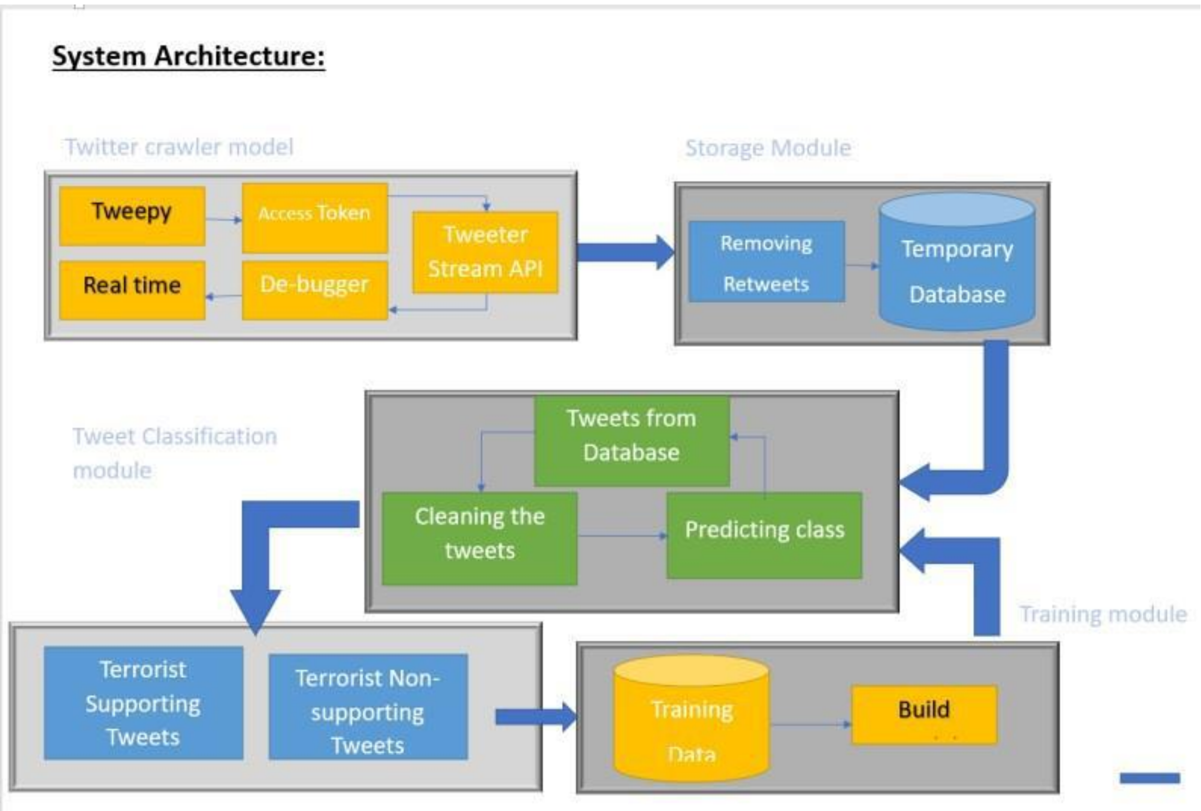
Conde-Cespedes et al. presented a classification method based on sentiment analysis and Word2vec, to detect accounts that conduct terrorist propaganda. To do so, they applied their analysis to a set of tweets collected during and before the terrorist attacks in Paris, in November 2015. They started by identifying common words and phrases used by PRO-ISIS. Then, they applied Word2vec for word representation. Finally, they used sentiment analysis and the words analyzed in the previous steps to define some characteristics to build their classifiers and detect a suspicious account.

Orero and Ngoge [9] developed a model that can help establish patterns associated with terrorist activities, using sentiment analysis of Twitter data. For this, after the classification of tweets, based on dictionary wordlist, they used an algorithm to map terrorist activities in real time.

Conde-Cespedes et al. presented a classification method based on sentiment analysis and Word2vec, to detect accounts that conduct terrorist propaganda. To do so, they applied their analysis to a set of tweets collected during and before the terrorist attacks in Paris, in November 2015. They started by identifying common words and phrases used by PRO-ISIS. Then, they applied Word2vec for word representation. Finally, they used sentiment analysis and the words analyzed in the previous steps to define some characteristics to build their classifiers and detect a suspicious account.

In the authors proposed a method to automatically detect the radical content published by terrorist groups on Twitter. For this, they used a supervised automatic learning approach, which can classify tweets as radical or non-radical. To achieve this, they collected pro ISIS tweets from known accounts, then they cleaned them, and finally they determine three types of features: stylometry-based functions, time and features based on sentiment. The algorithms used to classify the tweets are: SVM, Naive Bayes and AdaBoost.

SYSTEM ARCHITECTURE:

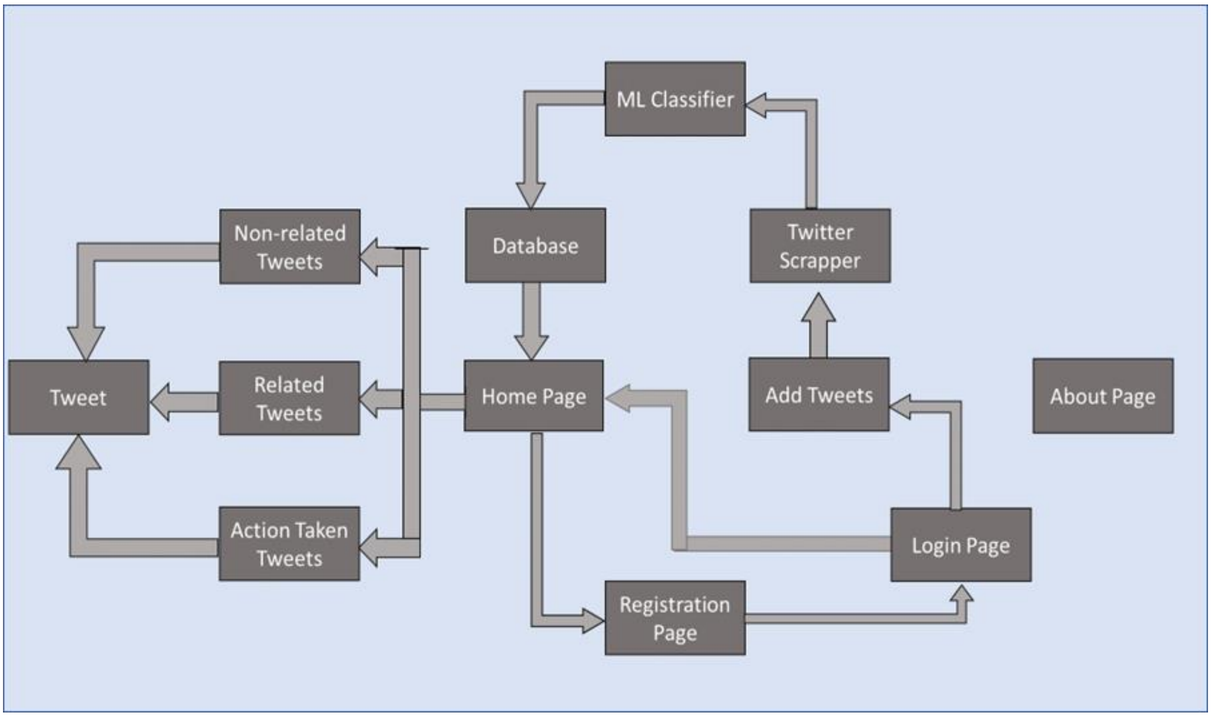


The system architecture of Twitter Terrorism Detection Framework comprises five basic modules:
1) Twitter data crawler module,
2) Storage module,
3) tweet classification module,
4) Output module and
5) Training module

Home Page: On this page we have displayed all tweets with their type. Every tweet has attributes like user of tweet, author of tweet, text of tweet and type of tweet. In this author name has link to single post where we have links to take actions like delete and report tweet as well as link to update type of tweet. At last, we have given links to pages. Each page of home contains 10 tweets.

About Page: This page contains data about website.

Register Page: This page is designed to register new user. On this page we have to fill a form which have to give attributes to user like user name, email id, consumer key, consumer key secret, access token, access token secret, password and confirm password. On click on sign in, registration form will be validated. In validation it will check user name and email id is used before this registration and length of password will check and then user will be added to database and redirected to login page.



Login Page: In login page we have to fill form which have two fields named email id and password. On click on log in user will get logged in. After login button we have link to reset password. If user is not registered, there is provided link to registration of user.
Add tweets page (+): This page will add tweets to database. In this page user will authenticate to twitter developer account using logged in user's credentials. Then It will create a stream of tweets and filled into database and type of the tweet will be predicted by machine learning model and stored into database.
Twitter Scrapper: This file used to retrieve tweets from twitter and fills into a python list which will be stored in the database in route.py folders.
Account Page: This page will provide information of account when we are logged in into account. It will show your username and email id of account and will provide you right to change email id and twitter developer account credentials. It will give you permission to change profile picture of your account.
Logout Page: This is the link to logout from website. That means current_user variable will be null. And user will be restricted to see tweets only. He/she will not have permission to modify or monitor anything in the website. When we entered to logout page, there will be change in the navigation bar.

When we open website, we are on homepage, if we have not inserted any tweets in the we will see blank content block. For doing any action, we firstly register user. Firstly, we fill all information and click on sign in our user will be added to User table in database. Then we will be redirected to login page. On login page will enter email id and password and logged in into application, that means our user will be stored in current_user object. Then we will get access to add and delete or report tweets as required. Firstly, we add tweets to database by clicking on + sign on navigation bar. Then content block will be filled with tweets in the database. When we click + tab consumer key, and consumer secret of current user will be loaded. And by using OAuthHandler, this user will be authenticated to twitter developer account. Then twitter developer API object will be created. And some tweets will be retrieved and filled into database and on the basis of text of tweet, type of tweet will be predicted and filled into database. These tweets will be displayed on content block.
We have provided pagination to each page of content block so that it will show only a few fixed numbers of tweets on page. At last, we have provided links to pages so that we can access pages by clicking on those page numbers.
When we click on the author's name of the particular tweet, we will be redirected to tweet page, on which we can only see that tweet. If we are authenticated user, with tweet data we will see two extra tabs named update and delete tab. Actually, delete tab is used to delete the tweet from main stream and automatically a report will be given to twitter using twitter API. Update tab is used to change type of tweet and again send to machine learning model to train it for better performance.
When we click related tab, all tweets will be filtered according to tweet type equal to "related" and stored into tweets variable, and then displayed using .html file.
When we click Extremism tab, all tweets will be filtered according to tweet type equal to "supported" and stored into tweets variable, and then displayed using .html file.

When we click on Action taken tab, we will be shown tweets on which we have taken any action like update type of tweet or add user of the tweet to blacklist. But these actions are possible only when the client of our website is logged in. Other client only can see tweets with their type.

When we log in to application, navigation bar will be changed at right side as +, account and logout tabs will be there instead of login and register. Account tab is clicked to see information about account. We can change their twitter developer account credentials and user's profile picture. We can change their email id. If we want to reset password, we are designing a tab for it.

EXPECTED OUTCOMES:

This project will differentiate between tweets as they are terrorism supporting or terrorism related or there is an unsupported or others. After the completion of this project, we expect that spread of terrorism on Twitter for different social media sites will be reduced for this project will help it to reduce. This project will expect that in this system or model will work efficiently to differentiate b tweets and report to Twitter as they contain any extremist thoughts.

CONCLUSION:

In this presentation tweeter terrorism detection framework to detect tweets that support terrorism from real tweet stream. Our framework will collect real time tweets by using tweeter streaming API and analyses them. It can be classified into three classes and based on category of tweet they shown in different screens of web applications.

REFERENCE

[1] A Framework for Analyzing Real-Time Tweets to Detect Terrorist Activities Akshay Karale1, Pranav Shinde1, Pushpak Patil1, Sanjay Parmar1, Prof. Niyamat Ujloomwale2 1,2Department of Computer Engineering, Dr. D. Y. Patil School of Engineering, Lohegaon, Maharashtra, India

[2] Understanding the Radical Mind:Identifying Signals to Detect Extremist Content on TwitterMariam Nouh_, Jason R.C. Nursey, and Michael Goldsmith_ _Department of Computer Science, University of Oxford, UK

[3] Detection and classification of social media-based extremist affiliations using sentiment analysis techniques

[4] Shakeel Ahmad1*, Muhammad Zubair Asghar2, Fahad M. Alotaibi3 and Irfanullah Awan4

- www.ieeexplore.ieee.org
- https://hcis-journal.springeropen.com/articles/10.1186/s13673-019-0185-6#Tab4
- Intelligent projects using python by packt.com
- https://www,tweepy.org
- https://www,devloper.twitter.org
- And many more newspapers, research papers, websites, magazines etc.