International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 02 | Feb 2020 www.irjet.net p-ISSN: 2395-0072

Design and Development of a System for Predicting Threats using Data Science

Gaurav Tripathi[1], Gaurav Rai[2], Sunny Singh[3]

[1,2,3]*Department of Computer Engineering, T.C.E.T, Mumbai*

Abstract - *Impact of Social Media on a human being is more severe and dangerous rather than their enemy as we are living in a world where social media is the second most essential part of our life, like with food and water humans cannot live, same can be said about the social media as it connects people who are living far away to those who are close to you. So making social media safe heaven is a difficult job but at-least taking a measure can be done before any chaos takes place due to social media activity could be done. As social media isn't bound by any boundaries of countries and government rules even sometimes international rules are also not enough to handle the threats, the government has many systems but not as integrated one to handle the public level threats which had already taken place in countries like India. The main aim of this research is to utilize natural language processing, data analytics, big data, machine learning ideas for the acknowledgement of upcoming threats and prediction of social media threats dependent on the real-time scenarios. Here in this paper, unsupervised natural language processing is applied to social media data, with a reason to predict, identify the people and source, responsible for spreading threats.*

Key Words: **Social media threats analysis, Threat prediction of social media, Natural Language Processing, Machine Learning, Data Analytics, Big Data, Convolutional Neural Network**

1. INTRODUCTION

Over the last few years, online communication has moved toward social-driven technologies, like social media networks such as WhatsApp, Facebook, Instagram, Twitter and other social media platforms such as blogging websites, YouTube Videos, online virtual communities, and online petition platforms such as change.org, etc. These social technologies have started a revolution in user-generated data, online global communities, and rich human behavior-related content. Understanding human preferences are important to the development of applications and platforms for predicting threats from social media data using data science. This paper discusses the role of social media data to understand the behavior of humans regarding their online activity based on data.

With the advancement in the latest technology about sentiment analysis and predictive analytics, it has opened many avenues for researchers and enterprises to understand the human mental state and their social behavior in a better way. The proposed challenge is to predict people's activities on social media, to help in eliminating and reducing the percentage of any incident (Stampede, Mob-lynching, etc.) before it's occurring. There are several existing systems which are currently used by the government of India they are listed as NETRA(Network Traffic Analysis System), CCTNS(Crime and Criminal Tracking & Networks System), Social Media Labs and where NETRA is most effective nowadays as it is used by intelligence bureau, India's domestic intelligence agency and research & analysis wing(RAW) as NETRA can analyze voice traffic passing through several social media platforms and can intercept messages with such keyword like 'attack', 'bomb', 'kill' in real-time system from the enormous amount of tweets, status updates, blogs, forums.

The purpose is to aim at making use of unsupervised natural language processing and machine learning algorithms (Clustering, Linear and Logistic regression) and data science (Data Pre-processing, Big Data, hybrid algorithm) in interpreting Social Web Data of various platforms like Facebook, Twitter, Google trends and other various sources. Start from fetching trending keyword from various social media platform and categorizing into single format and with the help of ensembling of clustering algorithm and using natural language processing to label the unstructured text for meaning and later on predict the sentiment of trending topic based on user's data & using that data our application will predict the threat level in three categories high, moderate & low region wise. Social Media Analysis will be monitoring, collecting, and analyzing the data of a Facebook page (posts, comments, likes, shares) and a Twitter profile (tweets, re-tweets, mentions, and public tweets containing one/two keywords only) and from other various sources for predicting the future or upcoming threats by taking each and every social media platforms data and categorizing them into six categories such as Toxic, Severe Toxic, Obscene, Insult, Threat, Identity Hate to predict the threat and toxicity level of that particular trending topic.

2. RELATED WORK

In the section below, the approach used has been discussed in detail.

2.1 Overview

So nowadays there are many tools and technology available in the market which can be used to analyze the social media data but most of it are using for marketing purpose that too even are not cost-effective, with the advances in technology about sentiment analysis and predictive analytics, it has opened many avenues for researchers and enterprises to understand the human mental state better. The proposed challenge is to predict people's activities on social media, to help in eliminating and reducing the percentage of any incident before it occurs. The purpose is aimed at making use unsupervised natural language processing and machine learning algorithms (Clustering, Linear and Logistic regression) and data science (Data Pre-processing, Big Data, hybrid algorithm) in interpreting Social Web Data of various platforms like Facebook, Twitter, Google trends and other various sources of machine learning algorithms (Clustering, Linear and Logistic regression) and data science (Data Pre-processing, Big Data) in interpreting Social Web Data of various platforms like Facebook, Twitter, Google trends and other various sources. Social Media Analysis will be monitoring, collecting, and analyzing the data of a Facebook page (posts, comments, likes, shares) and a Twitter profile (tweets, retweets, mentions, and public tweets containing one/two keywords only) and from other various sources for predicting the future or upcoming threats.

2.2 Algorithm

NLTK Classifier- NLTK Classifier to predict the sentiment of the comments and tweets that are either positive or negative by using a corpus dataset to cluster the data into categories.

Deep Learning Unsupervised Threat Classification algorithm using CNN and Word2vec.

CNN is a class of deep, feed-forward artificial neural networks (where connections between nodes do not form a cycle) & use a variation of multilayer perceptrons designed to require minimal preprocessing.

Word2vec - is a group of related models that are used to produce word embeddings. These models are shallow, two-layer neural networks that are trained to reconstruct linguistic contexts of words.

Pytrends is a library which is a substitute to google trends library for python to extract the data from google,youtube, images for trending topics and past information related data which were searched on google.

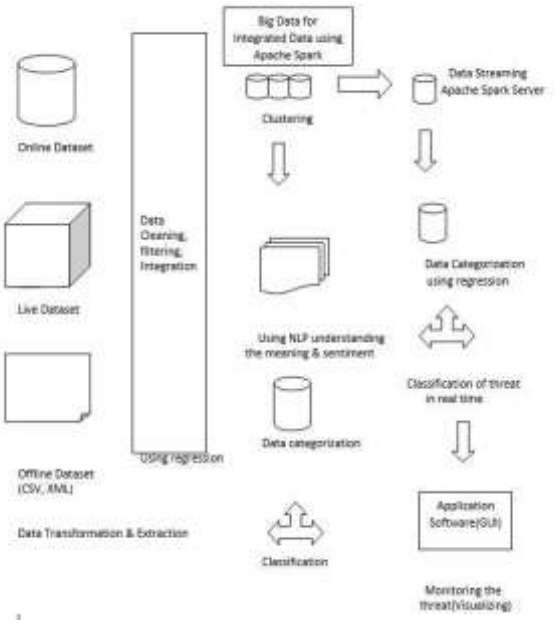

Fig I. Architecture

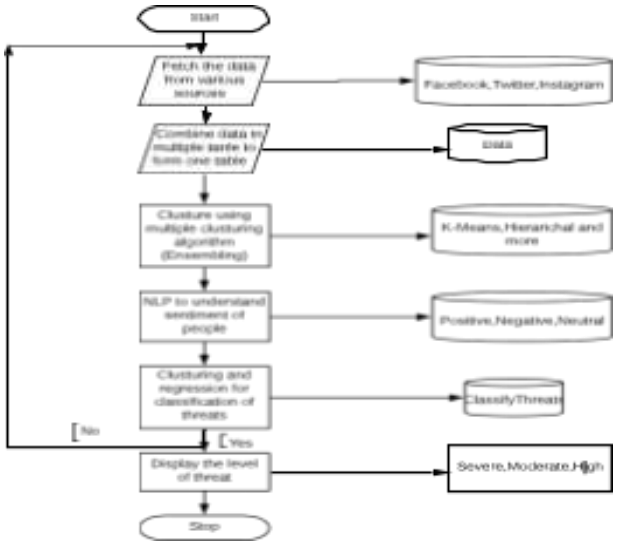

Fig II. Flowchart

2.3 Application

1. To predict the Real-time threat which is trending on social media in terms of the hashtag.
2. Individually track a detailed review of a person what kind of tweets and comments they usually do.
3. Streaming Real-time data from various sources and analyzing them for predicting the threat.
4. Real-time classification of threat in terms of low severe high depend on the threshold and population of the region

5. Region-wise classification of hashtags and tweets and further analyzing the particular hashtag to predict what kind of tweets are trending.

2.4 Expected Outcome

This project upon completion of full functionalities by having access to all social media platforms will easily predict the threat, fake news and categorizing the trending hashtag in terms of obscene, toxic, severe toxic, threat, identity hate, insult by analyzing every comments and tweet based on the region, this might help to categorize the trending hashtag or any trending topic whether they are worthy of sharing, re-tweets, comments, etc. A simple GUI will help to display the real-time threat analysis by streaming data from various social media platforms to check the threat level based on the region in three categories low, moderate and severe based on the population of that region.

3. METHODOLOGY

After the methodology being retrieving user's data from various social media platforms in structured/unstructured format most probably in JSON format using the APIs provided by these platforms. The extracted data is then cleaned and filtered for our specific requirements. The data then basically streamed to the client-side application using the SPARK Streaming and Socket Stream program with the help of Flask. Flask is a lightweight WSGI web application framework. It is designed to make getting started quick and easy, with the ability to scale up to complex applications. It began as a simple wrapper around Werkzeug and Jinja and has become one of the most popular Python web application frameworks. The main methodology for Threat Analysis is the Deep learning NLP Classification Unsupervised model using convolution 2D CNN for text classification method is used to train the dataset. And using NLTK Corpus data for sentiment analysis for the range of sentiments of people in terms of positive and negative statements. Every threat is classified in three categories low, moderate and severe based on classification of every tweet and comments into several terms like Toxic, Severe Toxic, Obscene, Insult, Threat, Identity Hate and by counting these and every trending can be tagged whether it is considered to re-tweet, repost, and comment.

4. TESTING ENVIRONMENT

In our research, we focus on a collection of data from various social media platforms such as Facebook, Twitter, and various other sources. In comparison with previous work, this paper focuses on more integrated retrieval of data and real-time analytics and prediction on the streaming data in a batch.

At first, the subject of this paper is the detection of hashtags, comments, trending topics on several social media platforms and posts which are posing threats in terms of "toxic", "insult", "identity hate", "obscene", "threat", "identity hate" to society in realtime .

Secondarily, in this paper, we utilized the 30% dataset of toxic comment challenge and which were published on Kaggle and 70% of data from various other platforms by categorizing them individually into threat category of several languages as every state in india has different context of writing the text, to identify the above-mentioned terms of the comments, posts, and tweets this may help us to achieve and identify the hashtags, trending topics and post which are posing threat on individual, society, and business in real-time.

5. CONCLUSIONS

Threat Analysis of Social Media Data of various platforms will be helpful to society as social media is the second most important aspect of human being so making it safe is a difficult task but somehow a particular security and safety in terms of security measures can be provided by achieving certain terms like trending hashtags, trending topics in terms of particular categories such as toxic, severe toxic, obscene, insult, threat, identity hate which could prevent users from commenting, tweeting, re-tweeting, reposting so many of the threats can be avoided just by tagging them as these will be implemented region-wise if it makes easier, efficient and using real-time prediction it can become one of the most accurate, efficient and effective ways to provide security to society and help reinforcement departments of countries and help people to understand the perspective and objective of trollers, criminals, fake news agencies and political interferences. This project provides a simple GUI which makes it easier to predict and analyze the threat of various platforms individually and predict the threat as well. This system is an integrated, real-time data processing of various social media platforms in an efficient way to detect the threats which are posed by social media.

ACKNOWLEDGEMENT

We wish to express sincere thanks to Mr. Vijay Jain and Dr. Anand Khandare for helping in our research and our Computer Engineering department for arranging such a platform. so that we can research all about Social Media Analytics from several platforms. We also extend our heartfelt thanks to our colleagues, family members, and well-wishers.

REFERENCES

[1] Michael Fire, Roy Goldschmidt, and Yuval Elovici 2014. Online Social Networks: Threats and Solutions(2014)

[2] Jennifer Golbeck, Zahra Ashktorab, Rashad O. Banjo, Alexandra Berlinger, Siddharth Bhagwan(2017, A Large Human-Labeled Corpus for Online Harassment Research(2017).

[3] Miltiadis Kandias, Vasilis Stavrou, Nick Bozovic, Dimitris Gritzalis 2013.Proactive Insider Threat Detection Through Social Media: The YouTube Case (2013)

[4] Racha Ajami, Noha Ramadan, Nader Mohamed, Jameela Al-Jaroodi(Security Challenges and Approaches in Online Social Networks 2011)

[5] Yakshi Sharma Dept of I.T, U.I.E.T Panjab University, Chandigarh, 160014 India Veenu Mangat2015. Mandeep Kaur practical approach to Sentiment Analysis of Hindi tweets(2015)

[6] Aristidis G. Vrahatis, Sotiris K. Tasoulis, Vassilis P. Plagianakos, Aristidis G. Vrahatis 2018.Convolutional Neural Networks for Toxic Comment Classification (2018)

[7] Weimin Luo, Jingbo Liu, Jing Liu 2009.An Analysis of Security in Social Networks (2009)

[8] Soumya.T.R, S. Revathy 2018. Survey on Threats in Online Social Media (2018)

[9] Yojana Goyal, Anand Sharma 2019. A Semantic Approach for Cyber Threat Prediction Using Machine Learning(2019)

[10] Anni Sapountzi, Kostas E. Psannis (Social Networking Data Analysis Tools & Challenges 2016)

[11] Harri Jalonen (Negative emotions in social media as a managerial challenge 2014)

[12] Natalya (Natalie) Bazarova, Yoon Hyung Choi, Victoria Schwanda Sosik, Dan Cosley (Social Sharing of Emotions on Facebook 2015)