

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 9, Issue 1, January 2020

A Survey on Rumour Detection and Emergency Response System Using Sentiment Analysis

Veeralahari Dandu¹, Pratiksha Chavan², Mohini Gore³, Pooja Chavan⁴, Shital Sungare⁵

U.G. Student, Department of Computer Engineering, DIT-Pimpri, Pune, Maharashtra, India¹⁻⁴

Assistant Professor, Department of Computer Engineering, DIT-Pimpri, Pune, Maharashtra, India⁵

ABSTRACT: Social media generates huge amount of sentiment rich data in the form of tweets, status updates and blog posts etc. In Sentiment analysis, this user generated data is very useful in knowing the opinion of the crowd. Twitter sentiment analysis is hard as compared to general sentiment analysis due to presence of slang words and misspellings. Knowledge base approach and Machine learning approach are the well known strategies used for analysing the text. Public and private opinion about vast variety of subjects are expressed and spread continuously through numerous social media. Twitter is one of the popular social media that is used nowadays. Twitter offers a fast and effective way to analyse customers' perspectives toward the crucial to success in the market place. This project uses knowledge base including various patterns for tweets along with multiple strategies to detect the sentiment expressed in a tweet and if a tweet is genuine or not. During natural or man made disasters, humanitarian organizations response look for useful information to support their decision making process. Based on fake news detection analysis, the humanitarian organizations provides help assistances in emergency situation.

KEYWORDS: Fake News, Twitter Data, Data Processing, NLP, Pattern Recognition, Response System.

I. INTRODUCTION

The Twitter has emerged as a major micro-blogging website, having over 100 million users generating over 500 million tweets every day. With such large audience, Twitter has consistently attracted users to convey their opinions and perspective about any issue, brand, company or any other topic of interest. Due to this reason, Twitter is used as an informative source by many organizations, institutions and companies. On Twitter, users are allowed to share their opinions in the form of tweets, using only 140 characters. This leads to people compacting their statements by using slang, abbreviations, emoticons and short forms. Along with this, people convey their opinions by using sarcasm and polysemy. Hence it is justified to term the Twitter language as unstructured. To extract sentiment from tweets, sentiment analysis is used. The results from this can be used in many areas like analyzing and monitoring changes of sentiment with an event, sentiments regarding a particular brand or release of particular product and analyzing public view of government policies.

A lot of research has been done on Twitter data in order to classify the tweets and analyze the results. In this project we aim to predict the sentiments from tweets by checking the polarity of tweets as positive, negative or irrelevant. Sentiment analysis is a process of deriving sentiment of a particular sentence. It is a classification technique derives opinion from tweets and formulates a sentiment and on the basis of which, sentiment classification is performed. Sentiments are subjective to the topic of interest. We are required to formulate that what kind of features will decide for the sentiment it embodies. In the programming model, sentiment were for to, is class of entities that the person performing sentiment analysis wants to find in the tweets. The dimension of sentiment class is critical factor in deciding the efficiency of the model. For example, we can have two-class tweet sentiment classification (positive and negative) or three class tweet sentiment classification (positive, negative and irrelevant). Sentiment analysis approaches can be broadly categorized in two classes – lexicon based and machine learning based. Lexicon based approach is unsupervised and it proposes to perform analysis using lexicons and a scoring method to evaluate opinions. Where machine learning approach involves use of feature extraction and training the model using feature

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 9, Issue 1, January 2020

sets and dataset. The basic steps for performing sentiment analysis includes data collection, pre- processing data, feature extraction, selection of baseline features, sentiment detection and performing classification using simple computation or else machine learning approaches. Features are like Parts-of speech features i.e. nouns, adverbs, adjectives, etc. in each tweets are tagged. We for the accuracy purpose of polarity will be detecting sentiments along with various knowledge base patterns and multiple machine learning strategies to evaluate the sentiments. Alongside we will be checking if the tweet is genuine or not or if has influenced by other tweets which can be very useful in rumors mitigation on social medias. This approach will produce the higher accuracy for polarity by considering POS factor and genuineness as well as can be used in various sectors such as analyzing product reviews or government policies, etc. where it can be found if negative influence is spread and if it affects people.

II. RELATED WORK

In research areas, various approaches have evolved, which propose methods to train a model and then test it to check its efficiency. Performing sentiment analysis is challenging on News data, as we mentioned earlier.

Here we define the reasons for this:

- Limited news size: with just 140 characters in hand, compact statements are generated, which results sparse set of features
- Use of slang: these words are different from English words and it can make an approach outdated because of the evolutionary use of slangs.
- News features: it allows the use of hash tags, user reference and URLs. These require different processing than other words.
- User variety: the users express their opinions in a variety of ways, some using different language in between, while others using repeated words or symbols to convey an emotion.
- Genuine or Influenced News: the users many times are under influence of various fake news that can be spread with social media and such news will get reposted many times to influence more.

Twitter is highly used social media and our system is take twitter as input. System will process all tweets using Natural Language Processing. All tweets are preprocessed. Tokenization is done in which all sentences are separated into words. The words which are not important they are removed from it. After that sentiment analysis is done in which polarity of every word is calculated. Every word having polarity in between 1 to -1. Using textblob algorithm system will find out emergency situation. After sentiment analysis rumor detection is done in which TF-IDF algorithm and pattern matching is used. If emergency situation is real then help assistance is provide to the needy people.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 9, Issue 1, January 2020

III. SYSTEM ARCHITECTURE

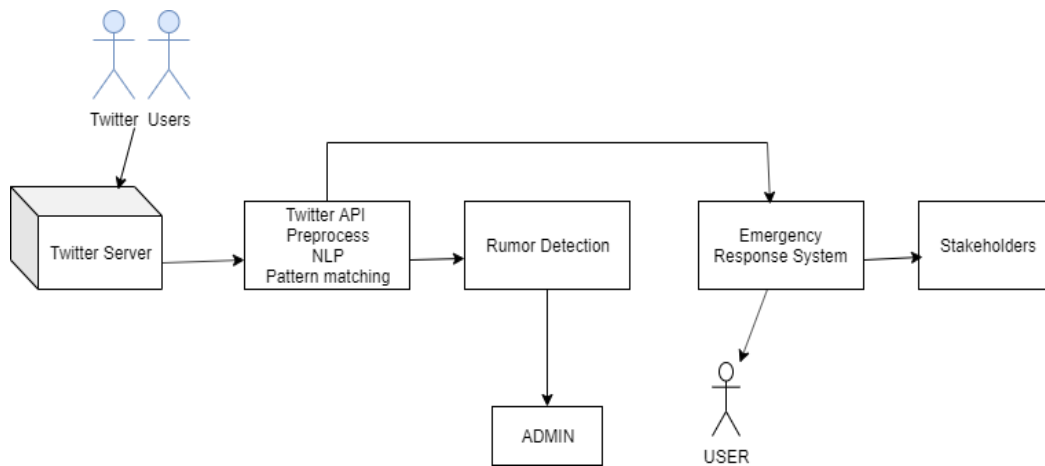


Fig. Working flow of rumour detection and emergency response system

Our system is divided into three modules :

1. GUI and mySQL connectivity.
2. Sentiment Analysis.
3. Rumor detection and Help assistance.

IV. MATHEMATICAL MODEL

- SYSTEM DESCRIPTION:

To understand issues and problems during the system fail.

Let 'S' be the system

Where,

$$S = \{I, O, P, Fs, Ss\}$$

Where,

I = Set of input

O = Set of output

P = Set of technical processes

Fs = Set of Failure state

Ss = Set of Success state

Identify the input data I1, I2, , In

$$I = \{(Twitter\ Data)\}$$

Identify the output applications as O1, O2,,On

$$O = \{(Rumors\ Detection, Fake\ News\ Detection, Sentiment\ Detection)\}$$

Identify the Process as P

$$P = \{(Data\ Processing, Natural\ Language\ Processing, Sentiment\ Analysis, Pattern\ Recognition)\}$$

Identify the Failure state as Fs

$$Fs = \{(If\ fake\ news\ not\ predicted)\}$$

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 9, Issue 1, January 2020

Identify the Success state as Ss

$$P = \{(Fake\ news\ detected\ successfully)\}$$

V. SOLUTION STRATEGY

In the proposed system, we will retrieve news from News using Twitter API based on the query. The collected news will be subjected to preprocessing. We will then apply the various patterns and strategic algorithms including some of machine learning algorithms for NLP to supervise the data. The results of the algorithms i.e. the sentiment and influence will be represented in graphical manner (pie charts/bar charts). The proposed system is more effective than existing system. This is because we will be able to know how the statistics determined from the representation of the result can have an impact in a particular field as well as influence of negativity spread by fake news.

VI. APPLICATIONS

Applications of this proposed system are :

- Fake News Analysis : In these system we will do the analysis of tweets and find out wether that specific news is fake or real using data preprocessing. Using that we clearly get idea of which news is true and which one is fake.
- Rumor Detection : Our system will also detect rumors. Algorithm which we are using into our system like TF-IDF help us in rumor detection. Sometimes people spread rumors over social media to detect that our system will help.using pattern matching we will do that in our system.
- Sentiment Analysis: Main aim of our system is to find out emergency situation and that is done by sentiment analysis. Using textblob algorithm we will find out polarity of each word and sentence. Depending upon word ranking system will detect critical situation to provide help to needy one.
- Emergency situation help assistance : System will dosentiment analysis, rumor detection using different algorithm and data processing to find out real emergency situation and provide help assistance to needy people who stuck in emergency situation.

VII.CONCLUSION

Rumors or Fake News spread on Online Social Network sometimes can cause serious negative social effects and it is usually impossible for humans to manually inspect millions of posts created everyday. Thus an automated technique to detect rumors on Online Social Network is of high practical value. Emergency response system is designed to provide help assistance during emergency situations.

REFERENCES

- [1] A. Zubiaga, M. Liakata, and R. Procter, "Learning reporting dynamics during breaking news for rumour detection in social media," arXiv preprint arXiv:1610.07363, 2016.
- [2] N. Ruchansky, S. Seo, and Y. Liu, "Csi: A hybrid deep model for fake news detection," in Proceedings of the 2017 ACM on Conference on Information and Knowledge Management. ACM, pp. 797–806,2017.
- [3] J. Gottfried and E. Shearer, "News Use A cross Social Media Platforms 2016," NUMBERS, FACTS TRENDS Shap.WORLD, 2016.
- [4] C. Silverman, "This Analysis Shows How Viral Fake Election News Stories Outperformed Real News OnFacebookSilverman, C. (2016). This Analysis Shows How Viral Fake Election News Stories Outperformed Real News On Facebook.k," BuzzFeed, 2016.
- [5] C. Silverman and J. Singer-Vine, "Most Americans Who See Fake News Believe It, New Survey Says," Buzzfeed, 2016.
- [6] C. Dewey, "Facebook fake-news writer: 'I think Donald Trump is in the White House because of me,'" The Washington Post, 2016.
- [7] N. Newman, R. Fletcher, A. Kalogeropoulos, D. a. L. Levy, and R. Nielsen, "Reuters Institute Digital News Report 2017," 2017.
- [8] N. Nic, D. A. L. Levy, and R. K. Nielsen, "Reuters Institute Digital News Report," Univ. Oxford, vol. 1, 2018.
- [9] E. C. Tandoc, Z. W. Lim, and R. Ling, "Defining 'Fake News': A typology of scholarly definitions," Digital Journalism, vol. 52, no. 1, pp. 1–17, 2017
- [10]X. Chen, G Elmes, X Ye, J Chang – GeoJournal, "Implementing a real time Twitter-based system fr resource dispatch in disaster management", 2016

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 9, Issue 1, January 2020

- [11]T Takahashi, N Igata – The 6th International Conference, Rumor detection on Twitter, 2012
- [12]S Hamidian, MT Diab, arXiv preprint arXiv:1912.08926, “Rumor detection and classification for twitter data”, 2019
- [13]Terry Traylor, Jeremy Straub, Gurmeer,” Classifying fake news articles using natural language processing to identify in-articale attribution as supervisd learning estimator” , 2019.
- [14] Neha Singh, Nirmalya Roy, Aryya Gangopadhyay”Analysing the Sentiment of Crowd for Improving the emergency responcy response services”, 2018.
- [15]Zeel Doshi, Subhash Nadkarni, Deparment of Information Technology”TweetAnalyzer : Twitter trend detection and visualization”, 2017
- [16]Sneh Paliwal, Sunil Kumar Khatri, Mayank Sharma, Amity University, ” Sentiment analysis and Prediction using Neural Network”, 2018
- [17]Sahar A El Rahman, Feddah Alhumaidi AlOtaibi, Wejdam AlShehri, ” Sentiment Analysis of Twitter data”, 2019
- [18]Bhavya Shah, Varun Agrawal, Utkarsh Dubey, Information Technology, ”Twitter Analysis for disaster management” , 2018
- [19]Quan Zou, School of Computer Information and Science, ” Prototype System using location based Twitter data for disaster management”, 2019
- [20]Roberto Interdonalo, Antoine Doucet, Jem-Loup Guillaume, ” Unsupervised crisis Information Extraction from Twitter data”, 2018