

Sentimental Classification on Social Media Text Using NLP Techniques

Suraj Prasad Keshri¹, Neelam Sahu²

¹Research Scholar, Dr. C.V. Raman University, Bilaspur (C.G.) India.

²Professors, Dept. of IT & CS, Dr. C.V. Raman University, Bilaspur (C.G.) India.

ABSTRACT

The idea of sentiment analysis has been receiving attention for the past few years. Major challenges in sentiment analysis are collection of huge data from sources, applying appropriate algorithms or techniques and splitting them into different sentiments. In this fast-expanding internet world, social media gives individuals a platform to express their views. In which we can say something. With the changing ways of things in different areas in our daily life, a person can easily express his opinion or idea. This idea does not say under any pressure but expresses itself. People tend to express themselves in their regional language or in a way convenient to them. Here social media has been created to play an important role in the personal feelings of the people. A huge amount of data gets accumulated in the network application, this data is the views and opinions of the person, based on their feelings, the best answer can be read. In this research paper, we will get the information about tweets posted by the customer to be positive, negative or neutral, for this the proposed Twitter model will first create a database of tweets from Twitter by using Twitter API, then we will write Using Text Blob, individual sentiment scores are assigned to the user's goals and using the text classification model they are classified as positive, negative or neutral.

Keywords: - Natural Language Processing (NLP), Text based classification Decision-making, Indian languages, Sentimental analysis, Social media analysis, , Text, blob Twitter.

I INTRODUCTION

At present, the rapid development of information through web applications is happening all over the world. The main problem is not to present the information on the screen but to be hungry for knowledge from the inputs. Sentiment analysis could be an important current research towards web page mining [1]. With the increasing complexity of social media, the number of people on social media has started changing rapidly. People review The number of people posting their opinion or review on the web is increasing day by day. User twitter contains a large number of sentiments and opinions. Twitter requires a set of appropriate techniques to extract data, which are helpful in forming Twitter data. Current research is built on Twitter data and from all angles of the field. Initially, user reviews are removed from Twitter, and user reviews and spoken words are used using algorithms and natural language processing techniques. based on the sentiment behind . The Highlight the importance of sentiment analysis and proof of concept. The rest of the paper is structured as follows: Part 2 highlights the survey conducted on the relevant area. That section 3 defines the expression of ideas, considers the discussion, and part 4 shows the implementation details, part 5 presents the summative results by analysis and comparison with the various works ends in part 6. Section 7 considers the discussion on the future work of the project

their own blog in their native language. During this project, the proposed work uses Kannada, Hindi, Tamil, Telugu, and Malayalam views. These reviews are written with respective opinion scores and are divided using certain techniques and algorithms [2]. Thought analysis is like a text in which the work of classification is done. In this, the data we take through social media, that model classifies a given text into two or more parts. Classes can be positive, negative or neutral.

II RELATED WORK

In this paper [3], by looking at the terms values Personal views are divided. The mathematical model simulated with the help of HADOOP proved to have 95% accuracy for the tweet. The results are displayed directly in the holding hash table. This model has been shown to be useful for the Twitter API for splitting tweets into positive or negative, which has been shown to work successfully for tweets collected in the research paper model [4] having a constant size.

III PROBLEM STATEMENT

The current time is very important in social media because now PD has commented on the main role. People mainly use social media for their entertainment purposes, and at the same time, they have also become accustomed to share their feelings. Include your opinion in those opinions which are going on in both the heart and the mind. Twitter is an online platform that has been sourced and does not contain any sentiment label. In sentiment analysis, parts of the rapid miner were performed using tools such as extensions as well as

validating and selecting the outputs from the results obtained.

In the research paper [5], the patient's prognosis is deciphered which were collected in the database records and emotional considerations are applied to the patient's thoughts. This paper used lexicon scoring and machine learning to evaluate performance using models. The proposed research model provided better understanding and visibility of patient problems and monitored health records, and observed patient response. Continuously evaluated patient records by collecting information from word cloud have gone.

The main objective [6] in this research is to find out the direction of tweets on social media. Through this research, the purpose of the author and his opinion or goal is to be determined. This part proposes fused feature vector and sentiment information. Our model works well for short Hindi text and tweets of limited length. This model is tested and classified with a set of classified techniques that also includes a Support Machine Vector (SVM) algorithm. Which tests the system with different functions and gives better results?

In the research paper [7] a technique was created to forecast the stock market in the Indian market which uses sentiment analysis to obtain the score from various social platforms of the research and various news channels.

The method in this technique provides accurate results. It had about 3.06 percent error which was less as compared to other methods. It is a technique based on algorithmic market share market data which helps people to make and make wise decisions while investing in share market.

This method and methodology of applying thought analysis has achieved a great level in recent times. Depending on the polarity of the individual, it is necessary to have classy principles for different types of emotions. The current method mainly draws attention to tweets and model text on various platforms.

In this research paper 8 it was observed that a person's mood style plays an important role in emotion. This study integrates Twitter text and thought to improve Twitter sentiment output [8].

In this research [9], the author treated emoji with texts, whether positive, negative or neutral, to derive the author's mood. It is often observed that the emoji play an

important role in the text to identify the author's self-reliance or feelings.

IV METHODOLOGY

Linguistics in Natural Language Processing (NLP) as a part of linguistics, it mainly collects on the interaction between the user's main language and the computer and considers those who are focused on this conversation and using a large amount of regional terms. Analyzes data. We review how to separate the pieces of data taken into positive, negative or neutral and how we analyze it from natural language databases and other mediums used in this research, mainly on political topics. Usually we focus is used to analyze the words of a phrase or a series of sentences.

(a) **Sentiment analysis in python with Twitter data:** - After the completion of the first step, the software is sent by the input data to the model. Then analyzing the polarity of the user and the individual we can obtain the emotional score in our structure. The polarity of text comments will be output as floating points within the range [-1.0,1.0]. . When the polarity value is less than 0, then the value of the statement will be negative and if the polarity value is greater than 0 then the value of the statement is obtained as positive, if the value is equal to 0 then the resultating value is not neutral. This happens when we comment Nuteral and Objectivity [0.0,1.0] when it falls in the range of 0. 0 represents a different objective term and 1.0 represents individual expressions. The proposed work illustrates the structure and character of NLP, the approach to the review obtained based on a keyword head or hash tag is shown in Figure 1.

- (i) **Twitter Authentication:-** To extract Twitter data at the current time, we require the Twitter API skew and hash tags: Consumer key
- Consumer secret
 - Access to ken and
 - Access to ken secret

These four classes are used to edit Twitter's; these sections can be analyzed by creating an application on Twitter. Using Required Sections Twitter's data requires a developer's data section.

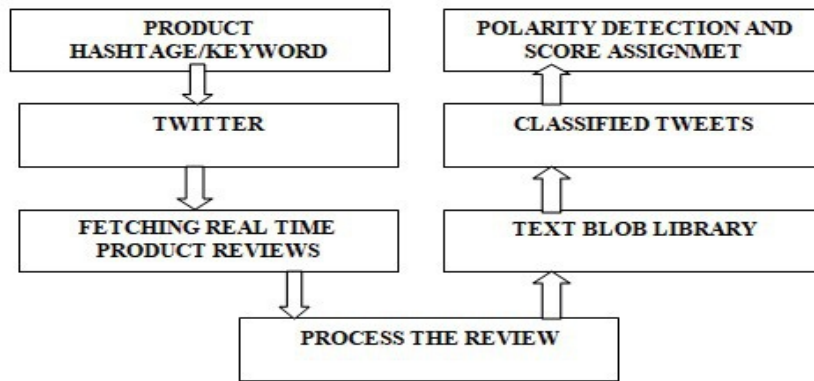


Fig.1. architecture of the proposed work.

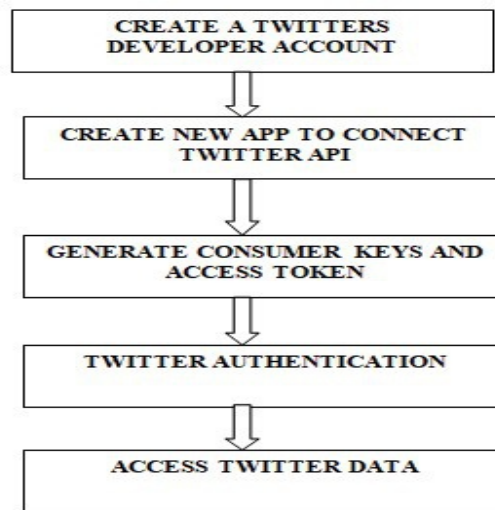


Fig.2. Twitter Authentication process

V TWITTER DATA COLLECTION

This research focused entirely on Twitter data because social media is very popular today and can be considered extremely popular at present. The comments

section of tweets we often extract from Twitter using hash-tags or keywords Hash tags Sour keywords are used to categorize user's tweets, making that data easier to identify and easier to see. can see from to remove tweets.

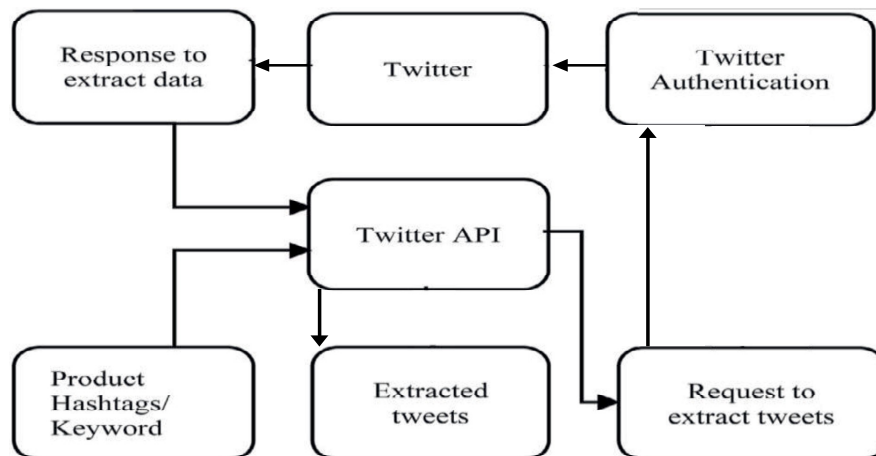


Fig.3. The Twitter data collection process.

(a) **Comparison:** - History sees a comparison of different works on emotional sentiment. The user understood the languages and problems by creating different datasets and time was analyzed using them. [3] In the paper, the impact of the corona virus in life in society is explained; using a technique based on this paper, the selection of words was done accordingly. These twitter ments are in English and this method achieved an accuracy of 92%. In the paper [5] it was observed that Twitter data is used in the English language. The result is obtained using machine learning and lexicon method techniques giving an accuracy of 96.81%. The research paper [6] covers the languages of Urdu, Roman and English. Extremis min analysis of expressions on social media. The linear support vector yields a result of 81%. The paper [10] contains research written in emotional analysis covid19 tweets in English. Achieved a result of 78.24% using very high learning techniques. The paper [22] used machine learning software to make social media comments in the English language and achieved a result of 73.33%.,

VI CONCLUSION

Twitter is a huge platform where a person's works, places, popularity etc. can be seen and shared sitting at home, where people give their opinion and share it with other's sentiments and keep criticizing, happy review etc. It is a difficult task to identify the sentiments behind each of the revisits separately because nowadays people are expressing their views in their regional languages. This data is being retrieved using text comments where text word user reviews are divided as positive, negative or neutral with polarity which can prove to be very helpful for us in analyzing the data and each Also calculates the sentiment score behind customer reviews. That's why by this Sentimental method users are able to analyze the reviews of other customers.

VII FUTURE WORK

The limits of this research paper of ours are only to the text data received from Twitter. In the future, for this dataset, data like picture, emoji and emotion can also be worked on. Only a few languages like Hindi, English and mixed Hinglish have been kept in this research paper.

REFERENCES

- [1] K.M.A. Kumar et al.(2015), Analysis of users' sentiments from Kannada web documents, Proc. Computer Science. 54247 256, doi:10.1016/j.procs.2015.06.029.
- [2] R. Thilagavathi and M. K. Krishnakumari, "Tamil English language sentiment analysissystem." [Online]. Available: www.ijert.org.
- [3] T.S. Raghavendra and K.G.Mohan, (2019), Web mining and minimization on frame work design non sentimental analysis for social tweets using machine learning, Proc. Computer. Science. page (152) , 230–235, doi:10.1016/j.procs.2019.05.047.
- [4] A.Sharma and U. Ghose, (2020) Sentimental analysis of Twitter data with respect to general elections in India, Proc. Computer Science.173, 325334 no.2019, doi:10.1016/j. procs.2020.06.038.
- [5] N.Chintalapudi et al.,(2021), Text mining with sentiment analysis on seafarers' medical documents, Int. J. Inf. Manag. Data In sights1, 100005, doi:10.1016/j.jjime.2020.100005.
- [6] A.I. Al-Ghadir et al.,(2019), A novel approach to stance detection in social media tweets by fusing ranked lists and sentiments, Inf. Fusion 67, 29–40 ,doi:10.1016/j.inffus.2020.10.003.
- [7] M.Kesavan et al (2020), Stock market prediction with historical time series data and sentimental analysis of social media data, in: Proceedings of the International Conference on Intelligent Computation Control Systems ICICCS, ,pp.477–482, doi:10.1109/ICICCS48265.2020.9121121.2020,no.1 cicc.
- [8] N.K. Khokale, (2021) engineering trends study of sentiment analysis of twitter users from their tweet's text and delusion patterns, Int. J. Eng. Innov. Technol. Expl. Eng.6192–195.
- [9] A. Alrumai et al.(2020), Sentiment analysis of comments in social media, Int. J. Electronics Computer. Eng. 10 (6), 5917–5922, doi:10.11591/ijece.v10i6.pp5917-5922.
- [10] M.Asif et al. (2020), Sentiment analysis of extreme is min social media from textual information, Telemat. Inform. 48 (2020) 101345 no. January, doi:10.1016/j.tele.2020.101345.

- [11] S. Geetha and V.K. Kaliappan (2018), Tweet Analysis Based on Distinct Opinion of Social Media Users, Springer Singapore.
- [12] M.B.Shelke and S.N.Deshmukh, (2020) Recent advances in sentiment analysis of Indian languages, Int. J.FutureGen.Comm.Net.13 (4)1656–1675.
- [13] M. Khan and K. Malik, (2018) “Sentiment classification of customer’s reviews about auto mobiles in Roman Urdu”.
- [14] E. Kumar, A real-time twitter sentiment analysis and visualization system: TwiSent,Int. Res. J. Eng. Technol. 3323 (2008) 3323–3330 no. June[Online]. Available:www.irjet.net.
- [15] P. Pandey and S. Govilkar (2015), A survey of sentiment classification techniques used for Indian regional languages, Int. J. Computer. Science. Appl. 5 (2) 13–26, doi:10.5121/ijcsa.2015.5202.
- [16] A.S.C. Chung,(2013) “BE havior Based aPproach forI Ntrusion, ”A simulation based approach for contingency planning for aircraft turnaround operation system activityes in airline hubs,vol.11,no.1,pp.1–6.
- [17] A.Dalmia et al. (2015): twitter sentimental alysis the good, the bad and the neutral, Sem Eval 520–526,doi:10.18653/v1/s15-2087.
- [18] S.Paliwal et al.(2018), Sentiment analysis and prediction using neural networks, in: Proceeding of the International Conference on Inventive Research in Computing Applications ICIRCA ,pp. 1035–1042, doi: 10.1109/ ICIRCA. 2018. 8597358. no. Icir