iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 7 **Issue:** V **Month of publication:** May 2019

DOI: https://doi.org/10.22214/ijraset.2019.5402

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue V, May 2019- Available at www.ijraset.com

Analysis of User Perception using Twitter Data

Syeda Sabreen Banu[1], Prof. Kavitha G[2]
[1, 2]*Computer Science and Engineering, UBDTCE*

Abstract: ***Social media have received more attention nowadays. Public and private opinion about a wide variety of subjects are expressed and spread continually via numerous social media. Twitter is one of the social media that is gaining popularity. Twitter offers organizations a fast and effective way to analyse customers' perspectives toward the critical to success in the market place. Developing a program for sentiment analysis is an approach to be used to computationally measure customers' perceptions. In today's world, opinions and reviews accessible to us are one of the most critical factors in formulating our views and influencing the success of a brand, product or service. With the advent and growth of social media in the world, stakeholders often take to expressing their opinions on popular social media, namely twitter. While Twitter data is extremely informative, it presents a challenge for analysis because of its humongous and disorganized nature. This project is a thorough effort to dive into the novel domain of performing sentiment analysis of people's opinions. Nowadays, Social media is becoming more and more popular since mobile devices can access social network easily from anywhere. Therefore, Social media is becoming an important topic for research in many fields. As number of people using social network are growing day by day, to communicate with their peers so that they can share their personal feeling every day and views are created on large scale. Social Media Monitoring or tracking is most important topic in today's current scenario. In today many companies have been using Social Media Marketing to advertise their products or brands, so it becomes essential for them that they can be able to calculate the success and usefulness of each product. For Constructing a Social Media Monitoring, various tool has been required which involves two components: one to evaluate how many user of their brand are attracted due to their promotion and second to find out what people thinks about the particular brand.***

Keywords: Sentiment Analysis, Opinion Mining, Natural Language Processing, Twitter.

I. INTRODUCTION

According to [1], millions of people are using social network sites to express their emotions, opinion and disclose about their daily lives. However, people write anything such as social activities or any comment on products. Through the online communities provide an interactive forum where consumers inform and influence others. Moreover, social media provides an opportunity for business that giving a platform to connect with their customers such as social media to advertise or speak directly to customers for connecting with customer's perspective of products and services. In contrast, consumers have all the power when it comes to what consumers want to see and how consumers respond. With this, the company's success & failure is publicly shared and end up with word of mouth. However, the social network can change the behaviour and decision making of consumers, for example, [2] mentions that 87% of internet users are influenced in their purchase and decision by customer's review. So that, if organization can catch up faster on what their customer's think, it would be more beneficial to organize to react on time.

Social Media has captured the attention of the entire world as it is thundering fast in sending thoughts across the globe, user friendly and free of cost requiring only a working internet connection. An important part of our behaviour has always been to find out what other people think. With the growing availability and popularity of opinion-rich resources such as online review sites and personal blogs, new opportunities and challenges arise as people now can, and do, actively use information technologies to seek out and understand the opinions of others. People are extensively using this platform to share their thoughts loud and clear. But it would be difficult to extract information from that vast platform so twitter plays an important part of our information-gathering. Twitter is one such well known micro-blogging site getting around 500 million tweets per day [3] .Each user has a daily limit of 2,400 tweets and 140 characters per tweet [4]. Twitter users post (or 'tweet') every day about various subjects like products, services, day to day activities, places, personalities etc. Hence, Twitter data is of Great germane as it can be used in various scenarios where companies or brands can utilize a direct connection to almost each of their client or user and thereby, improve upon their product. Consider a dis-satisfied costumer of a telecommunication company voicing out his/her grievances about a particular plan he/she is subscribed to. Twitter also serves as a huge platform for users to know more and get direct comments about a product or a service in which they are interested [5]. Opinions and reviews in the form of tweets from customers, potential users and critics can easily influence the image and consequently, demand of a product/service being provided by a company. Hence, whether the stakeholder's. Opinion is positive/negative about their offering becomes a crucial and pressing question for the organization to ask and monitor.

©IJRASET: All Rights are Reserved

A. Problem Statement

The problem in sentimental analysis is classifying the polarity of a given text at the document or aspect level. Whether the expressed opinion in a document, or a sentence or an entity feature is positive, negative, neutral or undecidable.

B. The Solution

The solution is to develop the intelligent system for detecting sentiment for dynamic tweets based on the threshold concept and display them in form of bar graph for better understanding for a user.
Different types of data are generated from twitter that needs to be organized and to monitor people's attitude towards products, gadgets, movie review etc. Sentiment analysis can help to develop valuable business insights from text based contents.

C. Objective of the project

1) To collect the datasets from the twitter using API.
2) To implement an algorithm for automatic classification of the collected datasets into positive, negative, neutral or undecidable.
3) Sentimental analysis to determine the attitude of mass is positive, negative or towards the subject of interest.

D. Advantages

1) Manual survey to extract opinion is avoided.
2) Saves time.
3) More accurate results.

E. Proposed System

The system is going to propose the solution for detecting sentiment for dynamic tweets based on the threshold concept. Based on the threshold value system is going to achieve accuracy in the project. In the proposed system initial expansion is done based on the topic selected. Based on the topic every word in the particular tweet, sentiment type of word is checked. Finally positive, negative or neutral count is incremented. The left out word which is not in any sentiment type ,that word sentiment is decided based on the positive ,negative and neutral count in that particular tweet. If positive count is more than negative and neutral them it will be considered positive sentiment only. Finally if left out word sentiment crosses threshold value, then that word sentiment is permanently considered.

II. LITERATURE SURVEY

In the existing system based on the initial expansion of the words they going to give sentiment process based on topics they adopted. Initially in the existing system iteration process is done (for example 1st iteration they are going to take 100 tweets, within that 100 tweets which words are coming with more positive or more negative count that words will be added as positive or negative before 2nd iteration). Here accuracy is less because after iteration immediately system considers positive or negative sentiment without considering left out words in tweets.

A. Opining Mining

Opinion mining refers to the broad area of natural language processing, text mining, computational linguistics, which involves the computational study of sentiments, opinions and emotions expressed in text [6]. Although, view or attitude based on emotion instead of reason is often colloquially referred to as a sentiment. Hence, lending to an equivalent for opinion mining or sentiment analysis.[7] stated that opinion mining has many application domains including accounting, law, research, entertainment, education, technology, politics, and marketing. In earlier days many social media have given web users avenue for opening up to express and share their thoughts and opinions [8].

B. Twitter

Twitter is a popular real time micro blogging service that allows users to share short information known as tweets which are limited to 140 characters [2,3], [9]. Users write tweets to express their opinion about various topics relating to their daily lives. Twitter is an ideal platform for the extraction of general public opinion on specific issues [9, 10]. A collection of tweets is used as the primary corpus for sentiment analysis, which refers to the use of opinion mining or natural language processing [1]. Twitter, with 500 million users and million messages per day, has quickly became a valuable asset for organizations to invigilate their reputation and brands by extracting and analysing the sentiment of the tweets by the public about their products, services market and even about

©IJRASET: All Rights are Reserved

competitors [10]. [2] Highlighted that, from the social media generated opinions with the mammoth growth of the World Wide Web, super volumes of opinion texts in the form of tweets, reviews, blogs or any discussion groups and forums are available for analysis, thus making the world wide web the fastest, most comprising and easily accessible medium for sentiment analysis.

Sentiment analysis, or opinion mining, aims at user's attitude and opinions by investigating, analysing and extracting subjective texts involving users' opinions, preferences and sentiment. Since Bo Pang put forth this concept in 2002, the academics have undertaken a diverse range of related research, due to its practicality in opinion monitoring and business competitive intelligence. Sentiment analysis on online reviews has become increasingly popular. A multidisciplinary research field in nature, sentiment analysis includes multiple fields such as natural language processing (NLP), computational linguistics, information retrieval, machine learning and artificial intelligence etc. As an astronomical quantity of sentimental subjective texts appear on Internet, researchers put more emphasis on complex sentimental sentences and texts instead of on words only.

Natural human language carries two types of information: objective information about facts and critical information with human subjective sentiment. The rapid development of forum, BBS, Blog, and review websites contributes to the exploding amount of such critical information, which reflects users' attitudes, viewpoints and opinions on products, policies, people and events, etc. Sentiment analysis has grown into a hot research field in natural language processing. Sentiment analysis first created a sensation at Text Retrieval Conference (TREC), and since 2006 related evaluation tasks have appeared every year. The annual multi-lingual opinion analysis task (MOAT) also started in 2006, including standard libraries in three languages, Chinese, English and Japanese. However, the research in China, particularly the sentiment analysis on Chinese has barely taken off. Initiated in 2008, Chinese opinion analysis evaluation (COAE) is the first sentiment analysis of the kind.

III.METHODOLOGY

A. *Modules*

1) *Tweets Import Module:* In this module, tweets are retrieved from the twitter API dynamically based on the college name input. To retrieve tweets from the twitter API account, first need to create twitter account in developer's console. After creating an account system will get consumers token key and access token key, with the help of generated keys, system is then going to communicate with twitter API to retrieve tweets. The retrieved tweets are imported into database.
2) *Pre-Processing Module:* In this module, the tweets which are imported to database from the twitter API, these tweets consist of unnecessary words, whitespaces, hyperlinks and special characters. First we need to do filtering process by removing all unnecessary words, whitespaces, hyperlinks and special characters.
3) *Self-Learning and word standardization System:* In this module, First system needs to initialize the dictionary (first iteration dictionary).In the dictionary generally system needs to initialize the positive, negative neutral and nouns. All big data and data mining projects based on the trained data, without trained data (initialization of words).So initialization of the trained data is very important.
4) *Sentiment Analysis Module:* In this module, pre-processed tweets are fetched from the database one by one. First system needs to check one by one keyword whether that keyword is noun are not, if noun then system will remove it from the particular tweet. After that the remaining keywords checked with sentiment type, whether that keywords are positive sentiment or negative sentiment or neutral sentiment. The remaining keywords in the tweet which does not belongs to any of the sentiment will be assigned temporary sentiment based on the more count of positive, negative and neutral. In the second iteration if the remaining word crosses the threshold of positive, negative or neutral, that keyword permanently added as expansion in the dictionary. Finally sentiment of the tweet is detected based on the positive, negative and neutral words in the particular tweet.

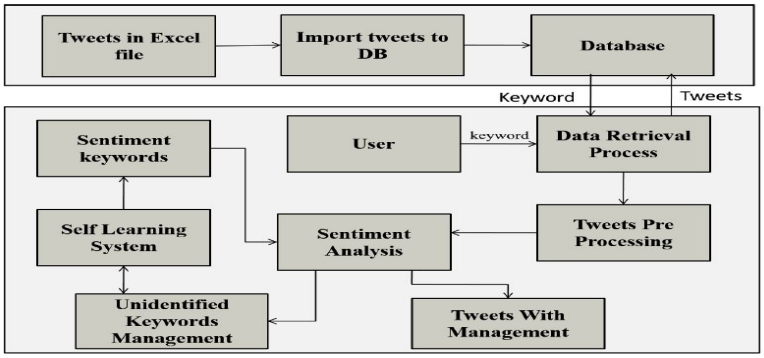


Fig 1: System Architecture

©IJRASET: All Rights are Reserved

Fig 1 shows the model for sentiment analysis of tweets. Tweets are fetched from twitter using twitter API and system can also use the stored tweets in excel style sheet file.

Excel-file is a temporary storage for the tweets, where system will implicitly store needed number of tweets. These tweets are imported to DB and stored in the database. From database keywords are fetched for data retrieval process. The tweets fetched in the data retrieval process are passed to tweets pre-processing. This module can be used for easier preparation of twitter data by removing hyperlinks, emoji's, special character and sentences containing single word.

Then the tweets are analysed and classified into positive, negative or neutral. The ones whose sentiment can't be determined are passed as unidentified keywords management to self-learning system where the keywords are attached with stored sentiments and again sent sentiment analysis process and finally the result obtained are sent to tweets with management and displayed using graphs and pie chart.

B. Algorithm

The sentiment analyser algorithm is shown below:

INPUT: Text File (comment or review) $\mathcal{T}$, The sentiment lexicon $\mathcal{L}$.
OUTPUT: S_{mt} = {***P***, ***Ng***, or ***N***} and Strength S, *where **P**: Positive, **Ng**: Negative, **N**: Neutral.*
INITIALIZATION: SumPos and SumNeg = 0, *where*
SumPos: accumulates the polarity of positive tokens $t_{i\text{-}Smt}$ in $\mathcal{T}$,
SumNeg: accumulates the polarity of negative tokens $t_{i\text{-}Smt}$ in $\mathcal{T}$
Begin

```
For each t_i ∈ T do
    Search for t_i in L
    If t_i ∈ Pos-list then
        SumPos ← SumPos + t_i-Smt
    Else if t_i ∈ Neg-list then
        SumNeg ← SumNeg + t_i-Smt
    End If
End For
If SumPos > |SumNeg| then
    S_mt = P
    S    = SumPos / (SumPos +SumNeg)
Else If SumPos < |SumNeg| then
    S_mt = Ng
    S    = SumNeg / (SumPos +SumNeg)
Else
    S_mt = N
    S    = SumPos / (SumPos +SumNeg)
End If
```

End

IV. RESULTS

The results of the system are given below:

A. Tweets details

The tweet details page is shown below in the Fig 2.

1) Keyword related fetched tweets are displayed here.
2) Ten tweets per page are displayed.
3) Total 10 pages are displayed after processing.

©IJRASET: All Rights are Reserved

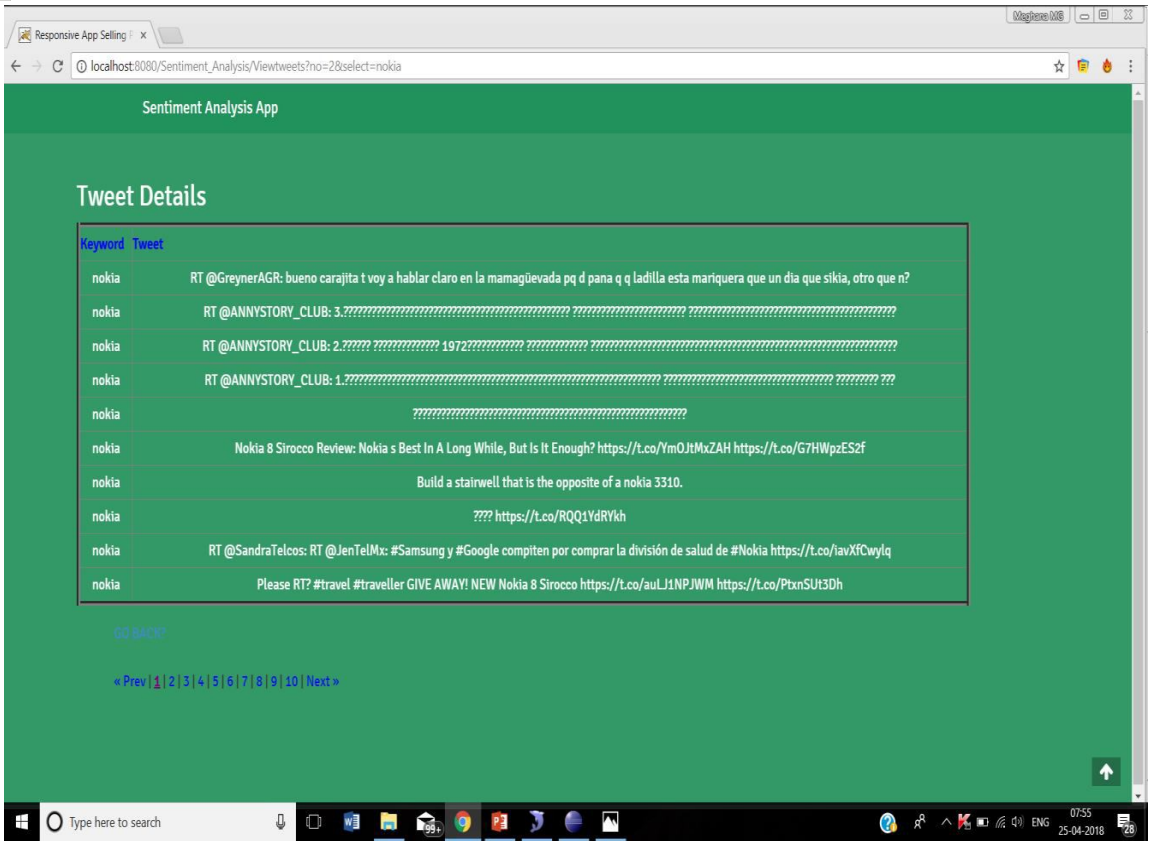


FIG 2: SNAPSHOT OF TWEETS DETAILS PAGE

B. Sentiment Analysis of Tweets

The sentiment analysis page is shown below in Fig 3.

1) Result of sentiment analysis is displayed on this screen.

2) Single tweet is shown with its sentiment.

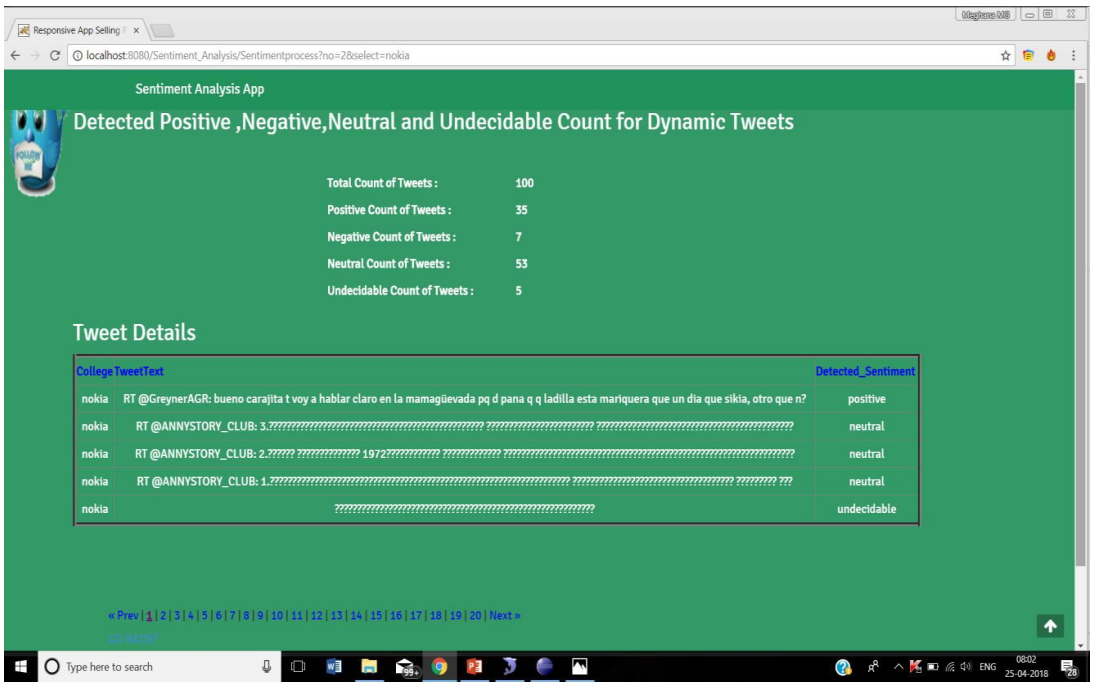


Fig 3: Snapshot of Sentiment analysis page

©IJRASET: All Rights are Reserved

C. Graphical representation of analysis

The graphical representation page is shown below in Fig 4.

1) The graphical representation of result.

2) Here bar graph is used to show the results.

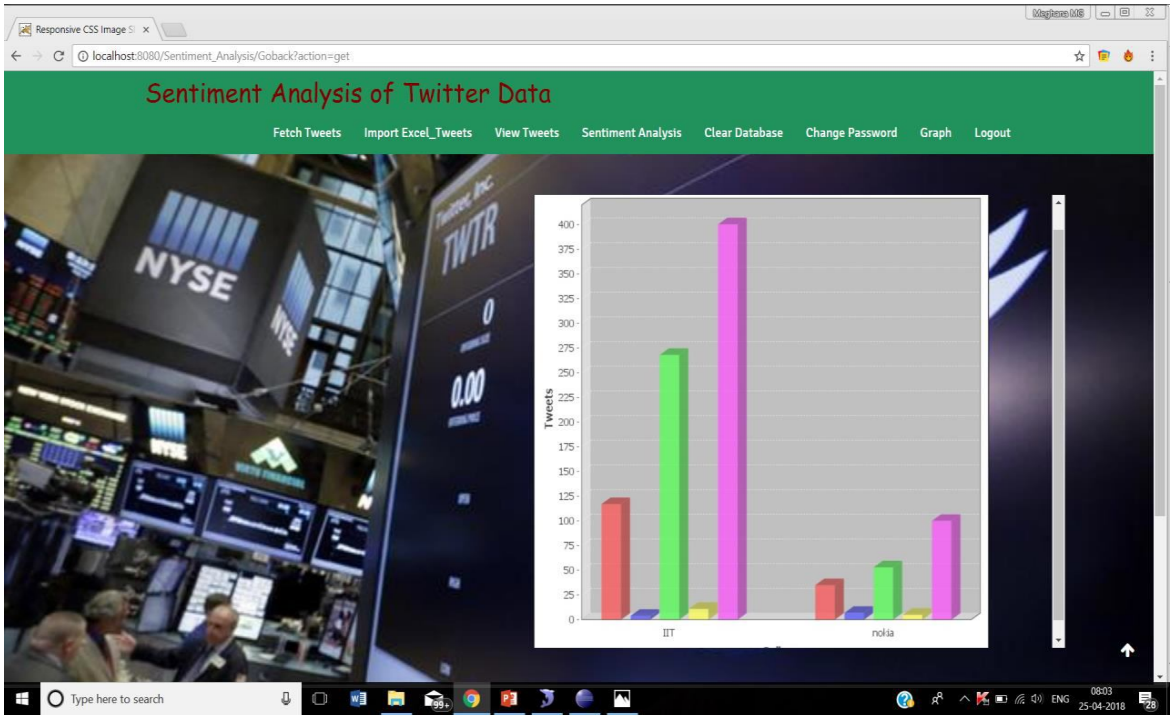


FIG 4: SNAPSHOT OF GRAPHICAL REPRESENTATION PAGE

V. CONCLUSION

Sentiment analysis is an effective way of classifying the opinions formulated by people regarding any topic, service or product. Automation of this task makes it easier to deal with the massive amount of data being produced by social websites like Twitter on a real-time basis. Compared to the existing system, based on the threshold concept, iteration concept, keywords are expanded. Based on the expanded words efficiency of the system is increased. Successfully collected all the required datasets for analysis process. Algorithm implemented for automatic classification of text into positive, negative, neutral or undecidable.

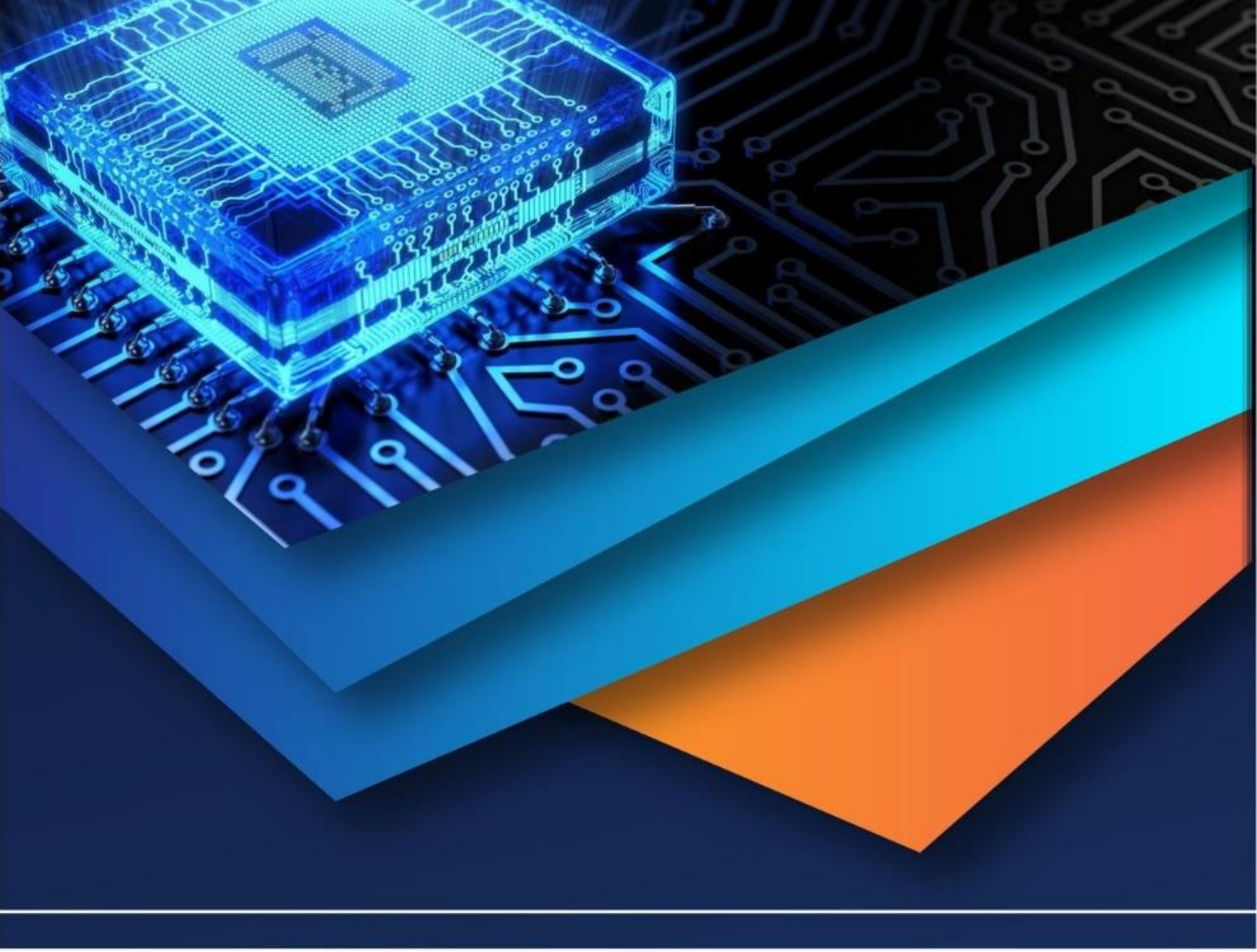
VI. FUTURE SCOPE

The sentimental analysis process can be developed as a product to know the reviews on various fields, such as movie, college, schools, and services so on…

REFERENCES

[1] M.Rambocas, and J. Gama, "Marketing Research: The Role of Sentiment Analysis". The 5th SNA-KDD Workshop'11. University of Porto, 2013.
[2] A. K. Jose, N. Bhatia, and S. Krishna, "Twitter Sentiment Analysis". National Institute of Technology Calicut, 2010.
[3] P. Lai, "Extracting Strong Sentiment Trendfrom Twitter". Stanford University, 2012.
[4] Twitter Usage/Company Facts, https://about.twitter.com/company.
[5] Posting a tweet, https://support.twitter.com/articles/15367-posting-a-tweet.
[6] King R. A., Racherla P. and Bush V. D., What We Know and Don't Know about Online Word-of-Mouth:A Review and Synthesis of the Literature, Journal of Interactive Marketing, vol. 28, issue 3, pp. 167-183, August 2014.
[7] T. Carpenter, and T. Way, "Tracking Sentiment Analysis through Twitter,". ACM computer survey. Villanova: Villanova University, 2010.
[8] D. Osimo, and F. Mureddu, "Research Challenge on Opinion Mining and Sentiment Analysis," Proceeding of the 12th conference of Fruct association, 2010, United Kingdom.
[9] A. Pak,and P. Paroubek, "Twitter as a Corpus for Sentiment Analysis and Opinion Mining," Special Issue of International Journal of Computer Application, France: Universitede Paris-Sud, 2010.
[10] S.Lohmann, M. Burch, H. Schmauder and D. Weiskopf, "Visual Analysis of Microblog Content Using Time-Varying Co-occurrence Highlighting in Tag Clouds," Annual conference of VISVISUS. Germany: University of Stuttgart, 2012.

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)