iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** V **Month of publication:** May 2022

DOI: https://doi.org/10.22214/ijraset.2022.43263

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: 7.538
Volume 10 Issue V May 2022- Available at www.ijraset.com

A Survey on Real Time Sentiment Analysis Based on Social Media Mining

Avani Vagadiya[1], Sapna Patel[2]
[1,2]*Assistant Professor in Computer Engineering Department, Parul Polytechnic Institute, Parul University, Vadodara, Gujarat, India*

Abstract: ***Social media is a platform where people from any corner of the world share their opinions with each other. Thus current era is called as social media era. Social networking sites such as Facebook and Twitter plays an important role in information retrieval and web data analysis. In a survey it is found that Twitter produces more than 500 million of tweets each day which is about 8 TB of data which can be mined and sentiments analysis can be carried out. The purpose of mining and exacting opinions is todiscover and categorize the positive and negative sentiments of society. So as there is a huge repository of data, clustering is efficient and quick way to study people's expression and get a conclusion. This paper surveys the different mining techniques to carry out opinions and sentiments analysis and represent it to the best way based on subjectivity and polarity.***

Keyword: Social Media, Opinion Mining, Sentiments Analysis, Clustering, Expressions, Social Networking.

I. INTRODUCTION

Social media is a web based services that allows people to create public or private profile and provides communicating connect with people to share. Thus social media has gained a huge attentionnow a days. So, social media mining come out as best way to conclude billions of people thinking and get a reliable result out of it which makes it an excellent marketing tool. It can be used for many business related activities for example to analyze customer experiences, one need to recognize theopinions and sentiments of the customers based on which product review can be concluded. As social media data have threecharacteristics they are having:

1) Very large amount of data
2) Dynamic data
3) Noisy data

These Characteristics of data make research difficult on dat. Hence various data miningtechniques are used which makes it possible to extract knowledge out of it. Social media mining techniques are as below:

a) Opinion Mining
b) Sentiment Analysis
c) Clustering
d) Text Mining
e) Web Mining

II. SOCIAL MEDIA MINING APPLICATIONS AND TASKS

Social Media Mining is an elaborative filed of a data mining task, which is used to representing, analyzing and identifying interesting actionable patterns from the raw social media data [1].

A. Social Media Mining Tasks

Social Media Mining performs following tasks:

1) Social Media Event Detection
2) Community Structure
3) Network Measures and Models
4) Social Search
5) Trust in Social Media
6) Sentiment Analysis in Social Media

©IJRASET: All Rights are Reserved

7) Social Spammer Detection
8) Distrust and Negative Links
9) Role of Social Media in Crises
10) Location based Social Network Mining
11) Information Cascade



Figure 1: Social Media Fields

B. Social Media Mining Applications

Social media mining is performed by Social media analysts. Hence, this process is called Social Media Analytics. Social Media Mining/Analytics can be used in:

1) Business Development
2) Social Science Research
3) Health Service
4) Educational Purpose
5) Influence Marketing
6) Weather Forecasting
7) Fraud Detection
8) Financial Banking
9) Price Detection in Share Market
10) Criminal Investigation

III. RELATED WORK

In this paper, we have focused on socialnetworking/media sites' data like Twitter. Twitter isa social networking site which provides us a platform to post real time reactions and opinions about everything. The messages of this platform areshort and often people do not maintain rules ofgrammar in their messages. That is why using the traditional methods for social media mining gives the poor results. Hence to get perfect result with better quality two methods are used:

1) *Sentiment Analysis:* Aim of this process isto mine opinions from raw data of social media at sentence and document level [1].
2) *Text Classification:* Aim of this process is grouping the similar text messages whichmake manageable information. It uses a framework that can a form cluster of similar texts which belong to the same topic and similar topics [1].

A. Sentiment Analysis

The process of sentiment analysis follows a specific structure which contains four stages [3]:

1) Data Collection Stage- Trending topics, topic definition and tweets are downloaded in this stage to generate a document.
2) Preprocessing Stage- In this stage some techniques are used to convert the raw social media data into an understandable form withbest quality.
3) Data Modeling Stage- Documents which hasbeen created from social media data in the first stage will be transformed into tokens.
4) Sentiment Analysis Stage- In this stage, tweets will be classified based on polarity scores.

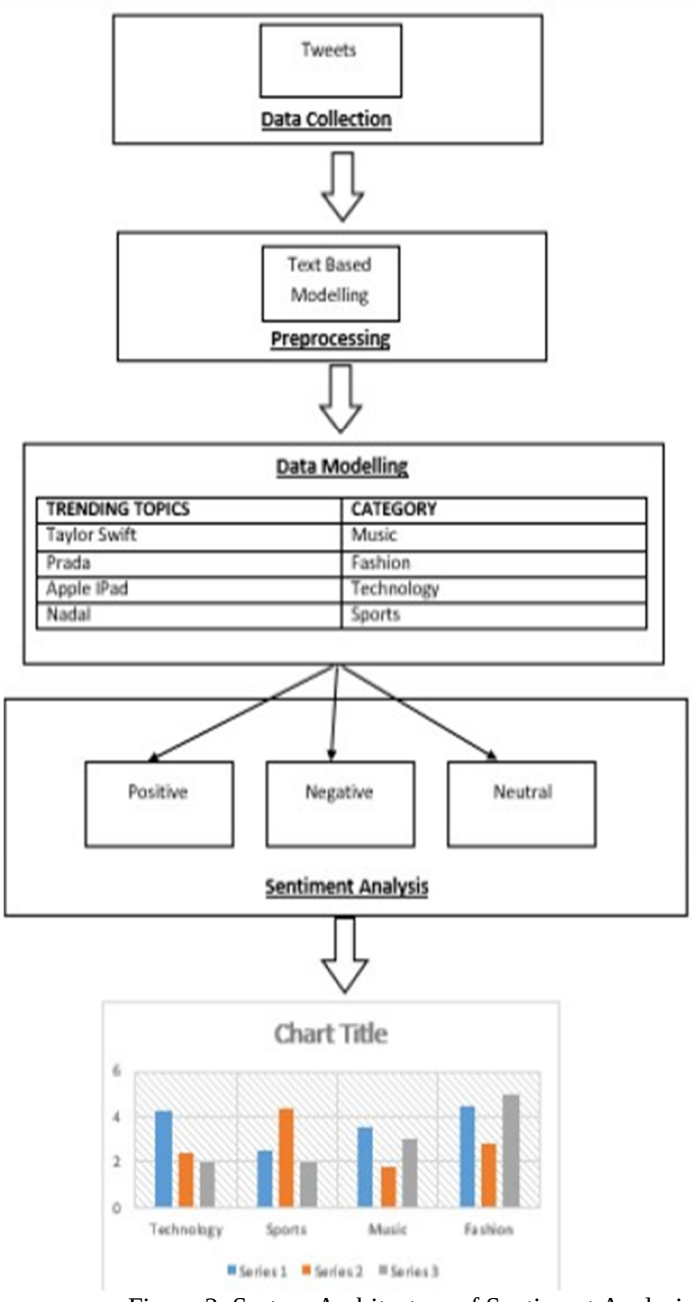


Figure 2: System Architecture of Sentiment Analysis

B. Working of Sentiment Analysis

1) Sentiment Intensity Analzer()- It returnssentiment intensity scores to the sentences.

2) polarity_scores()- It returns a float valuefor sentiment strength on basis of the input text. Positive value refers positive valence, negative value refers negative valence and the remaining refers neutral valence.

3) The polarity scored words will be compared with a value that has a maximum and minimum range.

4) Then, in next step charts will be created to plot the sentiment based results.

To understand the sentiment analysisprocess easily we can take the example of trending technologies.

a) Keywords passed are: 'Smartphone','Laptop', 'Tablet'.

b) Final result after Sentiment Analysis Process:

Total number of tweets collected- 5756
Tweets collected for **Smartphone- 2703**
Positive tweets-535
Negative tweets- 110
Neutral tweets- 2058
Tweets collected for **Laptop- 1899**
Positive tweets-437
Negative tweets- 302
Neutral tweets- 1143
Tweets collected for **Tablet- 870**
Positive tweets-144
Negative tweets- 52
Neutral tweets -593

Figure 3: Analysis of tweets by Sentiment Analysis

Created chart to plot the result based onSentiment Analysis:

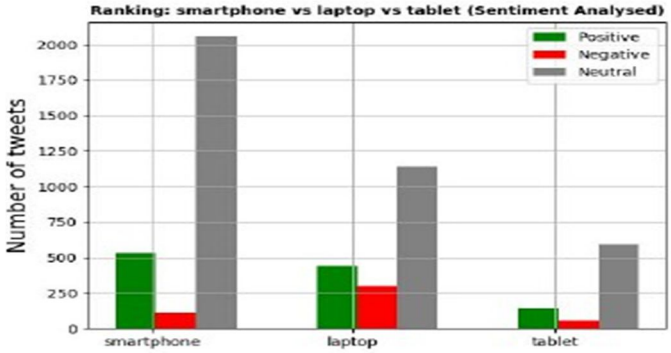


Figure 4: Final result of Sentiment Analysis

C. Opinion Mining

In this paper, manually labelled data have been used as the training data to build a new model for Sentiment Classification process. The new created architecture is known as 'Opinion Miner' [2].

First, the tweets from Twitter will becollected and the preprocessing step will be applied on those tweets. Preprocessing is used to indentify various properties of messages that users have posted on the social networking platform. Tweets are used to hold an opinion of a user. In next step, these tweets will be extracted and then tweets will be classified in labeled classifiers [2].

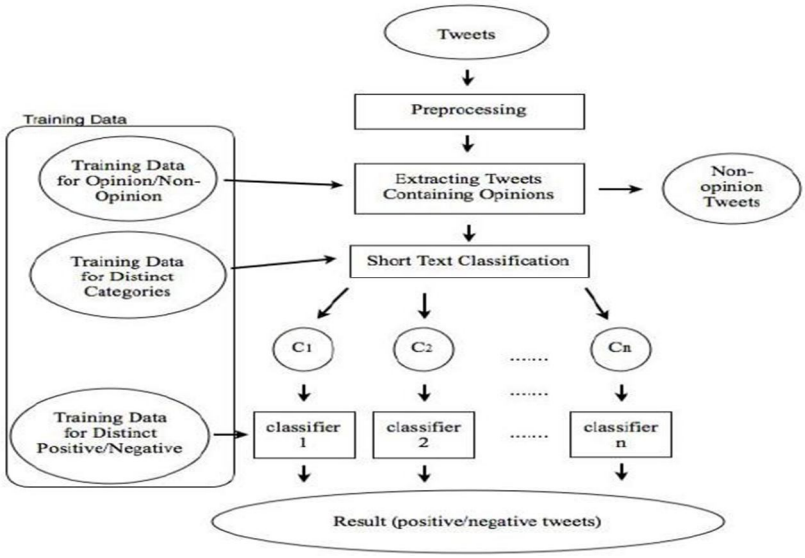


Figure 5: System Architecture of Opinion Mining

Unique Properties of messages that users post on social networking platform are like [3]:

1) *Usernames:* It refers to the Twitter user names that users use to direct their messages. For this '@' symbol will be used.
2) *Hash Tags:* Twitter facilitates the user to post the text/keywords with the use of hash tags, which maintains a specific form of "#<tagname>".
3) *RT (Re-Tweeting):* If a tweet of someone is interesting enough for others then other people can do re-tweeting and Twitter uses RT to represent the re-tweeting.

For the pre-processing step some of thetweets will be eliminated like [4]:

a) Tweets that are not in English language.
b) Tweets that have very few words (Thresholdvalue for length of tweet is set as five).
c) Tweets that have very few words apart fromthe greeting words.
d) Tweets that have just links/UniformResource Locations (URL).

D. Clustering

Clustering means grouping or separating particular set of objects having some similarities on its characteristics and aggregating them. This is the simpler and most effective way of handling large amount of data to get knowledge from it.

Clustering can be used in social media mining for analyzing and finding the textual similarities between the user contents. Main tasks of clustering in social media mining are:

1) *Pre Processing:* It includes tweetscollection process from the Twitter. It also contains some sub tasks like
a) *Data Extraction:* Data Collection process
b) *Stop Words Removal:* Removing stop words like "the", "and", "a", "as", "about", "at" etc.
c) *Stemming of the Text:* Reducing the inflected words like "talk", "talked" and "talking" etc. This can be performed byStemming algorithm.
d) *Lexical Analysis:* Nouns, verbs, adjectives and adverbs will be grouped as cognitive synonyms sets.

2) *Clustering Techniques:* Based on clusteringtechniques various algorithms are applied onset of objects which makes it easier to differentiate. Some of the algorithms used inclustering are:
a) *Centroid Based Algorithm:* In this algorithm every group is referred by a vector value.
b) *Distributed Based Algorithm:* This algorithm combines objects belonging to the same distribution.
c) *Connectivity Based Algorithm:* This algorithm has hierarchical representation based on the relation on distance between them.
d) *Density Based Algorithm:* This algorithm creates clusters based on the density. The idea is to expand cluster as long as it exceed the threshold value of neighboring clusters.

3) *Experimental Results:* This process will conclude the final result of analysis in the form of graphs or datasets.

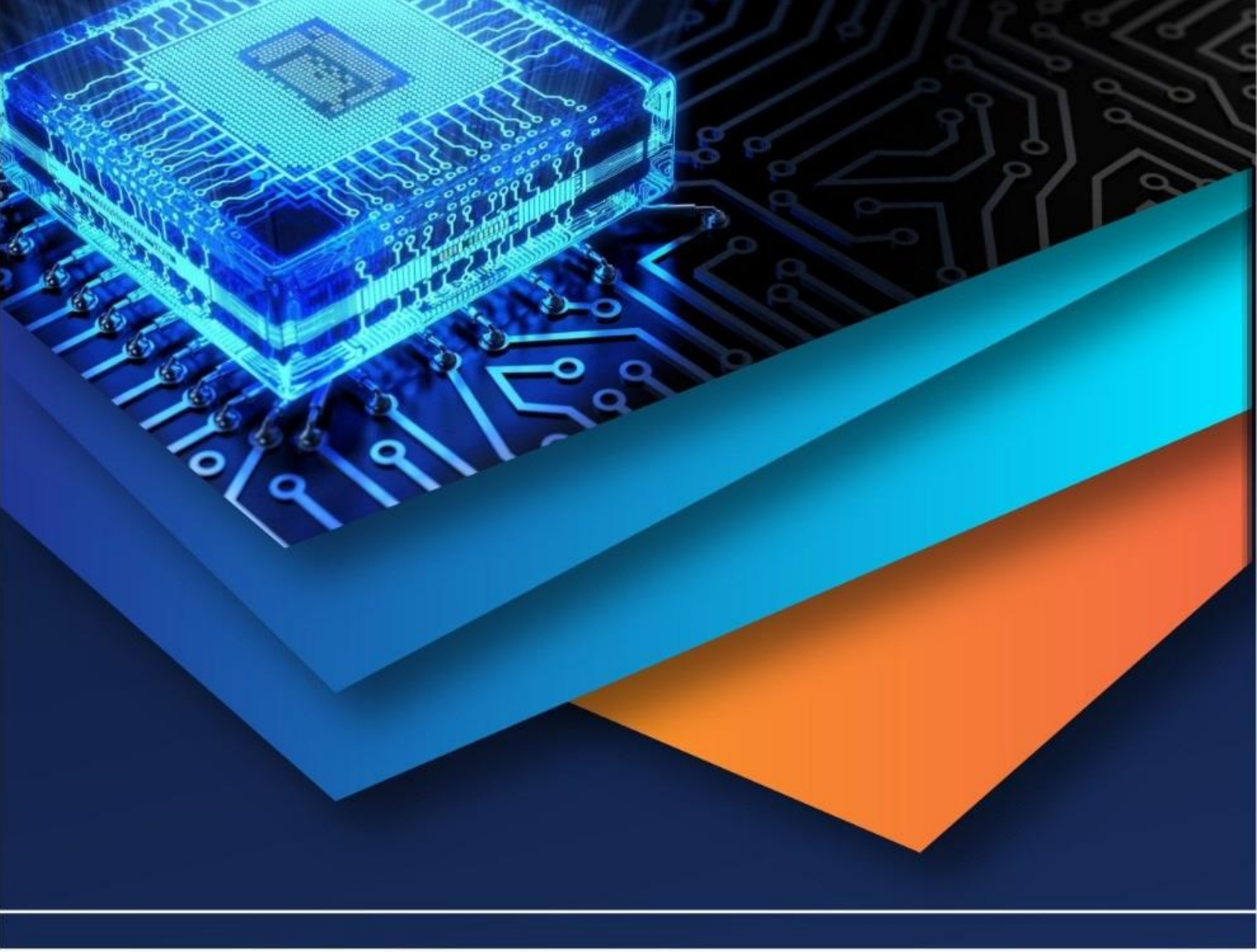
IV. CONCLUSION

In this paper, we have mined the raw data which has been collected from the microblogging sites like Twitter and FaceBook. On these raw data, techniques like Sentiment Analysis and Opinion Mining has been performed to classify the sentiments of users on the basis of their posts. By Sentiment Analysis the data like posts of users on Twitter and Facebook will be mined then Opinion Mining technique will be applied on that data to cluster the sentiments of users. The final result afterOpinion Mining process will be shown as a graph of different sentiments for better understanding. Thus it becomes easy and efficient to conclude the interesting behaviour pattern of the society which can be very much useful for brands, companies, businesses and even an individual.

REFERENCES

[1] Ardra, Blessy Merin Varughese, Merline Susan Joseph, Preethi Elsa Thomas, Sherly K K "Analyzing the Behavior of Youth to Sociality Using Social Media Mining" IEEE 2017.
[2] Donia Gamal, Marco Alfonse, EL-Sayed M. EL –Horbaty, Abdel Badeeh M.Salem "A Comparative Study on Opinion Mining Algorithms of Social Media Statuses" IEEE 2017.
[3] Po-Wei Liang, Bi-Ru Dai "Opinion Mining on Social Media Data" IEEE 2013.
[4] Andrei Pavel,Vasile Palade, Rahat Iqbal, Diana Hintea "Using Short URLs in Tweets to Improve Twitter Opinion Mining" IEEE 2017.
[5] Baocheng Huang, Guang YuS "Research on the mining of opinion community for social media based on sentiment analysis and regional distribution" IEEE 2016.

[6] Devendra K. Tayal, Sumit K. Yadav "Analysis of sentiments & polarity computation of opinions Sign In or Purchase" IEEE 2017.
[7] Shreya Ahuja, Gaurav Dubey "Clustering and sentiment analysis on Twitter data" IEEE 2017.
[8] Jainee Vora, Anu Mary Chacko "Sentiment analysis of tweets to identify the correlated factors that influence an issue of interest" IEEE 2017.
[9] Rajeshwari Dembala, S. Vagdevi "Conceptual notion for opinion mining from upcoming big data" IEEE 2017.
[10] Anitha Anandhan, Liyana Shuib "Social Media Recommender Systems: Review and Open Research Issues" IEEE 2018.



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)