



INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Microblogging On Twitter

M.Galeiah

Department of Information Technology
S.R.K.R. Engineering College(A)
SRKR Marg, Bhimavaram.

M.Hima Varshita

Department of Information Technology
S.R.K.R. Engineering College(A)
SRKR Marg, Bhimavaram.

N.Jhansi

Department of Information Technology
S.R.K.R. Engineering College(A)
SRKR Marg, Bhimavaram.

M.preethi

Department of Information Technology
S.R.K.R. Engineering College(A)
SRKR Marg, Bhimavaram.

K.Satwika

Department of Information Technology
S.R.K.R. Engineering College(A)
SRKR Marg, Bhimavaram.

Abstract—Microblogging is a new form of communication in which individuals write brief descriptions of their present condition, which are then shared via email, the internet, instant chats, or mobile devices. By examining the topological and geographic characteristics of the social network of Twitter, one of the most well-known microblogging platforms, we provide our observations of the microblogging phenomenon. We have discovered that people mostly use microblogging to discuss their daily activities and to seek out or exchange information. We provide a taxonomy that groups the fundamental purposes behind users' microblogging posts. We demonstrate that users with related intents connect with one another by averaging the apparent intentions of users in implicit communities generated from the data.

Keywords—Microblogging, Data analysis, social media, Twitter.

I. INTRODUCTION

The term "opinion mining" also applies to sentiment analysis. Social media, such as blogs, forums, wikis, review sites, social networks, tweets, and so on, are used by people to share their information, experiences, and ideas with the rest of the world. This has altered how people communicate and have an impact on others' social, political, and economic conduct in Web 2.0. In fact, Web 2.0 gives everyone a voice and promises to improve human collaboration on a global scale by enabling people to share thoughts through read-write Web and user-generated content. An opinion is essentially a sentiment, view, attitude, emotion, or appraisal regarding a thing or a part of it from the perspective bearer at a certain moment. The entity can be a thing (i.e., a product or service), a thing (i.e., an event), a person (i.e., an organisation), a thing (i.e., a topic), or something (i.e., a feature). Companies, legislators, service providers, social psychologists, researchers, and other players must assess user generated opinions due to the expansion of these viewpoints in order to make better decisions. Finding innovative methods to uncover and summarise the broad overview of a certain topic has grown more difficult as a result of the sheer volume of tweets. With the huge amount of

increase in the social networks, the number of people expressing their views and opinions via social networks are increasing. This information is useful for everyone like businesses, governments and individuals. The software requirements specification is produced at the culmination of the analysis task. Software Requirement Specification is preliminarily a user's manual, which is primarily concerned with the user and external view of the project. The software requirement specification analysis is the technical specification of the requirements of the software project. Requirement focuses on describing the purpose of the system. Users should have an easy interface to view and perform various operations.

II. LITERATURE SURVEY (RELATED WORK)

Microblogging on Twitter has been a topic of interest in the field of computer science, information science, communication, and social media. Here is a literature survey on microblogging on Twitter.

1. Boyd, D. (2010). The possibilities, dynamics, and ramifications of social networking sites as networked publics. *Networked self: Identity, community, and culture on social networking sites*, edited by Z. Papacharissi (pp. 39-58). Routledge.

2. Honeycutt, C., and Herring, S. (2009). Twitter conversation and collaboration go beyond microblogging. In the 40th Hawaii International Conference on System Sciences proceedings.

3. Java, A., Song, X., T. Finin, & B. Tseng (2007). Understanding the communities and usage of microblogging is why we use Twitter. Web mining and social network analysis were covered in Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop (pp. 56-65).

4. Kwak, H., Lee, C., Park, H., and Moon (2010). Is Twitter a news media or a social network? In the 19th international conference on the internet's proceedings (pp. 591-600).

5. M. Naaman, H. Becker, and L. Gravano (2011). Emerging trends on Twitter are described as hip and stylish.

62(5), 902-918, Journal of the American Society for Information Science and Technology.

6. Popescu, A. M., & Pennacchiotti, M. (2011). a method for classifying Twitter users using machine learning. The 5th International AAAI Conference on Weblogs and Social Media Proceedings.

7. Tumasjan, A., Sprenger, T., Sandner, P., and Welpe, I. M. (2010). Using Twitter to predict elections: What 140 characters can tell us about political emotion. Weblogs and Social Media: Proceedings of the Fourth International AAAI Conference.

8. Vieweg, S., Hughes, A., Starbird, K., and Palen, L. (2010). Twitter's potential to improve situational awareness amid two natural disasters. SIGCHI conference proceedings on human considerations in computer systems (pp. 1079-1088).

9. Yang, Q., Chen, E., Li, Y., Hu, Y., & Xue, G. (2013). using social connections to analyse sentiment in microblogging. 2013 Conference on Empirical Methods in Natural Language Processing Proceedings (pp. 1324-1335).

These studies cover various aspects of microblogging on Twitter, including its usage patterns, social network analysis, user classification, sentiment analysis, and its role in events and political sentiment analysis.

III. SYSTEM IMPLEMENTATION (METHODOLOGY)

There are various methodologies for implementing sentiment analysis such as using Machine learning algorithms like Natural Language Processing or standard libraries, Naive Bayes, that are already available. But we are implementing a Machine learning algorithm for sentiment analysis i.e., Natural Language Processing (NLP).

4.1 Implementation

We are developing a web app for twitter sentiment analysis which classifies the tweets using NLP Algorithm. So, we are using Jupyter Notebook as a framework for integrating the twitter data to our project. We are executing the following to complete our project.

a) 4.1.1 Implementation of Classifier

Following are the steps involved in implementing the classifier:

Step 1: Preparing the Test Set

1.1: Getting the authentication credentials

For getting the authentication credentials we need to perform the following steps:

1.1.1 Create a developer account:

We have to create a developer account which will take 10 days in order to get the account after submitting our request to twitter.

1.1.2 Create an app:

After getting the developer account. We should create an app with some by filling some details.

1.1.3 Getting keys and tokens for authentication:

After creating the app get the credentials from keys and tokens

1.2: Authenticating Tweepy with our Python script

After getting the credentials we add the following code to our python script for authenticating the twitter API. We first copied all the 4 generated keys in an excel sheet and named it 'config.xlsx'

#reading the file

```
config = pd.read_excel('./config.xlsx')
```

#assigning the data in the file to the variables

```
twitterApiKey = config['twitterApiKey']
```

```
twitterApiSecret = config['twitterApiSecret']
```

```
twitterAccessToken = config['twitterAccessToken']
```

```
twitterAccessTokenSecret = config['twitterAccessTokenSecret']
```

#authenticating the tokens

```
auth = tweepy.OAuthHandler(twitterApiKey, twitterApiSecret)
```

```
auth.set_access_token(twitterAccessToken, twitterAccessTokenSecret)
```

```
twitterApi = tweepy.API(auth, wait_on_rate_limit = True)
```

Step 2: With the help of the tokens, access a twitter user's tweets

This phase helps in retrieving the tweets of the user whose name we had mentioned in the code. Similarly, it would also retrieve the tweets related with the search term mentioned in the code. All the recent tweets were retrieved quite accurately.

Step 3: Clean up the tweets

The code for this phase is given in the coding section. In this first the tweets are broken into tokens and with the help of regular expressions, the tweets were cleaned. In this step, all the hashtags, user IDs, etc., were removed from the previously retrieved tweets.

Step 4: Getting the subjectivity and polarity

Polarity is a type of floating-point number that falls within the range [-1,1], where 1 denotes a positive statement, -1 a negative statement, and 0 a neutral statement. Objective statements pertain to factual information, but subjective sentences typically allude to personal opinion, emotion, or judgement. Moreover, subjectivity is a float with a [0, 1] range.

Step 5: Analyzing the tweets into positive, negative and neutral

It will depend on this variable "a" how the previously collected tweets were divided into positive, negative, and neutral tweets. A tweet is considered negative if the value of "a" is less than zero, neutral if the value of "a" is equal to zero, and positive if the value of "a" is larger than zero.

Step 6: Displaying the analysis in graphs

The code for this phase is given in the coding section. Here using the "matplotlib" libraries and "seaborn" various graphical patterns can be plotted. Firstly, we have shown the percentage of the tweets based on their classification. We have plotted a pie chart, bar graph, scattered plot, and a wordcloud for the tweets that were retrieved on giving a username and we have plotted a bar graph and a scattered plot for a search term.

IV. EXPERIMENTS & RESULTS

Here is the data frame in which the tweets are placed. After all the normalization of tweets based on the polarity and subjectivity the tweets were given sentiments. If the polarity is less than zero the tweets were given a sentiment of negative, if it is greater than zero the tweets were given a sentiment of positive and if it is equal to zero the tweets were given a sentiment of neutral.

	Tweet	Subjectivity	Polarity	Sentiment
0	Here goes the in/nVeerayyaTitleSong/n/n/nWalt...	0.00	0.00000	Neutral
1	విలక్షణమైన నటుడు, తనదైన శైలి తో తెలుగు పేక్షక...	0.00	0.00000	Neutral
2	Merry Christmas to All !! 🎄	0.00	0.00000	Neutral
3	మాకు జన్మనిచ్చి, క్రమశిక్షణతో పెంచి, జీవితపు ...	0.00	0.00000	Neutral
4	Rest in peace nNavarasa Natana Sarvabhouma n...	0.00	0.00000	Neutral
5	Wishing Sri a Very Happy 50th Birthday!! Many...	0.75	0.75000	Positive
6	Wishing all the very best for RangalMarthanda	0.39	1.00000	Positive
7	Here goes the SrideviChiranjeevi Song ... shot ...	0.00	0.00000	Neutral
8	Kudos to Argentina & LionelMessi on the sp...	0.90	1.00000	Positive
9	Sharing a little surprise for all of you n at ...	0.75	0.15625	Positive

Fig 4.1 Tweets with their Sentiments

Based on the sentiments of the each tweet, whole twitter account of a person or an organization can be easily evaluated using graphs and plots.

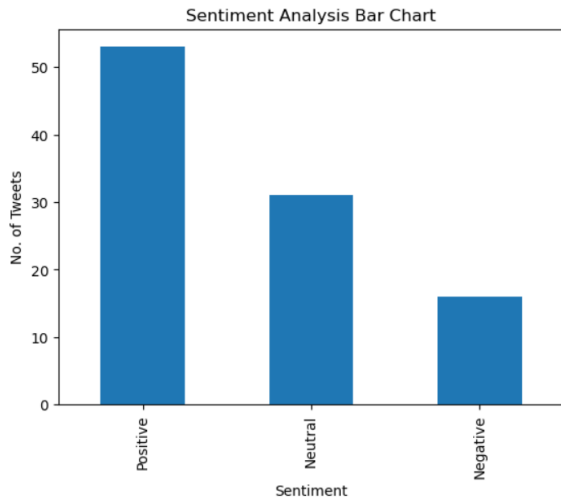


Fig 4.2 Bar chart of Sentiments

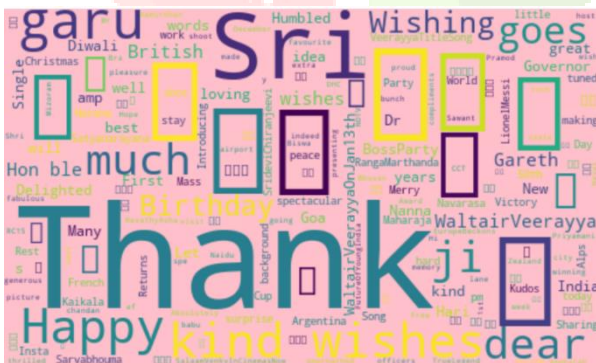


Fig 4.3 Most used words in the tweets

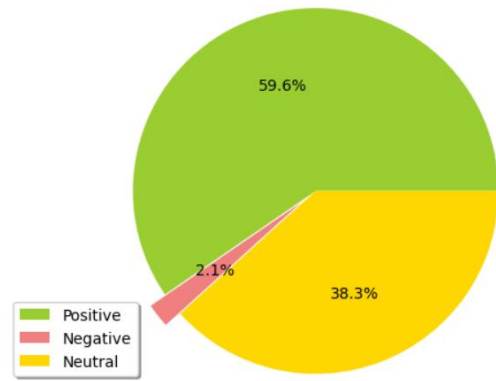


Fig 4.4 Pie chart of Sentiments

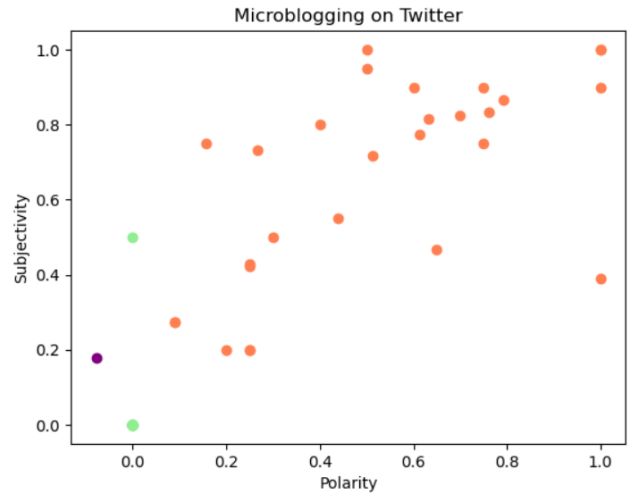


Fig 4.5 Scatterplot based on Subjectivity and Polarity

V. CONCLUSION

We have collected different data from the Twitter, our project had analyzed the tweets that were retrieved. We have used Tweepy (twitter api) for accessing tweets from twitter. Our methodology has extracted tweets from Twitter by the 4 keys that were generated after creating an app in the developer account and classified tweets into positive, negative and neutral. It also stated the subjectivity and polarity of the tweets extracted on giving a username as well as a search term. Then the retrieved data can be visualised using visualization techniques.

VI. FUTURE WORK

Microblogging on Twitter is a dynamic field, and there are several areas that could be explored for future research, such as:

- 1. **Analysis of user behavior and engagement:** Studying how users engage with Twitter and what motivates them to use the platform, including factors such as emotional expression, self-disclosure, and privacy concerns.
- 2. **Analysis of social networks:** Examining the structure of social networks on Twitter and how they influence the spread of information, opinions, and sentiment.
- 3. **Natural language processing:** Exploring new techniques and tools for more accurate and efficient natural language processing of Twitter data, including multilingual analysis, sarcasm detection, and irony detection.
- 4. **Event detection and analysis:** Developing methods to detect and track real-world events on Twitter, including natural disasters, political events, and public health emergencies, and analyzing the sentiment and impact of these events on Twitter users.

5. Machine learning and deep learning: Investigating new machine learning and deep learning techniques for analyzing and modeling Twitter data, including classification, clustering, and prediction tasks.

6. Ethical considerations: Examining the ethical and social implications of microblogging on Twitter, including privacy, bias, misinformation, and censorship, and developing guidelines and policies to address these issues.

ACKNOWLEDGMENT

We would like to acknowledge all the Colleagues and collaborators who contributed to the research, including data collection, analysis, and interpretation. We would acknowledge Twitter platform and its users for providing the data and enabling research in the field. We would also like to acknowledge Reviewers and editors who provided constructive feedback and suggestions for improving the quality of the research, previous researchers in the field whose work provided the foundation for the current study and the broader community of researchers and practitioners in the field who have contributed to the development of microblogging on Twitter research.



REFERENCES

- [1] Developing a Naive Bayes Classifier from in Python by Jason Brownlee on (2020)
- [2] ZahraRezaei.(2018).Sentimental analysis on Twitter using McDiarmid algorithm <https://ieeexplore.ieee.org/document/8167924>.
- [3] Daniele Cienne.(2019).Twitter vigilance: A multi-user platform for analysis, NLP and sentiment analysis.
- [4] Shital Anil Phand.(2017).Twitter sentiment classification using NLP <https://ieeexplore.ieee.org/document/8122138>
- [5] Jumadi.(2017).Opinion mining on Twitter microblogging using SVM <https://ieeexplore.ieee.org/document/7577569>
- [6] Fajri Koto.(2019).Inset lexicon: Evaluation of a word sentiment analysis in microblogs <https://ieeexplore.ieee.org/document/8300625>
- [7] Metin Bilgin.(2017). Sentiment analysis on Twitter data with semi-supervised Doc2Vec <https://ieeexplore.ieee.org/document/8093492>
- [8] Tanveer Ahsan.(2016). Combining a rule-based classifier with weakly supervised learning for twitter sentiment analysis <https://ieeexplore.ieee.org/document/7856499>
- [9] Networks".49th Hawaii International Conference on System ssscc Sciences(HIXSS),IEEE,2016. Poria, S., Cambria, E., & Gelbukh, A. (2015). Conference on Empirical Methods in NLP (pp. 2539-2544)