International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 11 | Nov 2019 www.irjet.net p-ISSN: 2395-0072

Sentimental Analysis of Twitter Data for Job Opportunities

Aaroh Baweja[1], Mr. Pankaj Garg[2]

[1]Student, Department of Information Technology, Maharaja Agrasen Institute of Technology, Delhi, India
[2]Assistant Professor, Department of Information Technology, Maharaja Agrasen Institute of Technology, Delhi, India

---

Abstract - *Informal communities are the fundamental assets to assemble data about individuals' feeling and notions towards various themes as they go through hours day by day on social Medias and share their opinion. Twitter can profoundly affect your pursuit of employment achievement – or scarcity in that department .Little advances can assist you with transforming Twitter into your very own pursuit of employment stage. We have gathered the tweets identified with employments utilizing python language and machine learning algorithms and sentimental analysis to secure the position openings. In this paper, we show the use of sentimental analysis and how to associate with Twitter also, run sentimental analysis queries. We run investigates various questions openings for work and show the intriguing results. We understood that in certain queries neutral notion for tweets are essentially high which unmistakably shows the confinements of the present work.*

Key Words: **Sentimental Analysis, Social Media.**

1. INTRODUCTION

This Conclusion and sentimental mining is a significant research region on the grounds that because of the tremendous number of day by day posts on social networks, removing individuals' opinion is a difficult undertaking. Around 90% of the present information has been given during the last two a long time and getting understanding into this enormous scale information is not trifling [17, 18]. Wistful investigation has numerous applications for various areas for instance in organizations to get criticisms for items by which organizations can get familiar with clients' input and audits on social medias. Supposition and nostalgic mining has been well considered in this reference and every single various approaches furthermore, look into fields have been talked about [10]. There are likewise a few works have been done on Facebook [19-23] sentimental analysis anyway in this paper we for the most part center on the sentimental analysis of Twitter. For a bigger writings one arrangement could be comprehend the content, abridge it and offer load to it whether it is strongly positive, weakly positive, positive, neutral or strongly negative, weakly negative or negative.

Two basic ways to deal with remove content rundown are an extractive and abstractive technique. In the extractive strategy, words and word phrases are separated from the first content to create a rundown. In an abstractive strategy, attempts to gain proficiency with an inside language portrayal and afterward produces synopsis that is progressively like the rundown done by human. Content comprehension is a critical issue to unravel. Some ML methods, including supervised and unsupervised learning, are being used. There are various ways to deal with produce outline. One approach could be rank the significance of sentences inside the content and afterward produce synopsis for the content dependent on the significance numbers. There is another methodology called start to finish generative models. In a few space like picture acknowledgment, discourse acknowledgment, language interpretation, and question-replying, the start to finish strategy performs better. As of now, assessment investigation is a point of incredible premium and improvement since it has numerous functional applications. Organizations use assumption investigation to naturally break down overview reactions, item audits, online networking remarks, and so forth to get important bits of knowledge about their brands, item, and administrations.

Content comprehension is a huge issue to illuminate. Some, including different directed and unaided calculations, are being used. There are various ways to deal with create synopsis. One approach could be ranking the significance of sentences inside the content and afterward produce outline for the content dependent on the significance numbers. There is another methodology called start to finish generative models. In a few areas like picture acknowledgment, discourse acknowledgment, language interpretation, and question-replying, the start to finish strategy performs better.

A few works have utilized a cosmology to get it the content [1]. In the expression level, nostalgic investigation framework ought to have the option to perceive the extremity of the expression which is talked about by Wilson, et.al [9]. Tree kernel and highlight based model have been applied for nostalgic investigation in twitter by Agarwal and et.al [11]. SemEval-2017 [12] too shows the seven years of nostalgic investigation in twitter undertakings. Since tweets in Twitter is a particular dislike a typical book there are a few works that address this issue like the work for short casual writings [13]. Nostalgic investigation has numerous applications in job opportunities [14].

In this paper we will discuss analysis of social network and its importance then we'll discuss twitter being a rich source for sentimental analysis and followed by significant level and conceptual of our execution. We will show a few inquiries on various themes and show the polarity of tweets.

2. TWITTER ANALYSIS

Social network examination is the investigation of individuals 'connections and interchanges on various points also, these days it has gotten more consideration. Millions of individuals give their assessment of various themes consistently on social medias like Facebook and Twitter. It has numerous applications in various territories of research from sociology to business [3].

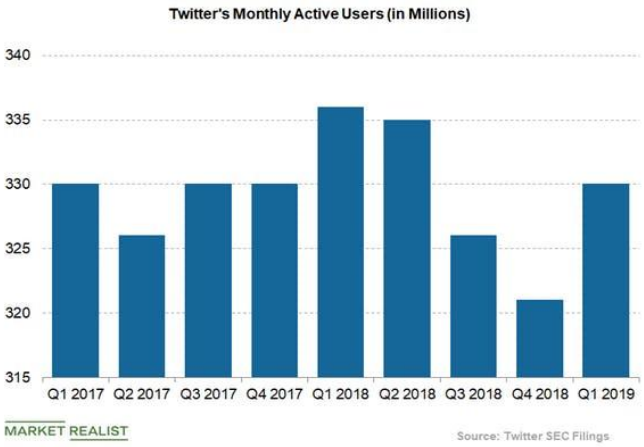


FIG1:Twitter monthly active users 2019

Twitter these days is one of the mainstream social media which according to the statistician [4] at present has more than 300 million active users monthly. Twitter is the rich source to find out about individuals' opinion and wistful investigation [2]. For each tweet it is essential to decide the feeling of the tweet regardless of whether is it strongly positive, weakly positive, positive, neutral, strongly negative, negative or weakly negative.

Another test with twitter is just 140 characters is the limitation of each tweet which cause individuals to utilize expressions and works which are definitely not in language preparing. As of late twitter has stretched out the content confinements to 280 characters for each and every tweet.

3. SENTIMENTAL ANALYSIS OF TWITTER

Informal organizations is a rich stage to find out about individuals' assessment and conclusion with respect to various themes as they can convey and share their supposition effectively on social medias including Facebook and Twitter. There are diverse supposition situated data gathering frameworks which mean to separate individuals' assessment with respect to various themes. The sentimental frameworks nowadays have numerous applications from business to sociologies.

Since interpersonal organizations, particularly Twitter contains little messages and individuals may utilize unique words and shortened forms which are hard to separate their feeling by current Common Language preparing frameworks effectively, in this manner a few analysts have utilized profound learning and machine learning procedures to concentrate and mine the polarity of tweet text [15].Some of the abbreviations are gr8 for great, FB for Facebook, Insta for Instagram,B4 for Before, OMG for Oh My god and many others. Hence Sentimental analysis of short messages like Twitter's posts is testing.

4. DESIGN AND IMPLEMENTATION

This specialized paper reports the execution of the Twitter estimation investigation, by using the API given by twitter.

There are incredible works and devices concentrating on content mining on interpersonal organizations. In this undertaking the riches of accessible libraries has been utilized.

The way to deal with separate sentiment from tweet is as per following

1. Start by downloading and caching the sentiment dictionary.

2. Download the testing data sets(tweets) and use them as input in program
3. Remove stop word to clean the test dataset of tweets.
4. Each word needs to be tokenized and then to be given as an input in the program.
5. Each word is then compared with positive, negative sentiment word and counter is incremented accordingly.
6. In the end based on positive and negative count polarity is extracted and is classified into strongly positive, positive, weakly positive, neutral, strongly negative, negative and weakly negative based on polarity.

Research has been done on sentimental analysis for different purposes for example work designed by Wang, et.al [5] is a real-time twitter sentimental analysis of the presidential elections.

Figure1.Shows high level abstract algorithm for sentimental analysis

There are different procedures to connect to twitter API (tweepy), fetch the tweets perform data preprocessing like removing stop words, tokenization, stemming and lemming. After that classify them on the basis of polarity obtained and return the final results.

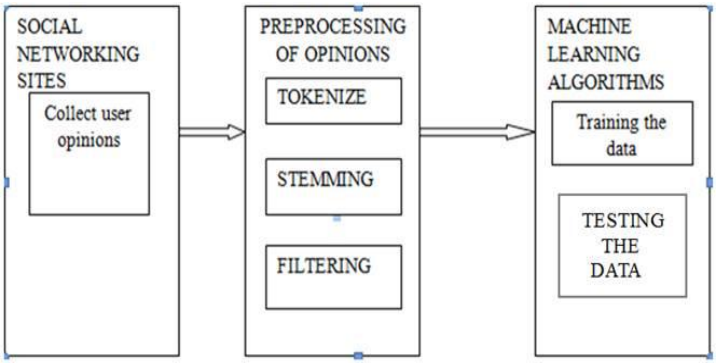


Fig.2 Algorithm for sentimental analysis

A. *IMPLEMENTAION*

For this paper we have used python as it has packages like TextBlob and tweepy that are used in sentimental analysis

Install these libraries by following commands.

- ***Python3 –m pip install textblob***
- ***Python3 –m pip install tweepy***

Next step is to download dictionary by using following command

Python3 –textblob.download_corpora

The python library for text processing textbolob uses NLTK for natural language processing [6].Corpora is large structured used to analyze tweets.

B. Connecting to Twitter API

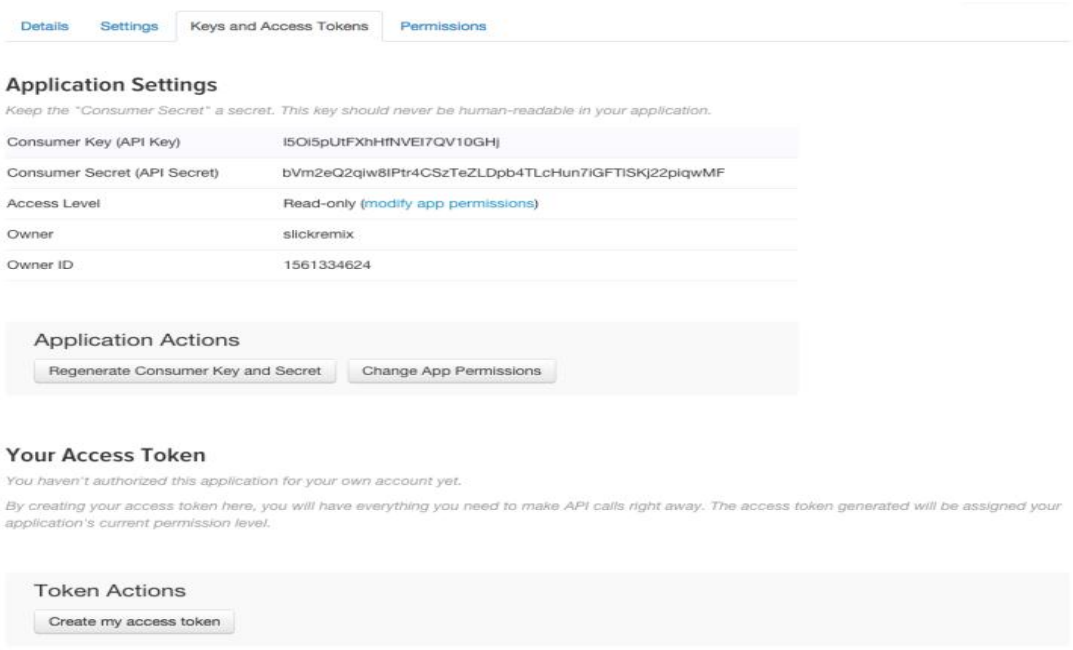


Fig2 Twitter Application Manager

After generating application using your twitter developer account we need to generate API keys there are 4 keys required consumer key, Consumer secret key, Access Token, Access token secret.

```
#Enter Keys
consumerKey=""
consumerSecret=""
accessToken=""
accessTokenSecret=""

auth= tweepy.OAuthHandler(consumerKey,consumerSecret)
auth.set_access_token(accessToken,accessTokenSecret)
api=tweepy.API(auth)
```

FIG3. Establishing connection with the API using keys

After generating keys we need to import them into our code so that connection xan be established with twitter API So that we can query the tweets regarding certain keywords and also no. of tweets we want to analyze.

C. Sample Output

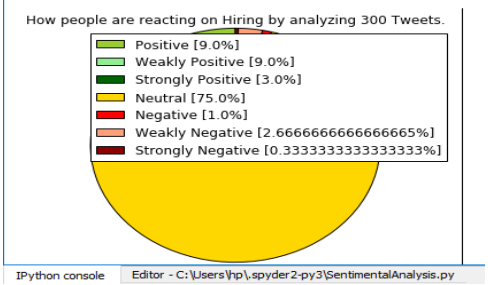
Following shows the sample outputs of the queries based on job opportunities

```
Enter keyword/hashtag to search about:Hiring

Enter no. of tweets to analyze:300
How people are reacting on Hiring by analyzing 300 tweets.

General Report:
Weakly Positive

Detailed Report:
9.0% people thought it was positive
9.0% people thought it was weakly positive
3.0% people thought it was strongly positive
1.0% people thought it was negative
2.6666666666666665% people thought it was weakly negative
0.3333333333333333% people thought it was strongly negative
75.0% people thought it was neutral
```

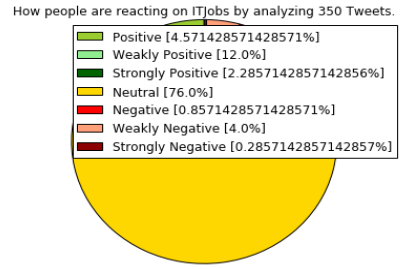



```
Enter keyword/hashtag to search about:ITJobs

Enter no. of tweets to analyze:350
How people are reacting on ITJobs by analyzing 350 tweets.

General Report:
Weakly Positive

Detailed Report:
4.571428571428571% people thought it was positive
12.0% people thought it was weakly positive
2.2857142857142856% people thought it was strongly positive
0.8571428571428571% people thought it was negative
4.0% people thought it was weakly negative
0.2857142857142857% people thought it was strongly negative
76.0% people thought it was neutral
```

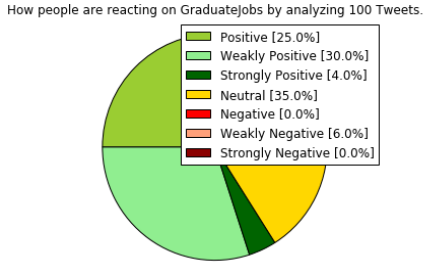



```
Enter keyword/hashtag to search about:GraduateJobs

Enter no. of tweets to analyze:100
.How people are reacting on GraduateJobs by analyzing 100 tweets.

General Report:
Weakly Positive

Detailed Report:
25.0% people thought it was positive
30.0% people thought it was weakly positive
4.0% people thought it was strongly positive
0.0% people thought it was negative
6.0% people thought it was weakly negative
0.0% people thought it was strongly negative
35.0% people thought it was neutral
```

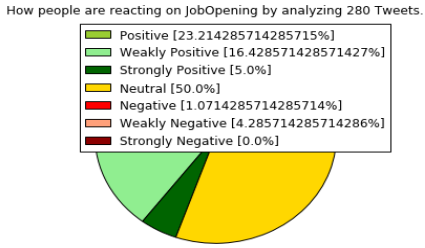



```
Enter keyword/hashtag to search about:JobOpening

Enter no. of tweets to analyze:280
How people are reacting on JobOpening by analyzing 280 tweets.

General Report:
Weakly Positive

Detailed Report:
23.214285714285715% people thought it was positive
16.428571428571427% people thought it was weakly positive
5.0% people thought it was strongly positive
1.0714285714285714% people thought it was negative
4.285714285714286% people thought it was weakly negative
0.0% people thought it was strongly negative
50.0% people thought it was neutral
```

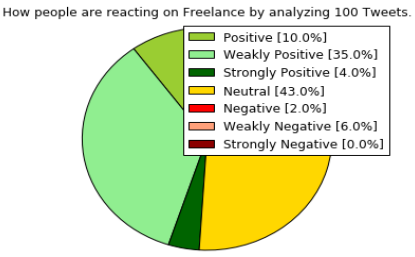



```
Enter keyword/hashtag to search about:Freelance

Enter no. of tweets to analyze:100
How people are reacting on Freelance by analyzing 100 tweets.

General Report:
Weakly Positive

Detailed Report:
10.0% people thought it was positive
35.0% people thought it was weakly positive
4.0% people thought it was strongly positive
2.0% people thought it was negative
6.0% people thought it was weakly negative
0.0% people thought it was strongly negative
43.0% people thought it was neutral
```

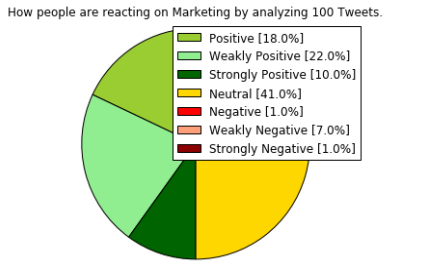



```
Enter keyword/hashtag to search about:Marketing

Enter no. of tweets to analyze:100
How people are reacting on Marketing by analyzing 100 tweets.

General Report:
Weakly Positive

Detailed Report:
18.0% people thought it was positive
22.0% people thought it was weakly positive
10.0% people thought it was strongly positive
1.0% people thought it was negative
7.0% people thought it was weakly negative
1.0% people thought it was strongly negative
41.0% people thought it was neutral
```



5. Conclusions

The pie chart illustrates sentiments of users regarding that tweet query.

It is observed that in above shown graphs percentage of neutral tweets are significantly higher in some queries.

This is because it depends upon the data under observation. We might get different results as opinions may change depending upon world circumstances.

In this paper we discussed importance of social network and their applications in different domains. We concentrated on Twitter as and have actualized the python program to actualize sentimental analysis.

We demonstrated the outcomes on openings for work related points and observed that neutral sentiments are essentially high which appears there is a need to improve twitter sentimental analysis.

REFERENCES

[1] Boguslavsky, I. (2017). Semantic Descriptions for a Text Understanding System. In Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference "Dialogue" (2017) (pp. 14-28).

[2] Pak, A., & Paroubek, P. (2010, May). Twitter as a corpus for sentiment analysis and opinion mining. In LREc (Vol. 10, No. 2010).

[3] Scott, J. (2011). Social network analysis: developments, advances, and prospects. Social network analysis and mining, 1(1), 21-26.

[4] Statista, 2017, https://www.statista.com/statistics/282087/number-ofmonthly-active-twitter-users/

[5] Wang, H., Can, D., Kazemzadeh, A., Bar, F., & Narayanan, S. (2012, July). A system for real-time twitter sentiment analysis of 2012 us presidential election cycle. In Proceedings of the ACL 2012 System Demonstrations (pp. 115-120). Association for Computational Linguistics.

[6] TextBlob, 2017, https://textblob.readthedocs.io/en/dev/

[7] Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. Foundations and Trends® in Information Retrieval, 2(1–2), 1- 135.

[8] Dos Santos, C. N., & Gatti, M. (2014, August). Deep Convolutional Neural Networks for Sentiment Analysis of Short Texts

[9] Wilson, T., Wiebe, J., & Hoffmann, P. (2005, October). Recognizing contextual polarity in phrase-level sentiment analysis. In Proceedings of the conference on human language technology and empirical methods in natural language processing(pp. 347-354). Association for Computational Linguistics.

[10] Liu, B. (2012). Sentiment analysis and opinion mining. Synthesis lectures on human language technologies, 5(1), 1-167.

[11] Agarwal, A., Xie, B., Vovsha, I., Rambow, O., & Passonneau, R. (2011, June). Sentiment analysis of twitter data. In Proceedings of the workshop on languages in social media (pp. 30-38). Association for Computational Linguistics.

[12] Rosenthal, S., Farra, N., & Nakov, P. (2017). SemEval-2017 task 4: Sentiment analysis in Twitter. In Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)(pp. 502-518).

[13] Kiritchenko, S., Zhu, X., & Mohammad, S. M. (2014). Sentiment analysis of short informal texts. Journal of Artificial Intelligence Research, 50, 723-762.

[14] Balahur, A., Steinberger, R., Kabadjov, M., Zavarella, V., Van Der Goot, E., Halkia, M., ... & Belyaeva, J. (2013). Sentiment analysis in the news. arXiv preprint arXiv:1309.6202.

[15] Poria, S., Cambria, E., & Gelbukh, A. (2015). Deep convolutional neural network textual features and multiple kernel learning for utterance-level multimodal sentiment analysis. In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (pp. 2539- 2544).

[16] Ortigosa, A., Martín, J. M., & Carro, R. M. (2014). Sentiment analysis in Facebook and its application to e-learning. Computers in Human Behavior, 31, 527-541.

[17] Bagheri, Hamid, and Abdusalam Abdullah Shaltooki. "Big Data: challenges, opportunities and Cloud based solutions." International Journal of Electrical and Computer Engineering 5.2 (2015): 340.

[18] Bagheri, Hamid, Mohammad Ali Torkamani, and Zhaleh Ghaffari. "Multi-Agent Approach for facing challenges in Ultra-Large Scale systems." International Journal of Electrical and Computer Engineering 4.2 (2014): 151.

[19] Ortigosa, A., Martín, J. M., & Carro, R. M. (2014). Sentiment analysis in Facebook and its application to e-learning. Computers in Human Behavior, 31, 527-541

[20] Feldman, R. (2013). Techniques and applications for sentiment analysis. Communications of the ACM, 56(4), 82-89.

[21] Dasgupta, S. S., Natarajan, S., Kaipa, K. K., Bhattacherjee, S. K., & Viswanathan, A. (2015, October). Sentiment analysis of Facebook data using Hadoop based open source technologies. In Data Science and Advanced Analytics (DSAA), 2015. 36678 2015. IEEE International Conference on (pp. 1-3). IEEE.

[22] Trinh, S., Nguyen, L., Vo, M., & Do, P. (2016). Lexicon-based sentiment analysis of Facebook comments in Vietnamese language. In Recent developments in intelligent information and database systems (pp. 263-276). Springer International Publishing. [23] Haddi, E., Liu, X., & Shi, Y. (2013). The role of text pre-processing in sentiment analysis. Procedia Computer Science, 17, 26-32.