|| Volume 6 || Issue 5 || May 2021 || ISSN (Online) 2456-0774
INTERNATIONAL JOURNAL OF ADVANCE SCIENTIFIC RESEARCH
AND ENGINEERING TRENDS

STUDY OF SENTIMENT ANALYSIS OF TWITTER USERS FROM THEIR TWEET'S TEXT AND DIFFUSION PATTERNS

Mr Nikhil Kumar Khokale[1], Prof. Rashmi Tundalwar[2]

PG Student, Computer Department, Dhole Patil College of Engineering, Pune, India[1]
Faculty, Computer Department, Dhole Patil College of Engineering, Pune, India[2]
Nikhilkhokale09@gmail.com[1], tundalwarrashmi@gmail.com[2]

--- ***--

Abstract: - A large number of people use social networking sites such as Twitter to express themselves. As a message of their human feelings, a tweet is considered. The sentiment analysis of user tweets was the subject of our research. Many studies have been conducted on this sentiment analysis, which only uses text found in the user's Tweets and produces good results from the small number of words in Twitter messages. A few studies have been done on this, showing that the emotions expressed in tweets are used to determine the personality of users and the polarities of tweets. In our research, we combined text message information from user tweets with sentiment distribution models to produce a more precise sentimental analysis from user tweets. We have used the miracle of mood change and mood-changing patterns to investigate the spread of emotions. We suggested random forest machine reading to predict the magnitude of the feelings and emotions that the user conveys in their tweets, using both textual information from the user's tweets and emotional and emotional distribution patterns. This is the first study to use emotional transmission methods to improve Twitter's emotional analysis, to our knowledge. Numerous studies in real-time dataset have shown that, compared with high-quality text analysis algorithms.

Keywords: - *Machine learning, Sentiment analysis, sentiment diffusion, Text Mining, Twitter*

--- ***--

I INTRODUCTION

Twitter is a social media platform that allows users to send and receive short messages called tweets. Individuals and organizations can use Twitter to send updates, known as tweets, telling subscribers or their followers what they think, do, or plan to do in the future. Users can communicate with other users on Twitter by retweeting their messages, which is similar to sharing someone's information with their followers. Since 2008, Twitter has become the most widely used and most influential online social networking platform on the planet. Because of the large volume of data available in the Twitter database due to the large number of users, the polarity of users' emotions and feelings conveyed in tweets has become a hot research subject due to its widespread usage. For example, analyzing the polarities of tweets from users in various sectors will be used to understand sector behavior and will be used for advertising and future improvement strategies. Similar private or government companies have used Twitter sentiment analysis as an important way to monitor consumer sentiment against their product and brand. The thoughts, emotions, and sentiments expressed in tweets were analyzed using textual details. Because of this, twitter data is sensitive because it stores important information. Data mining, using a variety of techniques, is required to apply this information from the data. Text-mining techniques, which can be done using natural language processing methods, can be used to obtain this information. Furthermore, the type of emotions must be used to evaluate critical data that must be mined. This is accomplished by the use of analytical emotions. Twitter is a social media platform where users can share their thoughts with the public through tweets. Twitter has over 330 million users worldwide and generates over 18000 pieces of data every second. Users' tweets may include statements, thoughts, news, and other types of sentences. This can be used as a Twitter feed with a lot of text and numbers. In general, other people's views are used as inputs to make decisions. By explicitly enquiring, this opinion can be verified. By enquiring about the decisions, because meeting the people who were supposed to ask requires time and effort. The other choice is to solicit feedback from Twitter users. Opinions can be expressed in the form of large-scale Twitter surveys. However, this viewpoint must be divided into three categories: neutral, positive, and negative. Furthermore, these tweets have not been organized into the groups you're looking for. As a result, it is still prevalent and appropriate.

II PROPOSED METHODOLOGY

We investigated mood swings, focused on tweets and their retweets with different types of emotions, and suggested a wide range of emotions in the data we collected from various Twitter users. Emotional modification was studied and suggested a predictive model for mood swings. We have developed a SentiDiff algorithm, which works repetitively, to

|| Volume 6 || Issue 5 || May 2021 || ISSN (Online) 2456-0774
INTERNATIONAL JOURNAL OF ADVANCE SCIENTIFIC RESEARCH
AND ENGINEERING TRENDS

predict the feeling of the strength of each message or tweet of a twitter message. This algorithm considered the interrelated relationship between tweet text, posts, and emotional distribution patterns when calculating algorithm performance. Limitations of predictable text-based views on emotion-based tweets were linked to the effect of mood swings, which increased the chances of tweets being properly separated from text-based data separator. In one case, the chances were slim. Emotional retrieval can be conveyed via text information from Twitter messages in this way.

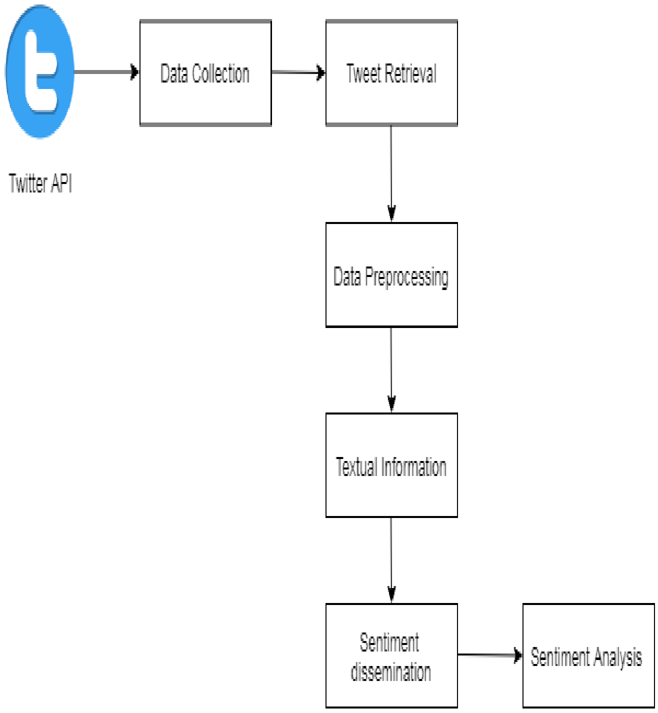

Figure -1: System Architecture

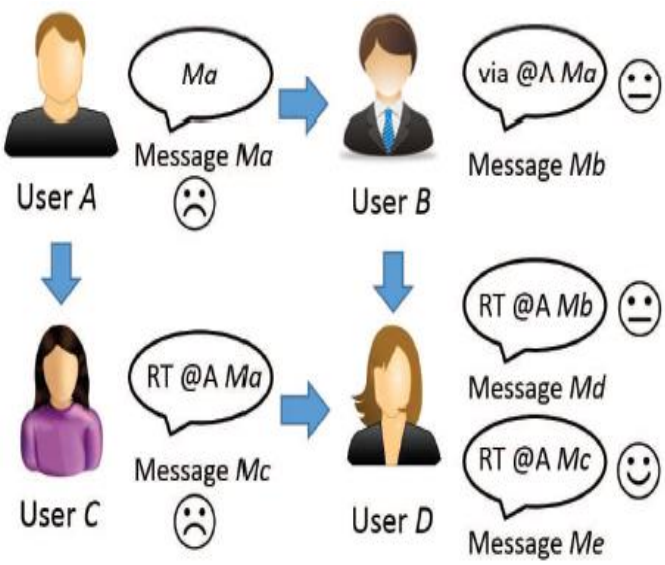

Figure -2: Example interactions among twitter users

A. Mathematical Model

For Twitter Sentiment Analysis, the mathematical model for integrating Textual Information and Sentiment Diffusion Patterns is as follows:

S= {I, F, O}

Where,

I = Set of inputs, input is the set of twitter tweets and retweets.

F = Set of functions

F= {F1, F2, F3}

F1: Textual Information, contents of tweets of user

F2: Repost Cascade Tree

The retweet cascade tree is a guided, acyclic labelled graph used to represent the relationships between tweets and their retweets.

T (V, E, l)

Where,

V – set of nodes,

E – Set of edges,

L – Function

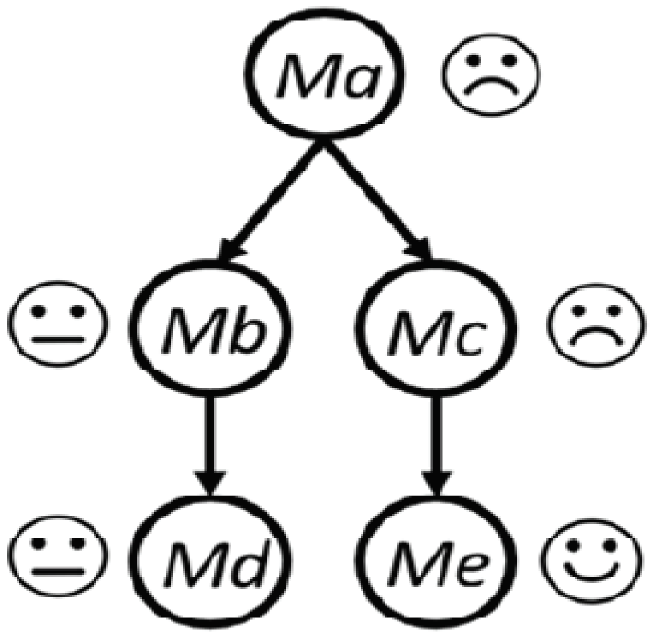

Figure -3: Report Cascade Tree

F3: Repost Diffusion Network

Repost diffusion networks are used to explain how Twitter users communicate with one another.

N (V, E)

Where,

V – set of nodes,

E – Set of edges,

|| Volume 6 || Issue 5 || May 2021 || ISSN (Online) 2456-0774
INTERNATIONAL JOURNAL OF ADVANCE SCIENTIFIC RESEARCH
AND ENGINEERING TRENDS

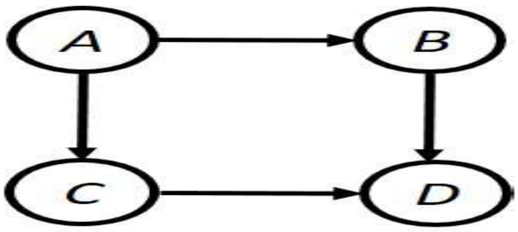

Figure -2: Report Diffusion network

F4: Sentiment Reversal

Sentiment reversal is the phenomenon in which the sentiment polarities of a tweet (parent tweet) and its retweet (child tweet) are reversed.

$$l(i) \neq l(j)$$

Where,

l – Functions

i – Parent tweet

j - Child tweet

O=Sentiment Analysis (i.e. Positive, Negative, Neutral)

TABLE 1 NOTATION USED IN THIS PAPER

Notation	**Meaning**
child(mi)	Twitter message mi, Child tweets
parent(mi)	Twitter message mi, Parent tweet
TL(mi)	Sentiment label of Twitter message mi anticipated through textual information-primarily based totally sentiment classifier
TP(mi, sl)	Probability of Twitter message mi to be classified with sentiment label sl efficaciously through textual records primarily based totally sentiment classifier
SP(mi , mj)	Probability that sentiment reversal happens among Twitter messages mi and mj anticipated through sentiment reversal prediction model
TSP(mi , sl)	Probability of Twitter message mi to be categorised with sentiment label sl after combining textual and sentiment diffusion information
FL(mi)	Sentiment label of Twitter message mi after combining textual and sentiment diffusion information

First, we use the named dataset to educate a sentiment classifier and a sentiment reversal prediction version primarily based totally at the textual statistics retrieved the use of herbal language processing. Following that, Algorithm 1 anticipated the sentiment polarity of every Tweeted message the use of a brand new dataset of Twitter messages from the equal cascade tree. If the predicted consequences among textual statistics primarily based totally sentiment classifier and sentiment reversal prediction version are in disagreement, the probability of Twitter messages being divided successfully with the aid of using textual content message primarily based totally sentiment classifier will decrease. Otherwise, the probability will rise. In the SentiDiff algorithm, the possibility that mi is split with sentiment label sl is first initialized as TSP(mi, sl), after which as TP(mi, sl), the possibility that mi is successfully categorized with sentiment label sl with the aid of using textual information based sentiment classifier. Then, the usage of an iterative method, we integrate textual data and sentiment diffusion data from mis discern and toddler tweets, in addition to the chances that sentiment reversals might arise amongst them. For Twitter message mi and its discern or toddler tweet mj, to be greater precise, If a textual data-primarily based totally sentiment classifier predicts that mi and mj have the identical sentiment polarity (i.e., TL(mi) = TL(mj) however a sentiment reversal prediction version predicts that mi and mj have one of a kind sentiment polarities (i.e., SP(mi, mj) >= 0.5), we agree with the consequences of the textual data-primarily based totally sentiment classifier are in battle with the consequences of the sentiment version. Similarly, we agree with the findings of the textual data-primarily based totally sentiment classifier and the sentiment reversal prediction version are incompatible while mi and mj are expected to bring one of a kind sentiment polarities (i.e., TL(mi) now no longer = TL(mj)), however the sentiment reversal prediction version predicts that sentiment reversal does now no longer arise among mi and mj. (i.e., SP(mi, mj) < 0.5). TSP(mi, TL(mi)) will lower because the probability of mi being successfully labelled with sentiment mark TL(mi) decreases.

TSP(mi, TL(mi)) 🡨 TSP(mi, TL(mi)) - TP(mi, TL(mi)) * TP(mj , TL(mj)) * SP(mi, mj)

where TL(mi) is the sentiment mark predicted by a textual information-based sentiment classifier, and SP(mi, mj) is the likelihood of sentiment reversal between mi and mj. Meanwhile, the likelihood of mi being correctly labelled with another sentiment mark sl -> -1, 0, 1 - TL(mi) will increase, TSP(mi, sl). It's worth noting that the number of probabilities of mi being listed with all possible sentiment labels must remain constant. As a result, the likelihood of mi being labelled with sentiment label sl increases for each of the other two potential sentiment labels for mi, sl -1, 0, 1 - TL(mi).

|| Volume 6 || Issue 5 || May 2021 || ISSN (Online) 2456-0774
INTERNATIONAL JOURNAL OF ADVANCE SCIENTIFIC RESEARCH
AND ENGINEERING TRENDS

TSP(mi, sl) ← TSP(mi, sl) + 0.5 * TP(mi, TL(mi))

* TP(mj , TL(mj)) * SP(mi, mj)

Similarly, if the sentiment polarities of mi and mj anticipated with the aid of using a textual information-primarily based totally sentiment classifier are like minded with the sentiment reversal prediction result, the chance of mi being effectively labeled with sentiment label TL(mi) increases, even as the chance of mi being labeled with different sentiment labels decreases.

As turned into carried out previously, we use the L∞ (max) norm of the distinction of VTS over consecutive iterations to be much less than β = 0:001 as our new release terminating state. Finally, the choice characteristic of the sentiment category trouble for Twitter message mi is seen. FL(mi) ← arg max (sl Ɛ{-1,0,1}) * TSP(mi, sl)

We'd like to point out that our proposed SentiDiff algorithm is a general system that can be used to combine different textual information-based sentiment classifiers.

III CONCLUSION

The polarity of mining sentiments expressed withinside the tweets is essential and tough to deal with. Because of the ordinary traits of Twitter messages, modern-day Twitter sentiment evaluation strategies rely entirely on textual content statistics from user-generated tweets and are not able to supply best results. Existing studies has proven that styles of emotion diffusion are carefully connected to the polarities of twitter messages; however, current strategies are mainly centered on textual content observed in tweet messages. However, it ignores the dissemination of emotional content. We have been prompted with the aid of using all the latest research and paintings at the convergence of information from numerous domains, and we took step one towards integrating textual statistics and spreading emotions to enhance Twitters sentiment evaluation output.

REFERENCES

[1] S. Symeonidis, D. Effrosynidis, and A. Arampatzis, “A comparative evaluation of pre-processing techniques and their interactions for twitter sentiment analysis,” Expert Systems with Applications, 2018

[2] J. Zhao and X. Gui, “Comparison research on text pre-processing methods on twitter sentiment analysis,” IEEE Access, vol. 5, pp. 2870–2879, 2017.

[3] K. Schouten and F. Frasincar, “Survey on aspect-level sentiment analysis,” IEEE Transactions on Knowledge and Data Engineering, vol. 28, no. 3, pp. 813–830, 2016.

[4] X. Zhang, D.-D. Han, R. Yang, and Z. Zhang, “Users participation and social influence during information spreading on twitter,” PloS one, vol. 12, no. 9, p. e0183290, 2017.

[5] Lei Wang, Jianwei Niu, Senior Member, IEEE, and Shui Yu, Senior Member, IEEE “SentiDiff: Combining Textual Information and Sentiment Diffusion Patterns for Twitter Sentiment Analysis” 2019