International Journal of Scientific Research in Science, Engineering and Technology

Print ISSN - 2395-1990
Online ISSN : 2394-4099

OPEN ACCESS

Available Online at : www.ijsrset.com
doi : https://doi.org/10.32628/IJSRSET

Covid-19 Related Sentiment Analysis on Twitter data using Machine Learning based Technologies

[1]T.Veena, [2]G. Medhaswi, [3]B. Vani, [4]P. Asha Jyothi, [5] S. Khaja Mohidden
[1]Assistant Professor, [2,3,4,5] UG Students
Department of CSE, Sri Vasavi Institute of Engineering And Technology, Nandamuru, Andhra Pradesh, India

ARTICLEINFO

Article History:

Accepted: 15 April 2024
Published: 22 April 2024

Publication Issue :
Volume 11, Issue 2
March-April-2024
Page Number :
417-421

ABSTRACT

During the crisis situation caused due to COVID-19 disease, managing mental health and psychological well-being is as important as physical health of people. As web based life is broadly utilized by individuals to communicate their feeling and supposition, our framework utilizes Twitter information posted by individuals during this emergency circumstance to dissect the feelings of individuals. For processing the cleaned data NRC Word-Emotion Association Lexicon (have aka EmoLex) is used. NRC Word-Emotion Association Lexicon is a list of English with real-valued scores of intensity for eight basic emotion words ns (anger, anticipation, disgust, fear, joy, sadness, surprise, and trust). The text content of tweeter dataset created by fetching tweets across the world have classified into basic emotions like anger, anticipation, disgust, fear, joy, sadness, surprise and trust. This analysis can be used by authorities to understand the mental health of the people and can take necessary measures to decide on policies to fight against coronavirus which is affecting the social well-being as well as economy of the whole world.

Keywords : COVID-19;Coronavirus; Emotions; Twitter; Tweets

I. INTRODUCTION

During the crisis situation caused due to COVID-19 disease, managing mental health and psychological well-being is as important as physical health of people. As web based life is broadly utilized by individuals to communicate their feeling and supposition, our framework utilizes Twitter information posted by individuals during this emergency circumstance to dissect the feelings of individuals. During this pandemic situation managing mental health and psychological well-being is as important as physical health. Understanding public opinion and emotions can help the Ministry of Health and family Affairs to take necessary actions. In this fast growing digital world, social networking services like Twitter, Facebook, Instagram are becoming a very popular medium of communication and a medium to express oneself. The five key characteristics of social media:

collectivity, connectedness, completeness, clarity and collaboration lend itself to be used increasingly to support crisis management function.

The scope of paper is restricted to only text content in twitter. But if closely view the social media content, many people are also using various emojis and graphics to convey their emotions. Further the system can be extended to consider such graphics and emojis in analyzing the sentiment.

II.RELATED WORK

Emotion analysis or sentiment analysis is becoming a wide area of research in different applications. In our proposed work, twitter data related to COVID -19 is analyzed. Many researchers have worked on emotional analysis on data collected from their local micro blogging site. Wang et al have performed an emotional analysis method on COVID19 data collected from Sina Weibo ,a microblogging site in China using Support Vector Machine, naïve Bayes and Random Forest classifiers. Mansur ALP et al classified Turkish Tweets into basic emotions using Deep Neural Network. Collected from twitter data across the world. Emotional analysis approach using social media is used by various researchers during natural disasters and for tracking pandemics. Hyo Jin Do et al used emotional analysis technique to understand the behavior of humans and the characteristics of the socio culture system during epidemic outbreak in Korea. Vasileios Lampos et al have used an approach of analyzing the twitter data to monitor the diffusion of pandemic in people. Rashid Kamal et al proposed a crowd source sensing technique using twitter data and word model which can classify the emotions and proved to be more efficient when analysis is done based on location. In many twitter dataset available, location from which tweet is sent is not mentioned. In such situations word models help in accurate classification of emotions. Twitter provides the service to people to share their emotions , Kiichi Tago et al proposed the influence of emotional behavior on user relationships based on Twitter data. Felipe Taliar performed emotional analysis on data collected from facebook and studied facebook reactions classified them into positive, negative and neutral and examined if they are related to Ekman's emotion classification. Mondher Bouazizi et al proposed a multiclass emotion classification using feature extraction. In their research they have used a tool called SENTA. Harshvardhan Achrekar et al proposed a method to predict influenza epidemic situations in the real world using Twitter data.

III.PROPOSED SYSTEM

Proposed work, twitter data related to COVID -19 is analyzed. Many researchers have worked on emotional analysis on data collected from their local microblogging site. Our framework utilizes Twitter information posted by individuals during this emergency circumstance to dissect the feelings of individuals. We used Multinomial naive Bayes and Random Forest classifiers for classifying the emotions. The text content of tweeter dataset created by fetching tweets across the world have classified into basic emotions like anger, anticipation, disgust, fear, joy, sadness, surprise and trust.

ADVANTAGES

- This analysis can be used by authorities to understand the mental health of the people.
- Understanding public opinion and emotions can help the Ministry of Health and family Affairs to take necessary actions.

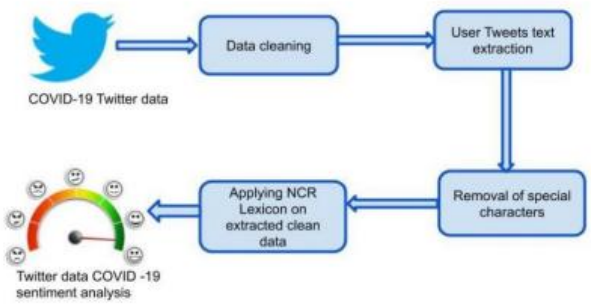

Fig 1: System Architecture

IV. RESULTS AND DISCUSSION

```
twitter_data.sentiment.value_counts().plot(kind = 'bar')
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f74d6140048>
```

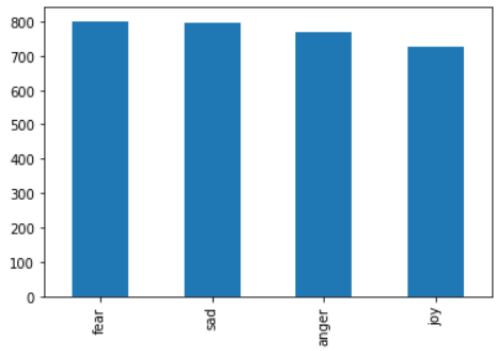

Fig 2. Result Screenshot

```
[81] print(new_twitter_data.emottions.value_counts())
     sns.countplot(x='emottions', data = new_twitter_data);
```

```
negative    1975
positive    1115
Name: emottions, dtype: int64
```

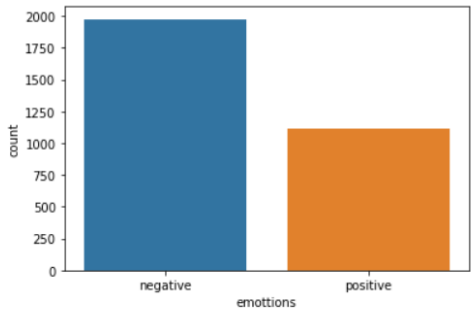

Fig 3. Result Screenshot

[69]

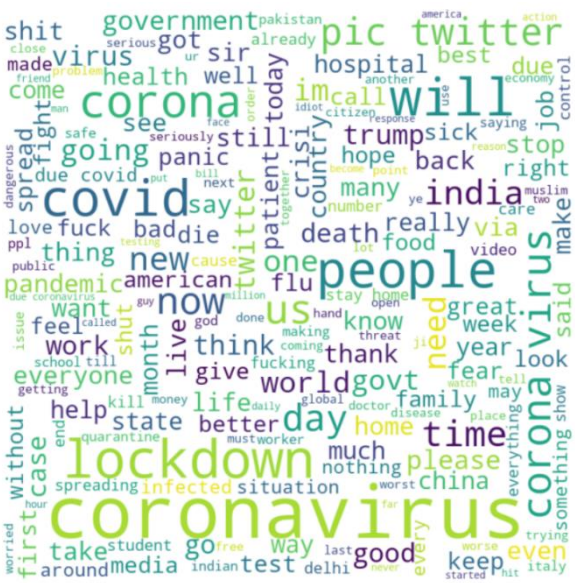

Fig 4. Result Screenshot

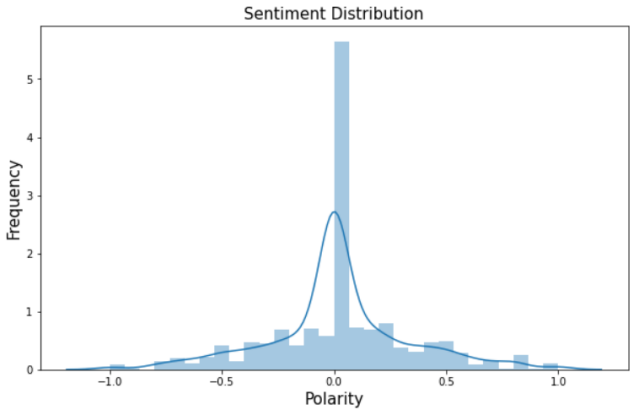

Fig 5. Result Screenshot

```
from sklearn.naive_bayes import MultinomialNB
mnb = MultinomialNB()
mnb.fit(X1_train, Y1_train)

pred = mnb.predict(X1_test)
from sklearn.metrics import confusion_matrix, classification_report
print(confusion_matrix(Y1_test, pred))
print('\n')
print(classification_report(Y1_test, pred))
print("Multinomial Accuracy : {}".format(mnb.score(X1_train, Y1_train)))
```

```
[[485 122]
 [104 216]]


              precision    recall  f1-score   support

    negative       0.82      0.80      0.81       607
    positive       0.64      0.68      0.66       320

    accuracy                           0.76       927
   macro avg       0.73      0.74      0.73       927
weighted avg       0.76      0.76      0.76       927

Multinomial Accuracy : 0.9736477115117892
```

Fig 6. Result Screenshot

```
[87] text1=["""agree the poor in india are treated badly their poors seek a
     living in singapore and are treated like citizens they are given free medical treatment given food daily sim cards to call home to
      tell their family that they are fine if covid 19 case treated foc in hospitals"""]

[88] text1=count_vect.transform(text1)

[89] mnb.predict(text1)

     array(['negative'], dtype='<U8')
```

Fig 7. Result Screenshot

```
from sklearn.naive_bayes import MultinomialNB
nb = MultinomialNB()
nb.fit(X_train, y_train)

preds = nb.predict(X_test)
from sklearn.metrics import confusion_matrix, classification_report
print(confusion_matrix(y_test, preds))
print('\n')
print(classification_report(y_test, preds))
print("Multinomial Accuracy : {}".format(nb.score(X_train,y_train)))
```

```
[[123  69   5  24]
 [ 40 168   7  19]
 [  6  15 154  49]
 [ 14  21  21 192]]


              precision    recall  f1-score   support

       anger       0.67      0.56      0.61       221
        fear       0.62      0.72      0.66       234
         joy       0.82      0.69      0.75       224
         sad       0.68      0.77      0.72       248

    accuracy                           0.69       927
   macro avg       0.70      0.68      0.69       927
weighted avg       0.70      0.69      0.69       927

Multinomial Accuracy : 0.9639389736477115
```

Fig 8. Result Screenshot

V.CONCLUSION

In this paper, Twitter Data is analyzed across the world to understand the mental health of people during the COVID-19 pandemic situation through emotion analysis and classified it into basic emotions. As seen on the official websites of WHO and the Health Ministry, a separate tab is created for behavioral and psychological health which includes various audios and videos to keep people mentally healthy. By using our analysis, authorities can better understand the mental health of the people and update the content accordingly. Analysis is done on data across the world as although people live in different countries, having different physical and cultural environments, they still have the same basic emotions.

VI. FUTURE WORK

This analysis will give the good results in emotional analysis The scope of paper is restricted to only text content in twitter. But if closely view the social media content, many people are also using various emojis and graphics to convey their emotions. Further the system can be extended to consider such graphics and emojis in analyzing the sentiments.

II. REFERENCES

[1]. H. J. Do, C. G. Lim, Y. J. Kim, and H. J. Choi, "Analyzing emotions in twitter during a crisis: A case study of the 2015 Middle East Respiratory Syndrome outbreak in Korea," 2016 Int. Conf. Big Data Smart Comput. BigComp 2016, pp. 415–418, 2016, doi: 10.1109/BIGCOMP.2016.7425960.

[2]. M. A. Tocoglu, O. Ozturkmenoglu, and A. Alpkocak, "Emotion Analysis from Turkish Tweets Using Deep Neural Networks," IEEE Access, vol. 7, pp. 183061–183069, 2019, doi: 10.1109/ACCESS.2019.2960113.

[3]. A. Giachanou and F. Crestani, "Like it or not: A survey of Twitter sentiment analysis methods," ACM Comput. Surv., vol. 49, no. 2, 2016, doi: 10.1145/2938640.

[4]. L. Li et al., "Characterizing the Propagation of Situational Information in Social Media during COVID-19 Epidemic: A Case Study on Weibo," IEEE Trans. Comput. Soc. Syst., vol. 7, no. 2, pp. 556–562, 2020, doi: 10.1109/TCSS.2020.2980007.

[5]. V. Lampos and N. Cristianini, "Tracking the flu pandemic by monitoring the social web," 2010 2nd Int. Work. Cogn. Inf. Process. CIP2010, pp. 411–416, 2010, doi: 10.1109/CIP.2010.5604088.

[6]. R. Kamal, M. A. Shah, C. Maple, M. Masood, A. Wahid, and A. Mehmood, "Emotion classification and crowd source sensing; A lexicon based approach," IEEE Access, vol. 7, pp. 27124–27134, 2019, doi: 10.1109/ACCESS.2019.2892624.

[7]. K. Tago and Q. Jin, "Influence analysis of emotional behaviors and user relationships based on Twitter data," Tsinghua Sci. Technol., vol. 23, no. 1, pp. 104–113, 2018, doi: 10.26599/TST.2018.9010012.

[8]. F. T. Giuntini et al., "How do i feel? Identifying emotional expressions on facebook reactions using clustering mechanism," IEEE Access, vol. 7, pp. 53909–53921, 2019, doi: 10.1109/ACCESS.2019.2913136.

[9]. M. Bouazizi and T. Ohtsuki, "Multi-class sentiment analysis on twitter: Classification performance and challenges," Big Data Min. Anal., vol. 2, no. 3, pp. 181–194, 2019, doi: 10.26599/bdma.2019.9020002.

[10]. H. Achrekar, A. Gandhe, R. Lazarus, S. H. Yu, and B. Liu, "Predicting flu trends using twitter data," 2011 IEEE Conf.