

Detecting Stress Based On Social Interaction In Social Networks

Manisha Mirchandani
Department of Computer
Engineering
VES Institute of Technology
Chembur, Mumbai
manisha.mirchandani@ves.ac.in

Shreya Narsapur
Department of Computer
Engineering
VES Institute of Technology
Chembur, Mumbai
shreya.narsapur@ves.ac.in

Gayatri Pawar
Department of Computer
Engineering
VES Institute of Technology
Chembur, Mumbai
gayatri.pawar@ves.ac.in

Kajal Rajani
Department of Computer Engineering
VES Institute of Technology
Chembur, Mumbai
kajal.rajani@ves.ac.in

Prof. Indu Dokare
Department of Computer Engineering
VES Institute of Technology
Chembur, Mumbai
indu.dokare@ves.ac.in

Abstract: With the increase in the pace of life, it can be safely said the the stress level amongst people is rapidly increasing. Timely help, if not provided to the stress problems of any individual, can cause dire results. In the last few decades, social media has turned out to be the most powerful means of communication. With the growing use of social media, the amount of emotions which an average person shares online has also risen swiftly. So, it would be right to assume that the person's stress level can be very well be deduced from his social media account. The way to do this would be through Sentiment Analysis. By analyzing a person's social media feed, i.e., going through the posts (and emoticons attached in the post) which a person shares, we can identify the state of mind of the person while updating those posts. This system would be hugely successful in determining the stress level of an social media user.

Keywords: Stress, Social Media, Sentiment Analysis.

I. INTRODUCTION

Social media can be considered as a goldmine for Sentiment Analysis. An active social media user shares the majority of the details of his/her life online. This accounts for a large database on which algorithms can be applied for analysis. There are various social networking sites these days namely, FaceBook, Twitter, LinkedIn, Pinterest, Instagram, Tumblr, Reddit, Snapchat, Whatsapp, Quora etc. The fact that there is a need for so many different platforms itself speaks volumes about the number of social media users.

The platform chosen by us is Twitter, the reason being its limited word count per tweet (140 characters). This will result

in not a smaller dataset, but a definite searching space for keywords. The tweet can be then used for performing Sentiment Analysis and the sentiment of the particular tweet can be extracted. Moreover, there will be certain intensity values or say scores which will be assigned to the keywords. On computation of the total of these scores, a final result will be determined, which will be then used to diagnose the stress level of the user.

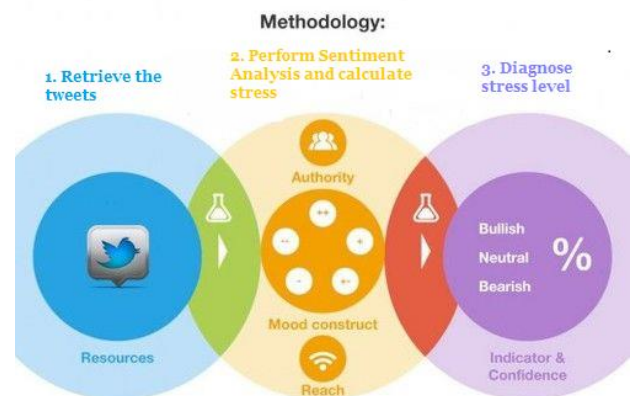


Fig 1: The process of Stress Detection

II. METHODOLOGY

Twitter has over 328 million monthly active users. This makes up for a very large database to perform Natural

Language Processing on. To study these tweets, we use an algorithm which is made up of 3 algorithms.

These algorithms are:

1. The first part of the algorithm aims at retrieving the tweets which have certain keywords in them.
2. The second part of the algorithm performs Sentiment Analysis on the tweets found in the first step.
3. The third and final step calculates the resultant score based on the intensity values assigned to the keywords.

The algorithm works as follows:

1. **Tweet retrieval:** There will be a dataset of keywords and emoticons which will be pre-saved into our system. This step will retrieve all those tweets which contain any one or more of these keywords and/or emoticons.

The reason for including emoticons is that they have become a very famous means of dialogue and quite a lot of times, people just reply or communicate only using them.

To carry out the process of retrieving the tweets from Twitter, a password, known as Twitter API key will be required which can be found from Twitter at request.

Given below is a sample of these keywords and emoticons:





































POSITIVE	NEGATIVE
Happy Excited Joyous Gleeful Satisfactory Funny Beaming	Unhappy Sorrowful Dejected Glum Gloomy Dismal Blue
POSITIVE	NEGATIVE
                 	                 

Fig 2: Keyword and Emoticon Table

2. **Sentiment Analysis:** The same words can be used in a variety of places to mean a lot of different things. We will need to determine what a particular keyword means in the context of the particular tweet or conversation.

We will also be required to stem the words in order to determine the root words which will then be used for the purpose of analysing.

The keywords will be segregated into three categories: positive, negative and neutral. Then, based on these scores, a final score will be computed which will be known as the complex result.

3. **Result declaration:** The complex result calculated in the above step will be the one which will be taken into account while declaring the final result about the stress of the particular user. There will be three categories in which these scores will be divided: unstressed, tensed and stressed. The highest scores will correspond to the unstressed category where the user, according to the tweets in analysis, is happy and satisfied. The tensed category will correspond to the moderate scores and it will mean the user is not stressed at the moment but proactive care needs to be taken. The stressed category will correspond to lowest scores and the users who will be in this category will be considered to tweet such messages or simply, tweets, which indicate that he is highly worried and stressed in his/her life.

III. ALGORITHM AND WORKING

1. The tweet retrieval step will involve getting all the tweets which contain the keywords. The input for this step are the keywords and the tweets along with the API key received from Twitter.
The twitter of the user according to the given username will be scanned and the tweets containing the keywords will be returned as the output.
Since ours will be an application based API authentication key, the number of tweets returned will be 15 per minute. These tweets will be given as an output and will be given to the next steps for further analysis.
2. The second step is the most important Sentiment Analysis step. Here, the input is the tweets which we have received from the previous step. As the keywords are classified into positive, negative and neutral, the scores will be assigned to each of these keywords.
The positive, negative and neutral keywords will have decimal valued scores ranging from 0-1. The positive and neutral keywords will be multiplied by a positive coefficient while the negative keywords will be multiplied by a negative coefficient and these resultant values will be added such that the final complex result will be a value between -1 and 1. This complex result, along with the positive, neutral and negative scores will be the output for this step.
3. The final step will take the complex result as the input and segregate the value into one of the three categories, unstressed, tensed and neutral.
After segregating the user into the particular category, reactive measures will be suggested for the users coming under the stressed and tensed categories.
A graph, indicating the stress levels for users of all the three categories will be given to them as an output.

IV. ADVANTAGES AND DISADVANTAGES OF THE SYSTEM

A. Advantages:

- Easy stress detection technique based on very little input by the user.
- Non-invasive in nature.
- Will help in detection of victims of cyber crime; eg: the recent Blue Whale Game fiasco.

B. Disadvantages:

- The user should possess a Twitter account and be an active user of the same.
- The application is not reliable in cases of very critical subjects.

V. CONCLUSION

In today's world, where mainly the youth and almost all of the population is suffering from surmounting stress, be it because of peer pressure, work load or other domestic tensions; it is very crucial to have a reality check about how stressed a person really is.

It is because of this reason that timely detection and prevention of stress is a dire need.

We have utilised both human as well as machine learning and applied the concepts of Sentiment Analysis. The main characteristic of this system is its non-invasiveness and fast-oriented implementation in detecting stress when compared with the previous approaches.

We have come up with this project which assists people in scrutinizing the problem of stress. This project will be very beneficial for those who are not so comfortable in opening up about their problems to others. It will help these people get a reality check and may prompt them to reach out and get medical help, just based on their social interactions.

It is because of this reason that timely detection and prevention of stress is a dire need.

We have utilised both human as well as machine learning and applied the concepts of Sentiment Analysis. The main characteristic of this system is its non-invasiveness and fast-oriented implementation in detecting stress when compared with the previous approaches.

VI. REFERENCES

- [1] S. Begum, M. U. Ahmed, P. Funk, N. Xiong, B. von Schelm Minor "Using calibration and fuzzification of cases for improved diagnosis and treatment of stress", in 8th European Conference on Case-based Reasoning workshop proceedings, pp. 113-122, September 2006.
- [2] L. A. Zadeh, "Fuzzy logic = computing with words", IEEE Transactions on Fuzzy Systems, vol. 4, no. 2, pp. 103-111, May 1996.
- [3] J. Andren, P. Funk, "A case- based approach using behavioral biometrics to determine a user's stress level", ICCBR Workshops, pp. 9-17, 2005.
- [4] Xiaojun Chang, Yi Yang, Alexander G Hauptmann, Eric P Xing, and Yao-Liang Yu. Semantic concept discovery for large-scale zero-shot event detection. In Proceedings of International Joint Conference on Artificial Intelligence, pages 2234–2240, 2015.
- [5] Dan C Ciresan, Ueli Meier, Jonathan Masci, Luca Maria Gambardella, and Jurgen Schmidhuber. Flexible, high performance " convolution neural networks for image classification. In Proceedings of International Joint Conference on Artificial Intelligence, pages 1237–1242, 2011.
- [6] Sung Ju Hwang. Discriminative object categorization with external semantic knowledge. 2013.