**AUTHOR:** S. T. Sadasivuni and Y. Zhang

**NAME:** Saahil Karnik

**CLASS:** CSC 4350

**SEMESTER:** Fall 2020

**Software Engineering for Machine Learning:** A New Method for Discovering Daily Depression from Tweets to Monitor People’s Depression Status

**SUMMARY**

Problem Statement

In today’s world, social media has become the over-arching mode of communication for everyone. Social Networks are that media that is involved in every person’s' life to share/exhibit their emotions and feelings. More people share emotion-related tweets indicate that a predominant feature occurred on that day or in that location. Studies have revealed that usage of the internet has a direct impact on user happiness and observed that heavy internet users are twice unhappy than others. Today, mental health affects more than 300 million (>4%) people, globally, and this number is rapidly climbing. Many mental health illness patients do not agree that they are ill, for fear of stigma, which makes proper identification of patients complicated.

Proposed Solution

Many users post information on social media and the authors considered that posting of information is uniform across all the days. Different people will post in different situations, and this depends on various parameters. The authors attempted to study day-to-day posts related to depression and those related to anti-depression and computed a parameter C1, which indicates the depressive level of that day. The authors chose Twitter for their data, because of its popularity.

Results

The authors conclude that higher the ‘CI’ values relate to greater depression levels.

**CRITIQUE**

Overall, my opinion about the paper is positive. The authors have touched upon a very relevant subject (mental health) and have made an apt observation about its co-relation to social media usage.

A strength of the paper is the method used for data mining (identifying negative/positive tweets using Kessler’s ten-point questionnaire). Another aspect I appreciate is the in-depth analysis of data using confusion matrices, that give a summary of prediction results on a classification problem and also insight on errors made by the classifier.

One minor flaw is that owing to sarcasm, some tweets with positive keywords may belie a negative state of mind of the poster, which might contribute to noise in the data. Even so, the use of Precision, Recall, F-1 Score, and Mathews Correlation Coefficient (MCC) help to minimize this noise.

I feel that there is a class of social media users who simply absorb information via social media, but refrain from posting any kind of information. It would be difficult to lock down their state of mental health with the method utilized in this paper. If a more comprehensive idea of mental health is to be tracked, I would suggest tracking the account activities of a sample of social media users. (tracking likes, dislikes, views, etc.)

**SYNTHESIS**

I would suggest that this work be carried forward by comparing diff C1 values (depression parameter) obtained at different areas and during different periods to track depression and/or mental health deterioration. Tracking user activity can also help pinpoint toward the causes of mental health deterioration. Interpretation of these results or patterns could give an insight to economists, visionaries to plan and execute societal programs in those areas and during those periods.

**BIBLIOGRAPHY**

S. T. Sadasivuni and Y. Zhang, "A New Method for Discovering Daily Depression from Tweets to Monitor Peoples Depression Status," *2020 IEEE International Conference on Humanized Computing and Communication with Artificial Intelligence (HCCAI)*, Irvine, CA, USA, 2020, pp. 47-50, doi: 10.1109/HCCAI49649.2020.00013.