IST 736: Text Mining: Topic Modeling of Tweets
Teresa Cameron
SUID: 386327244 tmcamero@syr.edu

Introduction

In the United States, tens of millions of people each year are affected by mental illness. It is estimated only half of people with mental illnesses receive treatment. Left untreated, mental illness can lead to disability and suicide. Part of the problem with treating mental illness is the pressure from society to keep the suffering private.

Over time, the stigma of mental health is slowing dissolving. People who suffer from mental illness and their loved ones are no longer kept silent about their suffering and have access to more resources than before. One of the resources available is social media.

Although evidence suggests excessive amounts of time on social media can increase mental health issues, it can also help people stay connected and supported. People can seek or offer emotional support, find outlets for self-expression, form networks with people, and discover valuable information. Material about mental health issues can be distributed to the public to further reduce the social stigma and guide people to proper resources.

In order to understand how people use social media for mental health, patients and health care providers can analyze tweets about mental health for topics.

Analysis & Models

About the Data

The data consists of 220 tweets that used the hashtag @mentalhealth, which were obtained using an API. The information gathered from Twitter was the UserName, Tweet, Time and Date, and Location. The tweets were tokenized, and a CSV file was created that had columns for the above information plus the original tweet and the tokenized tweet.

The top 20 usernames by number of tweets is displayed below as Figure 1. Most of the users had only one tweet and the user with the most tweets was ManinMortlake. This individual was conducting a fundraiser for mental health during the time period the tweets were gathered.

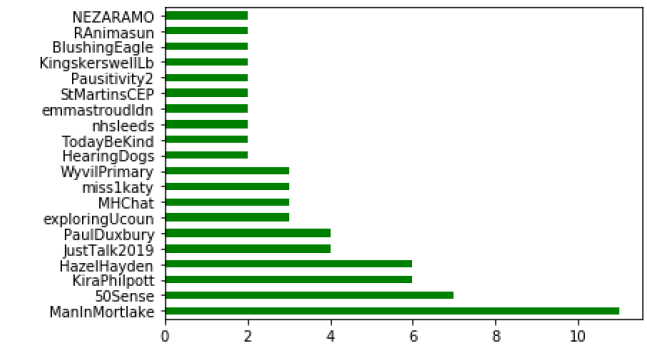

Figure 1

The locations of the tweets were analyzed using a bar graph shown as Figure 2. Most of the tweets were not labeled, but of those labeled, most were from United Kingdom.

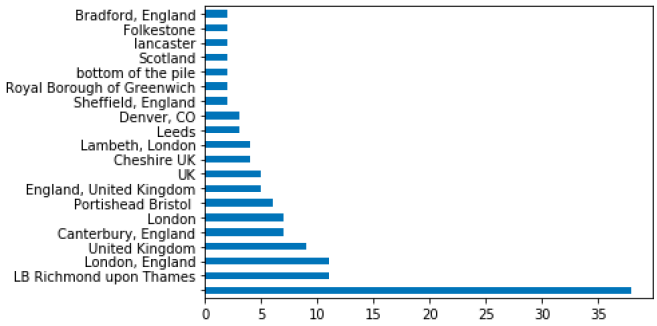

Figure 2

After analyzing the usernames and locations, a top word frequency bar graph was created with the tokenized tweet data. Figure 3 shows after mentalhealth, health, and mental, the next most frequent words were mindcharity, help, support, and positive words such as good, thanks, kindness, love, and working. Mind is a mental health charity in England and Wales.

The words children and women were also mentioned possibly showing a concern for how mental health affects people from these demographics.

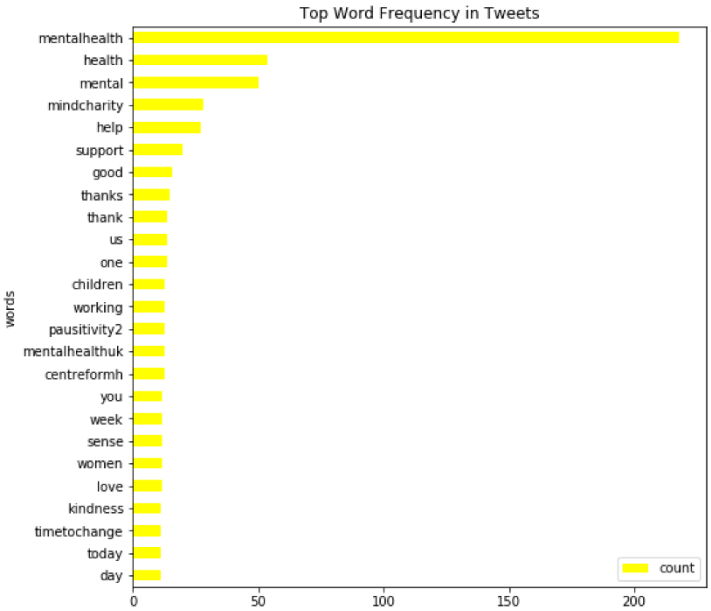

Figure 3

Models

Finally, the tokenized tweets were analyzed using Latent Dirichlet allocation to see what topics were being discussed. The model was created using the LDA function from genism.

In order to find the optimal number of topics, a topic coherence analysis was done. Figure 4 is a graph showing the coherence scores for various number of topics. The graph line spikes to -10 at both 3 and 25 before dropping to lower scores. Graphics were created with both number of topics plus additional ones for visual evaluation. This analysis would probably work better on longer documents or if all hashtags were removed.

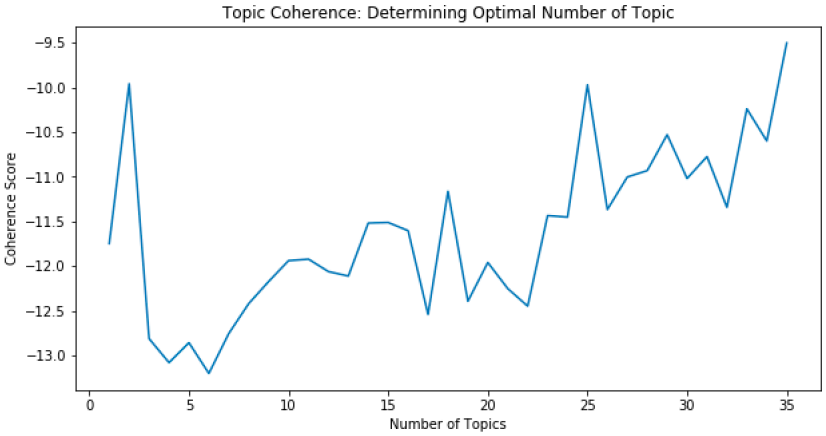

Figure 4

Results

Figure 5 shows the results of using the number of topics to be 2. The first topic seems to be about how help and love are important and it mentions the charity, Mind. The second topic has most of the words from the first, but also thanks, good, and working.

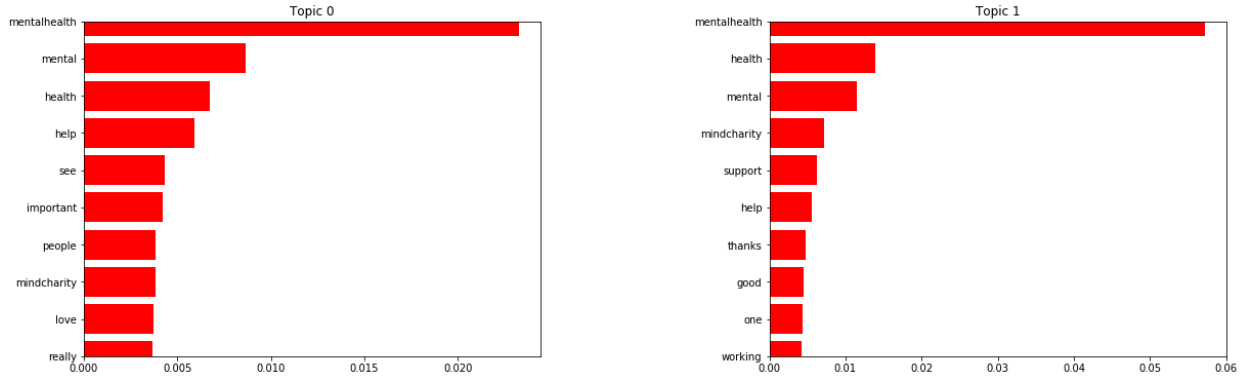

Figure 5

Figure 6 below shows the topic visualized when 8 topics were chosen. Many topics seem to show the multiple hashtags used when discussing mental health. If someone is using Twitter to

search for resources or to connect to people with similar interests or challenges, they could start following these hashtags.

Some of the more interesting topics would be the first, second, third, fifth, and seventh. The first seems to be about women's mental health during menopause. The second appears to be about children and receiving support, thanks, and advice. The third discusses coping with the lockdown during the pandemic and feeling lonely. The fifth topic mentions find, kindness, support, help, and mentions walked steps. This could be related to the user who was raising money for mental health by tracking steps. The seventh topic mentions research, help, sense, and time plus the Mind charity and centreformh, which stands for the Centre for Mental Health. This is another mental health charity in the United Kingdom.

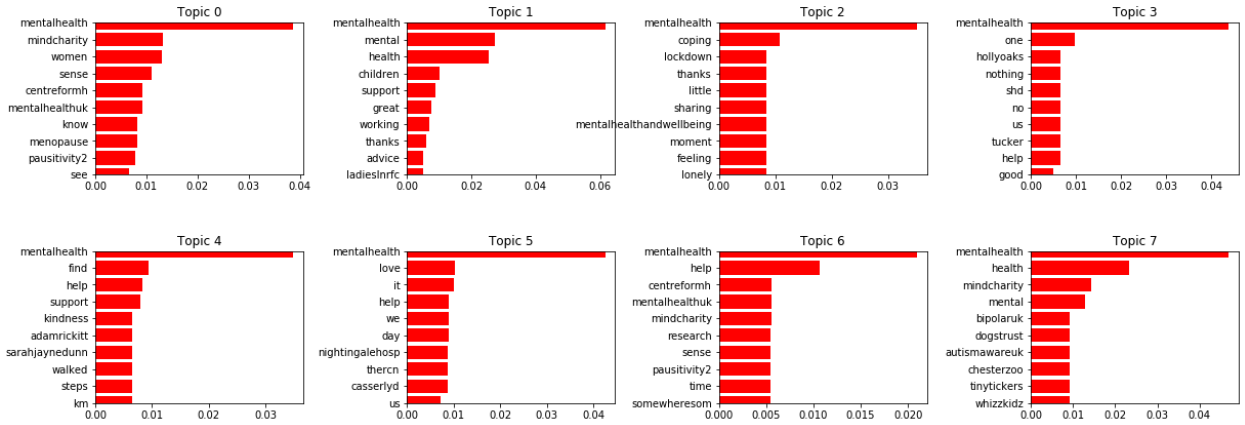

Figure 6

Twenty-five topics did produce some interesting topics such as concerns for employee health and engagement, raising money for a birthday towards mental health, anxiety and depression over COVID19, prejudice, and even going to the coast of wales and how the sound of waves soothes the mind. There were several tweets that discussed the benefits of nature and exercise for mental health.

Conclusion

The increase use of social media and the reduction of stigma on mental health has allowed people who would previously be silenced to have a voice and access to resources. If a person is affected by mental health, social media has the ability to give that person an outlet for connection and creativity. Topic analysis of Tweets reveal some of the options available.

The topic analysis of tweets revealed many resources such as charities that offer help to people with mental illness. These charities also assist in research and work with the government to allocate resources. People who are suffering would not only be able to reach out to them for help, but they could become hopeful things will be better due to the research and advocacy.

Additionally, health professionals could use this analysis to see up-to-date information regarding people's concerns with mental health. Some of the topics mentioned issues specific to women and children while others revealed how some were suffering due to the pandemic. Health professionals could then release information specific to these issues on social media to alleviate those concerns.

Social media has the ability to improve the lives of people and to make them worse. Important issues like mental health can be affected by social medial in both positive and negative ways. By analyzing Tweets for topics, health professionals will have the knowledge to increase the likelihood social media will be used to improve the lives of people affected by mental health.

Resources

National Institute of Mental Health https://www.nimh.nih.gov/health/statistics/index.shtml