Use of Twitter to Analyze Mental Health Topics

Ragheb Abunahla*, Praise Ayorinde*, Rahul Sharma*, Barret Tran*
*Lassonde School of Engineering, York University. Toronto, CA.
Email: raghebabn@gmail.com, praise.ayorinde@yahoo.com, rahul494@my.yorku.ca, tranbarrett@gmail.com

ABSTRACT

Mental health has been studied extensively through different avenues. More recently, social mediums like Twitter have been used to analyze peoples' attitudes, beliefs and opinions about mental health. In this paper, we look for interesting patterns in tweets by performing statistical, sentiment and similarity analysis on users of Twitter that have tweeted about mental health topics. Our results show that the trending mental health topics are on depression, abuse, suicide and stress. Also, most of the tweets on mental health originate from users in the United States of America. Our analysis also depict that more negative tweets on mental health topics are posted on Sunday in comparison with other days of the week. Additionally, we also show that users who tweet about experiences with suicide, are more similar to users who tweet about experiences with depression, in comparison with users who tweet about experiences with abuse. These results could help mental health organizations plan smarter campaigns, aimed at increasing mental health awareness through social media. Also these results could help policy makers in addressing trending mental health issues facing our population today.

CCS CONCEPTS

• Twitter • Psychology • Text Processing

KEYWORDS

mental health, depression, suicide, analysis, twitter

1 INTRODUCTION

Mental health is a very important aspect of our lives as it affects the way we approach our daily tasks [1]. A lot of research has gone into understanding how individuals seek mental healthcare, our attitudes and beliefs towards mental health [2], the stigma related with seeking mental health care [3], and how people talk about mental health issues on social media [4]. As many of these researches seek to understand peoples' opinions or beliefs about mental health, experiments are usually carried out with surveys where participants self-report on their beliefs and attitudes. A big issue with surveys is that participants might provide answers they think the researcher would like to hear, which are not necessarily their own opinions or attitudes [3].

Social media is one unobtrusive way to study a person's opinion, attitude or belief towards mental health [3]. Twitter is a social medium through which individuals post about their opinions, or their experiences, which provides very rich data that could be used to understand people's beliefs, opinions and attitudes. These short messages, called tweets, are usually made of sentences, and have to be at most 280 characters [8]. Twitter users usually interact with each other by retweeting and following other user's feeds (use of twitter). Also, users make use of hashtags to organize their tweets under certain topics [3]. Since tweets are public – unless users mark them as private – they provide very rich unstructured data that could be analyzed and could answer some really interesting questions [3]. For example, some researchers have analyzed Twitter content to tackle problem drinking [6].

In this project, we analyzed tweets concentrated around mental health topics. The questions we pose in this project are 1) the trending themes or topics in mental health 2) the distribution of these mental health themes across countries 3) the words that usually appear with mental health topics 4) the polarity of tweets over days of the week 5) similarity across users who have tweeted about a personal experience with a mental health issue. The answers to these questions would provide some quick statistics about mental health to inform the public on the trending mental health issues on social media. Also, these answers could help mental health organizations and public policy makers in planning campaigns to improve mental health awareness, target users who might need help with a mental health issue, identify how people tweet about mental health issues, and the topics to address at a given time. Lastly, analyzing tweets to answer mental health questions could help researchers extract raw opinions about mental health, and could be used in conjunction with surveys to control the biases that occur with surveys.

The remaining part of this paper are organized as follows: In section 2, we briefly discuss our data domain, processing domain, and the analysis we carried out to answer the questions mentioned above. Then, in Section 3, we provide the general architecture of our solution, and explain the limitations faced with this approach. In Section 4, we provide the results of our analysis on tweets and visualizations to answer each of our questions. In Section 5, we summarize our findings, and provide some ideas for future work.

2 DATA AND DATA ANALYSIS

Twitter provides data that is very rich but unstructured. We used a framework called tweepy, that can be used to access tweets either through streaming or batched reads in Python. To do more rigorous analysis, we decided to ingest and process data by both streaming and batch. We used streaming to answer questions that involved quick analysis to find the trending themes in mental health or interesting words being tweeted with mental health topics, while we used batch processing to answer tougher questions about analysing user patterns or sentiments across days. Therefore, we used a combination of Spark streaming and HDFS for storing and processing large amount of data. The overall architecture of our solution is shown in Figure 2.

In the speed (stream) layer, data from tweepy is coming at a very high velocity, as we get tweets every two seconds. In this layer, we do not store any tweets, but process the tweets in real-time and provide real-time results. This makes the source rate high, as the results serve a live dashboard that are updated in real-time. In tweepy, the configurations can be set to retrieve full tweets (280 characters), this makes it easy to get complete data.

In the batch layer, data is coming in at a high velocity as well, and then stored into csv files. These files are then loaded into and processed by HDFS. The source rate and sink rate are a lower in Hadoop, as it is a distributed file system, and requires processing on the distributed files. The sizes of the data being processed by both layers are large as Twitter has over 100 million users, and more than 340 million tweets are produced per day [8]. Tweepy allows us access to only about 1% of the entire tweets with the free version, which is still a lot of data. We use some key terms shown in Figure 1 to extract tweets related with mental health topics. Therefore, we are only provided with a subset of the 1% available to use by tweepy. Although, it is difficult to know exactly what percentage of this 1% are collected, we collected data both by streaming and batch reads, on three different computers to retrieve large amounts of data. Using this technique, we were able to get large gigabytes of data, which we processed in different ways.

A dissertation done by Zaydman [4] analyzed more than a million tweets and highlighted some key mental health themes that appeared more frequently than other themes in the mental health domain. We used these mental health themes depicted in Figure 1 to query twitter. Using these mental health hashtags, we were able to extract tweets centered around mental health to answer the questions we posed.

#abuse	#eatingdisorders	#nostigma	#presspause
#addiction	#endthestigma	#nostigmas	#mentalhealthmatters
#alzheimers	#IAmStigmaFree	#1SmallAct	#ocd
#anxiety	#mentalhealth	#psychology	#suicideprevention
#bipolar	#pts	#mhchat	#therapy
#bpd	#anxiety	#schizophrenia	#trauma
#Operationalstress	#therapy	#ptsd	#WMHD2015
#mhsm	#endthestigma	#psychology	#worldmentalhealthday
#trauma	#AA	#schizophrenia	#stress
#spsm	#mentalhealthmatters	#stigma	#wellbeing
#alcoholism	#mentalhealthawareness	#stopsuicide	#adhd
#depressed	#mentalillness	#suicide	#bpd
#depression	#MH	#shellshock	

Figure 1: Key themes in Mental health domain used to extract mental health related tweets.

We applied different forms of transformations and cleaning on all the data that we collected. Mostly, MapReduce functions were used to reduce the data to key-value pairs that was used for analysis. Depending on the questions we were answering, we used different filtering functions as well to extract appropriate tweets. In order to gather more meaningful information, we cleaned the tweets by removing stop words made available by the nltk library [10]. We also removed links, non-letter characters, emoticons and so on. In some cases, we excluded retweets and hashtags completely to avoid over representation of words in our analysis. We also transformed tweets into unigrams, bigrams and trigrams when needed to identify interesting phrases tweeted with mental health topics.

In the stream layer, tweets were processed at a very fast rate and processed at least once. In the batch layer, tweets were processed at a medium rate, as Hadoop takes the processing to the files, and are processed exactly once. The analytics we did were sentiment analysis to identify polarity of tweets across days of the week. We also used tf-idf to create user profiles consisting of top keywords across a user's tweets, and then we performed similarity analysis using Jaccard algorithm.

These different analyses answer the question posed by unobtrusively studying individuals' responses or opinions about mental health topics. Through these analysis, we can easily identify trending topics in mental health, interesting words that users are tweeting about as well as analyze the moods of users over days of the week by looking at the sentiment analysis. Additionally, by creating profiles for each user, we can discover how similar people that tweet about mental health issues are.

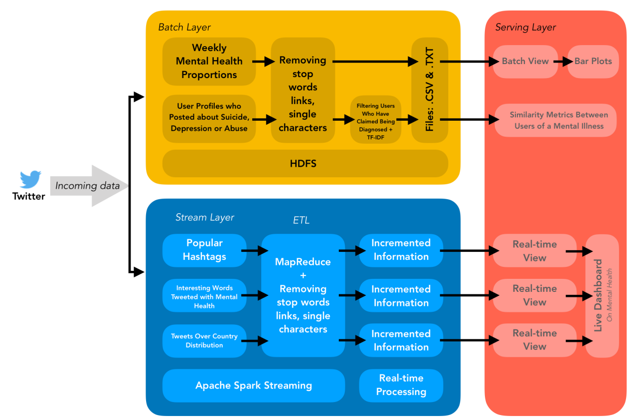

Figure 2: Overall architecture of solution that depicts the different data layers.

3 ARCHITECTURE

The architecture we employed to answer our questions is one that used both speed and batch layers. In each of those layers we ingest and process data in slightly different ways. In the speed layer, tweepy provides streams of tweets that are not stored, but processed real-time and serve a live dashboard. The live dashboard provides information about trending hashtags in mental health, distribution of users across countries that post about these topics, and interesting bigrams and unigrams that were posted with mental health topics. In the batch layer, we collect data in two ways as well. Firstly, we use the key themes discussed earlier to retrieve mental health themed tweets over 7 days and stored in csv files. These csv files were then loaded into HDFS, and processed using MapReduce functions, and Vader Python library to perform sentiment analysis of tweets. Secondly, we used some key phrases, which we will discuss later to extract users who had tweeted experiences about depression, suicide or abuse. We then collected tweets for each of those users and identified keywords among each of the users. We made depression, suicide and abuse unique groups, so that we could compare users among them as well. Using these keywords, we performed similarity analysis. The results of the batch layer were fed to a python script that populated visualizations to depict the information we discovered from the analysis we carried out.

As we were using batch data to process tweets across days of the week, it allowed us to perform more stable analysis on each day, as opposed to using stream data. To use stream data, we would need to stream tweets for a whole week, which could provide a lot of noisy data. We also decided on using stream data to identify key themes in mental health, instead of using batch data. This seemed logical to us, as mental health organizations can track which topics to raise awareness or change attitudes on, at a given period. A limitation with this approach though, is that organizations would not be able see how trends have changed over time, which could also provide very useful information as to how impactful campaigns are, or how people address mental health topics over a period of time. It would be important to point out that the streamed tweets could be stored in text files over a period of days or months, and mined later for more information, but would require a large system like HDFS that can store big datasets.

4 EVALUATION AND RESULTS

4.1 Trending Hashtags in Mental Health

We extracted only the hashtags in mental health related tweets, and discovered the trending topics in mental health during this period. We streamed tweets for close to thirty minutes, and used Spark processing to process the tweets in real-time. The results of the processing fed a live dashboard which we created following the tutorial in [9]. A sample bar graph showing the trending hashtags is shown in Figure 3. This analysis shows us that the most popular topics at this time are suicide, mental health, stress, and depression comes in as well.

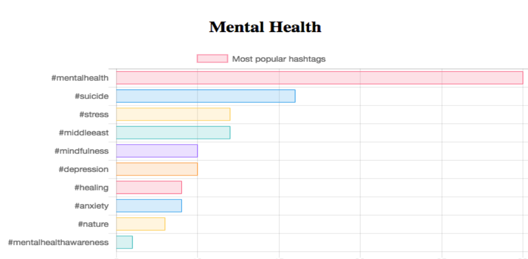

Figure 3: Bar graph showing the trending mental health hashtags

4.2 Distribution of Mental Health Topics Across Countries

Apart from providing tweets, and user information about the users that posted tweets, Twitter also makes available a tweet's location if the user made it public. We used this information to plot a pie chart shown in Figure 4 that depicts the distribution of mental health themed tweets across countries. It is noteworthy that to collect enough tweets that had a location attached to them, we had to stream for more than an hour. Therefore, our chart might not be a total representation of tweets across countries, but it does provide some sample information about the countries that tweet the most about mental health issues.

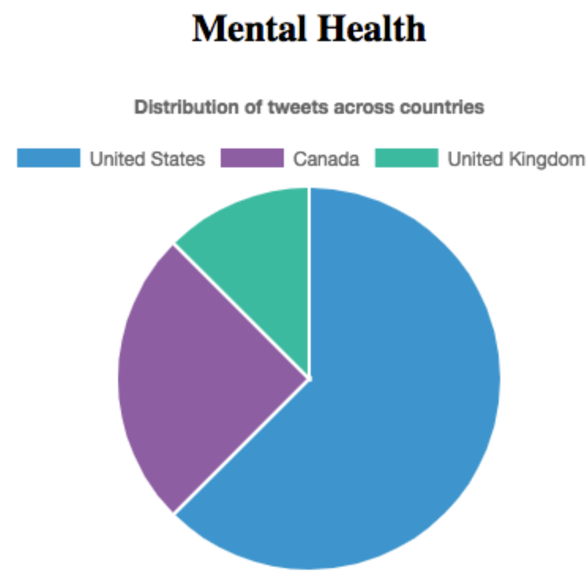

Figure 4: Distribution of mental health topics across countries

4.3 Interesting Words Associated with Mental Health Topics

We used Spark processing as well to process tweets into unigrams, bigrams and trigrams. We then streamed tweets to discover interesting phrases that are usually posted alongside mental health topics. Some interesting ones that were found were **lonely, school, climate change, money, I feel, men** among some other non-interesting unigrams and bigrams. We did not discover any interesting trigrams. Unigrams were gotten after removing stop words, single letter words and words from our key themes, since we already know that those words would be in those tweets. The cleaning steps done with unigrams were not done with trigrams and bigrams, so as to extract meaningful phrases that might have gotten a different meaning if stop words had been removed. As only one interesting bigram was discovered, we show only a sample of the unigrams graph below in Figure 5. We do not show the results of bigrams and trigrams as they were not significant.

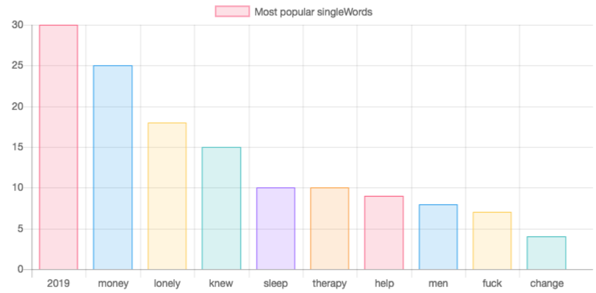

Figure 5: Interesting single words tweeted along side mental health topics

4.4 Sentiment Analysis of Tweets over Days of the week

We posed an interesting question to see peoples' polarity when they tweet about mental health topics across days of the week. To properly analyze the data, we collected tweets using tweepy's query function to get exactly 18,000 mental health related tweets per day. After visually scanning the data, we decided to remove retweets, so as to avoid over representation of certain tweets, which could skew our results. After removing retweets, we had roughly 4,000 tweets per day. In order to normalize the data for each day, we got the least number of tweets we had, and then used that number to retrieve tweets for the remaining days, so that each day has the same number of tweets. Additionally, we removed stop words, links and other non-letter characters. We then loaded the csv file in Hadoop, where we used MapReduce to calculate the sentiment analysis for each day.

Sentiment analysis can be done in multiple ways [4]. The common ways are to use supervised machine learning algorithms, the lexical approach or a hybrid approach [4]. We utilized the Vader Python library that allowed us to calculate the polarity of tweets, using a trained machine learning algorithm. We also used the lexical approach so that we could compare the results from both approaches. In the lexical approach, we retrieved mental health lexicons that had been coded positive, negative or neutral from [7]. A sample of these lexicons are shown in Figure 6. Then, we broke each day's tweets into a list of words, and calculated the polarity of tweets by the frequency of negative or positive or neutral words that they contained of the mental health lexicons. We got really interesting results as depicted in Figure 7, 8 and 9.

We got a lot of negative tweets on Sunday, which might show peoples' moods or attitudes towards the coming work or school day on Monday. We got the least amount of negative tweets on Friday, which we interpreted as people being happy that the week is over. Surprisingly, we got almost a similar amount of negative tweets between Monday and Friday, which might mean that people are too busy to tweet on Monday, but might be able to talk more on the weekends.

Also noteworthy, is that the lexical approach counted almost twice as more tweets as negative, compared to Vader. It is not clear which data was used to train the machine learning algorithm Vader uses, but it is interesting to see that it categorizes most of the tweets as neutral. We get very similar results for positive tweets, the differences seem to show in how each approach deal with negative or neutral sentences.

1	weaksubj	1	abandoned	adj	n	negative
2	weaksubj	1	abandonment	noun	n	negative
3	weaksubj	1	abandon	verb	y	negative
4	strongsubj	1	abase	verb	y	negative
5	strongsubj	1	abasement	anypos	y	negative
6	strongsubj	1	abash	verb	y	negative
7	weaksubj	1	abate	verb	y	negative
8	weaksubj	1	abdicate	verb	y	negative
9	strongsubj	1	aberration	adj	n	negative
10	strongsubj	1	aberration	noun	n	negative
11	strongsubj	1	abhor	anypos	y	negative

Figure 6: Mental Health Lexicons, marked either positive, negative or neutral

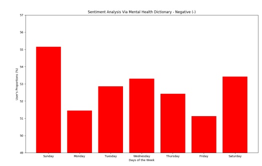

Figure 7a: Negative tweets across days of the week using lexical approach

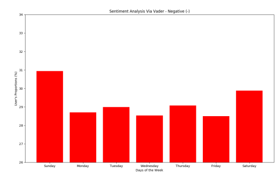

Figure 7b: Negative tweets across days of the week using Vader

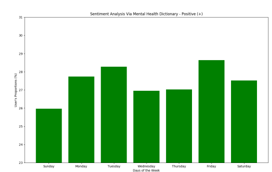

Figure 8a: Positive tweets across days of the week using lexical approach

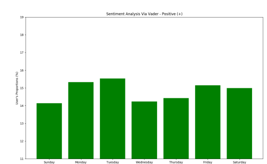

Figure 8b: Positive tweets across days of the week using Vader

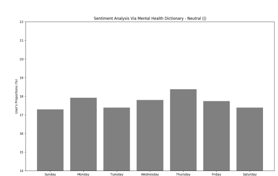

Figure 9a: Neutral tweets across days of the week using lexical approach

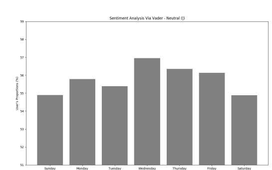

Figure 9b: Neutral tweets across days of the week using Vader

4.5 Similarities Across Users in Twitter

We also wanted to calculate user profile to find any interesting themes across users. Using results from the trending hashtags in mental health where it was discovered that suicide, depression and abuse were trending, we formed user groups under these topics. We made the assumption that anyone who tweeted with *I am…* or *I was…,* had experienced some kind mental health issue before. Using this assumption, we used *'I am suicidal'* to extract 800 users who had posted tweets containing that phrase, which formed the suicide group. We then did the same thing for depression and abuse by searching for *'I have depression'* and *'I was abused'* respectively. This gave us exactly 800 users in each group. Next, we extracted 200 tweets per user, and carried out tf-idf analysis to identify the top 100 words of a user, where a document is the collection of the tweets for that user, and the collection of documents, is the tweets of all the users in that group. We cleaned the tweets by removing stop words, single letter words, retweets, mentions, links, and non-letter characters. We also removed hashtags since we did not want over representation of keywords in the data. Also, to get more meaningful data, we filtered out tweets that did not contain any of our key themes included in Figure 1. This was done to remove tweets of a user that are not related to mental health topics. Instead of comparing users against each other, we decided to compare groups. To do this, we aggregated all the keywords of the users in a group, using MapReduce and chose the top words for that group. Using Jaccard Similarity, we calculated the similarity between groups and got some interesting results. The similarity between the suicide and abuse groups was 46%, while the similarity between abuse and depression groups was 45%. The similarity between the depression and suicide groups was 67%. From our results, it seems that users who tweeted about depression, used similar words with users who tweeted about suicide, in comparison with users that tweeted about abuse. We also created a word cloud that shows the keywords being tweeted about across these groups. The more frequent a word is across the groups, the bigger it would appear in the image shown in Figure 10.

Figure 10: word cloud showing keywords across tweets from users who have tweeted about suicide, depression or abuse

5 CONCLUSION

The results we obtained by performing analytics on Twitter provided some interesting information. The results confirmed our guesses that people would have more time to post on weekends. This information could help organizations know when to start campaigns, because it seems that they would get the most audience during the weekend. It made sense that negative tweets are higher on Sunday, as it feels like general truth that people usually don't feel like going to work or school on Monday. It was interesting to also find the data show those results for the different days of the week.

Also, identifying similarity across users in different mental health topics could help compare and contrast users within those different topics. For example, from the results of our analysis, organizations might want to focus on users who have tweeted about depression, as they were similar to users who had tweeted about suicide.

Additionally, identifying the trending hashtags over time could help public policy makers to address mental health issues when the public is concentrated or talking about it, to increase their reach. Researches could also use twitter to understand peoples' opinions and compare users, which could help them identify key attitudes across different types of people.

In the future, we could choose users from more than three groups, and look at the similarity among these groups to see if there are themes yet to be discovered. For example, if most users that tweet about stress also tweet about work or school, more subgroups could be made under these general mental health groups, to identify problems that might need to be addressed immediately. Another interesting analysis might be to look at the age distribution of the users that post about risky mental health issues like suicide. This analysis would provide useful information to better understand the population tweeting about mental health issues. Extracting the age of users though would require more rigorous cleaning and longer streaming, as users are not required to include their age in their description, but some do. Future analysis could be restricted to only users that included their age, so as to answer more interesting questions such as mental health awareness across age groups.

REFERENCES

[1] Mental Health Government. 2017. What is Mental Health? Retrieved from https://www.mentalhealth.gov/basics/what-is-mental-health
[2] Fahad R. Choudhry, Vasudevan Mani, Long C. Ming and Tahir M. Khan. 2016. Beliefs and perception about mental health issues: a meta-synthesis. Neuropsychiatric disease and treatment, 12. (October 2016), 12 pages. DOI:10.2147/NDT.S111543
[3] Nicola J. Reavley and Pamela D. Pilkington. 2014. Use of Twitter to monitor attitudes toward depression and schizophrenia: an exploratory study. PeerJ. 2014, 2, e647 (October 2014), 15 pages. DOI: 10.7717/peerj.647
[4] Mikhail Zaydman. 2017. Tweeting About Mental Health. Ph. D. Dissertation. Pardee Graduate School. Santa Monica, CA.
[5] Tech Crunch. 2017. Twitter officially expands its character count to 280 starting today. Retrieved from https://techcrunch.com/2017/11/07/twitter-officially-expands-its-character-count-to-280-starting-today/
[6] Joshua H. West, Parley C. Hall, Kyle Prier, Cark L. Hanson, Christophe Giraud-Carrier, Shannon E. Neely, Michael D. Barnes. 2012. Temporal variability of problem drinking on Twitter. Open Journal of Preventive Medicine, 2, 1, (December 2011), 6 pages. DOI: http://dx.doi.org/10.4236/ojpm.2012.21007
[7] Ashwanth Ramji. 2017. Depression Sentiment Analysis using Twitter Data. Retrieved from https://github.com/AshwanthRamji/Depression-Sentiment-Analysis-with-Twitter-Data
[8] Wikipedia. 2018. Twitter. Retrieved from https://en.wikipedia.org/wiki/Twitter
[9] Hanee Medhat. Apache Spark Streaming Tutorial: Identifying Trending Twitter Hashtags. Retrievd from https://www.toptal.com/apache/apache-spark-streaming-twitter
[10] NLTK library. Retrieved from https://pythonspot.com/nltk-stop-words/