

Tracking Suicidal Tendency using Twitter Data and Machine Learning Algorithms

Likitha Daiphule, Bhaskar Reddy, Avinash Savith, Apoorva T V, Kavyashree



Abstract: Social media analytics has a major part in a person's life in this scenario. It is used to obtain the thoughts and opinion, sentiments of People. In this world people are comfortable sharing their thoughts and feelings effectively on social media rather than sharing their happiness or problems to their friends, parents or siblings'. Cerebral health indicators, with depression, Depression and nervousness leads to high risk of people obligating to suicide. Digital knowledge plays a major role to find suicidal tendency of people and to help them out. The study or research about finding the amount of people who have suicidal tendency or not was carried over by many universities where they collected the data from twitter or any health organizations. Twitter data is the most easily available data when compared to Facebook or any other social media site. These observations help us to determine the percentage of people having suicidal tendency or not by many processes which includes data preprocessing, data augmentation, testing and training, and final result representation. We use machine learning concepts. Sentiment Analysis or opinion mining is used. There are many reasons for suicides across the world, using this digital or social data and with the help of machine learning we could also differentiate between the group of people who actually are depressed or people tweeting jokes, songs etc.

Index Terms: Bag Of Words, sentiment analysis, Natural Language Processing.

I. INTRODUCTION

People living with depression, anxiety, loneliness raised danger of suicide contrasted with the all-inclusive community. Teenagers with depression are almost four times more probable than people in the population to attempt suicide. Lifetime suicide hazard among people with depression ranges from 5% to 13%, speaking in this gathering to a major source of mortality. There is a serious need for fresh ways in which people suffering from depression can differentiate suicide hazard. Online life stages have developed as imperative computerized observing devices equipped for encouraging the identification and following of various maladies and general well-being concerns. A developing amount of studies have featured the

possibility and guarantee of utilizing well known internet based life for observing on the web discussions about suicide and for conceivably distinguishing those in danger of suicide. A few examinations have utilized Twitter information to describe suicide-related discussions and to screen suicide hazard [8,7]. An investigation of Twitter tweets establish that discussions regarding suicide were regularly imparted together to content about self-damage and discouragement [9]. Additional examination researched the mental attributes of online life clients in China who had put discussions re suicide on the Weibo micro blogging stage [10].

Study has demonstrated that considering online life movement can yield significant general well-being bits of knowledge about genuine mental disarranges. For example information caught from Facebook was utilized to portray familiarity with anxiety over the United States, while a different investigation established that discussions regarding suicide on Twitter were regularly negative, recommending the nearness of social disgrace. Various examinations have likewise exhibited that people living with depression utilize well known online life at practically identical rates as the overall public. Further, people with psychological instability seem to utilize online networking to share their sickness encounters or look for counsel from others with comparative conditions. A progression of on-going investigations have distinguished novel examples of correspondence on Twitter among clients who self-recognize as having depression and disturbed mind, as reflected by phonetic contrasts, contrasted with control clients changes toward increasingly positive notion following self-revelation of psychological maladjustment via web-based networking media and more prominent utilization of discussion terms about emotional wellness indications contrasted with control clients. Moreover, webbased life, for example, Twitter might be particularly significant for observing suicide hazard among people with Depression as these stages offer extraordinary chances to draw in youthful grown-ups. For example, all things considered, Twitter clients will in general be more youthful contrasted with the general populace. This is exceedingly important in light of the fact that suicide mortality among individuals with Depression is most prominent among more youthful grown-ups, where people 20-34 years old are more than multiple times as likely contrasted with youthful grown-ups from a similar age bunch from the overall public to kick the bucket because of suicide [2].

II. RELATED WORK

This section of the paper consists of a brief summary of the previous work done in the fields of sentiment analysis and twitter-based suicide tendency detection.

Manuscript published on 30 May 2019.

* Correspondence Author (s)

Likitha Daiphule, School of CSIT, REVA University, Bengaluru, India.
Dr Bhaskar Reddy, Associate professor, School of CSIT, REVA University, Bengaluru, India.
Avinash Savith, School of CSIT, REVA University, Bengaluru, India.
Apoorva T V, School of CSIT, REVA University, Bengaluru, India.
Kavyashree, School of CSIT, REVA University, Bengaluru, India.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC-BY-NC-ND license [http://creativecommons.org/licenses/by-nc-nd/4.0/](https://creativecommons.org/licenses/by-nc-nd/4.0/)



A writing survey brought out the danger of suicide and self-hurt is likewise exceptionally conspicuous among youthful people measured at very great danger of creating neurosis. Too, there is expanding acknowledgment about people with depression and disturbed mind are overwhelming clients of online networking and show one of a kind correspondence designs on these advanced stages. This mix features the possibility to use online networking for identifying suicide hazard and advising suicide counteractive action endeavours in this in danger quiet gathering. Notwithstanding, less is thought about whether individuals with depression and disturbed mind talk about suicide via web-based networking media.

A vital initial move toward creating systems to utilize internet based life for supporting the recognition of suicide chance among people with depression and disturbed mind is to more readily see how this objective populace discusses suicide on mainstream web based life. Hence, in this inspective examination, our points were to (1) research the recurrence of online correspondences about suicide among Twitter clients who self-distinguish as having Depression contrasted and a control gathering of run of the mill Twitter clients; (2) portray the planning of tweets about suicide among Twitter clients who self-recognize as having Depression contrasted and a controlled gathering of average Twitter clients; (3) decide if dialog about other basic psychological well-being side effects, including sorrow or nervousness, is prescient of online exchanges about suicide. We speculated that Twitter clients who self-recognize as having depression and disturbed mind would be essentially bound to post tweets comprising suicide terms when contrasted with Twitter clients from the all-inclusive community, in this way mirroring the raised danger of suicide saw among people with depression and disturbed mind in genuine settings.

III. PROPOSED APPROACH

A. Data Collection:

All information broke down in this examination was openly accessible and was gathered from the Twitter internet based life stage. Twitter is a famous micro blogging stage where clients post short statuses called "tweets" that contain a limit of 140 characters—since 2018, this has expanded to a limit of 280 characters for each tweet. It is evaluated that the in excess of 330 million dynamic Twitter clients post more than 500 million tweets for each day. This features a massive wellspring of spontaneous information with energizing potential to ponder different parts of human conduct and screen wellbeing conditions including dysfunctional behaviour. In particular, we chose this internet based life stage for this investigation since it has recently been utilized for leading examination on a few distinctive psychological well-being conditions, including discouragement, bipolar confusion, and posttraumatic stress issue. Imperatively, information caught from Twitter has been utilized in research portraying on the web discourses and frames of mind about Depression, investigating semantic markers of Depression, and supporting endeavors to recognize people with depression and disturbed mind. In conclusion, Twitter clients will in general be more youthful contrasted with the general populace, which is particularly imperative given the raised suicide chance among youthful people with depression [2]. In

this way, given Twitter can accomplish across the board reach, and that we can develop existing related work, we confirmed that Twitter would be a perfect stage to possibly fill in as a compelling computerized instrument for observing danger of suicide among individuals with depression and disturbed mind.

As well-known web based life stages have risen as an imperative wellspring of client created content that can yield important experiences for general wellbeing research, the moral contemplations with investigating and scattering this information have gotten more prominent consideration. While there remains an absence of accord over prescribed procedures for utilizing Twitter information in scholastic research, there is progressing discourse encompassing concerns identified with security, privacy, and educated assent. To limit potential dangers, we guaranteed that all information gathered in our investigation was accessible in the open space. Be that as it may, extra moral contemplations are justified, particularly with regards to socially slandering wellbeing conditions, for example, dysfunctional behaviour. For instance, scattering client created content gathered on Twitter could conceivably put a person in danger of mischief since delicate wellbeing data, for example, psychological maladjustment conclusion or indications, could be made recognizable in manners that were not planned by the first client who posted the substance on the web. In this way, to additionally ensure the personality of the Twitter clients whose information we analysed in this investigation, we expelled all usernames and recognizable subtleties from the substance that they posted on the web. Finally, we don't report particular tweets that could be utilized to distinguish the first Twitter client who posted the substance on the web, as this is a critical worry that has been talked about broadly in on-going writing on the morals of utilizing Twitter information for research.

B. Twitter Users and Characteristics:

We distinguished a comfort test of 250 Twitter clients who unequivocally self-recognized as having a Depression range issue in their profile or in a tweet. For instance, the clients' profiles could specify "individual living with depression and disturbed mind" or "I have Depression determination," while a tweet could make reference to "this is the manner by which I deal with my Depression" or "I was simply determined to have Depression." We displayed our information gathering strategies on earlier investigations that have utilized the Twitter stage for producing an accommodation test of clients with freely accessible records. We looked Twitter utilizing the accompanying terms: Depression, schizoaffective, schizotypal, and psychosis. We at that point affirmed oneself announced Depression finding by having one scientist create this underlying rundown of Twitter clients and after that a second specialist check the subtleties for each Twitter client on the rundown to guarantee right distinguishing proof of clients with a self-revealed Depression range issue. To make an all-inclusive community correlation gathering, we utilized the GET statuses/test highlight from the Twitter Developer Platform to gather an arbitrary example of all openly accessible tweets.

At that point, two research collaborators physically analyzed these tweets to affirm that the tweet had a place with a genuine individual, was from an ordinary client, and was in English. This procedure was proposed to guarantee that Twitter clients incorporated into the controlled bunch were genuine Twitter clients. To limit the danger of choosing any bot or spam clients, both research partners must be in understanding of a Twitter client on every one of these three standards. We avoided any Twitter clients where there was contradiction. Our objective was to make a gathering of clients that was agent of average Twitter clients. We perceived a case of 250 control customers. We decided sex for the example of Twitter clients in light of the fact that various investigations have recognized a connection among sexual orientation and suicide hazard too, psychological wellness indications, for example, discouragement and tension have a known relationship with sex. Furthermore, among people with Depression, death because of suicide is greater among men than in ladies [2]. We utilized a procedure for coding each Twitter client's sexual orientation as male, female, or obscure/inadequate information. Two scientists freely utilized these codes starting with each Twitter client's username, trailed by profile name, profile depiction, profile photograph, and after that tweets. The two analysts at that point audited their last sexual orientation codes for each Twitter client to guarantee steadiness and to determine any contradictions.

We likewise separated a few qualities for the Twitter clients incorporated into this examination. This included gathering metadata from the Twitter clients' records, including all out amount of tweets, tweets every day (i.e., absolute tweets isolated by days dynamic), tweets in most recent 200 days, amount of companions, amount of adherents, top choices every day, and amount of days the record has remained dynamic. We likewise estimated each Twitter client's effect, which is determined as a devotee's to-companions proportion where the client's amount of adherents is isolated by their amount of companions. It thus represents as just an estimate of Twitter's effect as well as influence since a higher ratio implies that such a user has a large amount of individuals who follow their account but follow tweets from few other certain users.

In our last example, incorporated into the investigations detailed here, we had an aggregate of 203 Twitter clients who self-recognized as needing depression and disturbed mind and 173 control clients. The last amount of clients changed on the grounds that a few records wound up difficult to reach (i.e., private, erased, prohibited, or disabled) or were inert (i.e., no posts amid the 200-day ponder period) at the season of information accumulation.

C. Tweets with Suicide Words:

We recovered all tweets displayed amid the 200-day time frame from the Twitter clients incorporated into this examination. Inside this accumulation of clients' tweets, we distinguished just tweets that contained the catchphrases suicide or self-destructive. Earlier examinations have appeared there is an assortment of terms utilized via web-based networking media that might be demonstrative of suicide hazard. The term suicide is every now and again contained in suicide-related discussions. In this manner, we purposefully constrained our pursuit to these two terms to

improvise the conviction that the dialog content caught in this investigation was unequivocally alluding to suicide. We likewise considered this critical in light of the fact that online talks about suicide have been associated with real suicide chance. For instance, an investigation from Japan demonstrated that announcements explicitly referencing the term suicide on Twitter were fundamentally connected with self-destructive ideation and conduct. Notwithstanding hunting down suicide-related terms, we additionally chosen watchwords for other emotional well-being side effects that are realized hazard factors for suicide. These incorporate the accompanying terms: wretchedness, discouraged, tension, and on edge.

IV. METHODOLOGY

We utilize a two-class Naive Bayes classifier with n-gram highlights for our assessment arrangement, bag of words with logistic regression and TF-IDF algorithm with KNN algorithm. Preparing information is named by a semi-regulated methodology that utilizes emojis as boisterous names. Our preparation technique is appropriate as portrayed for practically any language that utilizes spaces as word separators.

A. Tokenization and Features:

We utilize a Naive Bayes classifier on n-gram highlights to arrange opinion in tweets. To acquire n-gram highlights, we first need to tokenize the content info. Tweets represent an issue for standard tokenizes intended for formal and ordinary content. Tweet writings have one of kind qualities like missing whitespaces, unpredictable spelling and extraordinary tokens, for example, hash tags ("#tag"), notices of twitter clients ("@username") and retweet markers ("RT"). In this way the information is tokenized utilizing an all-inclusive rendition of O'Connor's normal expression based tweet tokenizer [O'Connor et al., 2010]. It utilizes principles to characterize strings that structure a token, for example emojis, urls and strings of accentuation, and parts the remainder of the strings at whitespaces.

B. Classifier:

We utilized the Naive Bayes classifier from the NLTK normal language handling toolkit for preparing. We changed the classifier with an expansion that enabled us to persistently retrain the classifier by refreshing the watched include recurrence check with the new preparing information. We classify tweets utilizing similar highlights we utilized for the preparation information. Class probabilities for the polarities are determined utilizing logarithmic probabilities. The classifier at that point relegates the mark of the most plausible class to the tweet. After the preparation, the classifier is streamlined by evacuating some uninformative highlights. All n-grams with equivalent consider are evacuated they don't impact the characterization.

C. Bag of Words:

This method for removing highlights from content for use in displaying, for example, with AI calculations. ... A sack of-words is a portrayal of content that depicts the event of words inside a report. It includes two things: A vocabulary of known words. It is known as a "sack" of words, on the grounds that any data about the request or structure of words in the archive is disposed of. The model is just worried about whether realized words happen in the record, not where in the report.

A typical component extraction strategy for sentences and records is the sack of-words approach (BOW). In this methodology, we take a gander at the histogram of the words inside the content, for example considering each word considers an element. In this methodology, we utilize the tokenized words for every perception and discover the recurrence of every token. Consider few tweets: "I like mango", "I love India". We try to consider every sentence as a different document and create a list of all the words from all of the two documents which do not include the punctuation. We get, "I" and "love," "like," "mango" and "India." The next step is vector development. Vectors transform text that the learning algorithm of the machine can use. We take the first document—"I like mango" and we check the frequency of words. The final vector obtained that is the output of this will be given the input to logistic regression algorithm.

D. TF-IDF Algorithm:

An issue with scoring word recurrence is that profoundly visit words begin to command in the record (for example bigger score), yet may not contain as much "educational substance" to the model as rarer however maybe space explicit words. One methodology is to rescale the recurrence of words by how regularly they show up in all reports, with the goal that the scores for incessant words like "the" that are additionally visit over all records are punished.

RESULTS

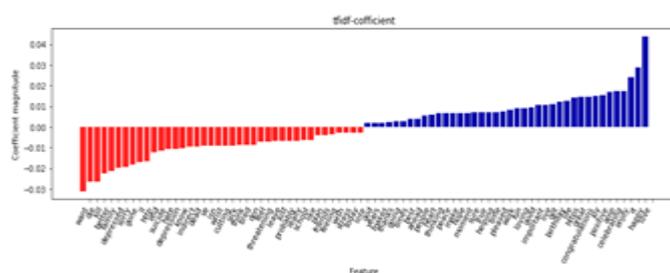


Fig 1: Result of TF-IDF Algorithm.

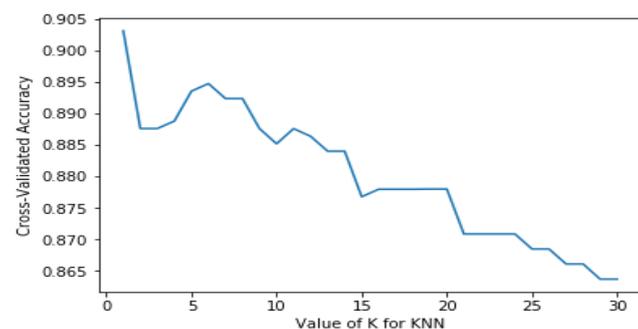


Fig 2: Results for KNN Algorithm.

V. CONCLUSION AND FUTURE ENHANCEMENT:

1. Amongst the numerous fields of investigation, there is one field where people have ruled the machines more than any – the capacity to examine feeling, or notion examination.
2. The eventual fate of this information investigation field is tremendous. This task examinations the feelings of the client as well as figures different outcomes like the client with greatest companions/devotees, top tweets and so on.
3. Further this project can be extended by creating an app which will be useful for people suffering from suicidal tendency.
4. Next also they can create a website where this predicts the users who have tweeted continuously which shows they have suicidal tendency and this website can provide them some help like giving information about the help centres and other.

REFERENCES

1. Girardi P, Kaplan K, Amador XF, Harkavy-Friedman J, Harrow M, Pompili M, et al. Suicide risk in Depression: Learning from the past to change the future. *Ann Gen Psychiatry* 2007 Mar 16;6:10.
2. Gerhard T, Olfson M, Crystal S, Huang C, Stroup TS. Premature mortality among adults with Depression in the United States. *JAMA Psychiatry* 2015 Dec;72(12):1172-1181.
3. Hor K, Taylor M. Suicide and Depression: A systematic review of rates and risk factors. *J Psychopharmacol* 2010 Nov;24(4 Suppl):81-90.
4. Madoff LC, Freifeld CC, Brownstein JS. Digital disease detection: Harnessing the Web for public health surveillance. *N Engl J Med* 2009 May 21;360(21):2153-2155, 2157.
5. Fisher S, Robinson J, Hetrick S, Bailey E, Rodrigues M, Cox G, et al. Social media and suicide prevention: A systematic review. *Early Interv Psychiatry* 2015 Feb 19.
6. Scourfield J, Colombo G, Burnap P, Hodorog A, Amery R. Multi-class machine classification of suicide-related communication on Twitter. *Online Soc Netw Media* 2017 Aug;2:32-44.
7. Matsubayashi T, Mori K, Sawada Y, Ueda M. Tweeting celebrity suicides: Users' reaction to prominent suicide deaths on Twitter and subsequent increases in actual suicides. *Soc Sci Med* 2017 Dec;189:158-166.
8. Barnes MD, Jashinsky J, Hanson CL, Burton SH, West J, Giraud-Carrier C, et al. Tracking suicide risk factors through Twitter in the US. *Crisis* 2014;35(1):51-59.
9. Krauss MJ, Cavazos-Rehg PA, Connolly S, Sowles SJ, Bharadwaj M, Rosas C, et al. An Analysis of Depression, Self-Harm, and Suicidal Ideation Content on Tumblr. *Crisis* 2016 Jul 22:1-9.
10. Zhu T, Guan L, Kwok CL, Cheng Q, Yip PS. Suicide communication on social media and its psychological mechanisms: An examination of Chinese microblog users. *Int J Environ Res Public Health* 2015 Sep 11;12(9):11506-11527.