

# Detection of Suicide Ideation on Twitter

Anonymous

## 1. INTRODUCTION

In today's world, social media is one of the platforms for personal expression of thoughts, ideas and to show our individuality. We post nearly all nitty-gritty details of our everyday life, from talking about what we eat to our opinions on political matters. Especially for young adults, social media is a form of catharsis. From these posts, we can assume human emotions and what a person goes through mentally. People who go through an unfortunate episode are likely to convey their feeling through social media indirectly[4]. In comparison with users who spent more time on social media like Twitter, Instagram, Facebook, and other platforms were shown to have a considerably higher rate of reported depression than those who spent less time, according to recent studies[2]. Recently a 16 year old girl committed suicide over Instagram poll results on users voting that she should die [3]. There are several triggers and indications online which can theoretically determine if a person is likely to be suicidal or not. Our project focuses on identifying individuals who are suicidal based on the content that they post. We wish to contribute to the ongoing research on detection of suicide ideation and measure the performance of four machine classifiers in their accuracy of correctly distinguishing suicide related and non suicide related posts.

Natural language processing helps us perform text classification. This helps classifying texts into various groups based on its content. There are certain things that you can do with NLP, like, sentiment analysis, topic detection and language detection. Sentiment analysis helps us understand if the topic that is being talked about is in a positive or a negative context [5]. We will be performing sentiment analysis on twitter data by users to understand and classify references to suicide. We will be comparing performance of machine classifiers on this data and provide the best possible recommendation for identifying suicide related posts.

## 2. DATA SETS

- **Twitter Live Data & Detect Depression In Twitter Posts:**[1]:

Twitter Live Data is one of the sources selected to analyze and apply mining techniques. This data contains all the tweets tweeted by any of the Twitter user.

The second source we have selected is specific to Suicidal/Depression data in the Twitter Data set.

1. Format of data (Twitter Live Data): JSON
2. Format of data (Detect Depression In Twitter Posts): CSV

The following are few attributes collected from the live Twitter data:

1. **id:** 'id' gives a numeric id for any individual Twitter user.
2. **created\_at:** 'created\_at' is a date-time based attribute, gives information about when the tweet was tweeted by a particular user.
3. **text:** 'text' is the actual tweet tweeted by the user.
4. **user:** 'user' is the user\_id of the person who started the particular tweet.
5. **source:** 'source' gives the utility where the tweet is posted.
6. **in\_reply\_to\_status\_id:** 'in\_reply\_to\_status\_id' is a integer type attribute which gives an integer value which indicates whether the tweet is replied to or not.
7. **friends\_count:** 'friends\_count' is a integer type attribute which gives the friends count of a particular user.
8. **country:** 'country' is a string type attribute which gives the country from where the tweet originated from.
9. **coordinates:** 'coordinates' is a integer-list type attribute which gives the coordinates of the country attribute.
10. **retweet\_count:** 'retweet\_count' is a integer type field which gives the count of retweets a users has done.
11. **favorite\_count:** 'favorite\_count' is a integer type attribute which tells us how many times the tweet has been liked.
12. **emoji:** 'emoji' field depicts any emoji's present in the tweet.

Most of the data present in both of our data sets are common except fields like 'emoji' but the major difference is that the second source only focuses on data pertaining to users having depressive or suicidal thoughts.

### 3. LIST OF DELIVERABLES

1. Data gathering
2. Data merging from live and existing datasets
3. Data cleaning
4. Data pre-processing
5. Sentimental analysis
6. Classification using :
  - (a) Naive Bayes
  - (b) Decision tree
  - (c) Random forest
  - (d) Support vector machine
7. Comparison of machine classifiers using measure scores like precision recall f-measure and accuracy

8. Visualization of the performance of the aforementioned machine classifiers

### 4. REFERENCES

- [1] Detect depression in twitter posts. <https://github.com/peijoy/DetectDepressionInTwitterPosts>. (Accessed on 09/23/2019).
- [2] Does social media cause depression? | child mind institute. <https://childmind.org/article/is-social-media-use-causing-depression/>. (Accessed on 09/20/2019).
- [3] Instagram urges users to report suicidal posts after 16-year-old kills herself over death or life poll results. <https://www.newsweek.com/instagram-urges-users-report-harmful-posts-after-16-year-old-kills-herself-over-death-or-life-poll-results>. (Accessed on 09/23/2019).
- [4] These 5 social media habits are linked with depression | live science. <https://www.livescience.com/62718-social-media-habits-depression.html>. (Accessed on 09/20/2019).
- [5] What is text classification? | monkeylearn. <https://monkeylearn.com/what-is-text-classification/>. (Accessed on 09/23/2019).