



American International University-Bangladesh (AIUB)

Suicidal Tendency Prediction Using Machine Learning

MD. RAZIB MOLLAH (20-42153-1)

EFFAT ARA TANZIN (20-42000-1)

JAKIR HASAN (18-38795-3)

MD. IFAZUL ISLAM (18-36180-1)

Supervised by

MD. FARUK ABDULLAH AL SOHAN

*A Thesis submitted for the degree of Bachelor of Science (BSc)
in Computer Science and Engineering (CSE) at
American International University Bangladesh in
May 2024*

Faculty of Science and Technology (FST)



Abstract

Suicidal propensity prediction using social media data is a novel factor influencing suicidal behavior. suicide is the 11th main reason of death worldwide. [11]. Perceptions of suicide are influenced by social, environmental, and personal factors. Recently, suicide has become a "social" threat. [12]. Twitter is rapidly being explored as a source for displaying and perceiving a person's state of mind. The purpose of this paper is to test whether tweets about suicide can be classified based on their content. This is achieved by iteratively applying machine learning algorithms and engaging programmers. Data from Twitter was analyzed for various suicidal words and comments using public APIs. [13]. Collect suicide-related tweets with suicide-related keywords and save them in a CSV file as a record. A recursive neural network is then used to classify these tweets to determine if they are likely suicides. The proposed model identifies suicidal users so that they can be approached and referred to mental health care. This will benefit them and potentially save lives.

People prefer modern technology and use online social media as a means of communication when they are considering suicide. [14]. The two fundamental risks that lead to successful suicide are suicidal attempts and suicidal thoughts, and primary identification and detection are seen as effective strategies to prevent both. This study illustrates several approaches for understanding suicidal thoughts in online user material, with a focus on Twitter data from the last two years as a form of early detection using sentiment analysis and supervised learning techniques.

Our result got 80% accuracy after implementing a labeling tweeter dataset. The results of this study may aid in the development of machine learning algorithms that may anticipate suicidal ideation and conduct. Eighty percent of suicidal situations were accurately identified by the algorithms. To create a rapid screening tool for use in primary care or community settings, our findings can be applied.



Declaration by author

This thesis is composed of our original work and contains no material previously published or written by another person except where due reference has been made in the text. We have clearly stated the contribution of others to our thesis as a whole, including statistical assistance, survey design, data analysis, significant technical procedures, professional editorial advice, financial and any other original research work used or reported in our thesis. The content of our thesis is the result of work we have carried out since the commencement of the Thesis.

We acknowledge that the copyright of all material contained in my thesis resides with the copyright holder(s) of that material. Where appropriate we have obtained copyright permission from the copyright holder to reproduce material in this thesis and have sought permission from co-authors for any jointly authored works included in the thesis.

MD. RAZIB MOLLAH
20-42153-1

EFFAT ARA TANZIN
20-42000-1

JAKIR HASAN
18-38795-3

MD. IFAZUL ISLAM
18-36180-1



Approval

This thesis titled Suicidal tendency prediction using machine learning has been submitted to the following respected member of the board of examiners of the Faculty of Science and Technology impartial fulfillment of the requirements for the degree of Bachelor of Science in Computer Science and Engineering on (May 2024) and has been accepted satisfactory.

MD. FARUK ABDULLAH AL SOHAN
Lecturer , Computer Science
Faculty of Science and Technology
American International University-
Bangladesh

DR. MD. ABDULLAH - AL - JUBAIR

Assistant Professor & Director
Department of Computer Science
American International University-
Bangladesh

PROF.DR. DIP NANDI

Professor & Associate Dean
Faculty of Science and Technology
American International University-
Bangladesh



Contributions by authors to the thesis

List the significant and substantial inputs made by different authors to this research, work, and writing represented and/or reported in the thesis. These could include significant contributions to the conception and design of the project; non-routine technical work; analysis and interpretation of research data; drafting significant parts of the work or critically revising it so as to contribute to the interpretation.

	MD. RAZIB MOLLAH	EFFAT ARA TANZIN	JAKIR HASAN	MD. IFAZUL ISLAM	Contribution (%)
	20-42153-1	20-42000-1	18-38795-3	18-36180-1	
Conceptualization	25%	25%	25%	25%	100%
Data curation	25%	25%	25%	25%	100%
Formal analysis	25%	25%	25%	25%	100%
Investigation	25%	25%	25%	25%	100%
Methodology	25%	25%	25%	25%	100%
Implementation	70%	0%	30%	0%	100%
Validation	25%	25%	25%	25%	100%
Theoretical derivations	25%	25%	25%	25%	100%
Preparation of Figures	25%	25%	25%	25%	100%
Writing – original draft	25%	25%	25%	25%	100%
Writing – review & editing	25%	25%	25%	25%	100%

If your task breakdown requires further clarification, do so here. Do not exceed a single page.



Acknowledgments

Acknowledgments recognize those who have been instrumental in the completion of the project. Acknowledgments should include any professional editorial advice received including the name of the editor and a brief description of the service rendered.

Keywords

Machine learning, Multinomial Naïve Bayes, Neural Network



Table Of Contents

Contents

SUICIDAL TENDENCY PREDICTION	1
USING MACHINE LEARNING.....	1
ABSTRACT	2
DECLARATION BY AUTHOR	3
CONTRIBUTIONS BY AUTHORS TO THE THESIS	5
ACKNOWLEDGMENTS	6
KEYWORDS.....	6
TABLE OF CONTENTS.....	7
LIST OF ABBREVIATIONS AND SYMBOLS	8
1 INTRODUCTION.....	9
1.1 GENERAL AIM.....	10
1.2 RESEARCH OBJECTIVE	10
1.3 RESEARCH QUESTION	10
1.4 OVERVIEW OF THE COMING CHAPTER	11
1.5 OUR THESIS TOPIC.....	11
2 LITERATURE REVIEW	12
3 METHODOLOGY.....	13
3.1 INTRODUCTION.....	13
3.2 PROCEDURES	13
4 RESULTS OR FINDINGS.....	15
5 DISCUSSION	16
6 CONCLUSION.....	17
7 CODE	18
REFERENCES.....	20



List of Abbreviations and Symbols

API	Application Programming Interface
ML	Machine Learning
MNB	Multinomial Naïve Bayes



Chapter 1

1 Introduction

700,000 individuals commit suicide each year, which is a significant societal issue. People are taking their own lives all across the world, particularly those who are in their 20s and older. [15]. The World Health Organization reports that he made an attempt at suicide in his 30s. Suicide Between the ages of 10 and 34, it is the second largest cause of death. Suicidal ideation, which is what it is often known as, includes the concept of taking one's own life. [16] Suicide may affect people of all ages. Ideas arise from a variety of diverse sources, such as surprise, rage, guilt, melancholy, and worry. Despite inadequate therapy, long-term depression can result in suicide. Most individuals who have suicidal thoughts don't really attempt suicide. Medical experts and medications can be used to treat suicidal thoughts. But the majority of individual's experience suicidal thoughts. Because it is stigmatized, avoid seeking therapy. instead, a lot of people decide to post about their plans to die on social media. [17] As it is mental disease diagnosis and treatment, early warning symptoms, and hazards, the element could be the best strategy for preventing suicide. Social media, which is mostly used by young people, In latest years it has evolved within a powerful aperture into the mental health and comfort of its users. It provides unidentified presence in various online groups and creates a forum for open discussion on socially taboo topics. In general, fifty percent of suicides and over twenty percent of those who attempt suicide write a suicide note. [18] Therefore, any written indication of suicidality is considered to be concerning, and the writer should be questioned about their own thoughts. Social media materials such as blog posts, conference post, tweets and other online notes are often still recorded and strong maintained. The study of social media and associated venues for discussing mental health has grown in the field of computational syntax. It provides a useful analysis environment for generating groundbreaking innovations that could revolutionize suicide risk reduction and suicide detection. [19] It can provide a useful entry point for intervention. investigated the Reddit Suicide Watch users' posting habits when they follow celebrity self-destruction news. He presented a strategy that would effectively stop suicides involving prominent figures We explored Reddit's social media shift from mental health discussion to suicidal ideation. To determine the characteristic signs of this transition, he developed a analytical method based on propensity score identical.

As we all know suicide is the second largest cause issue, [20] and we tried to prevent this occurs by using machine learning. This project can save lives. How? This project will trace people's suicidal thoughts throughout worldwide social media. By that, we can detect that particular person can save lives. If this process continues, we can ensure that this suicide issue will not be the world's 2nd largest cause of death.

We downloaded Tweet dataset records from a GitHub repository. We analyzed data from both routine tweets and messages by persons who had committed suicide. This information was kept in Excel. The label "1" in the dataset denotes suicidal behavior, while the label "0" denotes neutral behavior. There are 9119 samples in total, 4000 for suicidal samples and 5119 for neutral samples,



respectively. We obtain the dataset from Excel, perform any necessary preprocessing, train a **Multinomial naive Bayes classifier**, and then put the model to use via serving.

1.1 General aims

- As far as we know, the social media text data Conduct a comprehensive review of suicide Idea recognition and its methods and applications from a machine learning perspective.
- Present both classic content and discuss them analysis of machine learning and new technologies, Twitter data, Farewell letter and online social content.
- enumerate existing and less explored tasks, Talk about their limits. Combine existing documents and provide perspective for future research The direction of this field.

Our main and primary concern is preventing suicide by predicting suicide posts through social media and also locating papers using tools for prediction and computing PPV estimates for suicidal behavior.

1.2 Research Objectives

- Evaluating optimal suicide rate prediction models across continents.
- Visualize differences in key traits that influence suicide rates.
- Estimating the best suicide rate prediction model

1.3 Overview of forthcoming chapters:

- Chapter 1: contains an introduction to the work. This section defines research goals.
- Chapter 2: Add a literature review, and examples of previous studies relevant to this research.
- Chapter 3: Presents the overall model and proposed architecture.]
- Chapter 4: Describes experiments, results, and research evaluations.
- Chapter 5: Results, future scope, and limitations are described in this section of this chapter.
- Chapter 6: This chapter is conclusion part.
- Chapter 7: This is the last chapter of this project, here we will provide the code.



1.5 Our Thesis Topic

“Early Suicidal Tendency Predication by Using Machine Learning”



Chapter 2

2 Literature review

It is important to detect suicide tendency precisely, so that to take precautions and create strategies to deal with the situation and minimize the adverse effects. So far, some previous research studies on related fields have already conducted. Social networks have captured attention to the research group, Especially when trying to figure out their format, user wiring and interactions between them. People tend to use Twitter to express their affections and talk around their regular lives. Linguistic Survey and global Count Version 2007 rates specific words and sentiments. Cognitive and basic English language expressed by individuals orally and in writing. Can be recycled to analyze aim in impassioned posts.[1]. They addressed the depression debate by working with language markers and using a bulletin board system to gather information from depressed individuals and pre-existing non-depressive Internet forums. said people with depression were more likely to report medical problems. and who are overseen by regulations that ensure the protection of mortality study participants. General Guidelines Suicide prediction on AI-based social media and medical ethics platforms is mostly outside the healthcare organization. Almost completely separated, the pots always remain independent. Forecast style is considered an exclusive trade secret [2]. A new and different challenge is social networks such as Twitter, A place where users post feedback and conclusion about commodity in actual time. Some early Findings and Recent Analysis of Twitter depressed Data. Mining Opinion has two major research field.[3]. Both document-level and document-level Methods of Classification Sentences are commonly placed on definitive remark of opinion. D.Hotman found that there is room for significant improvement in Suicide forecast and determent entire a combination of novel analytic approach and accessory, such as the use of machine learning algorithms and data science They found that the majority (83%) of those who died by suicide contacted health services in the year prior to death, and 45% contacted those services in the month prior to death. and emphasized the need for better validation and follow-up [4]. Suicide is world`s maximum critical problems.. year by year total More and more people are committing suicide. expect that to happen. about it, 800 people died from various causes. Thousands of people die each year an inference model was created using suicide attempt data. By providing nine properties, researchers were able to predict linear relationships. 99% estimated accuracy. Podlogar et al (2018).[5]. In this task, we compute the affinity between the tweet dataset and the training data. Create a job using the Machine A learning algorithm. In this background, target on pessimistic, awful thoughts and facts, and consider suicide. Evaluate the entire system and present Twitter results with predictive data

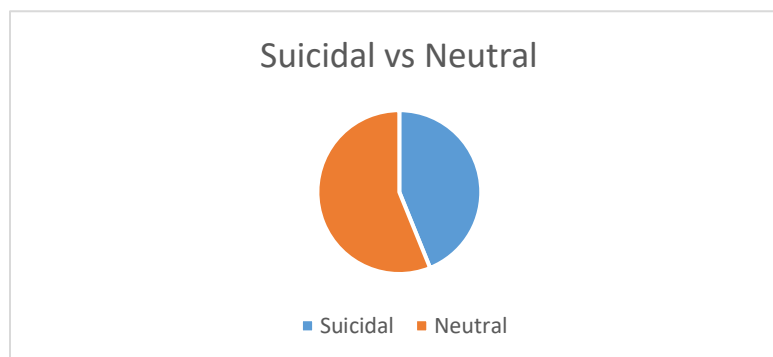
Chapter 3

3 Methodology

3.1 Introduction

After determination of doing analysis on Suicidal Tendency Prediction we planned to collect data from twitter posts. But instead of that we found a previous dataset from a Git Repository which was pre-processed with 9224 rows and 2 columns.

So, any pre-processing procedures was not necessary for this dataset. We started by checking the dataset either balanced or imbalanced and as a result we found that our dataset was imbalanced.



3.2 Procedures

3.2.1 Value Counts

In the beginning we started with value counts, as a result we get the data length: 8800

3.2.2 Missing Value Check

After getting the value counts result, we proceed to check missing values and there was no missing values.



```
In [51]: df.isnull().sum()
```

```
Out[51]: label    0  
        data     0  
        dtype: int64
```

3.2.3 Duplicate Value Check and Drop

We tried to find out the duplicate values & We dropped all the duplicate values from the dataset to make the result more accurate.

3.2.4 Compare Value

In this step we again count the dataset value after checking the missing values and duplicate values and compare with the previous value we got in the beginning.

So,

the previous value = 8800

the current value = 8800

Difference = $8800 - 8800 = 0$

3.2.5 Splitting Data

We separated the label and data into x & y arrays to split our data into train and test sets.

3.2.6 Data preprocessing

We converted the test sets in numerical data to use on the training data so that we can scale the training data and also learn the scaling parameters of that data and store the data in a multi-dimensional array.

3.2.7 Using Algorithm & Methods

We use Multinomial NB algorithm to analyzing text input and solving problems with numerous classes and supervised learning applications model to get statistics calculated across all of the models. And we used a python score method to measuring the accuracy of the model against the training data.

Finally, we forecast a certain future output based on trends found through historical data using the model prediction method.



Chapter 4

4 Results or findings

We fetched Tweet records published on a GitHub repository. We used data from posts by people who posted suicide and regular tweets. This data was stored in Excel. The label '1' in the dataset represents suicidal and the label '0' represents neutral. The total number of samples is 9119, distributed as 4000 and 5119 for suicidal and neutral samples respectively.

respectively. We fetch the dataset from excel, preprocess it, and We trained our model **multinomial navies bays classifier**, finally the model is used for serving.

The final accuracy of the model is 0.80 and the loss is 0.20.

	Precision	Recall	F1-Score
0	0.95	0.66	0.78
1	0.72	0.96	0.82
Accuracy	0.80		
Macro avg.	0.83	0.81	0.80
Weighted avg.	0.84	0.80	0.80

Fig: Accuracy Level



Chapter 5

Discussion

Our experiment has limitations due to incomplete data and biased annotations. One of the most important problems in today's research, which mostly uses supervised learning approaches, is the lack of data. Normally, they need to be manually annotated. To facilitate more study, there aren't enough labeled data, though. A further problem is a bias in annotation brought about by hand labeling using pre-established annotation criteria. In other instances, the annotation may cause labels to be biased, producing false evidence that supports the authors' decision to commit suicide. Identifying suicidal ideation in its early stages is an important and productive approach to early suicide. The majority of research on this subject has been conducted by a psychologist using analysis of the data. nonetheless, the bulk of computer scientists' research is done through feature engineering-based machine learning representation learning. Machine learning techniques have the potential to do this. Early detection of suicidal ideation and later early suicide prevention. In this study, we sought to test and rate various ML models to detect whether a user's tweets contained signs that would indicate suicide intent. Determine which model performs best. The major goal of the study is to identify the most reliable model for identifying suicidal thoughts on Twitter so that we can do so with a user with a high level of accuracy.

Suicidal inclinations are more effectively-recognized when regular linguistic patterns in social media material are identified. Using different techniques for various NLP methods frequently supports it. Desmet et al. [6] To identify suicidal thoughts, developed a suicide note analysis approach employing binary (SVM) classifiers. Huang et al. [21]. A psychological vocabulary based on a Chinese feeling dictionary was produced. He used the SVM technique to determine a classification for a system that would be used to detect suicidal thoughts in real time on Chinese Weibo. Braithwaite et al [7]. ML technique is effective in distinguishing between those at risk of suicide and those who are not. Sueki et al. [8] said that linguistic framing is crucial for spotting suicide signs in text after studying the suicidal intent of young Japanese Twitter users. For instance, the statement "want to suicide" is more usually associated with a lifelong desire to commit suicide than the phrase "want to die" demonstrates that automatic ML classifiers on TF-IDF characteristics may be used to automatically discriminate between different levels of worry among postings connected to suicide. Wood et al. [9] A total of 125 Twitter users were discovered by Wood et al., and their tweets previous to the data accessible before their suicide attempt were followed. Okhapkina et al. [10], But We download the dataset from Excel, perform preliminary processing on it, train a multinomial Naive Bayes classifier, and then use the model to serve. And it shows a promising result which is 80% accuracy.



Chapter 6

Conclusion

This research goes beyond predicting suicide posts, in this new age social media world, text data is generated every second. This data can be used in many ways. Through deep sentiment analysis, you can solve health-related problems or detect hate and post-toxicity to stop bullying.

In this research, we created and assessed a suicidal tendency prediction ML-based system. Our experimental Findings indicated that the Multinomial NB achieving model performed better with 80% accuracy in employing textual cues, as opposed to the latter's 91.5% accuracy, in diagnosing suicidal ideation. According to our analysis, suicidal postings score worse on attention, mind-thinking, perception, sociality, and cognitive processes than non-suicidal posts, scoring higher on authenticity, anxiety, mentality, depression, negativity, and sentimental melancholy. People today are more sensitive than ever before, which has a significant negative impact on their families, friends, and those around them. The suggested technique may assist in identifying people who need medical care and lowering suicide rates by recognizing suicidal intent in user posts.

Although the results of our research indicate that applied classification algorithms work quite well, the metrics' absolute value suggests that this is a difficult problem that merits additional investigation. We may attempt to obtain a bigger dataset with content connected to suicidal ideation as well as a brand-new dataset with comparable themes in our future study. We might look at the relationship between suicidal thoughts and factors like the family environment, the weather, etc. To better illustrate and evaluate our suggested hybrid approach, both sets of data will be acquired from various social media platforms. The dataset's efficacy will also be utilized to explore deeper learning classifiers in the future, including C-LSTM, RNN, and their combination models, as well as to evaluate parameter optimization.

We believe that our study will contribute to the development of a very successful suicide detection and reporting system that can be used in social media networks as a useful intervention tool between at-risk individuals and mental health providers.

Chapter 7

Code

```
jupyter suicide-ML Last Changelog: 12/26/2022 (autosaved)
File Edit View Insert Cell Kernel Windows Help
In [45]: # Import
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.feature_extraction.text import CountVecToizer

In [46]: # Import dataset
df = pd.read_csv('suicidal_data.csv')

In [47]: df

Out[47]:
   label      data
0      1  my life is meaningless i just want to end my l...
1      1  nothing i wanna die to myself daily for a fo...
2      1  work stress i really hate my myself daily...
3      1  i did something on the 3rd of december i overth...
4      1  i feel like no one cares i just want to die m...
...
9219  0  supermees nothing but can't wait start waiting...
9220  0  i m off too bad i gotta wake up hehe early...
9221  0  just picked up some cats from the market to as...
9222  0  i wish we would get it the same time at the c...
9223  0  need to sleep but cant

9224 rows x 2 columns

In [48]: df['data'].value_counts()

Out[48]:
1 dont want to die
165
1 hate revealing i hate them so much when i see them i just want to die i feel my blood boiling and my skin burning and my eyes
watering my heart drops and i feel like fire is consuming my entire body why are they so evil and cruel i hate them
26
1 want to die
19
1 clean me
18
1 hate revealing i hate them so much when i see them i just want to die i feel my blood boiling and my skin burning and my eyes
watering my heart drops and i feel like fire is consuming my entire body why are they so evil and cruel i hate them
15

...
my world has come crashing down and i cant see any other escape i also posted this in so and am sorry for double posting but i
feel like this thread was actually more suitable here's my story just been dumped by my boyfriend of 4 years for another girl
with who has been friends with since before we met he slept with her before we met and apparently has been talking to her witho
ut me knowing all throughout our relationship he hasn't physically cheated per se that i know of but i feel he has been emotionally
cheating on me for a while it hurts so fucking much that it actually like physical pain at the moment and i cant stop crying has
also refusing to talk to me we all live in the same place and i just heartbroken and horrified by the idea that i have to see
them together in the future and keep him then iwe also found out that my dad also had an affair with someone when i was 18 ye
rs old around 26 and when i had anorexia i was close to dying from anorexia and cheated on my man they've been married for 35
years and my dad has always been my rock who i've always looked up to their relationship was one i idolized and always wanted i
cant get my head around the idea he cheated on my man my sister died when she was 8 and i was 4 and i was and they managed to get through
that my man deserves so much better and i am so angry that he could ever do that to her at such a vulnerable time in their li
ve this year i've also lost my best friends after an incident happened at a hen party which made me reevaluate if they were actual
ly friends they were acting bitchy and awful and so i ended up telling the bride that i dont think i could be her bridesmaid a
nmore and she ended the friendship and all the other friends sided with her i lost my friends my partner and my dad in such
a short space of time i feel like i have nothing but my job but even though i love my job its not enough to keep me going i fee
l like no one wants me in their lives i cant stop crying and i just want to die i dont see what the point to anything is anymor
e my man is my only friend i have 1
1 indefectible anastasia down my preferred method right now cant even begin to describe the amount of shit on my backside suicid
e attempts later here i am about to make a fourth one and i amusing a throwaway to post because assured of potentially surviv
ing this outbreak in december i had a terrible injury from falling on my own body with my full then weight of 97 kg a follow up
to three nasty concussions from falling down stairs and getting kicked in the head a short nervous system chronic pain a fucked
up spine limb weakness tingling total loss of function in arms and legs i cant get anywhere alone in fear that i will lose control
of my legs in the middle of a crosswalk or something massive fucking medical bills and i amnt poor a alone and i have a ver
y illegal job ambling down at a small accounting company to help my parents pay my down and medical bills and regular lif
e expenses i recently had my glasses stolen so that another large sum of money and visual impairment to deal with until i save
up for a new pair llyrs of terrifying skin picking issues left me looking like a meth addict as mothers of small children in p
ublic places often like to tell me i cant eat anyone i haven't been able to sleep properly since i was an infant but not anore
xia king down from exhaustion i feel like more of a liability to everyone than a friend or a family member or a girlfriend or a da
ughter i cant tell me anything i cant tell my friends i cant tell my girlfriend if any of the secrets i keep i feel like i
to the public in going to be feeling certain homelssness discrimination from teachers like the other kids like i feel like i
school i will get fired cause already dragging a massive black bag of trauma from back then i dont want to do this any more anore
xia if my body still planning on killing me somehow i will do it by my own fucking self i cant even reach out to a therapist tried to
re of then two distict me so seriously when i started off by mentioning my skin as a legit issue i have mental and one just told
me i was a typical whitey child who should be happy to live with and not make a big deal out of it i was a big deal out of it i
ed in any way being actively suicidal for years now isnt smiling of dying thaty anorexia to everyone i know to ensure at least a mod
icum of a normal life of being abused by peers by my own body by my brain by myself fucking doing one of it hell maybe i want to
be myself off of everyone fucking back at least my parents wont have to spend so much money on a defective useless child who d
ont make the good grades she used to make and who cant even be healthy and happy like literally all the others cheers everyone
ill anorexia
1 when will it end and since i almost 26 now i've wanted to kill myself for so many times i've stopped counting its always the same i
feel there is something wrong i have a strong anxiety attack and then i start to think of committing a suicide i start to look
for the best way how to do it then i my calling pills start work i cry myself to sleep and the next day i want to visit a doctor
and tell her i need help sometimes i do that sometimes i dont she sends me to a psychiatrist and then to another psychiatrist
at and they all will prescribe some pills that help me to sleep but never to solve any of my obsessive thoughts of committing a
suicide then i live normally again for few days weeks if luckily even months then its back it will always come back right now
y the frequency of these attacks is increasing few months ago i moved and i live alone which i dont really handle well i tr
y to overcome my breakup that has happened almost 8 months ago after 3 years together until i still love him i have a new job w
hich is stressful on its own high school teacher here and i haven't still finished my uni programme and the time flies so fast
i about me doing anything productive so now amnting here writing this stuff so that i dont browse the suicide advice website
s and am asking myself what the hell is wrong with me anyone very attractive intelligent kind person who has the whole futur
e ahead of her but i adjust unable to appreciate it because there is something very wrong inside of my head people who so
uld never expect it as i come off as a confident popular extrovert and it makes it even more difficult to actually speak about
this with anyone in my life my parents know about it but we are not in contact and my friends have their own lives and
problems this is a vicious circle if it ever ends it will be not a happyed another worthless person who wishes nothing else th
an to go to sleep and never wake up one day it is going to happen hopefully it will be done by the hand of fate godfather and not
by my hand the odds right now are on the latter though
1
1 employed 26 i cant stop fantasizing about an attempt i have been battling serious depression for 3 years i have been getting
help for over 1 on fluoxetine as well as cannabiz i self have currently 2 months cleanie sitting in my apartment that my parent
s pay for with my long time friend getting stolen and watching tv all i can think about is cutting myself or killing myself i s
ven saw my counsellor today that my parents pay for high i will never be able to function i frequently fantasize i want to
ng i dont want to die i just desperately want more immediate help i just want to stop suffering in ontario because our mental he
alth support is terrible i failed and had to drop out of school i lost my friend counselling personality all at once i
and my transit pass i currently trying to fill out job applications and get my resume out ok i am i supposed to survive been
has stalized before and i was fucking horrible i was in the hospital for 72 hours until i will go into the hospital for 72 ho
urs and theyll show me full of fucking benzocaine counsellor wants me to toss the blades and eat obviously but neither have happ
ened yetty friends and family nor supportive yet am i useful as so tired i just want to live anyone it doesnt seem
worth it to me or even seemly alive to stop others from suffering
1
1 utternerv why why and to think that it is still on the show you kuterlall penn you've been the bright star in ho
1
Name: data, Length: 8880, dtype: Int64

In [49]: df.drop_duplicates(inplace = True)

In [50]: df['data'].value_counts()

Out[50]:
1 want to die

my life is meaningless i just want to end my life so badly my life is completely empty and i dont want to have to create meanin
```



```
poite way when i'll i stop feeling jealous of tragic characters like gonem pile for the swift and they were able to bring to t  
heir lives  
i  
my twitter always lagss damn  
i  
once tried the freezer method but didn't work for me since then i have a backup hd stored outside my home good luck anyway  
i  
why am i working today tuesdays are my days off  
i  
i  
i will end it on my 16 birthday i dont have any motivation to live! anself confident i have lots of friends! asintelligent and  
i am not ugly but i cant do this anymore our lives have no goal and we are so tiny compared to the whole galaxy life is too bor  
ing without weed or alcohol i will kill myself because i have absolutely no patience and motivation to do this for still around  
70 years and whats the difference to the world when i kill myself i can do it so quick and afterwards i cant even care about th  
e other people that are still in this world also! anself really curious about what happens after our lifes if we become just no  
thing its okay because i wont even realize it and if theres some kind of afterlife it wont be as bad as this boring world i gue  
ss! anself an egotistic arrogant lazy and depressed asshole that is really curious  
i  
what is the easiest way to kill yourself what is the easiest way for a 15 yo to kill himself without much pain  
i  
injection i attempted suicide a while back and made a cocktail of medicine and injected it in my arm it didn't work i got bette  
r temporarily now i need to go is there a way i can make it work this time  
i  
what is the fucking point what is the fucking point to be happy how why it doesnt last pain is the default and it always comes b  
ackive tried years ive tried pushed from one pill to the next so i can be numb enough to enjoy the rat raceguttered myself in fro  
nt of every therapist until they feign empathy before taking my crumpled bills since they never accept my insurances clothes a  
re made in sweatshops my food was tortured horrifically! lawmakers are corrupt soldiers work for profit law is an i vilution a  
nony is fake everything is a scam everyone lies everyone steals everyone hurts no one is innocent and it just keeps going and g  
oing and goingwere all just dumb pack animals who cling to routine and jump from one cycle to the next to keep ourselves busy a  
nd blindso why continue whats the point because it gets better why would anyone want to get better in such a crooked world  
i  
kutmerrrrr why why and to think that it is still on the show ugh kutnerkal penn you ve been the bright star in ho  
i  
Name: data, Length: 8880, dtype: int64
```

```
In [51]: df.isnull().sum()
```

```
Out[51]: label    0  
data      0  
dtype: int64
```

Seperating the x and y from dataset

```
In [52]: x = df.data.values  
y = df.label.values
```

```
In [53]: x
```

```
Out[53]: array(['no life is meaningless i just want to end my life so badly my life is completely empty and i dont want to have to creat  
e meaning in it creating meaning is pain how long will i hold back the urge to run my car head first into the next person conin  
g the opposite way when i'll i stop feeling jealous of tragic characters like gonem pile for the swift and they were able to br  
ing to their lives',  
"muttering i wanna die to myself daily for a few months now i feel worthless shes my soulmate i cant live in this horrib  
le world without her i am so lonely i wish i could just turn off the part of my brain that feels",  
"work class i really feel like my only purpose in life is to make a higher man money parents forcing me through collage  
and i have too much on my plate i owe a lot of money i know this is the easy way out but i am really tired all of these issues  
are on top of dealing with criminals in america as well i want to rest",  
...., "still 3 more days until my internet gets uncapped",  
"my throat is still really sore i was meant to be going on a 5 day camp from friday but not so sure now",  
"kutmerrrrr why why and to think that it is still on the show ugh kutnerkal penn you ve been the bright star in ho",  
dtype=object])
```

```
In [54]: y
```

```
Out[54]: array([1, 1, 1, ..., 0, 0, 0], dtype=int64)
```

```
In [55]: x_train,x_test,y_train,y_test = train_test_split(x,y,test_size = 0.2)
```

Data preprocessing

```
In [56]: cv = CountVectorizer()  
x_train = cv.fit_transform(x_train)
```

```
In [57]: x_train.toarray()
```

```
Out[57]: array([[0, 0, 0, ..., 0, 0, 0],  
[0, 0, 0, ..., 0, 0, 0],  
[0, 0, 0, ..., 0, 0, 0],  
....,  
[0, 0, 0, ..., 0, 0, 0],  
[0, 0, 0, ..., 0, 0, 0],  
[0, 0, 0, ..., 0, 0, 0]], dtype=int64)
```

```
In [58]: from sklearn.naive_bayes import MultinomialNB
```

```
In [59]: model = MultinomialNB()  
model.fit(x_train,y_train)
```

```
Out[59]: MultinomialNB()
```

```
In [60]: x_test = cv.transform(x_test)
```

```
In [61]: x_test.toarray()
```

```
Out[61]: array([[0, 0, 0, ..., 0, 0, 0],  
[0, 0, 0, ..., 0, 0, 0],  
[0, 0, 0, ..., 0, 0, 0],  
....,  
[0, 0, 0, ..., 0, 0, 0],  
[0, 0, 0, ..., 0, 0, 0],  
[0, 0, 0, ..., 0, 0, 0]], dtype=int64)
```

```
In [62]: model.score(x_test,y_test)
```

```
Out[62]: 0.7864848517319784
```

Matix

```
In [63]: from sklearn.metrics import confusion_matrix  
from sklearn.metrics import classification_report
```

```
In [64]: pred = model.predict(x_test)
```

```
In [68]: print(classification_report(y_test,pred))
```

	precision	recall	f1-score	support
0	0.95	0.65	0.77	970
1	0.69	0.96	0.80	791
accuracy			0.79	1761
macro avg	0.82	0.80	0.79	1761
weighted avg	0.83	0.79	0.78	1761

```
In [ ] :
```



References

1. Zhao, D. & Rosson, M.B., (n.d), Retrieved from http://research.ihost.com/cscw08-socialnetworkinginorgs/papers/zhao_cs_cw08_workshop.pdf
2. Marks, M. (2019). Artificial intelligence-based suicide prediction.
3. Wilson T., Wiebe J. and Hoffmann P., “Recognizing Contextual Polarity in Phrase- Level Sentiment Analysis”, In the Advanced Research and Development Activity (ARDA).
4. D’Hotman, D., & Loh, E. (2020). AI enabled suicide prediction tools: a qualitative narrative review. *BMJ Health & Care Informatics*, 27.
5. Podlogar, M. C., Gai, A. R., Schneider, M., Hagan, C. R., & Joiner, T. E. (2018). Advancing the prediction and prevention of murder-suicide. *Journal of Aggression, Conflict and Peace Research*.
6. Desmet, B.; Hoste, V. Emotion detection in suicide notes. *Expert Syst. Appl.* **2013**, *40*, 6351–6358. [[Google Scholar](#)] [[CrossRef](#)]
7. Braithwaite, S.R.; Giraud-Carrier, C.; West, J.; Barnes, M.D.; Hanson, C.L. Validating machine learning algorithms for Twitter data against established measures of suicidality. *JMIR Ment. Health* **2016**, *3*, e21. [[Google Scholar](#)] [[CrossRef](#)]
8. Sueki, H. The association of suicide-related Twitter use with suicidal behaviour: A cross-sectional study of young internet users in Japan. *J. Affect. Disord.* **2015**, *170*, 155–160. [[Google Scholar](#)] [[CrossRef](#)]
9. Wood, A.; Shiffman, J.; Leary, R.; Coppersmith, G. Language signals preceding suicide attempts. In Proceedings of the CHI 2016 Computing and Mental Health Workshop, San Jose, CA, USA, 7–12 May 2016. [[Google Scholar](#)]
10. Okhapkina, E.; Okhapkin, V.; Kazarin, O. Adaptation of information retrieval methods for identifying of destructive informational influence in social networks. In Proceedings of the 2017 IEEE 31st International Conference on Advanced Information Networking and Applications Workshops (WAINA), Taipei, Taiwan, 27–29 March 2017; pp. 87–92. [[Google Scholar](#)]
11. “Suicide.” *Suicide*, 17 June 2021, www.who.int/news-room/fact-sheets/detail/suicide.
12. Desmet, B.; Hoste, V. Emotion detection in suicide notes. *Expert Syst. Appl.* **2013**, *40*, 6351–6358. [[Google Scholar](#)] [[CrossRef](#)]
13. Sivakumar, C. *et al.* (2022) “An improvised method for anomaly detection in social media using Deep Learning,” *2022 International Conference on Electronics and Renewable Systems (ICEARS)* [Preprint]. Available at: <https://doi.org/10.1109/icears53579.2022.9751851>.
14. O’Dea, Bridianne, et al. “Detecting Suicidality on Twitter.” *Internet Interventions*, vol. 2, no. 2, Elsevier BV, May 2015, pp. 183–88. *Crossref*, <https://doi.org/10.1016/j.invent.2015.03.005>.
15. Sedgwick, Rosemary, et al. “Social Media, Internet Use and Suicide Attempts in Adolescents.” *Current Opinion in Psychiatry*, vol. 32, no. 6, Ovid Technologies (Wolters Kluwer Health), Nov. 2019, pp. 534–41. *Crossref*, <https://doi.org/10.1097/yco.0000000000000547>.

15. “Suicide.” *Suicide*, 17 June 2021, www.who.int/news-room/fact-sheets/detail/suicide. Desmet, B.; Hoste, V. Emotion detection in suicide notes. *Expert Syst. Appl.* **2013**, *40*, 6351–6358. [Google Scholar] [CrossRef]
16. “Facts About Suicide | Suicide | CDC.” *Facts About Suicide | Suicide | CDC*, 24 Oct. 2022, www.cdc.gov/suicide/facts/index.html.
17. “UpToDate.” *UpToDate*, www.uptodate.com/contents/suicidal-ideation-and-behavior-in-adults. Accessed 23 Jan. 2023.
18. “UpToDate.” *UpToDate*, www.uptodate.com/contents/suicidal-ideation-and-behavior-in-adults. Accessed 23 Jan. 2023.
19. Wongkoblap, Akkapon, et al. “Researching Mental Health Disorders in the Era of Social Media: Systematic Review.” *Journal of Medical Internet Research*, vol. 19, no. 6, JMIR Publications Inc., June 2017, p. e228. Crossref, <https://doi.org/10.2196/jmir.7215>.
20. “Suicide 2nd Highest Cause of Death in People Aged 15-29.” *The Daily Star*, 14 June 2018, www.thedailystar.net/health/suicide-2nd-highest-cause-death-people-aged-15-29-1591021. Du, J.; Zhang, Y.; Luo, J.; Jia, Y.; Wei, Q.; Tao, C.; Xu, H. Extracting psychiatric stressors for suicide from social media using deep learning. *BMC Med. Inform. Decis. Mak.* **2018**, *18*, 43. [Google Scholar] [CrossRef] [PubMed][Green Version]
21. Yao, H.; Rosenthal, R.W.F. Detection of Suicidality among Opioid Users on Reddit: A Machine Learning Based Approach. *J. Med. Internet Res.* **2019**. [Google Scholar] [CrossRef]
22. Hermann, K.M.; Kocisky, T.; Grefenstette, E.; Espeholt, L.; Kay, W.; Suleyman, M.; Blunsom, P. Teaching machines to read and comprehend. In Proceedings of the Advances in Neural Information Processing Systems, Montreal, QC, Canada, 7–12 December 2015; pp. 1693–1701. [Google Scholar]
23. Orabi, A.H.; Buddhitha, P.; Orabi, M.H.; Inkpen, D. Deep learning for depression detection of twitter users. In Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic, New Orleans, LA, USA, 5 June 2018; pp. 88–97. [Google Scholar]
24. Bhardwaj, T. et al. (2022) “A review on suicidal ideation detection based on machine learning and Deep Learning Techniques,” *2022 IEEE World AI IoT Congress (AIIoT)* [Preprint]. Available at: <https://doi.org/10.1109/aiiot54504.2022.9817373>.
25. Chadha, A. and Kaushik, B. (2019) “A survey on prediction of Suicidal Ideation using machine and Ensemble Learning,” *The Computer Journal*, 64(11), pp. 1617–1632. Available at: <https://doi.org/10.1093/comjnl/bxz120>.
26. Ramachandran, A. et al. (2020) “Performance evaluation of different machine learning techniques using Twitter data for identification of suicidal intent,” *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)* [Preprint]. Available at: <https://doi.org/10.1109/icesc48915.2020.9155747>.