Journal of Harbin Engineering University
ISSN: 1006-7043

Vol 45 No. 11
November 2024

AI/Ml Intervention Towards Detection and Prevention of Suicidality

[1]Rajanikanta Sahu, [2]Sumant Sekhar Mohanty, [3]Shubhashree Pattnayak
[1]Computer Science and Engineering
Gandhi Institute of Excellent Technocrats Bhubaneswar, India rajanikantsahu84@gmail.com
[2]Computer Science and Engineering
Gandhi Institute of Excellent Technocrats Bhubaneswar, India sumantmohanty72@gmail.com
[3]Computer Science and Engineering
Gandhi Institute of Excellent Technocrats Bhubaneswar, India shubhashree111994@gmail.com

Abstract: People's attitudes, emotions, and activities are the searchable archives on social media (SM), thus providing an excellent opportunity to capture the behavioural attributes. The research on the possibility of leveraging Artificial intelligence-based models on social media is in its infancy. Our work focused on investigating the possibility of automatically detecting suicide- related posts on social media. The research started with the first objective of collecting a large dataset from two online platforms, Twitter and Reddit, to prepare the machine learning frameworks. After collecting the relevant data, the research problem is divided into two parts. One is to differentiate between suicidal and non-suicidal content. After data collection, human annotation was performed as per the proposed annotation scheme. There has been a detailed analysis of the methodology based upon the proposed advanced feature engineering mechanism, extracts and identifies the most relevant features, and then delivered to machine learning algorithms in order to expand accuracy.

Keywords: - Suicidality, LIWC, API, Human annotation, Feature engineering, F1 score

INTRODUCTION

Detection of suicidality among people is one of the core concerns because of the unavailability of tests and social stigma related to mental illness. As people move towards social networking sites to express their feelings freely, it acts as a firehouse of emotions and feelings. A good amount of research is going on to analyse the feasibility and power of social media that could help probable suicidal sufferers get out of considering punishing move. Mental health is a broader term that includes both illness and wellness of an individual, but people usually focus on the illness part. Many people experience mental health concerns at times, but the concern is translated into the illness when the signs reflect the frequent stress or affect the ability of an individual. Mental illness is defined as as all the disorders identified by the change in mood, thinking, or behaviour associated with distress orimpaired functioning of the brain (American Psychiatric Association, 2013). Mental illness is treated as a distinct illness despite having a strong relation in developing and treating various chronic diseases like cancer, diabetes, obesity, etc. (Turner & Kelly, 2000). Suicide is one of the mental illnesses and a leading cause of death worldwide, consuming many lives. Suicidological research uses the term "Suicidality" to describe suicide in a larger context, encompassing completed suicide, attempted suicide, any idea of suicide, or communicating about suicide (O'Dea et al., 2015; Perepletchikova, 2020). When a person gets thoughts to end his/her life, it is referred to as suicidal ideation, while as, suicidal behaviour includes self-harming with the intent to end the life. (O'Dea et al., 2015).

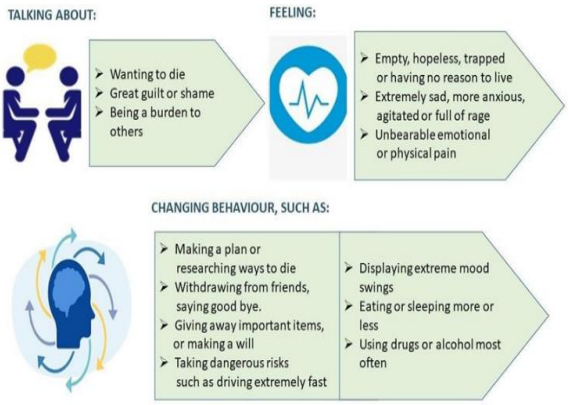


Figure 1. Warning Signs of Suicide

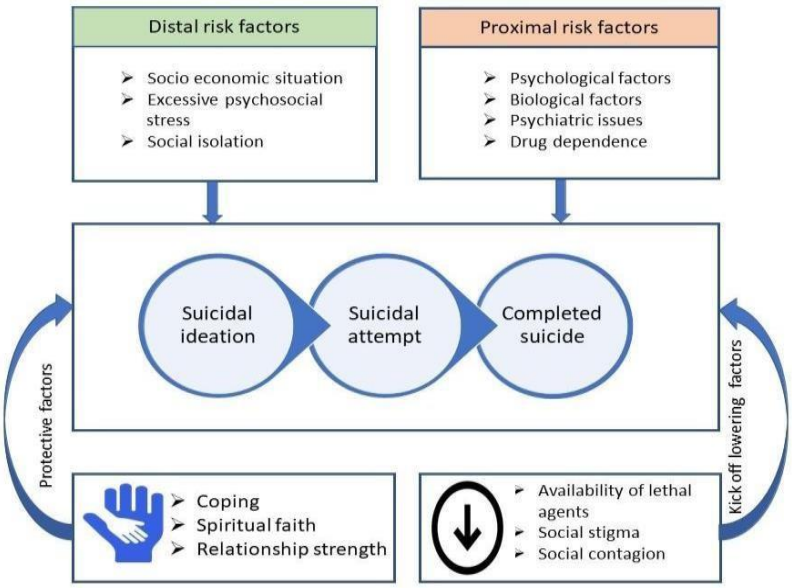


Figure 2. Process of Suicide and the Allied Risk Factors

The Word "suicide" should not be taken lightly, it can be a cry for help fromsomebody, and if its signs are detected early, many lives could be saved. Suicide is preventable, and suicide prevention should be the responsibility of everyone. However, the issue with this type of mental illness is that suicidality cannot be detected at labs, and people usually don't cooperate with clinicians. Researchers use other data sources, such as personal diaries (Peck, 1989) and suicide notes (Leenaars et al., 2001), to understand the causes of illness in a better manner.

1. Motivation

Based on the latest observation by the World Health Organization (WHO) (Organization, 2014), suicide is regarded as the 13th main cause of death worldwide and the prominent reason of demise among youths. Social media provides a platform to express feelings and provides an ample opportunity to capture behavioural attributes. Various studies found that the use of language, interacting and expressing oneself gives various indications about mental health. The well-known tool Linguistic Inquiry and Word Count (LIWC) is a tool that is usually used for the psychometricanalysis of the language. Recent researchers have provided enough evidence that individuals experience very contented towards sharing their moods via socialmedia that could be used as a prevention tool (Ahuja et al., 2014; Jashinsky et al., 2014; O'Dea et al., 2017). Peers or website moderators can notice suicidal sentiments, reply to the victim, or report it to a professional suicide prevention facility. As time plays a very vital role in the efficacy of suicide intervention, detecting and recognising the implicit suicidal statements at a proper time becomes a challenging task for moderators.

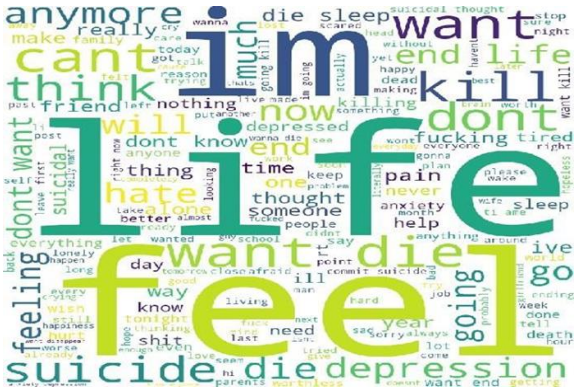


Figure 3. Word Cloud of the Suicidal Dataset

There arises a dire need where suicide prevention specialists need to monitor public user- generated content to detect the clues related to the Suicidality. However, there exist several roadblocks in the way of prevention workers monitoring user-generated material due to the information overload (Desmet, 2014). The

prevention workers can use the keyword filtering approach to uncover the relevant content and cut down the irrelevant content, but using the specific terms/ warning signs manually for detecting the posts cover only a specific explicit suicide statements and don't help in detecting the implicit statements related to the suicide. Moreover, using multiple keywords for querying the statements would lead to false alarms. Another issue with the manual keyword monitoring is that the chat word issues associated with the language on social media will always make it hard to retrieve the message and add false negative values. Thus there is a need for an intelligent computational mechanism that could help the prevention workers accurately identify the suicidal posts and alsothe level of risk in those posts without manually searching through various keywords for effective intervention. The mechanism can in turn, help the potential suicidal individuals to get the proper care in real-time, which wasotherwise very hard to detect using traditional approaches.

2. Objectives

The main goal of this endeavour is to advance and assess a methodology/models for detecting at-risk suicidal posts from individuals on social media. There has developed aTechnique/ Framework for extracting the data such as tweets/posts from social media, as the important challenge in this domain is the non-availability of public data. There is a need to collect various social media posts containing both suicidal and non-suicidal content. It is propose the human annotation scheme of data. The methodology will be presented and applied that will help is labelling the suicidal social media posts for the classification task. Pre-processing and refining the raw data by removing the irrelevant and unwanted data to make the dataset ready for Machine Learning. There has a mechanism / Models being developed which will automatically classify the at-risk tweets/posts in different levels of concern like suicidal and non- suicidal text and based on the severity of text concern to fatality like high, moderate and no risk by applying the various machine learning techniques.

3. Research Challenges

Analysing and detecting the post exhibiting suicidal ideation on social media remains the most challenging task due to the following issues.

- ***Extracting data and Human annotation:* -** The suicidal data or data of at-risk persons is not freely available on social media. Extracting the required data of suicidal users from social media is a daunting task as social networking sites usually do not provide access to their data. There is a need to investigate the various approaches to extract relevant data for research purposes by following the proper ethics. Another difficulty is validating such data from experts in psychology and psychiatry. Thus annotating thelarge corpus is another big challenge.
- ***Implicit nature of posts*: -** The posts indicating suicidal ideation are notalways explicit in nature. They may implicitly indicate about the suicide (e.g.) the post maybe "The life is full of miseries, Better is to get out as soonas possible" instead of "I am going to kill myself".
- ***Stigma and taboo linked with psychological sickness:* -** There is a humiliation allied with psychological disorder. Individuals usually feel shy or embarrassed totalk about their problem. It is thus a challenging task to analyse and evaluate the said illness out of the social media content of the affected persons/users.
- ***Linguistic Norm and chat word issues:* -** Social media data does not follow the linguistic norm. Misspellings, abbreviations are found commonly on social media. Grammatical errors are also found in the language of social media, e.g. Siecide in place of suicide. Chat words also follow different pattern making it hard to detect the emotions behind the text.
- ***Capturing sentiment:* -** Using social media to capture emotional tweets is another challenging task. It is difficult to distinguish between non-sarcastic posts and sarcastic posts. The hash tags describing the tweets like "Want to die" are usually used sarcastically. Hash tags describing the anxiety can describe the situation arising from some examination. It is not necessarily a suicidal concern.

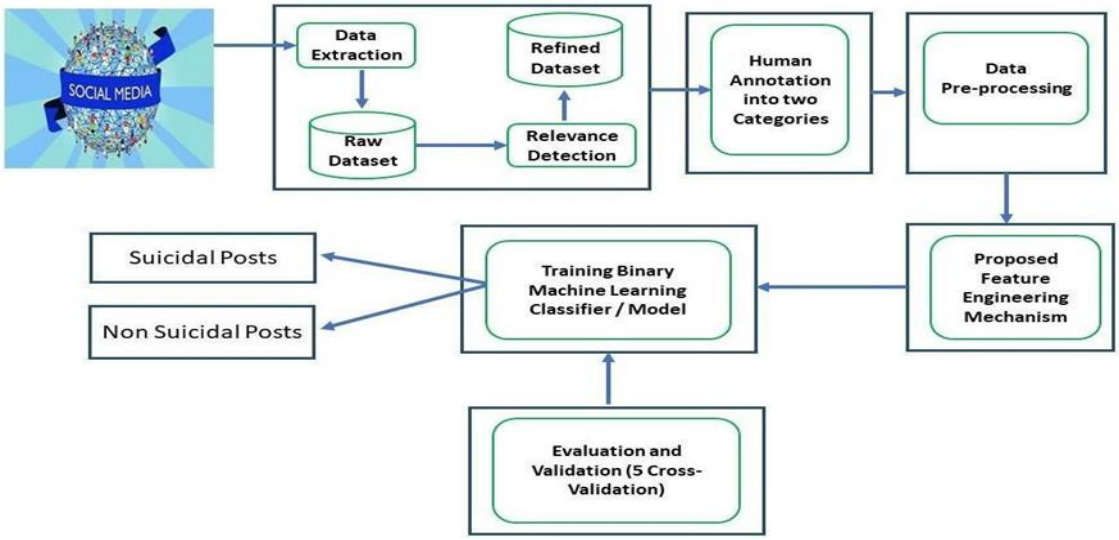


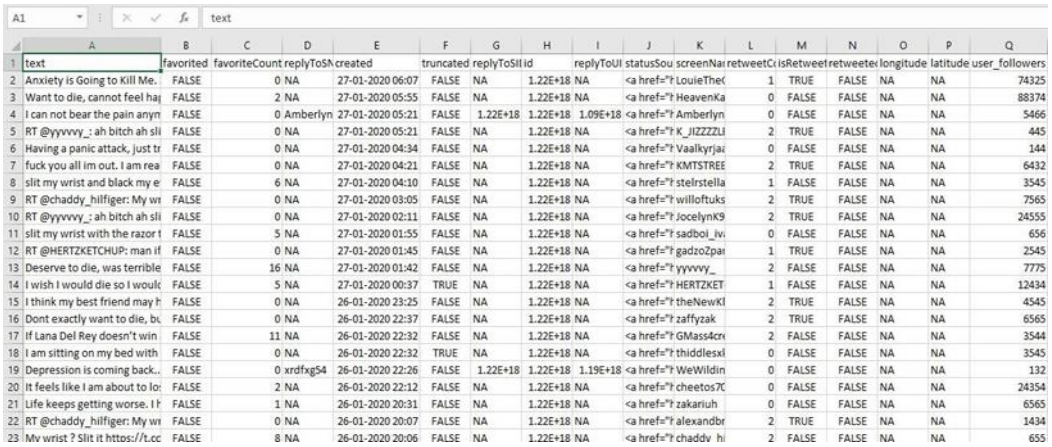
Figure 4. Methodology adopted for Classification of Suicidal and Non-Suicidal Posts

Methodology

The methodology works in five phases. At first there being develop of terms against variety suicide subjects considering diverse withdrawal approaches like application programming interfaces (API's) to cite the appropriate facts. Secondly it annotates the fact sheets. The next i.e. in third, most critical part of our methodology that deals with the preprocessing and proposed feature engineering mechanism. Forth part discusses about the execution of several machine learning procedures for building the model. In the final there becomes assessment and confirmation of the outcomes.

Two social networking sites, Twitter and Reddit, were used to extract the relevant data using their API's. The extracted data from Twitter contains various field like text, screen name, retweet count and various other fields. Our model needs only the text and the label field to train the machine learning model, and other fields are unnecessary for our purpose. Moreover, we also extracted the data from the subreddit r/SuicideWatch by using the Praw. Reddit instance to extract the particular posts from thesubreddit. A total of 14195 posts were used to train the machine learning model. Figure 5.2. shows the sample tweets extracted using the API of Twitter

3.1 Data Collection



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	text	favorited	favoriteCount	replyToSN	created	truncated	replyToSI	id	replyToUI	statusSou	screenNa	retweetCo	isRetweet	retweeted	longitude	latitude	user_followers
2	Anxiety is Going to Kill Me.	FALSE	0	NA	27-01-2020 06:07	FALSE	NA	1.22E+18	NA	<a href="h	LouieThe(	1	TRUE	FALSE	NA	NA	74325
3	Want to die, cannot feel ha	FALSE	2	NA	27-01-2020 05:55	FALSE	NA	1.22E+18	NA	<a href="h	HeavenKa	0	FALSE	FALSE	NA	NA	88374
4	I can not bear the pain anyn	FALSE	0	Amberlyn	27-01-2020 05:21	FALSE	1.22E+18	1.22E+18	1.09E+18	<a href="h	Amberlyn	0	FALSE	FALSE	NA	NA	5466
5	RT @yyvvvy_: ah bitch ah sli	FALSE	0	NA	27-01-2020 05:21	FALSE	NA	1.22E+18	NA	<a href="h	K_JIZZZLI	2	TRUE	FALSE	NA	NA	445
6	Having a panic attack, just tr	FALSE	0	NA	27-01-2020 04:34	FALSE	NA	1.22E+18	NA	<a href="h	Vaalkyrjai	0	FALSE	FALSE	NA	NA	144
7	fuck you all im out. I am rea	FALSE	0	NA	27-01-2020 04:21	FALSE	NA	1.22E+18	NA	<a href="h	KMTSTREE	2	TRUE	FALSE	NA	NA	6432
8	slit my wrist and black my e	FALSE	6	NA	27-01-2020 04:10	FALSE	NA	1.22E+18	NA	<a href="h	stelrstella	1	FALSE	FALSE	NA	NA	3545
9	RT @chaddy_hilfiger: My wr	FALSE	0	NA	27-01-2020 03:05	FALSE	NA	1.22E+18	NA	<a href="h	willoftuks	2	TRUE	FALSE	NA	NA	7565
10	RT @yyvvvy_: ah bitch ah sli	FALSE	0	NA	27-01-2020 02:11	FALSE	NA	1.22E+18	NA	<a href="h	JocelynK9	2	TRUE	FALSE	NA	NA	24555
11	slit my wrist with the razor t	FALSE	5	NA	27-01-2020 01:55	FALSE	NA	1.22E+18	NA	<a href="h	sadboi_iv	0	FALSE	FALSE	NA	NA	656
12	RT @HERTZKETCHUP: man if	FALSE	0	NA	27-01-2020 01:45	FALSE	NA	1.22E+18	NA	<a href="h	gadzoZpa	1	TRUE	FALSE	NA	NA	2545
13	Deserve to die, was terrible	FALSE	16	NA	27-01-2020 01:42	FALSE	NA	1.22E+18	NA	<a href="h	yyvvvy_	2	FALSE	FALSE	NA	NA	7775
14	I wish I would die so I would	FALSE	5	NA	27-01-2020 00:37	TRUE	NA	1.22E+18	NA	<a href="h	HERTZKET	1	FALSE	FALSE	NA	NA	12434
15	I think my best friend may h	FALSE	0	NA	26-01-2020 23:25	FALSE	NA	1.22E+18	NA	<a href="h	theNewKl	2	TRUE	FALSE	NA	NA	4545
16	Dont exactly want to die, bu	FALSE	0	NA	26-01-2020 22:37	FALSE	NA	1.22E+18	NA	<a href="h	zaffyzak	2	TRUE	FALSE	NA	NA	6565
17	If Lana Del Rey doesn't win	FALSE	11	NA	26-01-2020 22:32	FALSE	NA	1.22E+18	NA	<a href="h	GMass4cr	2	FALSE	FALSE	NA	NA	3544
18	I am sitting on my bed with	FALSE	0	NA	26-01-2020 22:32	TRUE	NA	1.22E+18	NA	<a href="h	thiddlesxl	0	FALSE	FALSE	NA	NA	3545
19	Depression is coming back..	FALSE	0	xrdfxg54	26-01-2020 22:26	FALSE	1.22E+18	1.22E+18	1.19E+18	<a href="h	WeWildin	0	FALSE	FALSE	NA	NA	132
20	It feels like I am about to los	FALSE	2	NA	26-01-2020 22:12	FALSE	NA	1.22E+18	NA	<a href="h	cheetos7(	0	FALSE	FALSE	NA	NA	24354
21	Life keeps getting worse. I h	FALSE	1	NA	26-01-2020 20:31	FALSE	NA	1.22E+18	NA	<a href="h	zakariuh	0	FALSE	FALSE	NA	NA	6565
22	RT @chaddy_hilfiger: My wr	FALSE	0	NA	26-01-2020 20:07	FALSE	NA	1.22E+18	NA	<a href="h	alexandbr	2	TRUE	FALSE	NA	NA	1434
23	My wrist ? Slit it https://t.co	FALSE	8	NA	26-01-2020 20:06	FALSE	NA	1.22E+18	NA	<a href="h	chaddy_hi	2	FALSE	FALSE	NA	NA	655

Figure 5. Screenshot of the Tweets Extracted through Twitter API

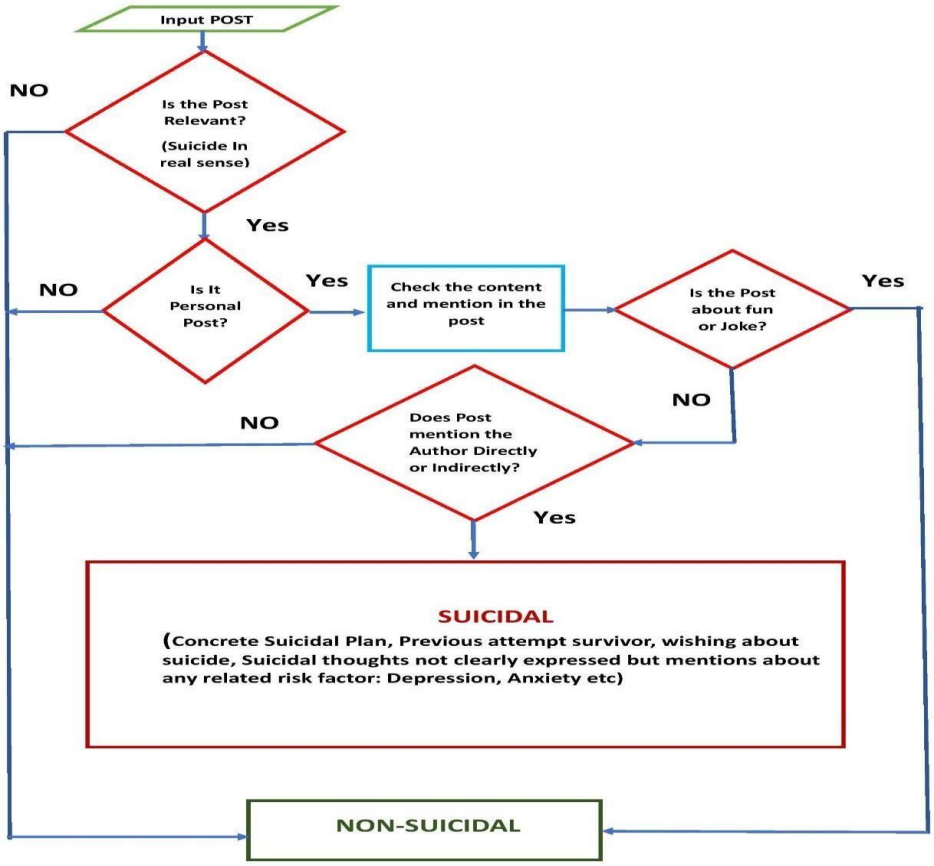


Figure 6. Annotation Scheme for Binary Classification

3.2 Pre-processing

Pre-processing the text data (Anand et al., 2018; Angiani et al., 2016) is the most critical step for preparing the dataset for machine learning algorithms. Textual posts usually contain the part of information that does not have anyrole to play in classification and, if included, can increase the complexity in analysis. Some pre-processing steps arecommon for text classification problems, but other steps depend upon the problem statement and impact the final output. So, in this step, we try to remove the noise from our dataset and only retain that part of the data that could be used for extracting the relevant features. Moreover, this step reduces the dimensions to a larger extent. We used the R language to pre-process our dataset through the following two broad steps.

3.3 Tokenisation

Tokenisation is the process of breaking textual data like posts/streams into individual, meaningful terms. Each of these terms is called a token. The individual, meaningful terms included the words, symbols, punctuation marks, abbreviations etc. These tokens became the input for further processing in the processing pipeline.

3.4 Normalisation

After the tokenisation, another essential step to preprocess our data was to apply the normalisation techniques. The first normalisation step was to remove the elements like punctuations, symbols and numbers that do not convey any meaningful information and unnecessarily take part in the curse of dimensionality if left unattended. For example, "depression" and "depression!" are treated as different features if punctuation is not removed. So is the case with symbols and numbers also. The next step in text standardisation was to convert all the letters into smaller cases. The reason for converting all the words in the smaller case is that two words with different cases are treated as different elements in the vector space model. (e.g.) "*Suicide, "suicide"* and "*SUICIDE*" are treated as two different

words resulting in an unnecessary increase in the dimensions. The stop words (*"is, this, then, a, an")* etc., which are the common words in all documents and do not play any part in distinguishing the documents, were also removed. Another essential step that was performed was the stemming of the documents. We transformed all the words to their root form. Various words convey the same semantic meaning and are the different variations of a common word called a Root word. For example, "*killing"* and "*kill"* reflect the same thing in opinion mining. So the concept is to strip out the words and reduce the inflection towards their root form. This process also helps to reduce the dimensions.

We implemented tokenisation and Normalisation using packages in the R language. Tokenisation is performed using the tokens ( ) function that takes many parameters. The first parameter is the text, the second parameter is "what", meaning how to tokenise the document. We set it to the "word" to perform the word-level tokenisation. The other parameters "remove_punct", "remove_symbols", "remove_numbers", "remove_url", "remove_separators" were all set to true to perform the normalisation of the data. Another function, tokens_tolower ( ), was used to perform the lower casing of the data.

tokens_wordstem ( ) was used to perform the stemming and remove the inflection towards the root word. Various other functions were also used to perform other preprocessing steps. An example below shows the original text before and after applying tokenisation is as under

(1)Original text: [b] I think my "deAth moment" is coming [sad]… and, I am getting CLOSE and, I am not sure what to do anymore.

(2)Cleaned text: I think my death moment is coming and I am getting close and I am not sure what to do anymore.

3.5 Feature Engineering

As machine learning algorithms deal with numerical data, the text must be changed into numerical features. The simple text feature extraction technique, the document-term matrix, can provide a solution. However, the issue with the textual data is that as the dataset grows, the curse of dimensionality increases and a large number of features often represent a small number of instances and confuse the model. Thus only the most dominating and relevant features need to be included to help the classifierdifferentiate between instances and provide an accurate prediction. Data sparsity also needs to be resolved, and dimensions of the feature space need to be projected to the lower-dimensional subspace. Each tweet/post is represented as a vector of numerical values called as features. An extra feature called "class label" is also added for the training data. The features help the model to learn and differentiate between classes. The below subsections discuss various features that collectively represent our hybrid article withdrawal mechanism that consists of features generated throughvarious techniques like Bag of Words, Term frequency-inverse document frequency, Latent semantic indexing, Average cosine similarity and Lengthof the posts.

Pseudocode1_EFASI: Enhanced Feature Engineering Approach for SuicidalIdentification (EFASI)

1: Necessitate: Clarified talks (Tweeter and Reddit talks) ($P_{input.csv}$), Classifier_Name, Classifier_Hyperparameters

2: Certify: Suicidal posts (P_S) and Non-Suicidal Posts (P_{NS})

3: for i makes 1 to n (Complete number of tweets) **do 4:** C [i] = P_{input}[i] $ Tag //Count tags

5:Text.csv = c [i] //CSV file holds tweets with respective tags

6: end for

7: Pro = tokens (Text.csv) //Tokenization and other text standardization's

8: Pro = tokens_tolower (Pro) // Minor issue

9: Pro = tokens_remove (Pro) // Manual artefact halt tweet exclusion

10: Pro = tokens_stem (Pro) // curtailing

11: Text2.csv = Pro // Processed file

12: for i from 1 to n **do**
13:P_length[i] = nchar (Text2 [i]) // span of tweet 14: end for
15: Tokens = tokens_ngrams (Text2, n= 1:3) //ngram structures upto 3 grams of complete dataset
16: Tokens.dfm = dfm (Tokens) //mark article piece medium
17: dfm_trimmed = dfm_trim (Tokens.dfm, min_docfreq, min_termfreq) // trimming the tokens with fewer significance
18: TFIDF_ Sorts = dfm_trimmed. Tfidf // Extracting TFIDF Sorts
19: unfinished. cases <- which (! whole. cases (TFIDF_ Sorts))
20: TFIDF_ Sorts [unfinished. cases,] <- rep (0.0, ncol (TFIDF_ Sorts))
// substituting partial cases
21: LSA_ Sorts = SVD (TFIDF_ Sorts, nv=50,100,150,200, 250)
// Utmost related kind removed via dimensionality drop considering unique magnitude disintegration of LSA
22: train.similarities = cosine (LSA_Features)
// Finding similarity of doc's based on Cosine measure
23: for i from 1 to n row **do**
24: ASS [i] = mean (train. similarities [i, Suicidal post index)) // Regular suicide resemblance built upon cosine function
25: end for
26: Optimal _structure Set = LSA_sort + P_length + ASS
27: ClASSIFIER (Classifier_Name, Classifier_ Hyperparameters, CV=5, Optimal _Feature Set)

3.6 Machine Learning Algorithms

Machine learning algorithms are computer programmes that adapt to new data and improve their performance. The "learning" aspect of machine learning refers to the fact that these algorithms modify how they analyse data over time, much like people do. The free lunch theorem (Wolpert & Macready, 1997) explains that no algorithm can be strictly called better than other algorithms. One algorithm performing better on a particular problem doesn't mean that it can perform the same on other problems. It is therefore necessary to thoroughly experiment the algorithms, fine-tune them and determine the best algorithm for Classification.

3.7 F1 Score

F1 Score is a metric that is used to measure the performance of the classifier/model when that model needs a balance between Precision and Recall & also when the dataset is imbalanced, having a large number of actual negative values. Usually, for learning models, false positives and false negatives provide an important role. The F1 score tries to give more weight to these values and contribute in minimizing the impact of true negative values.

$$F1\,Score = Precision * Recall\,Precision + Recall$$

Conclusion

Ultimately the overall pathology of suicide, the role of the internet in suicidal behaviour, various hurdles in online towards prevention of suicide have suitably demonstrated. The general proposed annotation scheme based upon various criteria used for labelling the tweets/posts have significantly laid out. A detailed analysis of the methodology based upon the proposed enhanced feature engineering mechanism that extracts and identifies the most relevant features, which are then supplied to machine learning algorithms to enhance accuracy and overall F1 Score.

Future Scope

Developing the machine learning model to analyse suicidal posts on social media is in its infancy because of the unavailability of data due to privacy and ethical issues. There is a dire need to train a Multi-class machine learning model on a larger dataset containing the real emotions of people suffering from suicidality. There is a need to use a rich feature engineering

mechanism for the extraction of relevant features.

References

[1] World Health Organization. *Suicide worldwide in 2019: global health estimates* Geneva: World Health Organization; 2021.

[2] Berrouiguet S, Courtet P, Larsen ME, Walter M, Vaiva G. Suicide prevention: towards integrative, innovative and individualized brief contact interventions. *Eur Psychiatry.* 2018; 47:25–6.

[3] Chesney E, Goodwin GM, Fazel S. Risks of all-cause and suicide mortality in mental disorders: a meta-review. *World Psychiatry.* 2014; 13(2):153–60.

[4] Nordentoft M, Mortensen PB, Pedersen CB. Absolute risk of suicide after first hospital contact in mental disorder. *Arch Gen Psychiatry.* 2011; 68(10):1058–64.

[5] Runeson B, Haglund A, Lichtenstein P, Tidemalm D. Suicide risk after nonfatal self-harm: a national cohort study, 2000–2008. *J Clin Psychiatry.* 2016;77(2):240–6.

[6] Parra-Uribe I, Blasco-Fontecilla H, Garcia-Parés G, Martínez-Naval L, Valero-Coppin O, Cebrià-Meca A, et al. Risk of re-attempts and suicide death after a suicide attempt: a survival analysis. *BMC Psychiatry.* 2017; 17(1):163.

[7] Observatoire national du suicide (France). Suicide: quels liens avec le travail? Penser la prévention et. les systèmes d'information: 4ème rapport; 2020.

[8] Torous J, Larsen ME, Depp C, Cosco TD, Barnett I, Nock MK, et al. Smartphones, sensors, and machine learning to advance real-time prediction and interventions for suicide prevention: a review of current progress and next steps. *Curr Psychiatry Rep.* 2018; 20(7):51.

[9] Fonseka TM, Bhat V, Kennedy SH. The utility of artificial intelligence in suicide risk prediction and the management of suicidal behaviors. *Aust N Z J Psychiatry.* 2019; 53(10):954–64.

[10] Lindh ÅU, Dahlin M, Beckman K, Strömsten L, Jokinen J, Wiktorsson S, et al. A comparison of suicide risk scales in predicting repeat suicide attempt and suicide: a clinical cohort study. *J Clin Psychiatry.* 2019; 80(6).

[11] Franklin JC, Ribeiro JD, Fox KR, Bentley KH, Kleiman EM, Huang X, et al. Risk factors for suicidal thoughts and behaviors: a meta-analysis of 50 years of research. *Psychol Bull.*. 2017; 143(2):187–232.

[12] Bernert RA, Hilberg AM, Melia R, Kim JP, Shah NH, Abnousi F. Artificial intelligence and suicide prevention: a systematic review of machine learning investigations. *Int J Environ Res Public Health.* 2020; 17(16):5929.

[13] Berrouiguet S, Billot R, Larsen ME, Lopez-Castroman J, Jaussent I, Walter M, et al. An approach for data mining of electronic health record data for suicide risk management: database analysis for clinical decision support. *JMIR Ment Health.* 2019; 6(5):e9766.

[14] Miller DD, Brown EW. Artificial intelligence in medical practice:the question to the answer? *Am J Med.* 2018; 131(2):129–33.

[15] Burke TA, Ammerman BA, Jacobucci R. The use of machine learning in the study of suicidal and non-suicidal self-injurious thoughts and behaviors: a systematic review. *J Affect Disord.* 2019; 245:869–84.