Exploratory analysis of depression pattern within twitter post and investigation of suicidal intensity

Md Iftekharul Mobin[1,2*], Second Author[2,3†] and Third Author[1,2†]

[1*]Department, Organization, Street, City, 100190, State, Country.
[2]Department, Organization, Street, City, 10587, State, Country.
[3]Department, Organization, Street, City, 610101, State, Country.

*Corresponding author(s). E-mail(s): iftekhar.mobin@gmail.com;
Contributing authors: iiauthor@gmail.com; iiiauthor@gmail.com;
†These authors contributed equally to this work.

Abstract

In this paper, exploratory data analysis is conducted with Reddit and Twitter dataset to investigate the suicidal intensity within a depressed person's post. These social meadia's Text datasets have become ideal choice for analyzing depression and suicide classification among the NLP researchers. Main objective is to estimate, if a depressed person has a propensity for suicide, gauge the severity of the depression or vice versa. Prior assumption that depressed people do posses some extent of suicidal thought. In this research, for training the classifier Reddit C-SSRS Suicide Dataset is used that contains the suicidal risk classification label and has defined features belongs to specific suicidal risk. Trained classifier, detects the suicidal risk within twitter's public post which contains depression vs suicide category. From the observed result, statistical measurements depicts the degree of suicide intensity within depressed label post. Different vectorization methods and machine learning estimators are used to obtain best possible results. From the analysis it is revealed that suicidal tendency within depression people post is extremely high. Depressed person's post in the twitter showed 60% similarities in various categories of suicidal intensity.

Keywords: keyword1, Keyword2, Keyword3, Keyword4

1 Introduction

Suicide is a significant cause of death worldwide. In india, USA and many other countries significant number of people die because of suicide [10, 32]. Only In USA approximately 46,000 people committed suicide in 2020 [32]. Many countries the second highest cause of death among teenagers and younger adults is suicide. Suicide and depression are major health hazards, resulting in the death of one person every 40s globally. More than 300 million people worldwide experience depression annually, 800,000 people die by suicide. These two are intertwined phenomena. According to [32] About 4% of individuals diagnosed with depression commit suicide, and more than half of the persons who attempt suicide meet the criteria of depression. Depression triggers suicidal risk. Several studies showed that depression depression patients are very prone to suicidal attempt [11, 22, 36]. To what extent of depression level triggers suicidal risk is a scrutiny.

Clinical depression severity estimation methods rely on interview based interrogation session where patient confront with psychologist. During the interrogation session patient may not be honest about expressing their thoughts. It is common phenomenon that emotionally distressed individual hides their feelings to others. More-often patients prefer not to disclose their emotions, often reluctant to seek help from psychotherapists, or doctor. Hence, conventional interview-based diagnosis is insufficient to accurately predict a psychiatric status.Also, It is hard to quantify the level of depression during suicidal attempt. It may varies based on various factors like society, religion, family bonding, emotional maturity and many others factors. Due to lack of confidence, fear of death, religious obligations, and societal stigma against this act, even severe depressed person may not consider making an attempt at suicide. But they seek empathy consciously or unconsciously in the social sites like twitter, reddit and facebook [7]. Shen et.al in [28] and Xu et al. [40], depicted how online users debate topics connected to depression in social networks and what is their language patterns. Choudhury et al. in 2013 [8] showed that there is possibility of detecting and diagnosing depression via social media. [23] conducted face-to-face interviews with 14 active Twitter users, to investigate the depressed behaviors in social media users. With the aforementioned research, it is clearly revealed that social media depression detection is not only possible but promising result can be observed. Severity level of depression can be determined by analyzing social media activities. Through data visualization we can explore various facts and clues among this two emotions. Our research focus on detection of suicidal tendency within a depressed person's post. Find out important features, explore different facts and hidden underlying information of depression and suicide.

2 Literature Review

Extensive research has been conducted before about depression and suicide. In 2020 Chancellor et. al conducted systematic literature review of the mental health status prediction using social media data [6]. More than 75 studies were taken into consideration between 2013 to 2018. It provides a detailed overview of data collection sources, data annotation methods, pre-processing and feature selection, model selection followed by accuracy estimation, cross validation and models' benchmarks for mental illness. In 2022 Zhang et. al claimed that 399 scientific research papers were reviewed and mental illness related research is increasing gradually [42]. There were other review papers on these two issues in which similar topics are analyzed such as Castilla et. al in 2020 [5] and Malhotra et. al 2022 [19]. All of these review based research studies analyzed mostly social media's Text data, and discussed NLP tools and techniques. In some research multimodal data samples are used such as: along with social website post Instagram images are taken into consideration. Along with status of the social website post, Electronic Health Record (EHR) has also been taken into consideration by zheng et. al in [43] and Paulo et. al in [20] in 2020. In 2021 Lang He et. al conducted research of Audio visual features [12] aiming how facial expression and voice can be used as an input features to determine mental illness effectively. Visual impact on individuals to detect depression also has been studied in some papers [41]. Martinengo reviewed 69 apps in 2019 that provides suicide preventive solutions and support [21]. Most of these apps have emergency calling features that can support mental patients over phone or app. It is observed that Text based expression depicts mental illness more clearly compared to other features and is dominant among researchers to detect mental issues effectively. In this study we will be focusing on the text based features only for mental illness and suicidal pattern detection.

2.1 Instrument for measuring Severity

From the very beginning questionnaire based suicidal/depression intensity measurements tool were available. These scale are applied for setting the questionnaire during the interrogation. This process provides weights to answers replied by individuals. Most renowned scales are PHQ-9, BSS [2], DSM-5, DASS-21 [10], Columbia Suicide Severity Rating Scale (C-SSRS) [13, 25] etc. Most of these scales are based on predefined questions and answers. [3, 10, 14, 15, 34, 39]. These scales are mostly specified questions and multiple choice questions and answers have weight 1 5 and according to the feedback from the user severity of symptoms are decided. Then statistical Machine Learning models are applied to investigate patterns [29]. [28] The Beck's Depression Inventory [2] consists of 21 questions regarding the users' physiological and mental states. Another example is the CES-D Scale [26], which has 20 questions concerning users' sleep patterns and guilty feelings. Users must either rate the severity of their circumstances in response to the questions, which either offer many answers with varying scores, or both. According to the scale of the overall score, the degree of depression is identified. In contrast, the Diagnostic and Statistical Manual of Mental Disorders (DSM) [38] offers nine different types of depressive indications, including low mood and impaired interest. Before making a final judgment, clinicians typically determine if these symptoms have been prevalent throughout time.

In [36] research study conducted in the City of Vantaa, Finland, for the age group of 20 to 69 years with 1119 primary-care patients using (PRIME-MD) questionnaire. Suicidal behaviour was investigated and present suicidal ideation was measured with the Scale for Suicidal Ideation (SSI) and afterwards suicide attempts were evaluated based on medical records. Survey is conducted on specific demographics of population, Mostly statistical machine learning methodologies are applied to determine risk, features and intensities within collected samples [29]. In [29] 4,882 medical students were surveyed on the basis of demographic and clinical records via WeChat app. In [9] collected dataset of 2181 redditors post which contains posts related to suicidal ideation, behavior, or attempt. Then assisted by professional practitioners, the psychiatrists prepared a dataset of only 500 redditors using tool called C-SSRS [25]. [35] Within the Vantaa Primary Care Depression Study (PC-VDS, 91 patients) and the Vantaa Depression Study (VDS, 153 psychiatric out-and 41 inpatients), suicidal ideation was assessed with the Scale for Suicidal Ideation (SSI), Hamilton Depression Scale (HAM-D) item 3 and Beck Depression Inventory (BDI) item 9, and by asking whether patients had seriously considered suicide during the episode. The positive and negative predictive values (PPV, NPV) for suicide attempts during a six-month follow-up were investigated. These standards have been successfully validated and used for many years in real-world situations, without a doubt. Till date many researchers are using these scales to determine suicidal symptoms [16]. However, these scale might not completely encompass introvert patient behaviors and symptoms like current social media. In this study we will be focusing on the text based dataset collected from the socal media sites only for mental illness and suicidal pattern detection.

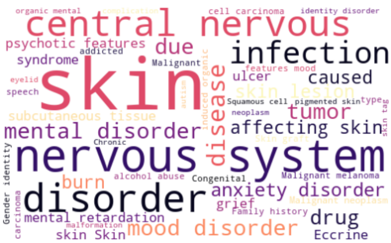
3 Dataset

Text based Suicide and depression detection classification task uses NLP techniques and methodologies which are highly dependent on the quality of dataset, accurate annotation of labels and size of samples. Text based samples are mostly collected from Twitter, Reddit [33], facebook, weibo etc websites donated by various institutions or researchers. Several researchers contributed publicly available dataset [27]. In 2021 the Computational Linguistics and Clinical Psychology CLPsych 2021 workshop organized a Task challenge for detecting suicidal risk [18]. It facilitated participants providing sensitive authentic dataset on the problem of predicting suicide risk from social media Twitter. The dataset for the task includes information who attempted suicide or succeeded along with some control who have not. After collecting dataset from social sites, proper labeling is crucial for training machine learning classifier models. In[17] research study used Twitter post collection API for collecting Tweets and collected Tweets of size 2509 were obtained, of which

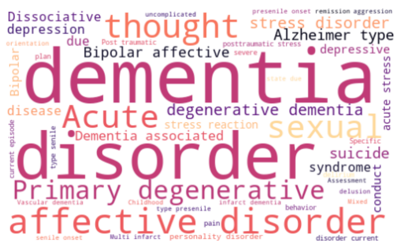
216 post were found relevant by 3 Expert psychologists evaluators. Furthermore, using LIWC, dictionary of the Linguistic Inquiry and Word Count [24], which is a linguistic feature analysis software that calculates the degree of positive and negative emotions across a wide spectrum of texts, Tweets were evaluated and results are statistically presented. In 2018 shing et. al [31], and in 2019 Gaur et. al [9] consulted with the professional practitioner psychiatrist to annnotate the dataset and segregated into several categories. [9] contains gold standard dataset of 500 redditors prepared from 2181 redditors post and validated by four practicing psychiatrists following the guidelines outlined in Columbian Suicide Severity Rating Scale (C-SSRS).

3.1 Training Dataset

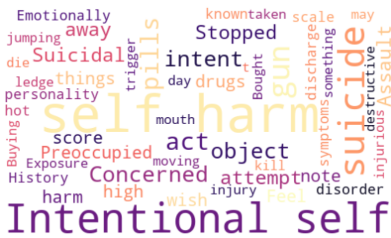
In this research we have used dataset from [9]. For training classifier in this research 2019's Gaur et. al [9] Reddit C-SSRS dataset is used. Compared to the existing four-label classification scheme (no risk, low risk, moderate risk, and high risk), this dataset introduced 5 lavel classification suicide indicator, ideation, behavior, attempt and another extra category incorporating supportive category. Supportive category represents whenever someone shows empathy and condolance for a suicidal post. It is not taken into account in the analysis section since supporting group does not belong to examined specimen individual. Word cloud is showed in figure **??** to depict each category and influence of dominant keywords based on frequency.



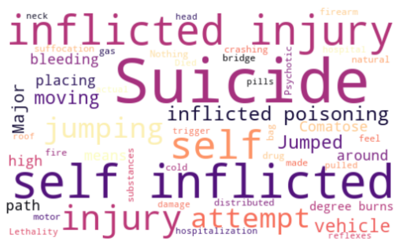
(a) Indicator category word cloud



(b) Ideation category word cloud

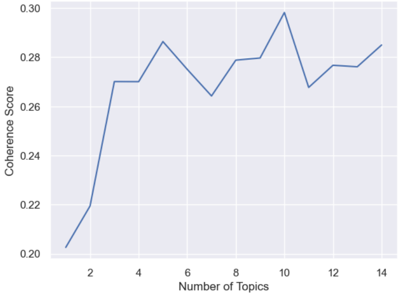


(c) Behavior category word cloud

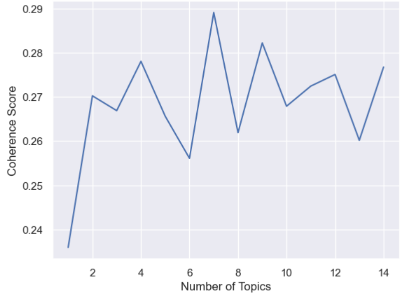


(d) Attempt category word cloud

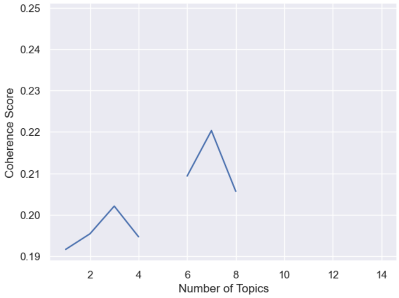
3.2 LDA model coherence to determine optimal topic



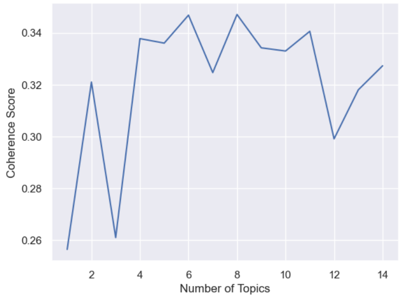
(a) Indicator category word cloud



(b) Ideation category word cloud



(c) Behavior category word cloud



(d) Attempt category word cloud

3.3 Relative importance vs term frequency

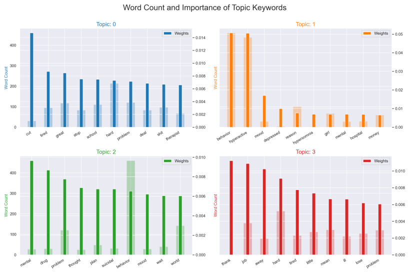
Relative importance is measured using LDA model. Then depicted in figure **??**.

3.4 LDA based feature exploration

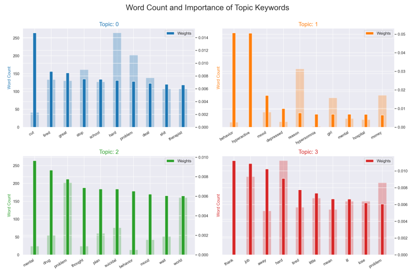
LDA is a versatile tool to investigate various keywords in subtle topic within context. LDA driven results are included in this study to observe the latent topic and terms.

3.5 Testing/Validation Dataset

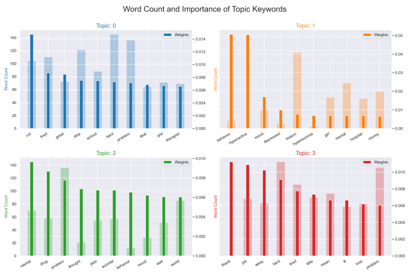
For testing dataset is collected from kaggle. It is an opesource dataset publicly available collected from reddit website by a pushshift API contained suicide and depression category. This publicly available Reddit datasets in Kaggle Website comprised of 232,074 post annotated for binary classification as suicidal or non-suicidal in [1] for detecting suicidal ideation. The dataset is a



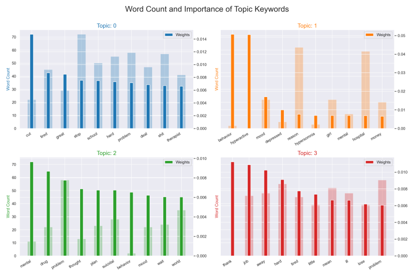
(a) Indicator category word cloud



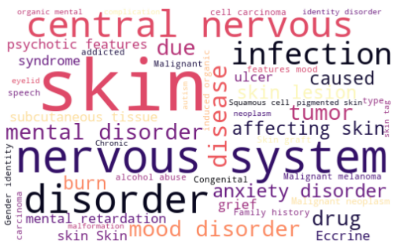
(b) Ideation category word cloud



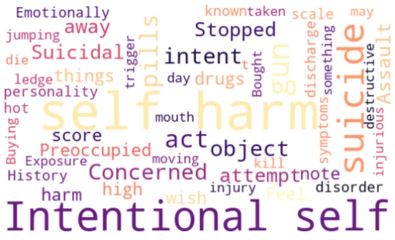
(c) Behavior category word cloud



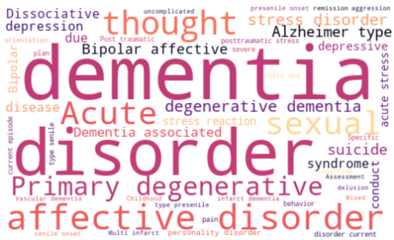
(d) Attempt category word cloud



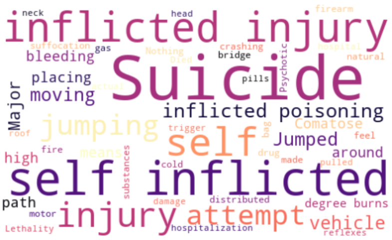
(a) Indicator category word cloud



(b) Behavior category word cloud

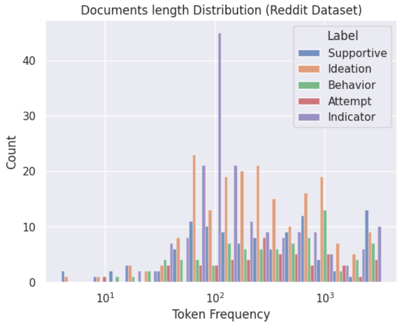


(c) Ideation category word cloud

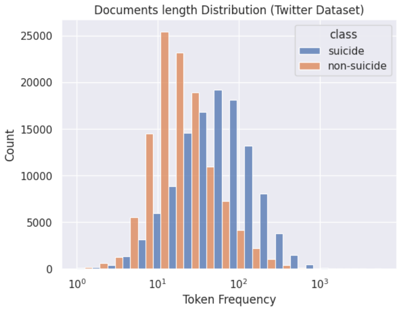


(d) Attempt category word cloud

collection of posts from the "SuicideWatch" and "depression" subreddits of the Reddit platform. All posts that were made to "SuicideWatch" from Dec 16, 2008(creation) till Jan 2, 2021, were collected while "depression" posts were collected from Jan 1, 2009, to Jan 2, 2021. In this research trained classifier is applied to detect class on this two category. Main objective is to determine the suicide categories (indicator, ideation, behavior, attempt, supportive) within this dataset. Document length frequency and token distribution is depicted in Figure 5. From the frequency distribution we can see some of the document sizes are very large. Hence, during the training process we chopped the sentences into multiple sentences keeping the label same.



(a) Reddit dataset distribution



(b) Twitter dataset distribution

Fig. 5: Train and Test dataset Twitter and Reddit dataset distribution

3.6 Data Processing and Models

3.6.1 Data Pre-processing

Social media dataset are mostly Text data which needs data pre-processing, cleaning, feature extraction and data mining related NLP tasks. NLP based text data contains noises such as: unnecessary quotes, special characters, punctuation etc. Moreover, morphological analysis is needed to retrieve root words followed by stemming, lemmatization. Then, sentences are divided into equal-length fragments, and null word padding is applied as needed. Words within a phrase are now referred to as tokens or features, and the dataset is shown as a corpus. Special features/tokens are further preprocessed and filtered using text data feature extraction tools and methods. Features are passed through a process in which features are converted to corresponding IDs and sentences which contains a series of IDs are represented a vector. The embedding is another term that is frequently used in relation to vector text analysis. Various Vectorization methods are present. Traditional vectorization method provide weights to terms/words mainly based on frequency of words within the sentence and documents, rather than its importance and contextual meaning. Also, how does a particular word or term create impact on the neighboring words is not taken into consideration, since these models does not have any prior knowledge of any words. Hence, various neural network based language models are proposed which are pretrained on massive amount of dataset. These models mainly carries weight which represents word to word relationships and most cases can provide contextual meaning of given sentence based on pre-trained dataset knowledge. Deep learning models recently showed remarkable achievements in this case representing corresponding knowledge.

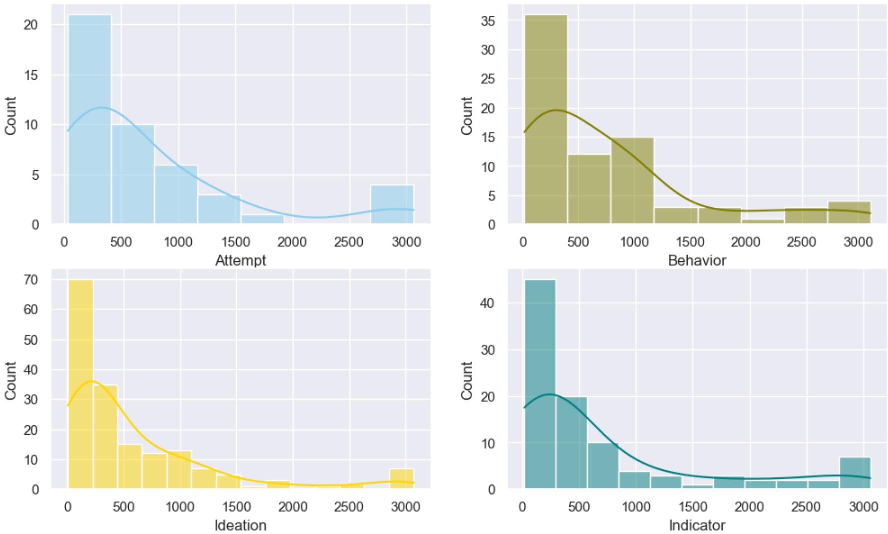


Fig. 6: Document size length in each class of suicidal category and samples frequency

Here, simple algorithm is applied to balancing this dataset. Since some specific features are already provide for each category. Categorical features are mixed together shuffled and then chopped into fixed sized tokens and considered as sample sentences for specific category. This process continues until we found desired number of samples. Shuffling happens randomly and all the samples features only sentence and specific feature sentence are piled together to train classifier.

3.6.2 Classification

Machine learning and Deep Learning models are particularly used for Text classification and ML for feature selection or extraction in several studies [5, 6, 42]. These extensive reviews reveal Deep learning methods receive more attention and perform better than traditional machine learning methods whereas in some cases when extracted or filtered features are fit into training process models are able to perform better. NLP techniques are applied for the annotated dataset collected from Twitter, Reddit, Facebook, instagram, Weibo [37] etc. After that various deep learning and machine learning models are trained for classification of suicide and depression. Then trained models are applied to determine the correct class of given text. [19] Provided a through investigation about passed research techniques, features, datasets, and performance metrics [6, 42].

3.6.3 Data Learning Models

Among Several Deep Learning approaches most successful NLP classifier for segregating Depression and suicidal task are CNN, LSTM, GRU, XLNET, BERT, Variants of BERT RoBERTa and variants of CNN such as: CNN-BiLSTM etc [1, 30, 37] also showed promising results.

3.6.4 Feature Selection

The feature selection procedure has a substantial impact on the performance of machine learning and deep learning models because it lowers noise in the trained dataset, enabling the model to accurately understand data patterns. LIWC, LDA, LSA, n-gram analysis [24, 33] etc are used as features analysis tools. Most dominant approaches are n-gram word frequency based approach TF-IDF. Apart from the Deep learning model n-gram Traditional feature retrieve based analysis conducted in some research papers. In 2017 Shen et. al [**?**] collected several forms of features comprised of six chorots, namely, social network features, user profile features, visual features, emotional features, topic-level features, and domain-specific features and prepared a feature rich dictionary. This multimodal depressive dictionary learning model was used to detect the depressed users on Twitter using machine learning models.

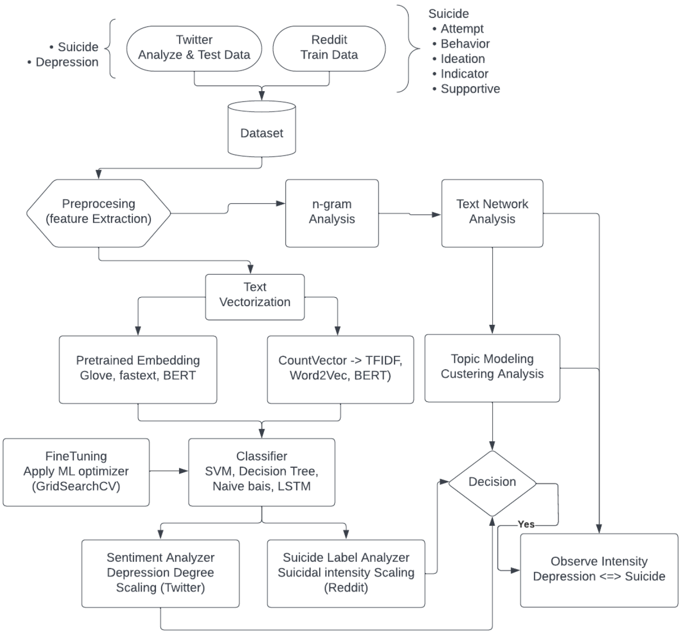
Most dominant approaches are Word2Vec, XGBoost, SVM, Random Forest, other regression MLs. Typical word embedding approaches TF-IDF and Word2Vec, and CNN–BiLSTM are applied in [1]. Using LIWC features, XGBoost ML together surpasses the accuracy of CNN–BiLSTM in [1]. Others have used machine text Summarization based feature extraction strategy followed by classification for depression detection is applied in [44]. [4]built a set of baseline classifiers using lexical, structural, emotive and psychological features extracted from Twitter posts. Then baseline classifiers are updated by building an ensemble classifier using the Rotation Forest algorithm and a Maximum Probability voting classification decision method. [6] This paper provided an excellent overview of 75 studies in between 2013 and 2018 outlining the methods of data annotation for mental health status, data collection and quality management, pre-processing and feature selection, and model selection and verification.

3.7 Other Tools or Methods

Apart from the Machine Learning or statistical model approaches there are some sites, helpline numbers and apps [21] that contains various facts, statistics, tips, tools, healthline numbers and ways to handle suicidal depression [**?**].

4 Methodology

In several research papers reddit dataset has become a good estimator to determine depression and suicide. Reddit dataset is used as a test dataset for classifiers for this research because of its availability and open access. Here classifier is trained with the suicidal intensity determinant Twitter's dataset which is used to determine the depression intensity within a suicidal post. It provides the suicidal risk classification dataset and has specific features belongs to specific category of suicidal risk. As test dataset for classification result Reddit's suicide vs depression dataset are chosen. Trained classifier, detects the suicidal risk within reddit post which contains depression vs suicide category.
,



In this research dataset is pro-processed using the most common techniques of NLP mentioned in the section 3.5.1. We cleaned the dataset by removing unwanted characters, symbols and stop words. Then further pre-processing is conducted which are followed for standard data cleaning process for NLP task.

Pre-processed dataset is converted to vector. Various vectorizer are used to convert the corpus into corresponding vector. Text is converted to meaningful feature vector then classifier is trained to determine depression degree within a suicidal post. Various vectorization methods are applied to get best possible outcomes. Then machine learning classifier is applied to estimate the class category. From the observed result statistical analysis explained the degree of depression within a suicidal post. In our analysis we used different classifiers to determine how suicide intensity is showed in depression vs suicide category dataset. Thereby we can infer degree of suicidal tendency within depression person's post.

Hence, visualizing the result we can determine the suicidal tendency within depression post. Also, N-gram based analysis is conducted and frequency of Terms and connections of words or phrases are analyzed in this research scope. More often topic modeling clustering is used to determine latent topic and understand latent text network. From the network various facts can be revealed related to suicide and depression. This whole process is depicted in Figure **??**

5 Results Analysis

First we started our experiment with document length distribution. The length of document and term frequency within the corpus is visualized in Figure 4. From the distribution we can see that some of the document length are excessive long and contains more than 1000 tokens (within Twitter and also Reddit both Dataset). Depression class document length are usually shorter in length. Depression document length are tend to be smaller than suicide document length.

Short sentence does not carry much terms and hence does not carry enough information to be classified confidently by classifier algorithms. We started reducing the numbers of samples based on document length. By reducing the samples based on numbers of tokens present in a document (see Figure 4). Documents length versus category frequency information is showed in this chart. This charts explains if we filter out the shorter comments suicide post become dominant class and depression post become outnumbered. The difference showed an exponential pattern as length of document increases. Test dataset Reddit data distribution among depression and suicide class distribution ratio is equal. Filtering the class we have seen an interesting fact that depressed people does not want to comment very long.

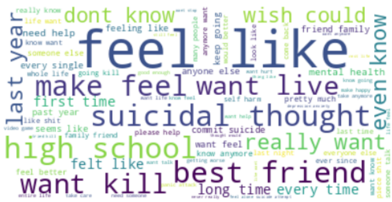
5.1 N-gram Analysis

5.1.1 Uni-gram

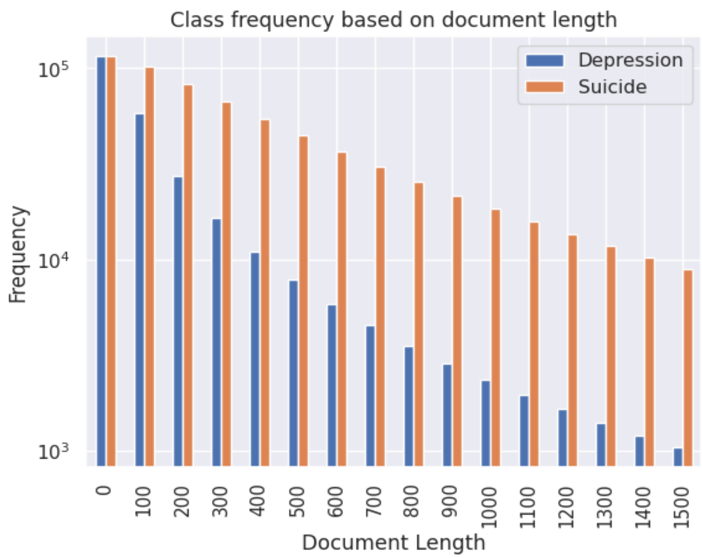
Dataset is split into separate tokens after preprocessing and uni-gram generated. Based on frequency of words wordcloud is generated from these unigrams. Frequency based comparison between two categories is conducted for depression and suicide for Test dataset in Figure 6. Main objective was to get top ranked words from Depression and Suicide corpus. After experiments we have seen There are similarities between the top ranked words those are occurring frequently. They tend to use slang and abusive terms compared to suicidal attempt thinking people. Rather suicidal depressed people want to share their thoughts with others using longer post. However, it does not reveals any clues in terms of hypothetical relationships between the two category. It is difficult find pattern in which we can determine the depression and suicidal thought. So far we found some pattern

1. short statements likely to be more depression category
2. Depressive statements tend to have slang
3. Suicidal thinking people's post having very high frequency of "kill" "die" these type of words or phrases.

(a) Wordcloud in Depression category



(b) Wordcloud in Suicide category



(c) Class frequency in different Document Length

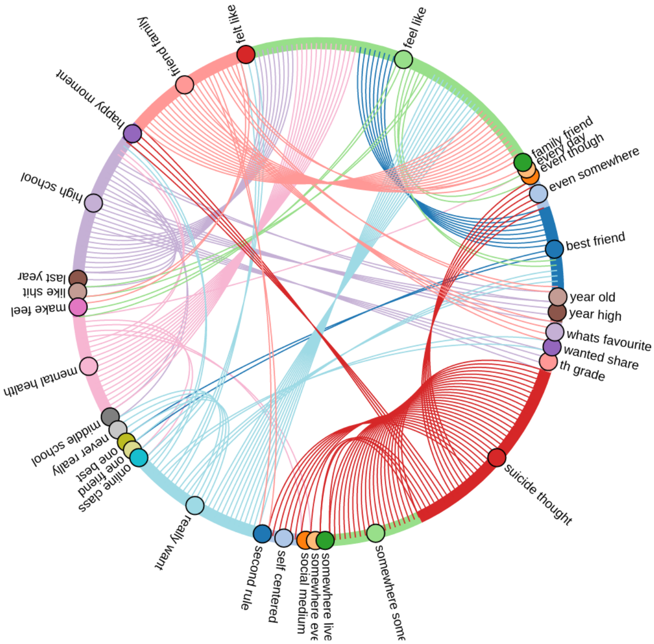
Fig. 7: Dataset features visualization and properties exploration

To understand the term occurring frequently in two different classes scatterplot library is used for visual analysis. From the above two scenario we can see that there is a pattern that people used to say more slang and abusive words when they are depressed. It is also interesting that there are many words have high frequency such as depression or depressed but belongs to suicide class. One important fact is revealed here is that we can see although suicide, suicidal these words has high frequency in Suicide class but depression, depressed also occurred in parallel with high frequency. Here several experiments can be conducted for exploratory analysis with scattertext library for terms significance. However, this library is computationally heavy for larger dataset for visualization. Another drawbacks is this library have significant focus on the terms based analysis. We have used simple vectorization methods by which we can have greater control on dataset and experiments code.

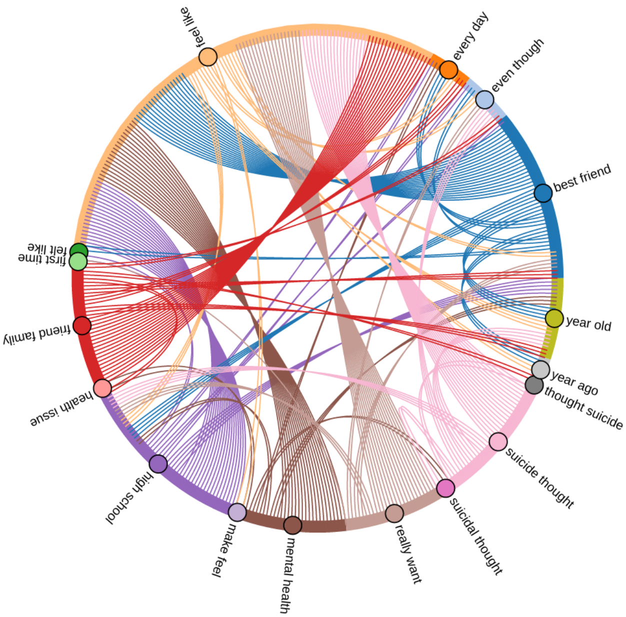
5.1.2 Bi-gram

First unigram is computed and analyzed then bigram is calculated for both categories. The bigram frequency showed there are some common terms like "mental health", "feel like", "make feel", "high school" etc showed high occurrences in the dataset. Hence, we started to understand its pattern in the corpus. For analysis we have considered ['high school', 'mental health', 'best friend', 'feel like', 'really want', 'suicide thought', 'friend family'] these bi-grams and wanted to explore its surrounding context for each category. We called this special bigrams since it showed importance in the suicidal and depression both categories appeared highly frequent matter. We want to analyze how these words have impact with its neighboring words.

To explore the impact of special bi-grams on the samples, special bi-gram terms containing samples are filtered from dataset. After that using lebel encoder bigrams are encoded as integers and then chord diagram is generated depicted in Figure 7 to find meaningful relationship within the samples between the bigram features.



(a) Depression Chord diagram



(b) Depression Chord diagram

Fig. 8: Bi-gram features relation exploration

From this two chord diagram interesting sentence can be inferred. Such as: from the depression class ? self centered person is depressed, having suicidal thought, want to go somewhere to live, spend happy moments and so on. For the suicidal class category suicidal attempt thinking people, have mental health issue, they want to share though with high school friends, best friends, friends and family members, having suicidal thoughts and so on.

Tri-grams or above did not reveals much meaning information, mostly does convey some meaningful information and therefore excluded for further experimental consideration.

6 Classification Results

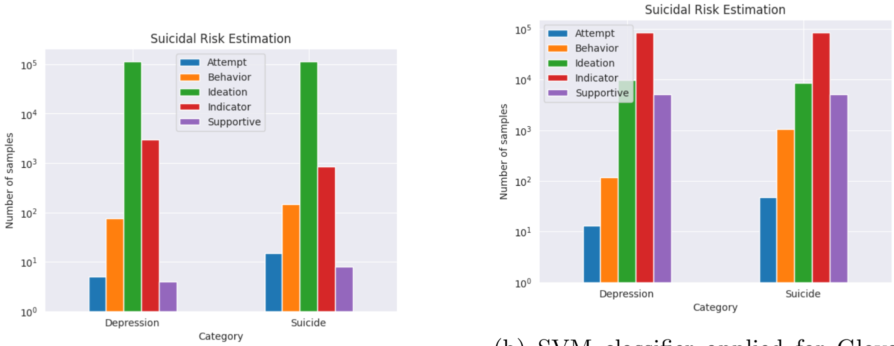
To segregate the Reddit suicide dataset into different categories of suicide first we have created a classifier using different classification techniques. Since our objective is not making highly accurate classifier. Following approach is applied in this study

- Pre-processed and useful features are used from Twitter's 500 post CSSR dataset for Training classifier
- Used count vectorizer and TFIDF transformer to generate vectors for the dataset
- Trained classifier to determine the categories of various suicidal intensities

We have used simple gridsearch technique of sklearn library and from a list of various classifiers applied on the dataset, we have chosen highest accurate classifiers to determine different label of suicidal risk. so that it can recognize the category. We have used classifiers "K Nearest Neighbors", "Linear and RBF SVM", "Gaussian Process", "Decision Tree", "Random Forest", "Vanilla Neural Net", "AdaBoost", "Naive Bayes". Using various set of parameters, from the result and experiments we found almost 60% accuracy for SVM model to predict the suicide intensity categories. From various set of values of SVM we found degree=2, gamma=0.7, kernel=rbf showed the highest accuracy.

6.0.1 Suicidal Intensities visualization

How much depression can trigger suicidal thoughts is an interesting question. In this study classifier is trained on the suicidal intensity. Then trained classifier is applied on the Depression/Suicide class dataset. From various machine learning models we have found SVM is a good performing model. SVM classifier is applied for the TFIDF vectorizer embedding (see results in figure 8a) and also for Word2vec pretrained vectorizer model. The results are shown in figure 8. From the results we can see that suicidal ideation between depression and suicidal categories number of samples are very similar. Within depression more number of samples are showed suicidal indicator category compared to suicide which is an interesting result. Suicidal behavior and attempt is comparatively high within the suicidal category than depression. Hence, figure 8a result seems to be pretty obvious, except for suicidal ideation category. Also for the suicidal indicator symptoms are higher within the depression category.



(a) SVM classifier applied for TFIDF vectorizer

(b) SVM classifier applied for Glove Word2Vec Pretrained model vectorizer

Fig. 9: Visualizing suicide intensities within Depression/Suicide class

For the word2vec vector embedding scenario supportive and indicator categories results are almost similar in depression or suicide both classes. There is slight difference is shown for suicidal ideation and within suicide class, suicidal ideation is slight higher. Except the behavior and attempt category for the rest categories depression and suicide showed almost similar number of samples.

7 Discussion

From the result it is revealed that suicide categories shown within depression and suicide class vividly. Specially suicidal ideation, indicator showed similar patterns. The number of samples within depression and suicide is almost similar for this two categories. Hence, we can infer depressed person comments showed suicidal ideation and suicidal indicating symptoms. Suicidal behavior and attempt showed higher number of samples within the suicide category compared to depression category. All these results seems very logical results. Although from the results mathematical formulas are not derived in this research study since results are susceptible to chosen classifier, chosen dataset, pretrained models vectors or embedding provided to the classifier.

8 Conclusion

Suicidal risk estimation task and classification samples to determine suicidal risk within social websites and blogs, techniques are discussed before. According to suicidal category previous work has been done before. However, to what extent depression level triggers suicidal risk is not yet discussed before. Also it

is difficult to determine since depression and suicide categorical variables are independent factor. There is not any underlying correlation. Several research conducted to segregate which post is suicidal and which one is depression various classifiers are proposed. Extensive work has been done to improve the classification accuracy by adopting most powerful vectorization techniques that uses cutting edge NLP models BERT and its various variants. Research has also been conducted on how much severity label of suicide within a post is studied.
The input format for the above table is as follows:

References

[1] Theyazn HH Aldhyani, Saleh Nagi Alsubari, Ali Saleh Alshebami, Hasan Alkahtani, and Zeyad AT Ahmed. Detecting and analyzing suicidal ideation on social media using deep learning and machine learning models. *International journal of environmental research and public health*, 19(19):12635, 2022.

[2] A Beck and M Kovacs. weisman a. *Handbook ofpsychiatric measures: Beck Scale for Suicide Ideation. American Psychiatric Association*, 2000.

[3] Aaron T Beck, Calvin H Ward, Mock Mendelson, Jeremiah Mock, and John Erbaugh. An inventory for measuring depression. *Archives of general psychiatry*, 4(6):561–571, 1961.

[4] Pete Burnap, Walter Colombo, and Jonathan Scourfield. Machine classification and analysis of suicide-related communication on twitter. In *Proceedings of the 26th ACM conference on hypertext & social media*, pages 75–84, 2015.

[5] Gema Castillo-Sánchez, Gonçalo Marques, Enrique Dorronzoro, Octavio Rivera-Romero, Manuel Franco-Martín, and Isabel De la Torre-Díez. Suicide risk assessment using machine learning and social networks: a scoping review. *Journal of medical systems*, 44(12):205, 2020.

[6] Stevie Chancellor and Munmun De Choudhury. Methods in predictive techniques for mental health status on social media: a critical review. *NPJ digital medicine*, 3(1):43, 2020.

[7] Xuetong Chen, Martin D Sykora, Thomas W Jackson, and Suzanne Elayan. What about mood swings: Identifying depression on twitter with temporal measures of emotions. In *Companion proceedings of the the web conference 2018*, pages 1653–1660, 2018.

[8] Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. Predicting depression via social media. In *Proceedings of the international AAAI conference on web and social media*, volume 7, pages 128–137, 2013.

[9] Manas Gaur, Amanuel Alambo, Joy Prakash Sain, Ugur Kursuncu, Krishnaprasad Thirunarayan, Ramakanth Kavuluru, Amit Sheth, Randy Welton, and Jyotishman Pathak. Knowledge-aware assessment of severity of suicide risk for early intervention. In *The world wide web conference*, pages 514–525, 2019.

[10] Jana M Havigerová, Jiří Haviger, Dalibor Kučera, and Petra Hoffmannová. Text-based detection of the risk of depression. *Frontiers in psychology*, 10:513, 2019.

[11] Keith Hawton, Carolina Casañas i Comabella, Camilla Haw, and Kate Saunders. Risk factors for suicide in individuals with depression: a systematic review. *Journal of affective disorders*, 147(1-3):17–28, 2013.

[12] Lang He, Mingyue Niu, Prayag Tiwari, Pekka Marttinen, Rui Su, Jiewei Jiang, Chenguang Guo, Hongyu Wang, Songtao Ding, Zhongmin Wang, et al. Deep learning for depression recognition with audiovisual cues: A review. *Information Fusion*, 80:56–86, 2022.

[13] Thomas E Joiner Jr, M David Rudd, and M Hasan Rajab. The modified scale for suicidal ideation: Factors of suicidality and their relation to clinical and diagnostic variables. *Journal of abnormal psychology*, 106(2):260, 1997.

[14] Sören Kliem, Anna Lohmann, Thomas Mößle, and Elmar Brähler. German beck scale for suicide ideation (bss): psychometric properties from a representative population survey. *BMC psychiatry*, 17(1):1–8, 2017.

[15] Kurt Kroenke, Robert L Spitzer, and Janet BW Williams. The phq-9: validity of a brief depression severity measure. *Journal of general internal medicine*, 16(9):606–613, 2001.

[16] Kuiliang Li, Xiaoqing Zhan, Lei Ren, Nan Liu, Lei Zhang, Ling Li, Ting Chen, Zhengzhi Feng, and Xi Luo. The association of abuse and depression with suicidal ideation in chinese adolescents: a network analysis. *Frontiers in psychiatry*, 13, 2022.

[17] Yolanda López-Del-Hoyo and Pedro Cerbuna. Exploring the risk of suicide in real time on spanish twitter: Observational study. *JMIR Public Health and Surveillance*, 8(5), 2022.

[18] Sean MacAvaney, Anjali Mittu, Glen Coppersmith, Jeff Leintz, and Philip Resnik. Community-level research on suicidality prediction in a secure environment: Overview of the clpsych 2021 shared task. In *Proceedings of the Seventh Workshop on Computational Linguistics and Clinical Psychology: Improving Access*, pages 70–80, 2021.

[19] Anshu Malhotra and Rajni Jindal. Deep learning techniques for suicide and depression detection from online social media: A scoping review. *Applied Soft Computing*, page 109713, 2022.

[20] Paulo Mann, Aline Paes, and Elton H Matsushima. See and read: detecting depression symptoms in higher education students using multimodal social media data. In *Proceedings of the International AAAI Conference on Web and social media*, volume 14, pages 440–451, 2020.

[21] Laura Martinengo, Louise Van Galen, Elaine Lum, Martin Kowalski, Mythily Subramaniam, and Josip Car. Suicide prevention and depression apps' suicide risk assessment and management: a systematic assessment of adherence to clinical guidelines. *BMC medicine*, 17(1):1–12, 2019.

[22] Alexander McGirr, Johanne Renaud, Monique Seguin, Martin Alda, Chawki Benkelfat, Alain Lesage, and Gustavo Turecki. An examination of dsm-iv depressive symptoms and risk for suicide completion in major depressive disorder: a psychological autopsy study. *Journal of affective disorders*, 97(1-3):203–209, 2007.

[23] Minsu Park, David McDonald, and Meeyoung Cha. Perception differences between the depressed and non-depressed users in twitter. In *Proceedings of the international AAAI conference on web and social media*, volume 7, pages 476–485, 2013.

[24] James W Pennebaker, Martha E Francis, and Roger J Booth. Linguistic inquiry and word count: Liwc 2001. *Mahway: Lawrence Erlbaum Associates*, 71(2001):2001, 2001.

[25] Kelly Posner, Gregory K Brown, Barbara Stanley, David A Brent, Kseniya V Yershova, Maria A Oquendo, Glenn W Currier, Glenn A Melvin, Laurence Greenhill, Sa Shen, et al. The columbia–suicide severity rating scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. *American journal of psychiatry*, 168(12):1266–1277, 2011.

[26] Lenore Sawyer Radloff. The ces-d scale: A self-report depression scale for research in the general population. *Applied psychological measurement*, 1(3):385–401, 1977.

[27] Esteban A Ríssola, Seyed Ali Bahrainian, and Fabio Crestani. A dataset for research on depression in social media. In *Proceedings of the 28th ACM conference on user modeling, adaptation and personalization*, pages 338–342, 2020.

[28] Guangyao Shen, Jia Jia, Liqiang Nie, Fuli Feng, Cunjun Zhang, Tianrui Hu, Tat-Seng Chua, Wenwu Zhu, et al. Depression detection via harvesting social media: A multimodal dictionary learning solution. In *IJCAI*, pages 3838–3844, 2017.

[29] Yanmei Shen, Wenyu Zhang, Bella Siu Man Chan, Yaru Zhang, Fanchao Meng, Elizabeth A Kennon, Hanjing Emily Wu, Xuerong Luo, and Xiangyang Zhang. Detecting risk of suicide attempts among chinese medical college students using a machine learning algorithm. *Journal of affective disorders*, 273:18–23, 2020.

[30] Nisha P Shetty, Balachandra Muniyal, Arshia Anand, Sushant Kumar, and Sushant Prabhu. Predicting depression using deep learning and ensemble algorithms on raw twitter data. *International Journal of Electrical and Computer Engineering*, 10(4):3751, 2020.

[31] Han-Chin Shing, Suraj Nair, Ayah Zirikly, Meir Friedenberg, Hal Daumé III, and Philip Resnik. Expert, crowdsourced, and machine assessment of suicide risk via online postings. In *Proceedings of the fifth workshop on computational linguistics and clinical psychology: from keyboard to clinic*, pages 25–36, 2018.

[32] Om P Singh. Startling suicide statistics in india: Time for urgent action, 2022.

[33] Michael M Tadesse, Hongfei Lin, Bo Xu, and Liang Yang. Detection of depression-related posts in reddit social media forum. *IEEE Access*, 7:44883–44893, 2019.

[34] Julio C Tolentino and Sergio L Schmidt. Dsm-5 criteria and depression severity: implications for clinical practice. *Frontiers in psychiatry*, 9:450, 2018.

[35] M Vuorilehto, HM Valtonen, T Melartin, P Sokero, K Suominen, and ET Isometsä. Method of assessment determines prevalence of suicidal ideation among patients with depression. *European Psychiatry*, 29(6):338–344, 2014.

[36] MS Vuorilehto, Tarja K Melartin, and ET Isometsä. Suicidal behaviour among primary-care patients with depressive disorders. *Psychological Medicine*, 36(2):203–210, 2006.

[37] Xiaofeng Wang, Shuai Chen, Tao Li, Wanting Li, Yejie Zhou, Jie Zheng, Qingcai Chen, Jun Yan, Buzhou Tang, et al. Depression risk prediction for chinese microblogs via deep-learning methods: content analysis. *JMIR medical informatics*, 8(7):e17958, 2020.

[38] Owen Whooley. Diagnostic and statistical manual of mental disorders (dsm). *The Wiley Blackwell Encyclopedia of Health, Illness, Behavior, and Society*, pages 381–384, 2014.

[39] Janet BW Williams. A structured interview guide for the hamilton depression rating scale. *Archives of general psychiatry*, 45(8):742–747, 1988.

[40] Ying Xu, Juan Qi, Yi Yang, and Xiaozhong Wen. The contribution of lifestyle factors to depressive symptoms: A cross-sectional study in chinese college students. *Psychiatry research*, 245:243–249, 2016.

[41] Jiayu Ye, Yanhong Yu, Qingxiang Wang, Wentao Li, Hu Liang, Yunshao Zheng, and Gang Fu. Multi-modal depression detection based on emotional audio and evaluation text. *Journal of Affective Disorders*, 295:904–913, 2021.

[42] Tianlin Zhang, Annika M Schoene, Shaoxiong Ji, and Sophia Ananiadou. Natural language processing applied to mental illness detection: a narrative review. *NPJ digital medicine*, 5(1):46, 2022.

[43] Le Zheng, Oliver Wang, Shiying Hao, Chengyin Ye, Modi Liu, Minjie Xia, Alex N Sabo, Liliana Markovic, Frank Stearns, Laura Kanov, et al. Development of an early-warning system for high-risk patients for suicide attempt using deep learning and electronic health records. *Translational psychiatry*, 10(1):72, 2020.

[44] Hamad Zogan, Imran Razzak, Shoaib Jameel, and Guandong Xu. Depressionnet: learning multi-modalities with user post summarization for depression detection on social media. In *proceedings of the 44th international ACM SIGIR conference on research and development in information retrieval*, pages 133–142, 2021.