ISSN (e): 2250 – 3005 || Volume, 14 || Issue, 3|| May. - June. – 2024 ||
International Journal of Computational Engineering Research (IJCER)

Depression Disorder based Suicide Cerebration detection using Machine learning and Deep Learning model.

Mohammad Shahmidul Islam
School. of Computer Science
University of Glasgow

Tanvir Ahmed
Dept. of Computer Science and Engineering Ahsanullah University of Science and Technology

Md Ashraful Hoque
Dept. of Computer Science and Engineering
South East University

Abstract— *Identifying suicidal ideation early in individuals with depression is crucial for timely medical intervention, potentially saving lives. While recent advancements in Natural Language Processing (NLP) have focused on distinguishing between suicidal and clinically healthy individuals based on linguistic analysis, there is a notable gap in differentiating between depression and suicidal ideation. Suicidal ideation, which involves the consideration of self-harm, is a severe concern in depressive disorders and requires immediate attention.This study advances NLP applications by developing methods to detect signs of suicidal ideation from social media content. Advanced feature extraction techniques are employed to isolate relevant features from an extensive set of feature vectors, utilizing a novel formula. This methodology improves upon the accuracy of traditional machine learning models previously used in similar studies.A variety of established machine learning algorithms, including Linear Support Vector Machine, Stochastic Gradient Descent Classifier, Multinomial Naive Bayes Classifier, Logistic Regression, and deep learning frameworks such as LSTM, BiLSTM, RNN, and BiRNN, are implemented to effectively classify the dataset. The research contributes to the field by proposing an innovative framework that combines insights from suicidality and depression to detect and understand suicidal ideation through linguistic patterns exhibited on social media platforms.*

Index Terms—*machine learning, deep learning, natural lan- guage processing, suicide ideation,depression disorders*

Date of Submission: 04-05-2024 Date of acceptance: 17-05-2024

I. INTRODUCTION

In this emerging world with thousands of developments in our daily life people still face mental breakdowns for various problems which sometimes even result into suicidal thoughts.Suicidal ideation, often known as suicidal thoughts, refers to having thoughts, ideas, or ruminations about taking one's own life. It is not a diagnosis, but rather a symptom of various mental diseases. It can also develop in response to negative experiences in the absence of a mental disor- der.Suicide ideation is linked to depression and other mood disorders; nevertheless, a variety of other mental conditions, life events, and familial events can all raise the likelihood of suicide thinking. A clinical assessment is required that focuses on the patients' mental condition in terms of death and suicide thoughts. A qualitative study of the motivations driving the suicide process, in particular, could serve as a helpful guidance for clinicians working in the tough field of preventing teenagers' suicidal gestures. The majority of instruments used to investigate suicidal motivation are self- report measures, which may result in a lack of sufficiently valid assessment of this area.Suicide rates continue to rise globally, with approximately 703,000 people tragically taking their own lives each year. These numbers are accompanied by a startling reality: for every suicide, there are roughly 20 suicide attempts, affecting a significant number of individuals. The majority of suicide deaths, around 90%, are linked to severe psychological problems, although additional factors can contribute, such as physical, social, or cultural influences. Individuals grappling with mental disorders, notably depression and schizophrenia, are particularly susceptible to suicidal tendencies. Often, those contemplating suicide face a combination of compromised mental well-being and chal- lenging life circumstances. Alarmingly, research indicates that about half of individuals with mental illnesses have attempted suicide at some point. Early detection of suicidal ideation plays a crucial role in providing timely assistance, offering hope in reducing the occurrence of suicide. Depression, a pervasive mental illness, is strongly associated with suicidal thoughts and actions. In fact, up to 90% of suicide victims have underlying mental health issues, with depression being one of the most prevalent. Addressing both depression and suicidal ideation comprehensively is vital for effective suicide prevention, as it allows for timely intervention and support, potentially saving lives.

In previous research paper which are solely based on the idea of suicide detection where there is hardly any existence of analysing depression disorders. Besides it is also very rare where we can see the same text can be considered in order to identifying whether a person is suicidal or mentally depressed.Our paper focuses on developing a co-relation be- tween suicidal ideation and depression detection. Moreover we have also made a categorical analysis on different types of depression disorders.We have also combined different types of techniques to classify the texts. For example the dataset is pre- possessed with different NLP techniques and the performance of classification measurement is done with the help of different machine learning and deep learning models.A literature survey regarding various Suicide Ideation Identification techniques have been conducted in section II. After that, the models related to identification of suicide ideation and depression disorder are explained in section III. Then properties of dataset are discussed in section IV. In section V, methodology of identification of suicide ideation and depression disorder are proposed. The information about hyperparameter setting is presented in section VI. The overall result analysis is explained in section VII. Finally ,the section VIII provides conclusion of this article and discusses future work.

II. RELATED WORKS

Many of works have been done on suicide ideation and de- pression disorder identification separately on different datasets using various models.

On the paper heading ”Neural network identification of high-risk suicide patients”-we discover The current study included 197 patients hospitalised at SMMHC (50 MSSA (medically serious suicide attempts) aged 40.3+9.3 years and 147 nn-MSSA aged 41.6+11.7 years). There are two algo- rithms in use: (1) A minimum-maximum algorithm using two clusters.(2) An N-cluster technique in which the remainder of the clusters are constructed once the first is identified.The first was a combination of BP-fuzzy logic trained using programmer input sets, yielding mean+SD findings of 97.25

+3.4% sensitivity and 69.25 +6.9% specificity. The second was the FALCON, which was trained with 15 suicide-related clinical factors and yielded mean+SD findings of 94% +6.9% sensitivity and 69% +6.9% specificity. [1]

On the paper heading ”Risk Prediction Nomogram for suicidal Attempts in Patients with major depressive disorder”- we discover The data for this investigation were collected cross-sectionally and retrospectively from a clinical database at Beijing Anding Hospital (a tertiary hospital for psychiatric diseases in Beijing, China). A logistic regression-based nomo- gram model. The ability to detect SA incidence trends was assessed to be 74% for marital status, 17% for degree, 4.9% for duration, 5.0% for time of onset, and 78% for number of episodes. [2]

On the paper heading ”Detecting Suicide Ideation from Sina Microblog”-we discover Sina Microblog has received 59999 microblog downloads in total. The final coding assignment was accomplished using a large sample of microblogs (n = 59046), which included 9123 microblogs with suicidal thoughts and 49923 without. Four alternative classifiers were used: Support Vector Machine (SVM), Multinomial Naive Bayes (MultiNB), Logistic Regression (LR), and Multi Layer Perception (MiP). MIP had the best accuracy (92.3%) and F1 (67%). MultiNB has a precision of up to 70% and a recall of 77%. MLP performs better in comparison. [3]

On the paper heading” Deep neural networks detect suicide risk from textual facebook posts ”-we discover This dataset had 83,292 posts created by 1002 Faccbook users (23.25 percent of whom were male) who published at least 10 posts. The average number of postings per user was 82.35 (standard deviation = 106.79), while the average number of words in each post was 31.14 (standard deviation = 66.56).

1. CWE (Contextualised Word Embeddings) Models based on ANNs. The addition of all risk factors in the MTM improved prediction accuracy significantly for general [AUC= 0.746, 95% CI 0.727, 0.765] and high suicide risk [AUC= 0.697, 95% CI 0.690,0.707]. [4]

On the paper heading ” Mental Health Disorder Detec- tion System based on Body-Area”- WMS data was collected from 74 adults at the Hackensack Meridian Health Carrier Clinic using an Empatica E4 SmartWatch and a Samsung Galaxy S4 Smartphone, which included features such as GSR (sympathetic nervous system arousal), IBI (heart rate), ST (Skin Temperature Reading), and 3 axis Accelerometer: Deep Neural Network Model (DNN).MHDeep obtains an average test accuracy of 90.4 %, 87.3 %, and 82.4 % for classifica- tions between healthy and schizoaffective disorder instances, healthy and major depressive disorder instances, and healthy and bipolar disorder instances, respectively, throughout the three data parti-tions. [5]

III. BACKGROUND STUDY

A. Natural Language processing

Tokenizer- A tokenizer is an essential component of natural language processing that divides text into smaller components known as tokens. It allows for efficient language processing by breaking down sentences, paragraphs, or entire documents into meaningful pieces like words, phrases, or characters. Tokenizers aid in activities such as language comprehension, sentiment analysis, and machine translation. They provide a structured representation of textual data and so serve as a crucial building component for many NLP applications.

GloVe Embedding- GloVe (Global Vectors for Word Rep- resentation) embedding is a common unsupervised learning technique that represents words in a high-dimensional space as dense vectors. To capture semantic links between words, it uses global word co-occurrence statistics. GloVe embeddings are excellent at capturing word similarities and have been widely employed in a variety of natural language process- ing tasks such as language modelling, sentiment analysis, and named entity recognition. GloVe embeddings improve the performance of NLP models by encapsulating semantic information, allowing for more accurate language processing and creation.

Bag of Words (BoW)- Bag of Words (BoW) is a straightfor- ward and widely used text representation technique in natural language processing. It ignores grammar and word order, focusing solely on the frequency of words in a document. BoW generates a sparse vector representation in which each dimension represents a distinct word and the value correlates to the frequency of that word. While BoW is computationally efficient, it lacks semantic knowledge and context, making it less successful for tasks that require more in-depth language comprehension.

TF-IDF- The statistical measure TF-IDF (Term Frequency- Inverse Document Frequency) is used in information retrieval and text mining. It assesses the significance of a word in a document in relation to a collection of documents. The TF- IDF algorithm takes into account both the frequency of a word in a document (TF) and its rarity across the full collection (IDF). TF-IDF assists in identifying key terms and improving the accuracy of text analysis, such as document similarity and keyword extraction, by allocating larger weights to rare words that are prevalent in certain documents.

B. Single Machine Learning Algorithm

LinearSVC- LinearSVC, a variant of SVM, is a versatile binary and multi-class classification method. It effectively handles high-dimensional and large datasets, seeking an op- timal hyperplane to separate classes. By employing a hinge loss function and regularization techniques, it determines the decision boundary. LinearSVC excels in handling linearly separable data and finds applications in text classification, image recognition, and other machine learning tasks.

SGD Classifier- The SGD Classifier is a linear classification algorithm that utilizes Stochastic Gradient Descent optimiza- tion. It efficiently handles large-scale and high-dimensional classification tasks. By updating model parameters iteratively with one training example at a time, it is computationally efficient and suitable for online learning. The algorithm of- fers flexibility through various loss functions and regulariza- tion techniques. Commonly used in text classification, image recognition, and sentiment analysis, it provides fast training and prediction times. SGD is a versatile and effective algo- rithm for linear classification problems in Machine Learning.

MultinomialNB- The Multinomial Naive Bayes classifier is a popular algorithm for text categorization. It assumes that words are conditionally independent given the class labels, following a multinomial distribution. It is effective for discrete features like word counts in text data. By evaluating the probability of classes based on word frequencies in training data, it quickly and accurately predicts class labels. Multino- mial Naive Bayes is widely used in spam filtering, sentiment analysis, and document classification, especially for tasks with numerous classes.

Logistic Regression- Logistic regression is a widely used statistical technique for binary classification problems. It pre- dicts the likelihood of an event based on input features. Unlike linear regression, it focuses on discrete outcomes. The logistic function (sigmoid function) transforms the output variable to a probability between 0 and 1. By training the model to minimize the discrepancy between predicted probabilities and actual labels, the coefficients are optimized using techniques like gradient descent. Logistic regression finds applications in various industries and tasks such as sentiment analysis, credit risk assessment, and disease diagnosis. Its simplicity, interpretability, and effectiveness make it important for binary classification.

C. Deep Learning Model

RNN- A recurrent neural network (RNN) is a form of artifi- cial neural network that is designed to process sequential data. It features feedback links that allow data to be retained over time. RNNs are useful in dealing with sequential data, such as natural language processing and speech recognition, since they capture context and connections between sequence items. They excel at tasks such as language modelling, machine translation, and sentiment analysis by sequentially processing data and keeping memory of previous information during computation. **BiRNN**- BiRNN, or Bidirectional Recurrent Neural Net- work, is a Recurrent Neural Network (RNN) extension that processes sequential data in both forward and backward direc- tions at the same time. BiRNNs capture a more comprehensive picture of the input sequence by combining information from both the past and the future context. As a result, they are well-suited for tasks such as sentiment analysis, named entity recognition, and machine translation, where contextual infor- mation from both previous and following words is required for correct prediction or classification.

LSTM- The LSTM (Long Short-Term Memory) architec- ture effectively addresses the vanishing gradient problem and captures long-term dependencies in sequential data. It employs specialised memory cells and gating mechanisms to selectively keep and discard information, allowing the network to recall key prior events while ignoring irrelevant facts. By efficiently modelling sequential dependencies over long periods of time, LSTMs have achieved substantial success in a variety of applications such as language modelling, speech recognition, and machine translation.

BiLSTM- BiLSTM (Bidirectional Long Short-Term Mem- ory) is an LSTM architecture extension that processes se- quential data in both forward and backward directions at the same time. BiLSTM obtains a more thorough knowledge of the input sequence by using information from the past and future context. It performs well in tasks including as sentiment analysis, named entity recognition, and machine translation, where contextual information from both preceding and subsequent words is critical for effective prediction or classification.

D. Confusion Matrix

The performance of a classification model on a set of test data for which the real values are known is frequently described using a confusion matrix.

True Positives (TP):- True positive represents the count of correctly predicted positive instances by a model. It indicates the number of positive samples that were accurately classified as positive, reflecting the successful detection of the target class.

True Negatives(TN):- True negative in a confusion matrix refers to the count of correctly predicted negative instances by a model. It indicates the number of negative samples that were accurately classified as negative, reflecting the successful exclusion of the target class.

False Positives(FP):- False positive in a confusion matrix is the count of negative instances that were incorrectly classified as positive by a model. It represents the instances where the model predicted a positive outcome, but the actual class was negative, indicating a type I error.

False Negatives(FN):- False negative in a confusion matrix is the count of positive instances that were incorrectly classi- fied as negative by a model. It represents the instances where the model predicted a negative outcome, but the actual class was positive, indicating a type II error.

IV. DATASET

A. Source

For this paper , We obtained the dataset for this prediction from Kaggle [**?**]. Dataset was web-scrapped from Reddit fol- lowing various SubReddits.In order to categorise the suicidal ideation the subReddit ”r/SuicideWatch” has been taken.In case of classifying different mental disorders the posts of few subreddits like ”r/depression” , ”r/Anxiety” , ”r/bipolar” , ”r/BPD” has been concerned.Reddit is a social news, content review, and conversation website based in the United States.

B. Description

Our dataset contains following features:
1.serial no
2.text
3.subreddit

4.class(Suicide or non-suicide)

The dataset has been stored in .csv format.It contains around 200000 data where 50.09% are suicide and 49.91% are non-suicide texts.Among them nearly 6500 data are labeled according to different depression disorders by the help of subreddits. We have split the dataset into 40:30:30 ratio and also with K-fold cross validation where K=5 ,in order to evaluate the model results for train, validation and test sets respectively.

V. METHODOLOGY

The goal of this study is to collect the posts of users blog-spot and then analyze the text. After processing the text the situation of suicidal behavior is to be detected.And the category of the depression disorder is to be identified. For this study we have used a number of Natural Language of processing techniques and Machine Learning models and Deep Learning models. We have used feature extraction method before training the models with suitable features.In case of natural language processing techniques Bag of words, TF IDF Vectorizer and GloVe embedding are used.For machine learn- ing LinearSVC, MultinomialNB,SGD Classifier and Logistic

Regression are applied.In case of deep learning we have used RNN, Bi-RNN, LSTM and Bi-LSTM.

A. Pre-Processing of Data

Data preprocessing is a crucial step in Machine Learning to enhance the quality of data. It is a data mining approach that transforms raw data into an understandable and readable format.

Data Cleaning-The dataset contains raw text from blogspot

. In order to clean unnecessary data Tokenization is used where

, this function automatically does 3 things: 1.Splitting the words by space.

2. Filtering out punctuation.
3. Converting the text to lowercase (lower=True).

Vectorizing Text-Vectorizing methodology is applied to map words or phrases from vocabulary to corresponding vector of real numbers which is used to find word prediction and similarities. For this study TF-IDF vectorizer is used to vectorize the cleaned text. We analyze the complete dataset to check for the presence of suicidal thoughts and compare lexical differences. We calculate the frequency of all the unigrams and bigrams in posts that indicate suicide as well as posts that do not.

B. Feature Extraction

GloVe embedding is used for feature extraction in this work to find essential features in data for coding by learning from the coding of the original data set to derive new ones. The Bag of Words (BoW) technique is also employed for feature extraction. These are natural language processing approaches that extract the words (features) used in a phrase and classify them based on frequency of use.

C. Executing Models

After the execution of feature extraction models, differ- ent classification algorithms -Linear Super Vector machine

,Stochastic Gradient Descent Classifier, Multinomial Naive bayes classifier,logistic Regression and Deep Learning models Like LSTM ,Bi-LSTM, RNN, Bi-RNN were applied on the dataset. Scikit-learn and Keras library were used for this purpose.

D. Performance Metrices

The metrics accuracy, precision, recall and F1 score are used to evaluate the model. Accuracy is measured by the formula given below :

1. $$\text{accuracy} = \frac{TP + TN}{(TP + FP + TN + FN)}$$

Precision is measured by the formula given below:

2. $$\text{precision} = \frac{TP}{TP + FP}$$

Recall is measured by the formula given below :

3. $$\text{recall} = \frac{TP}{TP + FN}$$

F1 score measured by the formula given below:

4. $$F1 = \frac{2 * \text{precision} * \text{recall}}{\text{precision} + \text{recall}}$$

E. Work Flow Diagram

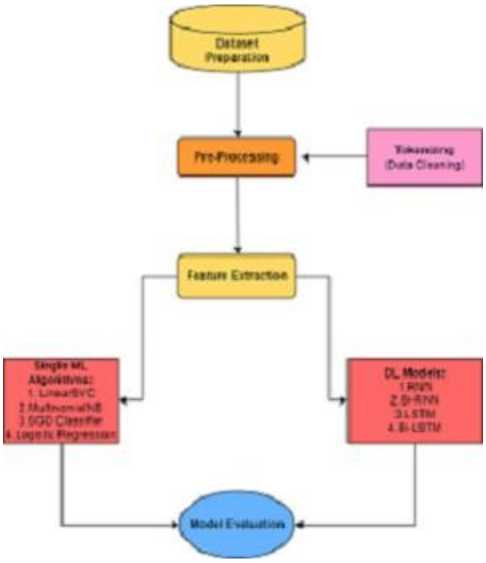

Fig. 1. Proposed Model

VI. EXPERIMENT

We have applied the machine learning models by splitting the dataset in two different ways. We spilt the train , validation and test sets into 40:30:30 ratio , Also with k-Fold cross validation where K=5.The performance metrics of different ML Classifiers are shown here:

Data splitting	Model	Accuracy	Precision	Recall	F1 Score
70:30 RATIO	Linear SVC	0.95	0.955	0.955	0.95
	SGD Classifier	0.93	0.92	0.93	0.93
	MultinomialNB	0.86	0.82	0.855	0.855
5 Fold Cross Validation	Linear SVC	0.93	0.92	0.93	0.92
	SGD Classifier	0.91	0.92	0.93	0.93
	MultinomialNB	0.87	0.82	0.81	0.84

We have run the RNN model for 20 epoch where batch size was 256.The total training ,validation and test data was splitted into 40:30:30 and k fold cross validation where k=5.When training, after 20 epochs the validation accuracy was 92.62%. After testing ,testing accuracy was 92.17%.

We have run the Bi-RNN model for 20 epoch where batch size was 256.The total training ,validation and test data was splitted into 40:30:30 and k fold cross validation where k=5.When training, after 20 epochs the validation accuracy was 93.04%. After testing ,testing accuracy was 93.04%.

We have run the LSTM model for 20 epoch where batch size was 256.The total training ,validation and test data was splitted into 40:30:30 and k fold cross validation where k=5.When training, after 20 epochs the validation accuracy was 92.71%. After testing ,testing accuracy was 92.90%.

We have run the Bi-LSTM model for 20 epoch where batch size was 256.The total training ,validation and test data was splitted into 40:30:30 and k fold cross validation where k=5.When training, after 20 epochs the validation accuracy was 93.26%. After testing ,testing accuracy was 93.26%.

Data splitting	Model	Accuracy
70:30 RATIO	LSTM	0.929
	BiLSTM	0.9317
	RNN	0.9217
	BiRNN	0.9304
5 Fold Cross Validation	LSTM	0.9312
	BiLSTM	0.9323
	RNN	0.9270
	BiRNN	0.9299

We have classified the text dataset on basis of different depression disorders according to the subreddits.The perfor- mance comparison of this labeling is measured by one machine learning model which is logistic regression and one deep learning model which is LSTM.

MODEL	ACCURACY
Logistic Regression	0.87
LSTM	0.866

We have analyzed suicidal rates on different types of depression disorders to measure how likely one often have a suicidal behavior.

DEPRESSION LEVELS	SUICIDAL RATES
Depression	8,5%
Anxiety	20%
Bipolar	25-50%
BPD	60-70%

VII. RESULT ANALYSIS

A. Result Analysis of ML Algorithms

Considering the available data, it is evident that the Linear SVC and SGD Classifier models have exhibited commendable performance in both the 40:30:30 and K-fold data splitting scenarios. These models have attained accuracy rates of 0.95 and 0.93 respectively in the 70:30 data splitting, while achieving accuracy rates of 0.93 and 0.91 in the K-fold data splitting. The precision and recall measures for these models also demonstrate high values, ranging from 0.92 to 0.955. Furthermore, the F1 score, an overall indicator of model performance, remains consistently high, ranging from 0.92 to 0.95. Contrarily, the MultinomialNB model's performance appears comparatively inferior. It achieves an accuracy rate of 0.86 in the 70:30 data splitting and 0.87 in the K-fold data splitting. Similarly, the precision, recall, and F1 score metrics for this model are relatively lower, ranging from 0.81 to 0.855 for precision, 0.81 to 0.855 for recall, and 0.84 to 0.855 for F1 score.In summary, based on the obtained results, the Linear SVC and SGD Classifier models emerge as the superior choices for this classification task due to their higher accuracy, precision, recall, and F1 score values.

B. Result Analysis of DL Models

When examining the data, it is evident that several deep learning models were evaluated using the 70:30 data splitting and K-fold methods. The LSTM model achieved an accuracy of 0.929 in the 40:30:30 data splitting and an accuracy of 0.9312 in the K-fold method. The BiLSTM model demon- strated slightly better performance with an accuracy of 0.9317 in the 70:30 data splitting and an accuracy of 0.9323 in the K-fold method. The RNN model achieved an accuracy of 0.9217 in the 70:30 data splitting and an accuracy of 0.9270 in the K-fold method. Lastly, the BiRNN model achieved an accuracy of 0.9304 in the 70:30 data splitting and an accuracy of 0.9299 in the K-fold method. In summary, based on the results obtained, the BiLSTM model showcased the highest accuracy in both data splitting techniques. It outperformed the other models, including LSTM, RNN, and BiRNN, which also achieved relatively high accuracy rates. These findings suggest that the BiLSTM model holds promise for the given classification task.

C. Result Analysis of Labelling Depression Disorders & defin- ing Suicidal rate

Table 5.1 shows evaluation accuracy on the dataset which has been labeled with the help of different categories of depression. Here the logistic regression model provides the accuracy of 87% and the deep learning architecture of LSTM provides the accuracy of 86.6%. Table 5.2 shows suicidal rate of different levels of depression where the depression category has around 8.5% rate , the anxiety category has around 20% , the bipolar category has 25-50% and Lastly BPD category has 60-70%.

VIII. CONCLUSION

In recent times, there has been an alarming surge in suicidal behaviors, impacting individuals of all age groups. These ac- tions may stem from severe depression disorders or impulsive decisions triggered by distressing situations. The objective of this research is to identify individuals experiencing suicidal thoughts and analyze their behavior, specifically focusing on the connection to depression disorders. By gaining a deeper understanding of this intricate relationship, we aim to develop effective approaches to prevent and discourage individuals from contemplating or engaging in self-harm. The significance of this research lies in its potential to save lives, promote mental well-being, and inspire proactive interventions that can provide hope and support to those in need.

By going through this research we have learned that it is a long term process to identify the suicidal behaviour or tendency within a person. It is a process where the person needs to be thoroughly communicated. So simply measuring a text to identify suicidal ideation is quite difficult. As in this paper we have used the post of Reddit platform , we considered the subreddits as the predefined labels whether it is to identify suicidal ideation or to categorize the depression disorders. Moreover our experiment is limited by a lack of data and annotation bias. One of the most pressing difficulties in present research, where supervised learning techniques are primarily used, is data scarcity. They frequently necessitate hand annotation. However, there is insufficient annotated data to support additional investigation. Another issue is annotation bias, which is produced by manual labeling with established annotation criteria. In other situations, the annotation may cause label bias, resulting in misleading evidence to confirm the authors' suicidal action. Our future goal is to cross check this labelling by the help of different type of suicide or depression rating scales which are popular in the field of clinical psychology and followed by the psychologist world- wide. This will contribute in the sector of medical intervention where it will work as a tool kit for the psychologist to make their task faster and taking necessary step for further mental rehabilitation.

REFERENCES

[1] H. Lin, J. Jia, Q. Guo, Y. Xue, J. Huang, L. Cai, L. Feng. Psychological Stress Detection from Cross-media Microblog Data Using Deep Sparse Neural Network. In proceedings of IEEE International Conference on Multimedia & Expo, 2014.

[2] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen,"MobileNet v2: Inverted residuals and linear bottlenecks," in Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition, 2018, pp. 4510–4520.

[3] MODAI, I., RITSNER, M., KURS, R., MENDEL, S. and PONI-
ZOVSKY, A., 2001, Validation of the computerized suicide.

[4] Desu, V., Komati, N., Lingamaneni, S., & Shaik, F. (2022, May). Suicide and Depression Detection in Social Media Forums. In Smart Intelligent Computing and Applications, Volume 2: Proceedings of Fifth International Conference on Smart Computing and Informatics (SCI 2021) (pp. 263-270). Singapore: Springer Nature Singapore.

[5] Barak-Corren, Y., Castro, V.M., Javitt, S., Hoffnagle, A.G., Dai, Y., Perlis, R.H., ..., Reis, B.Y., 2017. Predicting suicidal behaviour from lon- gitudinal electronic health records. Am. J. Psychiatry 174 (2), 154–162.

[6] M. Zaghloul, M. Salem, and A. Ali-Eldin, "A new framework based on features modeling and ensemble learning to predict query performance," *Plos one*, vol. 16, no. 10, p. e0258439, 2021

[7] T. H. Sheikh, "Text mining and its applications," *Int J Allied Pract Res Rev*, vol. 4, no. 11, pp. 1–8, 2017.

[8] G. Faggioli, T. Formal, S. Marchesin, S. Clinchant, N. Ferro, and B. Piwowarski, "Query Performance Prediction for Neural IR: Are We There Yet?," in *European Conference on Information Retrieval*, Springer, 2023, pp. 232–248

[9] M. Wen, D. K. Vasthimal, A. Lu, T. Wang, and A. Guo, "Building large-scale deep learning system for entity recognition in e-commerce search," in *Proceedings of the 6th IEEE/ACM International Conference on Big Data Computing, Applications and Technologies*, 2019, pp. 149–154

[10] C. Lee and M. Jung, "Predicting movie success from search query using support vector regression method," *International Journal Artificial Intelligence Application*, vol. 7, no. 1, pp. 1–10, 2016.