AI-Powered System Quantifies Suicide Indicators and Identifies Suicide-Related Content in Online Posts

P.Devendar babu[1], D. Pranathi[2], D. Pravalika[2], E. Supraja[2], G. Harika[2]

[1] Assistant Professor, Department of Information Technology, Mallareddy Engineering College for Women, (UGC-Autonomous), Hyderabad, India, devendar.p518@gmail.com.

[2] Student, Department of Information Technology, Mallareddy Engineering College for Women, (UGC-Autonomous), Hyderabad, India.

Abstract

Suicide is a serious public health concern worldwide, and the rise of social media and online platforms has brought new challenges in identifying and preventing suicidal behaviors. In the past, identifying suicide indicators and related content in online posts relied on human moderators or mental health professionals to manually review and categorize content. This manual approach was both labour-intensive and often lacked real-time capabilities, leading to delays in providing support to individuals in distress. Moreover, the scale of online content made it difficult for traditional methods to handle the ever-increasing volume of information. As a result, this project develops AI-powered system stems from the urgency to tackle the growing issue of suicide in the digital age, which can process data at a scale and speed that exceeds human capabilities, enabling it to analyze many posts, recognize patterns, and detect potential suicide indicators in real-time. This proposed AI-powered system's ability to process and analyze large-scale data in real-time allows for early detection and timely intervention, significantly improving the effectiveness of suicide prevention efforts. The main aim is to find a strong co-relation between components in the subsystem and compare the accuracies to build an alarming system. "Better late than never" the victim can be saved by the proposed method and immediate treatment can be started. Further, this AI-powered system holds great promise in revolutionizing the field of mental health care by enabling more proactive and personalized support for individuals at risk of suicide. Through continuous refinement and development, it is hoped that this technology will play a crucial role in saving lives and promoting mental well-being in an increasingly connected world.

1. Introduction

As per the world Health Organization (WHO), suicide is a primary cause of death among individuals between 15-29 years old across the world. 8, 00,000 of people commit suicide every year leading to increase in suicidal ideation. However, an individual person suicide plays an unsocial act that has overwhelming impact towards relations and families. [1] Several suicidal demises are inevitable and very significant to know the behaviour and the way how individual communicate thoughts and depression for inhibiting such deaths. Suicidal avoidance predominantly focuses on monitoring and observation of suicidal efforts and self-harm tendencies. The existence of content related to suicidal ideation plays a major role on the internet for the people seeking for help and offering support through the younger generation. [2] It is observed that, social media data from different blogs and websites (Facebook, Twitter etc.) are used to recognize the affected individuals instantly to offer help. Suicidal behaviour refers to all promising act of self-harm causing death, while suicidal ideation relates to depressive feeling of planning suicide or killing oneself. Although twitter deliver a chance to know the problem of an individual and to provide a potential way for the intervention of both in social level and individual for suicide prevention there exist no better practices using social media [3]. Suicide avoidance by suicidal identification is a best approach to radically reduce suicidal rates. The major experimental application of this work lies in its flexibility to any web-based social network that should be easily adaptable, wherein it tends to be utilized straightforwardly for breaking down text-based tweets posted by its clients and the tweets are flagged if its contents are related to suicidal thoughts. [4] In recent years, many existing studies focused on n-grams, like 3- grams and 5-grams that are used as keywords and phrases as search terms for suicidal prediction. [5] The objective is, to discover opinion, identify sentiment based on tweets posted by the people and classify them for decision making and suicidal prediction by lexicon and machine learning approach and to identify the suicide prediction level of the twitter users based on the twitter.

2. Literature survey

Metzler, et al. [6] proposed detecting potentially harmful and protective suicide-related content on Twitter: machine learning approach. The authors included a majority classifier, an approach based on word frequency (term frequency-inverse document frequency with a linear support vector machine) and 2 state-of-the-art deep learning models (Bidirectional Encoder Representations from Transformers [BERT] and XLNet). The first task classified posts into 6 main content categories, which are particularly relevant for suicide prevention based on previous evidence. These included personal stories of either suicidal ideation and attempts or coping and recovery, calls for action intending to spread either problem awareness or prevention-related information, reporting of suicide cases, and other tweets irrelevant to these 5 categories. Aldhyani, et al. [7] proposed detecting and analyzing suicidal ideation on social media using deep learning and machine learning models. initially, it is essential to develop a machine learning system for automated early detection of suicidal ideation or any abrupt changes in a user's behavior by analyzing his or her posts on social media. The authors propose a methodology based on experimental research for building a suicidal ideation detection system using publicly available Reddit datasets, word-embedding approaches, such as TF-IDF and Word2Vec, for text representation, and hybrid deep learning and machine learning algorithms for classification. A convolutional neural network and Bidirectional long short-term memory (CNN–BiLSTM) model and the machine learning XGBoost model were used to classify social posts as suicidal or non-suicidal using textual and LIWC-22-based features by conducting two experiments. To assess the models' performance, the authors used the standard metrics of accuracy, precision, recall, and F1-scores.

Haque, et al. [8] proposed A comparative analysis on suicidal ideation detection using NLP, machine, and deep learning. Initially, With the proper exploitation of the information in social media, the complicated early symptoms of suicidal ideations can be discovered and hence, it can save many lives. This study offers a comparative analysis of multiple machine learning and deep learning models to identify suicidal thoughts from the social media platform Twitter. The principal purpose of their research is to achieve better model performance than prior research works to recognize early indications with high accuracy and avoid suicide attempts. They applied text pre-processing and feature extraction approaches such as CountVectorizer and word embedding and trained several machine learning and deep learning models for such a goal. Jung, et al. [9] proposed Suicidality detection on social media using metadata and text feature extraction and machine learning. Metadata features were studied in great details to understand their possibility and importance in suicidality detection models. Results showed that posting type (i.e., reply or not) and time-related features such as the month, day of the week, and the time (AM vs. PM) were the most important metadata features in suicidality detection models. Specifically, the probability of a social media post being suicidal is higher if the post is a reply to other users rather than an original tweet. Moreover, tweets created in the afternoon, on Fridays and weekends, and in fall have higher probabilities of being detected as suicidality tweets compared with those created in other times. Sakib, et al. [10] proposed Analysis of Suicidal Tweets from Twitter Using Ensemble Machine Learning Methods. Initially, the main challenge is to prevent suicidal cases and detect a suicidal note from one's status, or tweet which will help to provide proper mental support to that person. The main motive of the proposed analysis is to anticipate whether a person's tweet contains suicidal ideation or not with the help of machine learning. To attain the objectives, the authors have used an accurate ensemble classifier that can identify content on Twitter that may potentially hint towards suicidal activity. In this research, they have also used several sets of word embedding and tweet features, and they have compared their model among twelve classifiers models.

Chandra, et al. [11] proposed Suicide Ideation Detection in Online Social Networks. Social network services allow its users to stay connected globally, help the content makers to grow their business, etc. However, it also causes some possible risks to susceptible users of these media, for instance, the rapid increase of suicidal ideation in the online social networks. It has been found that many at-risk users use social media to express their feelings before taking more drastic step. Hence, timely identification and detection are the most efficient approach for suicidal ideation prevention and subsequently suicidal attempts. The authors used a summarized view of different approaches such as machine learning or deep learning approaches to detect suicidal ideation through online social network data for automated detection, is presented. Mbarek, et al. [12] propose a new method that automatically detects suicidal users through their created profiles in OSNs. Their contribution consists in considering profiles from multiple data-sources and detecting suicidal users based on their available shared content across OSNs. They extract several types of features from the posting content of users to build a complete profile that contribute to high suicidal user prediction. They employ supervised machine learning techniques to distinguish between suicidal and non-suicidal profiles. Their experiments on a dataset, which consists of persons who had died by suicide, demonstrate the feasibility of identifying user profiles from multiple data-sources in revealing suicidal profiles.

3. Proposed System Model

This research focused on suicide prevention and sentiment analysis in online posts using machine learning and deep learning techniques. The primary objective of this work is to develop an AI-powered system that can identify suicide indicators and detect suicide-related content in online posts. This system aims to provide an early warning system for identifying individuals who may be at risk of self-harm or suicidal thoughts.

3.1 Data Preprocessing

Data pre-processing is a process of preparing the raw data and making it suitable for a machine learning model. It is the first and crucial step while creating a machine learning model. When creating a machine learning project, it is not always a case that we come across the clean and formatted data. And while doing any operation with data, it is mandatory to clean it and put in a formatted way. So, for this, we use data pre-processing task. A real-world data generally contains noises, missing values, and maybe in an unusable format which cannot be directly used for machine learning models. Data pre-processing is required tasks for cleaning the data and making it suitable for a machine learning model which also increases the accuracy and efficiency of a machine learning model.

- Getting the dataset
- Importing libraries
- Importing datasets

- Finding Missing Data
- Encoding Categorical Data
- Splitting dataset into training and test set

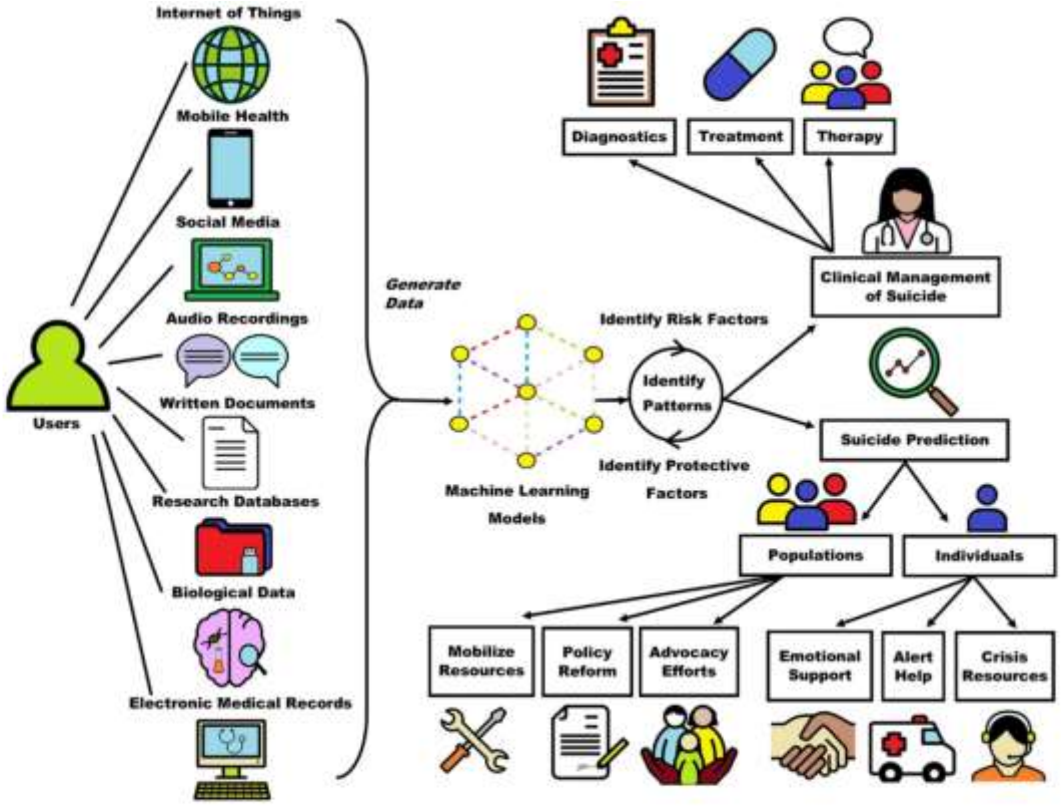

Figure 4.1: Overall design of proposed AI-powered suicide indicators.

Importing Libraries: To perform data preprocessing using Python, we need to import some predefined Python libraries. These libraries are used to perform some specific jobs. There are three specific libraries that we will use for data preprocessing, which are:

Numpy: Numpy Python library is used for including any type of mathematical operation in the code. It is the fundamental package for scientific calculation in Python. It also supports to add large, multidimensional arrays and matrices. So, in Python, we can import it as:

import numpy as nm

Here we have used nm, which is a short name for Numpy, and it will be used in the whole program.

Matplotlib: The second library is matplotlib, which is a Python 2D plotting library, and with this library, we need to import a sub-library pyplot. This library is used to plot any type of charts in Python for the code. It will be imported as below:

import matplotlib.pyplot as mpt

Here we have used mpt as a short name for this library.

Pandas: The last library is the Pandas library, which is one of the most famous Python libraries and used for importing and managing the datasets. It is an open-source data manipulation and analysis library. Here, we have used pd as a short name for this library. Consider the below image:

```
# importing libraries
import numpy as nm
import matplotlib.pyplot as mtp
import pandas as pd
```

Handling Missing data: The next step of data preprocessing is to handle missing data in the datasets. If our dataset contains some missing data, then it may create a huge problem for our machine learning model. Hence it is necessary to handle missing values present in the dataset. There are mainly two ways to handle missing data, which are:

- By deleting the particular row: The first way is used to commonly deal with null values. In this way, we just delete the specific row or column which consists of null values. But this way is not so efficient and removing data may lead to loss of information which will not give the accurate output.
- By calculating the mean: In this way, we will calculate the mean of that column or row which contains any missing value and will put it on the place of missing value. This strategy is useful for the features which have numeric data such as age, salary, year, etc.

Encoding Categorical data: Categorical data is data which has some categories such as, in our dataset; there are two categorical variables, Country, and Purchased. Since machine learning model completely works on mathematics and numbers, but if our dataset would have a categorical variable, then it may create trouble while building the model. So, it is necessary to encode these categorical variables into numbers.

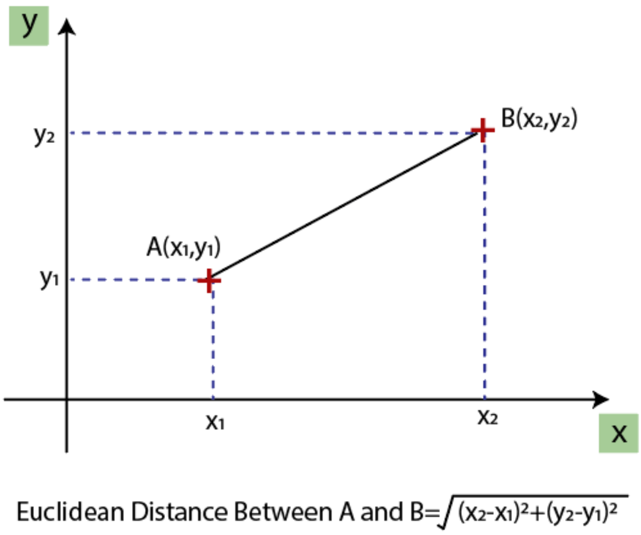

Figure 2: Feature scaling

Feature Scaling: Feature scaling is the final step of data preprocessing in machine learning. It is a technique to standardize the independent variables of the dataset in a specific range. In feature scaling, we put our variables in the same range and in the same scale so that no variable dominates the other variable. A machine learning model is based on Euclidean distance, and if we do not scale the variable, then it will cause some issue in our machine learning model. Euclidean distance is given as: If we compute any two values from age and salary, then salary values will dominate the age values, and it will produce an incorrect result. So, to remove this issue, we need to perform feature scaling for machine learning.

3.2 Splitting the Dataset

In machine learning data preprocessing, we divide our dataset into a training set and test set. This is one of the crucial steps of data preprocessing as by doing this, we can enhance the performance of our machine learning model. Suppose if we have given training to our machine learning model by a dataset and we test it by a completely different dataset. Then, it will create difficulties for our model to understand the correlations between the models. If we train our model very well and its training accuracy is also very high, but we provide a new dataset to it, then it will decrease the performance. So we always try to make a machine learning model which performs well with the training set and also with the test dataset. Here, we can define these datasets as:

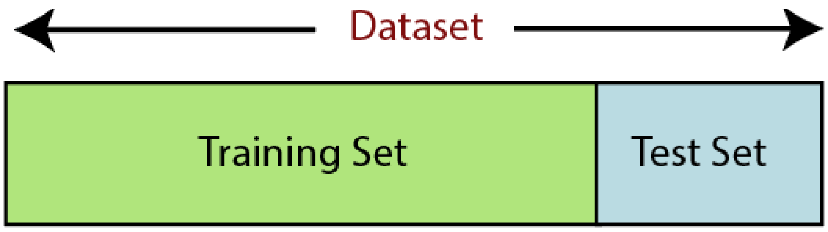

Figure 3: Splitting the dataset.

Training Set: A subset of dataset to train the machine learning model, and we already know the output.

Test set: A subset of dataset to test the machine learning model, and by using the test set, model predicts the output.

For splitting the dataset, we will use the below lines of code:

```
from sklearn.model_selection import train_test_split

x_train, x_test, y_train, y_test= train_test_split(x, y, test_size= 0.2, random_state=0)
```

4. Results description

Figure 4 is an illustration of the graphical user interface (GUI) application. It provides an overview of what the user interface looks like and demonstrates that it is designed for identifying suicide-related content in online posts.

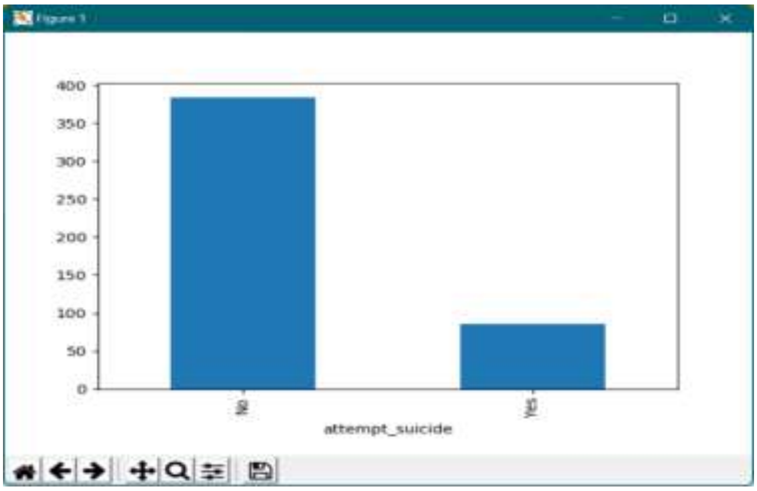

Figure 4: Count histogram plot with number of classes i.e., suicide attempt as label Yes, and No.

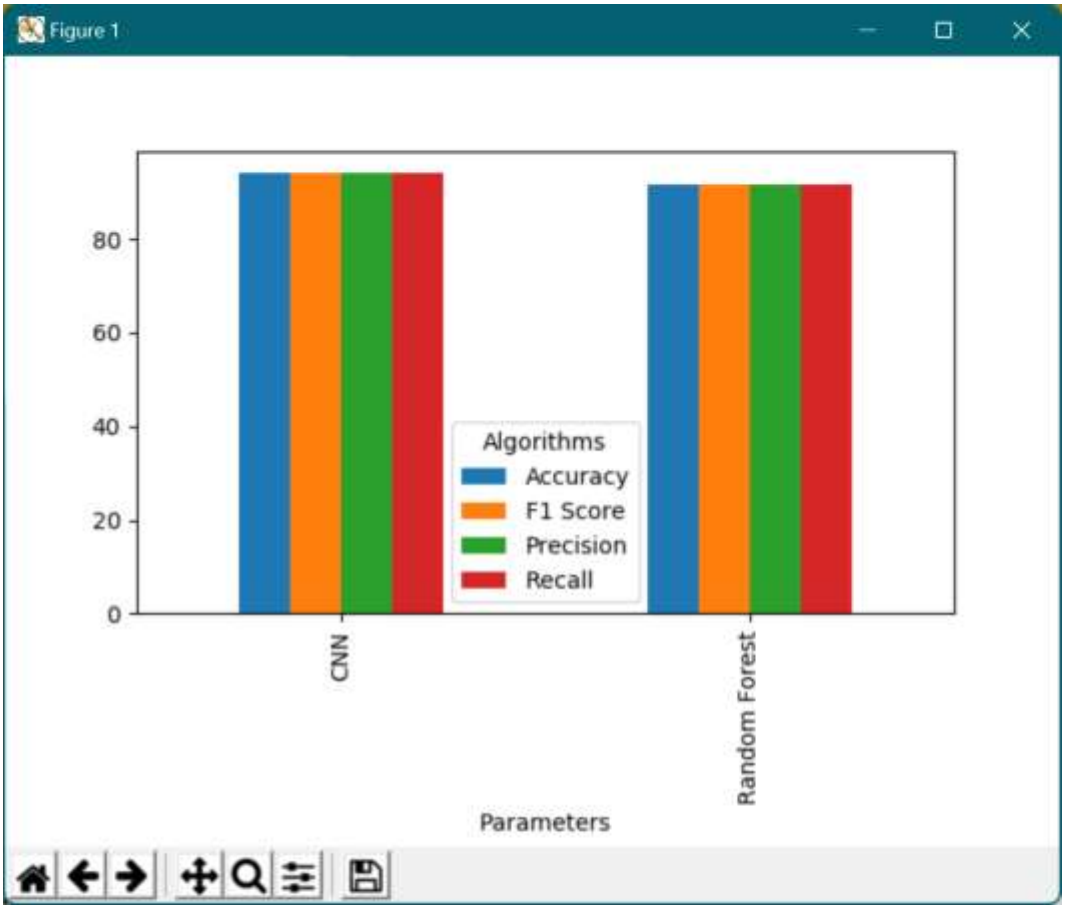

Figure 5: Performance evaluation of obtained quality metrics using random forest, and proposed deep CNN models.

Figure 2 displays a portion of the dataset that has been loaded into the GUI application before any preprocessing steps. It shows the raw or unprocessed data as it appears after uploading.

5.Conclusion

This work planned and assessed a novel way to deal with screen the psychological wellness of a client on Twitter. Working off existing examination, this research attempted to decipher and evaluate suicide cautioning signs in an online setting (client driven and post-driven social highlights). Specifically, this system concentrated on identifying trouble related and suicide-related substance and created two ways to deal with score a tweet: CNN-based methodology and a progressively conventional machine learning content classifier. To detect changes in enthusiastic prosperity, this research considered a Twitter client's action as a surge of perceptions and connected a martingale system to recognize change focuses inside that stream. Our examinations demonstrate that proposed CNN-based approach effectively isolates out tweets displaying trouble related substance and goes about as a ground-breaking contribution to the martingale structure. While the martingale esteems "respond" to changes in online discourse, the change point recognition technique needs enhancement.

References

[1] E. Rajesh Kumar, K.V.S.N. Rama Rao, Soumya Ranjan Nayak & Ramesh Chandra (2020) Suicidal ideation prediction in twitter data using machine learning techniques, Journal of Interdisciplinary Mathematics, 23:1, 117-125, DOI: 10.1080/09720502.2020.1721674.

[2] Abdulsalam, Asma, and Areej Alhothali. "Suicidal Ideation Detection on Social Media: A Review of Machine Learning Methods." arXiv preprint arXiv:2201.10515 (2022).

[3] Roy, A., Nikolitch, K., McGinn, R. et al. A machine learning approach predicts future risk to suicidal ideation from social media data. npj Digit. Med. 3, 78 (2020). https://doi.org/10.1038/s41746-020-0287-6.

[4] Rabani, Syed Tanzeel, Qamar Rayees Khan, and A. M. U. D. Khanday. "Detection of suicidal ideation on Twitter using machine learning & ensemble approaches." Baghdad Science Journal 17.4 (2020): 1328-1328.

[5] S. Ji, S. Pan, X. Li, E. Cambria, G. Long and Z. Huang, "Suicidal Ideation Detection: A Review of Machine Learning Methods and Applications," in IEEE Transactions on Computational Social Systems, vol. 8, no. 1, pp. 214-226, Feb. 2021, doi: 10.1109/TCSS.2020.3021467.

[6] Tadesse MM, Lin H, Xu B, Yang L. Detection of Suicide Ideation in Social Media Forums Using Deep Learning. Algorithms. 2020; 13(1):7. https://doi.org/10.3390/a13010007.

[7] Aldhyani THH, Alsubari SN, Alshebami AS, Alkahtani H, Ahmed ZAT. Detecting and Analyzing Suicidal Ideation on social media Using Deep Learning and Machine Learning Models. International Journal of Environmental Research and Public Health. 2022; 19(19):12635. https://doi.org/10.3390/ijerph191912635.

[8] Swain, D., Khandelwal, A., Joshi, C., Gawas, A., Roy, P., Zad, V. (2021). A Suicide Prediction System Based on Twitter Tweets Using Sentiment Analysis and Machine Learning. In: Swain, D., Pattnaik, P.K., Athawale, T. (eds) Machine Learning and Information Processing. Advances in Intelligent Systems and Computing, vol 1311. Springer, Singapore. https://doi.org/10.1007/978-981-33-4859-2_5.

[9] de Carvalho, V.F., Giacon, B., Nascimento, C., Nogueira, B.M. (2020). Machine Learning for Suicidal Ideation Identification on Twitter for the Portuguese Language. In: Cerri, R., Prati, R.C. (eds) Intelligent Systems. BRACIS 2020. Lecture Notes in Computer Science (), vol 12319. Springer, Cham. https://doi.org/10.1007/978-3-030-61377-8_37.

[10] A. Malhotra, R. Jindal, “Deep learning techniques for suicide and depression detection from online social media: A scoping review”, Applied Soft Computing, Vol. 130, 2022, 109713, ISSN 1568-4946, https://doi.org/10.1016/j.asoc.2022.109713.

[11] T. H. Sakib, M. Ishak, F. F. Jhumu and M. A. Ali, "Analysis of Suicidal Tweets from Twitter Using Ensemble Machine Learning Methods," 2021 International Conference on Automation, Control and Mechatronics for Industry 4.0 (ACMI), 2021, pp. 1-7, doi: 10.1109/ACMI53878.2021.9528252.

[12] Chandra, S., Bhattacharya, S., Banerjee(Ghosh), A., Kundu, S. (2021). Suicide Ideation Detection in Online Social Networks: A Comparative Review. In: Mandal, J.K., Mukhopadhyay, S., Unal, A., Sen, S.K. (eds) Proceedings of International Conference on Innovations in Software Architecture and Computational Systems. Studies in Autonomic, Data-driven and Industrial Computing. Springer, Singapore. https://doi.org/10.1007/978-981-16-4301-9_12.