

# Predicting Suicidal Behaviour Analysis Using Deep Learning Techniques

DR. N. SUGUNA<sup>1</sup>, DR. G. INDIRANI<sup>2</sup>

<sup>1, 2</sup> Associate Professor, Department of CSE, Government College of Engineering, Sengipatti, Thanjavur

**Abstract—** In 2019 survey, according to WHO, every year close to 8,00,000 people who attempt suicide. Suicide remains the second leading cause of death among a young generation with an overall suicide rate of 10.5 per 100,000 people. It is predicted that by 2020, the death rate will increase to one every 20. Almost 69% of the suicides occur in low and middle-income countries.

*It mainly aims to predict the suicidal behavior based on posts in social media websites like reddit, tweeter etc.,. The early detection of suicide through Deep learning [3] and classification approaches applied to Reddit social media. For such purpose, we employ an Bi-LSTM-CNN combined neural network to evaluate and compare to other classification models. Our experimental results shows the combined neural network with word embedding techniques can achieve the best result. Additionally, our result support and strengthen the ability of deep learning architectures to predict the suicidal behaviour.*

**Indexed Terms--** Suicidal ideation, Mental disorder

## I. INTRODUCTION

Suicide is a serious problem in today's world. To save people's lives, early diagnosis and avoidance of suicide attempts should be discussed. Suicidal ideation, also known as suicidal impulses, refers to people's plans to commit suicide. Suicidal thoughts can range from brief to lengthy, from systematic preparation to role playing to failed attempts. In recent years, social media has evolved as a strong "window" into the mental health and well-being of its consumers, who are mostly young people. According to the National Crimes Record Bureau (NCRB), more than 1.39 lakh Indians died by suicide in the year 2019, with 67 percent of them being young adults. As a result, a suicide prediction model is urgently needed to mitigate

the significant negative consequences of suicidal ideators.

Suicidal behaviour has been identified using a variety of suicidal prediction methods, including traditional learning and deep learning-based models [12]. Existing deep learning models have outperformed traditional ones when given enough checked posts on various incidents. This is attributable to their superior ability to retrieve features. They are, however, still unable to deal with the peculiar problem of suicide prediction. The validated posts are in text format due to the lack of word embedding techniques, resulting in unsatisfactory output of existing models. As a result, the aim of this work is to create an efficient model that can replace current models that perform poorly.

The first step in removing unsatisfactory aspects of an existing structure is to recognize them. The posts can be seen in text format here. As a result, since the post could not be fed into the algorithm, for that we proposed a model based on word embedding techniques, which will enable the text to be quickly fed into the model. second Because of the model, there is a lack of accuracy. This necessitates the proposed measuring mechanism being able to capture improvements in precision parameters and reliably include high-accuracy measurements. As a result of the lack of accuracy in this situation, we have a problem to solve.

We proposed a hybrid deep learning model that incorporates the Bi LSTM and CNN to solve the above challenges. The deep learning concept will be used in our proposed framework to predict suicidal behaviour. The large number of people and social media posts will be recorded as a dataset, which will be used to train a hybrid deep learning neural network to predict suicidal behaviour. The training of a deep learning model begins with vectorized data as input. As word embedding techniques are used for vectorization, the

model will continue to learn on its own and achieve the best possible outcome, which will improve in accuracy over time.

To our knowledge, we are the first to propose suicidal behaviour analysis prediction using a deep learning model based on a hybrid technique. We suggested a word embedding technique to predict suicidal behaviour.

The proposed word embedding strategies, such as glove and word2vector, translate text into vector format, which is then fed into our proposed combined deep learning model.

CNN and Bi LSTM are combined in our suggested hybrid model. This model is a general system for detecting suicidal behaviour.

The results show that the proposed hybrid model outperforms the conventional model in predicting suicidal behaviour. The model in this paper is focused on online and offline posts related to suicide ideation accessed from the Reddit API.

The remainder of the article is laid out as follows: Section 2 summaries the relevant literature analysis, Section 3 introduces the specifics of the proposed system, Section 4 presents experimental findings, and Section 5 concludes the report.

## II. RELATED WORKS

Here primarily work concentrates on the following two topics: traditional methods and hybrid neural networks.

There are many tasks related to suicidal ideation prediction, such as clinical methods i.e., patient-clinician interaction and medical signal sensing; textual content analysis, such as lexicon-based filtering and word cloud visualization; feature engineering, including tabular, textual and affective features; and deep learning-based representation learning, such as CNN and LSTM-based text encoders [13].

Another Research focuses on identifying suicidal ideation and different kinds of mental disorders for

early warning. Specifically, both user-level and post-level detection are considered to improve risk identification performance, An attentive RNN model is proposed with text representation and two sets of risk indicators encoded, I.e., lexicon-based sentimental state and latent topics within posts.

In the existing study, the algorithm is used to score online and offline tweets relating to suicidal ideation that were obtained using the Twitter API. They separated Emoticons and Synonyms features and used the n-grams model, which is a combination of Unigram, Bigram, and Trigram with half breed term comparison. Using machine learning algorithms, this model uses casual points to predict the sincerity of the messages.

Medical approaches, such as questionnaire-based methods, take longer and require more manpower. The most popular machine learning algorithms were used, but their accuracy was poor when compared to other techniques. Deep learning algorithms drew the attention of a small number of researchers. The impact of a hybrid model based on a deep learning algorithm is very limited. To address these issues, introduced a hybrid deep learning model is introduced with a word embedding textual function extractor. The performance of our proposed model is also compared to other current models. CNN has the advantage of attribute extraction, which is ideal for complex processing, and is the most commonly used deep learning model for text characterization tasks. Bi-LSTM can store information from past and future information sources for an extended period of time, making them more suitable for modeling sequential data. The integration of CNN and Bi- LSTM's functionality is an outstanding solution for suicidal ideation prediction.

## III. METHODOLOGIES

In this section, first how to integrate the three components of the proposed model is described: the word embedding techniques, the suicidal ideation predictor, and the comparative evaluation. This section also has a detailed algorithm flow.

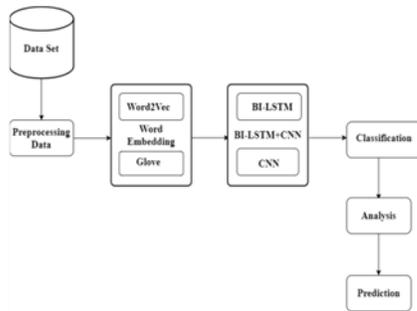


Fig 1. Various Components of the proposed model

### 3.1 Model Overview

Our model aims to learn transferable and discriminable function representations for predicting suicidal ideation. In order to achieve this, the proposed hybrid model integrates three major components: the word embedding techniques, the suicidal ideation predictor, and the comparative evaluation. First and foremost, since social media messages often contain information in several formats, (e.g., textual post and special characters), the word embedding method incorporates both textual and special characters before converting them into vectors to accommodate various types of inputs. The textual and special character feature representations are combined to form the final hybrid model feature after they have been learned. Both of them of the suicidal ideation predictor and the word embedding techniques are built on the pinnacle of the hybrid paradigm.

### 3.2 Word embedding textual feature Extractor

In natural language processing, word embedding refers to a set of language modeling and function learning techniques. Words from the language appear to translate onto a specific vector space of real numbers in a low-dimensional space by using word embedding techniques [16]. The techniques used are Word2Vector and Glove.

#### 3.2.1 Word2Vector:

Shallow neural network architecture is a simple feed forward neural network with one hidden layer. There are two approaches: • Skip Gram • CBOW (Continuous Back Of Words), Skip Gram Learn the Neighbor word based on the Focus word. CBOW Learn the Focus word based on the Neighbor word.

Example: Sentence = “Apple is a sweet fruit”

Context window = 2 , Focus Word = ‘a’

CBOW => Input = [Apple, is, sweet, fruit]  
Output = [a]

Skip gram => Input = [a] Output = [Apple, is, sweet, fruit]

#### 3.2.2 Glove

Using global data, determine the association between terms. Glove used Co-occurrence with global statistics. Matrix of Co-Occurrence: How many times do those word pairs appear together?

Example: “I play cricket, I love cricket , I love

	Play	Love	Football	I	Cricket
Play	0.0	0.0	0.0	1.0	1.0
Love	0.0	0.0	1.0	2.0	1.0
Football	0.0	1.0	0.0	0.0	0.0
I	1.0	2.0	0.0	0.0	0.0
Cricket	0.0	1.0	0.0	0.0	0.0

Example of co-occurrence matrix

### 3.3 Deep learning models

It's a form of machine learning inspired by the human brain's structure. This system is known as a Neural Network in DL (made up of layers of neurons). Between the input and output layers, there are other layers. Each layer reaches a decision, and the final layer makes the final decision.

#### 3.3.1 CNN

It's a kind of feed forward neural network. Go better for photographs or data that looks like a picture. It is similar to the basic neural network. CNN also have learnable parameter like neural network i.e, weights, biases etc. Convolution, Normalization, Pooling, Fully connected NN/any classifiers are the building blocks of CNN.

Convolution: Purpose is to extract features from the input

Input (signal) → image (always square image) [n\*n]  
Filter (kernel) → Detector (it has a pattern & move over to the i/p) => (mostly use odd no of filters) [f\*f]  
Output → convolved feature [n-f+1]

Pooling (optional): Generally, pooling means elements grouped by similarities. Resize the matrix size.

Input (signal) → Output of the convolution operation  
 Pooling window size [f\*f]  
 Output → resized matrix [n-f+1]

### 3.3.2 Bi-LSTM

The LSTM memory network is a recurrent neural network that can learn order dependency in a sequence problem. The LSTM cell will process data in a sequential manner while keeping it secret. The LSTM[10] predicts based on past memories. As a result, we travel to Bi-LSTM. It is a hybrid of Long Short-Term Memory (LSTM) and Bi-directional Recurrent Networks (BiRNN), a type of Recurrent Neural Network (RNN) used to process sequences and time series data. The advantages of LSTM in the form of storage in cell memory and Bi-RNN with access information from context before and after make Bi-LSTM. It causes the Bi-LSTM to have the advantage of LSTM with feedback for the next layer. However, on the other hand, Bi-LSTM can also handle data with dependence on long range. The forward function of Bi-LSTM with inputs of L units and H as the number of hidden units is calculated by the Equation 1 and 2:

$$a_h^t = \sum_{l=1}^L x_l^t w_{lh} + \sum_{h'=1, t>0}^H b_{h'}^{t-1} w_{h'h} \quad (1)$$

$$b_h^t = \sigma_h(a_h^t) \quad (2)$$

where {x<sup>t</sup>} is a sequence input, a<sub>h</sub><sup>t</sup> is the network input to LSTM of unit h at time t, the activation function of h at time t is symbolized by σ<sub>h</sub><sup>t</sup>. w<sub>lh</sub> is the weight of the input l towards h. w<sub>h'h</sub> is the weight of the hidden unit h towards the hidden unit h'. σ<sub>h</sub> is an activation function of the hidden unit of h. While the backward calculation of Bi-LSTM described by the Equation of 3 and 4.

$$\frac{\delta_{w_{hk}}}{\delta_{w_{hk}}} = \sum_{t=1}^T \frac{\delta_{o_t}}{\delta_{a_h^t}} b_h^t \quad (3)$$

$$\frac{\delta_{a_h^t}}{\delta_{a_h^t}} = \sigma_h' \left( a_h^t \left( \sum_{k=1}^K \frac{\delta_{o_t}}{\delta_{a_h^t}} w_{hk} + \sum_{(h'=1, t>0)}^H \frac{\delta_{o_{t+1}}}{\delta_{a_h^t}} w_{h'h} \right) \right) \quad (4)$$

### 3.3.3 Loss function

When working in a Machine Learning or a Deep

Learning Problem, loss/cost functions are used to optimize the model during training. The objective is almost always to minimize the loss function. The lower the loss the better the model. Cross-Entropy loss is a most important cost function. It is used to optimize classification models. Cross-Entropy Loss Function: Also known as loss, log loss, or logistic loss. Each forecast class probability is compared to the real class desired outcome, which is either 0 or 1, and a score/loss is computed that penalises the probability depending on how much it deviates from the actual expected result. The penalty is logarithmic, yielding a large score for large differences approaching 1 and a minimal score for small differences approaching 0. When changing model weights during testing, cross-entropy loss is used. The aim is to reduce the loss as much as possible; the less the loss, the better the model. The cross-entropy loss of a perfect model is zero.

$$L_{CE} = - \sum_{i=1}^n t_i \log(p_i) \quad \text{for } n \text{ classes} \quad (5)$$

where t<sub>i</sub> is the truth label and p<sub>i</sub> is the Softmax probability for the i<sup>th</sup> class.

## IV. EXPERIMENTS

In this part, the social media datasets used in the studies are introduced then cutting-edge suicidal ideation prediction methods are presented and finally the success of the proposed model is reviewed.

### 4.1 Datasets

To equally test the efficiency of the proposed model, are conducted tests on social network datasets obtained from Reddit, and dataset information are included accordingly.

#### Reddit Dataset

The reddit dataset comes from zenodo.org which is used to forecast suicidal behaviour. This dataset is divided into three sections namely User, Post, and Label. Here Post as the instruction set and Label as the evaluation set. The Reddit dataset contains text

material as well as additional social background information in the form of messages. In this paper, concentrates on predicting suicidal behavior using textual material. So here posts are deleted that lack textual detail.

Methods	Posts
Ideation	171
Supportive	108
Indicator	99
Behavior	77
Attempt	45

Table 1: Relation between methods And posts

#### 4.2 Experimental settings

Here the distributed Glove representation are used for words to extract textual features. Posts in the Raddit dataset that were not vectors, are converted into vectors to keep the data coherent. Here the Glove model is trained unsupervisedly on the testing dataset with a dimension size of 50.

The textual encoder is made up of Glove and Word2Vecs and has a dimension scale of 50. The suicidal ideation indicator has two entirely linked layers of CNN and Bi-LSTM with 64 and 32 layers, respectively. In our experiment, the vectorized data of 235 samples is used to train our model, and the remaining 157 samples are used as test data.

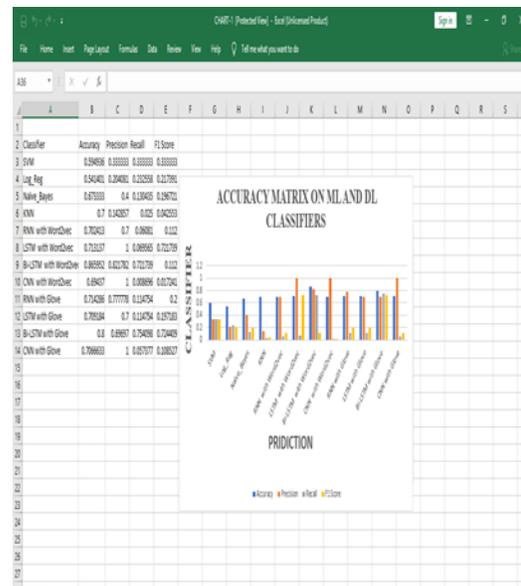
In order to train the whole network, we use a batch size of 128 instances. The model is conditioned for 45 epochs As non-linear activation functions, the Relu and Sigmoid functions are used to avoid over fitting are applied, binary\_cross entropy to our model's weights

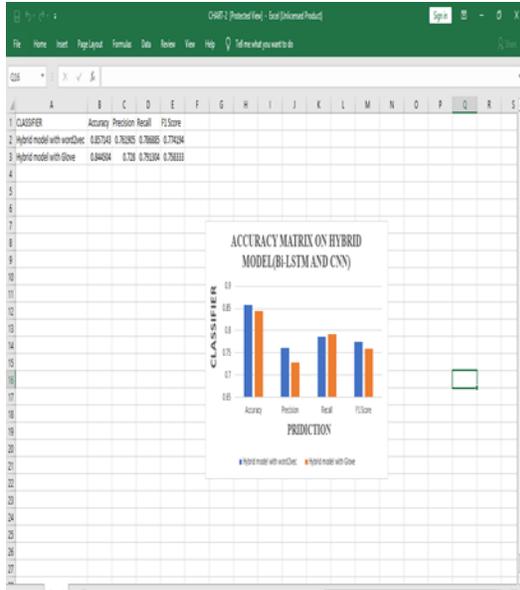
Word Embeddings	Classifier	Accuracy	Precision	Recall	F1 Score
Machine Learning	SVM	0.594936	0.333333	0.333333	0.333333
	Log_Reg	0.541401	0.204081	0.232558	0.217391
	Naive_Bayes	0.673333	0.400000	0.130435	0.196721
	KNN	0.700000	0.142857	0.025000	0.042553
Word 2 Vec	RNN	0.702413	0.700000	0.060810	0.112000
	LSTM	0.713137	1.000000	0.069565	0.721739
	Bi-LSTM	0.865952	0.821782	0.721739	0.112000
	CNN	0.694370	1.000000	0.008696	0.017241
Glove	RNN	0.714286	0.777778	0.114754	0.200000
	LSTM	0.709184	0.700000	0.114754	0.197183
	Bi-LSTM	0.821429	0.696970	0.754098	0.724409
	CNN	0.7066633	1.000000	0.057377	0.108527
Word2Vec	Hybrid (Bi-LSTM & CNN)	0.857143	0.761905	0.786885	0.774194
Glove	Hybrid (Bi-LSTM & CNN)	0.844504	0.728000	0.791304	0.758333

Table 2: Prediction Table

### V. RESULTS AND ANALYSIS

The experiment resulted in a hybrid model of CNN and Bi-LSTM. Machine Learning Approaches can be used to compare performance. The estimation of Hybrid model would be higher than the other approaches





## CONCLUSION

Suicidal ideation and mental health detection, aiming to automatically distinguish early-stage mental disorders and severe suicidality. In this paper, we proposed a novel hybrid deep learning models that incorporates the Bi LSTM and CNN for predicting suicidal behavior. Reddit Dataset to provide a more accurate classification on suicide risk levels and suicidal ideation. The system will be trained using Reddit data set in which we have 500 Reddit user posts with Labels were present in the dataset. Parts of the vectorised data are sent to the CNN while the other part is sent to the Bi-LSTM layer and their outputs after going through the internal representations of both the deep learning methods are concatenated and sent to the dense layer where relu ,sigmoid activation function is applied and the user is classified as either a suicide behavior or not suicide behavior. The model will be trained using the hybrid Neural Network Model while the error function to be adopted will be the Binary Cross Entropy Loss Function. Our experimental results shows the combined neural network with word embedding techniques can achieve the best performance on these four metrics with 0.85713 and 0.7613 improvement in terms of accuracy and precision respectively. Additionally, our result support and strengthen the ability of deep learning architectures to predict the suicidal behaviour.

## REFERENCES

- [1] David Delgado-Gomez, Hilario Blasco-Fontecilla, Federico Sukno, Maria Socorro Ramos-Plasencia, and Enrique Baca-Garcia. Suicide attempters classification: Toward predictive models of suicidal behavior. *Neurocomputing*, 92:3–8, 2012.
- [2] Umashanthi Pavalanathan and Munmun De Choudhury. Identity management and mental health discourse in social media. In *WWW*, pages 315–321. ACM, 2015.
- [3] Hong-Han Shuai, Chih-Ya Shen, De-Nian Yang, Yi-Feng Lan, Wang-Chien Lee, Philip S Yu, and Ming-Syan Chen. Mining online social data for detecting social network mental disorders. In *WWW*, pages 275–285, 2016.
- [4] David Raposo, Adam Santoro, David Barrett, Razvan Pascanu, Timothy Lillicrap, and Peter Battaglia. Discovering objects and their relations from entangled scene representations. *arXiv preprint arXiv:1702.05068*, 2017.
- [5] Verena Venek, Stefan Scherer, Louis-Philippe Morency, John Pestian, et al. Adolescent suicidal risk assessment in clinician-patient interaction. *IEEE Transactions on Affective Computing*, 8(2):204–215, 2017.
- [6] Adrian Benton, Margaret Mitchell, and Dirk Hovy. Multi-task learning for mental health using social media text. In *EACL. ACL*, 2017.
- [7] Glen Copper smith, Ryan Leary, Patrick Crutchely, Alex Fine, "Natural Language Processing Of Social Media As Screening For Suicide Risk", *SAGE*, 20jun 2018.
- [8] AliciaL.Nobies, JeffreyJ.Glenn, Kamran Kowsari, BethanyA.Teachman, LauraE.Barnes, "Identification Of Imminent Suicide Risk Among Young Adults using Text Message", *Proc SIGCHI Conf Hum Factor ComputSys*", April 2018.
- [9] G. Kassen, A. Kudaibergenova, A. Mukasheva, D. Yertargynkyzy, and K. Moldassan, "Behavioral risk factors for suicide among adolescent schoolchildren," *Elementary Education Online*, vol. 19, no. 1, pp. 66–77, 2019.

- [10] Shaoxiong Ji, Guodong Long, Shirui Pan, Tianqing Zhu, Jing Jiang, and Sen Wang. Detecting suicidal ideation with data protection in online communities. In DASFAA, pages 225–229. Springer, Cham, 2019.
- [11] Alejandro Gonzaloz Hevia, Rebeca Cerezo Menendez and Daniel Gayo-Avello, "Analyzing The Use Of Existing Systems For The CLP sych 2019 shared Task", University of Oviedo Spain, CLP sych Shared Task, 6 Jun 2019 .
- [12] Michael Mesfin Tadesse, Hong Fei Lin, Bo Ci, Liang Yang, "Detection Of Suicide Ideation In Social Media Forums Using Deep Learning", MDPI Article, 24 Dec 2019.
- [13] Shaoxiongji, Xie Li, Zi Hung, Erik Cambria, "Suicidal Ideation and Mental Disorder Detection With Attentive Relation Networks", OurWorldInData, 2020.
- [14] Yard E, Radhakrishnan L, Ballesteros MF, et al. Emergency Department Visits for Suspected Suicide Attempts Among Persons Aged 12–25 Years Before and During the COVID-19 Pandemic — United States, January 2019–May 2021.
- [15] Ji, Shaoxiong & Pan, Shirui & Li, Xue & Cambria, Erik & Long, Guodong & Huang, Zi. (2020). Suicidal Ideation Detection: A Review of Machine Learning Methods and Applications. IEEE Transactions on Computational Social Systems. PP.1-13.10.1109/TCSS.2020.3021467.
- [16] Han-Chin Shing, Suraj Nair, Ayah Zirikly, Meir Friedenberg, Hal Daumell and Philip Resnic, "Expert, Crowd Sourced and Machine Assessment of Suicide Risk Via Online Postings", Stanford Center For Population Health Science, Stanford University, Stanford CA, 5 June 2018