Journal of Information Systems Engineering and Management
2025, 10(6s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

A Natural Language Processing Model for Early Detection of Suicidal Ideation in Textual Data

Jocelyn M. Beriña[1*], Thelma D. Palaoag [2]

[1] Dean, Information Technology, Baao Community College, Baao Philippines ,[2]Graduate Program Coordinator, Information Technology, University of the Cordilleras, Baguio City, Philippines
[1]itsmejocelyn@gmail.com,[2] tdpalaoag@uc-bcf.edu.ph
[1]0000-0002-8732-7801,[2]0000-0002-5474-7260

ARTICLE INFO

Received: 09 Oct 2024

Revised: 10 Dec 2024

Accepted: 21 Dec 2024

ABSTRACT

According to the World Health Organization, suicide is one of the top 10 causes of death. An estimated 138 people's lives are significantly impacted for every suicide death, and practically every other statistic pertaining to suicide fatalities is concerning. The widespread usage of social media and the almost universal mobile devices used to access social media networks present new opportunities for preventative intervention as well as new forms of data for studying the behavior of those who (attempt to) take their own life. We show that it is feasible to identify suicidal risk individuals using social media data. Specifically, we propose concepts for an automated system for detecting measurable signals around suicide attempts using natural language processing and machine learning (particularly deep learning) approaches. The goal of this project is to improve the automatic identification and reporting of suicidal posts. It offers a method that examines Twitter as a social media platform to find warning indicators for suicide in people. The previously mentioned approach's main goal is to automatically detect anomalous changes in a user's online behavior. The comprehension and identification of intricate factors of risk or warning indicators that might precede the incident provide difficulties in the prevention of suicide. Many natural language processing (NLP) approaches are used to measure textual variations and pass them via a unique framework that may be utilized broadly to accomplish this objective. Deep learning and machine learning-based categorization algorithms are used to identify suicidal thoughts in the early stages by analyzing tweets on social networking platform Twitter. We first performed data pre-processing for both classifiers, followed by feature extraction, and then machine learning and deep learning classifiers, respectively. We use a CNN-LSTM model for this purpose to assess and contrast it with other classification methods. In comparison to earlier CNN-LSTM systems, the study demonstrates that the CNN-LSTM framework using embedding of words techniques achieves 94% classification accuracy.

Keywords: natural language processing; NLP; suicide ideation; early detection; LSTM, CNN

1. INTRODUCTION:

In the modern field of mental health, it has become imperative to identify and develop methods in preventing suicidal ideation among individuals. Alarmingly, the rate and prevalence of suicidal ideation across all ages continues to increase. In fact, the significant upsurge in suicide rates has become a critical public health issue worldwide [1]. Based on the data published by the Centers for Disease Control and Prevention (2023), in the year 2021 alone, approximately 12.3 million adults in the United States contemplated suicide, with 3.5 million people devising plans for committing suicide, and another 1.7 million actually making an attempt to end their lives. In the Philippines, a study led by the Department of Health and the UPPI (University of the Philippines Population Institute) identified that nearly 1 in every 5 Filipino adolescents from ages 15 to 24, have suicidal ideations[6]. This is an important cause for concern as more and more younger individuals are contemplating ending their lives. It is evident, therefore, that suicidal ideation is one of the most crucial mental health issues that need to be addressed and dealt with.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This research study is focused on developing a Natural Language Processing (NLP) model for the early detection of suicidal ideation using Long Short-Term Memory. Given the complexities of discovering the beginning of suicidal ideation among individuals, this study leverages on the use and power of Long Short-Term Memory (LSTM) networks. As stated by [3], the use of LSTM has revolutionized the realms of both neuro-computing and machine learning. For instance, this same model was utilized to enhance Google's speech recognition, elevate the responses of Amazon's Alexa and significantly improve the machine translations in Google Translate. Quite notably, the emergence in the use of LSTM for various technological purposes signified the departure from the limited performance that has been demonstrated by the traditional recurrent neural networks [3]. Utilizing the Long Short-Term Memory therefore presents a breakthrough and transformative approach in identifying and understanding the early signs of suicidal thoughts. As traditional assessment methods have been proven to be challenging in detecting people's thought patterns, NLP models, particularly the LSTM presents an opportunity to harness the vast amount of data available and decipher patterns that may elude traditional diagnostic approaches in determining suicidal ideation among individuals.

The main goal and purpose of this research study is not only to enhance the accuracy of early detection of suicidal ideation through the use and application of NLP models, but also contribute to a deeper understanding of the underlying factors and dynamics associated with a person's suicidal thoughts. Moreover, this study intends to not only gain a deeper understanding on the complex nature of suicidal ideation but also focus on the practical application of advanced technologies such as machine learning and NLP models for early intervention. By doing so, this study seeks to meaningfully contribute to the continuously evolving landscape of mental health care while helping save people's lives.

2. RELATED WORKS:

A. Studies on Natural Language Processing Models and Machine Learning for Suicide Detection and Prevention

The peer-reviewed research conducted by Arowosegbe and Oyelade (2023) is one of the studies that focused on assessing how natural language processing has been used in the mental health field for preventing suicide cases. Based on the findings of this study, it was validated that NLP has a potential in accurately detecting individuals who have suicidal thoughts even in its early stages. This study's findings also underscored the fact that the integration of machine learning and artificial intelligence presents promising opportunities in significantly improving the risk prediction & suicide prevention frameworks. Moreover, this particularly study emphasized that NLP can be employed for the purpose of creating cost-effective and resource-efficient alternatives to traditional suicide prevention methods. Indeed, this study has provided substantial evidences in supporting the beneficial role of NLP when it comes to identifying individuals with suicidal thoughts, thus providing unique possibilities for effective suicide prevention [4].

Another related study was that of Haque, et al. [5] which focused on employing both machine learning (ML) and deep learning (DL) approaches in detecting the early signs of suicidal ideation in the popular social media platform, Twitter. Similar to this study, the research conducted by Haque et al. [5] emphasizes the importance of timely identification for effective suicide prevention. Its key findings emphasize the successful comparison and analysis of various ML and DL models, with the Bidirectional Long Short-Term Memory (BiLSTM) model outperforming others and achieving an impressive 93.6% accuracy. This study also highlights the significance of utilizing diverse feature extraction techniques and emphasizes the need for comprehensive analyses, considering imbalanced datasets and exploring alternative classifiers for nuanced sentiment understanding. This research advocates for the future development of a real-world web application integrating these classifiers, aiming to provide mental health professionals with a practical tool for identifying online texts that evidently express suicidal thoughts and facilitating intervention. The study also outlined specific plans for further improving the performance of the model and exploring its application in other medical diagnoses using explainable AI and diverse data types [5].

Another study that explored the use of NLP for purposes of mental illness detection was that of Zhang, et al. [9]. In the said research, the authors presented a narrative overview of mental illness detection through NLP over the last ten years, aiming to comprehend the obstacles, patterns, methodologies, and the future direction for such a study. The analysis covered a total of 399 studies taken out of 10,467 records and the findings of the review revealed a consistently rising inclination in NLP research for mental illness detection, with deep learning techniques garnering increased focus and demonstrating superior performance compared to conventional machine learning methods [9].

The study by Zandbiglari, et al. [7] aimed to create an NLP algorithm that utilizes both machine learning and deep learning techniques for the recognition and categorization of suicidal behaviors documented among patients with ARDR (Alzheimer's disease and related dementia). Conventional ML and DL models were employed achieving results closely aligned with human performance, with up to 98% precision and 98% recall in identifying suicidal ideation within the given patient population. These findings indicate that the NLP model effectively replicated human annotation and this lays the groundwork for identifying and categorizing documentation related to suicidal ideation in the ADRD population. Moreover, this study contributes to the advancement of NLP techniques in healthcare, specifically in extracting and classifying clinical concepts, with a focus on suicidal ideation among ADRD patients. This emphasizes the robust capability of the NLP algorithm to accurately recognize and categorize documentation of suicidal behaviors in this population [7].

Similarly, the study by Coppersmith, et al. [8] explored the use of machine learning techniques (deep learning) and natural language processing models in detecting and quantifying signals or hints around suicide ideation among a given set of population. The focus of this study is on identifying the individuals who are at risk of suicide by means of automatically analyzing their use of language on social media. The findings derived from this study represent the foundations of a novel screening system that is widely talked about and yet has not been thoroughly explored and put into practice. As this study discovered, while the machine learning algorithms exhibit sufficiently high accuracy for use in an envisioned screening system, the remaining components of the system are not yet prepared for implementation. This particular study concluded by emphasizing the various innovative ways with which information from these algorithms may be utilized for effective intervention. However, the only point of contention is that which concerns the ethical aspect of using such a technology and its privacy implications to all users of social media platforms [8].

3. METHODOLOGY:

The goal of this experiment is to assess how well the refined model performs in identifying suicidal thoughts in written communications. The binary classification model that distinguishes between suicidal and non-suicidal thoughts is trained using qualitative and representative samples. However, due to a lack of medical experience overseeing the research, multi-classification to identify the causes or features of suicidal thoughts is not the goal of this study.

To achieve accurate binary classifications, it is crucial to guarantee that the positive data, which represents suicidal ideation, and the negative data, which represents non-suicidal ideation, are as similar as possible concerning text length, textual symbol variation (e.g., question marks, commas, and punctuations), and conversation topics during model training. If these factors are ignored during the training of the model, it might learn to distinguish between unrelated characteristics of the data.

In this project, testing and training are done using a qualitative dataset that is as similar to the previously specified attributes as is practical to assess the validity of the model.

A. LSTM: long short-term algorithm

A Long short-term memory (LSTM) could be a particular sort of artificially generated Recurrent Neural Network (RNN) design applied within the deep learning field. The LSTM has input linkages and is implied to maintain a strategic distance from long-term dependence. It can handle person information focuses as well as entirety information sequences. Voice recognition, not segmented data, linked recognition patterns, finding anomalies in network traffic, and detection of intrusions are a few applications where LSTM may be utilized.

LSTM consist of four elements inputs, output, forget gates and cells. The three gates equalize the exchange of data over the insides and the outside of the cell's recollections since the cell keeps up information for variable sums of time.LSTM systems are a supportive tool for categorization, preparing, Making forecasts based on data from time series is difficult because significant events in a temporal arrangement frequently occur after undetermined intervals of time. To illuminate the possible vanishing slope problem during standard RNN preparing, LSTMs were created. Since of its relative lack of concern to gap length, LSTM regularly outperforms RNNs, hidden Markov models (Hmm), and different other successive learning algorithms. The beginning stages of an LSTM's structure of cells are seen in Figure 1.

The uppermost horizontal line in a cell that represents the cell's status. With the use of structures known as gates, the LSTM may selectively remove or add data to the cell state. The structure of the cell is arranged in series. stage 1: The first stage in the LSTM method is to identify which information from the cell state should be discarded and whether any information is unneeded. A sigmoid layer known as the "forget gate layer" makes this determination.

$$f_t = \sigma \left(W_f . [h_{t-1}, x_t] + b_f\right) \qquad \textit{Equation 1}$$

Where h_{t-1} = output from pervious time stamp,

x_t = new input and b_f = bias

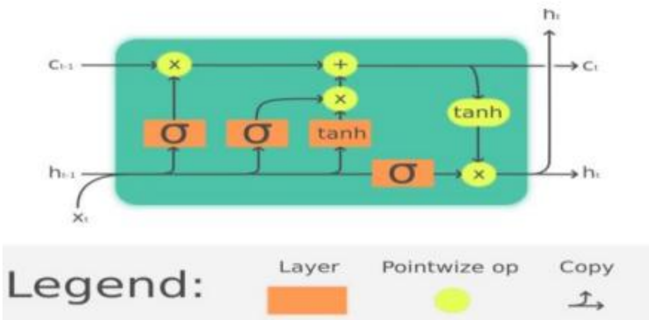

Figure 1: LSTM cell

Step 2: The next step is to choose which new data will be saved in the states of the cell. There are two components to this: The input gates layer, often known as the "sigmoid layer," determines which data should be updated first. Next, a "tanh layer" is used to create a vector of potential new candidate variables that could exist within the state. Combining these two will alter the position in the subsequent step.

$$i_t = \sigma \left(W_i . [h_{t-1}, x_t] + b_i\right) \qquad \textit{Equation 2}$$

$$\widetilde{C_t} = tan\ tan\ h \left(W_c . [h_{t-1}, x_t] + b_c\right) \qquad \textit{Equation 3}$$

Step 3: At this stage, update the former cell state, C(t-1), into the new cell state. It was decided upon earlier in the phases; its implementation is the only remaining task. First: We ignore what we previously decided to forget as time goes on and the old state becomes greater. Second: Raise the cell state after that. The revised candidate values are shown below, scaled in accordance with the proportion that will be applied to the update of each state value.

$$C_t = f_t * C_t + i_t * \widetilde{C_t} \qquad \textit{Equation 4}$$

Step 4: The output step is the final stage. This output, which could be a sifted form, will back our cell state. The sigmoid layer determines which parts of the cell's state will be output, thus run it first. Next, pass the cell state via tanh (which intensifies the specified parameters to be between 1 and -1) and multiply the outcome by the output of the sigmoid gate to ensure that, in a sense, the selected sections are output.

$$O_t = \sigma \left(W_o . [h_{t-1}, x_t] + b_o\right) \qquad \textit{Equation 5}$$

$$h_t = o_t * tan\ tan\ h \left(C_t\right) \qquad \textit{Equation 6}$$

B. Data processing

The dataset is made up of posts from the "depression" and "SuicideWatch" forums on the Reddit website. The posts are collected utilizing the Pushshift API. All posts made on "SuicideWatch" from December 16, 2008, when it was to begin with launched, and January 2, 2021; between January 1, 2009, and January 2, 2021, postings on "depression" were gathered.

Posts through the depression subreddit are categorized as depression, whilst all posts gathered from SuicideWatch are tagged as suicide. We gather postings from r/teenagers that do not include suicide.

There are two classifications in this model "suicide" and "non- suicide". The total data for suicide and non-suicide are 10000 as shown in figure 2.

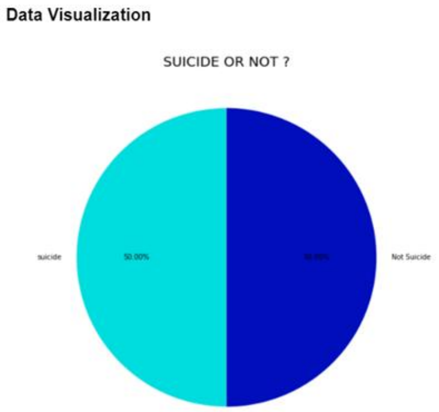

Figure 2 data visualization

Lowercase text is filtered at the preprocessing step by eliminating any items other than the alphabet, numerals, symbols, emoticons, and special characters. Next, the Natural Language Toolkit (NLTK) library is used to perform the stop word removal procedure. The NLTK Stemmer is then used to perform stemming to convert word categories into basic words.

C. *The stop-words*

A stop-words list is a collection of often used, unnecessary words that have little to no linguistics significance for text categorization. In order to lessen the quantity of irrelevant material in the content and increase the number of relevant details, we used the stopwords corpus from NLTK as shown in figure 3.. We also eliminated terms from the texts that were used rarely.

```
def clean_text(text):
    text_length=[]
    cleaned_text=[]
    for sent in tqdm(text):
        sent=sent.lower()
        sent=nfx.remove_special_characters(sent)
        sent=nfx.remove_stopwords(sent)
        text_length.append(len(sent.split()))
        cleaned_text.append(sent)
    return cleaned_text,text_length
```

Figure 3 Stop word

D. Tokenization

Feature extraction is a machine-learning technique that reduces dimensionality by mapping high- Dimensional information is organized into a collection of feature sets of low dimensions. The performance of machine learning models is enhanced while computing complexity is decreased by the extraction of important and relevant features. Thus, we will convert text into a vector (or series) of alternatives by applying highlight extraction algorithms. Word embedding and word count vectorizer are two of the most widely used feature extraction techniques that are frequently applied in ML and NLP for text categorization.

E. Word Embedding

Words are mapped into vectors using word embeddings. It can learn from a vast amount of data and comprehend the meaning behind a word in a document. When text_to_sequence is called, the words in a phrase are replaced with their corresponding numbers.

F. The padding

A pad sequence comprising the text sequence array with a 32-bit integer is appended to the converted sequence. Equivalent lengths of inputs are required by the model. To do this, invoke the pad_sequence function with a 50-length parameter as shown in figure 4.

```
train_text_seq=tokenizer.texts_to_sequences(cleaned_train_text)
train_text_pad=pad_sequences(train_text_seq,maxlen=50)

test_text_seq=tokenizer.texts_to_sequences(cleaned_test_text)
test_text_pad=pad_sequences(test_text_seq,maxlen=50)
```

Figure 4 padding

G. *Model Training and Tuning*

Create a binary - sequential model, with a sequence length of 40, Long-Short-Term-Memory, Dense layers output to 1 final layer.

In order to differentiate suicidal posts from non-suicidal posts using a DL neural network approach, each text sequence in the database has to have an equal real-value vector. This work was completed by using the post-padding sequence approach as shown in figure 5.

```
model=Sequential()
model.add(Input(shape=(40,)))
model.add(Embedding(v+1,300,weights=[embedding_matrix],trainable=False))
model.add(LSTM(20,return_sequences=True))
model.add(GlobalMaxPooling1D())
model.add(Dense(256,activation='relu'))
model.add(Dense(1,activation='sigmoid'))
model.compile(optimizer=keras.optimizers.SGD(0.1,momentum=0.09),loss='binary_crossentropy',metrics=['accuracy'])
```

```
r=model.fit(train_text_pad,train_output,validation_data=(test_text_pad,test_output),
          epochs=20,batch_size=256,callbacks=[early_stop,reducelr])
```

Figure 5 Model training

The model fitting or training will go thru 20 epochs, with callbacks to stop the training when a certain constraint is met, like stop when the model isn't improving.

4. RESULTS AND DISCUSSION

A. *Data Analysis Results*

The entire pre-processed text dataset to determine whether suicidal thoughts were common and to determine whether there were any differences in the vocabulary. We calculated the frequency of every unigram in texts that were suicidal as well as those that weren't. Using Python's WordCloud visualization tool, the top 200 unigrams from every group were selected in order to examine their characteristics and any connection to suicidal ideation. The top 200 unigrams from the dataset are shown in Figure 6 as a WordCloud representation split into two categories: a language that is both suicidal and non-suicide.

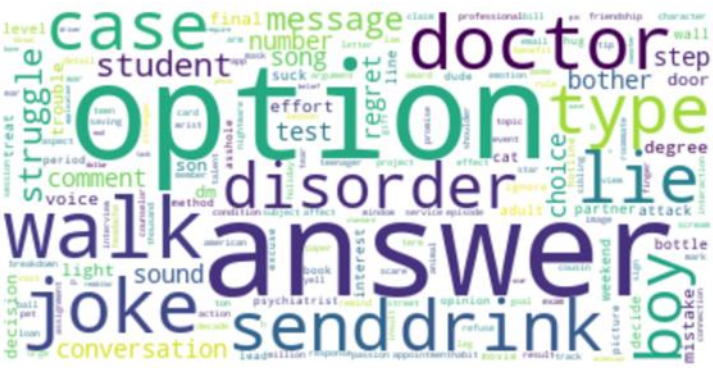

Figure 6 Word cloud

Suicidal intent texts contain terms like "regret," "shit," "pain," "hate," "option," "struggle," "I'm tired," and "session," "lie," as well as negation expressions like "test," "nothing," or "never," as the WordCloud of the suicidal class demonstrates. We then discovered that phrases like "death," "drink," "kill," "disorder," and "scream" that have a deathly effect also indicate the user's suicide thoughts. In contrast to the suicide texts, the unigrams evaluated in the non-suicidal messages primarily feature expressions of positive moments, upbeat attitudes, and sentiments ("want fun", "I'm happy", "beautiful feel", "laugh loud", "friendship"). People are also more inclined to engage in social activities and make an effort to maintain an optimistic mindset ("get better" or "job").

B. Validation and Testing

The trained model was saved into a tensorflow lite file (.tflite) and was loaded to a flask API, which is the back-end of the mobile application, which is the main job is receive text and identify the category whether it's suicidal or not-suicidal. Figure 7 depicts the code for testing the model.

```python
import tensorflow as tf

# Load the saved model
model.save("saved_model")
model = tf.keras.models.load_model("saved_model")

# Convert the model to TFLite format
converter = tf.lite.TFLiteConverter.from_keras_model(model)
tflite_model = converter.convert()

# Save the TFLite model to a file
with open("saved_model/model.tflite", "wb") as f:
    f.write(tflite_model)
```

Figure 7 testing

The API will return a PREDICTION value from (Not Suicidal) 0 to (Suicidal) 0.9xxx.

The researcher will set N which is the threshold to flag a certain percentage a suicidal:

PREDICTION >= N : Suicidal

PREDICTION < N : Not Suicidal

C. Evaluation Metrics

Several assessment criteria, such as accuracy (Acc) Equation 7, precision (P) Equation 8, recall (R) Equation 9, F1-score (F1) Equation 10, and testing duration, were employed to assess our suggested technique. It is predicated on a confusion matrix including details about each test sample's predicted outcome. The right rating is accuracy. The ratio of all cases accurately classified in the positive class to all cases classified in the positive class is known as precision. Stated otherwise, the proportion of instances categorized as positive is accurate. The ratio of all cases successfully categorized into the positive class to the total number of real positive class items is known as the recall. Stated differently, it indicates the proportion of affirmative cases that were accurately identified. The F1 score increases with the proximity of both values. False negative predictions (FN), false-positive predictions (FP), true negative predictions (TN), and true positive predictions (TP) are the rating scales. The accuracy determined as follows is the simple rating evaluation score:

$$Acc = \frac{TP+TN}{TP+TN+FP+FN} \quad Equation\ 7$$

$$Recall = \frac{TP}{TP+FN} \quad Equation\ 8$$

$$Precision = \frac{TP}{TP+FP} \quad Equation\ 9$$

$$F1 = 2\,\frac{Precision\ .Recall}{Precision+Recall} \quad Equation\ 10$$

Following is the result generated for Performance of Classifier.:

TESTING DATA CLASSIFICATION REPORT

	precision	recall	f1-score	support
non-suicide	0.92	0.94	0.93	23209
suicide	0.94	0.92	0.93	23206
accuracy			0.93	46415
macro avg	0.93	0.93	0.93	46415
weighted avg	0.93	0.93	0.93	46415

TRAINING DATA CLASSIFICATION REPORT

	precision	recall	f1-score	support
non-suicide	0.94	0.94	0.94	92828
suicide	0.94	0.94	0.94	92831
accuracy			0.94	185659
macro avg	0.94	0.94	0.94	185659
weighted avg	0.94	0.94	0.94	185659

Figure 8 Classification performance

D. Results

Following figure 9 shows the output result of the classifier. When the user gives inputs like "life is hard" or "I want to die" the classifier is returning the value with predication > 5 which indicated "Suicidal" and when user is giving input " Hello world" the predication value < 5 which means non-suicidal as shown in figure 9.

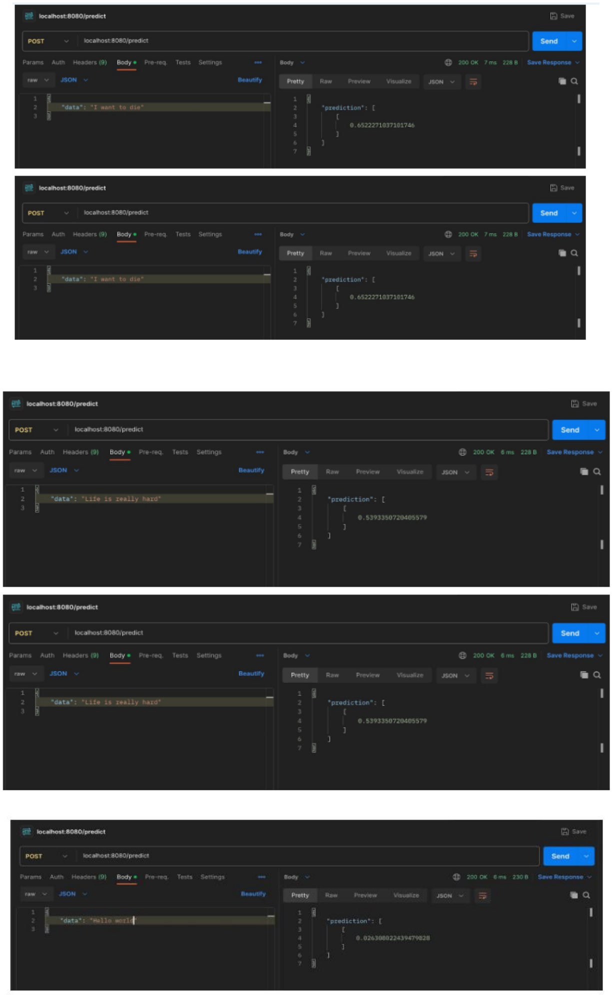

Figure 9 output results

E. Performance plot

We created learning curves to illustrate the CNN–LSTM model's performance in terms of each epoch's accuracy in training and validating to more accurately evaluate the experimental data.

Figure 10 displays the suggested model's validation performance for the identification of suicidal thoughts. These experiments were carried out to assess the suggested model's efficacy. Throughout an operational period of 20 epochs, the CNN–LSTM model attained a validation accuracy of 94%, having started at 92% accuracy. When compared to the initial value of 0.275, the validation loss was much improved, with cross-entropy measurements helping to get it down to as low as 0.20. Figure 11 shows the confusion matrix of testing the proposed model.

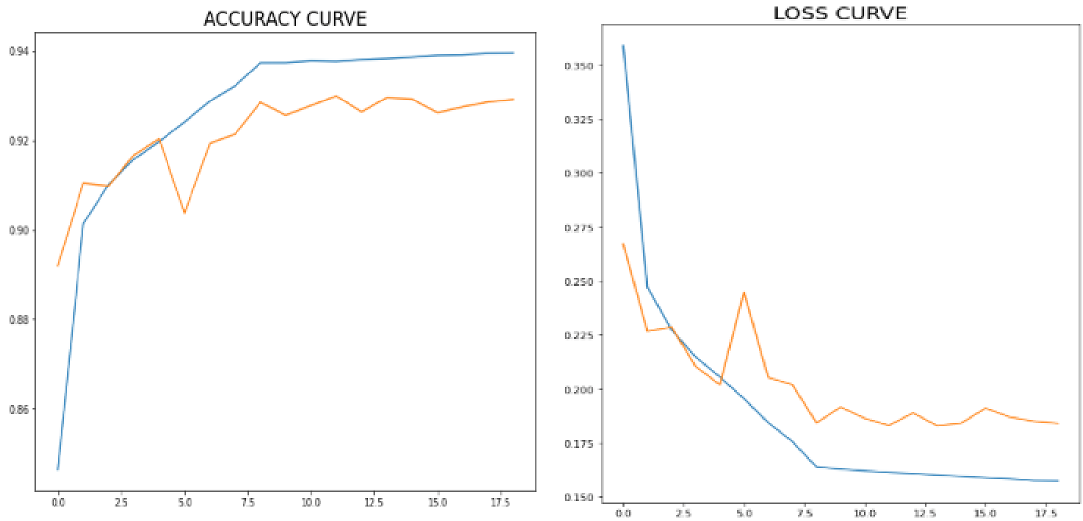

Figure 10 Training and validation (a) accuracy and (b) loss using textual features.

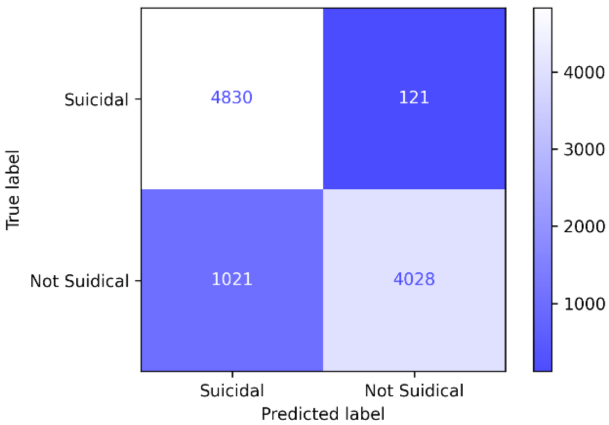

Figure 11 Confusion matrix

F. Discussion

Preventing suicide by identifying suicidal thoughts early on is a crucial and successful strategy. Psychologists have conducted the majority of this topic's research using statistical analysis. However, the majority of computer scientists' research employed DL representational learning and feature engineering-based machine learning. If medical experts can detect early suicidal intentions on microblogging sites like Twitter, their duties will be made a great deal simpler and perhaps adv. Modifies a verb, indicating a lack of certainty. By chance. noun an uncertainty. Report Word many lives spared. Techniques utilizing both deep learning and machine learning and ML have a chance to create new opportunities for improving early detection of suicidal idealization and early prevention. Table 1 shows the comparative analysis of the performance of proposed model with other existing methods.

Table 1 Comparative analysis of the performance of proposed model with other existing methods.

Paper Id	Dataset	Word representation	Model	Result
Ref [10]	6000 post	Word2Vec	LSTM-CNN	93 % accuracy
Ref [11]	7000 post	Word2Vec	LSTM	92% accuracy
Ref [12]	900 post	TF-IDF	SVM	92% accuracy
Proposed model	10000 post	Word2Vec	CNN-LSTM	94% accuracy

5. CONCLUSION

One important and successful policy for reducing suicide is the early diagnosis of suicidal ideation. The majority of this field's work has been conducted by psychologists employing statistical analysis, as opposed to computer scientists have employed representation learning based on DL and ML employing feature engineering. Early suicide ideation detected on microblogging platforms like Twitter would enable medical professionals to recognize and save a great number of lives. In order to enhance the identification of suicidal thoughts and prevent suicide before it happens, the DL and ML techniques may provide new possibilities.

In order to determine whether suicidal ideation indicators are present in user tweetsin this study, we compared and looked at many ML and DL models. The main objective of the investigation was to ascertain which model performed the best at identifying suicidal thoughts on Twitter users with remarkable precision. While there are numerous publicly accessible datasets for the job of suicidal ideation on various suicide forums or subreddits, an actual truth dataset isn't available to assess tweets posted online. Because of this, Using both suicide and non-suicidal phrases, users' real-time tweets served as the foundation for our experimental dataset. To train ML and DL algorithms, we next pre-processed the text using a variety of natural language processing approaches.

Word embedding and the CountVectorizer feature extractor were used, respectively, to train ML techniques. Our experimental results showed that the CNN-LSTM model performed better than the other models in validation, training, and testing. In our tests, the model performs better than both the ML and DL models, with a precision of 94%.. The rationale behind CNN-LSTM's enhanced performance is because it handles forward-backward interdependence from features sequencing handling disappearance of gradients and long-term reliance more skillfully, allowing it to extract pertinent information from longer tweets more efficiently.

6. ACKNOWLEDGEMENT:

We would like to express our gratitude to those who contributed to the success of this research. We also thank the organizers of the 3rd International Conference on Exploring Colossal Technological Advances of the 21st Century in Social Science, Education, Engineering, and Technology (ICECTASET) for providing a platform to present our findings and engage in meaningful academic discussions.

REFERENCES

[1] Arowosegbe, A., & Oyelade, T. (2023). Application of natural language processing (NLP) in detecting and preventing suicide ideation: A systematic review. Int J Environ Res Public Health, 20(2), 1514. https://doi.org/10.3390/ijerph20021514

[2] Centers for Disease Control and Prevention. (2023). Facts about suicide. https://www.cdc.gov/suicide/facts/index.html

[3] Coppersmith, G., Leary, R., Crutchley, P., & Fine, A. (2018). Natural language processing of social media as screening for suicide risk. Biomedical Informatics Insights, 10. https://doi.org/10.1177/1178222618792860

[4] Haque, R., Islam, N., Islam, M., & Ahsan, M. (2022). A comparative analysis on suicidal ideation detection using NLP, machine, and deep learning. Technologies, 10(3), 57. https://doi.org/10.3390/technologies10030057

[5] Mendes, L., Leonido, L., & Morgado, E. (2023). Correlation between suicidal ideation and addiction to various social media platforms in a sample of young adults: The benefits of physical activity. Societies, 13(4), 82. https://doi.org/10.3390/soc13040082

[6] University of the Philippines Population Institute. (2022). Pinoy youth in worse mental shape today, nationwide survey indicates. https://www.uppi.upd.edu.ph/news/2022/pinoy-youth-in-worse-mental-health-shape-today

[7] Van Houdt, G., Mosquera, C., & Napoles, G. (2020). A review on the long short-term memory model. Artificial Intelligence Review, 53(1). https://doi.org/10.1007/s10462-020-09838-1

[8] Zandbiglari, K., Hasanzadeh, H., Kotecha, P., Sajdeya, R., Goodin, A., & Jiao, T. (2023). A natural language processing algorithm for classifying suicidal behaviors in Alzheimer's disease and related dementia patients: Development and validation using electronic health records data. Medrxiv. https://doi.org/10.1101/2023.07.21.23292976

[9] Zhang, T., Schoene, A., Ji, S., & Ananiadou, S. (2022). Natural language processing applied to mental illness. Digital Medicine , 1-46.

[10] Tadesse M.M., Lin H., Xu B., Yang L. Detection of suicide ideation in social media forums using deep learning. Algorithms. 2020;13:7. doi: 10.3390/a13010007.

[11] Desmet B., Hoste V. Emotion detection in suicide notes. Expert Syst. Appl. 2013;40:6351–6358. doi: 10.1016/j.eswa.2013.05.050.

[12] Aladağ A.E., Muderrisoglu S., Akbas N.B., Zahmacioglu O., Bingol H.O. Detecting suicidal ideation on forums: Proof-of-concept study. J. Med. Internet Res. 2018;20:e215. doi: 10.2196/jmir.9840