Journal of Harbin Engineering University
ISSN: 1006-7043

Vol 46 No. 7
July 2025

Suicidal Post Detection on Reddit using Deep Learning Techniques

Swati Bansode [1], Dr. Bhavesh Patel [2], Dr. Vaishali Hirlekar [3]

[1, 2, 3] Shah and Anchor kutchhi Engineering College, Department of Computer Engineering, Mumbai, India
Email Id: [1] swati.17019@sakec.ac.in, [2] bhavesh.patel@sakec.ac.in, [3] vaishali.hirlekar@sakec.a.in

Abstract

One of the main causes of death is suicide worldwide, particularly among the young generation. The increase in suicidal posts on social media platforms such as Reddit has presented both challenges and opportunities for mental health intervention. Our work aims to use vast data is generated by individuals on Reddit by using advanced deep learning techniques to identify suicidal posts and non-suicidal posts. The dataset used is collected from the Reddit API called (PRAW) with the help of various subreddits (e.g., SuicideWatch, Anxiety, and Depression) and neutral topics (e.g., Jokes, Movies, Popular, Books). The proposed model uses long-short-term memory LSTM, bidirectional LSTM (Bi-LSTM), gated recurrent units (GRU), bidirectional GRU (Bi-GRU), and modified BERT-based transformers. The BERT-based model performed better in compared to other models with an accuracy of 98.5%, a precision of 98. 5%, and a recall of 98. 5%. These experimental results successfully verify the theoretical efficiency and adaptability of the proposed model in real-time suicidal post-detection.

Keywords: Suicidal post-detection, deep learning, natural language processing, Reddit, BERT, LSTM, Bi-LSTM, GRU and Bi-GRU.

1. Introduction

One of the biggest issues facing our age is suicide. Every year, 76,000 people harm themselves and many more attempt suicide. Every suicide is a tragedy that has a lifelong impact on survivors and affects families, communities, and even nations. Suicide affects people of all ages and was the third most common cause of death worldwide in 2021 among those aged 15 to 29 years [41]

Since social networking websites has evolved into a platform for individuals to express their emotions, such as tension, anxieties, sadness, and suicidal thoughts, it is now crucial to detect suicidal thoughts on these platforms. With the rise of social media sites like Reddit, deep learning and natural language processing (NLP) have demonstrated considerable promise in analyzing content generated by users to identify suicidal thoughts and assist individuals with mental health difficulties.

By using advance deep learning models such as Long Short-Term Memory (LSTM), Bidirectional LSTM (Bi-LSTM), Gated Recurrent Units (GRU), Bidirectional GRU (Bi-GRU), and a customized BERT-based transformer model, our model's main objective is to identify suicidal posts on the Reddit website.

The Reddit Dataset is used to train the models, which the Python Reddit API Wrapper (PRAW) [42] was used to gather. Posts from a variety of Reddit communities, or subreddits, such as SuicideWatch, Anxiety, and Depression, as well as general interest subreddits like Books, Movies, Popularity, and Jokes, make up our dataset. This varied dataset provides a balance for model training and model analysis by capturing not only posts on mental health but also natural and non-suicidal posts.

This study's contributions are threefold:

1. Knowledge Discovery:

This study analyzes linguistic and behavioral patterns suggestive of suicide ideation in online user material using deep learning techniques. Suicidal people often use personal pronouns, convey negativity, and explain their suffering in the present tense, while referring to their hopelessness and suicide intentions in the future tense, as each post reveals.

2. Dataset and Platform:

Our work uses a dataset that was curated from Reddit using the PRAW Reddit API, which included posts from neutral post from subreddits (Books, Movies, Popular, and Jokes) and subreddits like SuicideWatch, Anxiety,

and Depression. This dataset can be applied to general communication patterns in various circumstances as well as suicide post detection.

3. Deep Learning Models and Benchmarking:

The performance of different deep learning methods, such as LSTM, Bi-LSTM, GRU, Bi-GRU, and a customized BERT-based model, is compared in our work. Linguistic, contextual, and topic-specific information are used to instruct and evaluate these models. Future research in social media mental health detection can build on the baseline results, which can be used to detect suicide posts.

This work intends to provide reliable and scalable methods for identifying suicide posts on Reddit by fusing advanced methods of deep learning with a dataset. The findings could support real-time mental health monitoring and interventions to help identify those who are at risk early.

2. Related Work

Balaji et al. [1] used interviews with 47 survivors in an organized manner, ages 15 to 29, to investigate suicide attempts among the youngsters in India. The study finds that the main risk factors for suicide attempts are impulsivity, psychological discomfort, and interpersonal stressors.

Reynolds et al. [2] used machine learning to detect cyber-bullying on social media. The dataset utilized in this study was gathered from Formspring.me, labeled using Amazon Mechanical Turk, and trained using the WEKA toolbox for machine learning. Using a decision tree approach, they were able achieve a true positive rate of 78.5%. The study demonstrates how well machine learning algorithms can detect language patterns used in cyberbullying on form spring posts.

With a focus on depressive symptoms, Mowery et al. [3] created ML classifiers to identify symptoms of depression on Twitter, such as exhaustion, sleep disturbances, and low mood. This study used classifier models such decision trees, random forests, logistic regression, and support vector machines, which produce a moderate F1 score, using the SAD dataset.

Bi-LSTM and transformer-based pre-trained language models, such as BERT and ALBERT, ROBERTA, and XLNET, have been compared by Haque et al. [4]. Using suicidal intent posts on the Reddit platform, they refined their model and showed that, in terms of performance, transformer-based models outperform traditional deep learning models, with a highest accuracy of 95% of ROBERTa.

To determine whether a tweet contains suicide intent, Ananthankr- ishnan and Kumar [5] also employed BERT-based transformer models. Using a one cycle learning rate to maximize training, the study examined BERT-based models like Distil- BERT, ALBERT, RoBERTa, and DistilRoBERTa. In comparison to other BERT-based models, their work demonstrates that the Roberta model achieves a highest precision of 95.39% on the test dataset by fine-tuning the model using labeled tweets.

Using a multi-class classification strategy based on the RoBERTa language model, Murarka et al. [6] focused on categorizing mental diseases such depression, anxiety, bipolar disorder, ADHD, and PTSD on social media. By using Reddit data to train a classifier and behavioral tests for evaluating the model, this study advances the field. RoBERTa performs better than both LSTM and BERT models, according to the data.

The hybrid machine learning approach was created by Rabani and Khanday [7] to analyze suicidal thoughts on Reddit and Twitter. Mental health professionals annotate the dataset they use in their work to categorize posts into three groups: high risk, moderate risk, and no risk. They used several machine learning models, such as CNN, SVM, and Random Forest, and the hybrid model performed the best. In their evaluation of developments in PTSD detection tools,

Farrow and Jayarathna [8] discussed the latest developments in this field. They talked about the drawbacks of the conventional PTSD diagnosis process, which mostly depends on arbitrary interview-based techniques that may result in incorrect diagnoses. Their research looked at different strategies utilizing neurological and physiological indicators such eye tracking, EEG, and heart rate variability. These techniques improve the objectivity of PTSD symptom detection.

Ji et al. [9] applied supervised learning to identify suicidal thoughts on social media sites such as Reddit and Twitter. Support Vector Machines (SVM), Random Forests, Gradient Boosting Decision Trees (GBDT), XGBoost, and Long Short- Term Memory (LSTM) networks were among the ML and natural NLP methods used in this study. Word embeddings have been used to increase the precision to identify suicide-related

words by extracting linguistic and syntactic elements from web content.

In this study, Taunk, Singh, and Behera [10] deployed time series analysis and the Facebook Prophet forecasting model to analyze and predict suicide rates in India. The study found patterns and forecasted future suicide rates using historical suicide data from India's Ministry of Human Resource Development. In order to provide quarterly and annual estimates, the authors integrated holidays and seasonal factors into their model. According to the study's findings, suicide rates are rising in India, highlighting the need for more recent statistics and preventative measures.

In this study, Antonova et al. [11] explored how the COVID- 19 pandemic affected suicide rates and related risk variables across the globe. They analyze trends in suicide rates from 1986 to 2016 using data from the World Health Organization (WHO), extended by data from 2019 onward. To find trends and connections in suicide risk variables such economic down- turns, social isolation, and mental health issues, the study uses machine learning algorithms and data visualization tools in Python and R. Their analysis highlights the need for better data collection by indicating a connection between elevated stressors from the pandemic and an increased risk of suicide.

In their study, Chatterjee et al. [12] use ML and NLP to evaluate linguistic patterns related to suicide risk in order to predict suicidal ideation in Twitter data. Their study uses sentiment analysis, Latent Dirichlet Allocation (LDA) for modeling, and logistic regression to identify suicidal tendencies with 87% accuracy using a dataset of 188,704 tweets from California, USA. They focus how well sentiment characteristics, statistics, and emoticons work together to achieve high accuracy. According to their findings, social media data processed using NLP techniques can offer important insights into detecting early signs of suicide thinking among users.

In this study, Sharma and Kumar [13] have analyzed India's suicide rates as well as the effectiveness of the country's mental health programs throughout the previous ten years. They evaluate trends, causes, and demographic differences in suicide rates from 2010 to 2019 using data from the Health Management Information System (HMIS) and the National Crime Record Bureau (NCRB). Inconsistencies in data collection techniques and reporting quality are shown by their statistical analysis of growth rates and comparisons between NCRB and HMIS data. They discover that the majority of suicides among young people are caused by family troubles, unstable finances, and health challenges.

In order to predict suicidal tendencies, Chordia et al. [14] have put out a unique approach that combines text, audio, and video analysis. The model makes use of cosine similarity measures, the DeepFace library, and speech recognition, sentiment identification, and facial expression analysis. However, this method shows how multimodal data can be used to increase accuracy.

Imam et al.'s study [15] used natural language processing (NLP) techniques to detect suicidal thoughts in Reddit data. The highest possible accuracy of 93% was attained by the study using Random Forest classifiers and Long Short- Term Memory (LSTM) networks. The authors created a useful dataset by using the Pushshift API to gather data from the "Depression" and "SuicideWatch" subreddits.

To identify suicidal thoughts on Twitter, Sawhney et al. [16] have created a methodology that combines hyperbolic graph convolutions with social and emotional history. Their approach, which combined social ties and emotional history, produced an 81% recall and a 79% macro F1-score.

By using neural models that make use of social networks and historical user data, Sinha et al. [17] have put up multiple methods to detect suicidal intent on Twitter. Text classifiers, sequential models, stacking ensembles, and Graph Convolutional Networks (GCNs) were among the methods used. For this study, 34,306 tweets with hand annotations were used. However, the suggested techniques produced cutting- edge outcomes.

While resolving privacy and ethical considerations, Copper- smith et al. [18] showed how to use natural language processing (NLP) to evaluate Social media posts potential to increase the risk of suicide. OurDataHelps.org platform data was analyzed using models including GloVe embeddings, LSTM, and attention processes.

DiPietro et al. [19] presented the Robin dataset, which consists of more than 1.1 million Reddit posts and is the largest suicidal text corpus to date. Advanced machine learning models can be developed because of this dataset's availability of complicated types of suicidal content. With an accuracy of 97%, the BERT model outperformed Logistic Regression, which had an

accuracy of 94%. In addition to addressing ethical issues with privacy and Data's sensitive nature the study emphasized the value of varied datasets for generalization.

In their work, Malhotra and Jindal [20] systematically reviewed 96 studies on deep learning methods for social media suicide, self-harm, and depression detection. In addition to feature extraction methods like GloVe and faster Text, their review covers a variety of deep learning models, such as ANN, DNN, CNN, RNN, GRU, Bi-RNN, Bi-LSTM, Bi-GRU, BERT and XLNet. However, 99% accuracy was attained by BERT and other models.

In their study, Chadha et al. [21] employed machine learning algorithms to distinguish between posts that were suicidal and those which not. On a dataset gathered from Reddit, the authors used a number of techniques, such as SVM, Logistic Regression, and AdaBoost. According to their results, SVM achieved the maximum precision (80.72%), while Logistic Regression offered the best total accuracy (80.75

Ji et al.'s study [22] highlights the importance of feature engineering, namely in terms of lexical, syntactic, and se- mantic aspects that may be extracted from online text. Using feature engineering, well-known algorithms like SVM, Random Forest, Gradient Boosting Decision Trees (GBDT), and XGBoost have been utilized to categorize text data according to linguistic patterns and terms associated with suicidal thoughts. Furthermore, deep learning models such as Long Short-Term Memory (LSTM) networks—have demonstrated efficacy in extracting intricate contextual details from social media posts, improving model performance in detecting persons who are at danger. The study examined a number of datasets that were used to train these models, including those from Twitter, Reddit, the CLPsych, and the eRisk challenges. All of these datasets offered important real-world information about patterns of suicidal thought.

In their work, Jaideep Yadav et al. [23] developed a BERT- based approach for identifying cyberbullying on social media, with an emphasis on text from sites such as Wikipedia and Formspring. To categorize bullying content, the study used a single linear neural network layer and the pre-trained BERT model. Since of its transformer design, BERT is more accurate than standard models since it can capture context inside text. They obtained a high F1-score of 94% by oversampling the minority (bullying) class in the Formspring dataset.

In their study, Andrew Kipkebut and Emmanuel Chesire [24] employed BERT to detect suicidal thoughts in posts from Kenyan social media sites like Facebook and Twitter. By spotting linguistic patterns that could identify distress or suicidal thoughts, this methodology aims to prevent crises by focusing on early detection of suicidal intent. The BERT-based model obtained high precision and recall scores (91 and 87, respectively) after highly preprocessing and tokenization.

In order to identify suicidal ideation, Nandhini B. et al. [25] employed a number of ML and deep learning methods, such as CNNs and RNNs for text and image processing. With an emphasis on early intervention and real-time monitoring, this study examined a broader variety of suicidal ideation expressions covering text, visual, and multimedia data on platforms such as Twitter. Since there are usually far more non-suicidal messages than suicidal ones on social media, their work brought attention to the need to address data imbalances in these datasets.

3. Data Collection and Exploration:

The data was collected from social media platform Reddit through their APIs.

Reddit's extensive variety of user-generated content, which includes posts, comments, and debates on a wide range of subjects, made it the study main source of data. The Python Reddit API Wrapper (PRAW) [42], a Python framework for secure interaction with the Reddit API, was used to retrieve the data. Reddit's API requires authentication in order to access the data. Creating an application on Reddit's developer site and acquiring the client id, client secret, password, user agent, and username were the steps in this process.

1. Dataset Description:

Posts from subreddits with text fields such titles, body content, and unique identifiers to protect user anonymity created to the gathered dataset. Based on their intent, posts were divided into two categories: suicidal and non-suicidal. Some of the example of posts are shown in Table 1 below.

The following subreddits were the source of posts with suicide intent: Depression, Anxiety, and SuicideWatch The following sources provided posts classified as non-suicidal: r/Jokes, r/Movies, r/Books, and r/Popular.

Table 1: Rules and Example of suicidal and Non-suicidal Text separated by community

Categories	Community	Rules	Example
Suicidal Text	r/SuicideWatch r/Anxiety r/depression	Suicidal related Thought. Suicidal related though. Potential suicidal thought.	I will kill myself. I can't do it any- more. I think about suicide every day.
Non-Suicidal Text	r/Jokes r/movies r/books r/popular	Discussing about Death. Discussing of others suicide. Not related to suicide Not related to suicide	I was choking on a piece of food and it kind of scared some people. I almost died. Japanese Actress Sei Ashina Dies of Suicide at Age 36. Is anyone else hyped about the new Ron Cher now biography on Mark Twain? The US has the fourth highest suicide rate

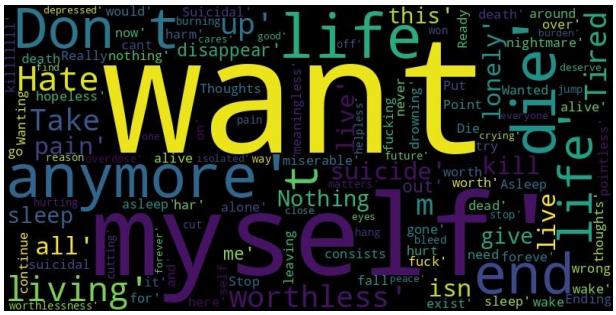

Figure 1. Keyword Based Word Cloud

2. Data Labeling:

A keyword-based search strategy was used to distinguish between posts that were suicidal and those that weren't. Posts with suicidal intent were identified using keywords that are suggestive of suicidal ideation, such as "disappear", "die", "worthless" and "pain" A word cloud was used to illustrate the keyword identification procedure Figure 1. According to this analysis, there were a lot of phrases associated with mental health issues and hopelessness.

3. Dataset Statistics and Imbalance:

The final dataset contained a total of 8,237 posts in all, which split into two groups 5,278 non-suicidal posts; 2,959 suicidal posts With more non-suicidal posts than suicidal ones, this distribution showed an imbalance in the dataset. Oversampling was used to improve model training performance and balance the dataset in order to solve this problem.

4. Word Cloud:

The increase in the use of words like "know", "think" and "reason" indicates a higher emphasis on cognitive and analytical conversations. The Figure 2 word cloud shows an emphasis on emotional vocabulary like "feel" and "die" and a greater emphasis on relational and action-oriented words like "help", "talk" and "friend" Both word clouds retain terms like "kill" and "end" indicating continuous references to extreme emotional states.

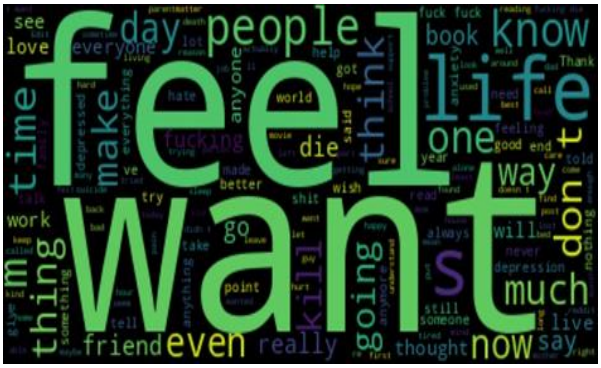

Figure 2. Word Cloud

4. Methodology

The aim of this study is to determine the top-performing algorithms that assist in detecting suicidal posts on Reddit. The various deep learning algorithms, including LSTM, Bi-LSTM, GRU, Bi-GRU, and one transformer that we customized in the layer BERT

model, have been compared in this study. Our suggested framework for this paper is shown in Figure 3.

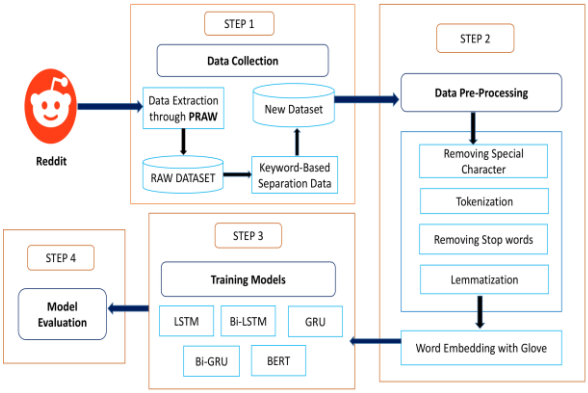

Figure 3. Proposed model

Data Pre-Processing and Feature Engineering

Due to its user-generated nature, the textual data gathered from Reddit is naturally noisy. It contains a number of inconsistent and unnecessary components, including slang, special characters, casual language, and excessive data. This data was subjected to a systematic pre-processing and feature engineering procedure to guarantee its suitability for computational analysis and deep learning models. The steps listed below describe the specific procedures used:

1) Removing Special Characters:

Special letters, symbols, and other non-alphanumeric components are frequently found in textual data, but they don't add to the text's semantic meaning. In this step, components including punctuation (.,,!,?), symbols (#, $,\@), emojis and other irrelevant characters had to be found and eliminated.

For example:

Original text: “I’m feeling #anxious! □□ Can’t seem to calm down @ all…”

After cleaning: “I’m feeling anxious Can’t seem to calm down”

2) Tokenization:

The practice of into discrete dividing text words, phrases, or significant subunits known as tokens is known as tokenization. This process simplifies the data and makes word-level analysis easier. Natural language processing (NLP) libraries like NLTK were used for tokenization in order to appropriately handle linguistic subtleties.

For example:

Text: “I am feeling overwhelmed today.”

Tokens: [“I,” “am,” “feeling,” “overwhelmed,” “today”].

3) Removing Stop Words:

Common English words known as "stop words" have minimal semantic significance when used in an analysis. The words “the”, “is”, “and”, “to”, “of” and “in” are among the examples. Although they are commonly used, these terms don't help individuals grasp the text's underlying meaning or intent.

These words were effectively eliminated from the dataset using a predetermined list of stop words from NLP libraries.

For example:

Original tokens: [“I,” “am,” “feeling,” “very,” “overwhelmed,” “today”]

After stop word removal: [“feeling,” “overwhelmed,” “today”]

4) Lemmatization:

A text normalization method called lemmatization breaks words down to their basic or root form. Lemmatization takes into account a word's grammatical context to ensure an accurate transformation, in contrast to stemming, which merely truncates words to a common prefix.

For example:

“running”, “ran” and “runs” → “run”,

“Better” → “good”

5) Word Embedding with GloVe:

A tool in Natural Language Processing (NLP), word embedding converts words into numerical representations so that text data can be processed by deep learning models. A popular technique for creating dense vector representations that capture the semantic links between words in a continuous vector space is called GloVe (Global Vectors for Word Representation).

Pre-trained GloVe embeddings were used for this study. These embeddings come in several dimensions, including 50, 100, 200, and 300-dimensional and Common Crawl. In these dimensions, each word is

represented as a dense vector that maintains relationships between words. Vectors, and were trained on extensive corpora like Wikipedia

5. Training Models

Batch processing is made possible by padding these tokenized sequences to ensure consistent length, which we describe in the previous step. In this stage, each word is converted into a multi-dimensional vector that captures its semantic meaning. The deep learning models use this embedding as their input for training models.

1) LSTM (Long Term short memory):

In order to handle sequential input, LSTM is made to learn both short-term and long-term dependencies. Word by word as it processes the sequence while preserving a concealed internal state that changes in response to new inputs. The LSTM generates a context vector that summarizes the sequence by carefully choosing whether information to keep, update, or delete using internal procedures called gates. A sigmoid activation function generates a probability indicating the presence of suicidal ideation once this context vector is sent to a fully connected (dense) layer.

2) B. Bi-LSTM (Bidirectional Long Term short memory):

By processing the sequence both forward and backward, the Bi-LSTM improves upon the LSTM. The backward LSTM records dependencies in reverse order, from the end to the beginning, whereas the forward LSTM records dependencies from the beginning to the end of the sequence. A richer representation of the text is produced by concatenating the outputs from both sides, which captures context from both past and future words. Because of this, Bi-LSTMs are especially good at detecting complex emotional meaning in post.

3) GRU (Gated Recurrent Unit):

It turns the input and forget gates into a single update gate, simplifying the LSTM architecture. This simplified architecture preserves the capacity to identify important connections in the text while lowering computing complexity. The sequence is efficiently processed by the GRU, which keeps relevant details while eliminating unnecessary ones. It produces a context vector that is sent to a dense layer for prediction, just as the LSTM.

4) Bi-GRU (Bidirectional Gated Recurrent Unit):

It turns the input and forget gates into a single update gate, simplifying the LSTM architecture. This simplified architecture preserves the capacity to identify important connections in the text while lowering computing complexity. The sequence is efficiently processed by the GRU, which keeps relevant details while eliminating unnecessary ones. It produces a context vector that is sent to a dense layer for prediction, just as the LSTM.

5) BERT (Bidirectional Encoder Representations from Transformers):

Installing the required libraries, such as PyTorch for deep learning, Transformers for working with BERT, and other crucial libraries, such as Pandas, NumPy, and Seaborn for data manipulation and visualization, is the first step in setting up the environment. The dataset is loaded and preprocessed when the environment is prepared. This involves collecting the Reddit dataset and then cleaning it by removing stop words, special characters, links, and additional symbols. After being tokenized with the BERT tokenizer, the text is classified as either suicidal or non-suicidal.

To effectively handle batches during training, the preprocessed data is transformed into a PyTorch dataset and supplied into a DataLoader. In order to improve the model's classification capabilities, a modified version of the previously trained BERT (Bidirectional Encoder Representations from Transformers) model is then created by adding extra layers such fully connected layers and dropout layers. Choosing an optimizer like AdamW to update the model's weights and defining an appropriate loss function, like Cross-Entropy Loss, are the first steps in the training process. After then, the model is trained over a number of epochs using Google Colab (GPU) to speed up calculations, and its performance is optimized by hyperparameter tuning.

Once training is complete, the model is evaluated to assess its ability to classify suicidal and non-suicidal posts.

6. Result

Five deep learning models—LSTM, Bi-LSTM, GRU, Bi-GRU, and BERT—were assessed for their ability to identify suicidal ideation in Reddit posts. The datasets were employed for the evaluation: Dataset, which was

a carefully selected collection from Reddit subreddits devoted to general issues and mental health. The outcomes show how BERT transformer-based architecture, which is excellent at capturing semantic and contextual details, allows it to perform better.

BERT obtained the highest accuracy of 98.5%, as indicated in Table 2, with matching precision, recall, and F1-score values of 98.5%. This result was marginally better than Bi-GRU (97.9% accuracy) and much better than LSTM (93.3% accuracy). The significance of utilizing forward and backward contextual dependencies in the classification task was underscored by the superior performance of the bidirectional models (Bi-LSTM and Bi-GRU) over their unidirectional counterparts (LSTM and GRU). With an accuracy of 97.4%, GRU outperformed LSTM by a small margin. This was probably because to its more straightforward architecture, which maximizes sequential learning.

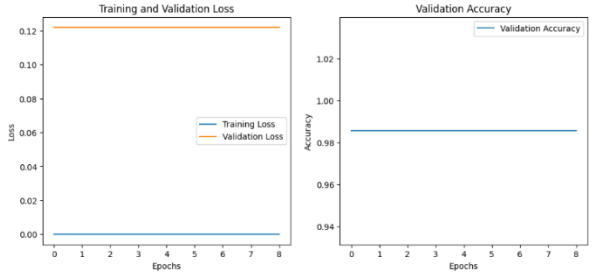

Figure 4. BERT model accuracy and loss curves for dataset

Table 2: Performance metrics

Models	Accuracy	Precision	Recall	F1 Score
LSTM	93.3	92.7	93.6	93.1
Bi-LSTM	97.6	96.2	97.6	97.6
GRU	97.4	98.2	96.5	97.3
Bi-GRU	97.9	96.3	96.8	97.9
BERT	98.5	98.5	98.5	98.5

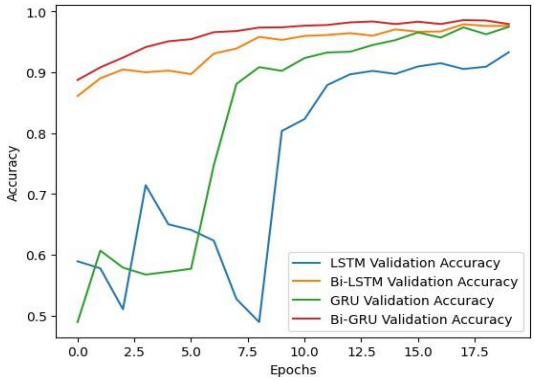

Figure 4(a). validation Accuracy of dataset

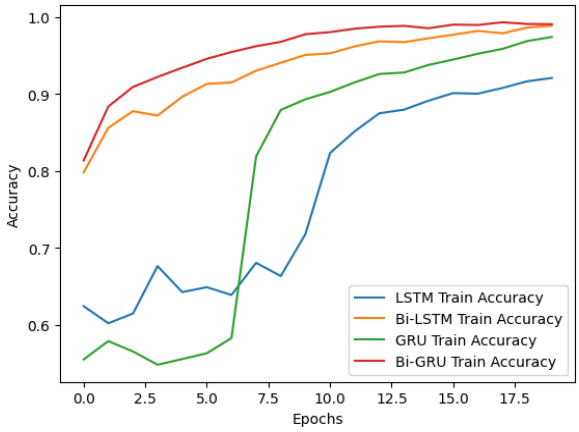

Figure 5(b). Training Accuracy of dataset

Figure 5: All of the pre-trained models accuracy and loss curves for dataset

7. Conclusion and Future Work

The goal of this research is to identify Reddit postings that are suicidal and those that are not. Social media is used by individuals to express their opinions, feelings, knowledge, stress, and support, especially young people who are contemplating suicide.

Deep learning methods for identifying suicidal posts in social media content, especially transformer models like BERT. Suicidal and non-suicidal posts are displayed in the results. We have split article using keyword-based search since our work assumes that posts from suicidewatch, anxiety, and depression are suicidal content, while posts from books, movies, popular culture, and jokes are not. However, this isn't always the case. This study categorizes English-language posts that are suicidal or non-suicidal. More supported languages could be added in the future, and text and images could be combined to enhance detection capabilities.

References

[1] Madhumitha Balaji, Kavita Mandhare, Kalyani Nikhare, Arjun K. Shah, Prajakta Kanhere, Smita Panse, Manjeet Santre, Lakshmi Vijayakumar, Michael R. Phillips, Soumitra Pathare, Vikram Patel, Katarzyna Czabanowska, Thomas Krafft, "Why young people attempt suicide in India: A qualitative study of vulnerability to action", SSM - Mental Health,Vol 3,2023,100216,ISSN 2666-5603

[2] Madhumitha Balaji, Kavita Mandhare, Kalyani Nikhare, Arjun K. Shah, Prajakta Kanhere, Smita Panse, Manjeet Santre, Lakshmi Vijayakumar,

Michael R. Phillips, Soumitra Pathare, Vikram Patel, Katarzyna Czabanowska, Thomas Krafft, "Why young people attempt suicide in India: A qualitative study of vulnerability to action", SSM - Mental Health,Vol 3,2023,100216,ISSN 2666-5603

[3] Mowery, Danielle and Bryan, Craig and Conway, Mike. (2017). Feature Studies to Inform the Classification of Depressive Symptoms from Twitter Data for Population Health. 10.48550/arXiv.1701.08229.

[4] F. Haque, R. U. Nur, S. A. Jahan, Z. Mahmud and F. M. Shah, "A Transformer Based Approach to Detect Suicidal Ideation Using Pre-Trained Language Models," 2020 23rd International Conference on Computer and Information Technology (ICCIT), DHAKA, Bangladesh, 2020, pp. 1-5, doi: 10.1109/ICCIT51783.2020.9392692.

[5] Ananthakrishnan, Gayathri and Jayaraman, Ashok Kumar and Trueman, Tina and Mitra, Satanik and Ak, Abinesh and Murugappan, Abirami. (2022). Suicidal Intention Detection in Tweets Using BERT-Based Transformers. 322-327. 10.1109/ ICCCIS56430.2022.10037677.

[6] Ankit Murarka, Balaji Radhakrishnan, and Sushma Ravichandran. 2021. Classification of mental illnesses on social media using RoBERTa. In Proceedings of the 12th International Workshop on Health Text Mining and Information Analysis, pages 59–68, online. Association for Computational Linguistics.

[7] Rabani, Syed and Khanday, Akib and Khanday, Ud and Khan, Qamar Rayees and Hajam, Ayoub and Imran, Ali and Kastrati, Zenun. (2023). Detecting suicidality on social media: Machine learning at rescue. Egyptian Informatics Journal. 24. 10.1016/j.eij.2023.04.003.

[8] Farrow, Bathsheba and Jayarathna, Sampath. (2019). Technological Advancements in Post-Traumatic Stress Disorder Detection: A Survey. 223-228. 10.1109/IRI.2019.00044.

[9] Ji, Shaoxiong and Yu, Celina and Fung, Sai-fu and Pan, Shirui and Long, Guodong. (2018). Supervised Learning for Suicidal Ideation Detection in Online User Content. Complexity. 2018. 1-10. 10.1155/ 2018/6157249.

[10] Taunk, K., Singh, P., and Behera, R.K. (2021). Suicide Trend Analysis and Prediction in India using Facebook Prophet. 2021 8th International Conference on Computing for Sustainable Global Development (INDIACom), 670-675.

[11] H. Antonova, E. C. S. Pamidimukkala, S. Liao and E. Ceesay, "Suicide Rate and Factors Analysis: Pre and Post COVID Pandemic Data Analysis," 2022 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS), Toronto, ON, Canada, 2022, pp. 1-8, doi: 10.1109/IEMTRONICS55184. 2022.9795714.

[12] M. Chatterjee, P. Samanta, P. Kumar and D. Sarkar, "Suicide Ideation Detection using Multiple Feature Analysis from Twitter Data," 2022 IEEE Delhi Section Conference (DELCON), New Delhi, India, 2022, pp. 1-6, doi: 10.1109/DELCON54057.2022. 9753295.

[13] S. Sharma and R. Kumar, "Suicide Rates in India and its Implications for National Mental Health Programme," Population Research Centre, Institute of Economic Growth, University of Delhi, 2022.

[14] Chordia, Gagan and Soni, Harshit and Mehta, Yash and Gupta, Aayush and Nayan, S and Kaliappan, Jayakumar. (2022). Suicidal Prediction Using Video, Audio, And Text Analysis. Cyber-Physical Systems. 10.1109/ICPS55917.2022.00015.

[15] Imam, Rafi and Jyoti, Oishi and Afrin, Zakia and Hossain, Md. Munawar and Mou, Tamanna. (2023). Suicidal Thought Detection Using NLP(Natural Language Processing) on Reddit Data. 1-6. 10.1109/ICCIT60459.2023.10441163.

[16] Ramit Sawhney, Harshit Joshi, Rajiv Ratn Shah, and Lucie Flek. 2021. Suicide Ideation Detection via Social and Temporal User Representations using Hyperbolic Learning. In Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pages 2176–219.

[17] Sinha, Pradyumna and Mishra, Rohan and Sawhney, Ramit and Mahata, Debanjan and Shah, Rajiv Ratn and Liu, Huan. (2019). #suicidal - A Multipronged Approach to Identify and Explore Suicidal Ideation in Twitter. 941-950. 10.1145/ 3357384.3358060.

[18] Coppersmith, Glen and Leary, Ryan and Crutchley, Patrick and Fine, Alex. (2018). Natural Language Processing of Social Media as Screening for Suicide Risk. Biomedical Informatics Insights. 10. 117822261879286. 10.1177/1178222618792860.

[19] DiPietro, Daniel and Hazari, Vivek and Vosoughi, Soroush. (2022). Robin: A Novel Online Suicidal Text Corpus of Substantial Breadth and Scale. 10.48550/arXiv.2209.05707.

[20] Malhotra, A., and Jindal, R. (2022). Deep learning techniques for suicide and depression detection from online social media: A scoping review. Applied Soft Computing, 130, 109713-. https://doi.org/10.1016/j.asoc.2022.109713

[21] A. Chadha, A. Gupta and Y. Kumar, "Suicidal Ideation Detection on Social Media: A Machine Learning Approach," 2022 2nd International Conference on Technological Advancements in Computational Sciences (ICTACS), Tashkent, Uzbekistan, 2022, pp. 685-688, doi: 10.1109/ICTACS56270.2022.

[22] 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.

[23] Yadav, J.P., Kumar, D., and Chauhan, D. (2020). Cyberbullying Detection using Pre-Trained BERT Model. 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), 1096-1100.

[24] Kipkebut, Andrew. (2023). DEEP LEARNING SUICIDE IDEATION DETECTION MODEL. Proceedings of the IRE.

[25] B, N., Jayanthi, Shrinivas, K., and Vinod, P. (2023). Machine Learning Approaches for Suicidal Ideation Detection on Social Media. 2023 4th International Conference on Computation, Automation and Knowledge Management (ICCAKM), 1-7.

[26] Mishra, Ajey and Srivastava, Akanksha and Modi, Aashi and Pandey, Ashish. (2023). Cyberbullying and Suicide Ideation Detection via Hybrid Machine Learning Model. 265-270. 10.1109/SPIN57001.2023.10117371.

[27] K. Nikhileswar, D. Vishal, L. Sphoorthi and S. Fathimabi, "Suicide Ideation Detection in Social Media Forums," 2021 2nd International Conference on Smart Electronics and Communication (ICOSEC), Trichy, India, 2021, pp. 1741-1747, doi: 10.1109/ICOSEC51865.2021.9591887.

[28] Darade, Vrushali Bhanudas. Deep Learning and Natural Language Processing for Suicidal Ideation Using Instagram Posts. Diss. Dublin, National College of Ireland, 2023.

[29] O'Dea, Bridianne and Wan, Stephen and Batterham, Philip and Calear, Alison and Paris, Cécileand Christensen, Helen. (2015). Detecting Suicidality on Twitter. Internet Interventions. 103. 10.1016/j.invent.2015.03.005

[30] Syed, Sobia and Amin, Imran. (2017). Prediction Of Suicide Causes in India Using Machine Learning. Journal of Independent Studies and Research - Computing. 15. 10.31645/jisrc/(2017).15.2.01.

[31] H. M G and A. Kaur, "A Review of "Machine Learning Approaches to Assess Psychological Health: A Comparative Analysis of Non-Meditators and Meditators"," 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), Delhi, India, 2023, pp. 1-7, doi: 10.1109/ICCCNT56998.2023.10306664.

[32] T. Bhardwaj, P. Gupta, A. Goyal, A. Nagpal and V. Jha, "A Review on Suicidal Ideation Detection Based on Machine Learning and Deep Learning Techniques," 2022 IEEE World AI IoT Congress (AIIoT), Seattle, WA, USA, 2022, pp. 027-031, doi: 10.1109/AIIoT54504.2022.9817373.

[33] Chiu, Chun and Lane, Hsien-Yuan and Koh, Jia-Ling and Chen, Arbee. (2021). Multimodal depression detection on instagram considering time interval of posts. Journal of Intelligent Information Systems. 56. 10.1007/s10844-020-00599-5.

[34] Zogan, Hamad and Razzak, Imran and Wang, Xianzhi and Jameel, Shoaib and Xu, Guandong. (2022). Explainable depression detection with multi-aspect features using a hybrid deep learning model on social media. World Wide Web. 25. 10.1007/s11280-021-00992-2.

[35] Ramírez-Cifuentes, D., Freire, A., Baeza-Yates, R., Puntí, J., Medina-Bravo, P., Velazquez, D.A., Gonfaus, J.M., and Gonzàlez, J. (2020). Detection of Suicidal Ideation on Social Media: Multimodal, Relational, and Behavioral Analysis. Journal of Medical Internet Research, 22.

[36] S. N, M. G, V. Mohammed Shaid, S. Purushothman and V. Gokul Sanjai, "Accurate Prediction and Detection of Suicidal Risk using Random Forest Algorithm," in 2024 4th International Conference on Pervasive Computing and Social Networking (ICPCSN), Salem, India, 2024, pp. 287-292, doi: 10.1109/ICPCSN62568.2024.00053.

[37] Cheng, Ju and Chen, Arbee. (2022). Multimodal time-aware attention networks for depression detection. Journal of Intelligent Information Systems. 59. 10.1007/s10844-022-00704-w.

[38] Sawhney, Ramit and Manchanda, Prachi and Singh, Raj and Aggarwal, Swati. (2018). A

Computational Approach to Feature Extraction for Identification of Suicidal Ideation in Tweets. 91-98. 10.18653/v1/P18-3013.

[39] Shrestha, Balkrishna and Rajbanshi, Bhim and Pandey, Shailesh. (2023). An Empirical Investigation of the World Happiness Report: The Effect of Reliogisity, Suicide Rate and Human Development. 344-348. 10.1109/ICAICCIT60255.2023.10466183.

[40] S. Faraz, S. S. Azhar Ali and S. Hasan Adil, "Machine Learning and Stress Assessment: A Review," 2018 3rd International Conference on Emerging Trends in Engineering, Sciences and Technology (ICEEST), Karachi, Pakistan, 2018.

[41] World Health Organization, “Suicide,” WHO, Nov. 17, 2021. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/suicide.

[42] PRAW Documentation, "PRAW: The Python Reddit API Wrapper,"[Online]. https://praw.readthedocs.io/en/stable/index.html.