https://doi.org/ 10.33472/AFJBS.6.5.2024. 3804-3815

African Journal of Biological Sciences

A HYBRID ATTENTION BASED DEEP LEARNING SYSTEM FOR SUICIDAL IDEATION DETECTION IN SOCIAL MEDIA

***K.SenilSeby, Dr.M.Elamparithi,** and **Dr.V.Anuratha,**

*Research Scholar, Department of Computer Science,
Kamalam College of Arts and Science, Anthiyur,
Bharathiar University, Coimbatore, Tamil Nadu, India.

*E-mail: edupassresea.r.ch@gmail.com

Associate Professors,
Department of Computer Science,
Kamalam College of Arts and Science, Anthiyur,
Bharathiar University, Coimbatore, Tamil Nadu, India.

E-mail: profelamparithi@gmail.com and profanuratha@gmail.com

Article History
Volume 6, Issue 5, Apr 2024
Received: 01 May 2024
Accepted: 09 May 2024
doi: *10.33472/AFJBS.6.5.2024. 3804-3815*

Abstract: Suicide is a critical issue in modern society. Many of the people who have the tendency to suicide share their thoughts and opinions through social media platforms. Suicidal ideation detection via online social network analysis has emerged as an essential research topic with significant difficulties in the fields of natural language processing and psychology in recent years. This paper proposes a hybrid attention-based convolution neural network and long short-term memory (HACNN-LSTM) for suicidal ideation detection of social media data. The proposed system consists of three phases, namely, data preprocessing, word embedding, and classification. To begin, the data preprocessing is carried out on the collected Reddit dataset. After that, the word embedding is performed on the preprocessed dataset by Term Frequency - Inverse Document Frequency (TF-IDF). Finally, the classification is done by using the HACNN-LSTM model, in which the hyperparameter is optimized by particle swarm optimization (PSO) algorithm. The findings proved that the proposed hybrid system achieves superior results than the existing techniques with 96.84% accuracy.

Keywords: *Suicidal ideation, social media, Reddit dataset, Data Preprocessing, Word Embedding,Classification, and deep Learning.*

1. INTRODUCTION

Social media webs are widespread platforms where anyone can share their thoughts, expressions, emotions, depression, and frustrations. Nowadays, social webs like Myspace, Facebook, Reddit, Instagram, and Twitter provide opportunities to communicate with diverse people [1]. Many users prefer to utilize social media networks to share their thoughts and emotions, and their daily experiences, problems, and issues [2, 3]. Suicidal ideation, death, and self-harming thoughts are among the most widely discussed themes on social media. Suicide is described as a person's deliberate attempt to take their own life [4]. Several factors can influence individuals to make the decision to end their lives (for example, emotional pain, marital problems, and biological, genetic, psychological, social, cultural, financial, and environmental factors) [5, 6]. Due to the complexity of the problem, no single risk factor can be reliably used to predict suicide. For instance, despite the strong association between suicide and depression, a depression diagnosis alone has a limited ability to predict suicide [7]. Their suicidal ideation can be cured with the help of healthcare experts and drugs, but most of them shun medical treatments owing to societal stigma. Instead, individuals prefer to convey their suicidal intentions via social media. However, since mental illness can be detected and addressed, early detection of depression may be the most effective strategy to avoid suicidal ideation [8].

The ML techniques such as naïve bayes, logistic regression, and support vector machine, etc. have improved accuracy of suicidal ideation analysis and expedite automatic evaluation of data this day. But the traditional ML algorithms consume more time and are incomplete. Deep learning algorithms are proven impressive in many computer vision and pattern recognition problems. They use neural networks that provide better results to complex problems than traditional ML algorithms [9]. In addition, DL model identified the social media language patterns that express suicidality [10]. This is motivated to propose a hybrid attention based deep learning system for suicidal ideation detection in social media. The main objectives of the proposed work are explained as follows:

- The proposed system uses TF-IDF to learn more about the relationship between inputs and to process the data more efficiently.
- The proposed system uses HACNN-LSTM for classification, a combination of CNN and LSTM network, and a proposed hybrid model of attention combines the best features to classify the suicidal ideation with higher accuracy.

The rest of our article is divided into various sections. Section 2 discusses the related work done in the domain of the detection suicide thoughts. Section 3 discusses the proposed methodology. Section 4 elaborates on the results and discussion and compares our work with the previous research. Section 5 provides the conclusion and future scope of the work.

2. LITERATURE SURVEY

This section surveys the recently published works related to our proposed work. **Jingfang Liu *et al*. [11]** presented suicidal ideation in social media based on an ensemble method based on feature fusion. Initially, the system manually extracted the language and time features for detection. Herein, the language features were extracted with the help of Chinese psychological analysis software 'Text Mind' to count the frequency of different word categories of posts. Finally, the system used the K-means algorithm based on the Word2vec word vector to cluster keywords highly related to suicide risk. The experimental results showed that the system achieved 80.61% accuracy and a 79.20% F1-score, which was comparatively higher than the existing methods for Weibo dataset. **Moumita Chatterjee *et al.* [12]** suggested a multi-modal feature-based technique for suicide ideation detection from online social media. Initially, the data preprocessing was performed on the collected dataset. After that feature extraction was carried out manually on the preprocessed dataset. Finally, the fused features were passed to the ML classifier such as logistic regression (LR), random forest (RF), support vector machine (SVM), and extreme gradient boosting (XGB) to predict the suicide ideation. The experimental results showed that an accuracy of 87% was obtained based on the LR classifier which outperformed other baselines for Twitter and Reddit dataset.

Pratyaksh Jain *et al.* [13] recommended a suicide analysis based on machine learning and natural language processing (NLP). To begin, the system converted the text in lower casing, removed the punctuation, removed numbers, tokenized the data, removed the stop words, stemmed the data and then finally passed it into machine learning model such as LR, SVM, RF, and Naïve Bayes classifiers to detect the suicide ideation. The model's effectiveness could be seen with 77.29 % accuracy and 0.77 f1-score of LR, 74.35 % accuracy and 0.74 f1-score of Naïve Bayes, 77.12% accuracy and 0.77 f1-score of SVM, 77.298% accuracy, and 0.77 f1-score of RF for Reddit dataset. **Syed Tanzeel Rabani *et al.* [14]** developed a multi-class machine learning classifier for identifying suicidal risk levels in social media posts from Twitter and Reddit platforms. The system involved three steps. To begin, relevant data was extracted from SNS. Next step was about annotating the posts into three levels of risk based upon the annotation scheme that was devised in consultation with mental health experts. The third step involved preprocessing the posts to remove irrelevant and redundant information and the feature extraction was carried out manually to extract the relevant features for training the multi-class machine learning model. Finally, the classification was performed based on the ML models such as SVM, RF, and XGB. The results illustrated that the XGB algorithm achieved an overall accuracy of 96.33% than SVM and RF models.

Tianlin Zhang *et al*. [15] proffered an automatic identification of suicide notes with a transformer-based DL model. Initially, the word embedding was performed based on the GloVe approach and it passed to the transformer encoder to capture features. Then, the features were given as input to the BiLSTM to obtain context-aware information by modeling sequences from forward and backward hidden states. Next, it passed to the max-pooling layer to reduce the dimensionality and finally the reduced features were passed to the classification layer to predict the suicide ideation. To test the effectiveness of the system, the system used subreddits dataset and the system achieved 95.0% precision, 94.9% recall, and 94.9% f-measure metrics respectively and therefore outperformed comparable state-of-the-art deep learning models.

The above-mentioned surveys provide satisfactory results, but it has some limitations. Most of the works uses ML algorithms to classify the suicidal thoughts. ML can ease social media suicidal thoughts analysis. It learns the affective valence of the words, so they do not require a pre-determined dataset. However, if the dimension of the gathered data is too small, they are not able to divide the training set and test set as conventional validation without the loss of significant modeling. To mitigate these problems most of the authors use DL-based algorithms. One of the biggest advantages of using the deep learning approach is its ability to execute feature engineering by itself. In this approach, an algorithm scans the data to identify features which correlate and then combine them to promote faster learning without being told to do so explicitly. However, improvement is still needed to get accurate detection results. Hence, this paper proposes a hybrid attention based deep learning system for suicidal ideation detection in social media.

3. PROPOSED METHODOLOGY

Figure 1 shows the workflow of the proposed work. Initially the tweets of the users are collected from the Reddit dataset. Then preprocessing processes such as removal of URL, hashtags, punctuation, and stop words, stemming, lemmatization and tokenization are performed on the collected dataset to get the more desired results. Then word embedding of the preprocessed data is done using the TF-IDF and Word2vec. Finally, the output of the embedding layer is given to HACNN-LSTM for suicidal detection.

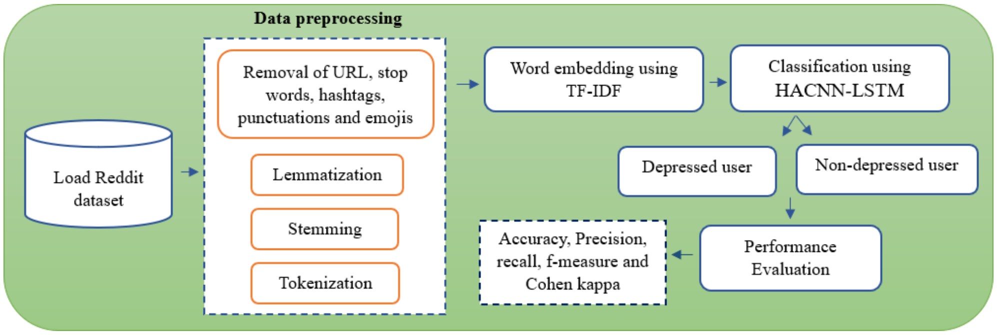

Figure 1: Workflow of the proposed methodology

3.1 Data Preprocessing

To begin, the tweets of the users are gathered from the Reddit dataset. After that, the preprocessing process is carried out to improve the quality of data. Herein, the proposed system performs some preprocessing operations such as removal of URL, hashtags, punctuation, and stop words, stemming, lemmatization and tokenization on the collected dataset to get the more desired results. These are shortly described as follows:

- **Removal of URL:** URL does not carry much information regarding the suicide ideation of the text. So, this process removes the URL link in the sentence.
- **Hashtags removal:** For this process, the hashtag is removed from the text and stored in a separate column. Then we can use the weight of the hashtag in a separate process.
- **Punctuation and Emojis removal:** This process removes the characters "?, !, :, ;, '," and Emojis to make the text easily processable.
- **Stop word removal:** This technique eliminates the frequent usage of words such as "the", "a", "an", and "in", which are meaningless and useless for the text classification. This reduces the corpus size without losing important information.
- **Stemming:** Stemming is the conversion of the words to their root meaning. It lessens the total words and facilitates computational speed.
- **Lemmatization:** Lemmatization takes the consideration of morphological analysis of the words. It reduces inflected words properly with the root words belonging to the sentences. It is also called a lemma which is the set of words in dictionary form, citation form and canonical form.
- **Tokenization:** This setting splits the documents into words/terms, constructing a word vector, known as bag-of-word.

3.2 Word Embedding

Once preprocessing is completed, word embedding is performed on the preprocessed dataset in this phase. Word embedding is the depiction of words in a real-valued vector format which encodes the word meaning, such that the words with similar meaning are nearer in the vector space. It converts each word and sentence of a given text into low-dimensional feature vectors. In this work, the proposed system uses TF-IDF to extract vector representations of words and sentences for suicidal/non-suicidal classification. Term Frequency - Inverse Document Frequency (TF-IDF) is an unsupervised term weighting scheme for information retrieval and text mining. TF-IDF represents the relative frequency of a word in a text document, and inverse document frequency scales with the number of documents. The higher TF-IDF value words imply they have a stronger relationship in the document in which they appear. Initially, it computes the frequency of occurrence of each word in each document (TF) using the following equation. $TF''_{tm} = 1 + \log \left(TF''_{tm}\right)$ (1)

Where, TF''_{tm} refers to the number of occurrences of term (tm). Next, the system counts the number of documents containing a specific word $\left(\overline{\overline{SD}}\right)$. After that, the system computes the inverse document frequency by using equation (2).

$$ITF'_{tm} = \log\left(\frac{TD}{\overline{\overline{SD}}_{tm}}\right) \tag{2}$$

Where, ITF''_{tm} refers the inverse document frequency, TD indicates the number of documents, and $\overline{\overline{SD}}_{tm}$ denotes the number of documents that contains term (tm). Finally, TF-IDF vectorizer or scores a word by multiplying the word's TF with the IDF, which is shown in equation (3).

$$\bar{\omega}_{tm,dc} = TF'_{tm} \times ITF'_{tm} \tag{3}$$

Where, $\bar{\omega}_{tm,dc}$ refers to the weight of term (tm) in document (dc) and after representing all the terms in the document into a vector form, they are given input to the classifier for further classification.

3.3 Classification

After obtaining word vectors using TF-IDF and Word2vec, classification of suicide ideation is done in this phase. In this work, the classification is done by hybrid attention-based convolutional neural network and long short-term memory (HACNN-LSTM). Hybrid CNN and LSTM are the types of DL algorithm that efficiently extracts the spatial-temporal features and long-term dependency features, but we well known that it is difficult to only use these features to accomplish accurate classification tasks, so deeper features often must be combined to accomplish these tasks. So, this paper uses the Multi-head Attention (MHA) model to extract deeper features. Also, hybrid network hyperparameters are tuned using the particle swarm optimization (PSO) algorithm to avoid the randomness and to increase the accuracy of the model. Thus, the combination of CNN, LSTM, and MHA is termed as HACNN-LSTM. The HACNN-LSTM mainly includes '6' layers such as convolutional layer, pooling layer, LSTM layer, MHA layer, fully connected layer, Dropout layer, and SoftMax classifier. These layers are briefly explained as follows:

a) Convolutional layer

The convolution layer convolves its input obtained by calculating the pointwise multiplication between all input channels and filters based on the structure of the feature map. The size of the filters is usually smaller than the actual data. Each filter convolves with the data and creates an activation map. The proposed system uses rectified linear unit (ReLU) $\left(\sigma^{*}\right)$ activation function. The ReLU activation function is a simple calculation that returns the value provided as input directly, or the value 0. It is defined as follows:

$$\sigma^*(\bar{\nu}) = \max(0, \bar{\nu}) \tag{4}$$

Where, $\bar{\nu}$ indicates the input word embedded vector.

b) Pooling layer

After convolution operation was performed, the pooling was taking place. This technique yields a lower-dimensional matrix by extracting values through the data which is convolved is termed a pooling layer. As a consequence, a layer generates matrices which are lower-dimension, which can be thought of as a simplified form of convolved features. As a result of the downsampling technique, the system will be more stable because slight changes in the input will not affect the pooled results.

c) LSTM layer

The pooled feature maps are then passed to the Long Short-Term Memory (LSTM) layer to extract features by maintaining a pair of long-term and short-term memories. LSTM is one of the best algorithms to work with sequence data along with an additional feature of having a memory element. This memory element enables LSTM to remember the previous sequence of steps. It overcomes the difficulty of vanishing gradient, faced with recurrent neural networks by a slight modification in the structure. An LSTM network has the input vector (pooled feature maps) $[h'_{\eta-1}, \bar{P}'_\eta]$ at time step η and the network new cell state is depicted as $\bar{c}_\eta$. The output vectors passed through the network between consecutive time steps η, and $\eta+1$ are denoted by h'_η. An LSTM network has three types of gates, the forget gate $(\bar{f}_\eta)$, input gate $(\bar{i}_\eta)$, and output gate $(\bar{o}_\eta)$, that help revise and control the cell states $(\bar{c}_\eta)$. The gates use ReLU activation function (σ^*) that is computed using equation (5). The LSTM executes certain pre-calculations before producing an output. The pre-calculations are mathematically expressed as follows:

$$\bar{f}_\eta = \sigma^*\left(ow^*_f \cdot \left[h'_{\eta-1}, \bar{P}'_\eta\right] + ob^*_f\right) \tag{5}$$

$$\bar{i}_\eta = \sigma^*\left(ow^*_i \cdot \left[h'_{\eta-1}, \bar{P}'_\eta\right] + ob^*_i\right) \tag{6}$$

$$\bar{c}_\eta = \sigma^*\left(ow^*_c \cdot \left[h'_{\eta-1}, \bar{P}'_\eta\right] + ob^*_c\right) \tag{7}$$

$$\bar{c}_\eta = \left(\bar{f}_\eta \cdot \bar{c}_{\eta-1}\right) + \left(\bar{c}_\eta \cdot \bar{i}_\eta\right) \tag{8}$$

$$\bar{o}_\eta = \sigma^*\left(ow^*_o \cdot \left[h'_{\eta-1}, \bar{P}'_\eta\right] + ob^*_o\right) \tag{9}$$

$$\bar{h}_\eta = \bar{o}_\eta \ \sigma^* \left(\bar{c}_\eta\right) \tag{10}$$

Where, $ow^*_f, ow^*_i, ow^*_o, and\ ow^*_c$ and $ob^*_f, ob^*_i, ob^*_o, and\ ob^*_c$ refers to the optimal weights and bias of the forget gate, input gate, output gate, and memory cell state, respectively and $\bar{h}_\eta$ indicates the hidden state and this layer gives an output sequence of values, that is $\bar{L} = \left\{ \bar{l}_0, \bar{l}_1, \bar{l}_2, ...\bar{l}_n \right\}$ and then it is passed to the Multi-head attention layer.

d) MHA layer

This attention layer is used to aid in the extraction of many significant deep features from the LSTM layer's output for the classification of suicidal thoughts. Herein, the proposed system uses Multi-head Attention (MA) to extracts deeper features from the LSTM's output sequence $\bar{L} = \left\{ \bar{l}_0, \bar{l}_1, \bar{l}_2, ...\bar{l}_n \right\}$. MA is a module for attention mechanisms which runs through an attention mechanism several times in parallel to extract features. The MA mechanism relies on scaled dot-product attention, which takes a query (q_μ), a key (k_μ), and a value (v_μ) from the input. These keys, values and queries are worked in parallel to get the values. First, the MA model transforms q_μ, k_μ, and v_μ into $\hat{S}$ sub-spaces, with different, learnable linear projections. It is expressed as follows:

$$q^{\hat{s}}_\mu, k^{\hat{s}}_\mu, v^{\hat{s}}_\mu = q_\mu \tilde{\varpi}^{q_\mu}_{\hat{s}}, k_\mu \tilde{\varpi}^{k_\mu}_{\hat{s}}, v_\mu \tilde{\varpi}^{v_\mu}_{\hat{s}} \tag{11}$$

Where, $q^{\hat{s}}_\mu, k^{\hat{s}}_\mu, and\ v^{\hat{s}}_\mu$ indicates the $\hat{s}^{-th}$ head of query, key, and value, respectively and $\tilde{\varpi}^{q_\mu}_{\hat{s}}, \tilde{\varpi}^{k_\mu}_{\hat{s}}, and\ \tilde{\varpi}^{v_\mu}_{\hat{s}}$ refers to the weight matrices of query, key and value. Furthermore, $\hat{S}$ attention functions are applied simultaneously to get the output features and it is mathematically expressed as follows:

$$\underline{\underline{MF}} = \overline{AT}^{\hat{s}} \ v^{\hat{s}}_\mu \tag{12}$$

$$\overline{AT}^{\hat{s}} = soft\max\left(\frac{q^{\hat{s}}_\mu \, k^{\hat{s}\ T}_\mu}{\sqrt{dm}}\right) \tag{13}$$

Where, $\sqrt{dm}$ refers the dimensionality of the model, $\overline{AT}^{\hat{s}}$ indicates the attention distribution produced by the $\hat{s}^{-th}$ attention head, and $\underline{\underline{MF}}$ refers to the first head output feature of the MA and then the MA concatenates each output state to produce the final deep feature maps.

e) Dropout layer

Feature maps obtained from the MHA layer are passed to the dropout layer. The dropout layer nullifies the connection of certain neurons in the next layer while leaving others unchanged. The purpose of this layer is to prevent overfitting and the coadaptation of hidden units by the random dropping out of noise that may be present in the training data.

f) Fully connected layer

In a layer that is fully connected each and every input is associated with every output, across the weights. Its objective is to combine the features from the attention layer and multiply with corresponding weights and produce final feature maps to the SoftMax classifier.

g) SoftMax classifier

This is an output layer that we used to estimate the probability of a post being suicidal or non-suicidal. To avoid vanishing problems, the SoftMax function uses a text feature vector acquired as a consequence of the fully connected layers and divides it into dataset classes. It provides the binary classification output. O indicates no risk and 1 indicates the risk. In addition, the whole network hyperparameters (weight and bias) are optimally chosen by the PSO algorithm to maximize the performance of the model on a validation set [16]. PSO solves a problem by trying to optimize a solution (hyperparameters) in an iterative way with respect to some measure of quality by enabling a group of particles (swarm) to scan the search space in a semi-random manner.

4. RESULTS AND DISCUSSION

In this section, the performance of the proposed hybrid attention-based deep learning system for suicidal ideation detection in social media is analyzed with the classical existing approaches in term of evaluation metrics. The system was implemented in the working platform of Python 3.7 with 12 GB of GDDR5 VRAM, and Intel Xeon Processor with two 2.20-GHz cores and 13 GB RAM. The descriptions of the dataset are shown in equation 4.1 and the performance analysis of the proposed work is given in section 4.2.

4.1 Dataset descriptions

To validate the efficacy of the proposed work, the system uses publicly available Reddit dataset downloaded from the Kaggle website, which is accessible through https://www.tensorflow.org/datasets/catalog/reddit. The dataset consists of 3,848,330 posts with an average length of 270 words for content and 28 words for the summary. Features include strings: author, body, normalized Body, content, summary, subreddit, subreddits_id. We split the entire dataset into training, testing, and validation subsets to analyze the performance of the proposed work. The total number of samples is 3,848,330, training samples are 2,693,831, testing samples are 769,666, and the validation data are 3848333. These Reddit posts were labeled using 4 divisions which include no risk, low risk, moderate risk, and severe risk.

4.2 Performance Evaluation

In this section, the performance of the proposed hybrid attention-based convolutional neural network and long short-term memory (HACNN-LSTM) is investigated against the existing convolutional neural network (CNN), recurrent neural network (RNN), artificial neural network

(ANN), and random forest (RF) in terms of accuracy, precision, recall, f-measure, and Cohen kappa metrics. These are shown in the following table.

Table 1: Results analysis of the proposed model

Metrics	Proposed	CNN	RNN	ANN	RF
Accuracy	96.84	93.56	92.14	89.64	85.12
Precision	97.85	94.32	93.26	90.52	86.03
Recall	95.74	92.23	91.56	88.79	84.56
F-measure	96.81	93.52	92.09	89.59	85.07
Cohen kappa	88.98	85.53	84.31	81.23	77.65

Table 1shows the outstanding outcomes than the existing methods. For example, considering the accuracy metric, the existing CNN, RNN, ANN, and RF attains accuracy of 93.56%, 92.14%, 89.64%, and 85.12%, respectively, which was lower than the proposed one, because the proposed one achieves maximum accuracy of 96.84%. Next, considering the precision, recall, and f-measure metrics, the existing CNN offers 94.32% precision, 93.52% recall, and 93.52% f-measure, the existing RNN offers 93.26% precision, 91.56% recall, and 92.09% f-measure, the existing ANN offers 90.52% precision, 88.79% recall, and 89.59% f-measure, and the existing RF yields 86.03% precision, 84.56% recall, and 85.07% f-measure, respectively. Based on these precision, recall, and f-measure metrics, the existing CNN, RNN, ANN, and RF offers less outcomes than the proposed approach, because our proposed approach achieves 97.85% precision, 95.74% recall, and 96.81% f-measure, respectively. In addition, the proposed one achieves best Cohen Kappa values of 88.98% that are better than the existing methods, because the existing methods attains low Cohen kappa values of 85.53%, 84.31%, 81.23%, and 77.65%, respectively. Thus, the overall experimental results show that the proposed one achieves superior outcomes when compared to the existing methods. Why the proposed work attains better outcomes than the existing methods means the proposed work initially performed preprocessing that improves the quality of the collected data. Also, the proposed system uses efficient deep learning algorithms such as CNN and LSTM to extract useful features, and also the proposed system uses MA mechanism to extract deeper features, which helps to accurate classification and attains better performance. This is graphically shown in figure 3.

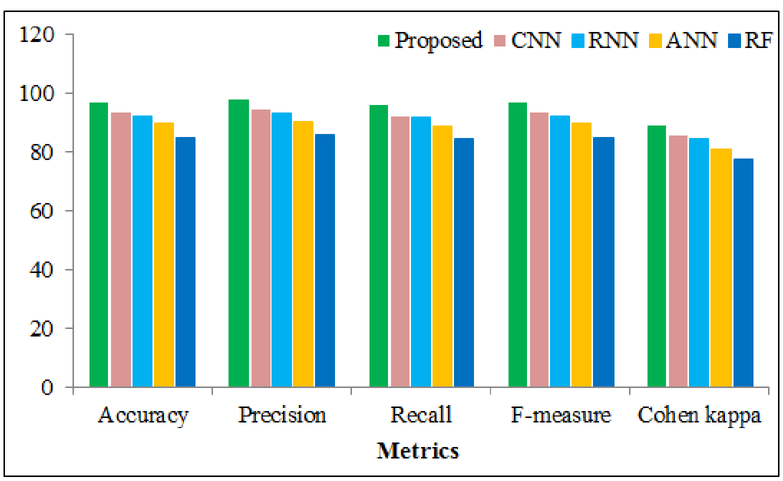

Figure 3: Performance analysis of the proposed work

5. CONCLUSION

In this paper, the proposed work proposes a hybrid attention based deep learning system for suicidal ideation detection in social media. The proposed system mainly comprises '3' phases, namely, data preprocessing, word embedding, and classification. To test and verify the effectiveness of the system, the proposed work uses Reddit dataset. The experimental analysis of the proposed work is done with the existing CNN, RNN, ANN, and RF classifiers. The evaluation is performed with the help of accuracy, precision, recall, f-measure, and Cohen kappa metrics. Herein, the proposed work achieves 96.84% accuracy, 97.85% precision, 95.74% recall, 96.81% f-measure, and 88.98% Cohen kappa, which were high-level outcomes than the existing methods. Thus, the overall experimental analysis shows that the proposed hybrid models gave better results compared to single-layer neural network models and also by detecting suicidal intent in users' posts, the proposed system may help identify individuals who require medical treatment and reduce suicide rates. In future work, this work will be prolonged to increase the accuracy by balancing the dataset using advanced algorithms.

CONFLICT OF INTEREST

I, K.SenilSeby, declares no conflicts of Interest to disclose.

RESEARCH INVOLVING HUMAN AND/OR ANIMALS

This article does not contain any studies with human participants or animals performed by any of the authors.

REFERENCES

1. Ghosal, S., & Jain, A. (2023). Depression and suicide risk detection on social media using fastText embedding and XGBoost classifier. *Procedia Computer Science*, *218*, 1631-1639.
2. Anderson, I. A., & Wood, W. (2021). Habits and the electronic herd: The psychology behind social media's successes and failures. *Consumer Psychology Review*, *4*(1), 83-99.
3. Weinstein, E., Kleiman, E. M., Franz, P. J., Joyce, V. W., Nash, C. C., Buonopane, R. J., & Nock, M. K. (2021). Positive and negative uses of social media among adolescents hospitalized for suicidal behavior. *Journal of Adolescence*, *87*, 63-73.
4. Sawhney, R., Joshi, H., Gandhi, S., & Shah, R. R. (2021, March). Towards ordinal suicide ideation detection on social media. In *Proceedings of the 14th ACM international conference on web search and data mining* (pp. 22-30).
5. Diniz, E. J., Fontenele, J. E., de Oliveira, A. C., Bastos, V. H., Teixeira, S., Rabêlo, R. L., ... & Teles, A. S. (2022, April). Boamente: A natural language processing-based digital phenotyping tool for smart monitoring of suicidal ideation. In *Healthcare* (Vol. 10, No. 4, p. 698). MDPI.

6. Swedo, E. A., Beauregard, J. L., de Fijter, S., Werhan, L., Norris, K., Montgomery, M. P., ... & Sumner, S. A. (2021). Associations between social media and suicidal behaviors during a youth suicide cluster in Ohio. *Journal of Adolescent Health*, *68*(2), 308-316.
7. Yeskuatov, E., Chua, S. L., & Foo, L. K. (2022). Leveraging Reddit for Suicidal Ideation Detection: A Review of Machine Learning and Natural Language Processing Techniques. *International journal of environmental research and public health*, *19*(16), 10347.
8. Haque, R., Islam, N., Islam, M., & Ahsan, M. M. (2022). A comparative analysis on suicidal ideation detection using NLP, machine, and deep learning. *Technologies*, *10*(3), 57.
9. Nikhileswar, K., Vishal, D., Sphoorthi, L., & Fathimabi, S. (2021, October). Suicide ideation detection in social media forums. In *2021 2nd International Conference on Smart Electronics and Communication (ICOSEC)* (pp. 1741-1747). IEEE.
10. Long, S., Cabral, R., Poon, J., & Han, S. C. (2022). A Quantitative and Qualitative Analysis of Suicide Ideation Detection using Deep Learning. *arXiv preprint arXiv:2206.08673*.
11. Liu, J., Shi, M., & Jiang, H. (2022). Detecting suicidal ideation in social media: An ensemble method based on feature fusion. *International journal of environmental research and public health*, *19*(13), 8197.
12. Chatterjee, M., Kumar, P., Samanta, P., & Sarkar, D. (2022). Suicide ideation detection from online social media: A multi-modal feature based technique. *International Journal of Information Management Data Insights*, *2*(2), 100103.
13. Jain, P., Srinivas, K. R., & Vichare, A. (2022). Depression and suicide analysis using machine learning and NLP. In *Journal of Physics: Conference Series* (Vol. 2161, No. 1, p. 012034). IOP Publishing.
14. Rabani, S. T., Khanday, A. M. U. D., Khan, Q. R., Hajam, U. A., Imran, A. S., & Kastrati, Z. (2023). Detecting suicidality on social media: Machine learning at rescue. *Egyptian Informatics Journal*, *24*(2), 291-302.
15. Zhang, T., Schoene, A. M., & Ananiadou, S. (2021). Automatic identification of suicide notes with a transformer-based deep learning model. *Internet interventions*, *25*, 100422.
16. Elgeldawi, E., Sayed, A., Galal, A. R., & Zaki, A. M. (2021, November). Hyperparameter tuning for machine learning algorithms used for arabic sentiment analysis. In *Informatics* (Vol. 8, No. 4, p. 79). MDPI.