

# A Deep Learning Model to Analyse Social-Cyber Psychological Problems in Youth

Ali Alqazzaz<sup>1</sup>, Mohammad Tabrez Quasim<sup>1,\*</sup>, Mohammed Mujib Alshahrani<sup>1</sup>, Ibrahim Alrashdi<sup>2</sup> and Mohammad Ayoub Khan<sup>1</sup>

<sup>1</sup>College of Computing and Information Technology, University of Bisha, 67714, Bisha, Saudi Arabia

<sup>2</sup>Department of Computer Science, College of Computer and Information Sciences, Jouf University, Sakaka, Aljouf, 72388, Saudi Arabia

\*Corresponding Author: Mohammad Tabrez Quasim. Email: tabrezquasim@gmail.com

Received: 08 April 2022; Accepted: 04 November 2022

**Abstract:** Facebook, Twitter, Instagram, and other social media have emerged as excellent platforms for interacting with friends and expressing thoughts, posts, comments, images, and videos that express moods, sentiments, and feelings. With this, it has become possible to examine user thoughts and feelings in social network data to better understand their perspectives and attitudes. However, the analysis of depression based on social media has gained widespread acceptance worldwide, other verticals still have yet to be discovered. The depression analysis uses Twitter data from a publicly available web source in this work. To assess the accuracy of depression detection, long-short-term memory (LSTM) and convolution neural network (CNN) techniques were used. This method is both efficient and scalable. The simulation results have shown an accuracy of 86.23%, which is reasonable compared to existing methods.

**Keywords:** Emotions; depression; social media

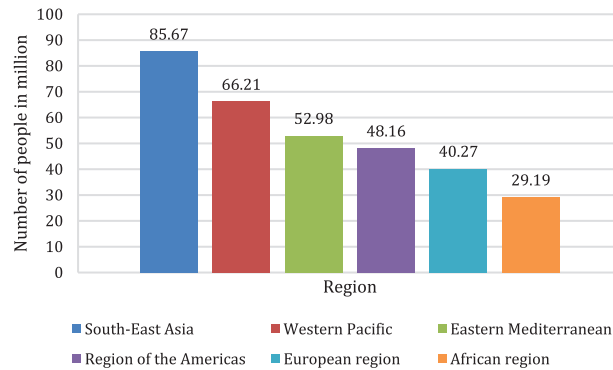
## 1 Introduction

Social network platforms like Facebook, Twitter, and Instagram have re-energized how individuals communicate and engage electronically. Users on social networking platforms such as Twitter, Instagram, Facebook, and the like may post their thoughts, feelings, and opinions about a specific topic or issue in real time. Users can share and comment on any discussion online on social networking platforms. While this could be a drawback but on the other side, it gives health professionals a window into what could be going on in the mind of someone who exhibits certain behaviour in response to specific topics. Using deep learning techniques, it may be possible to examine the particular hidden patterns embedded in the conversation and analyze them to disclose the mental state of use, like anxiety, happiness, sadness, anger, and depression. About 264 million people of all ages are depressed worldwide [1]. Depression is the world's prominent cause of impairment and contributes substantially to the worldwide disease burden as shown in Fig. 1. Research has shown that social network platforms can influence the organization of social interactions, including breakups, mental illness, smoking, sexual assault, and suicidal thoughts. The



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

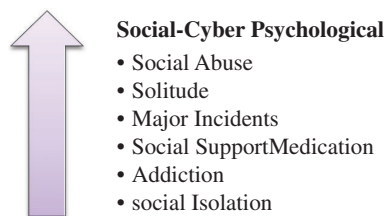
social-cyber psychological problems are the most common issues among youth. Most people don't know that they are suffering from depression and frustration, which leads to committing wrong actions such as suicide, consuming drugs, alcohol, and murder, are a few to name. These bad actions and misbehavior can be treated if we can detect the disease at an early stage.



**Figure 1:** Statistics of depression worldwide as of 2015, by WHO [2]

There are many reasons for these diseases, such as stress, gaming, unemployment, wrong companionship, and high ambition, etc. Out of these, stressful and pleasurable relationships are common triggers for depression. For youth, a pessimistic and stressful or unhappy family environment can affect their self-esteem and could lead to depression and consumption of drugs.

Family relations and conflict also lead to divorce and may result in feelings of insecurity, depression, anger/frustration, rejection, or a sense of loss. At the same time, poverty, unemployment, abuse, and violence within the family are also key contributors to these behaviors. Any abuse, solitude, social support, major life incidents, medication effects, internet addiction, and social isolation can lead to such problems, as shown in Fig. 2.



**Figure 2:** Causes of social-cyber psychological problem

The challenge is how to detect these social disorder problems. We understand that nowadays, every youth is connected with the cyber-world. Recently, social networking platforms have become a tool for social get-together. Every child is using social networking platforms to reflect their lifestyle. The users have different backgrounds, languages, ages, and educational levels. The posts of these users could be analyzed based on the semantics of natural semantics that forecast psychological factors such as frustration, depression, anxiety, involvement in gambling, or any other behavior that leads to social threats. Detecting the social disorder shall be in time to avoid any misshaping. Depression in adolescents is a significant mental health issue that results in a lack of motivation and enjoyment in previously enjoyed activities. This affects perception, feelings, and behavior which can cause cognitive problems. At the same time, depression may occur between adolescents and adults, which may be different. The issues

such as peer pressure, academic standards, family pressure, and changing relationships can lead to many ups and downs for young people. Teen depression is not a weakness but a severe psychological disease that can have dramatic effects and requires long-term counselling. Depression symptoms improve with therapy, including medicine and therapeutic counselling for most teenagers.

In this study, we'll look at Twitter data to see whether there are any signs where Twitter users are suffering from depression. In light of the study's primary goal, the following research issues were addressed in the work. What are the important things to look for in Twitter posts or comments when spotting depression? Is there a way to extract these characteristics from the comments on Twitter regarding depression? How are these things linked? Is it best to interact with a depressed Twitter user at a specific time of day? What are the essential deep-learning techniques for detecting sadness, anger, and depression in Twitter comments? There are several emotional states that Twitter users refer to when they express their feelings in posts or statements, such as happiness, sadness, fear, anger, and shock. We use deep learning algorithms to analyze numerous aspects of Twitter comments and make an overall classification of their various parts. In this study, we examined data from a Twitter page dedicated to bipolar depression and anxiety disorders that were made available to the public. The objectives of this research are as follows:

The primary objective is to investigate the technological sources to collect data for social-cyber psychological problems. Understand the scope, constraints, and format of collecting such data from devices. Another objective is to study and develop a deep learning algorithm for social-cyber psychological problems to improve civilization by building a knowledge base and advisories for society. The objective is to build a learning model based on a lexicon to detect social cyber psychological problems based on analyzing posts and behaviour in cyberspace to predict the traits.

The paper is organized as follows: Section 2 represents the background and related work. Section 3 details the proposed framework and deep learning algorithms. Section 4 presents a discussion of the results obtained from the simulation. Finally, concluding remarks are presented in Section 5.

## 2 Related Work

Trotzek et al. have emphasized that political, religious, and social instability triggers extremism among people, which can be seen as a social media thought written in English, the most popular language for expressing views on social media [3]. In this work, the authors have focused on dynamic analysis and multilingual textual data on social media to find the extreme strength of sentiment. Authors have categorized the incorporated texts as high, low, moderate, and neutral extremes. Authors claim that the lexicon is verified by subject experts, which have achieved 88% accuracy during the validation phase [3].

Wang et al. have focused on depression, a primary reason for suicide that can be analyzed through the natural language analysis of the written text [4]. In this work, the authors have used Reddit posts to determine the parameters that indicate depressions. Authors have used natural language processing (NLP) and machine learning (ML) methods to train the data and validate the efficiency of a lexicon frequently used by depressed people. Authors have utilized the best bigram and support vector machine (SVM) classifier to find the depression with 80% accuracy and 0.80 F1 score [4].

Shuai et al. have focused on the reason for depression and detection methods to avoid suicide [5]. Depression has been demonstrated to affect language use, and many depressed people turn to social media or the internet in general to get information or to deal with their issues. Neural networks have been used in machine learning models to detect depression in social media messages.

Syarif et al. has brought another exciting knowledge and relation of dietary supplements leading to mental disorders [6]. Twitter postings were analyzed in this study to find a link between dietary

supplement usage and mental health issues. The authors have seen several significant variations in depression, anxiety, and negativity levels among the two groups. The authors concluded that stress, despair, and mood swings are among the disorders that may be linked to consuming nutritional supplements [6].

Seah et al. have devised an exciting terminology for social network mental disorders (SNMDs), which includes an obsession with cyber relationships and net obsessions and compulsions, which are few to name [7]. Authors have argued to mine online social behaviour that can identify SNMDs at the earliest. Using a machine learning system that extracts attributes from social network data, the suggested study correctly identified the prospective victims of SNMDs. A multisource learning-based Tensor Model (STM) has been evaluated on 3,126 online social network users, which has achieved promising results for identifying users who are likely to have SNMDs.

Oyong et al. have presented a study on data extraction from social media to analyse mental disorders [8]. Computational language processes with linguistic and emotional aspects have been evaluated by the authors using Twitter as a source. SenticNet's four-dimensional emotional state of the main characteristics, such as self-reference and mental illness word count, has been proposed by authors.

Al-Ajlan et al. have presented a case study of Singapore to determine the suicide using social media [9]. Social sensing with digital traces has been obtained from Reddit to analyze the posts and comments posted in a state of depression leading to suicide.

Yadav et al. have focused on NLP and linguistic methods for depression analysis in Indonesia [10]. Weighting and depression score computations have been done using frequency lexicons and natural language processing on text resources. Text processing and Twitter-specific processing yielded an F1 score of 0.47 and a sensitivity value of 0.89 when the threshold was set to 0.

Lali et al. devised cyberbullying, a crime involving harassment and hate [11]. In this work, the authors have optimized Twitter cyberbullying analysis using deep learning and semantics of the word. The applied classification methods, deep understanding, and metaheuristic optimisation method tuning parameters have achieved good results.

In other existing literature [12–15], all works are focused on sentiment analysis but on other issues such as offensive text analysis or healthcare issues using social media only.

### **3 Proposed Framework**

The proposed framework consists of three major components as data aggregation, assessment and prediction, and expert data management system, as shown in Fig. 3. The framework suggests that data shall be collected from different possible sources. The number of sources plays a vital role in determining the accuracy. The second essential component is the assessment and prediction system. This is the central part of the framework, which consists of data extraction, text preprocessing, feature extraction from the dataset, and applying various deep learning algorithms, the expert database.

#### **3.1 Data Collection**

This component has data available on social media platforms, mobile phones, laptops, etc. The purpose of this component is to collect maximum data from the available platforms. Using a hybrid approach, data is extracted from Twitter, and some manual techniques are done. The lexical marks such as lonely, depressed, depressed, suicide, kill, anxiety, harm, anxious, threat, antidepressants, and hopeless have been used to extract data from Twitter. Moreover, some dataset has been manually collected from known users with Twitter accounts. The attributes of the dataset are shown in Table 1.

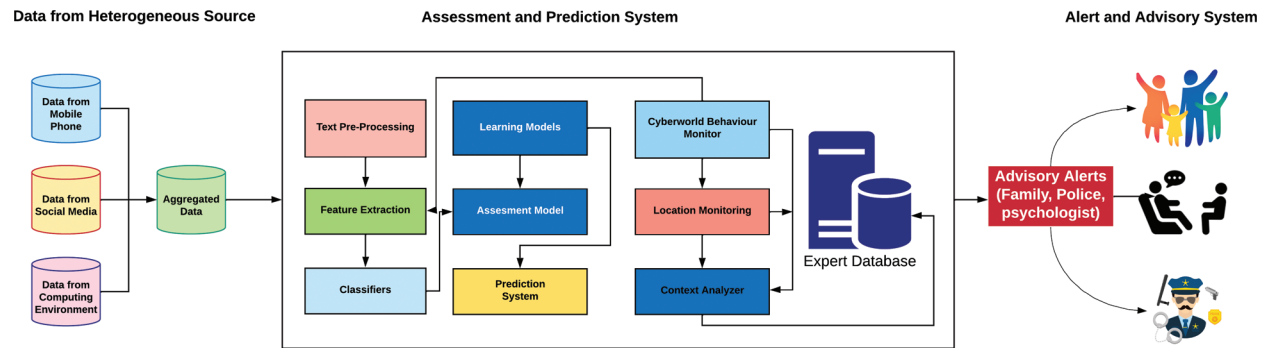


Figure 3: Proposed framework for social cyber problem prediction system

Table 1: Dataset attributes

User	String
Location	String
URL	Mixed
ID	Long
Hashtags	String
User mentions	String
Symbols	String

The unwanted lexicon has been removed, and a sparse matrix has been set to increase the efficacy of the model. We have conducted different Twitter scrapping functions in parallel to get the dataset for other lexical markers and finally merged it to get the large amount of dataset. An overview of the dataset has been shown in Fig. 4.

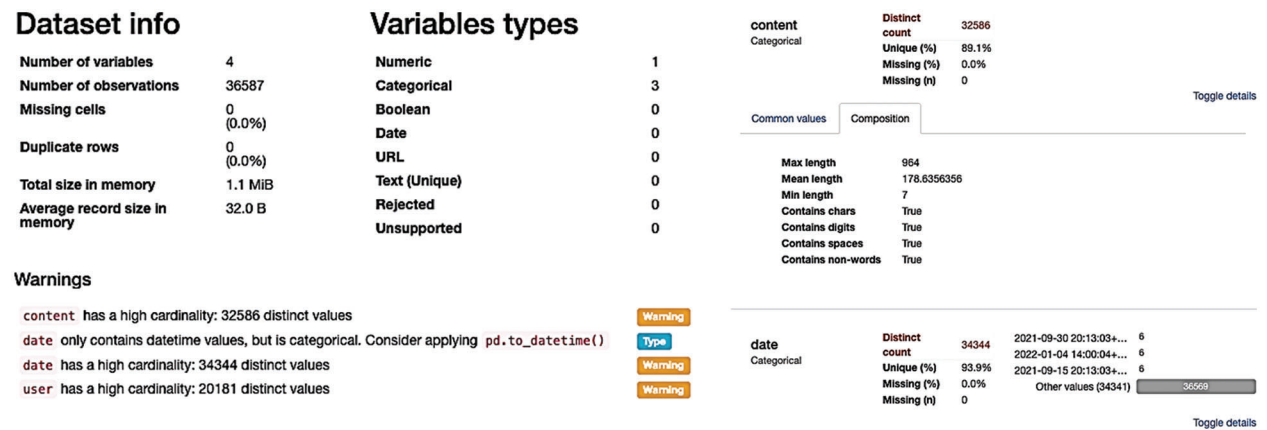


Figure 4: Dataset information

### 3.2 Prediction and Advisory System

- a) Pre-processing–This component will be responsible for processing the extracted data for noise removal, duplicate, and missing values.
- b) Feature extraction–This component will be responsible for extracting meaningful features from the collected data set, which will tune the performance of the algorithm.
- c) Classifier–This component will classify the extracted features based on prior knowledge and a trained data set.
- d) Machine learning and Prediction System–This component predicts the social-cyber psychological problems based on the trained data accuracy.
- e) Behavioral Monitoring/Location and Context Analyzer–This component will do continuous monitoring of the behaviour in the cyber world along with the context and location.
- f) Advisory Alters–This component will generate an advisory alter from the prediction system based on the classified criteria. The alert could be sent to family doctors, family members, and police, etc.

### 3.3 Methodology

Assume that  $P = \{p_1, p_2, p_3, \dots, p_n\}$ , is the set of web pages that contains the, post, comments or discussions. Generally, each page consists of textual data or a collection of posts written by a user. Usually, a page or a different representation of the page is considered as a single training instance during the training. Each page can be represented by Eq. (1).

$$\forall p_i, i \in P \quad (1)$$

Assume that the set of unique keywords that occur in the web page  $P$ . Thus, the set of keywords can be represented by Eq. (2).

$$K = \{k_1, k_2, k_3, \dots, k_n\} \quad (2)$$

The page  $p_1$  is the vector of size  $m$ , and  $f_i$  is he frequency of the keywords  $k_i$  in the web page  $p_i$ . Now, we can represent the  $P$  in a matrix dimension of  $n \times m$  as in Eq. (3).

$$\begin{bmatrix} f_{11} & f_{12} & f_{13} & \dots & f_{1m} \\ f_{21} & f_{22} & f_{23} & \dots & f_{2m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ f_{n1} & f_{n2} & f_{n3} & \dots & f_{nm} \end{bmatrix} \quad (3)$$

In the above matrix, if the keyword is present then  $f_{ij} = 1$  else  $f_{ij} = 0$ . Another common variation is to pass these frequencies through the function  $g(\cdot)$  as shown in Eq. (4).

$$g(f) \quad g_1(f) = 1 + \log_2(f) \quad \text{and} \quad g_2(f) = g_1(f) \times \log\left(\frac{n}{n_k}\right) \quad (4)$$

where  $n_k$  is the number of webpages in which the keyword  $k$  occurs.

### 3.4 Classifier

A model for depressive post comment recognition is built using psycholinguistic factors as input in this step. Taking into account our training dataset,  $\forall p_i \in P$ . Each  $p_i$  labelled as depressive or non-depressive. Therefore, the classifier function can be defined in Eqs. (5) and (6).

$$f: P \in L, L = l_1 | l_2 \quad (5)$$

$$f(p) \quad (6)$$

In regard to classification, the convolution neural network (CNN) is the prominent and effective model, and it consistently achieves the best results [16]. CNN has many connected weights. In the CNN, multiple maps can be merged as the output of the convolution. The CNN layers can be represented as shown in Eq. (7):

$$x_j^l = f\left(\sum_{i \in M_j} x_i^{l-1} * k_{ij}^l + b_j^l\right) \quad (7)$$

The selection of input maps is expressed using symbol  $M_j$ . The recurrent neural networks (RNN) utilise previous experiences based on internal recurrence. Unlike feed-forward networks, RNNs automatically project the dynamic aspects of the system. Thus, it becomes more powerful computationally for findings of chaotic time series prediction. It is possible to use RNN models with long short-term memory (LSTM) to process signals with low-and high-frequency components that have been delayed for a long time [17]. It takes a long time for RNN models to train because of their recurrent network design. In the proposed framework, first, word vectors of vocabulary terms are learned from a large Twitter dataset. To split a text into  $R = \{r_1, r_2, \dots, r_K\}$  regions, the regional CNN model divides each piece of input text into a single phrase. Once the word vectors have been consecutively passed through a convolutional layer and a max pooling layer, meaningful affective characteristics may be retrieved from each area. This text vector for disorder prediction is constructed by successively integrating regional information across regions using LSTM. A convolutional layer and max pooling layer may be used to extract a valuable emotional statement from the word vectors in each area. LSTM is used to generate a text vector for predicting psychiatric disorders by integrating these local variables across regions. The  $n$ -gram characteristics of each area are initially extracted using a convolutional layer. A region matrix has all word embeddings. The region matrix can be defined as  $M \in \mathbb{R}^{d \times |V|}$  where,  $|V|$  is the vocabulary size while  $d$  is the dimension of the word vector. The convolution filter can be denoted by  $L$ . The feature map can be defined as one as shown in Eq. (8):

$$y_n^l = f(W^l \circ x_{n:n+w-1} + b^l), W \in \mathbb{R}^{w \times d} \quad (8)$$

The  $b$  is the weight matrix, while  $w$  is the length of the filter. The rectified linear activation function is represented by  $f$ . Using the max-pooling, the convolutional layer's output may be subsampled. The most popular method for pooling data is using a max operation on each filter's output. Using a max-pooling layer makes sense for two reasons. In the first place, by removing non-maximal values, it decreases the computational complexity for higher layers [17]. In regard to data extraction, it is able to focus on local dependencies in distinct locations. After that, a sequential layer is given the region vectors. For each LSTM memory block, there is a cell state and a set of gates that include the input, forget, and output gates, which can be represented by Eq. (9):

$$I^t = \sigma(W_{LI}L_t + W_{HI}H_{t-1} + b_I) \quad (9)$$

$W$  represents the weight matrix,  $b$  is the bias vector, and  $\sigma$  is the logistics function. Now, we can represent the forget gate  $F$  as shown in Eq. (10):

$$F^t = \sigma(W_{LF}L_t + W_{HF}H_{t-1} + b_F) \quad (10)$$

The long-term memory is memorized by the state vector  $S$  as shown by Eq. (11):

$$S^t = F^t \cdot S^{t-1} + I^t \cdot (W_{LS}L_t + W_{HS}H_{t-1} + b_S) \quad (11)$$

The output gate, which generates the output can be represented by Eq. (12):

$$O^t = \sigma(W_{LO}L_t + W_{HO}H_{t-1} + b_O) \quad (12)$$

The hidden state  $H$  and attention layer can be represented by Eqs. (13) and (14):

$$H_t = O_t \tanh(S_t) \quad (13)$$

$$A(att)_t = LSTM(H_t, A(att)_{t-1}) \quad (14)$$

As a general rule, the attention approach is utilized to emphasize relevant information in the current activity instead of unnecessary information. Consequently, it may be used above the LSTM layer to improve the model's accuracy. Finally, as shown in Eq. (15), a softmax function can be used to decide the output value.

$$O = \text{softmax}(W_O A_O + b_O) \quad (15)$$

#### 4 Experimental Result and Analysis

In this work, we look at how well different depression detection classifiers perform when run in a shorter amount of time. With the training dataset, we build the CNN-LSTM model for qualitative parameters analysis and to evaluate the model's performance. We are looking towards an efficient implementation of the learning method that might executed on graphics processing units. Our trials are carried out on a Linux machine equipped with a graphics card and 16 GB memory, and Python 2.7 framework. The common words such as depression, anxiety, suicide, lonely, depressed, hopeless, mental health, and antidepressants, which are the frequency found in the tweets are shown in Fig. 5

Bigram has been utilized to tokenize comments, posts or tweets by a word that is important for the kind of sentiment and frequency analysis that is performed, as shown in Fig. 6.

In this study, we employed LR, KNN, SVM, and CNN-LSTM classifier algorithms for depression detection of emotional phrases to gain a better understanding of the general intuition behind depression. We demonstrated that all of these classification systems, which are based on language style, emotional process, temporal process, and all other variables, are capable of successfully extracting the emotional outcome associated with depression. The results of several characterizations, including precision, recall, F1, and accuracy, are shown in Table 2. This can be noted that CNN-LSTM produces a superior result in this case. As a result, we believe that this study has laid the groundwork for future research on inference and the discovery of additional information based on cause-event relationships, such as the detection of implicit emotions or causes, as well as the prediction of public opinion based on cause events, among other things. The linear regression has achieved an accuracy of 59.81, precision of 60.2, recall of 59.12, and F1 of 59.12. While as the SVV has obtained an accuracy of 60.62, precision of 97, recall of 61, and F1 of 75. The KNN has achieved an accuracy of 65.25, precision 77, recall 69, and F1 of 52. The decision tree algorithm has achieved an accuracy of 73, precision 58, remember 99, and F1 of 74. The lexical approach [10] has achieved an accuracy of 85, precision 57, recall 44, and F1 of 50, while our proposed method based on CNN and LSTM has received an accuracy of 86.23, precision 87, recall of 83, and F1 of 85. All results of various models have been shown in Table 2 for comparative analysis. Nevertheless, there is still potential for improvement. It is crucial to highlight that this study does not identify individuals who are suffering from depression; instead, it examines Twitter comments for signs of depression.



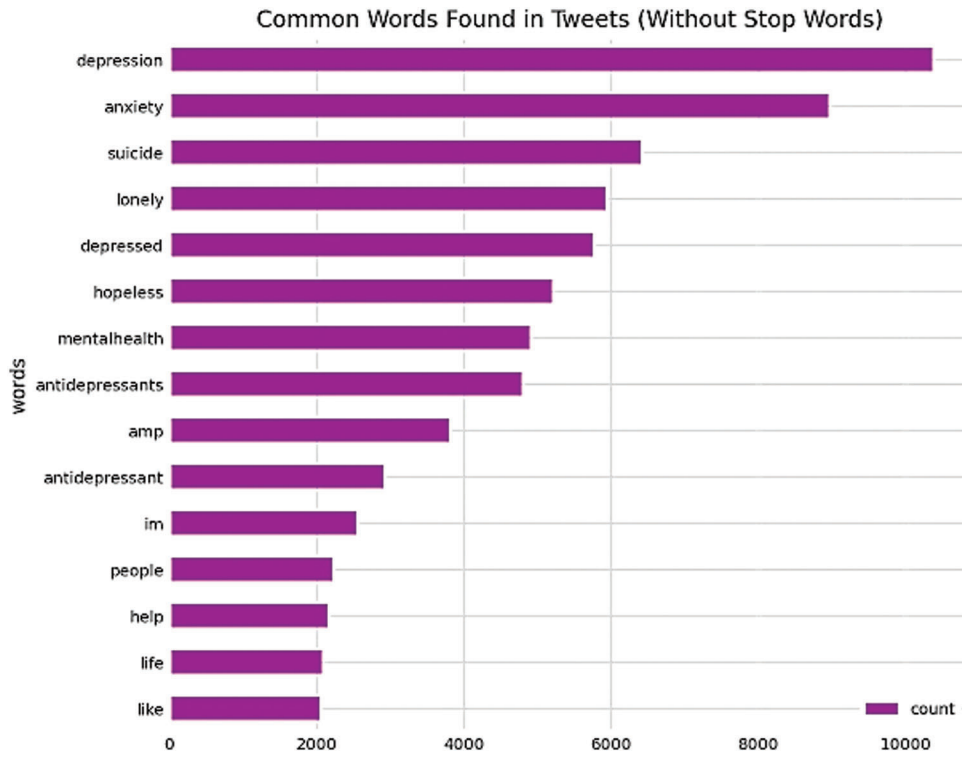


Figure 5: Common words in tweets

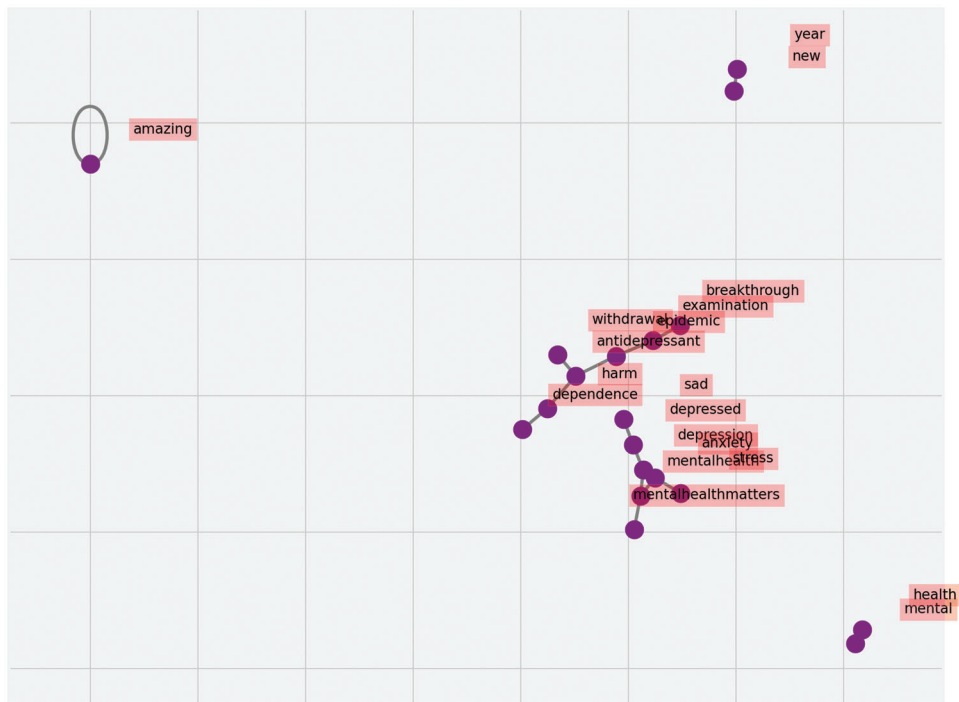
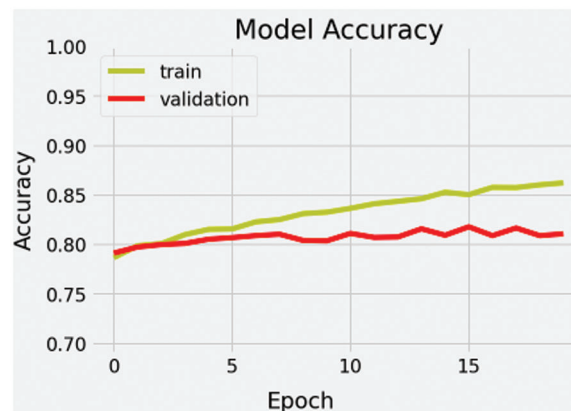


Figure 6: Bigram analytics

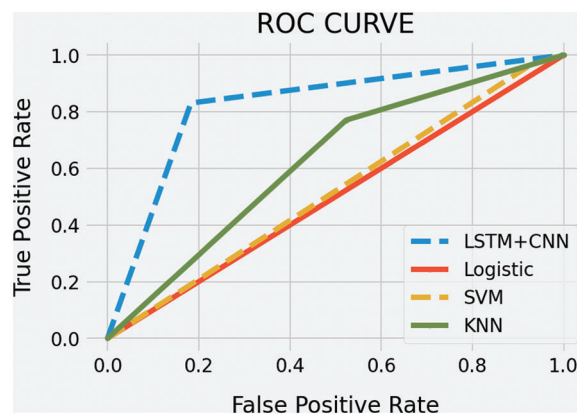
**Table 2:** Comparative analysis of performance

Technique	Precision	Recall	F1	Accuracy
Logistics regression	60.2	59.3	59.1	59.8
Support vector machine	97	61	75	60.6
K-nearest neighbors	77	69	52	65.2
Decision tree	58	99	74	73
Lexical approach [10]	57	44	50	85
CNN-LSTM (proposed)	87	83	85	86.2

We can observe the model accuracy of training and validation as shown in Fig. 7. The accuracy of the model increases with more epochs.

**Figure 7:** Model accuracy

The receiver operating characteristic (ROC) curve for KNN, SVM, logistic regression, and LSTM-CNN has been shown in Fig. 8. The LSTM-CNN has achieved a better ratio of true positive vs. false positive that shows the correctness of classification.

**Figure 8:** ROC Curve

## 5 Conclusion

This research has demonstrated the feasibility of employing Twitter to assess and identify significant depression among its users. To provide a clear understanding of the study, a list of research problems was included. Analysis of the selected dataset offers some insight into the research difficulties that have been identified. Depression may strike at any time and affect everyone. On the other hand, some stages or situations make us more susceptible to depression. Growing up, losing a loved one, starting a family, and retiring can all cause physical and mental changes that, in certain cases might result in depression for a large percentage of the population. What are the things to consider while looking for signs of depression on Twitter? When we post tweets on Twitter, it processes them on a “line-by-line” basis within and across the columns of a dataset. When we access a single text within a dataset, it analyses each line sequentially. It reads one target word at a time to determine which variables are associated with positive or negative attitudes regarding depression.

In this study, we reviewed 36587 Tweets to determine depression. The accuracy received for the proposed CNN-LSTM was 86.2%, which is superior than the existing models. In the future, we will deploy the model to take the data from other sources.

**Acknowledgement:** The authors extend their appreciation to the Deanship of Scientific Research, University of Bisha, Saudi Arabia for funding this research work through Promising Initiative Project under Grant Number (UB-Promising-01-1442).

**Funding Statement:** This project was funded by Deanship of Scientific Research, University of Bisha, Bisha, Kingdom of Saudi Arabia.

**Conflicts of Interest:** The authors declare that they have no conflicts of interest to report regarding the present study.

## References

- [1] M. Asif, A. Ishtiaq, H. Ahmad, H. Aljuaid and J. Shah, “Sentiment analysis of extremism in social media from textual information,” *Telematics and Informatics*, vol. 48, no. 101345, pp. 1–20, 2020. <https://doi.org/10.1016/j.tele.2020.101345>.
- [2] M. M. Tadesse, H. Lin, B. Xu and L. Yang, “Detection of depression-related posts in reddit social media forum,” *IEEE Access*, vol. 7, pp. 44883–44893, 2019.
- [3] M. Trotzek, S. Koitka and C. M. Friedrich, “Utilizing neural networks and linguistic metadata for early detection of depression indications in text sequences,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 32, no. 3, pp. 588–601, 2020.
- [4] Y. Wang, Y. Zhao, J. Bian and R. Zhang, “Detecting signals of associations between dietary supplement use and mental disorders from Twitter,” in *2018 IEEE Int. Conf. on Healthcare Informatics Workshop (ICHI-W)*, New York, N.Y, pp. 53–54, 2018.
- [5] H. H. Shuai, C. Y. Shen, D. N. Yang, Y. Lan, W. C. Le *et al.*, “A comprehensive study on social network mental disorders detection via online social media mining,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 30, no. 7, pp. 1212–1225, 2018.
- [6] I. Syarif, N. Ningtias and T. Badriyah, “Study on mental disorder detection via social media mining,” in *2019 4th Int. Conf. on Computing, Communications and Security (ICCCS)*, Rome, Italy, pp. 1–6, 2019.
- [7] J. H. K. Seah and K. J. Shim, “Data mining approach to the detection of suicide in social media: A case study of Singapore,” in *2018 IEEE Int. Conf. on Big Data (Big Data)*, Seattle, WA, USA, pp. 5442–5444, 2018.
- [8] I. Oyong, E. Utami and E. T. Luthfi, “Natural language processing and lexical approach for depression symptoms screening of Indonesian Twitter user,” in *10th Int. Conf. on Information Technology and Electrical Engineering (ICITEE)*, Kuta, Bali, pp. 359–364, 2018.

- [9] M. A. Al-Ajlan and M. Ykhlef, "Optimized twitter cyberbullying detection based on deep learning," in *21st Saudi Computer Society National Computer Conf. (NCC)*, Saudi Arabia, pp. 1–5, 2018.
- [10] S. H. Yadav and P. M. Manwatkar, "An approach for offensive text detection and prevention in social networks," in *ICIIECS 2015–2015 IEEE Int. Conf. on Innovations in Information, Embedded and Communication Systems*, Coimbatore, India, pp. 1–4, 2015.
- [11] M. I. U. Lali, R. U. Mustafa, K. Saleem, M. S. Nawaz, T. Zia *et al.*, "Finding healthcare issues with search engine queries and social network data," *Int. J. Semant. Web Inf. Syst.*, vol. 13, no. 1, pp. 48–62, 2016.
- [12] M. Barnidge, B. Kim, L. A. Sherrill, Ž Luknar and J. Zhang, "Perceived exposure to and avoidance of hate speech in various communication settings," *Telemat. Informatics*, vol. 44, pp. 101263, 2019.
- [13] S. Haider, M. Tanvir Afzal, M. Asif, H. Maurer, A. Ahmad *et al.*, "Impact analysis of adverbs for sentiment classification on Twitter product reviews," *Concurrency Computation*, vol. 33, no. 4, pp. 1–15, 2018. <https://doi.org/10.1002/cpe.4956>.
- [14] Y. Le Cun, L. Bottou, Y. Bengio and P. Haffner. "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 111, pp. 2278–2324, 1998.
- [15] K. Greff, R. K. Srivastava, J. Koutnik, B. R. Steunebrink and J. Schmidhuber, "LSTM: A search space odyssey," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 10, pp. 2222–2232, 2017.
- [16] M. R. Islam, M. A. Kabir, A. Ahmed, A. Kamal, H. Wang *et al.*, "Depression detection from social network data using machine learning techniques," *Health Information Science System*, vol. 6, no. 1, pp. 1–12, 2018. <https://doi.org/10.1007%2Fs13755-018-0046-0>.
- [17] M. Z. Uddin, K. K. Dysthe and A. Følstad, "Deep learning for prediction of depressive symptoms in a large textual dataset," *Neural Computing & Application*, vol. 34, pp. 721–744, 2022.