Applying Deep Learning Technique for Depression Classification in Social Media Text

Hussain Ahmad[1], Muhammad Zubair Asghar[1], Fahad M. Alotaibi[2], Ibrahim A Hameed[3*]

[1]Institute of Computing and Information Technology, Gomal University, D.I. Khan, Pakistan

[2]Faculty of Computing and Information Technology (FCIT), King Abdulaziz University, Jeddah, Saudi Arabia.

[3*]Department of ICT and Natural Sciences, Faculty of Information Technology and Electrical Engineering, Hovedbygget, B316, Ålesund, Norway

Corresponding Author: Ibrahim A Hameed

Email of author#1: hussainmarwat@yahoo.com

Email of author#2: zubair@gu.edu.pk

Email of author#3: fmmalotaibi@kau.edu.sa

Email of author#4 (corresponding author): ibib@ntnu.no

Abstract

In social media, depression identification could be regarded as a complex task because of the complicated nature associated with mental disorders. In recent times, there has been an evolution in this research area with growing popularity of social media platforms as these have become a fundamental part of people's day-to-day life. Social media platforms and their users share a close relationship due to which the users' personal life is reflected in these platforms on several levels. Apart from the associated complexity in recognising mental illnesses via social media platforms, implementing supervised machine learning approaches like deep neural networks is yet to be adopted in a large scale because of the inherent difficulties associated with procuring sufficient quantities of annotated training data. Because of such reasons, we have made effort to identify deep learning model that is most effective from amongst selected architectures with previous successful record in supervised learning methods. The selected model is employed to recognise online users that display depression; since there is limited unstructured text data that could be extracted from Twitter.

Keywords:

1. Introduction

Depression is considered as a mental disorder, and can be described as a ‘syndrome pertaining to an individual’s sadness, and anxiety, indicating a dysfunction with regards to the, psychological, or biological processes defining mental functioning’ (Gaspersz, et al., 2018, American Psychiatric Association, 2013). Studies pertaining to anxiety and depression have associated depression to different issues like high tendency of committing suicide and increased risk of cardiovascular dysfunction (Sonawalla, & Fava, 2001). As per statistics by World Health Organisation (2014), almost 20% of adolescents and children have had faced depression and other mental disorders.

Even though mental disorders are considered severe impacting one’s physical and mental health, the social stigma (such as there is no cure for mental disorders) or discrimination has resulted in overlooking of individuals by the community and discourage from taking necessary treatments. The literature describes about the inherent complexity pertaining to identifying depression via social media platforms, in which numerous researchers have put effort to highlight the key features by employing various machine learning approaches (Orabi et al. 2018).

Adequate amount of knowledge associated with the particular area of research needs to be acquired in order to facilitate extraction key characteristics to build an effective prediction technique. Even on extracting such features, it does not guarantee that those features have a key role in contributing to achieving higher accuracies. Because of this, we have evaluated the possibility of employing deep neural architectures since such features can be learned in the architecture itself (Kumar et al. 2019).

Here, we have assessed certain deep learning models in order to identify depression. We employed the textual content that have been acquired from Twitter in the form of benchmark dataset. The task includes binary classification of Twitter posts into normal vs. depressed. Our main goal is to identify depression by employing the deep learning technique based on BiLSTM method, provided the textual data of available text stream.

The current work is motivated by the similar works (Kumar et al. 2019, Orabi et al. 2018) conducted on detecting and classifying depression in Twitter posts. The key limitations of the prior studies are listed as follows: (i) The early studies on depression classification are based on lexical resources (Kumar et al. 2019) , and (ii) current lexicons have limited entries for depression-oriented terms. However, in last few years, deep learning neural networks have produced encouraging results for classifying disease-related information (Orabi et al. 2018). We take the problem of depression detection as a binary-classification task. We take the training set Tr=$\{tr_1,tr_2,tr_3....tr_n\}$, and class label depression=*{ yes, no}*. A depression-related label (depression vs normal) is allocated to each tweet. The objective is to develop a predicative model, trained on the training corpus and can classify a tweet as *normal* or *dressed.*

In this study, we address three research questions: *(i) RQ1: How to classify user's tweet as "depressed" or "normal", by applying deep learning approach? (ii) RQ2: How can we evaluate performance of of the Proposed Bi-LSTM (Deep Learning model) in connection with Traditional Machine Learning Classifiers? and (iii) RQ3*: *How can we evaluate performance of of the Proposed Bi-LSTM (Deep Learning model) in connection with other Deep Learning Classifiers?*

Summarisation of proposed technique as well as key contributions could be done as follows.

- *Depression Classification:* Classification of tweets into *normal* vs *depressed* using deep learning model
- Word embedding based feature engineering: we put forward a method to utilize word-embedding based feature engineering pertaining to depression classification with emphasis to recognise online users that are having depression with regards to their tweets. We have employed our technique to for detecting depression on the Twitter dataset.

• Performance evaluation: We evaluate and report the performance pertaining to various deep learning models usually employed in sentiment analysis tasks, specifically to identify depression. We have *also* expanded our study to incorporate various word embeddings as well as hyperparameter tuning.

2. Related Work

Recognising the cure needed for depression and anxiety is regarded to be a complex health-care phenomenon pertaining to particular treatments, patients' suffering with regard to the symptoms and disabilities associated with the patients' symptoms (American Psychiatric Association, 2013). It should also be noted that measuring the disorder's severity is challenging and can be achieved by a highly trained professional that can implement various techniques like clinical interviews and text descriptions, and also their judgements (American Psychiatric Association, 2013). By taking into account the complexity associated with the procedures as well as the level of skills needed to detect depression as well as the required treatments, identifying depression within social networking sites by employing sentiment analysis as well as emotion detection techniques could be regarded as a initial step for creating awareness. Respecting ethical aspects regarding the utilisation of social networks textual and visual data as well as its privacy is important. The research scientists who are dealing with such type of social networks data need to be wary about protecting the users' privacy as well as their ethical privileges in order to prevent depression and distress. In some studies necessary steps are taken to anonymise the social network data for the security of user privacy. Coppersmith et al. (2015b) employed an approach based on a whitelist technique to anonymise the data fed of participants. Although anonymisation has been maintained for screen names as well as URLs employing salted hash functions, there still exists the chance that could result in breach of user privacy. For such reason, a confidentiality agreement was signed by the researchers to ensure the data privacy.

Since social media interactions are associated with a more natural setting, it is also essential to identify to extent to which an individual has revealed regarding its personal information, and if the information being published is accurate and adequate that would help in determining whether the person is suffering from a depression. It has been recognised that the longitudinal data that have been published on social media platforms are valuable (De Choudhury, 2013, 2014, 2015) to a considerable level of self-disclosure (Balani and De Choudhury, 2015; Park et al., 2012).

Majority of the prior studies on identifying depression via social media networks have concentrated mostly on classical feature engineering techniques. While reviewing literature, it has been noticed that Linguistic inquiry word count (LIWC) lexicon has been widely used as a feature engineering resource and technique to detect and extract lexical characteristics. This lexicon includes more than thirty-two types of psychological traits (Pennebaker et al., 2007). The use of lexicons is regarded as a key feature extraction mechanism to recognise, depression (Schwartz et al., 2014, De Choudhury et al., 2013 Coppersmith et al., 2014a). In order to identify depression, scientists had no option but to identify characteristics that overlapped mutually, and also important for depression. For instance, using 2^{nd} and 3^{rd} person pronouns were used lesser than the first-person pronouns (Lehrman et al., 2012; De Choudhury, 2013) in order to identify persons that were likely suffering from depression and distress.

It was observed that it is challenging to work with the data that have been acquired from the social network sites because the users post text in an unstructured manner. When a message is composed on Twitter, it gets introduced with misspelled words, new terminologies, syntactical omissions, and limited character length. An intuitive method to deal with the challenges associated with unstructured data is Character n-gram models. By accounting for the effectiveness pertaining to aforementioned language models during classification that involve employing Twitter content, Coppersmith et al. (2014a,b) utilised different models like unigram and character n-gram models in order to harvest characteristics when recognising users suspicious of suffering from different mental problems like depression. In a similar manner, character n-grams have been regarded as a pivotal feature extraction technique to identify mental disorder like depression (Coppersmith et al., 2015a, Mitchell et al., 2015). Although approaches based on topic modelling like LDA (latent Dirichlet allocation) have been employed widely to improve the prediction capability of classifier (Mitchell et al., 2015), researchers also regard supervised learning-based topic modelling techniques (Resnik et al., 2015) as well as topics extracted based on clustering approaches like GloVe Word Clusters and Word2Vec (Preotiuc-Pietro et al., 2015b) to be more dependable when classifying online users that could be suffering from a depression. In addition to lexical features like word n-grams and character n-grams as well as syntactic characteristics like POS-tagged words, user’s interactions on social media like retweet rate and posting frequency, as well as other details like age, sex and personality (Preotiuc-Pietro et al., 2015a) were regarded as key indicators to identify depression.

In general, there has been an evolution in the research pertaining to detection of depression using lexicon-driven techniques to topic and machine learning. The latest research was aimed at improving models' performance by employing vector space representations as well as convolutional neural network model with bi-directional mechanism (Kshirsagar et al., 2017) to identify and describe posts depicting health hazards. In our research, we used a model that yielded promising results to identify depression from user's Twitter messages by skipping the need for any exhaustive feature engineering.

3. Proposed Methodology

A considerable performance improvement has been shown by the Deep learning model namely, Long Short-Term Memory (LSTM) in terms of capturing both semantic and syntactic information. However, it can only preserve the past information and can't utilize the future context. The Bidirectional LSTM is an extension of the traditional LSTM and is able to utilize related information from both the previous and future context. In the next section, we exploit the BiLSTM model for depression classification on a given input text (Cai et al. 2019).

The Simple Recurrent Neural Network (RNN) can retain the contextual information just for a small interval of time. This issue can be handled by introducing LSTM, which has a memory block instead of simple RNN unit. The LSTM allows to capture the past information (Vinayakumar et al. 2017). [Bidirectional Long Short-Term Memory (BiLSTM) is a type of recurrent neural network, which can process data in two directions: previous and next, because it works with two hidden layers. This is the main point of difference with LSTM. The BiLSTM has exhibited promising results in natural language processing applications like text processing, named entity recognition and others (Rhanoui et al. 2019) . Therefore,, we choose to employ a variant of the Simple RNN, calleds Bi-LSTM, to classify input text into two classes, namely: "depressed" or "normal".

Given a sequence of words $w = [w_1, w_2, w_3, \dots \dots, w_n]$, where n is the number of words, the BiLSTM layer calculates the forward and backward hidden vectors and generates an output vector which is denoted as $\overleftrightarrow{h}$. The architecture of the BiLSTM is shown in Fig. 1 (Zhang et al. 2018) [

Table 1 shows mathematical representation of both Forward and Backward LSTM used for classifying tweets as *"normal" or "depressed"*.

Table 1 Mathematical representation of the proposed Forward and Backward LSTM model

Forward LSTM Equations	**Backward LSTM Equations**	**Description**
$i_t = \sigma(W_i.[h_{t-1}, x_t] + b_i)$	$i_t = \sigma(W_i.[h_{t+1}, x_t] + b_i)$	• i shows input gate • f shows forget gate • o shows output gate • h_t is the hidden state • c_t is the current cell state • b is called bias • W denotes the weight metrics
$f_t = \sigma(W_f.[h_{t-1}, x_t] + b_f)$	$f_t = \sigma(W_f.[h_{t+1}, x_t] + b_f)$	
$o_t = \sigma(W_o.[h_{t-1}, x_t] + b_o)$	$o_t = \sigma(W_o.[h_{t+1}, x_t] + b_o)$	
$c\sim_t = \tau(W_c.[h_{t-1}, x_t] + b_c)$	$c\sim_t = \tau(W_c.[h_{t+1}, x_t] + b_c)$	
$c_t = f_t.c_{t-1} + i_t.c\sim_t$	$c_t = f_t.c_{t+1} + i_t.c\sim_t$	
$h_t = o_t.\tau(c_t)$	$h_t = o_t.\tau(c_t)$	

For depression classification task, we used the Anaconda based Jupyter notebook to implement Bidirectional LSTM with the Keras Deep learning library. For example, Fig. 1 shows the classification of input text: "I'*m just so tired of this. my body is tired, my mind is a mess. i just really want to lay in bed and never get up. i'm just so tired of life."* as "depressed".

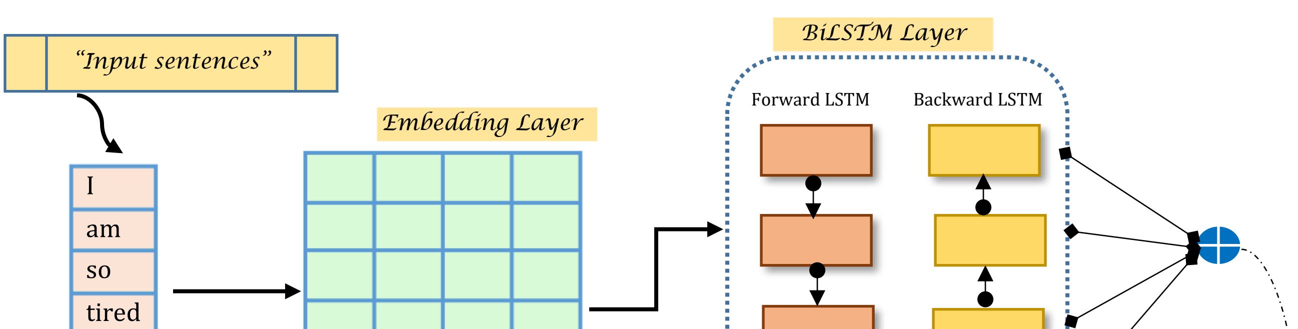

Fig. 1 Deep learning model for Depression Classification

3.5 Three-way split (Train, test and validation split)

To develop a deep learning model, the first step is to partition the depression-related dataset. We adopted the dataset used by (Orabi et al. 2018) for depression classification. We used a three- way split scheme by partitioning of the data into train, test and validation sets (Khan et al. 2019)). Their description is given as follows:

Train set: This sub-part of the dataset learns or uncover s the relationship between the input features and the target variables. We have used 80% of the data as a train set.

Validation set: This is another sub-part of the dataset, used for performing error analysis, model performance and parameter tweaking. It also identifies how accurately the model learns the relationship between the input and target outcome. We used 10% validation set (Asghar et al. 2019a).

Test set: This subset provides a final estimate of the model performance after it has been trained and validated. The remaining 20% of dataset is used for the test set.

For performing depression classification, a tweet is classified into two classes: “normal” or “depressed”. For this purpose, the BiLSTM model is experimented by tweaking the parameters as shown in Table **2.** The parameters whose values are tuned for the experimentation are; Vocab_size range: $[1000 - 4000]$, Batch size range: $[12 - 64]$, and The BiLSTM unit range is set as: $[20 - 170]$. The parameter whose values are kept constant are as follows; (i) embedding dimension, (ii) Epochs, (iii) activation function, and (iv) input size.

Table 2 Proposed BiLSTM Model with Parameter Setting

Model Parameter	**Value**	**Parameter**	**Value**
hidden layers	3	# of epochs	7
unit size (BILSTM)	20, 25,30,60,90,170	Batch size	12,16, 20, 32,35,64
Size of Vocabulary	1000,1500, 2000, 2500,3000,4000	Activation function	Sigmoid
Size of the input Vector	58	Embedding dimensionality	128

The used six variants of Bi-LSTM models with the selected parameters shown in Table 3. In these models, Bi-LSTM unit size is varied along with the filter size and number of filters size.

Table. 3 Bi-LSTM Models with Six Variants of Parameter Setting

Name of the model	**Filters no.**	**Unit size of BiLSTM**	**Size of filter**
BiLSTM (1)	32	50	3
BiLSTM(2)	32	120	2
BiLSTM(3)	12	127	3
BiLSTM(4)	12	170	2
BiLSTM(5)	10	200	3
BiLSTM(6)	10	215	2

4. Experimental Results and Discussion

This section presents an evaluation of results while conducting experiments to answer research questions and comparison of similar works with the proposed system.

4.1 Answer to RQ1

To find answer to RQ1: ***"How to classify user's tweet as "depressed" or "normal", by applying deep learning approach?",*** We conducted different experiments, detail is given as follows:

4.1.1 First Experiment

We performed experimentation with multiple variants of Bi-LSTM models based on different parameters setting, as shown in Table 4. It is obvious that a maximum accuracy of 93.5% is achieved by Bi-LSTM#6 model with the following hyperparameter tuning: Epochs=30, and batch size=300. The test accuracy, and training time are also reported for all Bi-LSTM models.

Table 4 Hyperparameter tuning for Bi-LSTM model

Models	**Epochs**	**Training Time(s)**	**Test Accuracy**	**Test Loss**
Bi-LSTM(1)	30	2s	82.32%	0.38
Bi-LSTM(2)	30	3s	81.41%	0.28
Bi-LSTM(3)	**30**	3s	**83.16%**	0.25
Bi-LSTM(4)	30	6s	83.91%	0.41
Bi-LSTM(5)	30	8s	84.5%	0.24
Bi-LSTM(6)	30	**18s**	**93.5%**	**0.28**
Bi-LSTM(7)	30	24s	90.4%	0.32

4.1.2 Second Experiment

Several deep learning models have been applied, namely RNN, CNN, LSTM, GRU and BiLSTM, to classify a tweet as "normal" or "depressed". The proposed Bi-LSTM model for depression classification outperformed the similar DL models (Table 5).

Table 5 Comparative results of proposed model with other Deep Learning Models.

Model	Accuracy	Precision	Recall	F1-Score
RNN	82.289%	0.83	0.82	0.83
CNN	0.87	0.84	0.86	0.82
LSTM	79.001%	0.81	0.77	0.79
GRU	0.84	0.81	0.84	0.83
Bi-LSTM	0.93	0.89	0.91	0.90

4.2 Answer to RQ3:

To find answer to RQ2: *"How can we evaluate performance of of the Proposed Bi-LSTM (Deep Learning model) in connection with Traditional Machine Learning Classifiers?*", to evaluate the efficacy of Bi-LSTM model with different supervised machine learning technique, we have conducted this experiment. The results are presented in Table 6, depicting the proposed model has outperformed all other comparing ML classifiers: RF, SVM, KNN, LR, and NB. While inspecting the performance of ML classifiers, SVM has performed better in respect to improved f-measure, recall, and precision.

Table 6 Comparison of proposed model with ML Classifiers

Technique	Classifier	Precision(%)	Recall(%)	F-score(%)	Accuracy(%)
Machine Learning (ML)	RF	0.82	0.80	0.81	0.80
	SVM	0.83	0.85	0.83	0.86
	KNN	0.76	0/72	0.70	0.72
	LR	0.82	0.80	0.81	0.81
	NB	0.83	0.70	0.76	0.84
Deep Learning	Bi-LSTM	**0.89**	**0.91**	**0.90**	**0.93**

4.3 Answer to RQ3

To find answer to RQ3, *"How can we evaluate performance of the Proposed Bi-LSTM (Deep Learning model) in connection with Baseline Studies?"***,** we conducted experiments to perform comparison of our approach with other similar works. Evaluation is performed on account of conducting experiments by inspecting values of different performance metrices like recall, accuracy, precision and f-measure. Table 7 presents the evaluation results of tour approach with other works, and it is evident that our system outperformed the comparing works with better results.

Table 7 Performance Evaluation with Baseline Studies

Kumar et al. 2019 (lexicon-based),				Orabi et al. 2018 **(supervised learning based))**				**Bi-LSTM(Proposed)**			
P(%)	R(%)	F(%)	A(%)	P(%)	R (%)	F(%)	A(%)s	P(%)	R(%)	F(%)	A(%)
0.69	0.74	0.71	0.81	0.71	0.82	0.82	0.81	0.89	0.91	0.90	0.93

Results obtained are promising, but still needs further improvement in accuracy and precision. However, it is an attempt to classify depression in Twitter posts and more enhancements can be made by launching merger of classifiers with deep reinforcement learning features.

4.3 User Interface of the proposed System

To predict, whether the given text reflects a person as normal or depressed, we developed a user-friendly web interface using Python-based Flask environment (Baker, A. 2018), and the trained deep learning model is deployed using Keras library. The front end (main page) of the web application is shown in Fig 2 An online user can enter the comment in the textbox (Fig. 3), then after pressing the predict button, an output is displayed as *"depressed"* or *"normal"* (Fig. 4).

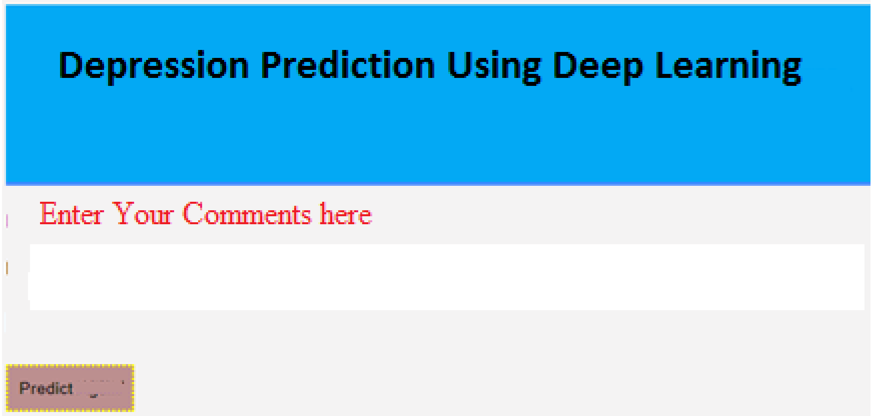

Fig.2 Front-end of the System

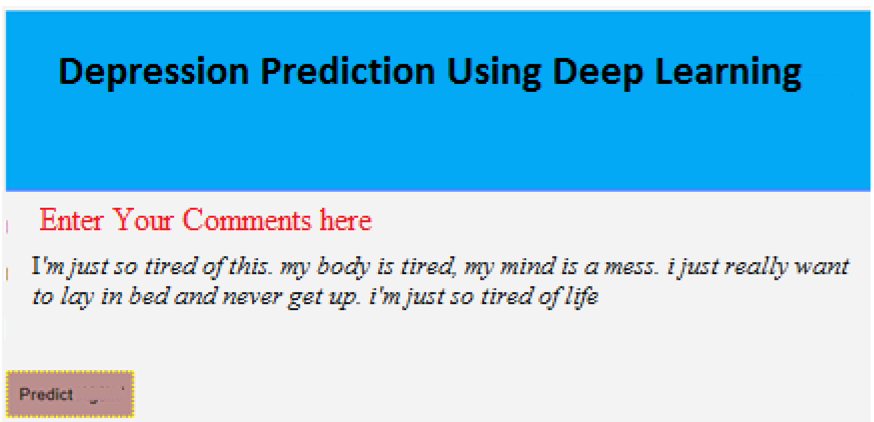

Fig. 3 User input to the system

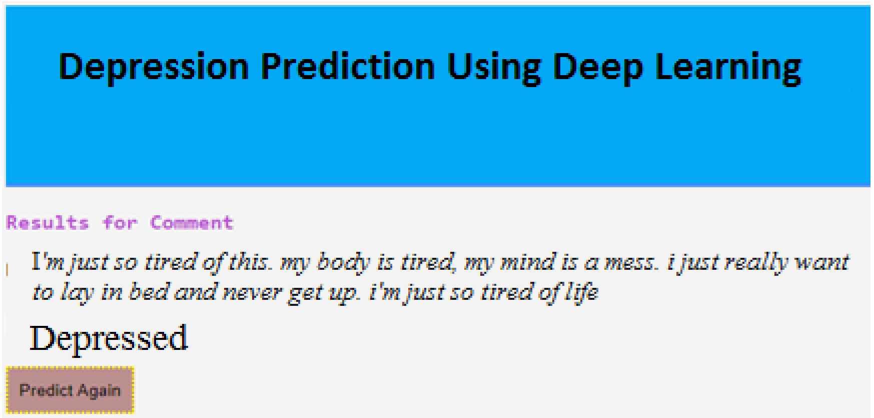

Fig. 4 System output

Conclusion and Future Work

The problem of depression classification from Twitter posts by exploiting a Deep neural network model, namely BiLSTM is presented. The proposed BiLSTM model maintain the sequential information in forward & backward directions, and finally, the tweet is classified as normal or depressed. We performed experiments with different machine, deep learning models and evaluated their performance on the public dataset. The results are promising, depicting that the BiLSTM yielded improved results with respect to other approached with improved f-measure (90%), recall (91%), accuracy (93%), precision (89%).

However, there are certain limitations like use of limited dataset and using only Twitter dataset for experimentation. Furthermore, we did not use pre-trained word embedding model and also the class labels in the dataset are imbalanced, that produce degraded performance of the system.
As future work, more datasets need to be evaluated by utilizing the deep learning based techniques to evaluate the effectiveness of the system. Using balanced classes in training data can upgrade the accuracy of the proposed appraoch, and finally, feature representation schemes such as Glove, Fasttext, and word2Vec can improve the system's performance.

References

Ahmad, S., Asghar, M. Z., Alotaibi, F. M., & Awan, I. (2019). Detection and classification of social media-based extremist affiliations using sentiment analysis techniques. *Human-centric Computing and Information Sciences*, *9*(1), 24.

American Psychiatric Association. 2013. Diagnostic and statistical manual of mental disorders (5th ed.), 5 edition. American Psychiatric Publishing, Washington.

Baker, A. (2018). Python and Flask Dev Environment Setup Guide. [online] Twilio.com. Available at: https://www.twilio.com/docs/usage/tutorials/how-to-set-up-your-python-and-flask-development-environment.

Cai, L., Zhou, S., Yan, X., & Yuan, R. (2019). A Stacked BiLSTM Neural Network Based on Coattention Mechanism for Question Answering. *Computational Intelligence and Neuroscience*, *2019*].

Daniel Preotiuc-Pietro, Johannes Eichstaedt, Gregory Park, Maarten Sap, Laura Smith, Victoria Tobolsky, H Andrew Schwartz, and Lyle Ungar. 2015a. The Role of Personality , Age and Gender in Tweeting about Mental Illnesses. In Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality, pages 21–30.

Daniel Preotiuc-Pietro, Maarten Sap, H. Andrew Schwartz, and Lyle Ungar. 2015b. Mental Illness Detection at the World Well-Being Project for the CLPsych 2015 Shared Task. In Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality, pages 40–45.

Gaspersz R, Nawijn L, Lamers F, Penninx BWJH. (2018). Patients with anxious depression: overview of prevalence, pathophysiology and impact on course and treatment outcome. *Curr Opin Psychiatry*. 2018;31(1):17-25.

Glen Coppersmith, Mark Dredze, and Craig Harman. 2014a. Measuring Post Traumatic Stress Disorder in Twitter. In In Proceedings of the 7th International AAAI Conference on Weblogs and Social Media (ICWSM)., volume 2, pages 23–45.

Glen Coppersmith, Mark Dredze, Craig Harman, and Kristy Hollingshead. 2015a. From ADHD to SAD: Analyzing the Language of Mental Health on Twitter through Self-Reported Diagnoses. In Computational Linguistics and Clinical Psychology, pages 1–10.

Glen Coppersmith, Mark Dredze, Craig Harman, Hollingshead Kristy, and Margaret Mitchell. 2015b. CLPsych 2015 Shared Task: Depression and PTSD on Twitter. In Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality, pages 31–39.

H Andrew Schwartz, Johannes Eichstaedt, Margaret L Kern, Gregory Park, Maarten Sap, David Stillwell, Michal Kosinski, and Lyle Ungar. 2014. Towards Assessing Changes in Degree of Depression through Facebook. In Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality, pages 118–125.

James W Pennebaker, Cindy K Chung, Molly Ireland, Amy Gonzales, and Roger J Booth. 2007. The Development and Psychometric Properties of LIWC2007 The University

Khan, A., Feng, J., Liu, S., & Asghar, M. Z. (2019). Optimal Skipping Rates: Training Agents with Fine-Grained Control Using Deep Reinforcement Learning. *Journal of Robotics*, *2019*.

Kumar, A., Sharma, A., & Arora, A. (2019). Anxious Depression Prediction in Real-time Social Data. *Available at SSRN 3383359*.

Michael Thaul Lehrman, Cecilia Ovesdotter Alm, and Rub´en A. Proa˜no. 2012. Detecting Distressed and Non-distressed Affect States in Short Forum Texts. In Second Workshop on Language in Social Media, Lsm, pages 9–18, Montreal.

Margaret Mitchell, Kristy Hollingshead, and Glen Coppersmith. 2015. Quantifying the Language of Schizophrenia in Social Media. In Computational Linguistics and Clinical Psychology, pages 11–20, Colorado. Association for Computational Linguistics.

Munmun De Choudhury. 2013. Role of Social Media in Tackling Challenges in Mental Health. In Proceedings of the 2nd International Workshop on Socially-Aware Multimedia (SAM'13), pages 49–52.

Orabi, A. H., Buddhitha, P., Orabi, M. H., & Inkpen, D. (2018, June). Deep learning for depression detection of twitter users. In *Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic* (pp. 88-97).

Philip Resnik, William Armstrong, Leonardo Claudino, and Thang Nguyen. 2015. The University of Maryland CLPsych 2015 Shared Task System. In CLPsych 2015 Shared Task System, c, pages 54–60.

Rohan Kshirsagar, Robert Morris, and Samuel Bowman. 2017. Detecting and Explaining Crisis. In Proceedings of the Fourth Workshop on Computational Linguistics and Clinical Psychology — From Linguistic Signal to Clinical Reality, pages 66–73, Vancouver. Association for Computational Linguistics.

Rhanoui, M., Mikram, M., Yousfi, S., & Barzali, S. (2019). A CNN-BiLSTM Model for Document-Level Sentiment Analysis. *Machine Learning and Knowledge Extraction*, *1*(3), 832-847

Sonawalla, S. B., & Fava, M. (2001). Severe Depression. CNS drugs, 15(10), 765-776.

Vinayakumar, R., Soman, K. P., & Poornachandran, P. (2017). Evaluation of recurrent neural network and its variants for intrusion detection system (IDS). *International Journal of Information System Modeling and Design (IJISMD)*, *8*(3), 43-63.].

World Health Organization. 2014. WHO — Mental health: a state of well-being.

Zhang, Y., Wang, J., & Zhang, X. (2018, June). YNU-HPCC at SemEval-2018 Task 1: BiLSTM with Attention based Sentiment Analysis for Affect in Tweets. In *Proceedings of The 12th International Workshop on Semantic Evaluation* (pp. 273-278).