IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 15 October 2024, accepted 2 November 2024, date of publication 12 November 2024, date of current version 22 November 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3496741

RESEARCH ARTICLE

A Hybrid Deep Learning Model for Predicting Depression Symptoms From Large-Scale Textual Dataset

SULAIMAN ALMUTAIRI[1], MOHAMMED ABOHASHRH[2], HASANAIN HAYDER RAZZAQ[3], MUHAMMAD ZULQARNAIN[4], ABDALLAH NAMOUN[5], (Member, IEEE), AND FAHEEM KHAN[6]

[1]Department of Health Informatics, College of Public Health and Health Informatics, Qassim University, Buraidah, Qassim 51452, Saudi Arabia
[2]Department of Basic Medical Sciences, College of Applied Medical Sciences, King Khalid University, Abha 62521, Saudi Arabia
[3]Jabir Ibn Hayyan University for Medical and Pharmaceutical Sciences, Najaf 54001, Iraq
[4]Department of Computer Science and IT, Cholistan University of Veterinary and Animal Sciences, Bahawalpur, Punjab 63100, Pakistan
[5]AI Center, Faculty of Computer and Information Systems, Islamic University of Madinah, Madinah 42351, Saudi Arabia
[6]Department of Computer Engineering, Gachon University, Seongnam-si 13120, South Korea

Corresponding authors: Sulaiman Almutairi (ssmtiery@qu.edu.sa) and Faheem Khan (faheem@gachon.ac.kr)

This work was supported in part by the King Salman Center for Disability Research through Research under Grant KSRG-2023-560, and in part by the Deanship of Research and Graduate Studies at King Khalid University through a Large Research Project under Grant RGP 2/566/44.

ABSTRACT A significant number of individuals are facing mental health issues due to a lack of timely treatment and support for detecting depression. This lack of early treatment is a primary factor contributing to conditions such as anxiety disorders, bipolar disorders, sleep disorders, depression, and, in severe cases, self-harm and suicide. Consequently, identifying individuals suffering from mental health disorders and offering prompt intervention is an extraordinarily challenging task. Therefore, this research introduced a novel hybrid deep-learning method for predicting depression at an early stage. In this study, we proposed a hybrid deep learning model for depression prediction, which mainly combines a Convolution Neural Network (CNN) and a Long Short-Term Memory (LSTM) model. An enhanced version of the LSTM approach, namely Two-State LSTM (TS-LSTM), is applied based on the feature attention mechanism. The proposed framework incorporates a feature attention mechanism into the TS-LSTM approach, which increases the ability to identify relationships and extract keywords for depression detection using the attention layer. This methodology is employed on a large dataset obtained from a publicly accessible online platform for young people. This dataset consists of text questions asked by young users on the platform. We extracted features through a one-hot encoding method from robust indicators of potential depression symptoms, which were predefined by medical and psychological experts. In comparative evaluations compared to conventional approaches, our system demonstrates superior performance. The experimental outcomes revealed that the proposed approach attained an accuracy of 97.23%, a precision of 98.57%, a recall of 97.13%, an F1-score of 97.84%, and a specificity of 97.93%, respectively. These results highlight the efficiency of the developed methodology that accurately predicts depression.

INDEX TERMS Deep learning, convolution neural network, long short-term memory, attention mechanism, depression detection.

I. INTRODUCTION

Social media platforms have emerged as a widespread method of communication, providing a platform where a considerable portion of the population conveys their feelings and shares their life understandings. These platforms not only contain general public information but also host a significant volume of content originating from individuals experiencing depression. Consequently, they can potentially produce sensitive social indications that hint at a person

The associate editor coordinating the review of this manuscript and approving it for publication was Rajeswari Sundararajan.

grappling with severe matters, such as self-harm, hopeless ideation, or an inclination towards illegitimate actions. Depression is a medically recognized disorder and stands as one of the most widespread mental health conditions impacting millions of people universally. This condition is considered hazardous as it impacts a person's mental well-being and physical repercussions. The harshness of depression is typically assessed based on the mental health status of the individual [1]. Among the most dominant mental health disorders are anxiety disorders, restiveness, sleep disturbances, eating disorders, addictive behaviors, depression, trauma-related conditions, and stress-related disorders [2]. Depression, in particular, is characterized by persistent feelings of hopelessness, demotivation, frequent mood swings, and a loss of interest in daily activities, encompassing physical, mental, and social dimensions, ultimately leading to emotional distress and physiological changes in the affected individual's body.

The World Health Organization (WHO) approximates that more than 295 million individuals globally wrestle with depression [3]. Depression can have a profound impact on individuals' well-being and their ability to function effectively in various aspects of life, including school/college activities and family life. Notably, adolescent depression has been linked to the development of attitude disorders and severe mental sickness in adulthood [4], [5]. According to WHO data, approximately 0.7 million people die by suicide annually, with suicide ranking as the fourth dominant cause of death among 16–20-year-olds [3]. Notably, mental illnesses, with depression at the forefront, constitute five of the major diseases responsible for disability or incapacity [6]. Therefore, the global burden of depression is extensive. The occurrence of depression in adult people spans almost 6% of different cultures, encompassing slighter forms of depression, including limited symptoms, insignificant depression, and probable depression, which impact about 21% of adults [7]. Individuals in middle age face the highest risk of developing depression. Moreover, there was a significant increase in the prevalence of depression worldwide, witnessing a surge of 19% between 2005 and 2015. It is crucial to emphasize that early proficient interference can meaningfully enhance mental symptoms by addressing issues like low self-confidence and rumination, as well as tackling somatic issues such as gastrointestinal concerns and sleep disorders in the majority of cases [8], [9]. Many existing prediction techniques for depression lack accuracy, primarily due to the absence of adequate diagnostic methods. However, social platform like Instagram, Snapchat, Facebook, and Twitter offers a promising avenue for predicting depression due to their vast user base and the extensive activities taking place on their platforms. Fig. 1 highlights that anxiety and depression rank among the most prevalent and significant issues for young people. The younger population frequently devotes a substantial amount of time to social media platforms, which allows scholars to gather and examine the content shared on these platforms [10], [11].

Detecting depressive symptoms at an initial phase, coupled with consequent valuation and intervention, can meaningfully enhance the forecasts of easing symptoms and track the root causes of the disorder. This technique can also help moderate the contrary significances on overall well-being, health, and several aspects of economic, personal, and social life [12], [13]. However, classifying depressive symptoms poses considerable issues in people's lives and needs significant resources. The primary approaches include scientific interviews and questionnaires accomplished by healthcare institutions or organizations [14]. These approaches often rely on the employment of psychological measurement tools to make predictions regarding the presence of a mental disorder, mainly depression. This traditional procedure primarily uses one-on-one questionnaires and offers a rough diagnosis of depressive psychological disorders. As a result, there is a need to develop an automated system that can categorize early signs of depression and tailor treatment strategies for those affected.

Furthermore, Structured Clinical Interviews have been accurately intended to predict the severity of symptoms and shared behavioral patterns shown by individuals grappling with depression [15]. Based upon prior inquiries in clinical psychology, it has been clarified that the connection between a language user (e.g., a speaker or writer) and their written or spoken expressions carries considerable meaning and holds promise for future applications [16]. Addressing mental health and its related challenges, which concern that spans all phases of life, from childhood through adolescence to adulthood. Depression often leads to a determined state of low mood, which can suggestively reduce creativity and interest in one's daily life [17]. Depression often establishes symptoms such as feelings of isolation, sleep disturbances, variations in appetite and sleep patterns, trouble concentrating on both personal matters and work in some cases, and a heightened risk of suicide [18], [19]. The recent study conducted by Havegerova´ et al., has shed light on the viability of text-based techniques for classifying individuals who may be at risk of depression. This research harnessed in-formal text samples associated with holiday experiences, signifying a promising avenue for detecting depression in individuals through analyzing their written content [20].

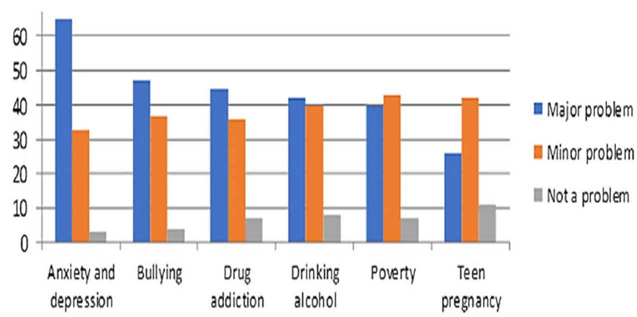


FIGURE 1. The prevalence of depression among young people in the united states in 2018 [25].

In previous studies, it has been established that the automatic analysis system of depressive symptoms within the text can have numerous practical applications, such as identifying sentiments in suicide notes and categorizing aggressive or miserable language in conversations or blog posts [21], [22], [23], [24]. However, identifying depression symptoms from patterns in textual content remains largely unexplored.

The key challenges comprise precisely capturing substantial indicators of depression within texts and identifying these symptoms in brief textual passages. To address these challenges, we aim to develop an automated hybrid deep-learning model to detect text depression symptoms. Our research is based on a dataset containing text-based interactions with young people seeking guidance concerning their self-perceived depressive symptoms. We believe that our research algorithms, which focus on natural language descriptions of user issues, can contribute to depression research. As a result, the present examination is dedicated to investigating how symptoms of depression are expressed through natural language text exploiting artificial intelligence technology.

II. RELATED WORK

Numerous studies in the area of mental health and spoken language were conducted, examining various theories that link depression to sociology and psychology. One cognitive theory, for instance [46], suggests that individuals experiencing depression employ pessimistic language when communicating and describing their surroundings. Their communication typically includes many negative words and a preponderance of first-person pronouns in both written and spoken discourse. Moreover, these individuals frequently withdraw from social interactions, leading to social isolation. Several researchers have investigated the relationship between mental disorders and language practice from different theoretical perspectives. For example, Adhikari and Tarusan [47] investigated linguistic patterns revealed by Alzheimer's disease patients in the context of the Nepalese language.

Similarly, Nick [48] explored the language used by college students who may be vulnerable to depression. Additionally, this research examined word usage in 300 poems written by eight poets who experienced suicidal tendencies and eight who did not, spanning dissimilar periods [49]. The results from these researches indicate that individuals anticipating suicide are inclined to utilize more self-oriented language, categorized by an abundance of self-referential words and personal pronouns while utilizing rarer collective or communal terms. These observations propose that such individuals may display social disinterest and an obsession with their thoughts and concerns.

Likewise, researchers have applied several deep learning algorithms to analyze brain data and accomplish high classification accuracy. Mahato and Paul reached a notable accuracy of 93.3% by employing features engineering techniques such as relative wavelet energy, wavelet entropy, band power, and alongside classifiers including multilayered perceptron neural network (MLPNN) and radial basis function network (RBFN) [50]. Ke et al. [50] referred to a CNN model, attaining an impressive accuracy of 98.81%. Kang et al. introduced independent component analysis (ICA) to eliminate artifacts and then utilized a CNN model, achieving an accuracy of 98.85% within the alpha band [51]. In separate research, Mahato and Paul [52] concentrated on band power and theta asymmetry features, accomplishing an accuracy of 88.3% using the SVM classifier. Mrazek et al. [53] reached an accuracy of 98.44% by capturing valuable features such as estimated entropy, sample entropy, wavelet packet decomposition (WPD), and band power using the Enhanced KNN method. Moreover, Dang et al. [54] illustrated innovation by merging the multivariate pseudo-Wigner distribution (MPWD) and frequency-dependent multilayer brain (FDMB) procedure with the CNN approach, achieving an accuracy of 97.27%. Movahed et al. [55] suggested extracting statistical and spectral wavelet functional connectivity, along with non-linear features, and utilized an SVM with an RBF kernel, realizing an impressive accuracy of 99%. Aydemir et al. [56] implemented a comprehensive technique for feature selection by utilizing neighbor-hood component analysis (NCA) and extracting 26 features from discrete wavelet transform coefficients. Their outcomes were outstanding, getting 99.10% accuracy with Weighted KNN and 99.04% accuracy with the Quadratic SVM classifier. Conversely, Loh et al. [57] referred to a CNN-based network, attaining a superior accuracy of 99.25%. Ibitoye et al. [58] explored two research utilizing supervised machine learning algorithms to predict emotional relations, using depression detection methods to recognize posts associated with depression on social media. Table 1 delivers a summary of research conducted by several scholars on algorithms for detecting depression from 2017 to 2021.

Trifan et al. [59] suggested a rule-based estimator employing the tf-idf weighting strategy to analyze bag-of-words features for categorizing depressive symptoms on the Reddit social media platform. Also, Kamil and Abbas [60], provided a solution for Twitter users for early depression detection by using the Bidirectional LSTM (BLSTM) CNN based on the XGBOOST machine learning technique for classified depression on the twitter dataset and achieved an accuracy of 94%. Thekkekara et al. [61] have connected deep learning algorithms such as CNN and BiLSTM to design a detective algorithm for depression by extracting data from social media and obtained an accuracy 96.71%. Furthermore, Lin et al., [89] developed a system known as SenseMood, which efficiently classifies and analyzes users with depression using the CNN-based architecture with BERT, and got 0.884 accuracy. Shen et al., [66] concentrated on the timely finding of depression through the analysis of social media by using labeled datasets on Twitter, distinguishing between

TABLE 1. Summary of relevant studies from 2017 to 2021. Abbreviations: NB—Naive Bayes, RF—Random Forest, SVM—Support Vector Machine, LR—Linear Regression, DT—Decision Tree, Decision Tree, GBDT—Gradient Boosted, CNN, LSTM, MLP—Multi-Layer Perceptron, WDL—Weighted Deep Learning, MSNL—Multiple Social Network Learning, GB—Gradient Boosting, SS3—A Text Categorization Approach.

Previous Studies	Approaches	Findings	Drawbacks
Farima et al. [64]	Multi-Layer Perceptron, SVM, Naïve Bayes	Superior performance by MLP	Limited dataset
Burdisso et al. [65]	SS3, Naïve Bayes, SVM, KNN	SS3 achieved higher accuracy	Time-consuming
Shen [66]	MDL, MSNL, NB, WBL	The best performance obtained by MDL	Concentrate on user confession
Chatterjee et al. [67]	GBDT, SVM, CNN, LSTM	CNN and LSTM achieved the best accuracy as compared to others	Work has a limited scope and more time consumed on training
De Souza et al. [68]	DT, SVM, RF, LR, GB	RF reached high accuracy	Concentrate on user comments
Nova [69]	LightGBM, Multinomial Naive Bayes, Multi-layer Perceptron	These approaches achieved satisfactory accuracy	Testing data is significantly less
Alsagri and Ykhlef, [70]	NB, SVM, DT	SVM provides superior accuracy	Cannot avoid over-fit data
Islam et al. [71]	KNN, DT, SVM	DT offers the best result	Data focus on user comments
Kim J. [72]	XGBoost, CNN	CNN achieved higher accuracy	Limited scope
Aldarwish and Ahmad, [73]	BN, SVM	NB reached the best accuracy	Dataset in in Arabic
Gaikar et al. [74]	Hybrid model, NB, SVM	The hybrid model obtained excellent accuracy	Time-Consuming
Nguyen T.L. [75]	SVM, Ensemble model, DT	DT provides superior accuracy	Data focus on user comments
Zia Uddin et al. [76]	RNNs and LIME with LSTM	Superior accuracy	This work is limited to the data of young people
Teck K. [77]	BERT	Performs depression detection for the Arabic language and provides better results	Due to limitations in the available data, it is advisable to further explore behaviors related to depression as manifested in social media.

depression and non-depression cases, and extracting six feature groups associated with depression, and attained 0.850 accuracy.

On the other hand, Tong et al., [91] introduced a new methodology known as Cost-sensitive Boosting Pruning Trees (CBPT), which showed robust classification performance on two publicly available Twitter datasets for depression detection and obtained 0.890 accuracy. Nadeem, [95] utilized a crowdsourcing approach to gather a list of Twitter users who claim to have been identified with depression based on a deep learning method with an integrated Bad of Words technique to quantify the content of each tweet and achieved 0.81% accuracy. While Samuel et al., [97] examined public sentiment associated with the pandemic by investigating coronavirus-specific Tweets employing R statistical software and its sentiment analysis tools. Furthermore, they provided a systemic review of two main machine learning classifiers and demonstrated an accuracy of 91%. Zogan et al., [1] have introduced a reasonable Multi-Aspect Depression Detection approach using a Hierarchical Attention Network (MDHAN) for the automatic identification of depressed users on social media and provided insight into the model's predictions, and got 85.20%, 86.90% accuracy respectively. Like our study, Yang [78] proposed a hybrid methodology to combine LSTM with CNN utilizing Keras to identify whether users on social platforms are depressive based on their Twitter posts. Moreover, for pre-trained vectors, they used the Word2Vec method using the twitter_sentiment dataset found from the Kaggle. The experimental results illustrated that the proposed hybrid model achieved an excellent accuracy of 98.90%, significantly outperforming to our study, which has an accuracy of 97.23%. In a recent study, Nareshkumar and Nimala [100] proposed a novel approach that efficiently and precisely detects posts associated with anxiety and depression. Furthermore, they used a knowledge distillation method to transfer insights from the large pretrained BERT model to enhance the model's performance and accuracy. Authors employed an enhanced BERT method with word2vec to efficiently examine and classify symptoms of depression and anxiety and attained an accuracy of 92%. Moreover, Orabi et al. [62] have established a CNN approach with notable performance in image processing and have extended its significance to text processing. Furthermore, Genest et al. [63] explored the usage of deep learning models to analyze TV show scripts and recognize the underlying emotions conveyed within the script.

III. PRELIMINARIES

The study of deep learning approaches for depressed symptoms classification includes exploring initial methods and techniques to evaluate their effectiveness in identifying depression-related patterns from data. The enhancement in classifying textual data has been mainly boosted by using progressive deep learning algorithms, such as RNNs and CNNs, coupled with incorporating nervousness's word embeddings [26], [27]. Table 2 illustrates an overview of some frequently observed symptoms. In contrast, RNNs are crafted explicitly for handling sequential data, making them adept at tasks like part-of-speech tagging and language translation [28]. CNNs are distinguished by their capacity to preserve translational invariance, allowing them to categorize exclusive features from images irrespective of their direction or strength in visual tasks. The integration of pooling methods has also made CNNs effective in processing textual data. By combining CNNs with word embeddings, they can capture both syntactic and semantic information from text.

TABLE 2. Examples of depression symptoms.

Symptom Type	Examples
Mood and emotional symptoms	Emotional distress, powerlessness, despair, discomfort, diminished self-worth, Guilt, Sadness, Anxiety, fear of insufficiency, unease, agitation, annoyance, frustration, handling one's emotions, humor, the inability to experience pleasure, a sense of emptiness, and a lack of emotional responsiveness, among other things.
Physical symptoms	Discomfort, Reduced appetite leading to weight loss, Sexual difficulties, Excessive daytime sleepiness, disrupted sleep patterns, Inability to sleep, Tearfulness, Inability to speak, slowed physical and mental activity, Lack of strength, Decreased vitality, Persistent tiredness, and more.
Cognitive symptoms	Lack of motivation, Absence of enthusiasm, perceived as incurable, negative outlook, Challenges in decision-making and finding solutions, memory loss, decreased ability to focus, Mental fog, Diminished social engagement, and so on.
Auto aggression symptoms	Wish to live, Suicidal thoughts, Self-harm, Suicidal attempt etc.
Functioning symptoms	Financial concerns, Capability to survive with life, Individual problems etc.
Complexity	Loss of job, Professional responsibilities, Stopping studies, etc.
Professional	Self-exclusion, Domestic life,
Social	Social connections, social seclusion, Expressing emotions, etc.

The LSTM represents a contemporary evolution of the conventional RNN, designed explicitly for sequential modeling. Utilizing a gating architecture, such as the long short-term memory, effectively addresses these challenges by enabling the preservation of memory over longer sequences compared to traditional RNNs [85]. Moreover, incorporating LSTM into this hybrid model improves its effectiveness. LSTM's capability to accomplish sequential inputs of varying lengths makes it particularly suitable for tasks like sequence labeling [29]. In recent years, deep learning algorithms have played a vital role in pattern recognition and artificial intelligence research [30], [31], [32], [33], [34]. However, these algorithms demonstrated two significant limitations. The first limitation is their computational intensity, which demands a lot of time for data modeling. In the early stages of deep learning, Restricted Boltzmann machines (RBMs) were utilized to accelerate training, showcasing their remarkable discriminatory capabilities [35]. At the start of deep learning, CNNs are traditionally focused on image and video analysis instead of time series.

The Word2Vec method employs neural networks to extract the associations between words in text corpora. This method comprises a single input layer, one hidden layer, and a final output layer. It takes a text corpus as input and produces an output vector. The resulting vector is structured so words with similar meanings are positioned closely to one alternative in the vector space. Within the Word2Vec framework, the neural network conducts mathematical operations on word tokens to identify similarities between words. There are two primary training procedures for Word2Vec: Continuous Bag of Words (CBoW), which uses context to predict a target word (w), and skip-gram, which employs a word (w) to predict the surrounding context or neighboring words. Word2vec generates numerical vectors representing word features, capturing data or information about individual words. It autonomously learns these features without human input and can make precise predictions near word meanings based on prior experiences [33].

In contrast, Recurrent Neural Networks (RNNs) excel in the inspection of consecutive text and patterns, outperforming CNNs in this domain [36]. Consequently, RNNs have not been widely implemented for time-sequential data analysis due to their superior discriminative power. However, typical RNNs often suffer from the vanishing gradient problem when handling high-dimensional and time-sequential data. To handle this issue, LSTM was presented as an enhancement to RNNs. As a result, this research leverages the advantages of LSTM-based RNNs to approach various emotional states within textual data.

A. DISABILITIES AND REHABILITATION

Recent statistics show a steady rate in the number of people with disabilities in America, reaching 13.5% of the total population in 2021 [37]. Globally, it is estimated that 1.3 billion persons (equating to a staggering 16%) suffer from a significant form of disability [38]. The diverse types of disabilities, including motor and cognitive impairments, often hinder the quality of citizens' lives and their ability to lead independent lives. Several modern technologies, such as ambient assisted living [39] and machine learning [40], are leveraged to overcome the challenges of human impairments. For instance, Scrutinio et al. [41] explored the effectiveness of tree-based machine learning models in predicting rehabilitation outcomes in people who experienced a stroke, which is the leading cause of disabilities. In post-stroke rehabilitation, wearable sensors and machine learning methods were utilized to assess rehabilitation movements and classify depressed patients' prescribed exercises [42]. In another study [43], authors apply unsupervised feature learning to evaluate the effectiveness of motor recovery treatments. A novel feature selection technique evaluating 17 kinematic features proved quite effective for upper limb rehabilitation of 41 stroke patients.

Furthermore, Tschuggnall et al. [44], the medical rehabilitation success of health professionals is predicted using a mix of regression and classification algorithms on a large-scale textual dataset of 1047 patients. One interesting research study investigated the effect of depression disorder on physical rehabilitation results [45]. XGBoost and random forest techniques were applied to 10 available features to assess depression in patients undergoing rehab sessions, showing a 92.6% classification accuracy using only five features. This study concentrates on text data processing, feature extraction, and recognition of depression symptoms in text, with the ultimate goal of implementing a chatbot as an intelligent application. Fig. 2 illustrates the layout of a text-based system for detecting depression symptoms within an intelligent application. In this system, a user inputs a query in text form, which is then processed by a server that performs feature extraction and utilizes deep learning algorithms. In this research, we have combined these deep learning algorithms to present our hybrid model.

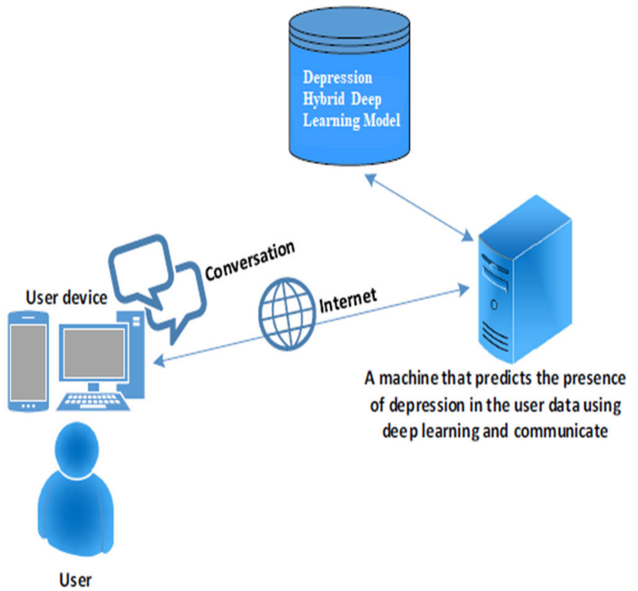


FIGURE 2. **A structured arrangement for categorizing textual content that includes signs of depression.**

B. GAPS OF EXISTING STUDIES

Online social media platforms, which are continually expanding, have established a means of facilitating communication in our everyday lives. Distinguishing between individuals who are depressed and those who are not based on linguistic differences presents a formidable challenge. The literature review has extensively investigated various methodologies for detecting and preventing depression using data mining techniques. Previous research has predominantly employed machine learning and deep learning algorithms [103], [104], [105], [106] using text data [107], medical data of healthcare [108], [109] for using different behavior [110]. Notably, most datasets created for these studies have included participants from a wide range of age groups, but none specifically targeted young adults between 18 and 25. Despite the challenges in extracting depression symptoms from textual data, text-based approaches still hold significant promise in this field.

However, based on recent studies and practical observations, we found one major limitation of typical LSTM it primarily focuses on the forward linguistic context when analyzing a word. This makes it essential for the LSTM model to learn from the backward context to ensure a full consideration of features. We noticed that a word's meaning is influenced by both forward and backward contexts in any model. Moreover, another limitation of LSTM is its inability to effectively identify rare and informative features during the feature selection process. As a result, the method often fails to capture subtle details and valuable information, eventually producing suboptimal solutions for classifying depressed symptoms from large-scale textual datasets.

Furthermore, In the diagnosis of depression, most scholars have primarily relied on a single machine-learning method, and their evaluation has often prioritized accuracy as the key metric for algorithm selection. Additionally, a uniform dataset size has been commonly applied across all algorithms. To address these gaps, this paper raises several essential research questions through Systematic Literature Review [111], [112] that warrant further investigation.

- RQ1: What challenges are associated with the implementation of text-based depression detection?
- RQ2: What constitutes the most efficient text-based method for detecting early-stage depression?
- RQ3: How can we investigate the most influential factors in identifying depression?

Nevertheless, the drawbacks highlighted in each article serve as both the challenge and opportunity to enhance the model's effectiveness. Therefore, to address the drawbacks of previous research, we integrated data pre-processing and hybrid deep learning approaches into our methodology to enable the early detection of depression from large-scale textual datasets.

IV. KEY CONTRIBUTION OF OUR STUDY

Deep learning models such as CNN and LSTM excel when abundant data is available, and with the abundance of open datasets, they yield superior results in categorizing social media contents as either depressed or non-depressed. In particular, RNNs are favored for their ability to incorporate external pre-trained embeddings. On the other hand, CNNs are well-suited for rapid text processing and excel in extracting local features from the text, making them especially valuable in Natural Language Processing (NLP) tasks. By combining CNNs and RNNs, we can take the strengths of both architectures. Our research focuses on assessing the effectiveness of the existing CNN and LSTM for predicting depressed symptoms using large-scale datasets. In this study, we introduce a hybrid architecture that leverages the strengths of both CNNs and advanced variants of traditional LSTMs. The key contributions of this research can be summarized as follows:

- A framework for detecting depression is introduced, leveraging deep learning methodologies to analyze textual content from social media.
- We proposed a hybrid deep learning methodology for depression detection, which mainly combines a comprehensive CNN-based network and an LSTM model is used. An enhanced version of the LSTM approach,

namely Two-State LSTM (TS-LSTM), is applied based on the feature attention mechanism. The proposed framework incorporates a feature attention framework into the TS-LSTM model, enhancing its ability to identify the relationships and extract keywords for depression detection by utilizing an attention layer known as CNN+TS-LSTM. To the best of our knowledge, this study utilizes semantic features, statistical analysis, and deep learning algorithms for the first time, and they are integrated with word embeddings for depression detection.

- Our research findings have demonstrated the superiority of our proposed hybrid approach over traditional studies conducted on a benchmark large-scale textual dataset accessed by social media.

V. PROPOSED FRAMEWORK

In this study, we briefly describe the proposed framework, which incorporates CNN and two-state LSTM based on a feature attention strategy. The developed architecture utilizes word embeddings as input, facilitating the extraction of high-level contextual word features across time steps.

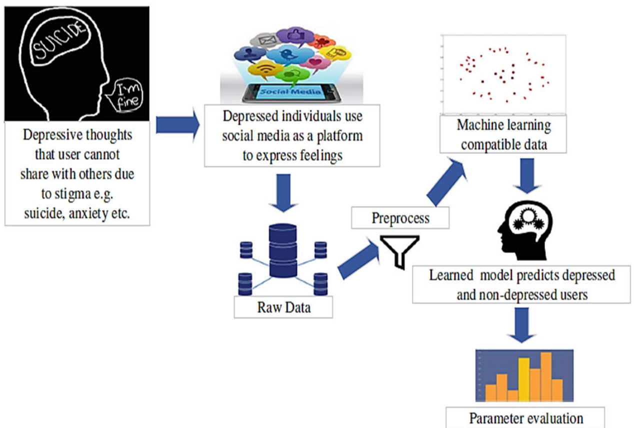


FIGURE 3. The proposed scheme framework.

Essentially, word embeddings serve as vectors or visual representations of words. Word2vec, "word to vector," is the predominant method for generating word embeddings. An embedding layer subsequently predicts these features before being fed into the CNN model.

The developed framework comprises six modules, illustrated in Fig. 3: (1) Data extraction from user-generated social media content. (2) Raw data analysis. (3) Cleaning and pre-processing of the raw data. (4) Extraction of features to create data suitable for machine processing. (5) Classification of depressive contents versus non-depressive contents. (6) Assessing and comparing the performance of the developed hybrid CNN+TS-LSTM with traditional depression detection approaches using various evaluation parameters. As shown in Fig. 4 a Rectified linear activation unit (Relu) is employed for the final classification.

The primary contributions of the developed mechanism lie in its ability to extract crucial features during two key phases:

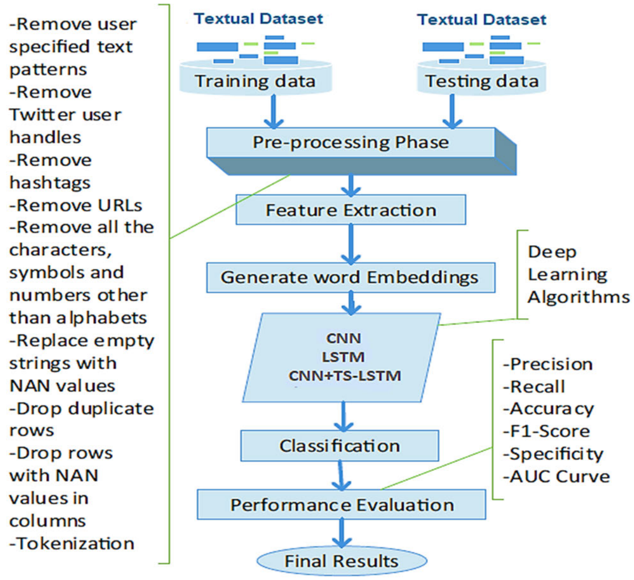


FIGURE 4. The proposed flowchart.

The Pre-Feature Attention TS-LSTM and the Post-Feature Attention TS-LSTM. The following steps visually represent the proposed approach, with its associated flowchart.

VI. DATASET DESCRIPTION AND PRE-PROCESSING

A. DATASET DESCRIPTION

Facebook and Twitter stand out as extensively used online social media platforms that offer accessible and user-friendly access to data. Data is the main requirement and key component for conducting these experiments. The superiority of the data serves as the essential factor in achieving precise results. One of the critical factors of this study lies in the Data Collection (DC) process, which has resulted in creating a pioneering dataset explicitly tailored to the problem under discussion. This dataset has been meticulously assembled from various sources and is publicly accessible, such as standard repositories and online social media platforms like Twitter and Facebook. As elaborated below, each source required distinct strategies to gather sufficient samples suitable for our experiments. Creating and verifying terminology utilized as a user-friendly data navigation vocabulary for individuals with mental health conditions is a time-consuming process of substantial importance. The text in the database was analyzed to summarize a collection of sentences and words that could indicate depression in individuals. A medical practitioner subsequently validated these sentences and words. Table 3 presents key features representing sentences and/or words that may appear in queries made by young individuals with depression. Table 4 illustrates the Nine phrases in Arabic (and its English translation) indicating depression [Aljazeera.net].

B. DC FROM EXISTING SOURCES

Due to the sensitive nature of content related to depression disorders, it is not readily available in sufficient quantity.

TABLE 3. Some key sentences and words are frequently employed in depressed and non-depressed content.

Depressed	Non-Depressed
Suicidal thoughts	Having fun
Feeling sad	Joyful
End my life	I am happy
Crying	Enjoying life
Grateful	Anguish
Nothing interests me	Feeling motivated
Always tired	Enjoying songs and movies
Negative thoughts	Positive thoughts
Loneliness	Connected
Unmotivated	Inspired
Excited	Miserable

TABLE 4. Nine phrases in arabic (and its english translation) indicating depression [aljazeera.net].

This always happens	يحدث هذا دائما
I will never be able to do that	لن أتمكن مطلقا من فعل ذلك
I can't	أنا لا أستطيع
I can not ..	لا يمكنني..
This is all my fault	كل هذا خطئي
I am stressed and feel tired	أنا مجهد وأشعر بالتعب
I know I'm alone	أريد ان أكون وحدي
No one cares	لا أحد يهتم
What's the point?!	ما الجدوى؟!

We encountered only a few well-suited datasets for our specific problem domain. These include a small and noisy depression dataset [78] found on a GitHub repository, which is referred to as 'dataset E1,' and a corpus of depressive tweet contents [79], which is referred to as 'dataset E2.' Dataset E1 consists of approximately 4,000 tweets expressing depressive moods by Twitter users, while dataset E2 contains 7,000 instances of users that are publicly written on social media, with identification of depressive and non-depressive disorders confirmed by medical experts. In dataset E1, the relevant columns of interest include 'id,' 'date,' 'content' (the topic of the text), 'text' (the user's text), 'user_name,' 'time_zone,' and 'label' (indicating depressive or non-depressive content). The second dataset, E2, comprises columns including user_id, u_name, label, and message, with our focus on the user_id, message, and label columns. It is important to note that both corpora are utilized to detect depression in online users, and therefore, they contain instances related to depression as well as those from normal individuals.

We observed that the two datasets contained a higher number of samples related to depression compared to non-depression cases. Consequently, this study addressed this imbalance by acquiring and integrating additional standard instances from a third dataset, the "SMILE Twitter Emotion Dataset" [80], denoted as E3.

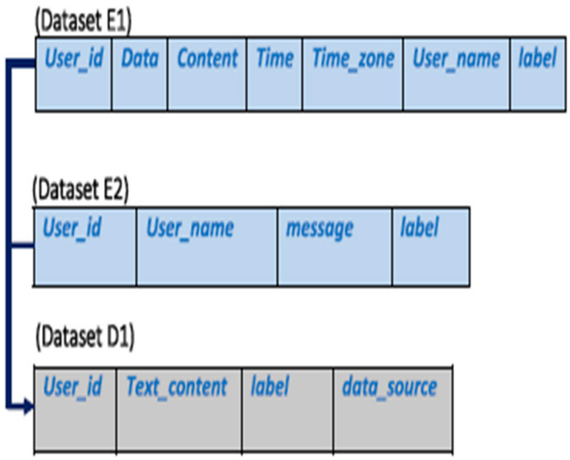


FIGURE 5. Converting the current datasets E1 and E2 into a normalized format as D1.

This dataset includes columns such as surprise, angry, happy, sad, crying, grateful, and irrelevant, apart from user_id. Since we focused on obtaining normal (non-depression related) samples, we filtered out instances with labels 'angry,' 'sad,' and 'not_relevant.' Subsequently, we incorporated the remaining 'happy' and 'surprise ' -correlated instances into our dataset D1, and we introduced an extra 'source' column to record the source information for each instance, as illustrated in Fig. 5.

C. DC FROM ONLINE SOCIAL MEDIA PLATFORM

To introduce diversity into the dataset, we gathered instances associated with depression symptoms from prominent social networking platforms. These platforms are rich sources of up-to-date user information. Online social networks (OSNs) have been identified as significant contributors to daily social interactions. Moreover, it has been noted that numerous individuals experiencing depression tend to articulate their emotions and sentiments on online platforms. As a result, these social networking platforms harbor significant social cues that can be employed to identify user moods and emotions. Our data collection efforts encompassed prevalent social media platforms, including Twitter and Facebook. We employed data scraping methods for Facebook and Twitter utilizing the Netvizz app [81] and the Tweepy and TWINT packages [82]. These methods were based on the most prevalent keywords found in typically depression-related datasets. Figures. 6, 7, and 8 exhibit word clouds that display the vocabulary used in the datasets from Facebook and Twitter, respectively. The size of each word in these visualizations corresponds to its frequency within the respective dataset.

After conducting separate exploratory data analyses (EDA) on both datasets, we discovered numerous vocabulary terms and expressions that serve as social signals for identifying depression symptoms in users. In addition to various distinctive dictionary terms, words like "Alone," "Loss," "Sad," "Hopeless," "Depress," "Stress," "Failure," "end_of_life,"

FIGURE 6. Word cloud displaying the most commonly utilized depression-related terms on social media, reflecting symptoms of depression(https://www.shutterstock.com/image-vector/vector-concept-conceptual-depression-mental-emotional -468973511).

etc., emerged as the most frequently utilized terms by individuals in both datasets. Table 2 shows examples from both datasets.

It was observed that individuals tend to exhibit greater openness and expressiveness on pure social networking platforms such as Facebook compared to those using Twitter. People on Facebook discuss various aspects of their lives, including family matters, association challenges, career-related issues, and health concerns.

Conversely, on blogging platforms such as Twitter, individuals tend to be less expressive and focus on a narrower range of topics, such as career issues, job-related matters, and unemployment problems. Analyzing the most frequently utilized words in the Facebook corpus (FC) reveals terms like 'jobless,' 'hopeless' 'cheating,' 'depressed,' 'God,' 'quit,' 'scary,' 'fake people,' 'empty,' 'loss,' 'crying,' 'directionless,' 'crisis,' 'painful,' 'low energy,' etc.

FIGURE 7. Word cloud on Facebook displaying the most commonly utilized vocabulary related to depression within the Facebook social platform (https://www.shutterstock.com/image-vector/vector-concept-conceptual-depression-mental- emotional-468973511).

In contrast, words such as 'aimless,' 'failure', 'jobless,' 'unemployment,' 'uselessness,' 'stressed,' 'sadness,' 'unmotivated,' and 'anxiety,' as shown in Table 5, demonstrate how we have categorized vocabulary terms into several categories, including Mood Vocabulary, Body Vocabulary, Relationship Vocabulary, Offensive Vocabulary, and Personal Pronoun Vocabulary (PP Vocabulary), based on their semantic information. Furthermore, additional information about the descriptions of the collected datasets utilized in this study can be found in Table 6.

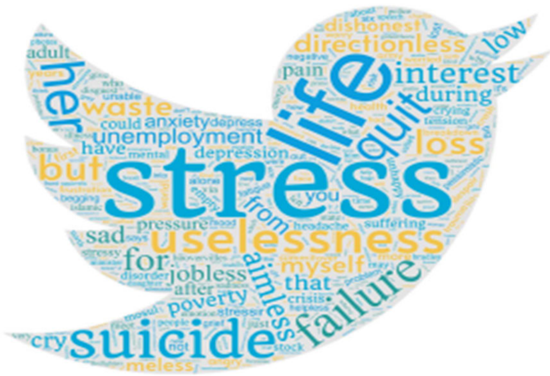


FIGURE 8. Twitter-generated word cloud showcasing the most commonly employed vocabulary associated with depression on the Twitter platform [101].

D. DATA PREPROCESSING

It is noticed that most conventional machine and deep learning approaches exhibit suboptimal presentation when dealing with raw data due to their lack of machine understanding. Therefore, it is essential to preprocess such data before inputting it into a learning algorithm. This research thoroughly preprocesses raw data, gathered from various channels is initially stored in separate databases (.csv sheets) and then undergoes overall data cleaning, normalization, and preprocessing filters. These filters include removing noisy data, handling missing data, and addressing inconsistencies. To eliminate unwanted elements like HTML tags, extra white spaces, special symbols, and digits, each instance undergoes filtering out these tokens. Handling the missing data involves overseeing empty cells, NaN values, and other associated issues. In the second phase, depicted in Fig. 10, the refined data undergoes another iteration through the natural language toolkit (NLTK) libraries for further analysis. The fundamental filters in the NLTK toolkit contain normalization, tokenization, lemmatization, and stop word removal. Normalization guarantees consistent case usage (lowercase in our investigation), while tokenization dissects the rare text into words and sentences, known as tokens.

These tokens are vital in grasping context or building networks for natural language processing (NLP). NLTK's tokenize module includes models for word and sentence tokenization. Generally, stop words in English such as "the," "is," and "and," are eliminated from the text as they contribute little meaningful information for classification tasks due to their frequent occurrence. Lastly, stemming simplifies words to their root forms, grouping different words under the same stem, even if the stem itself is not valid in the language. Fig. 9 illustrates these NLTK filters with an example sentence for clarity.

TABLE 5. Examples of sentences utilizing social cues to detect depressive symptoms on twitter and facebook.

Corpus	Vocabulary Terms	Example Sentence (From dataset)	Vocabulary Category
Facebook Corpus (FC)	Frustrated, Helpless, Disgust, Sad, Hopeless, Angry, etc.	I feel mentally exhausted and struggle to get through everyday life. The worker is no longer enjoyable. I cry every day. I am hopeless now. Looking for some sun to heal me up. What shall I do?	Mood_vocab
Facebook Corpus (FC)	Tongue, Eyes, Nose, Throat, Head, Chest,	"My head will explode with depression" "with all the pain in my chest"	Body_vocab
Facebook Corpus (FC)	Parents, Partner, Family, Mother, Girlfriend, Boyfriend,	"Life is beautiful with an honest partner, … without it is hell like mine"	Relationship_vocab
Twitter Corpus (TC)	Fatigue Trauma, Curse, F***, Hell, bloody, Fail/Failure	"Hell with this life…" "…F***ed up my entire career…" "I'm cursed with my success" "I have proved to be a failure in life"	Curse_vocab
Twitter Corpus (TC)	Vulgar, harm, Toxicity kill, obscene, illegal, end-of-life, etc.	"Naked like evil, life is ending me up…" "All alone in peace, checking on my blood color" "All issues vanish with you as you leave the world" "Death the end-of-life seems to be the only solution…"	Offensive_vocab
Twitter Corpus (TC)	I, me, myself, our, us, etc.	I feel so lonely because my friends left me. I kind of have no feelings. I am not happy. Am I the only one with such an unbearable pain in my chest. It is me the reason for all the bad things happening in my family. What's happening to me?	PP_vocab (Personal Pronoun)

TABLE 6. Datasets description after balancing and newly generated datasets.

Dataset/Corpus	Depressive instances	Non-Depressive Instances	Total #Instances
Facebook Corpus (FC)	3100	3400	6500
Twitter Corpus (TC)	6130	5660	11790
Dataset (E1) [31]	4112	4074	8186
Dataset (E2) [33]	5610	4600	10210
Smile Annotation Dataset (E3) [34]	N/A	3565	3565
Total	**18952**	**21299**	**40251**

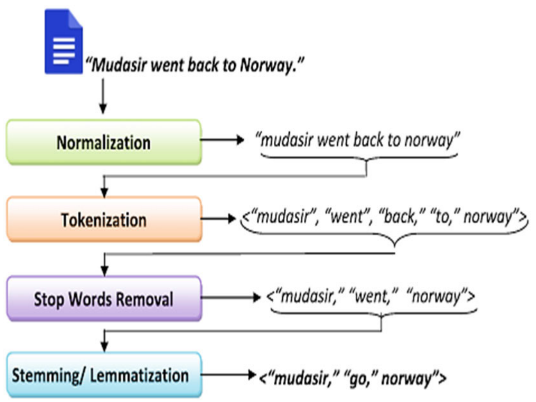


FIGURE 9. Data preprocessing using NLTK libraries.

VII. METHODOLOGY

A. WORD2VEC-BASED FEATURE REPRESENTATION

Word2Vec is a well-known technique in deep learning for generating word embeddings, dense vector representations of words that extract their semantic relationships. This technique creates a collection of word vectors, where words with parallel relationships are closely composed in the vector space, while words with distinct meanings are situated beyond. To illustrate, words like "crying" and "pain" exhibit proximity in the vector space, whereas words like "Mangoes" and "Pakistan" are noticeably distant from each other. The CBoW technique has been employed to identify the vocabulary terms employed within depressive tweets and regular non-depressive text. The collected depression-related content from platforms like Facebook and Twitter usually utilized terms including "depression" (4331 occurrences), "headache" (1209 occurrences), "disgust" (1199 occurrences), "quit" (3262 occurrences), "pain" (1790 occurrences), "negative" (1190 occurrences), and "helpless" (1341 occurrences) were identified.

Moreover, variations in vocabulary usage were noted across three different platforms. For instance, on the Facebook social network, the most frequently used dictionary terms include words like "fake" (7432 occurrences), "suicide" (6549 occurrences), "die" (7010 occurrences), "sadness" (4469 occurrences), "quit" (2438 occurrences), "pain" (3511 occurrences), "dishonest" (4018 occurrences), "pressure" (790 occurrences), "fatigue" (1000 occurrences), and more. In contrast, terms such as "stress" (4016), "life" (4092), "failure" (3900), "uselessness" (2917), "unemployment" (1698), "aimless" (992), "loss" (810), "waste" (676), and others have been observed as the most frequently used terms on the Twitter platform. Similarly, a distinct set of vocabulary terms emerged from YouTube comments, including words like "hell" (5019 occurrences), "trauma" (4967 occurrences), "curse" (4822 occurrences), "worse" (4020 occurrences), "death" (3978 occurrences), "naked" (3912 occurrences), "broken" (3814 occurrences), "her" (3009 occurrences), "mood" (2968 occurrences), "myself" (2771 occurrences), and more. Figures 6, 7, and 8 depict word clouds illustrating the most prevalent vocabulary terms for Facebook and Twitter corpus, respectively.

B. CNN MECHANISM

The initial use of CNN was focused on addressing computer vision tasks such as digital image processing and video analysis, but it was subsequently adapted for NLP tasks such as spam filtering, text categorization, question answering, and sentiment analysis. Over time, it has gained a significant reputation and is now one of the most extensively employed deep learning algorithms. CNN comprises two main components: an automatic feature-capturing segment and a categorizing segment. A conventional CNN architecture predominantly incorporates three distinct types of layers [84]:

- Convolution layers: In this layer, the input data undergoes a convolution operation using convolution kernels (filters), creating a feature chart.
- Pooling layers: The primary objective at this stage is to preserve as many pertinent features as possible while concurrently diminishing the spatial proportions of the input to prepare it for subsequent convolutional layers. Max and Average (max/avg pooling) are the most frequently employed pooling functions.
- Fully connected layers: After numerous rounds of convolution and pooling, these layers convert the 2D feature maps into a 1D feature vector. This vector is subsequently employed to classify the extracted features.

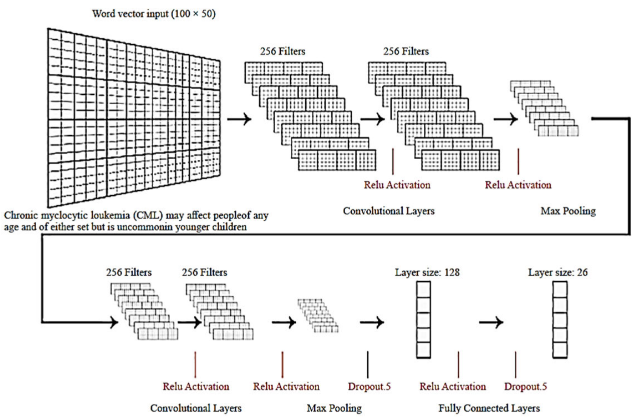


FIGURE 10. Traditional architecture of CNN model.

Fig. 10 refers to a visual representation of the standard CNN network architecture for depression detection. In this study, the computational operation of the Max pooling layer was utilized. The Relu activation function transforms the pooling layer vector, capturing its essential characteristics. This transformation helps abstract its vital properties, effectively mapping the text from the feature to the marker space. This research utilizes the CNN model to extract informative semantic features, and after extracting contextual semantic information from the CNN model, it is subsequently fed into the two-state LSTM model to identify essential features.

C. TWO-STATE LSTM MECHANISM

Based on the literature review, we found that recent study [102] has demonstrated a drawback of LSTM, as it mainly extracts only the forward linguistic context when analyzing words, making it unable to learn from backward contexts. Accordingly, our findings designate that the processing of a sentence is influenced not only by the forward context but also by the backward context in several language applications.

Consequently, we presented the Two-State LSTM (TS-LSTM) model to address the drawbacks. TS-LSTM is a modified structure that effectively extracts contextual information from both directions through two distinct processes: one is designated as the "forward pass" responsible for positive information flow, and the other is referred to as the "backward pass" accountable for handling negative aspects, as demonstrated in Fig. 11, This approach inspired by the bidirectional recurrent neural networks (BRNNs) [86].

TS-LSTM includes two separate recurrent networks, one for forward pass (from left to right) and the other for backward pass (from right to left) during the training phase. Eventually, these two passes are then combined into a single form of output layer. Equations (1) to (6) detail the computations for the forward pass, while equations (7) to (12) designate the calculations for the backward pass in the TS-LSTM model. Specifically, the equations for the input gate i_t, forget gate f_t, output gate o_t candidate state $\bar{C}_t$, memory cell C_{t-1} and final output activation state h_t for both the forward and backward LSTM processes are referred to as follows:

Forward Pass:

$$\overrightarrow{i_t} = \varphi(\overrightarrow{W_{xi}} \times [\overrightarrow{C_{t-1}}, \overrightarrow{h_{t-1}}, \overrightarrow{x_t} + \overrightarrow{b_i}) \tag{1}$$

$$\overrightarrow{f_t} = \varphi(\overrightarrow{W_{xf}} \times [\overrightarrow{C_{t-1}}, \overrightarrow{h_{t-1}}, \overrightarrow{x_t} + \overrightarrow{b_f}) \tag{2}$$

$$\overrightarrow{o_t} = \varphi(\overrightarrow{W_{xo}} \times [\overrightarrow{C_{t-1}}, \overrightarrow{h_{t-1}}, \overrightarrow{x_t} + \overrightarrow{b_o}) \tag{3}$$

$$\overrightarrow{\bar{C}_t} = tanh(\overrightarrow{W_{x\bar{c}}} * [\overrightarrow{h_{t-1}}, \overrightarrow{x_t} + \overrightarrow{b_{\bar{c}}}) \tag{4}$$

$$\overrightarrow{C_t} = \overrightarrow{f_t} * (\overrightarrow{C_{t-1}} + \overrightarrow{i_t} * \overrightarrow{\bar{C}_t}) \tag{5}$$

$$\overrightarrow{h_t} = \overrightarrow{o_t} * \overrightarrow{tanh}(\overrightarrow{C_t}) \tag{6}$$

Furthermore, we incorporated a backward pass into the proposed model to examine more meaningful insights.

Backward Pass:

$$\overleftarrow{i_t} = \varphi\left(\overleftarrow{W_{xt}} \times \left[\overleftarrow{C_{t-1}}, \overleftarrow{h_{t-1}}, \overleftarrow{x_t} + \overleftarrow{b_t}\right)\right. \tag{7}$$

$$\overleftarrow{f_t} = \varphi\left(\overleftarrow{W_{xf}} \times \left[\overleftarrow{C_{t-1}}, \overleftarrow{h_{t-1}}, \overleftarrow{x_t} + \overleftarrow{b_f}\right)\right. \tag{8}$$

$$\overleftarrow{o_t} = \varphi\left(\overleftarrow{W_{xo}} \times \left[\overleftarrow{C_{t-1}}, \overleftarrow{h_{t-1}}, \overleftarrow{x_t} + \overleftarrow{b_o}\right)\right. \tag{9}$$

$$\overleftarrow{\bar{C}_t} = tanh\left(\overleftarrow{W_{x\bar{c}}} * \overleftarrow{h_{t-1}}, \overleftarrow{x_t} + \overleftarrow{b_{\bar{c}}}\right) \tag{10}$$

$$\overleftarrow{C_t} = \overleftarrow{f_t} * \left[\overleftarrow{C_{t-1}}, \overleftarrow{l_t} * \overleftarrow{C_t}\right) \tag{11}$$

$$\overleftarrow{h_t} = \overleftarrow{o_t} * tanh\left(\overleftarrow{C_t}\right) \tag{12}$$

At time step t, for a given sequence $(x_1, x_2, \ldots, x_n)$ comprising n words, each represented as a dimensional vector, the activation of a word is denoted as $\vec{C}_t$, (representing the left-to-right context) and $\overleftarrow{C}_t$ (representing the right-to-left context) using forward and backward LSTM, respectively.

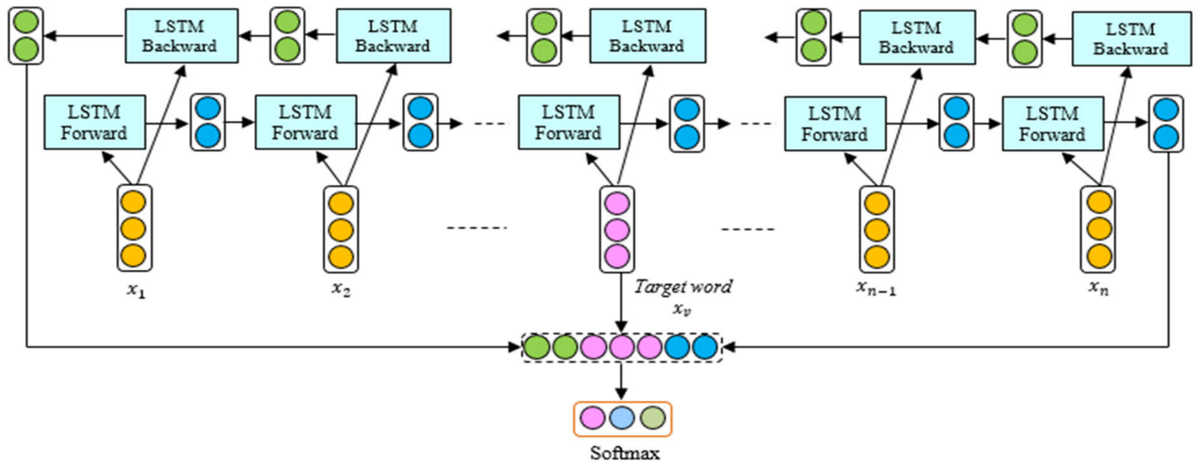


FIGURE 11. The proposed Two-State LSTM mechanism for depressed content detection.

These two context illustrations are then combined into a single context representation.

$$\bar{C} = f\left(V\left[\overrightarrow{C}_t : \overleftarrow{C_t}\right] + k\right) \quad (15)$$

D. SEQUENTIAL MECHANISM BY PRE-FEATURE ATTENTION TS-LSTM

The mechanism of Pre-feature attention in TS-LSTM is utilized to combine information from both preceding words and successive words, facilitating the initial comprehension step in depression detection. Within the framework of feature-attention TS-LSTM, the attention mechanism proves to be highly effective in extracting valuable insights from lengthy sentence reviews [87], aiding in categorizing emotions based on word-level characteristics. Furthermore, LSTM manages the flow of information within its units through a gating mechanism, and the two-state LSTM mechanism integrates contexts from both previous and subsequent words effectively within desirable contexts [88].

Moreover, the pre-feature attention strategy consists of both a forward and a backward sub-state. The forward sub-state systematically handles words from the embedding layer, commencing from the start. On the contrary, the backward sub-state conducts computations in reverse order, in contrast to the forward sub-state, which operates in the forward direction.

Naturally, at time step t, the input feature x^k is utilized to initialize both the forward candidate state $\vec{h}_{t-1}$ and the backward candidate state $\overleftarrow{h}_{t-1}$. These candidate states are subsequently employed in the pre-feature attention TS-LSTM. The previous and current states of the forward sub-state $\vec{C}_t$, and the backward sub-state $\overleftarrow{C}_t$ are also incorporated into this procedure and are mentioned in equations (13), (14), (15) as follows:

$$\bar{C}_t = tanh\left(\overrightarrow{W_x^{(\bar{C})}} * \left[\overrightarrow{h_{t-1}, x_t}\right] + \overrightarrow{b_{(\bar{C})}}\right) \quad (13)$$

$$\bar{C}_t = tanh\left(\overleftarrow{W_x^{(\bar{C})}} * [\overleftarrow{h_{t-1}, x_t}] + \overleftarrow{b_{(C)}}\right) \quad (14)$$

E. WORD-FEATURE SEIZING BY ATTENTION MECHANISM

After receiving the final output from the hidden state of the first layer, our architecture utilizes the attention mechanism in the feature-attention process to ascertain sentence polarity by concentrating on pertinent contexts at the word level. Fig. 12 depicts the detailed structure of the attention mechanism applied in our developed approach. Furthermore, Fig. 12 illustrates the configuration of o_t^k at k^{th} time step, which is produced by the attention mechanism in the following manner:

$$o_t^k = \frac{exp(e_t^k)}{\sum_{i=1}^{m} exp(e_t^i)} \quad (16)$$

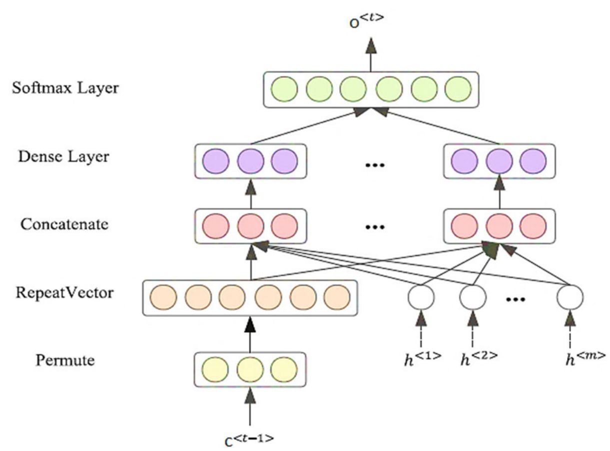


FIGURE 12. The intricate configuration of the attention mechanism.

In equation 17, the score fashion denoted by e_t^k for the memory cell at the time step k^{th} is denoted as e_t^k.

$$e_t^k = \left[c_{t-1}^T h_1, c_{t-1}^T h_2, \ldots, c_{t-1}^T h_m\right] \quad (17)$$

Let's express the hidden unit h_k of the pre-feature attention TS-LSTM at time-step, and also denote the memory cell

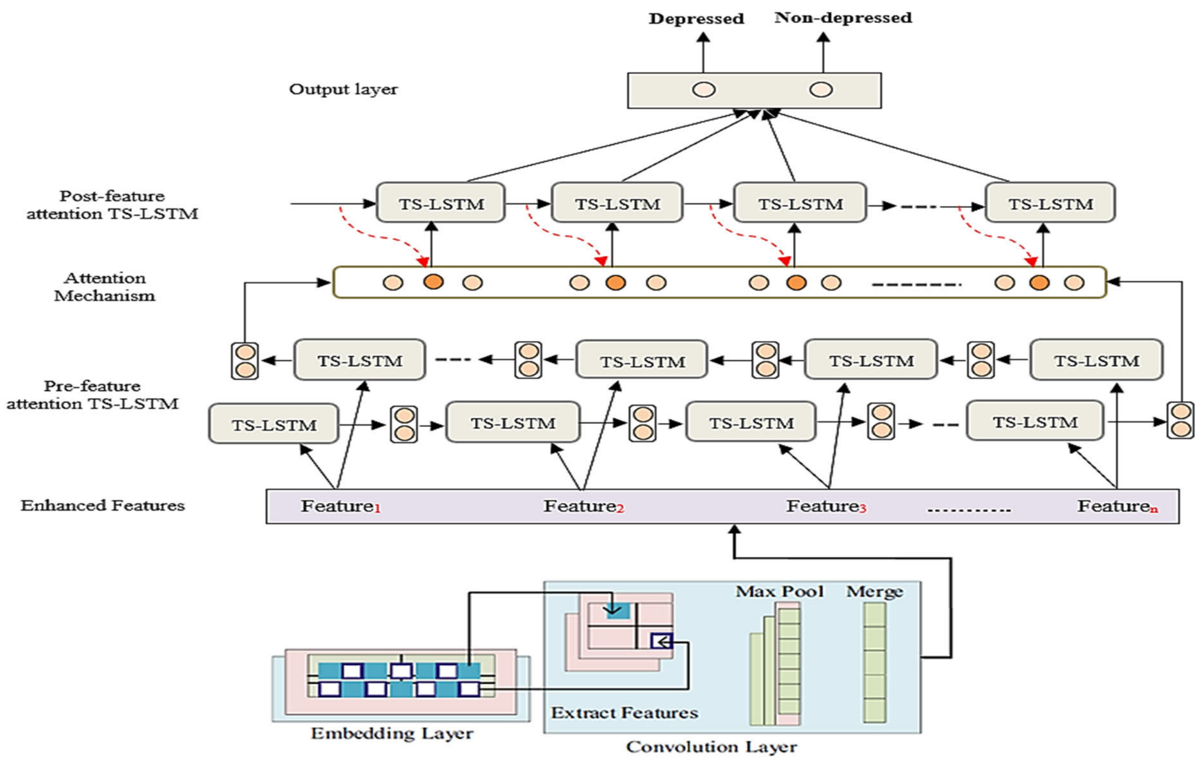


FIGURE 13. The overall architecture of the developed CNN+TS-LSTM based on the feature-attention mechanism.

illustrated by c_{t-1} at the k^{th} time-step in the post-feature attention LSTM. Finally, we can derive the attention output using equation (18) as follows:

$$o_t = \sum\nolimits_{k=1}^{m} o_t^k h_k \tag{18}$$

F. POST-FEATURE ATTENTION TS-LSTM

Utilizing the feature attention mechanism, we employ a post-feature attention TS-LSTM, which is subsequently accompanied by a word-feature seizing strategy. This combination is designed to capture sentence-level information by iteratively learning, intentionally mimicking human-like comprehension. Furthermore, during the second phase of feature attention "TS-LSTM, we incorporate a post-feature attention LSTM that mimics the decoding function, extracting the predicted features produced by the pre-feature attention LSTM and the attention mechanism layer. All the primary equations governing the post-feature attention TS-LSTM are identical to those of the conventional LSTM, except the absence of the candidate state is referred to in equation (19).

$$\bar{C}_t = tanh\left(W_x^{\bar{C}} * [h_{t-1}, x_t] + b_{\bar{C}}\right) \tag{19}$$

Subsequently, a Rectified linear activation unit (Relu) is employed for the final classification task to predict the depression class label ('Depressed' or 'Non-depressed') for the depression detection datasets. Consequently, we designed an LSTM mechanism enhanced with feature attention explicitly tailored for depression detection.

The overall architecture of the developed CNN+TS-LSTM based on the feature-attention mechanism is portrayed in Fig. 13. Fig. 14 demonstrates the parameters and variables of the CNN network composed of three convolution layers with the number of filters. The first two layers in CNN consist of 128 filters with a 3 × 3 kernel size, utilizing both sigmoid and ReLU as activation functions. The third convolution layer comprises 64 filters with a 3 × 3" kernel size and a sigmoid activation function. Subsequently, a Max-pool layer with a 4 × 4 kernel size is utilized. Lastly, a TS-LSTM is utilized based on a feature attention mechanism with slightly diverse hidden computations. The two-state nature of the computations, allowing information from both forward and backward directions, addresses the limitation of RNN by incorporating information from both preceding and subsequent steps. A "dropout layer" with a "keep probability" of 0.1 is employed to mitigate overfitting on the training data. The output layer employs ReLU activation, and the network is trained using "binary_crossentropy" loss and the "root mean squared propagation (RMSprop)" optimizer.

VIII. EXPERIMENTAL RESULTS AND DISCUSSION

This section describes the experimental configurations and assessment criteria employed in our research. Furthermore, this study examines the outcomes derived from implementing experiments utilizing the CNN+TS-LSTM methodology. A comparative analysis of our proposed approach and traditional approaches is also included. This section also provides insights into the statistical analysis and visual

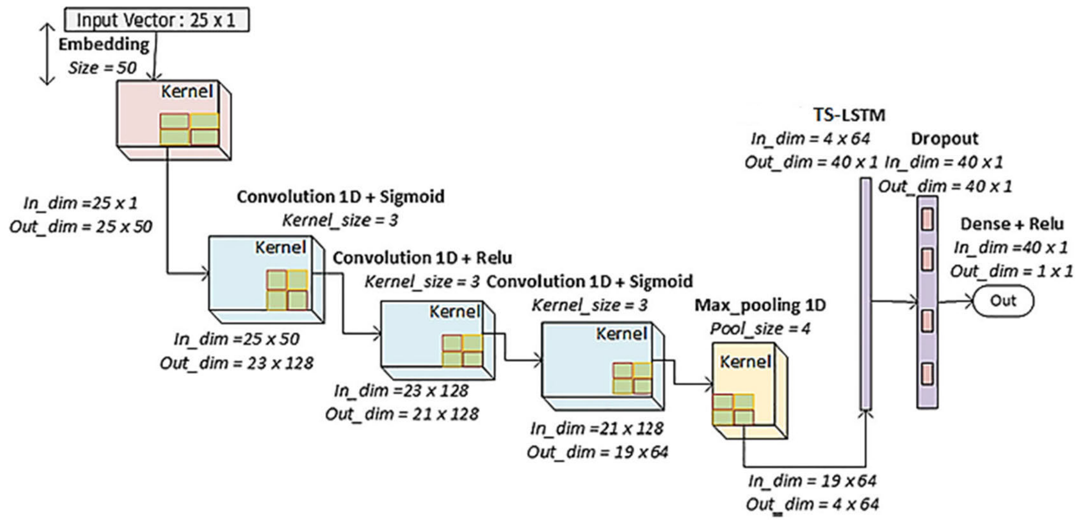


FIGURE 14. Parameters and layers used in the CNN+TS-LSTM hybrid model.

representations of the linguistic dialects utilized by depressed and non-depressed users.

A. EXPERIMENTAL DESIGN

Several experiments were performed to evaluate the effectiveness of the proposed study and existing research. The simulations of the proposed study were executed on a computer equipped with an Intel Core i7-3770CPU with a speed 3.40 GHz processor and 32 GB of RAM, operating on the Windows 10 OS. For data pre-processing and analysis, Python programming 3.12 compiler and Anaconda were employed as the development environment, supplemented by TensorFlow 2.4.1 and Keras 2.3, and for the word embedding process, the NLTK as essential libraries (deep learning tools). Furthermore, concise details regarding hyperparameter settings and evaluation metrics to enhance the developed mechanism are provided in subsequent subsections.

B. EVALUATION METRICS

Several evaluation metrics are utilized to assess the effectiveness of the developed model in predicting depressive symptoms, particularly in comparison to existing models. The assessment measures encompassed accuracy, precision, recall, and F-measure. Furthermore, the AUC curve can be described as the precise essential of the curve depicting changes in classification.

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \tag{20}$$

$$Precision = \frac{TP}{(TP + FP)} \tag{21}$$

$$Sensitivity = Recall = \frac{TP}{(TP + FN)} \tag{22}$$

$$\mathrm{F1} - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{23}$$

$$AUC = \frac{1}{2} = \left(\frac{TP}{TP + FN}\right) + \left(\frac{TN}{TN + FP}\right) \tag{24}$$

where TP, TN, FP, and FN refer to the True Positive, True Negative, False Positive, and False Negative, respectively.

IX. RESULTS ANALYSIS

In the section describing the experimental setup, it is evident that the gathered datasets have undergone normalization and amalgamation to be utilized in training, evaluating, and ultimately testing the Deep Learning (DL) approaches. The datasets consist of various categories, including texts related to depression.

A. FIRST DATASET AND EXPERIMENTS

The primary purpose of this study is to explore and evaluate the capability of the proposed model based on the feature attention mechanism to construct an effective classification model for detecting depressive content in user text. The performance of the developed model and comparative approaches were evaluated using accuracy, precision, recall, and F1 matrix. The table results indicate that all approaches have demonstrated notably favorable outcomes across two datasets when identifying depressive content in user text. Tables 7 illustrate the overall classification performance of the developed approach in terms of accuracy, precision, recall, F1, and support based on 100 epochs.

The existing studies have discussed Various deep learning approaches, and proposed methods have been applied to both small and large datasets. A dataset containing fewer than 10,000 records is considered "small," while those exceeding 10,000 records are deemed "large." The confusion matrix, error matrix, or contingency table is employed to evaluate and predict the performance of classification approaches. Figs. 15 depict the confusion matrix of the developed approach. For optimal assessment, this ratio, commonly

TABLE 7. Overall classification performance of the developed approach in terms of accuracy, precision, recall, F1, and support based on 100 epochs.

State	Accuracy	Precision	Recall	F1-score	Support
Depression	97.48	98.00	95.00	96.00	185
Non-Depression	97.16	97.00	98.00	98.00	991
Mean/Total	97.32	97.00	96.00	97.00	1179

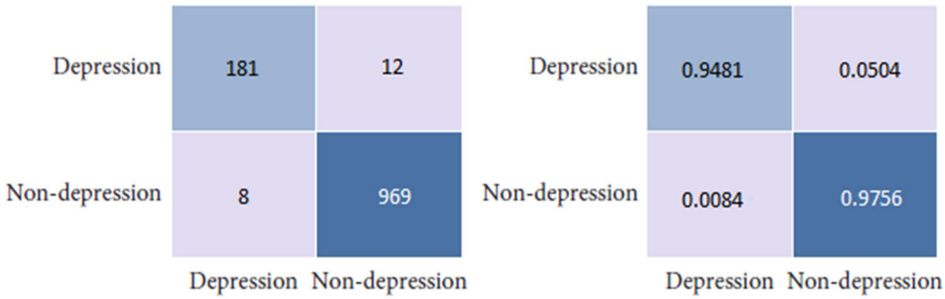


FIGURE 15. Confusion matrix of the proposed approach.

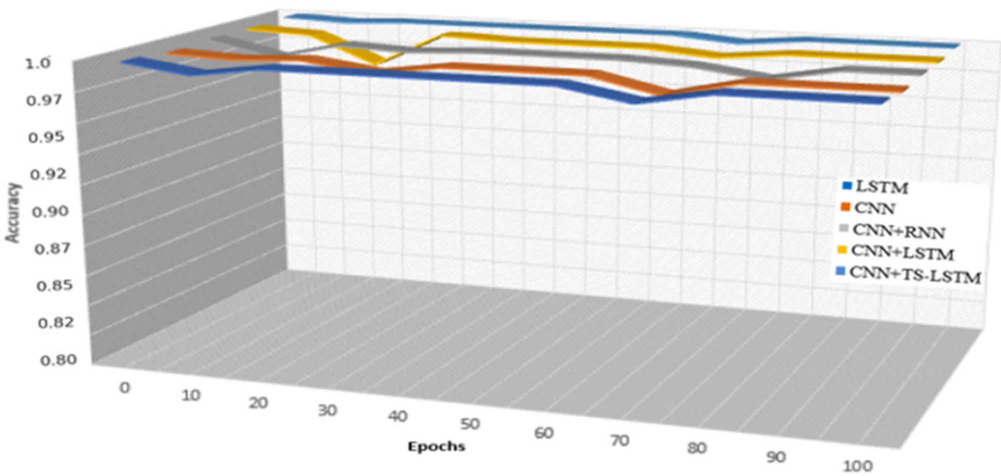


FIGURE 16. Accuracy results and epochs.

employed in many studies, dictates an 80:20 split for training and testing purposes. The assessment of the effectiveness of the depression prediction approach involves visualizing the prediction ratios' distribution across different classes in the test data with known actual values, represented as a confusion matrix. Fig. 15 display the confusion matrices for the proposed CNN+TS-LSTM model using the first dataset. CNN+TS-LSTM exhibits a notably higher count of true positive and true negative values, affirming its efficiency as a classifier for our dataset. Fig. 16 illustrates the accuracy and epoch of the proposed approach across the entire training dataset. The epoch denotes the training data cycle and the corresponding achieved accuracy, showcasing an upward trend where accuracy increases with each successive epoch.

B. SECOND DATASET AND EXPERIMENTS

The second dataset comprised 11,790 text samples, encompassing both depression and non-depression texts. Specifically, 1,530 samples were categorized as depression texts, while the remaining 10,260 were non-depression texts. In the second phase of experiments, we utilized five-fold cross-validation with the proposed approach, utilizing LSTM on robust features. Experimental results demonstrated that the proposed approach obtained superior performance using the Tweeter dataset (depicted in Fig. 17). Confusion matrices were utilized in the second set of experiments, including 80% data for training and the remaining 20% for testing. The experimental findings highlight the excellent performance of the developed CNN+TS-LSTM model.

C. COMPARATIVE ANALYSIS

The effectiveness of the proposed hybrid CNN+TS-LSTM approach in predicting depression is assessed by comparing it with LSTM and CNN models, using several evaluation metrics such as accuracy, precision, recall, F1-score, and specificity on the test set of the selected dataset. Table 8 illustrates that the proposed CNN+TS-LSTM approach outperforms both LSTM and CNN, demonstrating higher

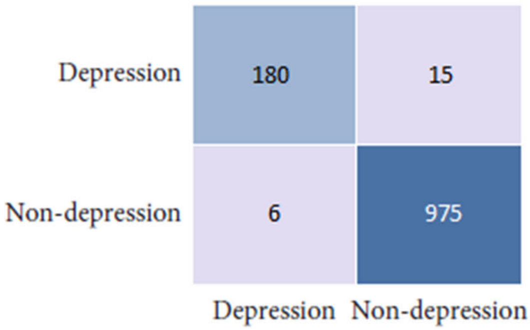


FIGURE 17. Confusion matrix of the proposed approach.

TABLE 8. Results from testing the proposed method and comparing it with other approaches, focusing on accuracy, precision, recall, F1-score, and specificity.

Evaluation Metrics	LSTM	CNN	CNN+TS-LSTM
Accuracy	92.71	92.89	97.23
Precision	92.64	93.42	98.57
Recall	95.21	93.96	97.13
F1-Score	93.91	93.68	97.84
Specificity	90.52	91.57	97.93

accuracy (97.23%), precision (98.57%), recall (97.13%) F1-score (97.84%), and specificity (97.93%). The primary objective of the developed method is to improve precision while keeping a consistently reliable recall, ensuring minimal inaccurate depression predictions. The performance metrics shown in Table 8 are visually depicted in Fig. 18, revealing that the LSTM and CNN approaches exhibit the lowest prediction performance compared to CNN+TS-LSTM.

Nevertheless, recent studies and practical observations have discovered that the LSTM analyzed a word in only the forward linguistic context. Hence, the LSTM must acquire contextual understanding through a backward pass. Furthermore, it is important to acknowledge that the forward pass and the backward pass shape the interpretation of a word in any language model. Therefore, this study employs "two-state LSTM to handle two directions, one for the positive time direction, referred to as the forward state, and another for the negative time direction, known as the backward state. Considering both sides, information is crucial as the semantics of a word in a sentence are influenced by previous historical information and subsequent information. Combining CNN's deep feature extraction capabilities with analyzing both state information provided by TS-LSTM, the proposed hybrid approach achieves the best results in depression prediction. To enhance the clarity of model performance analysis, Fig. 19 presents a comparative graph of the proposed model per epoch, focused on accuracy, loss, and AUC.

The accuracy for training and testing data across various approaches follows a typical Hilton curve, stabilizing around 0.90, as depicted in the graphs. Nevertheless, the loss function, mainly noticeable in test data, reveals an erratic noise output for LSTM and CNN. Notably, the developed hybrid CNN+TS-LSTM model reduces propagated noise compared to its counterparts. Regarding AUC scores, both

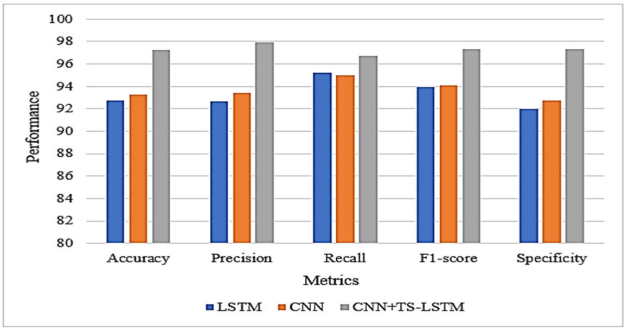


FIGURE 18. Comparison of LSTM, CNN, and CNN+TS-LSTM.

LSTM and CNN show an overall value of approximately 0.93, while the CNN+TS-LSTM model shows a slightly higher AUC value of 0.95214," demonstrating enhanced presentation. Consequently, the developed hybrid CNN+TS-LSTM approach enhances accuracy and reduces model loss in predicting depression analysis.

Table 9 compares F1-scores and accuracies between existing studies and the proposed hybrid CNN+TS-LSTM approach. Significantly, in comparison to previous research on predicting depression by using Twitter data, our hybrid model not only increases accuracy but also elevates the overall F1-score. The hybrid CNN+TS-LSTM model outperforms conventional methods, achieving an accuracy of 97.23. However, the CNN-LSTM model was proposed by Yang P.J. [78] achieved an excellent accuracy of 98.9%, which indeed suggests a strong performance as compared to our proposed approach in detecting depression from Twitter posts. This indicates that exploring hybrid deep learning approaches in the future could be a promising avenue for advancing depression analytics. The findings suggest that CNN is proficient at extracting specific features

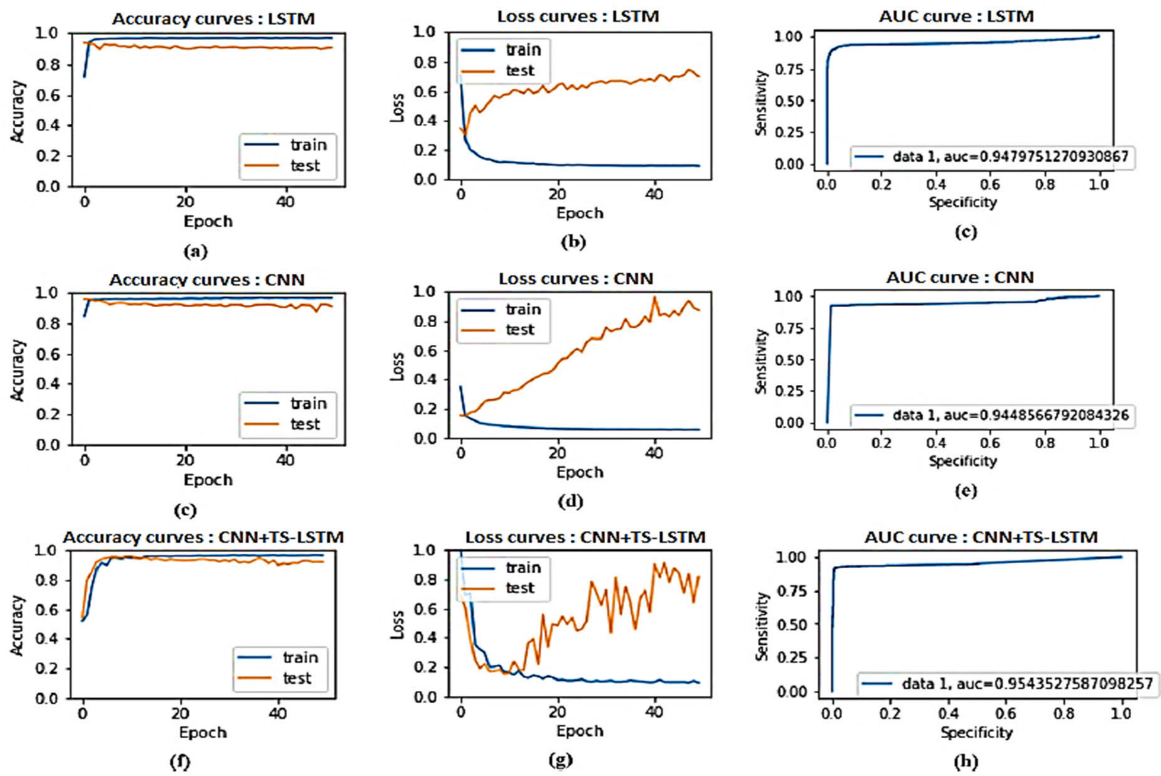


FIGURE 19. Comparing graphs includes: a graph depicting the accuracy of LSTM, b graph illustrating the loss of LSTM, c graph showing the AUC of LSTM, d accuracy graph for CNN, e loss graph for CNN, f AUC graph for CNN, g accuracy graph for CNN+TS-LSTM, h loss graph for CNN+TS-LSTM, and i AUC graph for CNN+TS-LSTM.

TABLE 9. Comparison of the performance of the proposed approach with traditional studies using accuracy and F1-score.

Existing studies/Evaluation metrics	Accuracy%	F1-score%
Adhikari et al. [46]	89.00	88.70
Mahato and Paul [50]	93.33	--
Ibitoye [58]	88.00	84.00
Trifan et al [59]	96.00	72.00
Kamil and Abbas [60]	93.80	93.74
Thekkekara et al. [61]	96.71	89.00
Farima et al. [64]	91.70	--
Filho et al. [68]	89.00	--
Nova [69]	77.00	--
AlSagri and Ykhlef [70]	83.12	79.63
Kim et al. [72]	96.96	98.44
Lin et al. [89]	88.40	93.60
Shen et al. [90]	85.00	85.00
Zhou et al. [91]	--	84.00
Tong et al. [92]	89.00	90.00
Tong et al. [93]	91.50	91.00
Naseem et al. [94]	94.00	--
Jamil et al. [95]	75.00	73.00
Nadeem et al. [96]	81.00	86.00
Shuai et al. [97]	90.40	--
Samuel et al. [98]	91.00	--
Nareshkumar [101]	92.00	--
Yang [78]	98.90	--
Zogan et al. [1]	85.20	85.20
Zogan et al. [1]	86.90	86.90
Proposed model	**97.23**	**97.84**

from various positions within a sentence but falls short in capturing the contextual aspects of individual word tokens. Combining convolution, pooling, and fully connected layers enables CNN to adjust and acquire crucial features through backpropagation. Convolutional operations facilitate weight sharing among neighboring positions, enabling kernels to extract local information within a specified space.

Furthermore, CNN utilizes pooling operations to discern relevant feature patterns, with the Max-pooling layer selecting the most significant feature from each input feature map and discarding others. This process effectively reduces the number of input features. On the other hand, LSTM effectively addresses challenges such as the 'vanishing gradient' and 'exploding gradient' issues and excels in contextual feature extraction; its limitation lies in being unidirectional. This suggests disregarding how the succeeding word in a sentence affects the present context. Conversely, the two-state LSTM approach, with its both-direction nature, focuses on capturing crucial contextual features within a sentence. Moreover, the embedding layer extracts word-level embeddings and includes character-level embedding vectors.

Therefore, the CNN+TS-LSTM approach, as proposed, effectively addresses the constraints of CNN, LSTM, and RNN by concurrently capturing local features and contextual information from the convolutional layer's outputs. This substantiates our hypothesis that the fusion of CNN and two-state LSTM enhances the extraction of localized features and utilizes the improved RNN functionality of TS-LSTM with its multidirectional capabilities.

D. STATISTICAL ANALYSIS

This investigation uses a t-test to assess the important distinction between the two categories. Specifically, tweets categorized tweets classified as depressed and non-depressed. The t-test [98]" is a parametric method used to determine dissimilarity between two sets. Its primary objective is to determine if a noteworthy distinction exists in the average string length between depression-related tweets and non-depression tweets. The t-test statistic is expressed by Eq. (25), wherein 't' refers to the calculated t-value, $\bar{x_1}$ and $\bar{x_2}$ denote the means of the two classes (depressed and non-depressed) being compared, $\bar{x_1}$ - $\bar{x_2}$ signifies the modification in sample means, s_1 and s_2 represent the standard error of the two distributions and n_1 and n_2 denote the respective number of comments in each class. For a t-test, the degrees of freedom ('df') are determined by the smaller of the two values ($n_1 - 1, n_2 - 1$).

$$t = \frac{\bar{x_1} - \bar{x_2}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \tag{25}$$

For the statistical analysis, we calculate the average length of both depressed and non-depressed tweets. The lengths of the tweets are measured and depicted in Fig. 20. Density distribution plots illustrating the string lengths are presented in Fig. 21a and 21b. Subsequently, the mean of the two distributions is determined: $\mu = 9.21$ for depressed strings and $\mu = 6.59$ for non-depressed strings. The null hypothesis assumes equality in the mean of two population sample circulations (H_0: $\mu = \mu 2$). A t-test is employed to evaluate the alternative hypothesis contradicting the null hypothesis.

The P-value plays a crucial role as a threshold and serves as a test statistic for the test results. It is pivotal in deciding whether to reject the null hypothesis (H_0: $\mu = \mu 2$) or assert that the two groups are distinct. The p-value obtained from the t-test is 0.000. Nevertheless, we compared this p-value with the chosen critical value α, which, in our research, is set at $\alpha = 0.02$.

1) When the p-value exceeds α (Critical value), the t-test does not reject the null hypothesis, indicating that the distributions share the same mean.
2) If the p-value is less than α (Critical value), the t-test rejects the null hypothesis, signifying a substantial change in the means of the sample distributions.

In this investigation, a two-tailed test was conducted, yielding a calculated t-test statistic value of 33.979. Comparing this to the tabulated value for t-test statistics with $\alpha = 0.02$ and degrees of freedom = 1 (31.783), as displayed in Table 10, we observe that the calculated value exceeds the tabulated one. The inference drawn from the t-test is the rejection of the null hypothesis, indicating a significant difference in the distribution of tweet lengths between depressed and non-depressed individuals at a 0.02 level of significance.

TABLE 10. T-test results.

t-test		
Test statistic	DF	Sig. (2 tailed)
33.979	1	0.000

tidy_tweets	len_tidy_tweets
do as diagnos things depress confirm wrong as adhd believe	10
hate friend med far wk diagnos liver failure depression mild paranoia chronic fatigu	13
bulli years torn apart bulli diagnos depress anxieti	8
diagnos depress before	3
self help tip for deal depress mental illness diagnos bipolar medicin	11

FIGURE 20. Length of tweet for non-depression users.

The frequency plot in Fig. 22a, b illustrates the occurrence of words in the first 20 positions. Words such as anxiety, severe, depressed, bipolar, help etc., commonly appear among depressed users. In contrast, non-depressed users frequently use words like love, like, know, think, etc. Analyzing the plot suggests that individuals using the former set of words are more prone to having a self-analysis indicative of depression, as actively conveyed on diverse social media platforms.

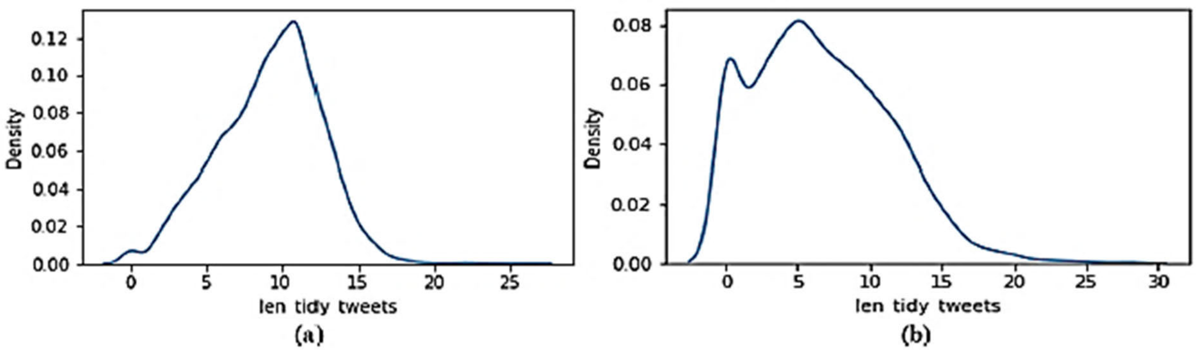


FIGURE 21. Plot depicting the density distribution for (a) depressed users and (b) non-depressed users.

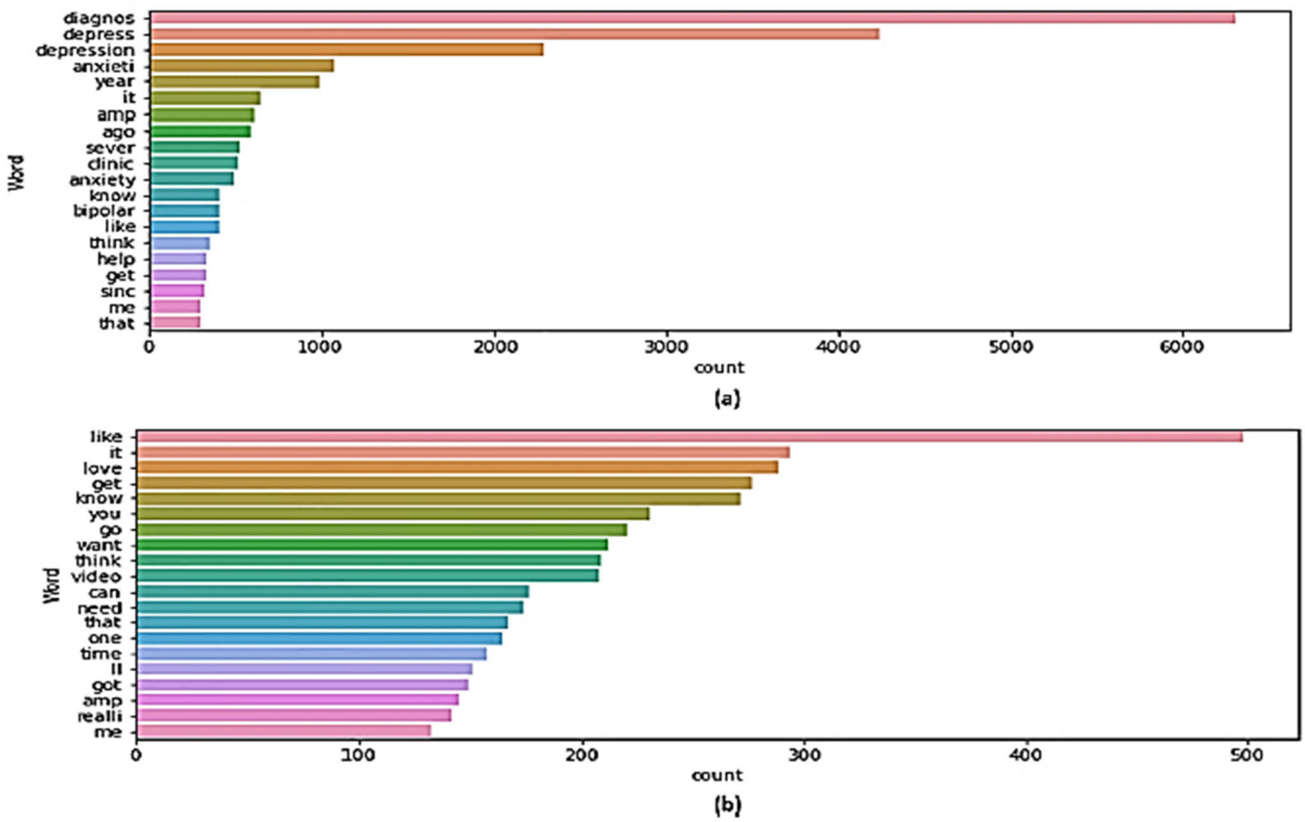


FIGURE 22. Word frequency distribution charts comparing a) depressed users and b) non-depressed users.

X. CONCLUSION AND FUTURE DIRECTION

Depression stands as a prevalent mental disorder that extends its reach globally. It is crucial to enhance our understanding of depression at the individual, community, and global levels. Prioritizing the identification of this issue and providing support to those affected by depression is paramount. This study introduces a novel approach for predicting depression using two datasets, Facebook and Twitter. The proposed model categorizes users into non-depressed and depressed groups based on a real-world dataset. In this study, we proposed a hybrid deep learning algorithm known as CNN+TS-LSTM for depression detection, which mainly combines a Convolution Neural Network (CNN) and a Long Short-Term Memory (LSTM) model. An enhanced version of the LSTM approach, namely Two-State LSTM (TS-LSTM), is applied based on the feature attention mechanism. The proposed framework incorporates a feature attention mechanism into the TS-LSTM model, enhancing its ability to identify the relationships and extract keywords for depression detection by utilizing an attention layer.

Furthermore, our methodology involves feature extraction through convolution layers and improved recurrent network architectures. When comparing the proposed CNN+TS-LSTM approach with traditional studies, it was observed that the developed approach exhibited superior performance across several evaluation metrics. Experimental studies concluded that the proposed hybrid CNN+TS-LSTM approach attained the highest accuracy of ''97.23%, precision of 98.57%, recall of 97.13%, F1-score of 97.84%, and specificity of 97.93%.

There is significant potential for further investigation into this work in the future. For example, one could investigate diverse combinations of neural network layers and activation functions to enhance the model's accuracy. Additionally, future research might benefit from pre-trained language methods such as Deep contextualized word representations (ELMo) and Bidirectional Encoder Representations from Transformers (BERT). These methods could be trained on a comprehensive dataset comprising depression-related content from Facebook, tweets,'' and YouTube. While utilizing such pre-trained language techniques poses challenges due to sentence sequence length restrictions, thoroughly examining these models in this context is crucial to uncover their strengths and weaknesses. Our forthcoming research aims to extend this analysis to include the detection of other mental illnesses alongside depression, providing a more comprehensive understanding of complex mental health issues affecting individuals.

DATA AVAILABILITY

The data is publicly available at Kaggle: https://www.kaggle.com/datasets/infamouscoder/mental-health-social-media; https://www.kaggle.com/datasets/sahasourav17/students-anxiety-and-depression-dataset.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

[1] H. Zogan, I. Razzak, X. Wang, S. Jameel, and G. Xu, ''Explainable depression detection with multi-aspect features using a hybrid deep learning model on social media,'' *World Wide Web*, vol. 25, no. 1, pp. 281–304, Jan. 2022.

[2] P. Meshram and R. K. Rambola, ''Diagnosis of depression level using multimodal approaches using deep learning techniques with multiple selective features,'' *Expert Syst.*, vol. 40, no. 4, May 2023, Art. no. e12933.

[3] WHO. (2021). *Suicide*. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/suicide

[4] M. Fang, S. Peng, Y. Liang, C.-C. Hung, and S. Liu, ''A multimodal fusion model with multi-level attention mechanism for depression detection,'' *Biomed. Signal Process. Control*, vol. 82, Apr. 2023, Art. no. 104561.

[5] G. Tasci, H. W. Loh, P. D. Barua, M. Baygin, B. Tasci, S. Dogan, T. Tuncer, E. E. Palmer, R.-S. Tan, and U. R. Acharya, ''Automated accurate detection of depression using twin Pascal's triangles lattice pattern with EEG signals,'' *Knowledge-Based Syst.*, vol. 260, Jan. 2023, Art. no. 110190.

[6] *Depression and Other Common Mental Disorders: Global Health Estimates*, document (No. WHO/MSD/MER/ 2017.2), World Health Organization, 2017.

[7] R. C. Kessler and E. J. Bromet, ''The epidemiology of depression across cultures,'' *Annu. Rev. Public Health*, vol. 34, no. 1, pp. 119–138, Mar. 2013, doi: 10.1146/annurev-publhealth-031912-114409.

[8] *Depression and Other Common Mental Disorders: Global Health Estimates*, World Health Org., Geneva, Switzerland, 2017, pp. 1–24.

[9] A. H. Weinberger, M. Gbedemah, A. M. Martinez, D. Nash, S. Galea, and R. D. Goodwin, ''Trends in depression prevalence in the USA from 2005 to 2015: Widening disparities in vulnerable groups,'' *Psychol. Med.*, vol. 48, no. 8, pp. 1308–1315, Jun. 2018.

[10] H. Almeida, A. Briand, and M. J. Meurs, ''Detecting early risk of depression from social media user-generated content,'' in *Proc. CLEF Working Notes*, 2022, pp. 1–12.

[11] A. Biradar and S. G. Totad, ''Detecting depression in social media posts using machine learning,'' in *Proc. Int. Conf. Recent Trends Image Process. Pattern Recognit.* Singapore: Springer, 2018, pp. 716–725, doi: 10.1007/978-981-13-9187-3_64.

[12] Z. N. Vasha, B. Sharma, I. J. Esha, J. Al Nahian, and J. A. Polin, ''Depression detection in social media comments data using machine learning algorithms,'' *Bull. Electr. Eng. Informat.*, vol. 12, no. 2, pp. 987–996, Apr. 2023.

[13] M. Zulqarnain, H. Shah, R. Ghazali, O. Alqahtani, R. Sheikh, and M. Asadullah, ''Attention aware deep learning approaches for an efficient stress classification model,'' *Brain Sci.*, vol. 13, no. 7, p. 994, Jun. 2023.

[14] S. Wang, G. Peng, Z. Zheng, and Z. Xu, ''Capturing emotion distribution for multimedia emotion tagging,'' *IEEE Trans. Affect. Comput.*, vol. 12, no. 4, pp. 821–831, Oct. 2021, doi: 10.1109/TAFFC.2019.2900240.

[15] U. Yadav, A. K. Sharma, and D. Patil, ''Review of automated depression detection: Social posts, audio and video, open challenges and future direction,'' *Concurrency Comput., Pract. Exper.*, vol. 35, no. 1, Jan. 2023, Art. no. e7407.

[16] M. Du, S. Liu, T. Wang, W. Zhang, Y. Ke, L. Chen, and D. Ming, ''Depression recognition using a proposed speech chain model fusing speech production and perception features,'' *J. Affect. Disorders*, vol. 323, pp. 299–308, Feb. 2023.

[17] A. Cahuas, Z. He, Z. Zhang, and W. Chen, ''Relationship of physical activity and sleep with depression in college students,'' *J. Amer. College Health*, vol. 68, no. 5, pp. 557–564, Jul. 2020.

[18] M. Grasdalsmoen, H. R. Eriksen, K. J. Lønning, and B. Sivertsen, ''Physical exercise, mental health problems, and suicide attempts in university students,'' *BMC Psychiatry*, vol. 20, no. 1, pp. 1–11, Dec. 2020.

[19] WHO Organization. (2019). *Suicide Data*. World Health Org. [Online]. Available: https://www.who.int/teams/mentalhealth-and-substance-use/data-research/suicide-data.

[20] K. Nomura, S. Minamizono, E. Maeda, R. Kim, T. Iwata, J. Hirayama, K. Ono, M. Fushimi, T. Goto, K. Mishima, and F. Yamamoto, ''Cross-sectional survey of depressive symptoms and suicide-related ideation at a Japanese national university during the COVID-19 stay-home order,'' *Environ. Health Preventive Med.*, vol. 26, no. 1, pp. 1–9, Dec. 2021.

[21] S. Ghosh, A. Ekbal, and P. Bhattacharyya, ''A multitask framework to detect depression, sentiment and multi-label emotion from suicide notes,'' *Cognit. Comput.*, vol. 14, no. 1, pp. 110–129, Jan. 2022.

[22] S. Ji, S. Pan, X. Li, E. Cambria, G. Long, and Z. Huang, ''Suicidal ideation detection: A review of machine learning methods and applications,'' *IEEE Trans. Computat. Social Syst.*, vol. 8, no. 1, pp. 214–226, Feb. 2021.

[23] A. H. Saffar, T. K. Mann, and B. Ofoghi, ''Textual emotion detection in health: Advances and applications,'' *J. Biomed. Informat.*, vol. 137, Jan. 2023, Art. no. 104258.

[24] (2019). *Major Depressive Disorder Among Teens*. Pew Research Center. Accessed: Mar. 10, 2021. [Online]. Available: https://www.pewsocialtrends.org/2019/02/20/most-u-s-teens-see-anxiety-and-depression-as-a-major-problem-among-their-peers/

[25] A. Sood, M. Hooda, S. Dhirn, and M. Bhatia, ''An initiative to identify depression using sentiment analysis: A machine learning approach,'' *Indian J. Sci. Technol.*, vol. 11, no. 4, pp. 1–20, Jan. 2018, doi: 10.17485/ijst/2018/v11i4/119594.

[26] M. Birjali, A. Beni-Hssane, and M. Erritali, ''A method proposed for estimating depressed feeling tendencies of social media users utilizing their data,'' in *Proc. Int. Conf. Hybrid Intell. Syst.*, Cham, Switzerland, 2016, pp. 413–420, doi: 10.1007/978-3-319-52941-7_41.

[27] A. H. Uddin, D. Bapery, and A. S. M. Arif, ''Depression analysis from social media data in Bangla language using long short term memory (LSTM) recurrent neural network technique,'' in *Proc. Int. Conf. Comput., Commun., Chem., Mater. Electron. Eng. (IC4ME2)*, Rajshahi, Bangladesh, Jul. 2019, pp. 1–4, doi: 10.1109/ic4me247184.2019.9036528.

[28] J. J. Stephen and P. P., ''Detecting the magnitude of depression in Twitter users using sentiment analysis,'' *Int. J. Electr. Comput. Eng. (IJECE)*, vol. 9, no. 4, p. 3247, Aug. 2019, doi: 10.11591/ijece.v9i4.pp3247-3255.

[29] M. Z. Uddin and J. Torresen, ''Activity recognition using smartphone sensors, robust features, and recurrent neural network,'' in *Proc. 13th Int. Symp. Med. Inf. Commun. Technol. (ISMICT)*, Oslo, Norway, May 2019, pp. 1–6.

[30] J. A. Castillo, Y. C. Granados, and C. A. F. Ariza, "Patient-specific detection of atrial fibrillation in segments of ECG signals using deep neural networks," *Ciencia e Ingeniería Neogranadina*, vol. 30, no. 1, pp. 45–58, Nov. 2019.

[31] J. Cai, G. Zhou, M. Dong, X. Hu, G. Liu, and W. Ni, "Real-time arrhythmia classification algorithm using time-domain ECG feature based on FFNN and CNN," *Math. Problems Eng.*, vol. 2021, pp. 1–17, May 2021.

[32] M. Zulqarnain, A. K. Z. Alsaedi, R. Ghazali, M. G. Ghouse, W. Sharif, and N. A. Husaini, "A comparative analysis on question classification task based on deep learning approaches," *PeerJ Comput. Sci.*, vol. 7, p. e570, Aug. 2021.

[33] A. Awate, H. Sharma, R. Hatkar, M. Shahade, B. Nandwalkar, Y. Jadhav, and M. Kulkarni, "Deep care: Smart depression counselling system through emotion recognition and opinion mining using deep learning," Tech. Rep., 2023.

[34] S. Dalal, S. Jain, and M. Dave, "Early depression detection using textual cues from social data' A research agenda," in *Proc. Int. Health Inform. Conf. (IHIC)*. Singapore: Springer, 2022, pp. 393–406.

[35] A. Vázquez-Romero and A. Gallardo-Antolín, "Automatic detection of depression in speech using ensemble convolutional neural networks," *Entropy*, vol. 22, no. 6, p. 688, Jun. 2020.

[36] A. Y. Kim, E. H. Jang, S.-H. Lee, K.-Y. Choi, J. G. Park, and H.-C. Shin, "Automatic depression detection using smartphone-based text-dependent speech signals: Deep convolutional neural network approach," *J. Med. Internet Res.*, vol. 25, Jan. 2023, Art. no. e34474.

[37] A. Houtenville, S. Bach, and S. Paul, *Annual Report on People With Disabilities in America: 2023*. Durham, NH, USA: Institute Disability, Univ. New Hampshire, 2023.

[38] *Global Report on Health Equity for Persons With Disabilities*, World Health Organization, Geneva, Switzerland, 2022.

[39] G. Cicirelli, R. Marani, A. Petitti, A. Milella, and T. D'Orazio, "Ambient assisted living: A review of technologies, methodologies and future perspectives for healthy aging of population," *Sensors*, vol. 21, no. 10, p. 3549, May 2021.

[40] T. S. Poornappriya and R. Gopinath, "Application of machine learning techniques for improving learning disabilities," *Int. J. Electr. Eng. Technol. (IJEET)*, vol. 11, no. 10, pp. 392–402, 2020.

[41] D. Scrutinio, C. Ricciardi, L. Donisi, E. Losavio, P. Battista, P. Guida, M. Cesarelli, G. Pagano, and G. D'Addio, "Machine learning to predict mortality after rehabilitation among patients with severe stroke," *Sci. Rep.*, vol. 10, no. 1, p. 20127, Nov. 2020.

[42] I. Boukhennoufa, X. Zhai, V. Utti, J. Jackson, and K. D. McDonald-Maier, "Wearable sensors and machine learning in post-stroke rehabilitation assessment: A systematic review," *Biomed. Signal Process. Control*, vol. 71, Jan. 2022, Art. no. 103197.

[43] L. Lu, Y. Tan, M. Klaic, M. P. Galea, F. Khan, A. Oliver, I. Mareels, D. Oetomo, and E. Zhao, "Evaluating rehabilitation progress using motion features identified by machine learning," *IEEE Trans. Biomed. Eng.*, vol. 68, no. 4, pp. 1417–1428, Apr. 2021.

[44] M. Tschuggnall, V. Grote, M. Pirchl, B. Holzner, G. Rumpold, and M. J. Fischer, "Machine learning approaches to predict rehabilitation success based on clinical and patient-reported outcome measures," *Informat. Med. Unlocked*, vol. 24, Jan. 2021, Art. no. 100598.

[45] H. Aarons and E. Willis, *The Sociological Quest: An Introduction To the Study of Social Life*. New York, NY, USA: Taylor & Francis, 2022.

[46] S. Adhikari, S. Thapa, U. Naseem, P. Singh, H. Huo, G. Bharathy, and M. Prasad, "Exploiting linguistic information from nepali transcripts for early detection of Alzheimer's disease using natural language processing and machine learning techniques," *Int. J. Hum.-Comput. Stud.*, vol. 160, Apr. 2022, Art. no. 102761.

[47] M. D. Aperocho and M. A. E. Tarusan, "Depressive language on a virtual common: A critical discourse analysis," *Int. J. Multidisciplinary: Appl. Bus. Educ. Res.*, vol. 3, no. 3, pp. 454–475, Mar. 2022.

[48] I. M. Nick, "Black and blue: Revealing suicidality in the poetry of the afro-German writer-activist, may ayim," in *The Contemporary Writer and Their Suicide*. Cham, Switzerland: Springer, 2023, pp. 261–278.

[49] S. Mahato and S. Paul, "Detection of major depressive disorder using linear and non-linear features from EEG signals," *Microsyst. Technol.*, vol. 25, no. 3, pp. 1065–1076, Mar. 2019.

[50] H. Ke, D. Chen, B. Shi, J. Zhang, X. Liu, X. Zhang, and X. Li, "Improving brain E-health services via high-performance EEG classification with grouping Bayesian optimization," *IEEE Trans. Services Comput.*, vol. 13, no. 4, pp. 696–708, Jul. 2020.

[51] M. Kang, H. Kwon, J.-H. Park, S. Kang, and Y. Lee, "Deep-asymmetry: Asymmetry matrix image for deep learning method in pre-screening depression," *Sensors*, vol. 20, no. 22, p. 6526, Nov. 2020.

[52] S. Mahato and S. Paul, "Classification of depression patients and normal subjects based on electroencephalogram (EEG) signal using alpha power and theta asymmetry," *J. Med. Syst.*, vol. 44, no. 1, pp. 1–8, Jan. 2020.

[53] V. Mrazek, S. Jawed, M. Arif, and A. S. Malik, "Effective EEG feature selection for interpretable MDD (major depressive disorder) classification," in *Proc. Genetic Evol. Comput. Conf.*, Jul. 2023, pp. 1427–1435.

[54] W. Dang, Z. Gao, X. Sun, R. Li, Q. Cai, and C. Grebogi, "Multilayer brain network combined with deep convolutional neural network for detecting major depressive disorder," *Nonlinear Dyn.*, vol. 102, no. 2, pp. 667–677, Oct. 2020.

[55] R. A. Movahed, G. P. Jahromi, S. Shahyad, and G. H. Meftahi, "A major depressive disorder classification framework based on EEG signals using statistical, spectral, wavelet, functional connectivity, and nonlinear analysis," *J. Neurosci. Methods*, vol. 358, Jul. 2021, Art. no. 109209.

[56] E. Aydemir, T. Tuncer, S. Dogan, R. Gururajan, and U. R. Acharya, "Automated major depressive disorder detection using melamine pattern with EEG signals," *Appl. Intell.*, vol. 51, no. 9, pp. 6449–6466, Sep. 2021.

[57] H. W. Loh, C. P. Ooi, E. Aydemir, T. Tuncer, S. Dogan, and U. R. Acharya, "Decision support system for major depression detection using spectrogram and convolution neural network with EEG signals," *Expert Syst.*, vol. 39, no. 3, Mar. 2022, Art. no. e12773.

[58] A. O. Ibitoye, R. F. Famutimi, D. O. Olanloye, and E. Akioyamen, "User-centric social opinion and clinical behavioral model for depression detection," *Int. J. Intell. Inf. Syst.*, vol. 10, pp. 69–76, Jan. 2021.

[59] A. Trifan, R. Antunes, S. Matos, and J. L. Oliveira, "Understanding depression from psycholinguistic patterns in social media texts," in *Proc. Eur. Conf. Inf. Retr.*, Apr. 2020, pp. 402–409.

[60] R. Kamil and A. R. Abbas, "Predicating depression on Twitter using hybrid model BiLSTM-XGBOOST," *Bull. Electr. Eng. Informat.*, vol. 12, no. 6, pp. 3620–3627, Dec. 2023.

[61] J. P. Thekkekara, S. Yongchareon, and V. Liesaputra, "An attention-based CNN-BiLSTM model for depression detection on social media text," *Expert Syst. Appl.*, vol. 249, Sep. 2024, Art. no. 123834.

[62] A. H. Orabi, P. Buddhitha, M. H. Orabi, and D. Inkpen, "Deep learning for depression detection of Twitter users," in *Proc. 5th Workshop Comput. Linguistics Clinical Psychol., From Keyboard Clinic*, Jun. 2018, pp. 88–97.

[63] P.-Y. Genest, L.-W. Goix, Y. Khalafaoui, E. Egyed-Zsigmond, and N. Grozavu, "French translation of a dialogue dataset and text-based emotion detection," *Data Knowl. Eng.*, vol. 142, Nov. 2022, Art. no. 102099.

[64] I. Fatima, B. U. D. Abbasi, S. Khan, M. Al-Saeed, H. F. Ahmad, and R. Mumtaz, "Prediction of postpartum depression using machine learning techniques from social media text," *Expert Syst.*, vol. 36, no. 4, Aug. 2019, Art. no. e12409.

[65] S. G. Burdisso, M. Errecalde, and M. Montes-y-Gómez, "A text classification framework for simple and effective early depression detection over social media streams," *Expert Syst. Appl.*, vol. 133, pp. 182–197, Nov. 2019.

[66] G. Shen, J. Jia, L. Nie, F. Feng, C. Zhang, T. Hu, and W. Zhu, "Depression detection via harvesting social media: A multimodal dictionary learning solution," in *Proc. IJCAI*, Aug. 2017, pp. 3838–3844.

[67] A. Chatterjee, U. Gupta, M. K. Chinnakotla, R. Srikanth, M. Galley, and P. Agrawal, "Understanding emotions in text using deep learning and big data," *Comput. Hum. Behav.*, vol. 93, pp. 309–317, Apr. 2019.

[68] E. M. D. S. Filho, H. C. V. Rey, R. M. Frajtag, D. M. A. Cook, L. N. D. de Carvalho, A. L. P. Ribeiro, and J. Amaral, "Can machine learning be useful as a screening tool for depression in primary care?" *J. Psychiatric Res.*, vol. 132, pp. 1–6, Jan. 2021.

[69] K. Nova, "Machine learning approaches for automated mental disorder classification based on social media textual data," *Contemp. Issues Behav. Social Sci.*, vol. 7, no. 1, pp. 70–83, 2023.

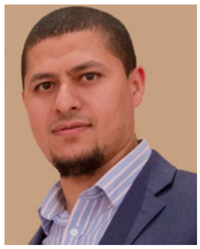
[70] H. S. Alsagri and M. Ykhlef, "Machine learning-based approach for depression detection in Twitter using content and activity features," *IEICE Trans. Inf. Syst.*, vol. E103.D, no. 8, pp. 1825–1832, 2020.
[71] M. R. Islam, M. A. Kabir, A. Ahmed, A. R. M. Kamal, H. Wang, and A. Ulhaq, "Depression detection from social network data using machine learning techniques," *Health Inf. Sci. Syst.*, vol. 6, no. 1, pp. 1–12, Dec. 2018.
[72] J. Kim, J. Lee, E. Park, and J. Han, "A deep learning model for detecting mental illness from user content on social media," *Sci. Rep.*, vol. 10, no. 1, p. 11846, Jul. 2020.
[73] M. M. Aldarwish and H. F. Ahmad, "Predicting depression levels using social media posts," in *Proc. IEEE 13th Int. Symp. Auto. Decentralized Syst. (ISADS)*, Bangkok, Thailand, Mar. 2017, pp. 277–280.
[74] M. Gaikar, J. Chavan, K. Indore, and R. Shedge, "Depression detection and prevention system by analysing tweets," in *Proc. Conf. Technol. Future Cities (CTFC)*, Mar. 2019, pp. 1–6.
[75] T.-L. Nguyen, S. Kavuri, and M. Lee, "A fuzzy convolutional neural network for text sentiment analysis," *J. Intell. Fuzzy Syst.*, vol. 35, no. 6, pp. 6025–6034, Dec. 2018.
[76] M. Z. Uddin, K. K. Dysthe, A. Følstad, and P. B. Brandtzaeg, "Deep learning for prediction of depressive symptoms in a large textual dataset," *Neural Comput. Appl.*, vol. 34, no. 1, pp. 721–744, Jan. 2022.
[77] Y. T. Kiong, "An initial study of depression detection on Mandarin textual through BERT model," in *Proc. 14th ACM Web Sci. Conf.*, Jun. 2022, pp. 459–463.
[78] P. J. Yang. (2020). *Detect Depression in Twitter Posts*. [Online]. Available: https://github.com/peijoy/DetectDepressionInTwitterPosts
[79] P. Girish. (2021). *Depression Detection Using ML Algorithm*. [Online]. Available: https://github.com/patilgirish815/Depression_Detection_Using_ML
[80] A. Rathee. (2021). *Smile Twitter Emotion Dataset*. [Online]. Available: https://www.kaggle.com/datasets/ashkhagan/smile-twitter-emotion-dataset
[81] B. Rieder, "Studying Facebook via data extraction: The Netvizz application," in *Proc. 5th Annu. ACM Web Sci. Conf.*, May 2013, pp. 346–355.
[82] J. Roesslein. (2009). *Tweepy Documentation*. [Online]. Available: http://tweepy.readthedocs.io/en/v3
[83] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," 2013, *arXiv:1301.3781*.
[84] M. Zulqarnain, R. Ghazali, M. G. Ghouse, Y. M. M. Hassim, and I. Javid, "Predicting financial prices of stock market using recurrent convolutional neural networks," *Int. J. Intell. Syst. Appl.*, vol. 12, no. 6, pp. 21–32, Dec. 2020.
[85] M. Zulqarnain, S. Abd, R. Ghazali, N. Mohd, M. Aamir, and Y. Mazwin, "An improved deep learning approach based on variant two-state gated recurrent unit and word embeddings for sentiment classification," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 1, pp. 594–603, 2020.
[86] S. Sachin, A. Tripathi, N. Mahajan, S. Aggarwal, and P. Nagrath, "Sentiment analysis using gated recurrent neural networks," *Social Netw. Comput. Sci.*, vol. 1, no. 2, pp. 1–13, Mar. 2020, doi: 10.1007/s42979-020-0076-y.
[87] P. Peng, W. Zhang, Y. Zhang, Y. Xu, H. Wang, and H. Zhang, "Cost sensitive active learning using bidirectional gated recurrent neural networks for imbalanced fault diagnosis," *Neurocomputing*, vol. 407, pp. 232–245, Sep. 2020.
[88] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A search space Odyssey," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 10, pp. 2222–2232, Oct. 2017.
[89] C. Lin, P. Hu, H. Su, S. Li, J. Mei, J. Zhou, and H. Leung, "Sensemood: Depression detection on social media," in *Proc. Int. Conf. Multimedia Retr.*, Jun. 2020, pp. 407–411.
[90] T. H. Zhou, G. L. Hu, and L. Wang, "Psychological disorder identifying method based on emotion perception over social networks," *Int. J. Environ. Res. Public Health*, vol. 16, no. 6, p. 953, Mar. 2019, doi: 10.3390/ijerph16060953.
[91] L. Tong, Z. Liu, Z. Jiang, F. Zhou, L. Chen, J. Lyu, X. Zhang, Q. Zhang, A. S. Senior, Y. Wang, L. Li, and H. Zhou, "Cost-sensitive boosting pruning trees for depression detection on Twitter," 2019, *arXiv:1906.00398*.
[92] L. Tong, Z. Liu, Z. Jiang, F. Zhou, L. Chen, J. Lyu, X. Zhang, Q. Zhang, A. Sadka, Y. Wang, L. Li, and H. Zhou, "Cost-sensitive boosting pruning trees for depression detection on Twitter," *IEEE Trans. Affect. Comput.*, 2022.
[93] U. Naseem, I. Razzak, M. Khushi, P. W. Eklund, and J. Kim, "COVIDSenti: A large-scale benchmark Twitter data set for COVID-19 sentiment analysis," *IEEE Trans. Computat. Social Syst.*, vol. 8, no. 4, pp. 1003–1015, Aug. 2021.
[94] Z. Jamil, "Monitoring tweets for depression to detect at-risk users," Ph.D. dissertation, School Elect. Eng. Comput. Sci., Univ. Ottawa, Ottawa, ON, Canada, 2017.
[95] M. Nadeem, "Identifying depression on Twitter," 2016, *arXiv:1607.07384*.
[96] H.-H. Shuai, C.-Y. Shen, D.-N. Yang, Y. C. Lan, W.-C. Lee, P. S. Yu, and M.-S. Chen, "A comprehensive study on social network mental disorders detection via online social media mining," *IEEE Trans. Knowl. Data Eng.*, vol. 30, no. 7, pp. 1212–1225, Jul. 2018.
[97] J. Samuel, G. G. M. N. Ali, M. M. Rahman, E. Esawi, and Y. Samuel, "COVID-19 public sentiment insights and machine learning for tweets classification," *Information*, vol. 11, no. 6, p. 314, Jun. 2020.
[98] C. W. Park and D. R. Seo, "Sentiment analysis of Twitter corpus related to artificial intelligence assistants," in *Proc. 5th Int. Conf. Ind. Eng. Appl. (ICIEA)*, Apr. 2018, pp. 495–498.
[99] J. Seo, K. Yoo, S. Choi, Y. A. Kim, and S. Han, "The latent learning model to derive semantic relations of words from unstructured text data in social media," *Multimedia Tools Appl.*, vol. 78, no. 20, pp. 28649–28663, Oct. 2019, doi: 10.1007/s11042-018-6211-2.
[100] R. Nareshkumar K. Nimala, "An enhanced BERT model for depression detection on social media posts," in *Proc. Int. Conf. Artif. Intell. Textile Apparel*, Singapore: Springer, Aug. 2023, pp. 53–64.
[101] R. Kanani, J. Fiaidhi, and V. Patel, "Depression sentiment analysis based on social media content like Twitter," *Authorea Preprints*, Dec. 2023.
[102] O. Wu, T. Yang, M. Li, and M. Li, "Two-level LSTM for sentiment analysis with lexicon embedding and polar flipping," *IEEE Trans. Cybern.*, vol. 52, no. 5, pp. 3867–3879, May 2022.
[103] Z. Ali, M. A. Khan, A. Hamza, A. I. Alzahrani, N. Alalwan, M. Shabaz, and F. Khan, "A deep learning-basedx-rayimaging diagnosis system for classification of tuberculosis, COVID-19, and pneumonia traits using evolutionary algorithm," *Int. J. Imag. Syst. Technol.*, vol. 34, no. 1, Jan. 2024, Art. no. e23014.
[104] M. Khan, P. M. Shah, I. A. Khan, S. U. Islam, Z. Ahmad, F. Khan, and Y. Lee, "IoMT-enabled computer-aided diagnosis of pulmonary embolism from computed tomography scans using deep learning," *Sensors*, vol. 23, no. 3, p. 1471, Jan. 2023.
[105] U. Islam, A. A. Al-Atawi, H. S. Alwageed, G. Mehmood, F. Khan, and N. Innab, "Detection of renal cell hydronephrosis in ultrasound kidney images: A study on the efficacy of deep convolutional neural networks," *PeerJ Comput. Sci.*, vol. 10, p. e1797, Jan. 2024.
[106] M. Meer, M. A. Khan, K. Jabeen, A. I. Alzahrani, N. Alalwan, M. Shabaz, and F. Khan, "Deep convolutional neural networks information fusion and improved whale optimization algorithm based smart oral squamous cell carcinoma classification framework using histopathological images," *Expert Syst.*, Jan. 2024, Art. no. e13536.
[107] S. Khan, I. Ullah, F. Khan, Y. Lee, and S. Ullah, "Historical text image enhancement using image scaling and generative adversarial networks," *Sensors*, vol. 23, no. 8, p. 4003, Apr. 2023.
[108] S. Ahmad, F. Mehmood, F. Khan, and T. K. Whangbo, "Architecting intelligent smart serious games for healthcare applications: A technical perspective," *Sensors*, vol. 22, no. 3, p. 810, Jan. 2022.
[109] A. A. Al-Atawi, F. Khan, and C. G. Kim, "Application and challenges of IoT healthcare system in COVID-19," *Sensors*, vol. 22, no. 19, p. 7304, Sep. 2022.
[110] H. U. Khan, M. Abbas, F. Khan, S. Nazir, A. Binbusayyis, A. Alabdultif, and W. Taegkeun, "Multi-criteria decision-making methods for the evaluation of the social Internet of Things for the potential of defining human behaviors," *Comput. Hum. Behav.*, vol. 157, Aug. 2024, Art. no. 108230.
[111] A. W. Khan, S. Zaib, F. Khan, I. Tarimer, J. T. Seo, and J. Shin, "Analyzing and evaluating critical cyber security challenges faced by vendor organizations in software development: SLR based approach," *IEEE Access*, vol. 10, pp. 65044–65054, 2022.
[112] S. U. Khan, A. W. Khan, F. Khan, M. A. Khan, and T. K. Whangbo, "Critical success factors of component-based software outsourcing development from Vendors' perspective: A systematic literature review," *IEEE Access*, vol. 10, pp. 1650–1658, 2022.

SULAIMAN ALMUTAIRI received the bachelor's degree in computer science from Bluefield State University, USA, the master's degree in information technology from RMU, USA, and the Ph.D. degree in biomedical informatics from Rutgers University, Camden, NJ, USA. With a tenure with Qassim University, since 2005, he has held various pivotal roles, including the Vice Dean of Planning, Development, and Quality. He is currently a Distinguished Data Governance Consultant, serving as the Director of Quality Scholarship Services with the Ministry of Education. He also holds the position of an Assistant Professor of health informatics with Qassim University. He has showcased his expertise as a Data Governance Consultant at prominent tech firms in Saudi Arabia, such as Elm and SITE. Additionally, he served as a cultural mission diplomat representing Saudi Arabia in Canada and contributed globally as a Data Consultant with the Development Gateway and UNICEF. His research interests include data governance, bioinformatics, information and policy, and software development.

MOHAMMED ABOHASHRH received the Doctor of Philosophy (Ph.D.) degree in biomedical informatics from Rutgers, The State University of New Jersey, Newark. He is currently an experienced Associate Professor with a demonstrated history of working in the Education Management Industry, King Khalid University, Saudi Arabia. He has published more than 40 research articles in international SCI journals and conferences as well. He is working on various research projects from KKU and Central Research Projects. His research interests include statistics, research, clinical research, data analysis, and healthcare.

HASANAIN HAYDER RAZZAQ received the bachelor's degree in computer science from Al-Rafidain University College, Baghdad, in 2007, and the master's degree in computer science and information systems from Osmania University, India, in 2015. He is currently pursuing the Ph.D. degree in image processing and deep learning with the University Tun Hussein Onn Malaysia (UTHM), Johor, Malaysia. He is an Assistant Professor with the Faculty of Medicine, Jabir Ibn Hayyan University for Medical and Pharmaceutical Sciences. He has successfully published more than ten articles in various international journals and conference proceedings. He is a reviewer for various journals and conferences. He has also served as a technical committee member for numerous international conferences. His research interests include deep neural networks, machine learning, natural language processing, data mining, and image processing.

MUHAMMAD ZULQARNAIN received the bachelor's and master's degrees in computer science from Islamia University of Bahawalpur (IUB), Pakistan, in 2012 and 2014, respectively, and the Ph.D. degree in deep learning from the University Tun Hussein Onn Malaysia (UTHM), Johor, Malaysia, in 2020. He is currently an Assistant Professor with the Faculty of Computer Science and Information Technology, Cholistan University of Veterinary and Animal Sciences (CUVAS), Pakistan. He has an approved Ph.D. and Master's Supervisor with the Higher Education Commission of Pakistan. He has successfully supervised several Ph.D. and master's students and published more than 30 articles in various international journals and conference proceedings. He is a reviewer for various journals and conferences and a co-editor in a Springer conference proceeding. He has also served as a technical committee member for numerous international conferences. His research interests include deep neural networks, machine learning, natural language processing, data mining, and time series prediction.



ABDALLAH NAMOUN (Member, IEEE) received the bachelor's degree in computer science, in 2004, and the Ph.D. degree in informatics from The University of Manchester, U.K., in 2009. He is currently a Full Professor of intelligent interactive systems and the Head of the Information Technology Department, Faculty of Computer and Information Systems, Islamic University of Madinah. He has authored more than 90 publications in research areas spanning intelligent systems, human–computer interaction, machine learning, smart cities, and technology acceptance. He has extensive experience in leading complex research projects (worth more than 23 million Euros) with several distinguished SMEs, such as SAP, BT, and ATOS. His recent research interests include integrating state of the art artificial intelligence models in the development of interactive systems and smart spaces.

FAHEEM KHAN received the Ph.D. degree in computer science from the University of Malakand, Khyber Pakhtunkhwa, Pakistan. He was an Assistant Professor in Pakistan for four years and supervised many articles and students. Since April 2021, he has been an Assistant Professor with the Department of Computer Engineering, Gachon University, South Korea. His research interests include computer networking, wireless networking, MANET, sensor networking, the IoT, artificial intelligence, and AI in healthcare systems.