Indonesian Journal of Electrical Engineering and Computer Science
Vol. 40, No. 1, October 2025, pp. 388~396
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v40.i1.pp388-396

Interrogative insights into depression detection via social networks and machine learning techniques

Chaithra Indavara Venkateshagowda[1], Roopashree Hejjajji Ranganathasharma[2], Yogeesh Ambalagere Chandrashekaraiah[3]
[1]Department of Computer Science and Engineering, GSSS Institute of Engineering and Technology for Women, Visvesvaraya Technological University, Belagavi, India
[2]Department of Artificial Intelligence and Data Science, GSSS Institute of Engineering and Technology for Women, Visvesvaraya Technological University, Belagavi, India
[3]Department Computer Science and Engineering, Government Engineering College, Visvesvaraya Technological University, Belagavi, India

Article Info

Article history:

Received Jan 4, 2025
Revised Jun 13, 2025
Accepted Jul 3, 2025

Keywords:

Deep learning
Depression
Machine learning
Mental health
Social network

ABSTRACT

As users on social networks (SNs) interact with one another by exchanging information, giving feedback, finding new content, and participating in discussions; thus, generating large volumes of data each day. This data includes images, texts, videos and can be used to help the user find out how they have been doing, when they were depressed, how not to be depressed, and other similar insights. Depression is one of the most common chronic illnesses and it has emerged as a global mental health problem. But the lack of these data is incomplete, sparse and sometimes inaccurate, and so the task of diagnosing depression using automated systems is still proving a challenge. Various techniques have been used to detect depression through the years however, machine learning (ML) and deep learning (DL) techniques offer better ways. In the context of that, this study reviews state-of-the-art ML and DL approaches for the detection of depression using systematic literature review (SLR) method as well as highlight fundamental challenges in literature, which future works can focus on. We hope that this survey will provide a better understanding of these strategies for the readers and researchers in the ML and DL fields, when it comes to diagnosis of depression.

Corresponding Author:

Chaithra Indavara Venkateshagowda
Department of Computer Science and Engineering,
GSSS Institute of Engineering and Technology for Women, Visvesvaraya Technological University
Belagavi, India, KRS Rd, Metagalli, Mysuru, Karnataka - 570016
Email: chaithra.iv@gmail.com

1. INTRODUCTION

Mental issues are the most common global phenomenon while one out of five people suffers from a mental health condition that can disrupt multiple areas of their lives [1]. Depression is a chronic disease or condition found with many of the people. It is one of the most unrecognized and least treated types of mental health disorders. This has turned out to be a major issue in the economic, social and specifically health sector. It is among the top killers and is becoming one of the most common diseases with no effective treatment. It is frequently observed in the presence of other underlying mental health conditions, and can trigger dormant mental disorders. Several studies have reported that depression is one of the top cause of death, many of them due to suicide. Depression is irrational, so hard to understand. The idea that the words of one person may push another to self-destruction is a sensitive topic that many choose to ignore. By applying machine learning

(ML) and deep learning (DL), the detection process can be enhanced, hence being able to reduce such unfortunate consequences. With the ubiquitous nature of social media social network (SN) platform such as Facebook, Twitter, or Reddit have encouraged people to share information about their daily life, personal experiences, opinions, and challenges. These platforms provide new windows of self-expression and social interaction. Status updates, tweets and posts and comments, all of these laid bare the thoughts and feelings, the lives that the users lived. More importantly, these digital traces can reflect the mental conditions and the personal lives of users, although some limited. Analyzing the techniques of ML and DL can enhance the ability to identify certain characteristics that assist SN users in determining mental states [2]. SN platforms collect a vast amount of data from users, while by implementing ML and DL algorithms, specific features can be identified in data for the purposes of identifying mental states for SN users. Mental health challenges, like depression, can present as hopelessness, guilt, irritability, self-deprecation, loneliness, isolation, and thoughts of suicide. While these warning indicators are not independently definitive, posts exhibited on social media that incorporate these behaviors can act as precautionary signs of depressive disorders.

Various related work has been referred to understand the significance and scope of study. New method to detect the difference between people with depression have been evolved and without using their facial expression patterns and ML classification algorithms. As an example, the random forest (RF) classification method has slightly less accuracy (80.1%) for nonverbal visual compliance analysis of automated depression diagnosis [3]. Future improvements on the applied ML models may offer more reliable means of determining the importance of the features and degree of depression. A multitasking learning model incorporating deep neural network (DNN) classification and FusionNet has been proposed to address word vector classification and statistical categorization [4]. In one example, XLNet is utilized to extract textual features, enabling depression diagnosis through hierarchical attention networks combined with clinical interviews. These methods integrate affective information and external features into the attention mechanism, boosting model performance. Ksibi *et al.* [5] employed two DL architectures—1-dimensional convolutional neural network (1DCNN) and long short-term memory (LSTM)—to automate the diagnosis of unipolar depression using electroencephalogram (EEG) data. Existing system has also reportedly used a CNN-based and recurrent neural network (RNN)-based model; the CNN-based model showed higher performance than the RNN-based model [6]. A new method based on inverted vocal tract parameters was also suggested in [7] for detecting depression. The classification employed generalized CNN dilated models. The study created variety features by pooling speech data across two separate datasets of depression, allowing for improved generalization of CNN models for depression detection. The model proposed showed about 10% more accuracy in comparison to the model trained on one dataset. Dialog interaction is a challenge for this model though. An additional challenge for DL approaches regarding large scale social data is to perform effectively without losing vital information required for classifying accurately.

The identified problems from the review are mainly related to implementation methodology being carried out by existing ML/DL approaches for diagnosing depression. The significant challenges are: i) due to the complexity of big data, existing models often have difficulty extracting significant features from unstructured datasets, ii) the presence of unstructured or unlabeled data can cause a degradation in the accuracy of the model or in the worst case, the loss of data. Processing such data well is still a big problem. iii) The irrelevant or unexpected data can make the things tough on the part of data labeling and data classification which can cause hindrance to model performance. Apart from these, there are also issues pertaining to quality data and biased focus on the local research problems where global optima is challenging to accomplish or even investigated. The goal of this paper is to present a comprehensive review of the literature concerning depression detection methods based on ML and DL algorithms. The value-added contribution are: it provides a broad overview of depression detection, discussing the range of opportunities and challenges, as well as possible research avenues. Further on, we expound on need for ML and DL techniques in depression detection, elucidating the caveats in existing works and areas requiring investigation as per the fundamentals.

2. METHOD

The main goal of the proposed study is to perform a meta-analysis of the efficiency of the ML and DL algorithms to explore the level of effectiveness in existing depression detection approaches. To this end, this review paper employed a desk research methodology as illustrated in Figure 1. Based on adopted research methodology google scholar along with reputed technical journal publications were undertaken as preliminary search. Once the method is adopted, the next step is an initial filtering based on inclusion and exclusion criteria. The main inclusion criteria of filtering consist in that the published article should predominantly contain methodological details with results obtained for having comprehending through the presented approach. The same time span for literature was considered for the second inclusion criteria as well, leading to the consideration of the published research articles containing technical implementation between the years 2019–2024. The third criterion of eligibility states that research articles are only related to either ML or DL as the primary implementation. The inclusion criteria are theoretical papers with results or an implementation model, and studies published from 2019 onward. Further, any paper which has focused on a non-learning-based approach towards improving internet of things (IoT) security is also not considered as the aim for this paper is to reflect upon the potential of learning-based schemes. The next phase in the operation is to do a removal of the duplicates. Any two research articles which were written based on same study model will be labelled as duplicated; this can also be split into i) two multiple extended study model from same author convey duplicate; ii) usage of exactly similar concept in 2 different research articles also will be carried as duplicated as the idea is to obtain complete unique study model. The initial search returned about 1,257 results and the first screening yielded 183 papers. Finally, 32 number of research articles have been considered after removing the duplicates and reviewed in this manuscript. In addition to that, this learning outcome of the proposed study provides more toward new research gap, research trend, and unique findings of previous studies.

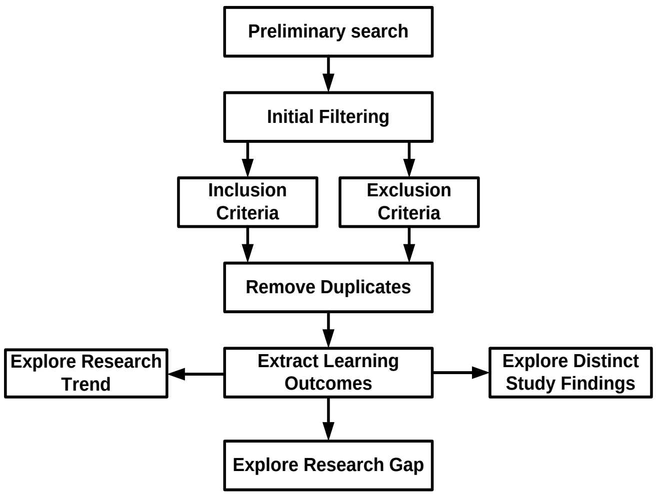

Figure 1. Method adopted in proposed study

3. RESULTS

This section presents compact discussion of existing learning-based approaches towards diagnosing depression. The overview of ML and DL approaches to depression detection is illustrated in Figure 2. This article examines the current state of various ML algorithms using the systematic literature review (SLR) methodology, specifically for detecting depression.

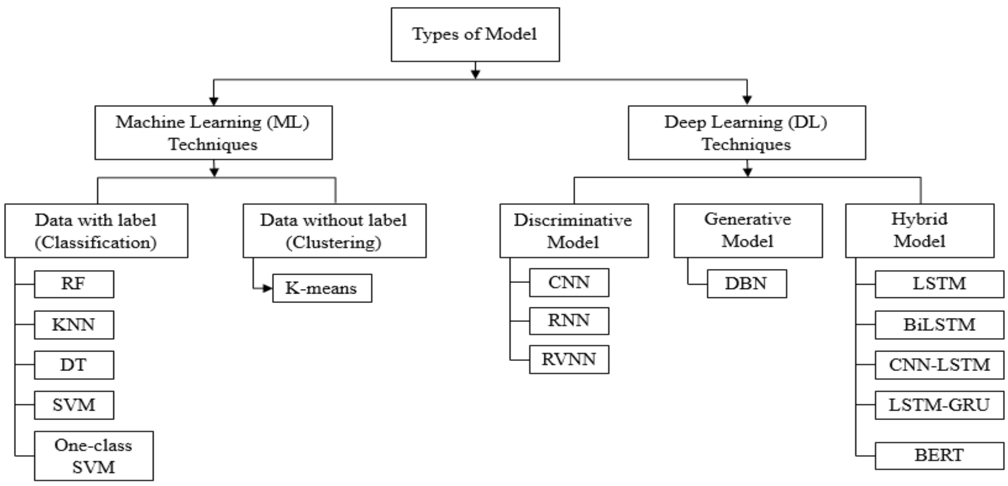

Figure 2. Taxonomy of ML and DL approaches for diagnosing depression

3.1. Machine learning for depression detection

Ahmed [8] examined how online communities could aid in spotting and gauging serious depression, by tracking social behavior, mood, language use, self-perception and the need for antidepressants, all inferred from users' online writing. Similarly, Castillo-Sánchez *et al.* [9] studied the use of Twitter for monitoring psychological states such as depression and suicide risk. Their findings indicated that concerns expressed in violent or disturbing texts could be identified using a combination of human analysis and automated classifiers. Multiple studies have shown how using user-generated content (UGC) strategically can help psychologically. Bokolo and Liu [10] conducted examinations of mental health issues and behaviors on SNs to assess whether individuals were experiencing depression. Angskun *et al.* [11] discussed more psycholinguistic feature types with greater potential to improve the classification as well as further decrease error rates. RF were one of the ML techniques which performed most accurately, specifically to identify depression.

Salma and Bhuiyan [12] analyzed users' Facebook behaviors as reflecting five-factor model of personality in efforts to describe individual identity formation process. They investigated correlations between user IDs and profile information—related to a user's screen name; the number of photos shared or received; frequency of attendance at events; and other similar activities. Titla-Tlatelpa *et al.* [13] proposed a new method to analyze sentiment on social media, arguing that users' messages provide insight into their emotional evaluations. This approach also revealed users' typical emotional responses and flagged extreme emotional changes. Alenezi *et al.* [14] also provided the first evidence regarding the major impact of depression on the women's health burden, specifically vein Facebook data. Ramamoorthy *et al.* [15] show that both word frequencies and topic modeling are informative features for prediction models. Using the bag of words approach, Salazar *et al.* [16] was able to improve depression diagnosis by evaluating tweet text and assigning categories through multiple classification. In addition, using content support vector machine (SVM) and training of a grading system. Nusrat *et al.* [17] enhanced results for sad tweets and percentage of users who are depressed among multiple users obtained significant betterment in recall and precision score.

Additional studies were conducted by Safa *et al.* [18] on Twitter and Weibo, utilizing sentiment analysis strategies that integrated human-made rules and vocabulary to quantify depression-related tweets. This approach offered a more comprehensive feature analysis compared to previous work. Their findings highlighted the significant role that text-based characteristics play in diagnosing online depression. Sharma and Verbeke [19] introduced a model using the extreme gradient boosting (XGBoost) ML algorithm to identify and predict depression biomarkers across diverse samples. The ML model, which dealt with an imbalanced dataset, tended to predict the absence of depression despite its presence. To address this issue, the authors employed several resampling techniques. The XGBoost algorithm was applied to each sample to classify mental health cases in healthy instances.

People with mental illness look online for help, in specialty communities or in general, such as Twitter and Reddit. Benjachairat *et al.* [20] considered the use of SNS data for risk assessment of selfharm. This approach was mostly focused on major depression. They examined behavioral patterns derived from what users wrote. In a successful public health initiative, Twitter has been used to create statistical models to determine the effect of childbirth on new mothers' behavior and mood. Microblogs and other means of social media platforms can also be used to detect if a person is at high risk for suicide, so that preventative measures can be taken before the person goes through with their intentions. Cao *et al.* [21] showed that complex intervention framework could be modified to protect those at risk of suicide via online microblog. While their study found suicidal ideation to be predictable, the prediction accuracy needs further improvement. Furthermore, suicidal ideation is a relatively rare event, making it a novel and pragmatic effort to develop models specifically trained to identify high-risk ideation.

3.2. Deep learning for depression detection

A large body of prior work has primarily centered around generating different methods for implementation, which are frequently based on using apps to identify depression. Moreover, in detecting mental health problems through social media, until recently, the use of ML methods, more specifically deep neural networks, was constrained. That limitation was mainly because training set annotation is notoriously tricky to achieve in large quantities. For example, Nijhawan *et al.* [22] in notable examples through designs that were effectively well-acclaimed in processing natural language, as well as skimming the most efficient DNN architectures for this purpose.

A graphical attention model was proposed combining multimodal depression detection knowledge for text, audio, and images [23]. Although they used temporal CNN models, they incorporated smaller questionnaire data into their experiments as part of a larger architecture. While the authors had not tested their model on short and noisy social media data, the scalability of their approach to processing such large datasets is still unexplored. Chen *et al.* [24] train a focus model for detecting depression from clinical and online interviews. The key finding was that those diagnosed with depression will use more emotional language than those who are not. Bucur *et al.* [25] in a recent study examined sadness among users during COVID-19 for using natural language processing (NLP) based approaches. Yin *et al.* [26] have used a RNN based multimodal model to predict the depression was trained on datasets of online users forum. Ahmed *et al.* [27] concentrated on DL approaches with multiple word insertions to improve training for early sorrow identification in SNs. They trained a new word to use the task-specific content for this purpose instead. Although employing a neural network model to derive high-quality features, the low efficiency of numeric data was largely due to missing attention to temporal dynamics and latent topics.

Many people with depression and emotional distress, however, are not correctly diagnosed and treated and do not seek help. It is important, therefore, that we build a system that can reliably and proactively assess these people. Following the design science principles, Singh *et al.* [28] presented the LSTM, a new design using DL relevant to alleviate depression and emotional discomfort. In parallel, the model combines general knowledge and domain-specific methods through word embedding and also incorporates several parallel LSTM units, a type of DL network. A tool that has been well researched and developed so far for detecting depression is voice recognition. Since there are changes in neuromotor coordination associated with psychomotor slowing, a core feature of major depressive disorder, articulatory coordination features (ACFs) are described. For example, Seneviratne and Espy-Wilson [29] suggested creating a general classifier for detecting depression by using a dilated convolutional network trained on ACFs from two datasets for depression. ACFs, derived from vocal tract parameters, proved to be effective features for depression detection whose notable efficacy was demonstrated in their work.

3.3. Internet of medical things for depression detection

It is important to identify with yourself early symptoms of stress and take steps to mitigate them in this world. It works with changes in physical conditions in the human body under stress, such as the decrease in the palm temperature, the increase in the physical activity, and the increase of sweat rates [30]. Cortisol, the main stress hormone, spikes under duress. With the right sensors, these physiological discrepancies could be monitored to see if an individual is actually stressed or not. Everyone reacts to stress differently, and it often disrupts their regular physiologic functions. An internet of medical things (IoMT)-based early-stage stress detection system can detect stress by analyzing cortisol levels, body temperature, movement rate, and sweat rate [31]. One example of such devices can be found in [32], which implements a Mamdani-type fuzzy logic controller, that is, a neural network-based approach forming more than 200 rules. This data is then sent to the cloud for monitoring stress levels in real-time. This enables regular evaluations and necessary measures to mitigate any health threats.

The early-stage stress detection model categories stress level in three ways i.e., low, normal and high. It is classified based on the perceptible alteration in the physical condition of the body. One practical example is tracking a student’s mental health by their parent or guardian. Running in real-time and with low power consumption, the system boasts a number of advantages such as high precision, low risk of systematic errors, and cost-effectiveness.

3.4. Frequently adopted dataset

In this section, we review studies involving ML and DL methods for depression detection and the datasets they utilized. There are several datasets used for depression detection, some of which are listed in Table 1 and the most well-known DAIC-WoZ dataset. From the 2020s we find two papers using the DAIC-WoZ dataset, and from 2018 one paper. This contains clinical interviews of 189 subjects in the form of audio/video and transcripts. The data is split into training, testing and development set containing 107, 35, and 47 subjects respectively. The CLPsych 2015 dataset consists of filtered public Twitter statuses from between 2008 and 2013 collected via the Twitter API along with up to 3000 most recent tweets for the users. The dataset includes 477 users with depression, 396 users with PTSD, 873 control users, and those users who have anxiety caused by frightening, stressful, or upsetting events. The GPC data set is a large clinical data set containing [over 1,300 tran-] scribed psychotherapy notes. Unlike the above datasets, the reddit automated depression diagnosis (RSDD) dataset is based on posts written by two groups of Reddit users (a depression group of 9,210 users and a control group of 107,274 users). We demonstrate that the DAIC-WoZ and CLPsych 2015 datasets are popular for depression detection in our research. The Avec2013, Avec2014, and Avec2016 datasets have also been extensively used across the aforementioned sub-challenges for depression detection.

Moreover, an open twitter dataset primarily concentrating on depressed users has gathered data from 500 different users, more than 1 million tweets in total, 334 of users whose condition is classified as sad. A human annotator reads through these tweets and verifies that they are in fact the individuals' own experiences of depression, not those of their friends or family. Randomly selected, human-analyzed non-depressed individuals are guaranteed never to have posted a tweet with the term “depress”. We then remove users with less than five tweets to remove noise and enhance data validity.

Table 1. Dataset adopted in existing implementations

Dataset	Problem tackle	Content type
DAIC-Woc	Depression detection	Audio, video, text
AVEC-2013	Depression detection	Video
WU3D	Depression detection	Picture, text
AVEC-2016	Depression detection	Audio, video
eRisk 2017	Depression detection	Text
Stressors associated with depression (SAD)	Depression detection	Text
Reddit self-reported depression diagnosis	Depression detection	Video
General psychotherapy corpus	Depression detection	Text
Open sourcing mental illness	Screening depression	Text
CLPsych 2015	Depression detection	Text
AVEC-2013 and AVEC-2014	Depression detection	Video
MR, Subj, CR, MPQA, SSST-1, SST-2, TREC	Sentence semantic classification	Text
eRisk	Early risk detection of depression and anorexia	Text
SST, CVAT	Dimensional sentiment analysis	Text
DAIC-WOZ, RAVDESS, AVi-D	Clinical depression recognition	Audio
DASS-21	Depression, anxiety, and stress prediction	Text
SAD	Depressive symptoms classification	Text
DAIC-WOZ	Depression classification	Audio
EEG (HydroCel Geodesic Sensor Net)	Mild depression recognition	EEG
AVEC-2017, DAIC-WOZ	Depression scale recognition	Audio, video, text

3.5. Research gap with tentative solution

The identified research gap are as follows:

- Data availability and quality: generally, this could be due to the both limited and imbalanced datasets. Data scarcity, particularly for a minority of populations and outgroup (e.g., different age groups and cultures) is a common problem. The majority of available data are cross-sectional; they reveal only a moment in time of this complex condition. Longitudinal data — data that follows people for long periods — might offer more detailed information on how depression begins and changes, which in turn could lead to improved models of how to detect it early. Most of the publicly available datasets are based on self-reported questionnaires that may take subjective forms and introduce biases to data collection. The main challenge is to obtain proper labeling and adding multi-modal data (e.g., behavioral data, voice analysis, or physiological signals).
- Multimodal data integration: depression does not only impact individuals at a psychological level; it may also be observed at behavioral level (e.g. different speech patterns, types of facial expressions) and also a physiological level (e.g. HR and sleep patterns). Research gap: the effective integration of these data types (e.g., text, voice, and biometric data) into ML models is a significant research gap that can help improve detection accuracy. Integrating multiple data sources (e.g., electronic health records, social media activity, and wearable sensors) necessitates the use of powerful data fusion techniques. Methods that can accommodate the complexity and heterogeneity of such data sources need to be developed through research.
- Sensitivity to early detection and predictive accuracy: current ML models are largely focused on detecting depression post-diagnosis, while pre-symptomatic (pre-diagnosis) detection remains a major gap. Models capable of predicting the onset of depression, which related to identifying subtle early symptoms and behavioral changes. The symptoms of depression, particularly in its early stages, can be quite mild and easily mistaken for other illnesses. Work is still needed to develop ML models that more accurately distinguish between depression and other mental health disorders and from normal variations in mood and behaviors.

The tentative solution to address the above-identified research gap are as follows:

- The issues pertaining to data availability and quality can be addressed by harnessing the potential of NLP. This adoption can significantly address constrained for model transparency as well as restricted annotated data.
- With proven accuracy of DL models, the performance towards early detection can be further improvised by investigating and adopting transformer model with attention mechanism. A pre-trained language model (PTLM) can be deployed to acquire contextual information for precise and granular interpretation of language patterns with complexities.
- Developing a novel form of framework with PTLM will promote understanding of dynamic discussion on social post with all evolving phrases and term with better probability of subtle depression cues.
- Finally, an extended modelling can be developed for forecasting any possibility of re-occurrence of depression in future by amending PTLM. For better accuracy, preprocessing and labelling can be carried out to offer more better classification performance.

This paper provides a thorough analysis of DL and ML methods for detecting sadness in SN data. The main conclusions show that by examining user-generated information, including text, photos, and behavioral data on websites like Twitter, Facebook, and Reddit, both ML and DL techniques have shown promise in detecting depression. Notably, DL models like transformer models and RNNs as well as methods like RF and SVM have demonstrated efficacy in raising classification accuracy. One possible method for improving the diagnosis of depression is the integration of multimodal data, which combines physiological, audio, and textual information. But there are still issues that impede the field's advancement, mainly with regard to data quality, unavailability, and the difficulty of acquiring labeled datasets.

This study expands on earlier investigations into the use of online behavioral data and social media for the identification of depression. The usefulness of UGT for tracking psychological states like depression and suicide risk has been shown in previous studies, including those by Ahmed [8] and Castillo-Sánchez *et al.* [9]. Our results support prior investigations while offering fresh viewpoints, especially with the use of multimodal data and DL methods to improve prediction accuracy. Our study highlights the possibility of using physiological data and voice patterns in depression diagnosis, which can offer a more comprehensive approach to understanding mental health through online interactions, in contrast to previous efforts that mostly concentrated on text-based analysis.

4. CONCLUSION

In conclusion, there is a lot of potential for using ML and DL to detect depression; however, obstacles including data quality, integration, and early detection need to be addressed. We can get closer to constructing proactive and efficient mental health diagnostic tools by increasing multimodal integration, developing more precise early detection models, and improving data collection techniques. Future studies in this area could completely change how we identify and comprehend depression, which will eventually improve mental health outcomes for people all across the world. Depression is an ambiguous disease hence SN data can not provide direct data to diagnose these kinds of diseases. Researchers in the previous few decades have proposed several approaches to detect depression using conventional ML and DL approaches. However, such work has yet to be systematically collated to provide a clearer picture. The main aim of this article is to provide a review on DL and ML approaches for depression diagnosis and to highlight the research efforts made in this area alongside the challenges in existing research. However, reviewing all previous studies is out of the scope of this survey. Recommendations: the following considerations will guide future investigations: i) unused, irrelevant feature elimination can increase system performance during training, ii) more efficient selection of the categorical techniques will help in future categorizations, iii) for output assessment in single or hybrid classifiers, ensemble becomes a more feasible option, iv) it is essential to have a validated feature selection, v) however, without additional data, the dataset size is much too small, vi) meanwhile, to tackle large scale data, DL should embrace unsupervised and semi-supervised learning. To conclude, novel techniques where size and diversity of data are considered to be the essences of concern can be presented as potential business optimization for DL methods and traditional ML applications to tackle class imbalance. Future studies should cross examine results with different metrics of success and share statistical results. Future exploration on modern hybrid-cluster approaches to ML and DL can be beneficial in large-scale psychosocial data on depression.

A number of important areas are yet unexplored and need more research. First, it's critical to overcome the constraints on the quality and availability of datasets. In order to reflect the dynamic nature of depression, future research might concentrate on developing more varied and longitudinal datasets. Furthermore, integrating multimodal data—text, audio, visual, and physiological signals—offers intriguing possibilities. However, more research is required to create reliable data fusion methods that can manage such a wide range of data kinds. Moreover, while current models focus on post-diagnosis detection, future studies should prioritize early detection methods to identify subtle, pre-symptomatic signs of depression. Lastly, applying unsupervised and semi-supervised learning techniques to large-scale data could help address issues related to labeled data scarcity and improve model generalization.

ACKNOWLEDGEMENTS

We wish to confirm that no known conflicts of interest are associated with this publication and all the authors have contributed equally.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Chaithra Indavara Venkateshagowda	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	
Roopashree Hejjajji Ranganathasharma		✓		✓		✓		✓	✓	✓	✓	✓		
Yogeesh Ambalagere Chandrashekaraiah	✓	✓				✓		✓	✓	✓	✓	✓		

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are available on request from the corresponding author.

REFERENCES

[1] M. Swathi, S. Manjusha, I. J. Vadakkiniath, and A. Gururaj, "Prevalence and correlates of stress, anxiety, and depression in patients with chronic diseases: a cross-sectional study," *Middle East Current Psychiatry*, vol. 30, no. 1, Aug. 2023, doi: 10.1186/s43045-023-00340-2.
[2] M. Squires *et al.*, "Deep learning and machine learning in psychiatry: a survey of current progress in depression detection, diagnosis and treatment," *Brain Informatics*, vol. 10, no. 1, Apr. 2023, doi: 10.1186/s40708-023-00188-6.
[3] M. Bader, M. Abdelwanis, M. Maalouf, and H. F. Jelinek, "Detecting depression severity using weighted random forest and oxidative stress biomarkers," *Scientific Reports*, vol. 14, no. 1, Jul. 2024, doi: 10.1038/s41598-024-67251-y.
[4] T. M. Quan, D. G. C. Hildebrand, and W. K. Jeong, "FusionNet: a deep fully residual convolutional neural network for image segmentation in connectomics," *Frontiers in Computer Science*, vol. 3, May 2021, doi: 10.3389/fcomp.2021.613981.
[5] A. Ksibi, M. Zakariah, L. J. Menzli, O. Saidani, L. Almuqren, and R. A. M. Hanafieh, "Electroencephalography-based depression detection using multiple machine learning techniques," *Diagnostics*, vol. 13, no. 10, 2023, doi: 10.3390/diagnostics13101779.
[6] Vandana, N. Marriwala, and D. Chaudhary, "A hybrid model for depression detection using deep learning," *Measurement: Sensors*, vol. 25, p. 100587, Feb. 2023, doi: 10.1016/j.measen.2022.100587.
[7] S. A. Almaghrabi, S. R. Clark, and M. Baumert, "Bio-acoustic features of depression: a review," *Biomedical Signal Processing and Control*, vol. 85, p. 105020, Aug. 2023, doi: 10.1016/j.bspc.2023.105020.
[8] J. O. Ahmed, "Social media psychology and mental health," *Middle East Current Psychiatry*, vol. 30, no. 1, Nov. 2023, doi: 10.1186/s43045-023-00362-w.
[9] G. Castillo-Sánchez, G. Marques, E. Dorronzoro, O. Rivera-Romero, M. Franco-Martín, and I. De la Torre-Díez, "Suicide risk assessment using machine learning and social networks: a scoping review," *Journal of Medical Systems*, vol. 44, no. 12, Nov. 2020, doi: 10.1007/s10916-020-01669-5.
[10] B. G. Bokolo and Q. Liu, "Deep learning-based depression detection from social media: comparative evaluation of ML and transformer techniques," *Electronics (Switzerland)*, vol. 12, no. 21, p. 4396, Oct. 2023, doi: 10.3390/electronics12214396.
[11] J. Angskun, S. Tipprasert, and T. Angskun, "Big data analytics on social networks for real-time depression detection," *Journal of Big Data*, vol. 9, no. 1, May 2022, doi: 10.1186/s40537-022-00622-2.
[12] N. Salma and F. Bhuiyan, "Facebook addiction and its impact on depression: a cross sectional study," *Journal of Health, Population and Nutrition*, vol. 43, no. 1, Jun. 2024, doi: 10.1186/s41043-024-00556-w.
[13] J. de J. Titla-Tlatelpa, R. M. Ortega-Mendoza, M. Montes-y-Gómez, and L. Villaseñor-Pineda, "A profile-based sentiment-aware approach for depression detection in social media," *EPJ Data Science*, vol. 10, no. 1, 2021, doi: 10.1140/epjds/s13688-021-00309-3.
[14] A. Alenezi, W. Hamed, I. Elhehe, and R. El-Etreby, "Association between Facebook addiction, depression, and emotional regulation among women," *Healthcare (Switzerland)*, vol. 11, no. 12, p. 1701, Jun. 2023, doi: 10.3390/healthcare11121701.
[15] T. Ramamoorthy, V. Kulothungan, and B. Mappillairaju, "Topic modeling and social network analysis approach to explore diabetes discourse on Twitter in India," *Frontiers in Artificial Intelligence*, vol. 7, Feb. 2024, doi: 10.3389/frai.2024.1329185.
[16] L. M. G. Salazar, O. Loyola-González, and M. A. Medina-Pérez, "An explainable approach based on emotion and sentiment features for detecting people with mental disorders on social networks," *Applied Sciences (Switzerland)*, vol. 11, no. 22, p. 10932, Nov. 2021, doi: 10.3390/app112210932.
[17] M. O. Nusrat, W. Shahzad, and S. A. Jamal, "Multi class depression detection through tweets using artificial intelligence," *arXiv [cs.CL]*, 2024. http://arxiv.org/abs/2404.13104.
[18] R. Safa, P. Bayat, and L. Moghtader, "Automatic detection of depression symptoms in twitter using multimodal analysis," *Journal of Supercomputing*, vol. 78, no. 4, pp. 4709–4744, Sep. 2022, doi: 10.1007/s11227-021-04040-8.
[19] A. Sharma and W. J. M. I. Verbeke, "Improving diagnosis of depression with XGBoost machine learning model and a large biomarkers dutch dataset (n = 11,081)," *Frontiers in Big Data*, vol. 3, Apr. 2020, doi: 10.3389/fdata.2020.00015.

[20] P. Benjachairat, T. Senivongse, N. Taephant, J. Puvapaisankit, C. Maturosjamnan, and T. Kultananawat, "Classification of suicidal ideation severity from Twitter messages using machine learning," *International Journal of Information Management Data Insights*, vol. 4, no. 2, p. 100280, Nov. 2024, doi: 10.1016/j.jjimei.2024.100280.

[21] L. Cao *et al.*, "Latent suicide risk detection on microblog via suicide-oriented word embeddings and layered attention," in *EMNLP-IJCNLP 2019 - 2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing, Proceedings of the Conference*, 2019, pp. 1718–1728, doi: 10.18653/v1/d19-1181.

[22] T. Nijhawan, G. Attigeri, and T. Ananthakrishna, "Stress detection using natural language processing and machine learning over social interactions," *Journal of Big Data*, vol. 9, no. 1, Mar. 2022, doi: 10.1186/s40537-022-00575-6.

[23] Z. Li, Z. An, W. Cheng, J. Zhou, F. Zheng, and B. Hu, "MHA: a multimodal hierarchical attention model for depression detection in social media," *Health Information Science and Systems*, vol. 11, no. 1, Jan. 2023, doi: 10.1007/s13755-022-00197-5.

[24] Z. Chen, J. Deng, J. Zhou, J. Wu, T. Qian, and M. Huang, "Depression detection in clinical interviews with LLM-empowered structural element graph," in *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, vol. 1, pp. 8174–8187, 2024, doi: 10.18653/v1/2024.naacl-long.452.

[25] A. M. Bucur, A. C. Moldovan, K. Parvatikar, M. Zampieri, A. R. KhudaBukhsh, and L. P. Dinu, "On the state of NLP approaches to modeling depression in social media: A post-COVID-19 outlook," *IEEE Journal of Biomedical and Health Informatics*, vol. 29, no. 6, pp. 4439–4451, Jun. 2025, doi: 10.1109/JBHI.2025.3540507.

[26] S. Yin, C. Liang, H. Ding, and S. Wang, "A multi-modal hierarchical recurrent neural network for depression detection," in *AVEC 2019 - Proceedings of the 9th International Audio/Visual Emotion Challenge and Workshop, co-located with MM 2019*, Oct. 2019, pp. 65–71, doi: 10.1145/3347320.3357696.

[27] S. Ahmed, S. Rakin, M. W. I. Waliur, N. B. Islam, B. Hossain, and M. M. Akbar, "Depression detection from social media Bangla text using recurrent neural networks," *arXiv [cs.HC]*, 2024. http://arxiv.org/abs/2412.05861.

[28] J. Singh, M. Wazid, D. P. Singh, and S. Pundir, "An embedded LSTM based scheme for depression detection and analysis," *Procedia Computer Science*, vol. 215, pp. 166–175, 2022, doi: 10.1016/j.procs.2022.12.019.

[29] N. Seneviratne and C. Espy-Wilson, "Generalized dilated CNN models for depression detection using inverted vocal tract variables," in *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, Aug. 2021, vol. 1, pp. 601–605, doi: 10.21437/Interspeech.2021-1960.

[30] C. Huang, J. Wang, S. Wang, and Y. Zhang, "Internet of medical things: a systematic review," *Neurocomputing*, vol. 557, p. 126719, Nov. 2023, doi: 10.1016/j.neucom.2023.126719.

[31] A. A. Al-Atawi *et al.*, "Stress monitoring using machine learning, IoT and wearable sensors," *Sensors (Basel, Switzerland)*, vol. 23, no. 21, p. 8875, Oct. 2023, doi: 10.3390/s23218875.

[32] I. Jeya Daisy, B. V. Kumar, and M. Krishnamoorthy, "Early-stage depression detector using IoMT," in *2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation, ICAECA 2021*, Oct. 2021, pp. 1–5, doi: 10.1109/ICAECA52838.2021.9675629.

BIOGRAPHIES OF AUTHORS

Chaithra Indavara Venkateshagowda received the Master's degree in computer science and engineering from Visvesvaraya Technological University, Belagavi, Karnataka, India in 2013 and is currently working towards Ph.D. degree. She joined the Computer Science and Engineering Department, Adichunchanagiri Institute of Technology, Chikkamagaluru as an Assistant professor, in 2015. Her research interests include machine learning, deep learning, artificial intelligence, sentiment analysis, and large language models. She can be contacted at email: chaithra.iv@gmail.com.

Roopashree Hejjajji Ranganathasharma completed B.E. (E&C) and M.Tech. (CS&E) from VTU, Belagavi, Karnataka, India and Ph.D. from CHRIST (Deemed to be University) Bengaluru, Karnataka, India. She has around 13 years of Industrial experience and 5 years of teaching experience. She is presently working as a Professor and Head of, the Department of Artificial Intelligence and Data Science at GSSSIETW, Mysuru. She can be contacted at email: roopashreehr@gsss.edu.in.

Yogeesh Ambalagere Chandrashekaraiah has completed B.E, M.Tech, and Ph.D. from Visvesvaraya Technological University Belagavi, Karnataka, India. Currently working as an Assistant Professor in CS&E, Government Engineering College, Chamarajanagar, Karnataka, India. His area of interest are wireless sensor network, IoT, and machine learning. He can be contacted at email: yogeesh13@gmail.com.