A review and meta-analysis of machine intelligence approaches for mental health issues and depression detection

Ravita Chahar, Ashutosh Kumar Dubey[*] and Sushil Kumar Narang
Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India

Abstract
Cases of mental health issues are increasing continuously and have sped up due to COVID-19. There are high chances of developing mental health issues such as depression, anxiety, schizophrenia, and dementia after 2–3 months of COVID-19 diagnosis. In this paper, a review and meta-analysis of machine intelligence approaches—namely, machine learning, deep learning (deep learning with hybrid boosting), and machine vision methods—for mental health issues and depression detection were presented. Meta-analysis was performed in four parts. The first part focused on the publication trends, criteria for inclusion and exclusion, and the current methodological scenario. The second part was intended for the methods and their advantages and limitations. It covered mental health issues and depression detection techniques along with the challenges. The third part focused on the discussion and applicability of datasets. The fourth part focused on the complete analysis and discussion along with suggestive measures; moreover, it covered the overall analysis, including the methodological impact, result impact, current trends, and some suggestions based on the limitations and challenges.

Keywords
Machine intelligence, Machine learning, Deep learning, Mental health issues, Depression detection.

1.Introduction
Mental health issues such as depression, anxiety, and stress are common nowadays. The prevalence of mental health issues has increased considerably due to the COVID-19 pandemic, as it caused psychological distress [1–4]. According to the World Health Organization (WHO), depression is a common mental disorder [5]. As per the WHO report in 2017, over 300 million people (equivalent to 4.4% of the world's population) worldwide are estimated to suffer from depression [5]. An increase of 11% from the previous year was noted in the US alone due to the COVID-19 pandemic [6, 7]. The best way to detect depression is to analyze the behavior of the person. Depression can be caused by many factors, such as work pressure, social responsibilities, strained relationships, sleeping disorder, tiredness, lack of concentration, complex interactions between social and competitive markets, and biological factors [5, 8–10]. A WHO report published a decade ago stated that the number of deaths from suicide was 1 million per year [5].

This gives an idea of how a depressed person may perceive life. The data presented in various studies indicate that if appropriate measures are not taken on time, a lot of people may suffer from severe depression, which can be determined by various symptoms [11–13]. People express themselves on social media platforms, through quotes, Shayari, etc., which reveals their likes and dislikes [12]. Factors such as genes and family history can also lead to depression. People who are deeply engaged in social media platforms keep posting regularly. This data can be used for identifying depression among people with the help of text extraction for the text quoted by people or through feature extraction of the images posted by them [13]. Addiction to smartphones and students' learning habits can also be an important factor [14–16]. Text data can be used for determining characteristics in order to analyze the behavior. In addition to text, some studies suggested withdrawal of emotional features from social media sites such as Facebook and Twitter can be analyzed for detecting depression symptoms [17]. The main benefits of text data analysis are easy analysis and preprocessing, data availability, less memory requirement, and easy pattern conversions [17].

*Author for correspondence

The main motivation behind this study was to present a systematic review and analysis of the machine intelligence approaches employed for mental health issues and depression detection. The major machine intelligence approaches considered here are machine learning, deep learning with hybrid boosting, and machine vision. Machine learning approaches can be further divided into three categories [18–24]: unsupervised, supervised, and reinforcement learning [18–24]. The process of identification of patterns from nonlabelled data is called unsupervised learning. Some of the methods in this category are principal component analysis (PCA), linear discriminant analysis (LDA), singular value decomposition (SVD), k-means (KM), and fuzzy c-means (FCM). The process of learning and training using labeled data is called supervised learning. Some of the methods in this category are linear regression (LR), naïve Bayes (NB), support vector machine (SVM), decision tree (DT), neural network (NN), k-nearest neighbors (KNN), and random forest (RF). The process of working with intelligent agents is called reinforcement learning. Some of the methods in this category are deep Q neural network (DQN) and recurrent neural network (RNN). A deep learning network is a subset of machine learning with multiple layers of NN [25]. It can be either generative and discriminative. Examples of deep learning networks are convolutional neural network (CNN) and restricted Boltzmann machine (RBM). We included these algorithms in this study as these algorithms can detect mental illnesses, especially depression [21–26]. Moreover, these algorithms are helpful in problem solving, knowledge representation, machine vision, and in social intelligence and information processing. The complete categorization is shown in *Figure 1*.
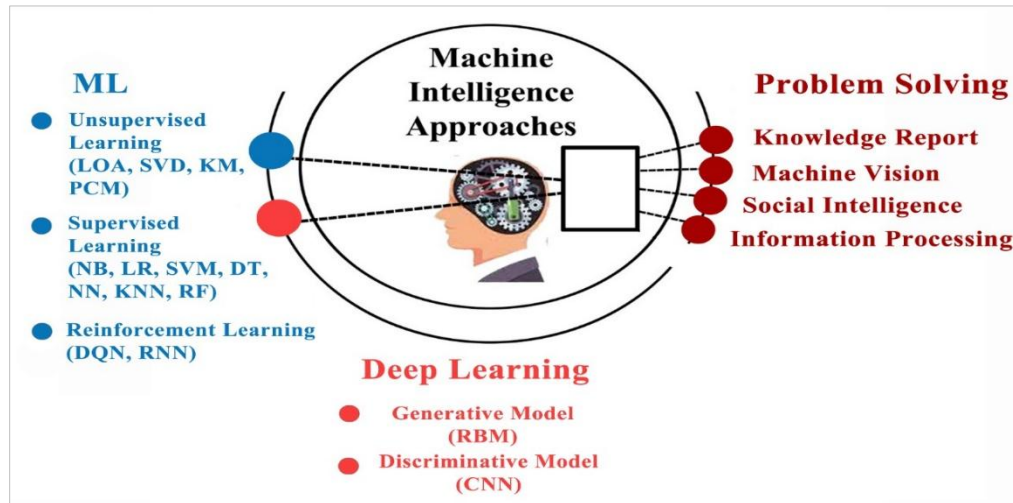

Figure 1 Pictorial representation of machine intelligence approaches and knowledge mapping

In recent years, several studies have employed machine learning, deep learning, and hybrid boosting methods to detect mental health issues and depression. However, several challenges still remain to be addressed, such as appropriate combination of models [27], validation of hybrid models in different environments [27], and insufficient resources allocated toward mental disorders in developing countries. Symptoms of mental illness vary from person to person and case by case; thus, the detection and prediction strategies should be modified accordingly. As such, implementing an appropriate computational method is a major challenge [28, 29].

The study objectives of this paper are as follows:

1. To provide a systematic review and analysis of the related work on mental health issues and depression detection.
2. To cover machine intelligence approaches such as machine learning and deep learning for the detection of mental health issues and depression.
3. To analyze the approaches in terms of performance, impact, and applicability.

The contributions of this paper are as follows:
- Complete discussion and analysis of published papers of the past seven years.
- Parametric and impact analysis of the approaches along with their advantages and limitations.

- Discussion on the solutions based on the researcher's recommendations and an analysis of the review.

This paper consists of six sections. Section 1 presents the introduction. In Section 2, the mapping methodology of the systematic review is discussed. In Section 3, the mapping report for the systematic review is presented, which covers the mapping strategy for the complete work. In Section 4, the approaches used for mental illness and depression, along with the advantages and limitations, are discussed; moreover, the datasets used in the literature are presented. Discussion and analysis are covered in Section 5. Finally, conclusions and future recommendations have been provided in Section 6.

2. Mapping methodology of systematic review

In this section, the mapping methodology of the systematic review is presented. The review and meta-analysis were categorized into four classifications. The first part focused on publication trends, criteria of the selection of papers and the methodological scenario for mental illness and depression detection. The second part was intended for machine intelligence methods and their advantages and limitations for mental illness and depression detection. The third part focused on the dataset's discussion and applicability. The fourth part focused on the complete analysis and discussion along with suggestive measures; it also covered the overall analysis, including the methodological impact, result impact, current trends, and some suggestions based on the limitations and challenges.

For the systematic review, the following mapping questions (MQ) and research questions (RQ) were prepared and are discussed in the subsequent section.
MQ1: What are the inclusion and exclusion criteria for paper selection?

MQ2: Which major publishers were considered for the study?
MQ3: What is the current scenario of mental illness and depression detection?

RQ1: What are the approaches used in the detection of mental illnesses and depression disorder?
RQ2: What are the major advantages and challenges of the approaches used in the detection of mental illness and depression disorder?
RQ3: What are the most commonly used datasets in the same domain?

3. Mapping report for the systematic review

3.1 MQ1: What are the inclusion and exclusion criteria for the paper selection?

A total of 777 papers on mental health problem published between 2015 and 2021 were considered. Finally, 114 papers were selected for this study. Database considered for the selection of the papers include IEEE Explore, Springer, Science Direct, Nature, Hindawi, MDPI, and Wiley. After excluding duplicate records (N = 56) from the papers (N = 777), only 721 papers remained for screening. After screening of the papers by title and abstract, only 523 records were left. *Figure 2* shows the PRISMA [30] style of representation adopted for paper selection in the study. Based on the inclusion and exclusion criteria, only 114 papers were selected for the final study. The inclusion and exclusion criteria are shown in *Table 1*.

Number of papers considered for each year along with the percentage distribution is presented in *Table 2*. We considered 20 papers from the year 2021, whereas from the years 2020 and 2019, we included 27 and 26 papers, respectively. We reviewed 23, 11, 5 and 2 papers from 2018, 2017, 2016 and 2015, respectively.

Table 1 Inclusion and exclusion criteria

|	**Inclusion Criteria (IC)**
IC1	Papers from the mental illness, especially depression detection based on machine intelligence approaches were considered.
IC2	Papers with result and some of systematic reviews in the same direction were considered for this study.
IC3	Only published papers with full online details were considered.
	Exclusion Criteria (EC)
EC1	Only English language papers were considered for the study.
EC2	Excluded the papers which are not available electronically.
EC3	Excluded full clinicians' papers and outcomes.
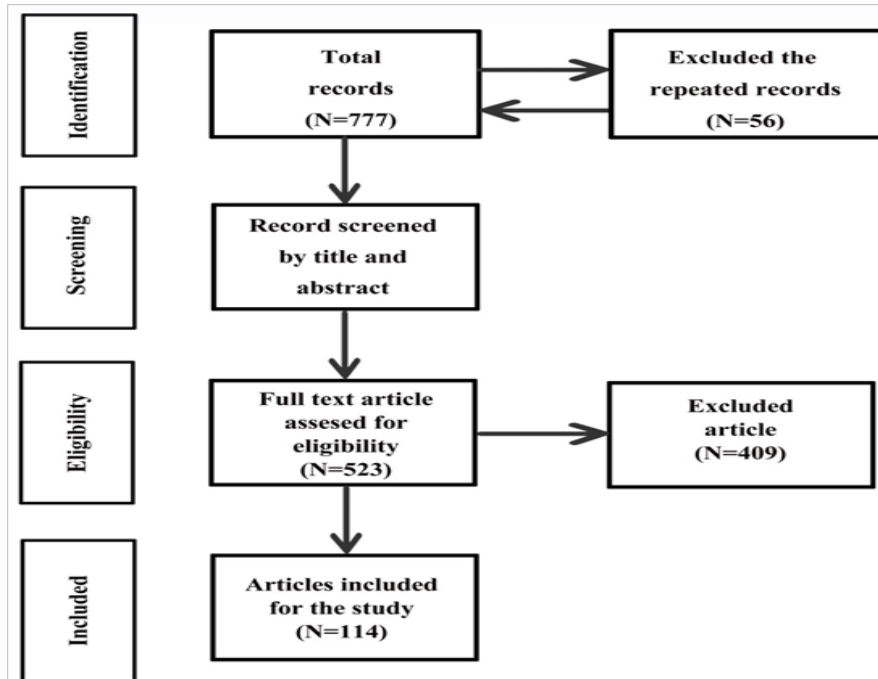
Figure 2 PRISMA style of representation adopted in case of paper selection for the study

Table 2 Number of papers considered for each year along with the percentage distribution

Sr. No.	Year	Number of papers	Percentage (%)
1	2021	20	17.54
2	2020	27	23.68
3	2019	26	22.81
4	2018	23	20.18
5	2017	11	9.65
6	2016	5	4.39
7	2015	2	1.75
		114	100.00

3.2 MQ2: What are the major publishers considered for the study?

For the study, we filtered out papers from a total of 777 papers from various publishers such as Springer, Science Direct, IEEE Xplore, MDPI, Hindawi, Wiley, and Nature. The paper distribution in terms of publishers is presented in *Table 3*.

Table 3 Distribution of papers based on different publishers

Sr. No.	Publication	Number of papers	Percentage	Database
	IEEE Conference	36	31.57	IEEE Xplore
2	IEEE Journals	19	16.66	IEEE Xplore
3	Elsevier Conference	3	2.63	ScienceDirect
4	Elsevier Journal	19	16.66	ScienceDirect
5	Springer Journal	6	5.26	Springer
6	Others (MDPI, Hindawi, Wiley, Nature etc.)	31	27.19	PubMed, Google Scholar, Wiley, Hindwai
Total		114	100	

3.3MQ3: What is the current scenario for the mental illness and depression detection?

Many people are connected through social media; this can be utilized for affective computing [31]. Affective computing refers to the study of emotions. With the help of these social platforms, human emotions, which reveal the interest of individuals, can be identified. Human-to-human interaction is dwindling whereas human-to-device interaction is increasing [32]. Depression detection can also be performed by reading facial expressions. This employs image processing, which is performed using various classification and clustering techniques [33]. Similarly, blinking, a special feature of the face, can help detect depression; this is useful in detecting depression from video visuals [34]. Several such modalities are available nowadays. Each modality helps detect features related to mental illness. Modalities can be audio-based, video-based, or text-based. All these modalities need some extraction techniques based on which the level of depression can be identified [35]. When a person interacts with a human or a device, they typically express their views either in a silent way or in aggressive ways. Various techniques can be used for detecting the scope of treatment and suggest remedies for depression.

According to the WHO, the fight against depression is difficult because of limited resources and trained personnel. Depression is predicted to a major health issue in the coming years [36]. In addition to audio and video features, speech recognition can be used to detect depression. The speech power of a normal person differs from that of a depressed person. Thus, vocal analysis can be employed to detect depression [37].

There are various features of depression, which vary according to the gender and age of a person. Some features are easily detectable in adolescents and some features are noted specifically in females. So, there is an impact on both on the disease like mental disorder [38]. An example is post-partum depression in females. As there are multiple forms of depression, the features depend upon the characteristics of individuals; for example, detection can be performed by analyzing interactions, observation of behavior, and from medical parameters. Medical parameters such as heart rate, pulse rate, blood pressure, and electroencephalography can be used to detect depression [39]. Depression is commonly detected by observing mood disorders as it is directly related to the emotions of a person. It directly affects the health of a person, both physically as well as mentally.

Personality attributes can also be considered for behavior prediction. Numerous surveys and questionnaires, such as Patient Health Questionnaire-9 (PHQ-9), have been used for such analyses [40]. Analysis reports helped in improving the platform and further enhanced the outcomes based on the selection of input parameters.

There are numerous factors that contribute to depression. However, various treatments, such as through medications or by mental therapy procedures, are available to help individuals come out from this depressive behavior. A positive attitude helps in recovery and the support from their loved ones can help immensely.

In this paper, we discuss the types of mental disorders and the relevant research. This paper mainly focuses on identifying the challenges in the mental health disorder detection.

4.Systematic review based on machine intelligence approaches

Review and analysis of the approaches used for mental illness and depression disorder detection are presented in this section. It also covers the major advantages and limitations of the approaches.

4.1RQ1: What are the approaches used for the mental illness and depression disorder?

In 2015, Demirci et al. [41] identified geriatric depressive symptoms and stated use of phone in severe depression, sleep quality, and anxiety as major factors. They used a 33-item smartphone addiction scale (SAS) and a six-point Likert-type self-rating. They assessed the sleep quality using Pittsburgh sleep quality index (PSQI), which is a measure of the subjective sleep quality. The Beck Depression Inventory (BDI) and Beck Anxiety Inventory (BAI) were used for determining the severity of depression. To examine the association between smartphone use and depression/anxiety and sleep quality, they used the LR. They performed data analysis by using the statistical package for the social science (SPSS) 15.0 for Windows. Their results revealed that there were 248 smartphone users and 71 were non-smart phone users. They found that the SAS scores of females (80.3%) were slightly higher than that of males (19.7%).

In 2016, Shim et al. [42] presented the NeuroScan SynAmp2 amplifier for recording the scalp electroencephalogram (EEG) data. The EEG data was processed using the Scan 4.3 software. They applied

a sensor-level feature set, a source-level feature set, and a combined feature set. Furthermore, for feature selection (sensor, source, and combined), they compared Fisher's scores. They applied leave-one-out cross-validation (LOOCV) with SVM for obtaining the classification accuracy. A high accuracy of 88.24% was obtained when combined (sensor + source) feature set was used. They used the combined features (sensor + source) for machine learning-based diagnosis of schizophrenia.

In 2017, Lins et al. [43] applied a regression model based on multilayer feedforward backpropagation artificial neural network (ANN) by using the error backpropagation learning method for classifying individuals with different stages of neurodegenerative diseases. The ratios considered for training, validation, and testing was 50%, 25%, and 25%, respectively. They designed a nonlinear regression model to classify healthy and pathological aging. Furthermore, they used NN, RF, SVM, and stochastic gradient boosting tree (SGBT). For comparison, they used simple LR model. RF method outperformed the other methods in all indicative variables.

In 2018, Al Jazaery and Guo [44] performed video preprocessing by subsampling the data. For this purpose, they used face detection and alignment. They used a combination of 3D convolutional network (C3D) and RNN (RNN-C3D) for obtaining spatiotemporal information and further sequencing. Audio/visual emotion and depression recognition challenge-13 (AVEC-13) and audio/visual emotion and depression recognition challenge-14 (AVEC-14) data were used in the experiments and for obtaining the spatiotemporal information. Finally, the sequence of spatiotemporal information was extracted. Using the global spatiotemporal information obtained from face expressions, they proposed an approach of detecting depression by using RNN-C3D.

In 2019, Iliou et al. [45] developed a data processing method for detecting depression. They considered a group of 249 students. They applied seven classifiers, namely fuzzy logic, SVM, nearest-neighbor, RF, Java-based implementation of repeated incremental pruning to produce error reduction (RIPPER) (JRip), multilayer perceptron (MLP), and C4.5, on the collected dataset. They found complete percentage (100%) in the classification and prediction part.

In 2019, Tariq et al. [46] studied the use of social media networks for detecting depression. They considered attention deficit hyperactivity disorder (ADHD), anxiety, bipolar, and depression as mental illnesses and applied the SVM, NB, and RF classifiers with and without co-training methods. The data of 1000 posts for each of the mentioned category of mental illness were downloaded using Python for Reddit. A total of 3922 posts were considered in the study. They adopted F score, recall, and precision as the performance parameters of the classifiers. Better and more effective results were found for the classifiers employed with co-training. The F score calculated for SVM was 0.84 in case of anxiety and 0.67 in the case of ADHD, whereas NB and RF yielded a F score of 0.70 and 0.71 for depression and bipolar, respectively. The RF, NB, and SVM classifiers employed with the co-training techniques outperformed classifiers used without co-training.

In 2019, Ay et al. [47] proposed a hybrid model that used EEG signals to detect depression. They considered 9116 EEG records, among which 4798 corresponded to depression and 4318 were normal records. A total of 30 subjects were included in the analysis. They developed a hybrid model consisting of CNN with long-short term memory (LSTM) for detecting depression by using EEG signals. Tenfold cross-validation was used in the hybrid model. An accuracy of 99.12% and 97.66% was obtained for right and left hemispheres, respectively.

In 2019, Li et al. [48] attempted to detect depression by using EEG features with machine learning methods. They studied the EEG data of 28 subjects recorded using the Net Station software, which used 128-channel HydroCel Geodesic Sensor Net (HCGSN). They employed ensemble and deep learning for processing the features. SVM was used as the classifier for ensemble learning and CNN was adopted for deep learning. They achieved an accuracy of 89.02% and 84.75% in the ensemble and deep learning models, respectively.

In 2019, Besteher et al. [49] applied relevance vector regression (RVR) on 743 healthy controls (HCs) in the Information eXtraction from Images (IXI) and Open Access Series of Imaging Studies (OASIS) databases. They also analyzed the 3D high-resolution structure of magnetic resonance imaging (MRI) scans in 38 patients with depression and 40 HCs; moreover, they tested the hypothesis of accelerated aging in major depression.

In 2019, Xing et al. [50] applied SVM and heart rate variability (HRV) characteristics for constructing a simple and efficient detection model. Accuracy,

sensitivity, specificity, area under the curve (AUC), and precision were considered as the performance parameters. An accuracy of 89.66% was obtained using the detection model.

In 2020, Yang et al. [51] proposed a framework for analysis of big data for detection of depression by using social networks. They used the dataset of myPersonality, a Facebook application, and applied various supervised machine learning algorithms: KNN, NB, SVM-linear RF, classification and regression trees (CART), logistic regression, and NN. The proposed framework was evaluated based on accuracy, balanced accuracy, precision, recall, F score, receiver operating characteristic curve (ROC), AUC, and Kappa values. They applied tenfold cross validation on all the algorithms. RF yielded the best performance: accuracy = 71.21%, 68.42% of balanced accuracy, 70.21% of precision and a recall of 86.84%. In contrast, LR yielded the worst performance.

In 2020, Altameen et al. [52] proposed a method for assisting robots in diagnosing mental illness by observing different parameters. They used the Collaborative Psychiatric Epidemiology Survey (CPES) dataset. The deep reinforcement learning (DRL) method was used for robot assistance. They compared the error rate of DRL (0.083) with that of Bayesian network (BN) (0.288), MLP (0.204), and deep learning neural network (DLNN) (1.131). The other performance parameters were F1 score, recall, and precision. They attained a precision of 99.689%, a recall value of 99.35%, and an accuracy of 98.42% the in DRL process.

In 2020, Abd et al. [53] performed a review analysis on the detection of mental health issues in the data sources available in online social networks (OSNs). They studied the various data sources and methods used for feature extraction, as well as the machine learning techniques. In total, 2770 papers were collected between 2007 and 2018. They found that the studies considered data from different application programming interfaces (APIs), Facebook, Twitter, and the online data from online communities, microblogs, etc.

In 2020, Kumar et al. [54] considered two datasets named Depression Anxiety Stress Scales 42 (DASS42) and DASS21. They applied eight machine learning algorithms—RF, radial basis function network (RBFN), local nearest neighbor (LNN), NB, J48, BN, and MLP—and grouped them into four

categories: probabilistic, NN, tree-based, and nearest neighbor. They used the Waikato Environment for Knowledge Analysis (WEKA) tool for applying the machine learning algorithms and classified mental health illnesses into three psychological disorders having five severity levels. They applied fivefold cross-validation for improving the accuracy. The performance parameters considered in these algorithms were recall, Kappa, F1 score, precision, error rate, and accuracy rate. In case of DASS42 dataset, RBFN produced a high accuracy 96.17%, 96.02%, and 97.48% in detecting depression, stress, and anxiety. The DASS21 dataset yielded 100% accuracy in anxiety detection by using RF, 96.55% accuracy in depression detection by using RBFN, and 96.55% accuracy in stress detection by using MLP method.

In 2020, Yue et al. [55] proposed an approach for identifying depression by sensing data collected from smartphones. The data used in this study was taken from Internet traffic. They applied SVM, multivariate regression, and RBFN to predict depressive features. Their result showed that Internet usage can be used as the most effective depression prediction factor. In this case, the F1 score was also high (80%).

In 2020, Baek and Chung [56] proposed the context deep neural network (DNN) consisting of information to predict situations and environment. To predict the context influencing the risk of depression based on regression, each context information related to the predictor variables of depression was considered as an input.

In 2020, Cai et al. [57] proposed a multimodal approach to recognize depression by using the fusion technology on negative and positive audio stimuli. They recorded the EEG signals for patients with depression and HCs. Genetic algorithm was employed for feature weighting in order to improve the performance of the recognition framework. They compared the accuracy of SVM, DT, and KNN classifiers and used tenfold cross-validation for the evaluation. A high accuracy of 86.98% was achieved by KNN for fusion modality of negative and positive audio.

In 2020, Qin et al. [58] established a model for the prediction of mental health status. They applied SVM and LR for missing values in the dataset. Furthermore, they employed the DT, eXtreme Gradient Boosting (XGBoost), KNN, and LR models. They used fivefold cross-validation as the

verification approach. They used accuracy for performance comparison and mean squared error (MSE) for model comparison in regression experiment. For the classification of self-reported health, the LR, XGBoost, SVM, ANN, and RF classifiers were used. ANN yielded the highest accuracy (69.9%), whereas LR yielded the lowest MSE (0.0630).

In 2020, Sharma and Verbeke [59] studied the depression detection mechanism. They balanced the dataset by creating multiple samples by performing original sample (OR), over sampling (O), under sampling (U), ROSE (R) sampling, and over-under (OU) sampling. Then, they used five XGBoost algorithms —Xgb.U, Xgb.OR, Xgb.OU, Xgb.R, and Xgb.O—on each sample for identifying the important biomarkers. The balanced F1 score, recall, precision, and accuracy were used as the performance parameters. They obtained the best performance results for XGBoost model on Xgb.O, with the highest accuracy of 0.9729, a balanced accuracy of 0.9765, a recall value of 0.9987, an F1 score of 0.9762, and a precision value of 0.9548.

In 2020, Morrow et al. [60] discussed machine learning methods for identifying the factors for ADHD. They took the diagnosis data of 6630 participants with ADHD in 2016–2017 from the National Survey of Children's Health. They applied three machine learning models: CART, an ensemble model called RF, and a deep multilayer neural network. For the descriptive statistics, they used SPSS v26. Tenfold cross-validation was used for characterizing the efficiency of the models. The deep multilayer neural network attained the highest AUC (0.71) and 95% Confidence Interval (CI).

In 2020, Lin et al. [61] utilized machine learning methods to identify the reason behind high psychological stress and suicide risks among military personnel. They studied the history of a cohort of 3546 military personnel. They used two-suicide ideation for training and testing. They applied DT, gradient boosting tree (GBM), SVM, LR, MLP, and RF as the six machine learning techniques for suicide ideation. The metrics used for the evaluation purpose were sensitivity, accuracy, specificity, recall, ROC AUC, specificity, precision-recall area under the curve (PR-AUC), F1 score, and Pearson's score. They used tenfold cross-validation for the training and testing datasets. The synthetic minority over sampling technique (SMOTE) was employed for minority data. An accuracy of 100% was achieved

using SVM and MLP in suicide ideation1 and 99.9% using SVM and MLP in suicide ideation.

In 2020, Su et al. [62] developed a machine model for the prediction of suicidal behavior in children and adolescents. They used the data of 129,485 patients from Connecticut Children's Medical Center (CCMC) HER database from October 1, 2011 to September 30, 2016. They employed the L1 regularization (L1-LR) model for univariate screening. They evaluated the predictive performance by examining the out-of-sample performance, and used specificity, AUC, sensitivity, and positive predictive value (PPV) as the metrics. They grouped selected predictors and obtained AUC values of 0.81–0.86. The proposed model exhibited an accuracy of 53%–62% for positive subjects of suicide and a specificity of 90%.

In 2020, Priya et al. [63] focused on psychological issues. They used the DASS21 dataset for evaluating mental health. They applied the RF, NB, DT, KNN, and SVM classifiers and considered depression, stress, and anxiety as the three psychological mental issues. A total of 348 participants were included. They applied these classifiers by using R programming language in RStudio version 3.5. The metrics selected were error rate, accuracy, specificity, sensitivity, recall, F1 score, and precision. They achieved F1 scores of 0.711 in RF for stress, 0.836 in NB for depression, and 0.527 in KNN for anxiety.

In 2020, Frassle et al. [64] employed generative embedding to predict the clinical trajectory for 85 patients suffering from major depressive disorder (MDD). They used generative models with the discriminative classifiers and predicted chronic depression in patients with fast remission by combining the generative model with SVM. The performance evaluation metrics used were PPV, accuracy, specificity, AUC, negative predictive value (NPV), sensitivity, and balanced accuracy for three binary classifications of different MDD groups. They found that generative embedding outperformed the classification based on SVM. They achieved a balanced accuracy of 79% in chronic depression versus fast remission and a balanced accuracy of 61% in gradual improvement versus fast remission.

In 2020, Tate et al. [65] developed a model to predict mental health problems by using machine learning. Initially, they included 474 variables in their dataset. The machine learning algorithms used were SVM, NN, XGBoost, RF, and LR; the complete analysis was performed in R. The data was split into three

sets: training (60%), tuning (10%), and testing (30%). SMOTE was used for the training set and it was implemented in R package. They found that SVM and RF yielded the highest AUC value (0.739) with 95% CI of 0.708–0.769 whereas SVM yielded an AUC value of 0.735 with 95% CI of 0.707–0.764.

In 2020, Betts et al. [66] developed a model to identify the risk of chronic depression after the occurrence of postpartum depression. They included the data of administrative health for patients giving birth in an Australian state from January 2009 to October 2014. They applied the boosted trees, LR, and ElasticNet models. They found that the boosted trees algorithm yielded the best results, that is, an AUC of 0.80, followed by the ElasticNet model, which yielded an AUC of 0.78, and the LR with lowest discrimination (AUC: 0.76).

In 2020, Lin et al. [67] analyzed the data gathered from CPES. They compared seven categories of drugs. They applied the PCA and LR in the proposed model. Their model yielded a mean accuracy of 83% and an internal validity of 72%.

In 2020, Cho et al. [68] developed a model to predict depression over a period of 18 months. They considered the data of 433,190 participants. They split the data into 70% and 30% for training and testing sets, respectively. They applied RF with fivefold cross-validation and achieved a sensitivity of 0.737, an accuracy of 0.822, a PPV of 0.097, a specificity of 0.824, a NPV of 0.992, and an AUC of 0.839.

In 2020, Goncalves et al. [69] applied the data mining (DM) technique by using the cross industry standard process for data mining (CRISP-DM) methodology. The RapidMiner software was used. They used different machine learning methods, namely KNN, gradient boosting tree (GBT), DT, and RF, for constructing DM models. The evaluation metrics used were specificity, accuracy, and sensitivity. They applied tenfold cross-validation and split the data in 70% and 30% for training and testing, respectively. RF yielded the best results, whereas KNN performed the worst. RF yielded the best accuracy of 94.24%, sensitivity of 98.78%, and specificity of 95.22%.

In 2020, Martínez-Monteagudo et al. [70] studied the role of cybervictimization in causing depression, stress, anxiety, and suicidal thinking. They used the DASS21 dataset and the European cyberbullying intervention project questionnaire. They included

1282 university students in their study. Their results suggested that being a victim of cyberbullying can increase the chances of suicidal thoughts. They highlighted the high prevalence of cyberbullying in universities. They found that 18.6% were victims of cyberbullying, among which 72.2% reported anxiety, 68.1% reported depression and 75.2% reported stress.

In 2020, Sağbaş et al. [71] used the rank of the features calculated using the gain ratio feature selection algorithm. For information acquisition, DT was used. To overcome bias, tenfold cross-validation was used to evaluate the classifier model. The results obtained using KNN, DT, and BN methods were compared.

In 2020, Castillo-Sánchez et al. [72] performed a comprehensive review on suicide detection by using social networks. They analyzed various data collection techniques, development processes, and validation metrics. They adopted a framework in which scoping of reviews was divided into various categories. Nearly 75% of the studies proposed the model for classifying the suicide categories. Approximately 50% of studies used the techniques based on DM where they performed feature extraction and identification. Other methods such as linguistic inquiry and word count (LIWC), LDA, latent semantic analysis (LSA), PCA, and nonnegative matrix factorization (NMF) were also used. In contrast, they found that SVM was mostly implemented i.e., approx. 62.5%. The second technique, which was mostly used was DT i.e., approx. 43.75% further, it was followed by LR and RF with 31.25% and 25%. These machine learning models were implemented using Python. The performance metrics used were recall, F score, and precision.

In 2020, Guntuku et al. [73] collected two datasets from Twitter and applied machine learning models on overall sentiments, anxiety, stress, and expressions with loneliness. They used Cohen's d for quantization. They compared the data for 2020 and 2019 and found that Cohen's d was higher in 2019 for all the cases. Cohen's d was –0.97 for sentiment, 1.5 for stress, 4.4 for anxiety, and 1.58 for loneliness.

In 2020, Elhai et al. [74] proposed a model for problematic smartphone use (PSU) by using machine learning algorithms. They used the Smartphone Addiction Scale-Short Version (SAS-SV), DASS, Fear of Missing Out Scale (FoMOS), and Ruminative Responses Scale (RRS) in the survey. Next, they

applied least absolute shrinkage and selection operator (LASSO), RF, SVM, ElasticNet, ridge, and eXtreme gradient boosting (XGBoost) classifiers. They found that FoMOS has a major contribution in the construction of PSU.

In 2020, Oyebode et al. [75] determined the factors associated with mental health. They applied natural language processing (NLP) for data preprocessing and vectorized the reviews by using the term frequency-inverse document frequency (TF-IDF) algorithm. In the next stage, they applied multinomial naïve Bayes (MNB), RF, LR, stochastic gradient descent (SGD), and SVM. SGD yielded the best performance, with a F1 score of 89.42%, a precision of 90.2%, a recall value of 88.7% for positive reviews, a recall value of 90.6% for negative reviews, and a precision of 89%.

In 2020, Zhang et al. [76] proposed a framework for the prediction of postpartum depression (PPD) among women. The framework model was trained using regularized LR, MLP, DT, XGBoost, and RF. They used the Seaborn and Scikit-learn libraries in Python for the training and evaluation, respectively, of the algorithm. They found that the machine learning and electronic health record (EHR) can be used to predict PPD. In the development phase, the AUC and CI of the model were 0.937 and 0.912–0.962, respectively; in the validation phase, the AUC and CI were 0.886 and 0.879–0.893, respectively.

In 2021, Hong et al. [77] adopted the SVM-recursive feature elimination (SVM-RFE) classifier for studying the features associated with suicide attempts and suicide ideation. They used the LOOCV classifier for training and testing in order to avoid the overfitting problem. The performance metrics considered were PPV sensitivity, NPV, specificity, and accuracy. Their SVM-RFE algorithm yielded a sensitivity of 73.17%, an accuracy of 78.59%, a specificity of 84.0%, a PPV of 88.24%, and a NPV of 65.63%.

In 2021, Zulfiker et al. [78] investigated six classifiers—KNN, adaptive boosting (AdaBoost), gradient boosting (GB), XGBoost, bagging and weighted voting (with GB, bagging, and AdaBoost)—by using various sociodemographic and psychological information to detect depression. For improved accuracy and depression prediction, they used SMOTE as it helps in reducing the class imbalance of the training data. The AdaBoost classifier outperformed others, with an accuracy of 92.56%.

In 2021, Parghi et al. [79] determined, which machine learning algorithm can better predict suicidal behavior. They compared the predictive power of various machine learning algorithms such as GB, LR, and RF. They employed split sample, enhanced bootstrap, and SMOTE as the sampling methods. The highest area under the precision-recall curve (AUPRC) values of 0.710 and 0.705, precision values of 0.980 and 0.940, balanced accuracy of 0.669 and 0.744, and recall values of 0.339 and 0.489 were obtained using RF and GB, respectively, for enhanced bootstrap.

In 2021, He et al. [80] proposed an integrated framework based on deep local global attention CNN and used it for detecting depression by adopting CNN with attention mechanism. For learning a deep and global representation, they used weighted spatial pyramid pooling. They achieved a reduced root mean square error (RMSE) value.

In 2021, Wardenaar et al. [81] used the SuperLearner (SL) algorithm for the identification of an optimally weighted parameter. They applied various learners such as ElasticNet, RF, GB, SVM, and SL and calculated the corresponding probability for chronic course, partial recovery, and full recovery. Their results revealed the prediction from low to high for the range of 0.41–0.91.

In 2021, Uddin et al. [82] used LSTM based on RNN to identify the symptoms of depression. They applied their method on two datasets. Initially, they considered the data from ung.no of Norway and then applied the one-hot method. They visualized the features by using LDA. The features extracted from one-hot method were applied to deep RNN (DRNN) to identify the symptoms of depression. Moreover, they compared the results obtained using their method with those for other machine learning algorithms such as DT, LR, ANN, SVM, CNN, and deep belief network (DBN). Their approach achieved a high accuracy of more than 98% in the first dataset. They applied tenfold cross-validation and in the first and second datasets, respectively and used RNN on robust features, achieving a mean accuracy of 99%. They utilized the local interpretable model-agnostic explanations (LIME) and explainable artificial intelligence (XAI) algorithms for the validation process.

In 2021, Wang et al. [83] constructed an architecture based on functional near-infrared spectroscopy (fNIRS) for depression detection. They considered the frequency and time domains for the extraction of features and used the DNN. The data were classified into 70% and 30% for the training and testing sets, respectively. They trained their model on the raw data from the Alexnet and ResNet 18 networks. ResNet18 yielded a high accuracy of 0.76. Furthermore, SVM was used for training. After performing the testing of data, AlexNet yielded a higher accuracy than the SVM. AlexNet yielded an accuracy of 0.83, a precision of 0.79, a recall value of 0.83, and F1 score of 0.80. AlexNet yielded the best results on correlation data, with a high accuracy of 0.90 and a precision of 0.91.

In 2021, Shin et al. [84] considered machine learning approaches such as LR, Gaussian naive Bayes (GNB), SVM, and MLP. They considered 70% and 80% for training and the remaining 30% and 20% for prediction. MLP yielded superior results, with an AUC of 65.9%, a sensitivity of 65.6%, and a specificity of 66.2%.

In 2021, Shi et al. [85] used the structural magnetic resonance imaging (sMRI) and fMRI for multimodal imaging and multilevel characterization with multiclassifier (M3) to distinguish patients with schizophrenia from healthy controls. They identified the influence of global signal regression (GSR) in the classification of schizophrenia. They utilized the Brainnetome-246 atlas and Power-264 atlas with and without GSR and achieved the highest specificity and accuracy with GSR. Without GSR, for Brainnetome-246 atlas, a specificity of 93.75%, an AUC of 0.8491, an accuracy of 83.49%, and a sensitivity of 68.69% were obtained. Overall, M3 was found to be an effective tool.

In 2021, Richter et al. [86] used a RF-based algorithm and the cross-validation technique and categorized the prediction results as anxiety vs. control and anxiety group vs. depression group. They achieved a prediction success of 76.81% and 69.66% in the healthy controls and clinical patients groups, respectively.

In 2021, Na et al. [87] used different machine learning algorithms (SVM, GBM, RF, regularized LR, and ANN) with HRV. To determine the best model, they used tenfold cross-validation. L1-regularized LR yielded the best results: an accuracy of 0.784, a sensitivity of 0.83, a specificity of 0.73, a

Matthews correlation coefficient (MCC) of 0.57, and a F1 score of 0.79.

In 2021, Aljameel et al. [88] performed a study on patients of COVID-19. They analyzed the data by using XGBoost, RF, and LR classifiers and applied SMOTE for data balancing and tenfold cross validation for data partitioning. They performed the experiments by using 20 clinical features. RF achieved the highest accuracy of 0.95 and an AUC of 0.99.

In 2021, Cho et al. [89] developed a predictive model to detect depression. They used the Scikit-learn library in Python 3.7. Two datasets of 2014 and 2016 were considered; the dataset of 2014 was considered for training and validation, and that of 2016 was considered for testing. SMOTE was used for reducing the data imbalance and tenfold cross-validation was used tuning the hyperparameters. They used LR for reducing the overfitting problem. They applied L1 regularization LASSO and L2 regularization (ridge regression) for the overfitting problem. The performance metrics considered were MCC, sensitivity, specificity, accuracy, and precision. Their model achieved a ROC value of 0.903 and an accuracy of 0.828.

In 2021, Gokten and Uyulan [90] divided the dataset of patients into MDD and post-traumatic stress disorder (PTSD) groups. They used SPSS 22 for data analysis. They used RF as an ensemble model for solving the regression and classification problems. DT was used for the classification and regression of continuous datasets. The implementation of RF was performed in Scikit-learn. They used ten-k-fold cross-validation for testing the performance of the classifier. Their model achieved an accuracy of 82%, a F1 score of 81%, a precision of 0.81, a recall value of 0.80, and an AUC of 0.88 in MDD, and an accuracy of 72%, a F1 score of 71%, a precision of 72%, a recall value of 71%, and an AUC of 0.76 in PTSD.

In 2021, Huberts et al. [91] compared the performance of different algorithms such as one-class SVM (OC-SVM), LR, RF, and ElasticNet with that of the XGBoost prediction algorithm. XGBoost has been used to extend the range of the model for proper estimation. LR produced a slightly higher probability for person-weeks, with a ratio of 1.226. Overall, XGBoost performed better compared to SVM, LR, and RF.

In 2021, Du et al. [92] recruited 300 patients of MDD. They assessed the Chinese version of the Clinically Useful Depression Outcome Scale supplemented with Diagnostic and Statistical Manual of Mental Disorders-5 (DSM-5) mixed subtype (CUDOS-M-C), Hamilton Anxiety Scale and Montgomery–Asberg Depression Rating Scale and Young Mania Rating Scale. The Cronbach's alpha, PCA, and intraclass correlation coefficient (ICC) were used for assessing the validity and reliability of CUDOS-M-C. A PCA of 54.82% was obtained as the cumulative variance, Cronbach's alpha was 0.892, and ICC was 0.853. The AUROC of CUDOS-M-C was 0.927, sensitivity was 91.65, and specificity was 79.9%.

In 2021, Liu et al. [93] applied machine learning algorithms in a study on mental health centers. They found that ElasticNet yielded a more focused result and reduced the overfitting problem. Python 3.6 with Scikit-learn 0.221 was used for analyzing the machine learning algorithms, and tenfold cross-validation was used for the external validation. The model yielded an accuracy of 72%, a specificity of 69.8%, and a sensitivity of 74.2%.

In 2021, Li et al. [94] performed an experiment to detect the condition of patients by applying various algorithms such as GB, LR, RF, and SVM. The ROC curve, specificity, accuracy, confusion matrix,

sensitivity, precision, and recall were used as the performance metrics to evaluate the methodologies adopted using different classifiers. They found that GB yielded the best results (an accuracy of 76.92%). This method yielded an accuracy of the 66.67% of in the male group and 71.73% in the female group. Similarly, they obtained high accuracy in the older (76.92%) and younger (53.85%) groups.

In 2021, Salankar et al. [95] presented an effective approach for identifying stress markers at the temporal, occipital, frontal, and central lobes. For data preprocessing and decomposition, they used variational mode decomposition (VMD). They derived the Poincare plots from eight variational modes. Furthermore, the Wilcoxon test was applied for the extraction of statistical significance. Lastly, SVM and MLP were used for the classification of nonstress and stress categories. They split the data into 70%, 10%, and 20% for the training, validation, and testing sets. They applied tenfold k cross-validation and chose specificity, accuracy, and sensitivity as the performance metrics. They achieved an accuracy of 100% for the frontal and temporal lobes by using MLP.

Some of the results, along with the approach and methods, are shown in *Table 4.*

Table 4 Results of some of the selected methods and approach

S.No.	Reference	Methods	Approach	Results
1	Liu et al. (2015) [96]	They collected the data and tried to identify the depressive features using Voice recording. The methods used for the pre-processing was KNN and SVM.	All types of subjects filled a questionnaire which was further evaluated by DSM-IV. As a filter approach they adopted the minimal-redundancy-maximal-relevance criterion and the Sequential Forward Floating Selection (SFFS) algorithm as the wrapper approach.	They analysed the result based on the dataset available.
2	Rukavina et. al (2016) [97]	They prepared a questionnaire and chose SVM as the classifier for performing classification.	This paper recorded the Physiological parameters of the subjects and evaluates according to gender and age.	They found the difference of 20% classification accuracy in gender-specific.
3	Cai et al.(2016) [98]	They applied the SVM classifier as a classification. The filter method is used for ranking purposes while the wrapper method is used as the performance for the predictors on the questionnaire.	Their study provides an insight that there is the existence of distinct factors in the difference of gender in the case of depression assessment.	The classification accuracy is higher than 90% in the proposed rule.
4	Li et al. (2016) [99]	This paper extracted the features from eye movement to analyze the problem. They applied RF algorithm for identifying the states. They used five different types of classifiers.	The extracted eye features were further applied to five different classifiers named KNN, SVM, RF, NB, and LR. For automatic classification, they used ROC.	In silence data RF gave the highest accuracy of 78.7% and in emotional data KNN gave the highest accuracy of 81%.
5	Alghowinem et al. 2016 [100]	They used the eye active appearances model for detecting eye movement. The SVM classifiers are used. Lastly,	In this paper they extracted the relevant features and applied SVM for classification. Different T-	The average accuracy is 88% based on feature fusion.

S.No.	Reference	Methods	Approach	Results
		multimodal fusion was introduced for combining the modalities.	Tests were applied for result comparison.	
6	Deshpande and Rao (2017) [101]	This paper focuses on data collected from the tweeter. They used NB classifier. Further the SVM was used as supervised learning for the linear and non-linear type of classification.	This paper followed the modular approach. The classifier used for NB classifier was MNB as it is good in distributed data. Further, SVM is used as supervised learning. The extracted data was further pre-processed by NLP.	MNB: Accuracy of 83% SVM: Accuracy of 79%
7	Jan et al. (2017) [102]	In identifying the facial expression, they used deep learning or hand-crafted techniques on AVEC-2014 datasets. Besides, in facial expression they used CNN models. It had 300 video clips which helped identify the depression.	The audio and video frameworks used the unimodel approaches. For Feature extraction they used hand-crafted image and then created the Visual Geometry Group (VGG) to create the groups of pre-trained CNN models. Further these features create the combination of both by using Feature Dynamic History Histogram (FDHH). They applied partial least squares and LR.	**Unimodel:** MAE of 6.68 and RMSE of 8.01 **Bimodal:** MAE of 6.14 and RMSE of 7.43
8	Zhu et al. (2017) [103]	They tried to identify depression automatically. They used the AVEC-2013 and AVEC-2014 data set. Their study used the facial appearance and dynamics in two steps using Deep Convolution Neural Networks (DCNN) in each approach.	The video data is used for the recognition of depression. They used separately Appearance DCNN and Dynamics DCNN on single video data. Further, a joint approach was used so that they can able to identify the fine-tuning of the image.	**AVEC-2013:** MAE: 7.58 RMSE: 9.82 **AVEC-2014:** MAE: 7.47 RMSE: 9.55
9	Pampouchidou et al. (2017) [104]	They used machine learning algorithm for feature extraction which was composed of classification, cross-validation methods, and regression.	They used the classification or regression-based approaches in a quantitative way.	They concluded that a video-based system is somewhat better in identifying the depression as it was missing in earlier studies.
10	Kiss and Vicsi (2017) [105]	This paper pre-processed the voice sample data by normalizing the pressure amplitude. They extracted the Low-Level Descriptors (LLD) of speech samples. They were classified with the help of SVM.	The feature they extracted with the help of LLD. Further, they statistically analyzed the difference in samples with the help of T-Test and paired sample T-Test. They used the SVM for different classification and evaluation parts for both males as well as for females. They applied the independent sample test for it.	Detection accuracy in read speech is 83% and detection accuracy in spontaneous speech is 86%.
11	Dibeklioğlu et al. (2017) [106]	They conducted an interview on the Hamilton Depression Rating Scale-24 (HDRS) for 21 weeks on intervals of 7 weeks for both Male and females. They gave the rating at three different scales after interviewing the patients like severely depressed, mildly, or remitted scale.	The researchers applied Logistic regression, classifiers on the collected data after interviewing them properly. Deep learning-based methods were used to discover non-linear features. Further, both these audio-video samples were merged with the mRMR algorithm to perform feature selection.	An accuracy of 78.67% were obtained based on the modality with head, face and voice.
12	Zucco et al. (2017) [107]	Emotion recognition was considered for sentimental detection text. PDF type data was taken as input which was pre-processed with the help of text mining and NLP such as speech tagging, stemming, word removal, tokenization, etc.	The pre-processed data was used to extract the features through N-programs negations, etc. An architecture was proposed where they can monitor the depression but its implementation, testing was proposed for future implications.	This study gave an insight into the steps of sentiment analysis and affective computing measures.
13	Madhav et al. (2017) [108]	This paper identified the depression among adults who watch TV or use computers for long hours. They performed self-analysis with the help of PHQ-9.	They found that watching TV and using the computer for more than 6 hrs. are responsible for getting the situation of depression. They used the PHQ-9 for this analysis. They performed the statistical	Higher odds of developing depression among female than male: 2.055 (CI= 1.547–2.729, p < 0.0001) was observed in the final model.

S.No.	Reference	Methods	Approach	Results
			analysis with the help of SAS 9.4 software.	
14	Pampouchidou et al. (2017) [109]	This paper extracted the features from videos. They used OpenFace application. They applied PCA for classification.	This paper extracted the motion images from the videos and compared them with history. The extracted features were compared and classified with the help of SVM. Finally, deep learning was applied for the decisions. They proposed the VGG feature fusion approach	**Proposed (VGG feature fusion):** Accuracy: 89.0% Average Recall: 88.6% F1 score: 87.4%
15	Ingle et al. (2017) [110]	This paper gives a new insight in detecting the depression in outpatient with the help of handheld devices. They discussed the comorbid depression. They surveyed some questions about PHQ-9 in their native language.	In this paper the researcher did their work out on the questions framed in PHQ-9 in the Hindi language. With the help of this they performed the self-administered analysis on the hand-held devices in the Hindi language. They used Epi info for entering the data and further it is analyzed with the help of iSurvey software.	Overall prevalence of depression was found to be 25% (95% confidence interval 16.6–34.8).
16	Kulkarni and Patil (2018) [111]	This paper considerd the input samples from AVEC-2013 dataset. They used the feature extraction process and applied the Fisher Vector (FV) method to the extracted feature. They used KNN classifier for the classification.	This paper pre-processed the image with Local Tetra Patterns (LTrP). The calculated pattern further used by FV for encoding. For this encoding they used Gaussian Mixture made because they found it an effective and efficient classifier for the image.	Accuracy of 91% achieved in LTrP.
17	Al-gawwam and Benaissa (2018) [112]	They considered videos sample and tried to capture the image from the videos. SVM classifier was used to identify the features.	They used the SVM, NB, AdaBoost and Bagging as the classifiers.	AdaBoost gave the accuracy of 92.95% for north-wind task and 88% for freeform task.
18	Stepanov et. al. (2018) [113]	This paper used machine learning techniques for effectively identifying depression detection automatically. They used LLD using openSMILE PHQ-8 was used as a tool for identifying the severity.	They used AVEC-2017 data set and withdrawn acoustic features and some characteristics based on behavioural. Regression was applied to the input. They choose SVM algorithm for the regression.	Behavioural cues achieved smaller errors and score of MAE=4.73 and RMSE of 5.54.
19	He et al. (2018) [114]	This paper analyzes the automatic depression with the help of dynamic appearance of facial. They proposed a novel approach that was based on the Dirichlet Process (DP) and the FV approach to aggregate the features for local descriptors. So that they can analyze the depression automatically.	They used the FV approach while detecting depression and DP on the AVEC-13 and AVEC-14 dataset. They used the openface framework. They applied the Dirichlet Process FV (DPFV) for organizing the features and then further adopted statistical aggregation methods for every video clip. A Support Vector Regression (SVR) was employed for the prediction of BDI-II	**Propose approach for visual data:** **AVEC-2013 (Test data):** MAE: 7.55 RMSE:9.20 **AVEC-2014 (Test data):** MAE: 7.21 RMSE:9.01
20	Kacem et al. (2018) [115]	They tried to detect the depression based on facial, head movement. They calculated the facial shape in motion with respect to their head poses movement with the help of kinetic measures.	They used the FV encoding. After extracting they used the mRMR algo for feature selection. Further, they applied SVM with gaussian kernel for different levels.	Multi-class SVM achieved the accuracy of 70.83% whereas logistic regression achieved the accuracy of 62.02%.
21	Islam et al. (2018) [116]	They tried to provide the insight on level of depression found in Facebook users. They collected the data from Facebook with the help of Ncapture. The collection was processed by LIWC 2015. The KNN was used for classification.	After processing the data, the classification technique applied was the Cubic KNN, Fine KNN, Cosine KNN, Cubic KNN, Weighted KNN, Medium KNN and Coarse KNN, so that through emotions they can detect	Coarse KNN is the best performing model. The results are as follows: Precision: 59% Recall: 77% F-measure:67%

S.No.	Reference	Methods	Approach	Results
			depression.	
22	Kiss et al. (2018) [117]	They found the speech modality helps in detecting depression. They used LLD to balance the pressure of the speech sample. Linear-SVM was used with two classification experiments.	As speech may vary, so LLD was appliedfor adjusting the amplitude pressure. From LLD they calculated the transient parts. In continuation, they calculated the Ratio of the Transient (RoT). Classification SVM was used with a linear kernel.	RoT can be considered as the good parameter. Accuracy of 81% was achieved for healthy and non-healthy speakers.
23	Katchapakirin et al. (2018) [118]	The researchers used the techniques of NLP on Facebook data. Machine learning algorithms were used.	They used RF, deep learning, and SVM algorithms for depression classifiers. WEKA tool was used for input assessment.	Deep learning model is found to be prominent in all. The result are as follows: Accuracy: 85% Precision: 80% Recall: 100% F-Measure: 88.9%
24	Nakamura et al. (2018) [119]	This study took the participants from different fields or departments. The participants' data were evaluated using the HDRS. They recorded the voices which were spoken by participants by just reading the phrase.	The recorded voices of participants were analyzed by openSMILE software. The Hamilton Depression Rating Scale (HAM-D) was used to evaluate the level of depression. They monitored it through MDD.	Estimation index has a high correlation of 0.832 with HAM-D score.
25	Nakamura and Mitsukura (2018) [120]	The EEG was recorded for the depression detection. Two EEG recordings were used for comparison purposes. They used the GRID-Hamilton Depression Rating Scale (GRID-HAMAD) rating scale.	The researchers tried to use the depression EEG and further analyzed it on GRID-HAMD rating scale. These comparisons were performed with the help of a two-sample T-Test. The frequency was converted by using Fourier transform.	The researchers continuously notice the EEG for healthy and depressed patients. By applying two independent T-sample tests they found the difference.
26	Song et al. (2018) [121]	This paper proposed that human behavior can be used for detecting depression automatically. They used CNN and SVM for feature selection.	They used the database of distress analysis interview which was provided by AVEC-2016. They used head pose, gaze direction, and emotion. The PHQ-8 was used for training and testing. They used the SVM and CNN as the classifiers and Support Vector Regressor (SVR) as the regressor.	**SVR Results:** The MAE is 4.37 and RMSE is 5.84 with the combination of Au+ Gaze+ HP as the modalities for gender independent. The value of MAE is 4.39 and RMSE is 5.75 for SVR in female under gender specific estimation result for the Au+ Gaze+ HP as the modalities.
27	Li et al. (2018) [122]	They performed detection by speech recognition. A deep model was presented. They did the voice extraction with the help of Mel-Frequency Cepstral Coefficients (MFCCs). Then they acquire the data from the Multiscale Audio Delta Normalization algorithm.	This study was performed on AVEC-2014 dataset. Multiscale Audio Delta Normalization (MADN) algorithm was used for local voices. They predicted the BDI scale for automatically detecting the depression. MFCCs was used for feature extraction and they proposed that deep learning was better for detecting the depression.	They achieved the MAE as 7.07 and RMSE as 9.15.
28	Yang et al. (2018) [123]	They considerd PHQ-8 questionnaire and applied a multi-model framework on it. A SVM classifier is used for classification. A new visual description Histogram of Displacement Range (HDR) was used.	This paper used the AVEC-2016 dataset for detecting the depression with the text-based data. OpenSMILE tool kit was used for the audio features. The DCNN and DCN were used for depression estimation. They used a single modal and multimodal framework for the depression estimation.	The hybrid model shows the F1 score up to 74.6%.
29	Maridaki et al. (2018) [124]	They used Gabor Motion History Image (GMHI) and Motion History Image (MHI). They extracted the	Further, Histogram of oriented Gradients was applied for edge information. PCA was applied to	With SVM Classifier they achieved the F1 score of 81.93% for both GMHI and MHI.

S.No.	Reference	Methods	Approach	Results
		combinations of appearances and applied SVM classifier, histogram of oriented gradients algorithm was applied for the edge information. They used the KNN classifier in both MHI and GMHI.	remove undesired features. They applied RF, RNN, SVM, and NB algorithm of machine learning.	
30	Oyong et al. (2018) [125]	They considered tweeter tweets to identify the depression symptoms by using NLP. They included major depressive disorder. They rated the input on the scale suggested by the Center for Epidemiologic Studies Depression scale.	For detecting depressive symptoms and the MDD they considered the time of 2 weeks to 60 days according to symptoms of depression. The data collected from the tweets was refined using Python. The other parts of text were also filtered by using InaNLP features.	**Standard Text Result at threshold 5:** Accuracy: 67% Precision: 32% Sensitivity: 89% Specificity: 63% F1 score: :47% **Twitter text evaluation score at threshold 0.8:** Accuracy: 85% Precision: 57% Sensitivity: 44% Specificity: 93% F1 score: 50%
31	Song et al. [126]	They proposed a method for the emotion recognition through EEG based on Dynamic Graph Convolutional Neural Network (DGCNN). They applied it on 15 subjects of Shanghai Jiao Tong University (SJTU) emotion EEG dataset (SEED) and 23 subjects of DREAMER dataset.	They created the graph and spectral graph filtering. The DGCNN model for the recognition of EEG emotions. Further the back propagation method was used to optimize the network parameters. This method was applied to both SEED and DREAMER databases. They found the better results in proposed method as compared to state-of-art method.	On SEED database they achieved, accuracy of 90.4%. The accuracy of 79.95% was achieved for subject dependent independent cross validation. Whereas, the accuracy of 86.23% was achieved in case of DREAMER database for valence, 84.54% for arousal and 85.02% for the dominance classification.
32	Cong et al. (2018) [127]	They used the dataset of social networking sites and applied XGBoost-Attention-Bidirectional LSTM (X-A-BiLSTM) in two different components. One is to reduce the imbalance of data and the second one helped to enhance the classification.	This paper applied the approach of deep learning in detecting depression. They chose the dataset from Reddit Self-reported Depression diagnosis.	By applying the X-A-BiLSTM deep learning they achieved: Precision: 0.69 Recall:0.53 F1-score:0.60
33	Martínez-Borba et al. (2018) [128]	This paper was based on the use of information communication technology in detecting depression among prenatal patients. They used the telephone and internet as the basic tools. They used the Edinburgh Postnatal Depression Scale (EPDS) for measuring the depression.	They aimed to identify the depression among patients with the help of phones and the internet. They recorded their activities and measure them on the scale of the EPDS and the postpartum depression screening scale.	They used ICT as a tool for recoding behavior.
34	Eichstaedt et al. (2018) [129]	They used the electronic medical facility to store medical records. This Electronic Medical record (EMR) stored the values of happiness, sadness, mood swings, etc. They used the python-based analysis. Further, they applied machine learning models for the predictions	This EMR was further comprised of various things under which they extracted the features based on the recorded dataset. The prediction model and LR were applied as the machine learning algorithms as the information communication technology tools.	AUC of 0.69 was achieved.
35	Shah et al. (2018) [130]	This paper did a study on older adults with the help of audio and computer-based therapies. They increase the depression treatment options among patients with the help of technological interventions. They used the Geriatric Depression Scale and the Hamilton scale	A very common technology method i.e., audio and video-based are used here for the detection. A small questionnaire was also prepared for the same, with the help of which it was to make an analysis. The Liner mixed approach of modeling is used for estimating the missing values.	They found these audio, as well as video behavioural therapies for cognition, were better to use for older adults.

S.No.	Reference	Methods	Approach	Results
36	Tan et al. (2018) [131]	They applied Computer Adaptive Testing (CAT) on the 1,135 participants taken as samples for depression. They used the model of Graded response.	This study divides the sample data into two formats: one is for construction and the other is for simulation purposes. They did data analysis according to it. For the different validity and predictivity they applied different criteria for depression. The Graded response model was used for the samples.	CAT depression Sensitivity: 93.7% Specificity:85.1%
37	Basu et al. (2018) [132]	This study used the Portable Personality Recognizer (PPR) for accessing the individual state. They used the hidden Markov Model (HMM) with different states. SVM was used for classification.	This paper recorded the expression by using PPR. They used cameras, LCD, and sensors. Different states of emotion were recorded which were further classified by SVM.	Achieved the emotion classification accuracy of 87.14% and personality recognition accuracy as 87.87%.
38	Jaiswal et al. (2019) [133]	This paper used different questionnaires like PHQ-9, Big Five Inventory (BFI), or Generalized Anxiety Disorder (GAD-7). They applied Correlation feature selection for finalizing the images. They calculated the Z-score and finalized the relevant feature. They applied DNN as a regression model on the score of PHQ-9 and GAD-7.	The researchers conducted various Questionnaire database and they applied PHQ-9, GAD-7, and BFI for the interview purposes of selected databases. They prepared a Virtual Human Questionnaire (VHQ). The analysis of this questionnaire was followed by the regression model i.e., DNN.	The combination of spectral features from GAD7 interview videos shows the highest Pearson Correlation Coefficient (PCC) of 0.59 and lowest RMSE and MAE of 3.64 and 2.73.
39	Gerych et al. (2019) [134]	They used the PHQ-9 and applied machine learning algorithms. They used the autoencoder and further SVM algorithm was applied. This SVM used the One Class (OC) method of its algorithm. ANN was used for autoencoder system.	This study took the data of student life with the classification of PHQ-9. They used ANN based autoencoders to classify the depressed and non-depressed classes. They used the one-class SVM (OC-SVM) algorithms. Comparison was performed with RF, MLP, and GB classifier methods.	Their method OC-SVM with autoencoder for feature projection shows the AUC-ROC of 0.92 and weighted F1 of 0.91.
40	Ma and Wang (2019) [135]	This paper was based on data collected from Twitter with the help of text mining. They tried to detect the depression symptoms from it with the help of extraction algorithm that is Rapid Automatic Keyword Extraction (RAKE). Further, they applied the Automatic Extract Keyword algorithm (AEKA) for extracting the specified terms.	This paper did a study on data collected from the tweeter and applied automatic extract keyword algorithm. They applied the values of RAKE and AEKA algorithm for creating the semantic graph.	In this paper researchers tried to create a result for an ordinary as well as for expert people so that they can get the idea of depression.
41	Lam et al. (2019) [136]	This paper is the study about identify the new approach for detecting depression automatically. They used Naïve Topic modeling and topic-modeling for patient's text and audio data. In continuation of it they used fine tuner for transforming.	This paper focused on the approaches used for detecting depression. They performed context-aware analysis and context-independent analysis. Modeling of deep one-dimensional CNN was applied so that they can do modelling of the acoustic feature. Finally, they applied the fusion of both because text, audio & video were used.	Multi-modal Model: Feedforward Network for the combination of Full Transcript + CNN-Augm shows the F1 score of 0.87, precision of 0.91 and recall of 0.83.
42	Kwon et al. (2019) [137]	This paper extracted some features and wants to train them automatically for performing pre-screening of depression. Two different deep learning Model i.e., was used. One is CNN and the other is the VGG16 model.	They used the deep learning model approach for performing the pre-screening of patients with the help of an asymmetry-based image of the prefrontal brain.	CNN shows the accuracy of 75% and higher accuracy of 87.5% in VGG16.
43	Guo et al. (2019) [138]	This paper collected data from the questionnaire round. They asked a few questions from participants and noted	The questionnaires used in this study were from PHQ-9 and International New psychiatric	Depression recognition gave the confidence interval of 95% in gender difference.

S.No.	Reference	Methods	Approach	Results
		their expressions in terms of the facial feature point and facial action units. They collected the data through Kinect.	Interview. The observations they got were recorded camera named Kinect.	
44	De Melo et al. (2019) [139]	The methods adopted for this study based on the architecture of deep learning which can predict the depression accurately. They used facial images for depression detection. They used the Multi-Task Cascade Convolution Networks (MTCNN) for facial images.	The researchers used AVEC-2013 and AVEC-2014 dataset. They extracted the facial features from the dataset and applied CNN and MTCNN as they are performing on both audio and video parts. Moving to the next part they compared the MAE and RMSE ratio.	Their proposed method in combination with the expectation loss shows the RMSE of 8.25 and MAE of 6.30 for AVEC-2013 dataset. Whereas the RMSE of 8.23 and MAE of 6.15 for AVEC-2014 dataset was achieved.
45	Pan et al. (2019) [140]	They discussed regarding the depression identification by using eye-movement and on that eye-movement they noticed the reaction time of the patient. After performing this they get the features individually. The features were fused and used for classification. The SVM classifier was used by them.	This paper proposed experimental base research of image cognition. They recorded the eye-movement with its reaction time; corresponding to this they extracted the features.	Their proposed method shows the accuracy of 86%.
46	Fang et al. (2019) [141]	This paper did a study on university students which they found are prominent to depression. They used two questionnaires as tools named PHQ-9 and the other is Zung Self-Rating Depression Scale (SDS). Other scale-like generalized anxiety disorder and also self-rating anxiety scale.	The models they applied were KNN, RF, SVM, LR, and LDA.	By selecting 8 features among all RF shows the best accuracy of 91.58%, sensitivity as 86.05%, specificity value of 96.15% and AUC of 0.8994.
47	Yuan and Wang (2019) [142]	This paper collected the data from 18-60 years old patients of a normal domain as well as of depressed domain. They did the extraction of data and further perform the transformation of data according to the eye-tracking trajectory. Then they applied ANN.	This paper simply collects the data and then applied the feature extraction process which they got from eye movement. The extracted data was converted to 2 dimensional from 3 dimensional. Further they applied ANN with the help of Python-based Keras.	Experiment shows the accuracy of 83.17% with less-neurons.
48	Tadesse et al. (2019) [143]	This study used the data collected from online users. They used NLP for data pre-processing. They applied Tokenization and stemming for removing the unnecessary part.	They used various algorithms of machine learning classifiers like SVM, RF, AdaBoost, LR, and MLP Classifier.	MLP shows the highest accuracy of 91% and F1 score of 0.93. SVM gave the accuracy of 80% and F1 score of 0.80.
49	Sumali et al. (2019) [144]	This paper used the features of landmarks, movement of eye pupils, and feature near to mouth area. They used the data obtaining from the clinical interview. They did a statistical comparison and feature selection. Further they applied machine learning on the selected feature.	This paper took the clinical interview. The pre-processed data was used for feature extraction. They applied a two-sample t-test. Further filter, wrapper methods, and Embedded methods were used in feature selection. They applied SVM as the machine learning algorithm.	They found a significant difference between depressed and healthy patients. The machine learning algorithms were given a better result in identifying the levels of depression.
50	Kim and Lim (2019) [145]	They used the HRV and a new-fuzzy algorithm for detecting the depression. The HRV was further divided by the Multimodal Affective Content (MAC). Non-overlap area distribution for the feature selection method was used and a neuro-fuzzy algorithm was used for identifies the HRV feature based on the content. A supervised machine learning algorithm was used	This paper was used for determining the optimal contents for analyzing the depression. The HRV division they got 22 features with the help of frequency domain and Poincare transforms methods. They applied a non-overlap area for the feature selection method.	Mean accuracy achieved was 86.6% and maximum accuracy achieved was 86.9%
51	Zhou et al. (2019) [146]	This paper studies the facial expression of the videos. They took the content from the adaptive feature by extracting. They used built-it on CNN and regression network to learn	The researchers applied the representations taken from facial expressions of videos. They gathered it from CNN and applied regression model on the	Proposed method shows the MAE of 6.37 and RMSE of 8.43.

S.No.	Reference	Methods	Approach	Results
		discriminative representation of depression.	depression feature of content adaptive.	
52	Niu et al. (2019) [147]	The researchers found that facial expressions are a better modality to identify depression detection. They proposed a Local Second-Order Gradient Cross Pattern (LSOGCP) to identify the facial dynamics. Once they extracted the features, they applied different classification and regression methods.	They extracted the images from the videos and applied the Gradient Cross Pattern. They found it gave good accuracy results. From the images extracted they applied three orthogonal planes. For calculating the combined result, they adopted the hierarchical method by combing the information between the groups and within-group regression values.	The proposed model shows the RMSE of 9.17 and MAE of 6.97 in AVEC-2103 dataset. Whereas the RMSE of 9.10 and MAE of 7.19 in AVEC-2014 in the BDI-II score prediction.
53	Al-Naggar et al. (2019) [148]	This paper tried to detect depressive symptoms at an early stage. They divided the approach of scanning the brain two halves: the first one is about structural analysis and the other is about functional analysis.	The approach followed in the study to work on the data received from MRI. With the help of this they tried to identify MDD among the patients. Initially they applied functional connectivity and functional connectivity after doing MRI.	This paper helped in proving the early depression detection problem. They suggested using some machine learning algorithms as the tool of MRI for upcoming generations.
54	Kumar and Subha (2019) [149]	They discussed regarding the prediction of depression symptoms from the EEG signal. They used the recorded data and performed feature extraction on it. They used deep models of learning from LSTM. Further they implemented the RNN in their LSTM model.	While implementing the LSTM they used the RNN. Their proposal for using LSTM was further compared with the Convolution LSTM (ConvLSTM).	LSTM shows the RMSE of 0.005 whereas in case of CNN-LSTM it is 0.007 and in case of ConvLSTM it is 0.0088. It means LSTM can predict the next instants of depression better than others.
55	Hao et al. (2019) [150]	This paper used the data of 1425 studies of questionnaires from different places. They applied different scales on the collected data. The scales used by the research are like a self-rating scale for depression or anxiety. A cluster analysis was performed on the collected data and various observations were counted.	Two questionnaires were prepared one for anxiety and the other is for depression. They applied the self-rating scale to these questionnaires. Further after applying the Gaussian mixture model, they got the difference value. With the help of the hierarchical regression model, they were easily able to perform the statistical analysis.	Their results suggested that the risk of depression can be determined or reduced through self-study.
56	Syarif et al. (2019) [151]	Social media data was collected for evaluation purposes. They used a self-declared corpus used for identifying the mental illness of a person. They used two categories of features one is linguistic-style and the other is expression-sentiment. The data taken from tweeter was pre-processed with python package. The natural language toolkit was applied to the sentiment feature extraction.	They used the language filter and post-activity filter for making the diagnosed group. They used the natural language toolkit and python package for pre-processing the data. Further they classified the data with the help of machine learning methods. They applied various clustering rules with parameters supporting the attributes.	They successfully divide the tweets into different levels of depressions (Low, Moderate, and High). Further they suggested exploring it more related to other features.
57	Liang et al. (2019) [152]	Their study aimed to present low virtual reality which may influence mental illness issues. They performed the experimental setup in the laboratory. A demographic questionnaire was used by them. They implemented it on intelligent Virtual Reality (VR).	They used the filled questionnaire on Intelligence VR where they calculated the alpha and bets, frequency bands. They applied the correlation analysis and compared the values.	There result showed that VR simulation is available for detecting the depression with the help of EEG signals. They classified 55 subjects at high risk of depression by using (Symptom's checklist) SCL-90. They found the higher Beta frequency for both right and left forehead as p=0.036, P=0.033 than the non-depression group.
58	Ware et al. (2020) [153]	They used the PHQ-9 and Quick Inventory of Depressive Symptomatology questionnaire. They	This study tried to identify the early detection features of depression with the help of the use	They found that smartphone data can be useful to predict the symptoms in case of behavioural

S.No.	Reference	Methods	Approach	Results
		extracted the features from the Density-based Spatial Clustering of Applications with Noise (DBSCAN) method. As a classification method they used the SVM with Radial Basis Function Kernel-Support Vector Machine Kernel (RBF-SVM).	of smartphone features. Two different questionnaires i.e., patient health and Quick Inventory of depressive type of questionnaires were used. They used the data of smartphones and WIFI metadata. The depressive symptoms they collected were further classified with the help of SVM.	as well as cognitive scases. They obtained the F1-score of 0.86.
59	Gadassi et al. (2020) [154]	This study divided future concerns into positive and negative aspects of occurrences. They used the Structured Clinical Interview for Disorders (SCID) questionnaire and divided it into two sections.	They used the data of two sessions recorded by SCID. It was used for ranking purpose.	They did a comparison in objective and subjective ratings. They analyzed based on occurrences and did a comparison with self and one another.

4.2 RQ2: What are the major advantages and limitations of the approaches used in the mental illness and depression disorder?

Based on the analysis of the studies conducted in 2015–2021, the following advantages and limitations were encountered in the approaches used for the detection of mental illness and depression disorders.

Advantages of the machine intelligence approaches

1. Approaches such as SAS are helpful in the monitoring of depression and anxiety scores.
2. SVM is useful when applied using the combined feature set (sensor-level feature set and source-level feature set).
3. The combination of CNN and RNN can be used to extract spatiotemporal information.
4. The preprocessing method is helpful in boosting the performance of the algorithms.
5. RF, NB, and SVM are helpful in the timely identification of various mental illnesses.
6. Hybrid models such as CNN with LSTM are more efficient and accurate in depression detection.
7. Combinations of machine learning and CNN algorithms can be employed for quick diagnosis.
8. The obtained automated scores can help in the quantitative assessment of therapeutic effects.
9. Machine learning frameworks can be used to supplement information from social networking platforms and can be scaled for the identification of social users.
10. SVM, multivariate regression, and RBFN perform better depression prediction compared to the volume-based features, that is, user-based features perform better at predicting the PHQ-9 scores and depression status.
11. DNN accomplishes more accurate learning based on the learning rate.
12. Genetic algorithm can be applied for feature weighting. KNN may be helpful for the fusion modality of negative and positive scenarios.
13. SVM and LR can be used for determining the missing values. New prospective can be visualized by using boosting algorithms with machine learning.
14. XGBoost algorithm can use the predictive power of biomarkers data, which can help in the diagnosis of depression.
15. Several psychological stresses can be detected using machine learning. Machine learning models can help achieve superior performance in the prediction of suicide ideation.
16. Relevant information can be stored; short-term and long term risks can be identified easily.
17. Medical check-up data were helpful in utilizing the results for direct measurements. Some of the data used in the study included the information taken from diagnosis and prescribed drugs.
18. This study will help increase awareness about educational and social policies.
19. Machine learning can be used with PHQ-9 and other combined features for screening depression. It can be used as an effective tool for enhancing the predictive accuracy of MDD. Composite measures showed the improvement in accuracy for detecting MDD.
20. DT can be used for information acquisition.
21. The sensor information is collected only during typing with a keyboard. Thus, no battery problem occurs. There is no need for a long period of time to reach a definitive decision.
22. Machine learning techniques are capable of assessing suicide risk.
23. Machine learning techniques can be applied to model the severity of PSU.
24. This study will help in raising clinical awareness regarding depression.

25. Machine learning and boosting algorithms are helpful in the extraction of sociodemographic and psychological factors responsible for causing depression.
26. This study showed the need for an enhanced bootstrap sampling approach with ensemble tree algorithms.
27. Valuable and discriminative features for depression analysis can be determined using the deep local global attention CNN. It is also efficient in describing the discriminated patterns.
28. The SL-based optimal algorithm can be used in the prediction of anxiety and depression.
29. A voice change study in depression was covered by us, which suggests the possibility of detecting depression through voice.
30. M3 was found to be an effective tool for the detection of schizophrenia.
31. Personalized evaluations and interventions may have potential applications in the clinical field.
32. The grid search technique can be used for optimizing the parameters.
33. LASSO, a regression analysis method, is helpful in feature reduction.
34. The depression data based on kinematic skeletal can be used for performing depression analysis.

Limitations of the discussed approaches:
1. The size of the data was small. The availability of the data was also limited.
2. The cross-sectional design was used in some of the studies, which is not the best way to design the causal relations. All used scales were self-rated.
3. In case of medication, the confounding effects cannot be controlled.
4. There is a need for exploration of the impact of genetic markers.
5. The effectiveness of RNN-C3D on human behavior problems should be checked.
6. Default parameters were used in the classifiers. More effectiveness can be achieved by altering the tuning parameters.
7. Number of subjects are required from diverse background. Complexities can be reduced by using graphical processing units (GPUs).
8. More pathological EEG signals are required. There was no clarity in the effect of acceleration of comorbidities.
9. The number of samples must be increased. Optimal efficacy can be achieved by increasing the number of samples. The dynamic change pattern may be helpful.

10. Some of the studies considered the single category of intention for each status update. Additional studies from medical professionals are required.
11. The assistance process can be improved by using metaheuristic optimized techniques.
12. The time required must be optimized as the time taken by the hybrid system is found to be more than that taken by the single system.
13. Balanced proportion of male and female subjects are required, and different populations must be considered.
14. External and internal influence factors must be determined. Personalized health management is needed for classifying the context.
15. Machine learning algorithms must be combined with the traditional approaches.
16. Inadequate information was provided by parents in the psychosocial treatment for ADHD. Outcome variable used, shared the common method as the critical predictor variable. A strategy is required to narrow down the redundancy of predictor variables.
17. In some cases, the data were restricted to only a single clinical setting. A limited number of suicidal events were considered. Data was collected from a single institution; thus, it may be biased. Patients' clinical notes were not available. De-identification is more complex in clinical text than in the structured EHRs data. Chances of biasedness to commercially insured patients.
18. The dynamic causal modeling approach was restricted to small networks.
19. In some of the studies, as the important factors were recorded by the parent, the reporting could be biased. The performance of the model can be improved by considering a larger sample size.
20. Some of the studies considered the effectiveness for sadness only; thus, results cannot be compared with other overall symptoms of depression. It is difficult to determine the causality because the study was cross-sectional.
21. Some of the models did not consider variables such as stress and negative experiences of childhood. Individual mental illness was not considered, which can be the reason for depression.
22. Different timeframes were considered in different studies, preventing the establishment of the causal relationship.
23. In some of the cases, there are chances of error and bias as the questionnaires were completed by

24. The rank of the features can be improved using motion and position sensors and can be extended using feature extraction algorithms and devices with different internal sensors.
25. In some of the cases, the participants were from a single country. Some studies considered only the measures related to self-reporting, whereas the interview methodology should be employed. Smart phone assessments were found more reliable than self-reporting. As it was a cross-sectional study, causal inferences may be unreliable.
26. The use of class probabilities made the result hard to interpret. Large classes can be chosen for the selection of latent class growth analysis (LCGA) model. SL included the complex interaction which makes it black-box; thus, it was hard to be interpreted.
27. The degree of anxiety may have meditated the change of voice to a greater extent than the depression. It is unable to confirm the relationship between the severity of depression and voice features.
28. In a study, the cerebellum was not covered during resting state fMRI (rsfMRI) scanning. The head motion of patients with schizophrenia was greater than that of healthy controls. Their study did not account for the head motion.
29. In some cases, the available data did not reflect the interpersonal relationships of respondents.
30. Some of the studies did not consider the depression history for the depressive group. The HDRS-24 scale does not reflect the complex and rigorous diagnosis of depression. The study was limited to integrate various forecasting capabilities. The kinematic features do not completely provide the proofs for the reflection of participants suffering from depression.

4.3 RQ3: What are the most commonly used datasets in the same domain?

Datasets used in the literature for the review has been discussed in this section along with the nature, participants, sample size, attributes and source.

Students of Süleyman Demirel University [41]

They considered the four hundred students to participate in this study. Out of 400, 52 students refused to participate and twenty-nine were having a low scale score. So, in total 319 students (203 females and 116 males) were considered from the Suleyman Demirel University. They divided the students into three groups; a smartphone non-user group (71 users), a low smartphone user group (121 users) and high smartphone user group (127 users). They assessed the participants who were smartphone users by using PSQI, BDI and BAI. In the smart phone non-user group, participants were assessed on the scale of SAS.

Psychiatry Department of Inje University Ilsan Paik Hospital [42]

This study considered the 34 patients (20 males and 14 females) of Schizophrenia and 34 patients Healthy Controls (HC) (14 males and 20 females). They recruited this data from the Psychiatry Department of Inje University Ilsan Paik Hospital. They diagnosed the people based on the Structural Clinical Interview for the Diagnostic and Statistical Manual of mental Disorder (DSM-IV). They measured the Clinical symptoms by using the Positive and Negative Syndrome Scale (PANSS).

At Laboratrio de Biologia Molecular do Centro de Oncohematologia Peditrica of the Hospital Oswaldo Cruz at University of Pernambuco (CEONHPE/HUOC/UPE) [43]

They considered 151 individuals from which 126 with the Mild Cognitive Impairment (MCI) in which 70 with MCI and 56 with Dementia and individual controls were 25. The data were collected from Laboratrio de Biologia Molecular do Centro de Oncohematologia Peditrica of the Hospital Oswaldo Cruz at University of Pernambuco (CEONHPE/HUOC/UPE)

AVEC-13 and AVEC-14 [44]

AVEC-13: This dataset considered 82 subjects which contains 150 videos having range between 18 to 63 years. The complete dataset was splitted into three parts having 50 videos in each of training, development and testing phase. The videos were having a resolution of 640×480 with 24-bit color and the sampling rate was 30 Hz. The BDI-II was estimated for each video.

AVEC-14: This dataset used "Freeform" recorded sessions for the evaluation. A total of 150 videos was used in this dataset. It was divided in three parts of training, development and testing having 50 videos in each. The average length of videos collected in AVEC-14 was 2 minutes. One- and four-times recording was performed for the participants with an interval of two weeks.

Office of the Special Consulting Health Services, of Patras University [45]

The selected the group of 249 patients from the office of special consulting Health services from the Patras University. The age considered was from 18-30 years. The BDI questionnaire was used for the assessment.

Python Reddit API Wrapper (PRAW) [46]
The data were downloaded by the use of python (API) for the Reddit, The Python Reddit API Wrapper (PRAW). They downloaded the 1000 posts for each category. A total of 3922 posts were used for the study.

Psychiatry Department, Medical College, Calicut, Kerala, India [47]
This study considered the 15-15 subjects of each class of depressed and normal. The total data considered for the study was 9116. They used the data of 4798 for depressed patients having record as 2398 for right hemisphere and 2400 for left hemisphere. Similarly, the data of 4318 were used for normal patients divided into 2159 for right hemisphere and 2159 for the left hemisphere.

Lanzhou University Second Hospital [48]
They recruited the 28- subjects at Lanzhou University Second School. The subjects were divided as 14 patients of depression and 14 patients of normal subjects.

MDD patients from Jena University hospital and MRI scans of IXI and OASIS [49]
They calculate the BrainAGE on 743 healthy controls (males: 316 and female: 427) on the dataset of IXI and OASIS.
IXI: https://brain-development.org/ixi-dataset/
OASIS: https://www.oasis-brains.org/
For DSM-IV Criteria: They recruited the 38 patients of major depression and 40 subjects of healthy control.

Affiliated Hospital of Chengdu University of Traditional Chinese Medicine (TCM) and Sichuan Provincial People's Hospital [50]
The recruited the 14 depressed patients from the Affliated Hospital of Chengdu University of Traditional Chinese Medicine (TCM) and Sichuan Provincial People's Hospital and 15 normal participants from the community and the University of Electronic Science and Technology of China. The average of depressed patients was 47._+10.26 years and normal participants was 41.47_+11.67 years.

myPersonality [51]
The dataset 'myPersonality' a Facebook Application was used for the study. It contained the approx. 6,000,000 responses. The metrics considered for a social network were NumStatus, AvgLength, Insomnialdx, Betweenness, Density, Transitivity, Brokerage and Network size.

collaborative psychiatric epidemiology survey (CPES) [52]
The data for the study was collected from the CPES. They found some missing values in the gathered data so the RF approaches was applied for the elimination.

APIs of Twitter, Sina Weibo (available in China), Facebook, Online sources like microblogs, web fora, crowd-sourcing, Amazon's Mechanical Turk (MTirk) and online communities

Depression Anxiety Stress Scale (DASS42 and DASS21) [54]
They used two DASS21 and DASS42 as the dataset. DASS42 consists of 42 questions. They collected the 39,776 instances between the years 2017 to 2019 through the questionnaires. Another dataset DASS21 was used constructed with the help of Google forms. It was completed by 349 participants whose age was between 18 years to 60 years. The DASS21 had 21 items and each category of stress, anxiety and depression had the seven questions.

University of Connecticut [55]
The data were collected from the total of 79 participants from October 2015 to May 2016. The students were full time from the University of Connecticut. The age considered between 18 to 25 years. 73.9% were female and 26.1% were male participants. Each participant used smartphones either iOS or Android to participate in this study. The data were collected from four types of sources, network traffic data, Patient Health Questionnaire-9 (PHQ-9), Clinical Assessments and Screen On/OFF events. Only Android phone users have the accessibility of screen ON/OFF events. These data were collected from clinical assessments as well as scree on/off events.

Korea National Health and Nutrition Examination Survey at the Centers for Disease Control and Prevention Korea [56]
They have considered the raw data from the Korea National Health and Nutrition Survey. It consists of 600 variables of examination health questions and nutrition. Examination survey includes the index of Obesity, Hyper Tension, Diabetes, dyslipidaemia, lung disease, liver disease, kidney disease, anaemia, oral disease, eyesight, grip and heavy metals.

National Key Research and Development Program [57]
They identified the participants by using the criteria of screening for 973 national Key Research and development Program. The dataset of 86 depressed patients and 92 normal controls were created. They included the age groups between 18 years to 55 years for both the healthy and depressed groups. The score greater or equal to 5 was considered for the depressed patients. The patients should be of higher education level. The depressed patients were not allowed to be on medication from last two weeks before this trial.

China Health and Retirement Longitudinal Survey (CHARLS) [58]

The data were used between year 2013 to year 2015 from the China Health and Retirement Longitudinal Survey (CHARLS). They took the sample size of 29,477 from the combination of CHARLS 2013 and 2015. The age considered was 45 and above.

Lifelines Cohort study database [59]

They considered the dataset of 11,081 as the samples from Dutch citizen. In this study 4,587 were male and 6,494 were female subjects. Out of 11,081 cases the 570 cases were of depression with self-reported. They divided the dataset under three categories of train, validation and test dataset.

https://www.lifelines.nl/researcher.**National Survey of Children's Health [60]**

The data of 6630 participants were considered for this study was taken from 2016 to 2017 by participants of the National Survey of Children's Health (NSCH). The age considered for youth was taken from 3 to 17 years and it has the 68.5% of male participants.

https://childhealthdata.org/learn/NSCH/resources/survey-instruments

Cardiorespiratory fitness and hospitalization events in the armed forces (CHIEF) study performed in the Hualien Armed Forces General Hospital, the military referral hospital of Eastern Taiwan [61]

This study used history of cohort of 3,546 of military personnel from the study performed in Hualien Armed Forces General Hospital of Eastern Taiwan. The age considered for the study was from 18 years to 50 years. The Brief Symptom Rating Scale (BSRS-5) was used as the questionnaire for the participants where they used the five domains like depression, anxiety, hostility, insomnia and interpersonal sensitivity. They used the Likert type of five-point scale from 0 to 4.

Connecticut Children's Medical Center [62]

This study considered the data from Connecticut Children's Medical Center (CCMC) Electronic health records (EHRs) database. It contains the record of 129,485 patients who had the visits of 641,708 from October 1, 2011 to September 30,2016. The age of patients lies between 10 to 18 years for the first attempt of suicide. They included the record 41,721 as young patients for the analysis purpose. Out of 41,721, there were 180 patients who have attempted the suicide and labelled as positive subjects and 41,541 were the patients who haven't attempted the suicide and labelled as negative subject.

DASS21 [63]

They used the DASS 21 questionnaire. This study was conducted on total 348 participants (both male and female) whose age was between 20 years to 60 years. This DASS 21 questionnaire consists of 21 questions where 7 questions were assigned to the individual scale of depression, stress and anxiety.

functional magnetic resonance imaging (fMRI) data NEtherlands Study of Depression and Anxiety (NESDA) [64]

They acquired the data of 85 patients from a Netherlands Study of Depression and Anxiety (NESDA). Total 2981 participants of age 18 years to 65 years were recruited from the different primary care, specialized organizations of mental health and communities. They included only those participants who were diagnosed with MDD, have reported symptoms, having two years of follow up for depression symptoms. So, final 85 patients were considered after performing the screening on 301 participants, which were the part of MRI experiment and in which 156 had the MDD diagnosis.

https://www.nesda.nl/nesda-english/
https://gitlab.ethz.ch/tnu/code/generative-embedding-nesda

Child and Adolescent Twin Study in Sweden (CATSS) [65]

The participants were considered from Child and Adolescent Twin Study in Sweden (CATSS). It contains the 15,156 twin pairs which were born in Sweden. They found that total participants, 7,638 were born and whose age was between the years 1994 to 1999. The age considered to be from two waves i.e., one wave was from 9 years or 12 years and second wave considered the age of 15 years. The age of second wave considered the outcome measures taken from Strengths and Difficulties Questionnaire (SDQ). Initially, they included the 474 variables which were included in the dataset.

All live births by the Australian state of Queensland (QLD) [66]

The data were taken from all live births linked with administrative of health from the Queensland (QLD) in the year from January 2009 to December 2015. The sample gathered consists of 435,912 babies and samples of 306,799 of the mothers.

Collaborative Psychiatric Epidemiology Survey (CPES) through the Interuniversity Consortium for Political and Social Research (ICPSR) [67]

The data were considered from the CPES. The CPES had total 20,013 values out of which only 1162 participants were found who took one of the drugs from the seven categories of mentioned drugs. Out of these 1162 participants they found the 814

instances in which the drugs were taken for sadness. The study was carried out from 2001 to 2003.

National Health Insurance Sharing Service (2019) in Korea [68]

Initially the data of 514,866 were collected from year 2002 to year 2015 for those who were having a national check-up. There were several participants who were lost before 2009 so they were excluded and the 494,192 participants were selected thereafter from year 2009 to the year 2015. After that, they excluded those participants who were without medical check-up so selected 449,685 participants with medical check-up. Further, after applying the exclusion criteria they selected the total 433,190 participants. Out of 433,190 participants they include the 10,824 participants in the depression group.

Michael Corley in 2018 using the Survey Monkey tool [69]

The data used in this study followed CRISP-DM methodology. The categorized the data into six phases like Data preparation, Data Understanding, Business Understanding, Evaluation, Deployment and Modelling. The RapidMiner tool was applied for the study.

European Cyberbullying Intervention Project Questionnaire [70]

The 1328 students from a Spanish university took participation in this study whose age was between 18 years to 46 years. They excluded the 46 participants due to some errors in their responses. Finally, they considered the 1282 students having 688 women and 594 men.

Data collected from android mobile application [71]

Dataset collected from the motion sensors and touch screen panels. It determines the stressful and non-stressful situation based on the keyboard typing behaviour. Accelerometer sensors measure the acceleration applied to the device and Gyroscope sensors gives the angular velocity of the smart phone. The collection was based on four stages. These are non-stressful state, stressor task, case of stress and ground truth survey.

Dashboard created to monitor and analyse (https://bit.ly/penncovidmap) [73]

They collected the two datasets which contains approximately 5 million tweets per day. These datasets contain 1% as the random sample for US tweets and second datasets of tweets which had the keywords related to Covid-19. They considered the four measures like stress, loneliness, anxiety and sentiment from 13th March to 6th May when the emergency was declared in the year 2020.

An online survey conducted in Eastern China [74]

They constructed the online survey in April 2019 at a large university in the Eastern China. Total 1238 students participated in the survey. They excluded the 141 participants for careless responding. So, finally 1097 participants participated where it had the majority of women than the men. The women were 898 and men were 199.

Collected data from Android Apps and iOS apps [75]

In this study, they identified the 183 apps from Google Play and 254 apps from app Store where they further excluded those apps which were not related to mental health and having less than five reviews. So, finally the selected the 104 apps related to mental health available on App Store and Google Play. They performed the 88125-sentiment analysis on the user reviews by using the ML techniques and then conducted the thematic analysis on reviews.

Weil Cornell Medicine (WCM) and New York City Clinical data Research Network data (NYC-CDRN) [76]

They considered the two EHR datasets. The one dataset was used from Weil Cornell Medicine (WCM) from January 2015 to June 2018. The validation of the dataset was performed from August 2014 to October 2017 from the data taken from New York City Clinical data Research Network data (NYC-CDRN). The included deliveries from WCM were 15,197 and from NYC-CDRN were 53,972.

Department of Psychiatry, the First Affiliated Hospital of Chongqing Medical University [77]

In 2021, Hong et al. [77] recruited the patients from the First Affiliated Hospital of Chongqing Medical University under the department of Psychiatry. They considered the 42 patients who attempted the suicide and 25 individuals of suicide ideation. The age considered was from 14 to 25 years. The assessment of patients was performed by the Structured Clinical Interview for DSM-IV Diagnosis (SCID). They excluded those patients who had the history of head trauma, DSM-IV Axis I comorbidities, substance or alcohol abuse, neurological disorders and any history of psychiatric disorders.

Online Survey of Bangladeshi citizens [78]

The dataset was obtained from the survey. It was performed for collecting the data from Bangladesh citizens of different age ranges. For this a questionnaire was prepared which consists of 55 questions. It consists of complex psychological and socio-demographic information. It was conducted in the period between April 2020 and August 2020.

Psychiatric inpatient unit at the Mount Sinai Health System for suicidal ideation or suicide attempt [79]
They recruited the patients from of Mount Sinai Health System taken from 10th January 2016 to 10th January 2019 who attempted for suicide. They collected the total 591 as the inpatients. Out of 591, 20 performed the follow-up within one month and others 571 were not the part of this.

Netherlands Study of Depression and Anxiety (NESDA) [81]
The data were considered from the Netherlands Study of Depression and Anxiety (NESDA). They recruited the patients from centres of primary and secondary care. The data sample used was of 2,981 where 2,329 was having lifetime diagnosis. These participants had the follow up after every, 2,4,6 and 9 years. They have the majority of female participants and mean age of total participants was 41.3 years.

Public website of Norwegian [82]
This approach was applied to two datasets. They obtained the large dataset from ung.no, it is a website of public Norwegian. In this website public can posts their questions related to problems and challenges. So, here the text based on mental health and emotions was included. In the first dataset they considered 11,807 for the experiments in which 1820 texts were of depression and another 9987 texts were of non-depressed. In the second dataset they collected the 21,470 samples of text in which 1470 texts were of depression and 20,000 texts were of non-depression.

Renmin Hospital of Wuhan University, Wuhan, China [83]
They considered the data from Renmin Hospital of Wuhan University, Wuhan China. This study was considered from March 2018 to October 2018. In all they acquired 96 samples, out of which 17 samples belong to non-depressive group and 79 samples belongs to depressed group. The PHQ-9 was used as the measurement scale.

Seoul National University, Hospital [84]
Patient population has been considered for the subjects. It has been collected from Seoul National University, Hospital for depressive symptoms. The age range considered to be between 9 years to 65 years. The inclusion criteria were read and understandability capability. The exclusion criteria were a history of brain surgery or head trauma and dementia diagnosis.

Center for Biomedical Research Excellence (COBRE) [85]
They collected the open dataset of neuroimaging from Center for Biomedical Research Excellence (COBRE). This dataset has 74 HCs and 71SZs. They collected the diagnostic information from the Structured Clinical Interview for the DSM Disorder (SCID). The exclusion criteria they used were like, neurological disorder history, severe head trauma, intellectual disability. According to this criterion, they excluded the 36 subjects. At the end they included the 109 subjects of 45 SZs and 64 HCs.
http://fcon_1000.projects.nitrc.org/indi/retro/cobre.html

Mental health department of Rambam Health Care Campus in Haifa [86]
There was total 111 participants who used Hebrew as their main language. The measure participants were women. It was approved Rambam health care campus. It covers depression, dysphoria, anxiety, panic attacks etc.

Primary data [87]
They obtained the data from medical chart which was recorded from January 2011 to December 2017. They included the participant's age from 20 years to 65 years, they must be diagnosed with panic disorder, according to the International Classification of Disease-10 (ICD-10) other anxiety disorders, measurement of HRV. Similarly, the exclusion criteria consist of any cardiovascular disease, disorders other than anxiety disorders. In all they recruited 60 patients of panic disorder and 61 patients of anxiety disorders after following the inclusion and exclusion criteria.

Data were collected from the King Fahad University Hospital, Dammam, Kingdom of Saudi Arabia (KSA)} [88]
They conducted the study in Imam Abdulrahmanbin Faisal University (IAU) under the department of computer science. The data were collected from the King Fahad University Hospital, Dammam, Kingdom of Saudi Arabia (KSA). The demographic features of patients were collected from April 30, 2020 to July 24. 2020. In total they recorded the data on the 287 patient's suffering from COVID.

Korea National Health and Nutrition Examination Survey (KNHANES) [89]
They considered the data from Korea National Health and Nutrition Examination Survey (KNHANES). The data were taken from year 2014 to the year 2016. Their data of year 2014 were considered for the training set and the data of year 2016 was considered for testing set. Total they used the data of 9488 patients. In which the data of year 2014 consisted of 7550 and the data of 8150 belongs to the year 2016.

Bursa Higher Specialization Training and Research Hospital Paediatric [90]
The recorded the 482 children as patients from the Bursa Higher Specialization Training and Research

Hospital Paediatric follow up centre between October 2012 and October 2014. They excluded the 58 children as they were having incomplete information. Further, the patients were divided into two groups, one was, according to the DSM-V of PTSD and other one was of MDD. They grouped the 287 participants as 177 participants of non-MDD and 110 participants with MDD in group 1. The second group consisted of the total 372 participants, which includes 177 as non-PTSD and 195 patients with PTSD.

Data provided by Statistics Netherlands [91]
They gathered the mental health record from the Dutch insurance act. It covers all specialist mental health treatment of Schizophrenia for all 17 million residents of the Netherland between 2010 and 2014.

First Affiliated Hospital of Zhejiang University, School of Medicine [92]
They performed the cross-sectional study on the patients November 2017 to April 2019 from the First Affiliated Hospital of Zhejiang University, School of Medicine. They considered the inclusion criteria as the patients should be diagnosed with major depressive episode according to the Mini International Neuropsychiatric Interview for DSM-IV-TR (MINI Plus 5.0), age criteria were from 16 years to 65 years, 5 years education was required, they should have the mixed features as specified by DSM-V. In total, they included the 300 participants in which 180 patients had the Bipolar, 49 had the MDD and 71 had the major depressive episodes.

earlydetect questionnaire [93]
The participants were asked to fill the earlydetect questionnaire. They performed this study on 955 participants from June 2016 to October 2018 which was approved by University of Alberta Research Board.

Shandong Mental Health center[94]
They collected the dataset of the total 170 participants, which was divided into two groups of datasets one was depression and other was non-depression. The paired them in the group of 85 patients of depression and 85 pairs of non-depressive participants. They conducted this study in the Shandong Mental Health center. They used the Kinect V2 device for the capturing of kinematic skeletal features.

Collected EEG and ECG signals [95]
They evaluated the total 36 participants for the evaluation. They conducted the two trials for each subject and capture the EEG and ECG as the physiological signals.
https://physionet.org/content/eegmat/1.0.0/.

5.Discussion
Based on the complete literature review and systematic meta-analysis the following points are indicated. It includes the methodology used, result analysis, advantages and challenges found along with the use of the datasets for the mental illness.
1. The machine learning approaches and frameworks can be used to supplement information from social networking platforms and can be scaled for the identification of social users. SVM is useful when applied using the combined feature set. SVM, multivariate regression, and RBFN perform better depression prediction compared to volume-based features. RF, NB, and SVM are helpful in the timely identification of various mental illnesses. The combination of CNN and RNN can be used to extract spatiotemporal information. Hybrid models such as CNN with LSTM are more efficient and accurate in depression detection. It is also indicated from the review and analysis that the combinations of machine learning and CNN algorithms can be employed for quick diagnosis.
2. Machine learning and boosting algorithms are helpful in the extraction of sociodemographic and psychological factors responsible for causing depression.
3. It is observed that the voice change study may be helpful in the depression detection.
4. It is found that the personalized evaluations and interventions may have potential applications in the clinical field.
5. It is observed that the default parameters were used mostly. More effectiveness can be achieved by altering the tuning parameters.
6. Some of the studies considered the single category of intention for each status update. Additional studies from medical professionals are required.
7. External and internal influence factors must be determined to enhance the diagnosis accuracy.
8. In some cases, the data were restricted to only a single clinical setting. A limited number of cases were considered. Data was collected from a single institution; thus, it may be biased. So, there are the chances of business to commercially insured patients.
9. In some of the studies, as the important factors were recorded by the parent, the reporting could be biased. The performance of the model can be improved by considering a larger sample size.
10. Some of the studies considered the effectiveness for sadness only; thus, results cannot be

compared with other overall symptoms of depression. It is difficult to determine the causality because the study was cross-sectional.

11. In some of the cases, the participants were from a single country. Some studies considered only the measures related to self-reporting, whereas the interview methodology should be employed. Smart phone assessments were found more reliable than self-reporting.

12. In some cases, the available data did not reflect the interpersonal relationships of respondents.

13. Some of the studies did not consider the depression history for the depressive group. So, it can be said that some of the studies were limited to integrate various forecasting capabilities.

Limitations
Only English language publications were considered and examined. Only machine intelligence approaches, such as machine learning, deep learning, and hybrid boosting, and machine vision methods' parameters were considered in the analysis and discussion. Papers with high clinicians' content were not considered in the review. A complete list of abbreviations is shown in *Appendix I*.

6.Conclusion
In this paper, a systematic review and meta-analysis of the mental health issues were provided using machine intelligence methods. A detailed discussion based on machine intelligence methods, way of applicability, advantages, and limitations was also presented. The methodological mapping covered the publication trends, mapping aspects, current scenarios, and a comprehensive discussion of the results. Several computational aspects based on these approaches were also considered and discussed. Our review and analysis suggested that machine intelligence approaches have potential application in the detection of mental health issues and depression in a timely manner.

1. Insufficient resources for detecting and treating patients with mental health issues are a major weakness of developing countries.

2. The gap in the treatment of mental disorders is another challenge. According to the WHO, approximately one-third of patients with schizophrenia do not receive any treatment and care.

3. No mental health policy has been formulated. There is a lack of awareness among parents, social stigma, and other external factors.

4. The association between symptoms and the internal and external factors should be determined.

5. The suitability of each method for different mental problem scenarios should be validated based on some simulation. This can help in handling real events efficiently.

6. The association between risk factors and mental health problems should be compared and analyzed for early-stage prediction. Most of the mental disorders are macro issues, but the risk associated may be high; thus, there is a need for proper intervention for the prevention and prediction of mental disorders [65].

7. There is a need for the removal or addition of symptoms that may affect the causes and the risks associated with positive and negative emotions.

References
[1] Rajkumar RP. COVID-19 and mental health: a review of the existing literature. Asian Journal of Psychiatry. 2020.

[2] Salari N, Hosseinian-far A, Jalali R, Vaisi-raygani A, Rasoulpoor S, Mohammadi M, et al. Prevalence of stress, anxiety, depression among the general population during the COVID-19 pandemic: a systematic review and meta-analysis. Globalization and Health. 2020; 16(1):1-11.

[3] Ustun G. Determining depression and related factors in a society affected by COVID-19 pandemic. The International Journal of Social Psychiatry. 2021; 67(1):54-63.

[4] Ghazali N, Zain NH, Fesol SF, Moketar NA, Odzaly EE, Teo NH. Relationship between learning habits and socioeconomic status: a COVID-19 pandemic study. International Journal of Advanced Technology and Engineering Exploration. 2021; 8(74):102-13.

[5] World Health Organization. Depression and other common mental disorders: global health estimates. World Health Organization; 2017.

[6] Abbott A. COVID's mental-health toll: how scientists are tracking a surge in depression. Nature. 2021; 590:194-5.

[7] Office for National Statistics (UK data); Centers for Disease Control and Prevention (US data), 2020

[8] Satinsky EN, Kimura T, Kiang MV, Abebe R, Cunningham S, Lee H, et al. Systematic review and meta-analysis of depression, anxiety, and suicidal ideation among Ph.D. students. Scientific Reports. 2021; 11(1):1-12.

[9]   Charles ST, Karnaze MM, Leslie FM. Positive factors related to graduate student mental health. Journal of American College Health. 2021:1-9.

[10] Ibrahim EN, Jamali N, Suhaimi AI. Exploring gamification design elements for mental health support. International Journal of Advanced Technology and Engineering Exploration. 2021; 8(74):114-25.

[11] Cook BL, Progovac AM, Chen P, Mullin B, Hou S, Baca-garcia E. Novel use of natural language processing (NLP) to predict suicidal ideation and psychiatric symptoms in a text-based mental health intervention in Madrid. Computational and Mathematical Methods in Medicine. 2016.

[12] Rambocas M, Gama J. Marketing research: the role of sentiment analysis. University of Porto, Faculty of Economics of Porto; 2013.

[13] Baden M, Lu L, Drummond FJ, Gavin A, Sharp L. Pain, fatigue and depression symptom cluster in survivors of prostate cancer. Supportive Care in Cancer. 2020; 28(10):4813-24.

[14] Alhassan AA, Alqadhib EM, Taha NW, Alahmari RA, Salam M, Almutairi AF. The relationship between addiction to smartphone usage and depression among adults: a cross sectional study. BMC Psychiatry. 2018; 18:1-8.

[15] Stanković M, Nešić M, Čičević S, Shi Z. Association of smartphone use with depression, anxiety, stress, sleep quality, and internet addiction. Empirical evidence from a smartphone application. Personality and Individual Differences. 2021.

[16] Matar BJ, Jaalouk D. Depression, anxiety, and smartphone addiction in university students-a cross sectional study. PloS ONE. 2017; 12(8).

[17] Havigerová JM, Haviger J, Kučera D, Hoffmannová P. Text-based detection of the risk of depression. Frontiers in Psychology. 2019.

[18] Chahar R, Kaur D. A systematic review of the machine learning algorithms for the computational analysis in different domains. International Journal of Advanced Technology and Engineering Exploration. 2020; 7(71):147-64.

[19] Dubey AK, Narang S, Kumar A, Satya MS, Garcia-diaz V. Performance estimation of machine learning algorithms in the factor analysis of COVID-19 dataset. Computers, Materials, & Continua. 2021; 66(2):1921-36.

[20] Zhao M, Feng Z. Machine learning methods to evaluate the depression status of Chinese recruits: a diagnostic study. Neuropsychiatric Disease and Treatment. 2020:2743-52.

[21] Kwakernaak S, Van MK, Cahn W, Janssen R, Investigators G. Using machine learning to predict mental healthcare consumption in non-affective psychosis. Schizophrenia Research. 2020; 218:166-72.

[22] Kim J, Lee D, Park E. Machine learning for mental health in social media: bibliometric study. Journal of Medical Internet Research. 2021; 23(3):1-17.

[23] Gao S, Calhoun VD, Sui J. Machine learning in major depression: from classification to treatment outcome prediction. CNS Neuroscience & Therapeutics. 2018; 24(11):1037-52.

[24] Shatte AB, Hutchinson DM, Teague SJ. Machine learning in mental health: a scoping review of methods and applications. Psychological Medicine. 2019; 49(9):1426-48.

[25] Chen L, Xia C, Sun H. Recent advances of deep learning in psychiatric disorders. Precision Clinical Medicine. 2020; 3(3):202-13.

[26] Govindasamy KA, Palanichamy N. Depression detection using machine learning techniques on twitter data. In international conference on intelligent computing and control systems 2021 (pp. 960-6). IEEE.

[27] Cacheda F, Fernandez D, Novoa FJ, Carneiro V. Early detection of depression: social network analysis and random forest techniques. Journal of Medical Internet Research. 2019; 21(6).

[28] Zhu H, Han G, Shu L, Zhao H. Arvanet: deep recurrent architecture for PPG-based negative mental-state monitoring. IEEE Transactions on Computational Social Systems. 2020; 8(1):179-90.

[29] Laden D, Jayasundara S, Kahanda I. RDoCer: information retrieval and sentence extraction for mental health using research domain criteria. In international conference on semantic computing 2020 (pp. 226-9). IEEE.

[30] Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. Journal of Clinical Epidemiology. 2009; 62(10):e1-34.

[31] Dai W, Han D, Dai Y, Xu D. Emotion recognition and affective computing on vocal social media. Information & Management. 2015; 52(7):777-88.

[32] Marín-morales J, Higuera-trujillo JL, Greco A, Guixeres J, Llinares C, Scilingo EP, et al. Affective computing in virtual reality: emotion recognition from brain and heartbeat dynamics using wearable sensors. Scientific Reports. 2018; 8(1):1-15.

[33] Perveen N, Ahmad N, Khan MA, Khalid R, Qadri S. Facial expression recognition through machine learning. International Journal of Scientific & Technology Research. 2016; 5(03):91-7.

[34] Ibrahim BR, Khalifa FM, Zeebaree SR, Othman NA, Alkhayyat A, Zebari RR, et al. Embedded system for eye blink detection using machine learning technique. In Babylon international conference on information technology and science 2021 (pp. 58-62). IEEE.

[35] Ray A, Kumar S, Reddy R, Mukherjee P, Garg R. Multi-level attention network using text, audio and video for depression prediction. In proceedings of the international on audio/visual emotion challenge and workshop 2019 (pp. 81-8).

[36] Thorstad R, Wolff P. Predicting future mental illness from social media: a big-data approach. Behavior Research Methods. 2019; 51(4):1586-600.

[37] Rajeswari A, Sowmbika P, Kalaimagal P, Ramya M, Ranjitha M. Improved emotional speech recognition

algorithms. In international conference on wireless communications, signal processing and networking 2016 (pp. 2362-6). IEEE.

[38] Neitzke AB. An illness of power: gender and the social causes of depression. Culture, Medicine, and Psychiatry. 2016; 40(1):59-73.

[39] Jung Y, Yoon YI. Multi-level assessment model for wellness service based on human mental stress level. Multimedia Tools and Applications. 2017; 76(9):11305-17.

[40] Farhan AA, Yue C, Morillo R, Ware S, Lu J, Bi J, et al. Behavior vs. introspection: refining prediction of clinical depression via smartphone sensing data. In wireless health 2016 (pp. 1-8). IEEE.

[41] Demirci K, Akgönül M, Akpinar A. Relationship of smartphone use severity with sleep quality, depression, and anxiety in university students. Journal of Behavioral Addictions. 2015; 4(2):85-92.

[42] Shim M, Hwang HJ, Kim DW, Lee SH, Im CH. Machine-learning-based diagnosis of schizophrenia using combined sensor-level and source-level EEG features. Schizophrenia Research. 2016; 176(2-3):314-9.

[43] Lins AJ, Muniz MT, Garcia AN, Gomes AV, Cabral RM, Bastos-filho CJ. Using artificial neural networks to select the parameters for the prognostic of mild cognitive impairment and dementia in elderly individuals. Computer Methods and Programs in Biomedicine. 2017; 152:93-104.

[44] Al JM, Guo G. Video-based depression level analysis by encoding deep spatiotemporal features. IEEE Transactions on Affective Computing. 2018.

[45] Iliou T, Konstantopoulou G, Ntekouli M, Lymperopoulou C, Assimakopoulos K, Galiatsatos D, et al. ILIOU machine learning preprocessing method for depression type prediction. Evolving Systems. 2019; 10(1):29-39.

[46] Tariq S, Akhtar N, Afzal H, Khalid S, Mufti MR, Hussain S, et al. A novel co-training-based approach for the classification of mental illnesses using social media posts. IEEE Access. 2019; 7:166165-72.

[47] Ay B, Yildirim O, Talo M, Baloglu UB, Aydin G, Puthankattil SD, et al. Automated depression detection using deep representation and sequence learning with EEG signals. Journal of Medical Systems. 2019; 43(7).

[48] Li X, Zhang X, Zhu J, Mao W, Sun S, Wang Z, et al. Depression recognition using machine learning methods with different feature generation strategies. Artificial Intelligence in Medicine. 2019.

[49] Besteher B, Gaser C, Nenadić I. Machine-learning based brain age estimation in major depression showing no evidence of accelerated aging. Psychiatry Research: Neuroimaging. 2019; 290:1-4.

[50] Xing Y, Rao N, Miao M, Li Q, Li Q, Chen X, et al. Task-state heart rate variability parameter-based depression detection model and effect of therapy on the parameters. IEEE Access. 2019; 7:105701-9.

[51] Yang X, Mcewen R, Ong LR, Zihayat M. A big data analytics framework for detecting user-level depression from social networks. International Journal of Information Management. 2020; 54:102141.

[52] Altameem T, Amoon M, Altameem A. A deep reinforcement learning process based on robotic training to assist mental health patients. Neural Computing and Applications. 2020:1-10.

[53] Abd RR, Omar K, Noah SA, Danuri MS, Al-garadi MA. Application of machine learning methods in mental health detection: a systematic review. IEEE Access. 2020; 8:183952-64.

[54] Kumar P, Garg S, Garg A. Assessment of anxiety, depression and stress using machine learning models. Procedia Computer Science. 2020; 171:1989-98.

[55] Yue C, Ware S, Morillo R, Lu J, Shang C, Bi J, et al. Automatic depression prediction using internet traffic characteristics on smartphones. Smart Health. 2020.

[56] Baek JW, Chung K. Context deep neural network model for predicting depression risk using multiple regression. IEEE Access. 2020; 8:18171-81.

[57] Cai H, Qu Z, Li Z, Zhang Y, Hu X, Hu B. Feature-level fusion approaches based on multimodal EEG data for depression recognition. Information Fusion. 2020; 59:127-38.

[58] Qin FY, Lv ZQ, Wang DN, Hu B, Wu C. Health status prediction for the elderly based on machine learning. Archives of Gerontology and Geriatrics. 2020.

[59] Sharma A, Verbeke WJ. Improving diagnosis of depression with XGBOOST machine learning model and a large biomarkers Dutch dataset (n= 11,081). Frontiers in Big Data. 2020.

[60] Morrow AS, Campos VAD, Zhao X, Liriano MM. Leveraging machine learning to identify predictors of receiving psychosocial treatment for attention deficit/hyperactivity disorder. Administration and Policy in Mental Health and Mental Health Services Research. 2020; 47:680-92.

[61] Lin GM, Nagamine M, Yang SN, Tai YM, Lin C, Sato H. Machine learning based suicide ideation prediction for military personnel. IEEE Journal of Biomedical and Health Informatics. 2020; 24(7):1907-16.

[62] Su C, Aseltine R, Doshi R, Chen K, Rogers SC, Wang F. Machine learning for suicide risk prediction in children and adolescents with electronic health records. Translational Psychiatry. 2020; 10(1):1-10.

[63] Priya A, Garg S, Tigga NP. Predicting anxiety, depression and stress in modern life using machine learning algorithms. Procedia Computer Science. 2020; 167:1258-67.

[64] Frässle S, Marquand AF, Schmaal L, Dinga R, Veltman DJ, Van DWNJ, et al. Predicting individual clinical trajectories of depression with generative embedding. NeuroImage: Clinical. 2020.

[65] Tate AE, Mccabe RC, Larsson H, Lundström S, Lichtenstein P, Kuja-halkola R. Predicting mental health problems in adolescence using machine learning techniques. PloS ONE. 2020; 15(4).

[66] Betts KS, Kisely S, Alati R. Predicting postpartum psychiatric admission using a machine learning approach. Journal of Psychiatric Research. 2020; 130:35-40.

[67] Lin A, Stolfi A, Eicher T, Neeley S. Predicting second-generation antidepressant effectiveness in treating sadness using demographic and clinical information: a machine learning approach. Journal of Affective Disorders. 2020; 272:295-304.

[68] Cho SE, Geem ZW, Na KS. Prediction of depression among medical check-ups of 433,190 patients: a nationwide population-based study. Psychiatry Research. 2020.

[69] Gonçalves C, Ferreira D, Neto C, Abelha A, Machado J. Prediction of mental illness associated with unemployment using data mining. Procedia Computer Science. 2020; 177:556-61.

[70] Martínez-monteagudo MC, Delgado B, Díaz-herrero Á, García-fernández JM. Relationship between suicidal thinking, anxiety, depression and stress in university students who are victims of cyberbullying. Psychiatry Research. 2020.

[71] Sağbaş EA, Korukoglu S, Balli S. Stress detection via keyboard typing behaviors by using smartphone sensors and machine learning techniques. Journal of Medical Systems. 2020; 44(4):1-12.

[72] Castillo-sánchez G, Marques G, Dorronzoro E, Rivera-romero O, Franco-martín M, De LTDI. Suicide risk assessment using machine learning and social networks: a scoping review. Journal of Medical Systems. 2020; 44(12):1-5.

[73] Guntuku SC, Sherman G, Stokes DC, Agarwal AK, Seltzer E, Merchant RM, et al. Tracking mental health and symptom mentions on twitter during covid-19. Journal of General Internal Medicine. 2020; 35(9):2798-800.

[74] Elhai JD, Yang H, Rozgonjuk D, Montag C. Using machine learning to model problematic smartphone use severity: the significant role of fear of missing out. Addictive Behaviors. 2020.

[75] Oyebode O, Alqahtani F, Orji R. Using machine learning and thematic analysis methods to evaluate mental health apps based on user reviews. IEEE Access. 2020; 8:111141-58.

[76] Zhang Y, Wang S, Hermann A, Joly R, Pathak J. Development and validation of a machine learning algorithm for predicting the risk of postpartum depression among pregnant women. Journal of Affective Disorders. 2021; 279:1-8.

[77] Hong S, Liu YS, Cao B, Cao J, Ai M, Chen J, et al. Identification of suicidality in adolescent major depressive disorder patients using sMRI: a machine learning approach. Journal of Affective Disorders. 2021; 280:72-6.

[78] Zulfiker MS, Kabir N, Biswas AA, Nazneen T, Uddin MS. An in-depth analysis of machine learning approaches to predict depression. Current Research in Behavioral Sciences. 2021.

[79] Parghi N, Chennapragada L, Barzilay S, Newkirk S, Ahmedani B, Lok B, et al. Assessing the predictive ability of the suicide crisis inventory for near- term suicidal behavior using machine learning approaches. International Journal of Methods in Psychiatric Research. 2021; 30(1).

[80] He L, Chan JC, Wang Z. Automatic depression recognition using CNN with attention mechanism from videos. Neurocomputing. 2021; 422:165-75.

[81] Wardenaar KJ, Riese H, Giltay EJ, Eikelenboom M, Van HAJ, Beekman AF, et al. Common and specific determinants of 9-year depression and anxiety course-trajectories: a machine-learning investigation in the Netherlands Study of Depression and Anxiety (NESDA). Journal of Affective Disorders. 2021; 293:295-304.

[82] Uddin MZ, Dysthe KK, Følstad A, Brandtzaeg PB. Deep learning for prediction of depressive symptoms in a large textual dataset. Neural Computing and Applications. 2021: 1-24.

[83] Wang R, Hao Y, Yu Q, Chen M, Humar I, Fortino G. Depression analysis and recognition based on functional near-infrared spectroscopy. IEEE Journal of Biomedical and Health Informatics. 2021.

[84] Shin D, Cho WI, Park C, Rhee SJ, Kim MJ, Lee H, et al. Detection of minor and major depression through voice as a biomarker using machine learning. Journal of Clinical Medicine. 2021; 10(14):1-15.

[85] Shi D, Li Y, Zhang H, Yao X, Wang S, Wang G, et al. Machine learning of schizophrenia detection with structural and functional neuroimaging. Disease Markers. 2021.

[86] Richter T, Fishbain B, Fruchter E, Richter-levin G, Okon-singer H. Machine learning-based diagnosis support system for differentiating between clinical anxiety and depression disorders. Journal of Psychiatric Research. 2021; 141:199-205.

[87] Na KS, Cho SE, Cho SJ. Machine learning-based discrimination of panic disorder from other anxiety disorders. Journal of Affective Disorders. 2021; 278:1-4.

[88] Aljameel SS, Khan IU, Aslam N, Aljabri M, Alsulmi ES. Machine learning-based model to predict the disease severity and outcome in COVID-19 patients. Scientific Programming. 2021.

[89] Cho SE, Geem ZW, Na KS. Predicting depression in community dwellers using a machine learning algorithm. Diagnostics. 2021; 11(8):1-12.

[90] Gokten ES, Uyulan C. Prediction of the development of depression and post-traumatic stress disorder in sexually abused children using a random forest classifier. Journal of Affective Disorders. 2021; 279:256-65.

[91] Huberts LC, Does RJ, Ravesteijn B, Lokkerbol J. Predictive monitoring using machine learning algorithms and a real- life example on schizophrenia. Quality and Reliability Engineering International. 2021.

[92] Du YL, Hu JB, Huang TT, Lai JB, Ng CH, Zhang WH, et al. Psychometric properties of the clinically useful depression outcome scale supplemented with DSM-5 mixed subtype questionnaire in Chinese patients with mood disorders. Journal of Affective Disorders. 2021; 279:53-8.

[93] Liu Y, Hankey J, Cao B, Chokka P. Screening for major depressive disorder in a tertiary mental health

centre using earlydetect: a machine learning-based pilot study. Journal of Affective Disorders Reports. 2021; 3:1-6.

[94] Li W, Wang Q, Liu X, Yu Y. Simple action for depression detection: using kinect-recorded human kinematic skeletal data. BMC Psychiatry. 2021; 21(1):1-11.

[95] Salankar N, Koundal D, Mian QS. Stress classification by multimodal physiological signals using variational mode decomposition and machine learning. Journal of Healthcare Engineering. 2021.

[96] Liu Z, Hu B, Yan L, Wang T, Liu F, Li X, et al. Detection of depression in speech. In international conference on affective computing and intelligent interaction 2015 (pp. 743-7). IEEE.

[97] Rukavina S, Gruss S, Hoffmann H, Tan JW, Walter S, Traue HC. Affective computing and the impact of gender and age. PloS One. 2016; 11(3):1-20.

[98] Cai J, Wang ZJ, Appel-cresswell S, Mckeown MJ. Feature selection to simplify BDI for efficient depression identification. In Canadian conference on electrical and computer engineering 2016 (pp. 1-4). IEEE.

[99] Li X, Cao T, Sun S, Hu B, Ratcliffe M. Classification study on eye movement data: towards a new approach in depression detection. In congress on evolutionary computation 2016 (pp. 1227-32). IEEE.

[100] Alghowinem S, Goecke R, Wagner M, Epps J, Hyett M, Parker G, et al. Multimodal depression detection: fusion analysis of paralinguistic, head pose and eye gaze behaviors. IEEE Transactions on Affective Computing. 2016; 9(4):478-90.

[101] Deshpande M, Rao V. Depression detection using emotion artificial intelligence. In international conference on intelligent sustainable systems 2017 (pp. 858-62). IEEE.

[102] Jan A, Meng H, Gaus YF, Zhang F. Artificial intelligent system for automatic depression level analysis through visual and vocal expressions. IEEE Transactions on Cognitive and Developmental Systems. 2017; 10(3):668-80.

[103] Zhu Y, Shang Y, Shao Z, Guo G. Automated depression diagnosis based on deep networks to encode facial appearance and dynamics. IEEE Transactions on Affective Computing. 2017; 9(4):578-84.

[104] Pampouchidou A, Simos PG, Marias K, Meriaudeau F, Yang F, Pediaditis M, et al. Automatic assessment of depression based on visual cues: a systematic review. IEEE Transactions on Affective Computing. 2017; 10(4):445-70.

[105] Kiss G, Vicsi K. Comparison of read and spontaneous speech in case of automatic detection of depression. In international conference on cognitive infocommunications 2017 (pp. 213-8). IEEE.

[106] Dibeklioğlu H, Hammal Z, Cohn JF. Dynamic multimodal measurement of depression severity using deep autoencoding. IEEE Journal of Biomedical and Health Informatics. 2017; 22(2):525-36.

[107] Zucco C, Calabrese B, Cannataro M. Sentiment analysis and affective computing for depression monitoring. In international conference on bioinformatics and biomedicine 2017 (pp. 1988-95). IEEE.

[108] Madhav KC, Sherchand SP, Sherchan S. Association between screen time and depression among US adults. Preventive Medicine Reports. 2017; 8:67-71.

[109] Pampouchidou A, Pediaditis M, Maridaki A, Awais M, Vazakopoulou CM, Sfakianakis S, et al. Quantitative comparison of motion history image variants for video-based depression assessment. EURASIP Journal on Image and Video Processing. 2017:1-11.

[110] Ingle VK, Pandey I, Singh AR, Pakhare A, Kumar S. Screening of patients with chronic medical disorders in the outpatient department for depression using handheld computers as interface and patient health questionnaire-9 as a tool. International Journal of Applied and Basic Medical Research. 2017; 7(2):129-33.

[111] Kulkarni PB, Patil MM. Clinical depression detection in adolescent by face. In international conference on smart city and emerging technology 2018 (pp. 1-4). IEEE.

[112] Al-gawwam S, Benaissa M. Depression detection from eye blink features. In international symposium on signal processing and information technology 2018 (pp. 388-92). IEEE.

[113] Stepanov EA, Lathuiliere S, Chowdhury SA, Ghosh A, Vieriu RL, Sebe N, et al. Depression severity estimation from multiple modalities. In international conference on e-health networking, applications and services 2018 (pp. 1-6). IEEE.

[114] He L, Jiang D, Sahli H. Automatic depression analysis using dynamic facial appearance descriptor and dirichlet process fisher encoding. IEEE Transactions on Multimedia. 2018; 21(6):1476-86.

[115] Kacem A, Hammal Z, Daoudi M, Cohn J. Detecting depression severity by interpretable representations of motion dynamics. In international conference on automatic face & gesture recognition 2018 (pp. 739-45). IEEE.

[116] Islam MR, Kamal AR, Sultana N, Islam R, Moni MA. Detecting depression using k-nearest neighbors (knn) classification technique. In international conference on computer, communication, chemical, material and electronic engineering 2018 (pp. 1-4). IEEE.

[117] Kiss G, Takács AB, Sztahó D, Vicsi K. Detection possibilities of depression and parkinson's disease based on the ratio of transient parts of the speech. In international conference on cognitive infocommunications 2018 (pp.165-8). IEEE.

[118] Katchapakirin K, Wongpatikaseree K, Yomaboot P, Kaewpitakkun Y. Facebook social media for depression detection in the Thai community. In international joint conference on computer science and software engineering 2018 (pp. 1-6). IEEE.

[119] Nakamura M, Shinohara S, Omiya Y, Higuchi M, Mitsuyoshi S, Takano T, et al. Feasibility study for

estimation of depression severity using voice analysis. In international conference on bioinformatics and biomedicine 2018 (pp. 2792-4). IEEE.

[120]Nakamura R, Mitsukura Y. Feature analysis of electroencephalography in patients with depression. In life sciences conference 2018 (pp. 53-6). IEEE.

[121]Song S, Shen L, Valstar M. Human behaviour-based automatic depression analysis using hand-crafted statistics and deep learned spectral features. In international conference on automatic face & gesture recognition 2018 (pp. 158-65). IEEE.

[122]Li J, Fu X, Shao Z, Shang Y. Improvement on speech depression recognition based on deep networks. In Chinese automation congress 2018 (pp. 2705-9). IEEE.

[123]Yang L, Jiang D, Sahli H. Integrating deep and shallow models for multi-modal depression analysis—hybrid architectures. IEEE Transactions on Affective Computing. 2018; 12(1):239-53.

[124]Maridaki A, Pampouchidou A, Marias K, Tsiknakis M. Machine learning techniques for automatic depression assessment. In international conference on telecommunications and signal processing 2018 (pp. 1-5). IEEE.

[125]Oyong I, Utami E, Luthfi ET. Natural language processing and lexical approach for depression symptoms screening of Indonesian twitter user. In international conference on information technology and electrical engineering 2018 (pp. 359-64). IEEE.

[126]Song T, Zheng W, Song P, Cui Z. EEG emotion recognition using dynamical graph convolutional neural networks. IEEE Transactions on Affective Computing. 2018; 11(3):532-41.

[127]Cong Q, Feng Z, Li F, Xiang Y, Rao G, Tao C. XA-BiLSTM: a deep learning approach for depression detection in imbalanced data. In international conference on bioinformatics and biomedicine 2018 (pp. 1624-7). IEEE.

[128]Martínez-borba V, Suso-ribera C, Osma J. The use of information and communication technologies in perinatal depression screening: a systematic review. Cyberpsychology, Behavior, and Social Networking. 2018; 21(12):741-52.

[129]Eichstaedt JC, Smith RJ, Merchant RM, Ungar LH, Crutchley P, Preoţiuc-pietro D, et al. Facebook language predicts depression in medical records. Proceedings of the National Academy of Sciences. 2018; 115(44):11203-8.

[130]Shah A, Morthland M, Scogin F, Presnell A, Dinapoli EA, Decoster J, et al. Audio and computer cognitive behavioral therapy for depressive symptoms in older adults: a pilot randomized controlled trial. Behavior Therapy. 2018; 49(6):904-16.

[131]Tan Q, Cai Y, Li Q, Zhang Y, Tu D. Development and validation of an item bank for depression screening in the Chinese population using computer adaptive testing: a simulation study. Frontiers in Psychology. 2018; 9:1-13.

[132]Basu A, Dasgupta A, Thyagharajan A, Routray A, Guha R, Mitra P. A portable personality recognizer based on affective state classification using spectral fusion of features. IEEE Transactions on Affective Computing. 2018; 9(3):330-42.

[133]Jaiswal S, Song S, Valstar M. Automatic prediction of depression and anxiety from behaviour and personality attributes. In international conference on affective computing and intelligent interaction 2019 (pp. 1-7). IEEE.

[134]Gerych W, Agu E, Rundensteiner E. Classifying depression in imbalanced datasets using an autoencoder-based anomaly detection approach. In international conference on semantic computing 2019 (pp. 124-7). IEEE.

[135]Ma L, Wang Y. Constructing a semantic graph with depression symptoms extraction from twitter. In conference on computational intelligence in bioinformatics and computational biology 2019 (pp. 1-5). IEEE.

[136]Lam G, Dongyan H, Lin W. Context-aware deep learning for multi-modal depression detection. In ICASSP international conference on acoustics, speech and signal processing 2019 (pp. 3946-50). IEEE.

[137]Kwon H, Kang S, Park W, Park J, Lee Y. Deep learning based pre-screening method for depression with imagery frontal EEG channels. In international conference on information and communication technology convergence 2019 (pp. 378-80). IEEE.

[138]Guo W, Yang H, Liu Z. Deep neural networks for depression recognition based on facial expressions caused by stimulus tasks. In international conference on affective computing and intelligent interaction workshops and demos 2019 (pp. 133-9). IEEE.

[139]De MWC, Granger E, Hadid A. Depression detection based on deep distribution learning. In international conference on image processing 2019 (pp. 4544-8). IEEE.

[140]Pan Z, Ma H, Zhang L, Wang Y. Depression detection based on reaction time and eye movement. In international conference on image processing 2019 (pp. 2184-8). IEEE.

[141]Fang J, Wang T, Li C, Hu X, Ngai E, Seet BC, et al. Depression prevalence in postgraduate students and its association with gait abnormality. IEEE Access. 2019; 7:174425-37.

[142]Yuan Y, Wang Q. Detection model of depression based on eye movement trajectory. In international conference on data science and advanced analytics 2019 (pp. 612-3). IEEE.

[143]Tadesse MM, Lin H, Xu B, Yang L. Detection of depression-related posts in reddit social media forum. IEEE Access. 2019; 7:44883-93.

[144]Sumali B, Mitsukura Y, Tazawa Y, Kishimoto T. Facial landmark activity features for depression screening. In annual conference of the society of instrument and control engineers of Japan 2019 (pp. 1376-81). IEEE.

[145]Kim M, Lim JS. Finding and evaluating suitable contents to recognize depression based on neuro-fuzzy algorithm. In international conference on information

Ravita Chahar et al.

and communication technology convergence 2019 (pp. 478-83). IEEE.

[146]Zhou X, Huang P, Liu H, Niu S. Learning content-adaptive feature pooling for facial depression recognition in videos. Electronics Letters. 2019; 55(11):648-50.

[147]Niu M, Tao J, Liu B. Local second-order gradient cross pattern for automatic depression detection. In international conference on affective computing and intelligent interaction workshops and demos 2019 (pp. 128-32). IEEE.

[148]Al-naggar AN, Bamashmos SH, Wadaane M, Abou AM, Hamawy L. Major depressive disorder early detection. In first international conference of intelligent computing and engineering 2019 (pp. 1-7). IEEE.

[149]Kumar SD, Subha DP. Prediction of depression from EEG signal using long short term memory (LSTM). In international conference on trends in electronics and informatics 2019 (pp. 1248-53). IEEE.

[150]Hao F, Pang G, Wu Y, Pi Z, Xia L, Min G. Providing appropriate social support to prevention of depression for highly anxious sufferers. IEEE Transactions on Computational Social Systems. 2019; 6(5):879-87.

[151]Syarif I, Ningtias N, Badriyah T. Study on mental disorder detection via social media mining. In international conference on computing, communications and security 2019 (pp. 1-6). IEEE.

[152]Liang G, Li Y, Liao D, Hu H, Zhang Y, Xu X. The relationship between EEG and depression under induced emotions using VR scenes. In MTT-S international microwave biomedical conference 2019 (pp. 1-4). IEEE.

[153]Ware S, Yue C, Morillo R, Lu J, Shang C, Bi J, et al. Predicting depressive symptoms using smartphone data. Smart Health. 2020.

[154]Gadassi PR, Tran TB, Joormann J. "What has been is what will be"? autobiographical memory and prediction of future events in depression. Cognition and Emotion. 2020; 34(5):1044-51.

Ms. Ravita Chahar is an Assistant Professor in Computer Science and Engineering Department at Chitkara University, Punjab. She is having a rich experience of 12 years in the teaching field. She holds the degree of Master in Engineering in Computer Science and Engineering from Maharshi Dayanand University, Rohtak. Her area of research interests are Databases, Data Mining and Machine Learning. Currently, she is pursuing Ph.D in Computer Science and Engineering from the Chitkara University, Punjab.
Email: ravita.chahar@chitkara.edu.in

Dr. Ashutosh Kumar Dubey is currently in the department of Computer Science and Engineering, Chitkara University, Rajpura, Punjab, India. He received his PhD degree in Computer Science and Engineering from JK Lakshmipat University, Jaipur, Rajasthan, India. He is the Senior Member of IEEE and ACM. He has more than 14 years of teaching experience. He has authored a book name Database Management Concepts. He has been associated with many international and national conferences as the Technical Program Committee member. He is also associated as the Editor/Editorial Board Member/ Reviewer of many peer-reviewed journals. His research areas are Data Mining, Health Informatics, Optimization, Machine Learning, Cloud Computing, Artificial Intelligence and Object-Oriented Programming.
Email: ashutosh.dubey@chitkara.edu.in

Dr. Sushil Narang is an Associate Professor in the Department of Computer Science and Engineering at Chitkara University, Rajpura, Punjab (India) since 2019. From 2006-2019, He was head of IT department at SAS Institute of IT & Research, Mohali, Punjab (India). From 1996-2006, He was Assistant Professor at Department of Computer Science Applications, MLN College, Yamunanagar, Haryana (India). He Completed his Ph.D. at Panjab University, Chandigarh (India). His Research on "Feature Extraction and Neural Network Classifiers for Optical Character Recognition for Good quality handwritten Gurmukhi and Devnagari Characters" focused on various image processing, machine as well as deep learning algorithms. His research interests lie in the area of programming languages, ranging from theory to design to implementation, Image Processing, Data Analytics and Machine Learning. He has collaborated actively with researchers in several other disciplines of computer science, particularly Machine Learning on real world use cases. He is a certified Deep Learning Engineer from Edureka. He possesses expertise in Object-Oriented Analysis; Design and Development using Java and Python programming using OpenCV in Image Processing and Neural Network construction. He has strong knowledge of C++ and Java with experience in component architecture of the product interface. With Solid training and management skills, he has demonstrated proficiency in leading and mentoring individuals to maximize levels of productivity, while forming cohesive team environments.
Email: sushilk.narang@chitkara.edu.in

Appendix I

S. No.	Abbreviation	Description
1	AdaBoost	Adaptive Boosting
2	ADHD	Attention Deficit Hyperactivity Disorder
3	AEKA	Automatic Extract Keyword algorithm
4	AI	Artificial Intelligence
5	API	Application Programming Interface
6	ANN	Artificial Neural Network
7	ASRS	ADHD Self- Report Scale
8	AUPRC	Area under the precision-recall curve
9	AUC	Area Under the Curve
10	AUDIT	Alcohol Use Disorders Identification Test
11	AVEC-13	Audio/Visual Emotion and depression recognition challenge-13
12	AVEC-14	Audio/Visual Emotion and depression recognition challenge-14
13	BAI	Beck anxiety Inventory
14	BDC	Bums Depression Checklists
15	BDI	Beck Depression Inventory
16	BDI-II	Beck Depression Inventory-II
17	BFI	Big Five Inventory
18	BN	Bayesian Network
19	BSRS-5	Brief Symptom Rating Scale
20	CART	Classification and Regression Trees
21	CATSS	Child and Adolescent Twin Study in Sweden
22	CCMC	Connecticut Children's Medical Center
23	CEONHPE/HUOC/UPE	Centro de Oncohematologia Peditrica of the Hospital Oswaldo Cruz at University of Pernambuco
24	C3D	Three Dimensional Convolutional Network
25	CNN	Convolutional Neural Network
26	CHARLS	China Health and Retirement Longitudinal Survey
27	CI	Confidence Interval
28	CPES	Collaborative Psychiatric Epidemiology Survey
29	CRISP-DM	Cross Industry Standard Process for Data Mining
30	COBRE	Center for Biomedical Research Excellence
31	CUDOS-M-C	Clinically Useful Depression Outcome Scale Supplemented with DSM-5 Mixed Subtype
32	DASS21	Depression Anxiety Stress Scales-21
33	DASS42	Depression Anxiety Stress Scales-42
34	DBN	Deep Belief Network
35	DCNN	Deep convolution neural networks
36	DCM	Dynamic Causal Modeling
37	DL	Deep Learning
38	DLNN	Deep Learning Neural Network
39	DRL	Deep Reinforcement Learning
40	DRNN	Deep Recurrent Neural Network
41	DSM-IV	Diagnostic and Statistical Manual of Mental Disorder-IV
42	DT	Decision Tree
43	DM	Data Mining
44	DNN	Deep Neural Network
45	DSM-V	Diagnostic and Statistical Manual of Mental Disorder-V
46	DP	Dirichlet Process
47	DPFV	Dirichlet Process Fisher Vector
48	EEG	Electroencephalogram
49	EPV	Events-per-variable
50	ECG	Electrocardiography
51	EHRs	Electronic Health Records
52	EPDS	Edinburgh Postnatal Depression Scale
53	EMR	Electronic Medical Record
54	FDHH	Feature Dynamic History Histogram
55	FOMO	Fear of Missing out
56	fNIRS	Functional Near-Infrared Spectroscopy
57	FV	Fisher Vector
58	GA	Genetic Algorithm
59	GAD-7	Generalized Anxiety Disorder scale-7
60	GBT	Gradient Boosting Trees
61	GB	Gradient Boosting
62	GE	Generative Embedding
63	GPUs	Graphical Processing Units
64	GSR	Global Signal Regression
65	GMHI	Gabor Motion History Image
66	HCGSN	HydroCel Geodesic Sensor Net
67	HC	Healthy Controls
68	HDRS-24	Hamilton Depression Rating Scale-24
69	HRV	Heart rate variability
70	HAM-D	Hamilton Depression Rating Scale
71	HDR	Histogram of Displacement Range
72	HMM	Hidden Markov Model
73	IAU	International Classification of Disease-10
74	ICD-10	International Classification of Disease-10
75	ICC	Intraclass Correlation Coefficient
76	ICT	Information Technology
77	IXI	Information eXtraction from Images
78	KNHANES	Korea National Health and Nutrition Examination Survey
79	KNN	K-nearest neighbors
80	KSA	Kingdom of Saudi Arabia
81	LASSO	Least Absolute Shrinkage and Selection Operator
82	LCGA	Latent class growth analysis
83	LGBP-TOP	Local Gabor Binary Patterns from Three Orthogonal Planes
84	LIME	Local Interpretable Model-Agnostic Explanations
85	LNN	Local Nearest Neighbour
86	LOOCV	Leave-one-out cross-validation
87	LDA	Linear Discriminant Analysis
88	LIWC	Linguistic Inquiry and Word Count
89	LLD	Low-Level Descriptors
90	LR	Logistic Regression
91	LSA	Latent Semantic Analysis
92	LSOGCP	Local Second-Order Gradient Cross Pattern
93	LSTM	Long-Short Term Memory
94	MAC	Multimodal affective content
95	MADN	Multiscale Audio Delta Normalization
96	MATLAB	MATrix LABoratory
97	MCI	Mild Cognitive Impairment

98	MCC	Matthews correlation coefficient
99	MDD	Major Depressive Disorder
100	MDQ	Mood Disorder Questionnaire
101	MFCCs	Mel-frequency cepstral coefficients
102	MHH	Motion history histogram
103	MHI	Motion History Image
104	MINI	Mini International Neuropsychiatric Interview
105	ML	Machine Learning
106	MLP	Multilayer Perceptron
107	MRI	Magnetic resonance imaging
108	MRLBP-TOP	Median Robust Local Binary Patterns from Three Orthogonal Planes
109	mRmR	Minimal Redundancy maximal Relevance
110	MSE	Mean Squared Error
111	MNB	Multinomial Naïve Bayes
112	M3	multimodal imaging and multilevel characterization with multiclassifier
113	MTCNN	Multi-task Cascade Convolution Networks
114	NB	Naïve Bayes
115	NESDA	Netherlands Study of Depression and AnxietyNN
116	NMF	Non-negative Matrix Factorization
117	NPV	Negative Predictive Value
118	NLP	Natural Language Processing
119	NSCH	National Survey of Children's Health
120	NYC-CDRN	New York City Clinical data Research Network
121	O	Over Sampling
122	OASIS	Open Access Series of Imaging Studies
123	OR	Original Sample
124	OSN	Online Social Networks
125	OU	Over-Under
126	PANSS	Positive and Negative Syndrome Scale
127	PCA	Principal Component Analysis
128	PHQ-9	Patient Health Questionnaire-9
129	PP	Pioncare Plots
130	PPD	Predict Postpartum Depression
131	PPR	Portable Personality Recognizer
132	PPV	Positive Predictive Value
133	PRAW	Python Reddit API Wrapper
134	PR-AUC	Precision-Recall area under the curve
135	PRC	Precision-Recall Curve
136	PSQI	Pittsburgh Sleep Quality Index
137	PSU	Problematic Smartphone Use
138	RAKE	Rapid Automatic Keyword Extraction
139	RBFN	Radial Basis Function Network
140	RF	Random Forest
141	RMSE	Root mean square error
142	RNN	Recurrent neural network
143	RNN-C3D	Three-dimensional Convolutional Network and Recurrent Neural Network
144	ROC	Receiver Operating Characteristics Curve
145	RoT	Ratio of the transient
146	RRS	Ruminative Responses Scale

147	rsfMRI	Resting state fMRI (rsfMRI)
148	RVR	Relevance Vector Regression
149	SAS	Smartphone Addiction Scale
150	SAS-SV	Smartphone Addiction Scale-Short Version
151	SCID	Structured Clinical Interview for the DSM Disorder
152	SDQ	Strengths and Difficulties Questionnaire
153	SDS	Self-rating Depression Scale
154	SFFS	Sequential Forward Floating Selection
155	SGBT	Stochastic Gradient Boosting classifiers
156	SGD	Stochastic Gradient Descent
157	SPSS	Statistical Package for the Social Science
158	SMO	Sequential minimal optimization
159	SMOTE	Synthetic Minority Over Sampling Technique
160	sMRI	Structural Magnetic Resonance
161	SL	Superlearner
162	SVM	Support Vector Machine
163	SVR	Support Vector Regression
164	SZs	Schizophrenia
165	TF-IDF	Term Frequency-Inverse Document Frequency
166	TCM	Traditional Chinese Medicine
167	U	Under Sampling
168	VGG	Visual Geometry Group
169	VHQ	Virtual Human Questionnaire
170	VMD	Variational mode decomposition
171	WCM	Weil Cornell Medicine
172	WEKA	Waikato Environment for Knowledge Analysis
173	XAI	Explainable Artificial Intelligence
174	X-A-BiLSTM	XGBoost-Attention-Bidirectional Long Short-Term Memory
175	XGBoost	eXtreme Gradient Boosting