International Journal for Innovative Engineering and Management Research

PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL

www.ijiemr.org

xx

COPY RIGHT

IJIEMR Transactions, online available on 31th May 2024. Link https://www.ijiemr.org/downloads/Volume-13/ISSUE-5

10.48047/IJIEMR/V13/ISSUE 05/58

TITLE: Exploring Machine Learning for Depression Detection: A Systematic Review and Analysis

Volume 13, ISSUE 05, Pages: 531-545

Paper Authors **Rajesh Saxena, Dr. Ashish Saini ,Dr. Rishi Kumar Sharma**

USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per UGC Guidelines We Are Providing A Electronic Bar Code

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Exploring Machine Learning for Depression Detection: A Systematic Review and Analysis

Rajesh Saxena[1], Dr. Ashish Saini[2] ,Dr. Rishi Kumar Sharma[3]
1 Research Scholar, CSE Department, Quantum University, Roorkee, India
Email: mr.rajeshsaxena24@gmail.com
2 Assistant Professor ,CSE Department, Quantum University, Roorkee, India
Email: ashishsainigkv7@gmail.com
3Associate Professor ,CSE Department, Quantum University, Roorkee, India
Email: risi.rishi1526@gmail.com

Abstract

Depression, a prevalent global mental health challenge, exerts profound impacts on individuals' physical well-being, emotional states, and social interactions. The World Health Organization estimates that approximately 450 million people worldwide currently grapple with mental health disorders. Timely identification of mental illnesses holds significant potential for enhancing outcomes. However, conventional diagnostic methods are characterized by time-intensive procedures, high costs, and prolonged patient monitoring requirements. To overcome these limitations, contemporary approaches leverage machine learning (ML) algorithms to analyze patient behavior, facilitating early detection. This review explores the landscape of depression detection models utilizing ML algorithms. The investigation commences by delineating distinct stages of mental illness manifestation. Subsequently, the diverse types of ML algorithms and their applications in mental health assessment are examined. The focus extends to elucidating the methodologies employed in constructing models for anxiety, stress, and depression detection through ML techniques. Critical analysis is a central aspect of this exploration, scrutinizing existing models based on objectives, ML algorithms employed, input-output parameters, databases, performance metrics, and tools. The aim is to provide comprehensive insights into the strengths and limitations of current approaches. Additionally, the review offers detailed insights into the commonly utilized tools, databases, and performance metrics in these models. In conclusion, the identification of outstanding research challenges is coupled with recommendations to enhance existing model efficacy. By shedding light on areas requiring further exploration and development, the review seeks to catalyze advancements in depression detection through ML. Ultimately, it emphasizes the transformative potential of ML in revolutionizing mental health assessment, paving the way for more efficient, cost-effective, and accessible avenues for identifying and addressing mental illness through data-driven insights.

Keywords: Depression, Machine Learning, Mental illness, feature selection

1. INTRODUCTION

Depression is widespread, impacting around 280 million people globally and ranking as a top cause of disability. The World Health Organization predicts that by 2030, depression will surpass other conditions to become the foremost cause of non-fatal diseases worldwide [1]. Depression is prevalent among young adults, notably university students, posing a significant public health issue [2]. A UNICEF study across 21 countries in the first half of 2021 found that around 19% of young people aged 15-24 self-reported experiencing depression [3]. A worldwide study on depression rates among college students revealed a prevalence ranging from 10% to 85%, with an average of 30.6% [4]. University students often face pivotal moments in their lives, navigating both positive and challenging changes. Living in hostels, forming romantic connections, striving for independence, and managing self-care can become daunting, triggering heightened stress, anxiety, and depression for many. These multifaceted experiences shape the student journey, demanding resilience and adaptability. It's crucial for universities to provide support systems and resources that address the emotional well-being of students during this transformative period, ensuring a healthier and more successful academic and personal life [5]. In addressing the demand for better mental health services, digital technology plays a crucial role by enhancing accessibility, engagement, and treatment outcomes. This growing significance has led to the development of a wide array of health technologies and applications, contributing to more effective mental health care [6]. Recently, there has been a rise in literature reviews and research surveys within medical and clinical psychology exploring Machine Learning (ML) applications for mental health. Shatte et al. [7] studied and review 300 literature records to explore the applications of Machine Learning (ML) in mental health. The researchers delineated four primary domains of focus, namely: (a) the identification and diagnosis of conditions, (b) prognosis, treatment, and supportive measures, (c) public health, and (d) research and administration. The empirical investigations demonstrate the capacity of machine learning (ML) to augment both clinical and research methodologies, thereby providing novel perspectives on mental health and overall well-being. The influence of Machine Learning (ML) on mental health hinges on how ML systems are designed. Our thoughts about machine learning are influenced by our hopeful belief that it can be helpful in this area and by our focus on making technology that puts people first and helps society. The hardest part of making a machine learning program is getting lots of good data that represents different types of people and getting permission to use that data to make better and fairer models. To understand this impact, we delve into current research in computing. This study examines the contemporary terrain of machine learning applications within the field of mental health, aiming to elucidate pivotal inquiries and considerations therein. It investigates the types of models and applications under development, exploring motivations driving ML use in mental health. The study examines the data utilized, considering types, scales, and access methods for ML analysis. Additionally, it delves into techniques employed and challenges faced during model development and evaluation. The paper synthesizes key findings from the literature and discusses their real-world implications. Lastly, it assesses the extent to which ethical challenges and implications are considered in the examined papers.

2. LITERATURE REVIEW

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.o

Several academic papers have outlined the use of machine learning in recognizing and evaluating manifestations of mental health issues. A considerable amount of this research focuses on promptly identifying and continuously monitoring depression or related symptoms, often by analyzing acoustic features in speech or Twitter posts. Furthermore, textual analysis has been utilized to automatically extract diagnostic information from written narratives or psychiatric records. Additionally, researchers have examined questionnaire data to distinguish between different mental health states or diagnoses, such as differentiating between patients diagnosed with bipolar I or bipolar II disorders. Utilizing advanced machine learning and text-mining techniques, the authors [8] discerned the ability to predict, within a specified timeframe, individuals most susceptible to suicidal tendencies upon referral to mental health services. Optimal predictive outcomes were achieved by amalgamating structured and free-text medical information, underscoring the significance of incorporating unstructured medical notes. However, employing tf-idf (term frequency - inverse document frequency) textual features yielded a slight reduction in model performance, potentially attributed to the brevity of clinical notes. The authors also observed a marginally inferior performance with topic composition features from the LDA model compared to the bag-of-words (BoW) model. While acknowledging the promise of their findings, the authors stressed the imperative for additional studies with expanded datasets, encompassing information from living patients, to augment predictive efficacy. They view their research as an initial stride toward enhancing suicide assessment in clinical settings through machine learning, advocating for further exploration before clinical deployment consideration. The research [9] delves into the dynamics of the doctor-patient relationship within healthcare, with a specific focus on psychiatry. Its objective is to discern a universal guideline for physicians to cultivate positive connections with their patients. Employing data mining techniques such as C4.5, See5.0, HIDER, and ECL, the study scrutinized patient satisfaction levels and distinguished between psychiatric and non-psychiatric patients. Information was gathered from both groups, centering on satisfaction levels and demographic characteristics, and systematically organized into a dataset for comprehensive analysis. The study underscores the escalating significance of patient satisfaction in healthcare, emphasizing the evolving landscape of the doctor-patient relationship. The findings hold the potential to advance comprehension of psychiatric elements and improve the doctor-patient relationship within healthcare environments. The study [10] illustrates how mobile phone metadata can be utilized to gauge mood disturbance and severity among individuals with bipolar affective disorder. Employing an end-to-end deep architecture called DeepMood, the research employs late fusion to capture diverse metadata views for mood score prediction. This investigation underscores the promise of electronic communication mediums like mobile phones for probing mental health issues and understanding how psychiatric conditions manifest in patients' everyday experiences. Leveraging mobile phone typing dynamics metadata offers a discreet and user-friendly means of gathering data that mirrors illness activity, facilitating the examination and treatment of psychiatric disorders in real-world contexts. Experimental findings indicate that the DeepMood model achieves a depression score prediction accuracy of 90.31% using session-level mobile phone typing dynamics, typically completed in under a minute. The research [11] aimed to figure out if we can predict someone's daily happiness using their smartphone's GPS data. It involved 33 participants, and the average level of reported depressive symptoms was 12.68. Preliminary findings suggest a potential link between an algorithm's ability to predict wellbeing and a person's varied location behavior. The study looked at how smartphone GPS data relates to predicting daily happiness, exploring factors like the accuracy of phone location measurements, the amount of data for learning, depression levels, changes in daily location patterns, and variations in emotional wellbeing. This research [12] focuses on people who've faced depression with over 5 specific symptoms. Its aim is to find the key factors contributing to depression by studying 91 patients with psychological issues. The study used fuzzy logic and Bayesian networks to pinpoint the main psychological factors leading to thoughts of death or suicide in depressed patients. The crucial factors include mood depression, loss of interest, guilt, urban living, and concentrated thoughts. The study's models demonstrated effective classification and high scores for each symptom. The study [13] delves into using text mining to detect communication patterns indicating higher risks of suicidality and depression in young adults. Employing supervised machine learning, researchers crafted classifiers distinguishing depression from suicidality based on a day's text messages. The top-performing classifier, a Deep Neural Network (DNN), achieved 70% accuracy, 81% sensitivity, 56% specificity, and a 44% false alarm rate. The research discusses potential clinical applications, ethical concerns, and merging text data with other smartphone streams to refine risk models. It emphasizes the necessity for innovative tools to gauge acute suicide risk and behavioral interventions to mitigate it. Contributions encompass SMS collection within a broader dataset and a predictive model with clinical implications for prevention. The study also addresses concerns regarding alert recipients and the balance between benefits and risks in deploying predictive monitoring tools. This research [14] explores the practicality of employing mobile sensors for ongoing evaluation of mental and physical health. The focus lies on tracking physical activity and mental health cues through wearable devices and smartphones. It identifies drawbacks in conventional survey methods and suggests mobile sensors as a remedy. The study outlines a methodology involving older adults in a retirement community, using waist-mounted devices with various sensors. Collected data encompasses detailed motion and privacy-conscious audio data, enabling continuous assessment of well-being. The document advocates for automated sensing techniques in mental health assessment, illustrating this with a case study highlighting the limitations of traditional surveys. The research establishes strong correlations between sensed speech and mental health scores, as well as physical activity scores and survey metrics. It underscores the potential of sensor-based systems for continuous health assessment, facilitating early detection. The conclusion stresses the promise of mobile sensors for ongoing mental and physical well-being assessment, urging enhancements in scalability, usability, and robustness. Additionally, there's a discussion about implementing the sensing system on smartphones for improved scalability and data collection in diverse environments. This research [15] investigates stress levels in social media posts after gun-related violence on college campuses, utilizing a machine learning classifier. Examining linguistic markers and psychological attributes, the study reveals heightened

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.o

stress expression, focusing on personal and social concerns, biological aspects, and linguistic style. The findings emphasize the potential for technology in supporting mental health during crises on campuses. The study suggests that technology-assisted tools aid administrators in crisis communication and student support. The stress classifier achieved 82% accuracy, surpassing the 68% baseline. Top predictive features included action-based nouns and verbs. Temporal analysis showed increased High Stress posts following incidents, validated by expert raters. Statistical analysis affirmed a significant rise in stress expression, reinforcing the impact of gun violence on college communities. The findings underscore the importance of addressing stress through technological interventions during crisis events on campuses. The document [16] outlines a study exploring mood prediction using mobile sensing and psychological traits from questionnaires. It analyzed a dataset spanning three years with 17,000 participants. Most reported feeling relaxed, and the method discerned relaxed from non-relaxed individuals with 75% accuracy, boosted by 5% with passive sensing data. Future plans include data imputation techniques, feature importance analysis, and continuous prediction models. They aim to integrate multi-modal approaches for static features. The study's implications extend to mobile health applications, emphasizing passive sensing and psychological traits for mood prediction. It details classifiers, feature extraction, and experimental setups. By combining passive sensing with traditional surveys, mood prediction improves. The study contributes to data exploration, supervised learning for mood detection, and large-scale dataset evaluation. It discusses related work and challenges in mood prediction systems. In conclusion, it highlights the potential of passive sensing and psychological traits for broad mood prediction and mobile health application development. The study [17] compared different Machine Learning techniques to find mental health problems. Stacking was the best, predicting accurately 81.75% of the time. All five methods in the study gave good results, with accuracy above 79%. The research shows how important it is to use computer learning to find mental health issues early, helping patients live better and get better treatment. The study suggests that using computer learning can be really helpful in mental health, helping predict and understand mental illnesses and how to treat them. The paper [18] introduces a new machine learning system that predicts suicide risk more accurately than current methods used by doctors. The system includes a unique way of extracting features, selecting important information, and using risk classifiers. It transforms a patient's medical history into a timeline image and evaluates responses based on specific criteria. The study, conducted on mental health data, confirms the system's effectiveness. In summary, the paper shows that this machine learning framework has great potential in enhancing suicide risk prediction, surpassing the capabilities of current tools. A study [19] discovered that using data from smartphones can reveal connections between students' mental well-being, academic performance, and daily activities. Researchers at Dartmouth College followed a term lifecycle, revealing that as the term progressed, stress increased while positive feelings, sleep, conversations, and activity levels decreased. The study involved 48 students over a 10-week term and is available for public access. The researchers plan to replicate the study at other universities to see if similar patterns emerge. While the study suggests potential benefits in providing feedback to students, privacy concerns must be addressed. The professor overseeing the study used the data to intervene and support students, preventing some from failing. StressMon [20] represents an adaptable system designed to discerning severe stress and depressive episodes among individuals, negating the necessity for specific devices and applications. Leveraging coarse-grained location data obtained from WiFi infrastructure, the system extracts features from individuals' routine behaviors and physical interaction patterns. Demonstrating commendable accuracy, StressMon achieved an AUC (Area Under Curve) score of 0.97 for stress detection and 0.88 for depression detection. Rigorous evaluation encompassed three semester long user studies with 108 students, substantiating its efficacy. Survey and interview data collected during the studies served as a reliable ground truth for assessing student stress levels and the primary causes, thereby validating the system's outcomes. StressMon holds promise for advancing Collaborative and Social Computing research, especially in the realm of mental health management. The research paper [21] introduces a binary classifier ensemble to identify depression through Quality of Life (QoL) scales, enhancing machine learning's ability to recognize depressive cases by analyzing connections between QoL scale elements and mental illness. The classifier achieved an impressive F1 score of 0.976 in predictions, boasting a high accuracy rate of 95.4% and a mere 4% misclassification rate for depressed cases. It surpassed baseline algorithms across all metrics and experiments, showcasing its efficacy in diagnosing Major Depressive Disorder (MDD) using health survey data. The study emphasizes the significance of incorporating psychosocial elements and mental well-being into early MDD diagnosis and prediction, offering a promising avenue for enhancing mental health screening and intervention. The study [22] employed machine learning models, namely sparse logistic regression (SLR), support vector machine (SVM), and random forest (RF), to predict depression risk in Korean college students. Three models' prediction accuracy was compared to assess their performance in identifying depression risk factors. The analysis utilized the family triad dataset, comprising 171 fathers, mothers, and college students. Various performance metrics, including AUC, were employed to evaluate model performance on the test dataset. Logistic regression was also used to identify significant factors associated with depression risk, such as father's cancer severity, mother's respiratory diseases severity, college students' self-perceived mental health, conscientiousness, and neuroticism. The results shed light on machine learning models' ability to accurately predict depression risk and identify crucial family and individual factors linked to depression among Korean college students. The study [23] found that Logistic Regression achieved 77.29% accuracy, Naïve Bayes 74.35%, Support Vector Machine 77.12%, and Random Forest 77.298%. Precision, recall, and f1-score for text classification were 0.80, 0.79, and 0.79 respectively, with an accuracy of 0.79. Logistic Regression led in accuracy, followed by Random Forest, Support Vector Machine, and Naïve Bayes. The research [24] employed machine learning methods, such as logistic regression with lasso regularization, to forecast depression risks in China's elderly. The LSTM+ML model-driven decision support system holds promise for early identification and intervention by healthcare professionals. Logistic regression with lasso regularization exhibited the highest predictive accuracy for depression in the elderly, as indicated by the area under the ROC curve. Lasso regression yielded the greatest net benefit among machine learning models, according to decision curve analysis.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Key predictors for elderly depression encompassed ADL/IADL, self-rated health, marital status, arthritis, and cohabitation frequency, identified consistently by various models like Random Forest and Gradient Boosting Decision Trees. In the study [25] six machine learning classifiers were examined to identify depression. The AdaBoost classifier, employing the SelectKBest feature selection method, achieved the highest accuracy at 92.56%. Bagging classifier yielded the best overall result with 89.26% accuracy, without feature selection. KNN had the lowest accuracy at 66.94%, while AdaBoost, GB, XG-Boost, and Weighted Voting classifiers had accuracies of 87.60%, 86.78%, 85.95%, and 86.78%. Utilizing different feature selection techniques significantly boosted all classifiers' accuracies. With SelectKBest, AdaBoost led with 92.56%, and with mRMR, it achieved 91.74%. AdaBoost also excelled with the Boruta feature selection technique.

3. DEPRESSION: CAUSES, SYMPTOMS, TYPES AND SCALES OF MEASUREMENT

Depression, clinically recognized as major depressive disorder (MDD), entails enduring feelings of profound sadness or pervasive disinterest in life. The DSM-5 (Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition), guides diagnosis, requiring five or more symptoms sustained over two weeks. Symptoms encompass pervasive low mood, fatigue, guilt, hopelessness, concentration issues, sleep disturbances, reduced interest, suicidal thoughts, restlessness, and weight changes. [26]. Depression arises from a complex interplay of biological, environmental, and psychological factors. Familial clustering of depression implies a genetic component, evident in a 70 percent concordance rate in biological twins. Structural and neurotransmitter variations in the brain, hormonal fluctuations (e.g., during menstruation or postpartum), personal or familial mental health history, trauma, substance abuse, medical conditions, and specific medications heighten depression risk. Identified risk factors encompass emotional stressors, societal issues, and physiological conditions, reflecting the multifaceted nature of depression etiology [27]. Major Depressive Disorder (MDD) and Persistent Depressive Disorder (PDD) are common forms of depression. MDD typically entails major depressive episodes with at least five symptoms persisting over two weeks, while PDD, also known as dysthymia, presents as chronic low-level depressed mood. Bipolar disorder shares major depressive episodes, and unipolar depression denotes recurrent major depressive episodes. Postpartum depression affects one in nine women after childbirth, while bipolar disorder involves alternating depressive and manic states. Seasonal Affective Disorder (SAD) occurs mainly in late fall and winter, with psychotic depression blending psychotic features with major depressive symptoms like guilt and worthlessness. [28]. Various scales are employed by medical professionals to gauge depression levels. The Beck Depression Inventory (BDI), applicable to individuals aged 13 to 80, comprises 21 self-report items with multiple-choice responses, necessitating approximately 10 minutes for completion. It has demonstrated global validity and reliability. The Patient Health Questionnaire (PHQ), specifically the PHQ-9, efficiently screens depressive symptoms in 1 to 5 minutes and is available in multiple languages. The Center for Epidemiologic Studies Depression Scale (CES-D), used in primary care, involves 20 self-report items rated on a 4-point scale, suitable from age 6 onward, and takes 20 minutes for administration. The Hamilton Rating Scale for Depression (HDRS or HAM-D) assesses depression across treatments, scored from 17 items in 15 to 20 minutes. Lastly, the 10-item Montgomery-Åsberg Depression Rating Scale (MADRS) evaluates depression severity in adults over 18 within 20 to 30 minutes, demonstrating heightened sensitivity to temporal changes, adapted from the Hamilton Depression Rating Scale [29].

4. MACHINE LEARNING ALGORITHMS USED FOR DEPRESSION DETECTION

Depression poses a global challenge, stressing the need for swift detection and effective treatment. Conventional diagnostic methods are often slow and cumbersome. Machine learning (ML) offers hope by analysing behavioral patterns for early detection. Machine learning (ML), a subset of artificial intelligence (AI), facilitates computer systems in acquiring knowledge from data without necessitating explicit programming instructions. It predicts future outcomes based on past data through algorithms. Machine Learning (ML) comprises supervised, unsupervised, and reinforcement learning paradigms. Within the domain of supervised learning, algorithms acquire knowledge from labeled datasets, enabling them to formulate predictions based on the provided information. It's vital in various fields like speech recognition and document classification. ML aids in improving diagnostic processes, decision-making, and outcomes for those facing depression and other mental health issues. Some important Machine Learning Model used in mental health detection are:

4.1 Support Vector Machine

The Support Vector Machine (SVM) represents a supervised learning paradigm employed for tasks pertaining to both classification and regression analysis. It finds the best hyperplane to separate classes in high-dimensional space, ensuring a maximum margin between the decision boundary and the nearest data points. SVM is robust in high-dimensional scenarios, widely applied in tasks like image and text classification, as well as bioinformatics. Unlike genetic algorithms and perceptrons, SVM returns the same optimal hyperplane parameters consistently, making it more reliable for classification tasks. Genetic algorithms and perceptrons, on the other hand, vary in results due to dependency on initialization and termination criteria, aiming only to minimize training error [30]. Support Vector Machine (SVM) algorithms employ a collection of mathematical functions known as kernels. The commonly utilized kernel is the linear kernel, represented as:

$$K(x,y) = e^{-\left(\frac{\|x-y\|^2}{2\sigma^2}\right)}$$

The data points that lie in close proximity to the hyperplane within a Support Vector Machine (SVM) framework are referred to as support vectors. The margins, in the context of SVM, denote the distances of these support vectors from the hyperplane. Support Vector Machines (SVM) in depression detection analyze patterns in brain imaging or behavioral data, classifying individuals based on features to aid accurate and early diagnosis.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

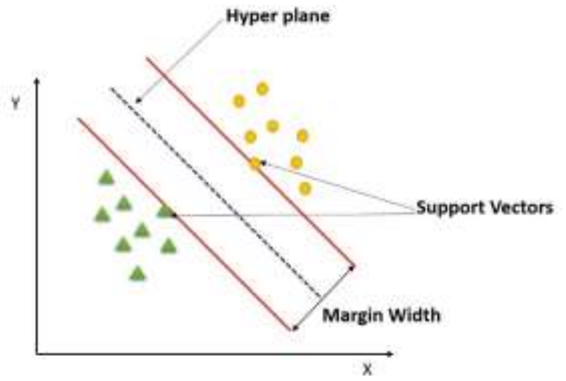

Figure 1. Hyper plane and Support Vector

4.2 k-Nearest Neighbours

kNN, or k-Nearest Neighbours, is a straightforward yet potent non-parametric classification approach. Its effectiveness hinges on selecting an appropriate k value, making it crucial for successful classification in various scenarios [31]. Within this algorithm, a cluster of neighbors, denoted by the parameter \(k \), is strategically chosen to facilitate the classification of a given data point. The assignment of the data point is determined by the majority class among the selected neighbors. The selection of \(k \) is pivotal, as an appropriate choice contributes to enhanced accuracy. This process of selecting the optimal \(k \) value is termed parameter tuning. The model discerns proximity based on the distance between neighboring points and the new data point. This distance is quantified using formulas associated with Euclidean or Minkowski distance, elucidating the spatial relationships that guide the determination of nearest neighbors in the context of the algorithm.

Euclidean Distance = $\sqrt{(x2 - x1)^2 + (y2 - y1)^2}$

Minkowski Distance = $(|x2 - x1|^p + |y2 - y1|^p)^{\frac{1}{p}}$

kNN analyzes features from depressed and non-depressed individuals, determining depression in new cases based on similarity to existing data.

4.3 Random Forest

Random Forests can be used for both classification and regression tasks, accommodating categorical or continuous variables. They efficiently handle high-dimensional problems, can be parallelized, and offer outlier detection and unsupervised learning. Robust to transformations and outliers, they provide a built-in generalization error estimate, making them suitable for diverse data analyses [32]. A Random Forest ensemble typically comprises 10 to 100 Decision Trees, each contributing a distinct decision. In the event of differing decisions among the trees, the majority decision is adopted. The primary merit of employing Random Forest lies in its ability to mitigate the common issue of overfitting encountered in various Machine Learning models. Overfitting arises when a model exhibits high performance during the training phase but fails to generalize well to unseen data during testing. Random Forest effectively addresses this challenge by incorporating the Ada Boost technique internally. This technique facilitates the transfer of misclassified abnormal data between Decision Trees, enabling iterative refinement of the model. Ultimately, this collaborative process enhances the overall predictive accuracy of the Random Forest ensemble. Random Forest employs multiple decision trees to analyze diverse features, enhancing depression detection accuracy through ensemble learning and robust classification.

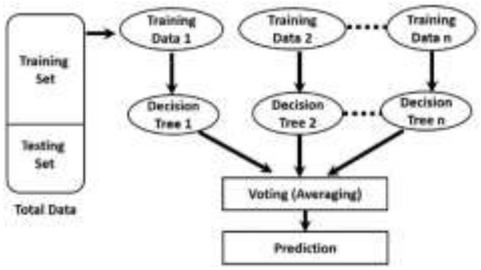

Figure 2. Working of Random Forest Model

4.4 Logistic Regression

Logistic Regression (LR) is commonly used in data mining for binary classification, offering probabilities. It extends to multi-class problems, following linear regression principles. Optimization methods like unconstrained methods enhance LR, with Truncated Newton methods, like TR-IRLS, proving effective for large-scale LR issues, outperforming Support Vector Machines (SVM). LR may face challenges with imbalanced data, small samples, or specific sampling strategies, causing inconsistency. Correction techniques, like prior correction and weighting, help tackle these issues [33]. The sigmoid function is recognized as an activation function within the context of logistic regression and is formally characterized by the following expression:

$$f(x) = \frac{1}{1 + e^{-x}}$$

The sigmoid function, characterized by its S-shaped curve, serves to transform real-valued inputs into a bounded range spanning from 0 to 1.

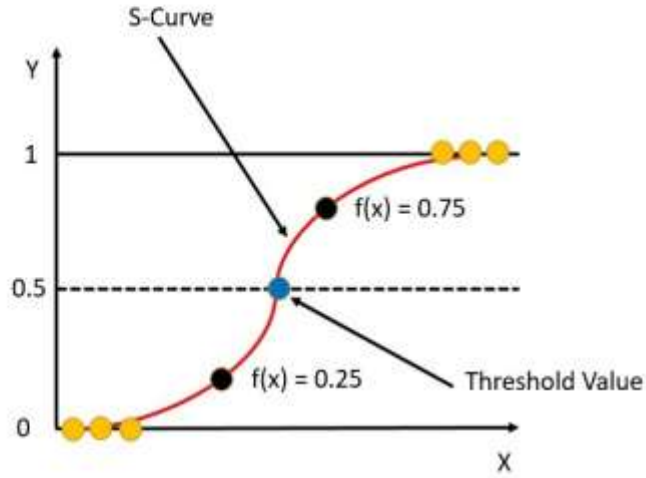

Figure 3. Sigmoid Function

Furthermore, when the output of the sigmoid function, representing the estimated probability, exceeds a predefined threshold, the model classifies the instance as belonging to a particular class. Conversely, if the estimated probability falls below the specified threshold, the model determines that the instance does not pertain to the class in question. Logistic Regression analyzes input features to predict the probability of depression, aiding in early detection and intervention strategies.

4.5 Decision Tree

A decision tree in machine learning is a model that categorizes or predicts data using a tree-like structure of choices. It splits features based on importance, making predictions from complex datasets. It's flexible and improves with multivariate criteria, aiming to minimize errors or other defined targets [34]. This model helps decide how to separate data based on certain rules and come to conclusions. It builds a tree using the dataset's columns as branches at different levels. The end branches give either YES or NO answers. The Decision Tree uses Entropy or Gini Index to figure out the order of branches, starting from the main one. Decision trees analyze psychological data to detect depression by evaluating symptoms, providing a structured approach for diagnostic decision-making.

4.6 XGBOOST

XGBoost, derived from the term Extreme Gradient Boosting, constitutes a machine learning algorithm that utilizes an ensemble of decision trees to address predictive tasks. Through a sequential process, individual trees are constructed successively. XGBoost meticulously assesses the significance of diverse variables by allocating weights to them. These weights are applied to all variables, guiding the data through decision trees for prediction purposes. In instances where a tree inaccurately predicts certain variables, the corresponding weights are adjusted accordingly. Subsequently, data incorporating these adjusted weights is channeled into the subsequent decision tree. This iterative process continues across multiple trees, culminating in the amalgamation of their predictions to form a robust and precise model. XGBoost exhibits versatility, accommodating a spectrum of problem domains, including numerical prediction, categorical classification, ranking, and tailored forecasting. The incorporation of ensemble techniques, exemplified by XGBoost, serves to enhance model efficacy. Positioned as a novel gradient boosting method, XGBoost demonstrates competitive performance across domains such as credit assessment, molecular prediction, and sentiment analysis. [35].

4.7 Naïve Bayes Classifier

Naive Bayes, a quick and effective probabilistic classification algorithm, relies on Bayes' theorem. Despite assuming feature independence, it excels in tasks like text classification and spam filtering. Surprisingly accurate in practice, it often makes correct decisions even with inaccurate probability estimates. Performance is influenced by feature distribution entropy, with lower entropy leading to better results. Naive Bayes performs well with certain types of feature dependencies, particularly in cases of complete independence or functional dependence. The algorithm's accuracy isn't directly tied to the degree of feature dependencies measured by class-conditional mutual information [36]. This model operates on the principles of Bayes' theorem, coupled with the Naïve assumption. Bayes' theorem calculates the probability of an event's occurrence given the probability of another event that has already transpired. Mathematically, Bayes' theorem is expressed through the following equation:

$$P\left(\frac{A}{B}\right) = \frac{P\left(\frac{B}{A}\right) * P(A)}{P(B)}$$

The fundamental premise of the Naïve Bayes classifier lies in the assumption that each feature exerts an independent and uniform influence on the output. Naive Bayes methodology employs a probabilistic model that relies on features to effectively classify instances of depression, leveraging textual or behavioral patterns with efficiency.

4.8 Convolutional Neural Network

A Convolutional Neural Network (CNN) is a neural network that understands the features and patterns in the given data and useful in classification of images, audios or videos. It is a sophisticated type of deep learning framework often employed in tasks related to recognizing images, such as identifying facial expressions [37]. It uses specific layers known as convolutional layers to automatically and systematically grasp patterns and features from the input data. CNN analyzes facial expressions, voice tone, and body language for depression detection with high accuracy.

4.9 Long Short-Term Memory

LSTM, a recurrent neural network, is crucial in Deep Learning for long-term dependency capture. It comprises three gates: Forget, Input, and Output, managing information flow in LSTM cells. LSTM along with natural language processing predicts users' emotions by analyzing their behavior across various applications for depression [38]. LSTM networks are utilized in depression detection by analyzing sequential patterns in individuals' linguistic and behavioral data. LSTMs capture long-range dependencies, identifying nuanced shifts in language and behavior, crucial for discerning depressive states. This deep learning approach enhances accuracy in identifying subtle indicators of depression, contributing to more effective diagnostic tools and interventions.

5. FEATURE SELECTION TECHNIQUES

Selecting relevant features in machine learning is crucial. Many variables in a dataset might be unnecessary, negatively affecting model performance. Feature selection helps identify and use only essential features. Choosing vital features in machine learning is key, excluding unnecessary ones to enhance model performance [39]. Data mining simplifies data by using dimensionality reduction, employing techniques like feature extraction and selection [40]. Feature selection and extraction serve the same goal but differ. Selection picks from original features, while extraction forms new ones. Selection reduces overfitting. Feature selection involves choosing the best and most relevant features for model construction, done either automatically or manually to enhance model performance. Feature selection is an active research area in machine learning because it is an important preprocessing step that leads to success in a variety of real-world applications. In general, there are three types of feature selection algorithms: supervised, semi-supervised, and unsupervised. This review exclusively examines supervised feature selection methods. There are mainly three types of supervised feature selection techniques, which are:

5.1 Filters Method

In the Filter Method, we pick features based on statistics. It doesn't rely on the learning process and picks features before starting. This method removes unnecessary and duplicate features from the model by ranking them using various measures. Filter methods are good because they're fast and don't make the data fit too closely, so they're less likely to be wrong. Filtering involves two main steps. First, the classification model ranks features one by one using criteria like distance, Pearson correlation, or entropy. Then, it picks the top-ranked features based on a set threshold. Features that don't make the cut are considered unimportant. The selected subset of features is then given to the induction classifier for further processing [41]. Filter methods for selecting features are different from wrapper and embedded methods. Unlike wrapper and embedded methods, filter methods are not tied to the classifier algorithm. This separation is good because filter methods are not influenced by the classifier's bias, which helps prevent overfitting. However, the downside is that filter methods don't take into account interactions with the classifier during the feature selection process [42]. There are two main types of filter methods: univariate and multivariate. Univariate methods look at one feature at a time, while multivariate methods look at groups of features together. Univariate methods like the $\chi 2$ test, Fisher's exact test, information gain, Euclidean distance, Pearson correlation, t-test etc. are popular because they are fast and simple, especially in fields dealing with lots of data. Univariate methods look at each feature on its own. They only care about how important each feature is and don't consider if features are similar or interact with each other. This can make models less accurate because they might include too many similar features or miss important interactions. Advanced multivariate filter techniques like Mutual Information Feature Selection (MIFS), Conditional Mutual Information Maximization (CMIM), Minimal-Redundancy-Maximal-Relevance (mRMR), and Fast Correlation-Based Filter (FCBF) have been created to fix this. They find important features while removing similar ones, keeping the most useful information for the model.

5.2 Wrappers method

Unlike filter methods, wrapper methods rely on the performance of a chosen classifier algorithm to pick the best set of features. This means wrapper methods help identify the most effective features for a specific classifier. This advantage has been proven to give better predictive results compared to filter methods [43]. In simple terms, the wrapper method uses a machine learning model like a tool to check different groups of features. It tries out various combinations, trains the model on each, and then picks the one that works best based on how well the model performs. Although it takes a lot of computer power and might memorize the data too much, it's a strong method that's great for finding tricky connections between features and making accurate predictions. People often use the wrapper method in things like recognizing images, understanding speech, and sorting text into categories. Forward feature selection, Backward feature elimination, Exhaustive feature selection are some important wrapper feature selection techniques. The forward selection method is a step-by-step process in which we begin with no features and add the one that most enhances our model in each round. We continue this process until adding a new variable no longer improves the model's performance. On the other hand, the backward elimination method is also iterative but starts with all features and removes the least significant one in each iteration. The process stops when there is no improvement in the model's performance after removing a feature. The exhaustive selection technique is a thorough approach to evaluating feature subsets. It examines all possible subsets, constructs a learning algorithm for each, and selects the subset that yields the best model performance.

5.3 Embedded method

In contrast to the wrapper strategy, which relies on a heuristic search guided by the classifier's outcomes, the embedded strategy represents a middle ground between filter and wrapper approaches. This approach conducts feature selection and constructs an optimized classifier within the learning process. The embedded method, positioned between filters and wrappers, selects features that arise during learning, guided by the classifier's evaluation criteria. This results in reduced computational costs compared to wrappers. Regularization approaches are the prevalent types of embedded methods. In embedded methods, feature selection is part of the classifier algorithm itself. As the classifier learns from the training data, it automatically adjusts its internal settings to assign the right importance to each feature for accurate classification. This means that the process of finding the best features and building the model happens together in one step [44]. Embedded methods are techniques used in machine learning that involve integrating feature selection into the model-building process itself. Examples of embedded methods include decision tree-based algorithms like decision trees, random forests, and gradient boosting. Another approach is feature selection using regularization models such as LASSO or elastic net. Regularization techniques typically operate alongside linear classifiers such as Support Vector Machines (SVM) or logistic regression. They achieve this by reducing the influence or size of coefficients associated with features that do not significantly contribute to the model's effectiveness [45]. Regularization is a technique in machine learning that introduces a penalty to certain parameters of the model to prevent it from overfitting. This method employs Lasso (L1 regularization) and Elastic Nets (L1 and L2 regularization) for feature selection. The penalty is imposed on the coefficients, leading some coefficients to become zero. Features with zero coefficients can be excluded from the dataset. Additionally, tree-based methods like Random Forest and Gradient Boosting offer feature importance, indicating which features have a significant impact on the target feature.

Table 1: Comparison between feature selection techniques

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Feature Selection Method		Strengths	Weaknesses	Example
Filter	Univariate	Rapid, scalable, classifier-agnostic, and mitigating the risk of overfitting.	Feature dependencies and Interaction with classifier not modeled	Chi-squared test, Fisher's exact test, Pearson correlation, t-test
	Multivariate	Less risk of overfitting	Slower and not as scalable as univariate	FCBF (Fast Correlation Based Filter), mRMR (Minimal Redundancy Maximal Relevance)
Wrapper		Better performance than filter method, Model interaction with classifier	Slower than filter and embedded methods, more prone to overfitting, selected features are classifier dependent	Genetic Algorithm, Randomized hill climbing
Embedded		Model feature dependencies, Faster than wrapper method, Model interaction with classifier	Slower than filter methods, Selected features are classifier dependent	Random forest, LASSO (L1) or elastic net regression

6. METHODOLOGY FOR DEPRESSION DETECTION USING MACHINE LEARNING

Detecting depression involves extracting data, processing it, selecting features, and using machine learning classifiers to categorize input data. Apart from machine learning methods, deep learning techniques are also widely used in depression detection. This part explains the steps, methods, and ways to carry out each step in a straightforward manner.

6.1 Pre-Processing Algorithms

Pre-processing in machine learning is the process of applying methods to unprocessed data before it is fed into an algorithm with the goal of improving its quality for more insightful analysis. For efficient pattern recognition and precise forecasting, tasks include handling, cleaning, and normalizing outliers. Some important algorithms used for pre-processing are LDA, SMOTE, LIWC, HMM. LDA (Linear Discriminant Analysis) makes data simpler by cutting down extra details and putting them in a smaller space. It keeps important things and makes it easier to see differences between groups [46]. SMOTE (Synthetic Minority Oversampling Technique) helps make datasets fairer by making more examples of the smaller group[47]. LIWC (Linguistic Inquiry and Word Count) helps understand emotions and structure in writing or talking [48]. HMM (Hidden Markov Model) is a way to understand sequences of symbols by guessing how they were made. It thinks about what you can see to figure out what happened inside [49].

6.2 Feature Extraction Techniques

In machine learning, feature selection is the process of carefully choosing a subset of relevant and significant features from a large feature set. Selecting the most informative and differentiating features that significantly influence the target variable's prediction will help to increase the model's efficacy. Some important feature extraction methods are SelectKBest, mRMR, Boruta, RELIEFF. SelectKBest is a feature extraction technique that keeps relevant features in the input data while removing unnecessary ones. It combines a statistical test with the selection of K features based on the statistical findings among variables using univariate statistical analysis [50]. mRMR (Maximum Relevance Minimum Redundancy) is a feature selection technique handling multivariate temporal data without compressing prior data. It prioritizes features with high relevance to class and minimal redundancy between classes, enhancing predictions in large datasets [51]. Boruta, a feature selection method within Random Forest classification, identifies and preserves pertinent variables while discarding less crucial ones. It iteratively compares feature importance to randomly generated shadows, retaining features with significantly higher importance for improved model performance [52]. RELIEFF is a popular algorithm in machine learning that removes unimportant features, improving model accuracy. It assesses feature importance and updates weights based on contributions, prioritizing discriminative features for better models [53].

6.3 Selection of appropriate Machine Learning Model

Machine learning models are broadly classified into three types: supervised learning, unsupervised learning, and reinforcement learning. Each type serves a different function and is appropriate for a variety of applications. Here's a quick summary of each:

6.3.1 Supervised Learning:

Supervised Learning is a machine learning paradigm that acquires the input-output relationship of a system from paired training samples. The output serves as the label, making the data labeled or supervised. Also known as learning with a Teacher, it aims to build a system predicting the output for new inputs. If the output is discrete, it leads to classification; if continuous, regression. Model parameters represent the input-output relationship and require estimation when not directly available. In contrast to Unsupervised Learning, Supervised Learning demands labeled data. Semi-

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.c

supervised Learning uses both labeled and unlabeled data. Active Learning involves the algorithm querying a user/teacher for labels during training [54]. Examples include Linear Regression, Decision Trees, Support Vector Machines, and Neural Networks.

6.3.2 Unsupervised Learning:

In unsupervised learning, data scientists provide images, and the system autonomously analyzes them to identify whether they contain this particular images or not. Unsupervised learning requires substantial data. Unlike supervised learning, unsupervised learning involves datasets without labeled points, aiming to discover patterns and categorize data points. The challenges include clustering, association, anomaly detection, and autoencoder issues [55]. Clustering, association, anomaly detection, and autoencoder are algorithms commonly used for unsupervised learning tasks in machine learning. Clustering: K-Means, Hierarchical Clustering; Association: Apriori, FP-growth; Anomaly Detection: Isolation Forest; Autoencoder: Principal Component Analysis (PCA).

6.3.3 Reinforcement Learning:

Reinforcement Learning is suited for sequential decision-making, learning through direct interaction with the environment for long-term goals. The agent balances exploration (trying new routes) and exploitation (maximizing rewards). It doesn't require a large dataset and learns by trial and error, receiving positive rewards for correct actions and negative for incorrect ones [56]. Q-Learning, Deep Q Networks (DQN), Policy Gradient methods are algorithms used for reinforcement learning.

6.4 Deep Learning Techniques

Apart from machine learning methods, deep learning techniques are also widely used in depression detection. Deep learning, a subset of machine learning, empowers computers to understand the world by organizing concepts in hierarchical layers. This structure allows machines to grasp intricate ideas by constructing them from simpler ones. In fields like image processing and computer vision, image segmentation is pivotal for applications such as scene understanding and robotic perception. The success of deep learning models in various vision applications has spurred the development of image segmentation approaches. Deep learning involves multi-layer neural networks, and its "deep" nature refers to the multiple stages in data processing, making it a foundational technology for advanced AI and automation, elevating AI to a level termed "Smarter AI"[57]. Deep learning is facilitated by artificial neural networks, drawing inspiration from the structural intricacies of the human brain. Neural networks (NNs) or artificial neural networks (ANN) serve as classifiers, mimicking the operations of human brains and neurons. These networks consist of processing units, such as nodes or neurons, organized into layers. Each processing unit receives signals from other neurons, integrates them, undergoes transformation, and produces corresponding outcomes. This emulation of neural processes within artificial systems contributes to the effectiveness of deep learning methodologies. There are various types of neural networks designed for specific tasks and applications. Some common types include: Feedforward Neural Networks (FNN) operate unidirectionally, conveying information from input to output via hidden layers. Recurrent Neural Networks (RNN) specialize in sequential data, maintaining information through recurrent connections. Convolutional Neural Networks (CNN) excel in image processing, using convolutional layers to identify patterns and spatial hierarchies. Long Short-Term Memory Networks (LSTM), a variant of RNN, address the vanishing gradient issue in extended sequences [58].

Table 2: Some Studies for Machine Learning based detection of depression

Ref	Data Domain	ML Approach	Result	Limitations
[59]	A 106-question survey was created, distributed to Bangladeshi university students, yielding 684 responses in total.	Created a hybrid depression scale with a voting algorithm using eight scales, applying 10 ML and 2 DL models for prediction.	Automatic depression assessment system yields 98.08%, 94.23%, and 92.31% accuracies with Random Forest, Gradient Boosting, and CNN.	Limited generalizability due to focusing on Bangladesh university students, excluding diverse age groups.
[60]	DAIC-WoZ database: contains depressed patient responses, including audio, video, and questionnaire texts from depressed and non-depressed.	The paper suggests a hybrid depression detection model, integrating text and audio features through specialized CNN and LSTM components.	Hybrid depression detection model, combining deep learning, achieved 98% audio CNN accuracy, surpassing 92% textual CNN.	The paper lacks discussion on the hybrid model's generalizability beyond DAIC-WoZ, omitting dataset details affecting results' applicability.
[61]	The study utilizes campus social platform data to identify depression, forming an experimental corpus through preprocessing.	Deep mining analyzes campus social platform data for depression using Python, employing feature selection and LIBSVM tool.	DI-CNN, SVM, and CNN algorithms were tested; SVM outperformed, revealing shallow features' enhanced impact on depression recognition.	Paper notes varied impact of features on depression recognition, suggesting weighting in future research. CNN limitations highlighted.
	Collected data from Kuala Terengganu institute students,	Used ML algorithms: Decision Tree, Neural	Decision Tree excelled in stress, Support Vector	Analyzed mental health in college students,

[62]	determining stress, depression, and anxiety levels using DASS-21 scores.	Network, SVM, Naïve Bayes, logistic regression for classification and prediction.	Machine in depression, and Neural Network in anxiety prediction. Visualization depicted student distribution.	recognizing unexplored aspects; employed machine learning with limitations.
[63]	Used 4 datasets from US public universities, representing diverse regions, with 20,000+ student enrollments.	Study includes ML models like Naïve Bayes, SVM, KNN, Logistic Regression, Decision Tree, Random Forest, XGBoost, NGBoost.	XGBoost excelled in all datasets, showing superb accuracy over 0.99. Social support, learning environment, and childhood adversities strongly impacted mental health.	The research examined survey-based behavior data. Future studies may include diverse data sources for a comprehensive evaluation of students' mental health.
[64]	Paper explores brain study in psychiatric disorders using MRI, EEG, and kinesics diagnosis techniques.	AI analyzes disease data, brain imaging, using methods like multitask-multimodal learning, classification, kernel, and deep learning.	Research emphasizes better machine learning for EEG data, explores kinesics for psychiatric diagnosis.	The paper lacks a thorough examination of AI app performance versus traditional psychiatric diagnostic methods in accuracy and effectiveness.
[65]	The paper utilizes anatomical and physiological data acquired from neuroimaging to create models for studying depression using machine learning methods.	Machine learning, like SVM and regression, is used for studying depression through imaging data analysis.	Few studies used parameter selection to optimize. Linear learning suits when samples < features. Larger datasets benefit from unsupervised methods.	Past studies had limited sample size, hindering optimal machine learning and generalizability of findings.
[66]	Surveyed employed and unemployed individuals globally via DASS 21. Analyzed Twitter data to identify potential suicidal tweets.	Used five machine learning methods to forecast anxiety, depression, and stress across varying severity levels.	Random Forest best predicted anxiety, depression, and stress levels among various classifiers for the study.	Paper overlooked ethical concerns in using machine learning for mental health prediction, neglecting algorithm effectiveness and generalizability.
[67]	Study used Dutch data with 11,081 cases, mainly 570 self-reported depression cases, posing class imbalance.	XGBoost algorithm used on data with various resampling methods to handle class imbalance issue.	XGBoost and biomarkers may improve depression diagnosis, surpassing traditional interviews with efficiency and accuracy.	Limited to Dutch citizens, the study's results may not apply broadly to diverse groups or nationalities.
[68]	AVEC2013 database from AVDLC comprises 340 clips of 292 subjects, labeled using BDI-II, divided into three sets.	Paper fuses manual and deep-learned features to gauge depression severity, incorporating LLD, MRELBP, and DCNN methods.	Paper contrasts LBP and MRELBP features, evaluating handcrafted versus deep-learned features for depression prediction on AVEC2014.	The paper's experiments, confined to AVEC2013 and AVEC2014 depression databases, restrict broader applicability.
[69]	Enrollment data for 637 individuals with IMID, including IBD, MS, and RA, were utilized.	LR, NN, RF models tested; AUC and Brier scores assessed; ten-fold cross-validation used.	LR model: 4 PROM items had AUC 0.91 for MDD; 2 PROM items had AUC 0.83 for anxiety disorder. Sensitivity/specificity: MDD 0.94/0.75; anxiety disorder 0.84/0.70.	Study on immune-mediated disease, small sample, self-reported outcomes, limits generalizability and reliability.
[70]	Survey gathered data from 604 Bangladeshi citizens of various ages with 55 questions.	Evaluate AdaBoost, Bagging, GB, XGBoost, Weighted Voting, and KNN with dataset. Use SelectKBest, mRMR, Boruta for feature selection.	AdaBoost, SelectKBest scored 92.56% predicting depression; Bagging, GB, XGBoost, Weighted Voting, 85.95-89.26%.	Paper lacks discussion on SMOTE limitations, and evaluation metrics aren't specified.
[71]	Data from mental health app articles was analyzed for ML, type, size. Study sizes unspecified.	Studies applied ML techniques: SVMs, decision trees, neural networks, LDA, and clustering.	ML aids mental health in diagnosis, treatment, research, and clinical administration, showing significant benefits.	Paper lacked thorough analysis on ML technique limitations: accuracy, generalizability, and potential biases.
[72]	671 participants' audio, video, speech responses, and PHQ-9	ML sorts data, explores ensembles, detects	Paper outlines depression diagnosis: data extraction,	Paper lacks comprehensive ML algorithm comparison

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.o

	data.	depression via extraction, pre-processing, and feature extraction.	ML training, classification, and future research.	for depression, neglects ethical concerns, dataset analysis.
[73]	Analyzed user emotions on social networks using publicly sourced Facebook data for insights.	Used machine learning (DT, SVM, KNN, Ensemble) to detect depression in Facebook data.	Study classifiers detected depression with 60-80% accuracy, showcasing varied but promising results.	Paper analyzes Facebook data for depression without considering broader social network diversity.
[74]	Studied university students' depression using PHQ-9 and Zung SDS, with additional scales like GAD and self-rating anxiety.	Used KNN, RF, SVM, LR, and LDA models for analysis and prediction in their study.	RF achieves 91.58% accuracy, 86.05% sensitivity, 96.15% specificity, and AUC of 0.8994 using 8 features.	Study limited to Chinese postgraduates; findings may not apply to other populations due to lacking demographic data.
[75]	Reddit posts were analyzed using NLP techniques like Tokenization and stemming for language usage assessment.	SVM and MLP classifiers assess the proposed method's performance, utilizing machine learning techniques in evaluation.	Combining LIWC, LDA, and bigram features with MLP classifier achieves top performance: 91% accuracy, 0.93 F1 score.	Study focuses on Reddit, might not reflect population, neglects biases in machine learning for depression detection.

7. LIMITATIONS AND CHALLENGES

Using machine learning for diagnosing behavior has come a long way, but there are still big challenges. One major issue is that different techniques aren't always accurate, and they use different algorithms and data. We need more research to combine findings and create standard techniques that mental health professionals can trust. These standard methods could help researchers, data scientists, and clinicians work together better, which is really important for making these tools useful in real-life clinical settings. Right now, most machine learning research happens in labs, not real clinical settings. To really test how useful these tools are, we need more studies in real clinics. For instance, researchers should compare how well clinicians diagnose patients with how well machine learning does, then see if the treatment matches up with the diagnosis. One problem with machine learning-based diagnosis is that they're compared to biased and sometimes inaccurate questionnaires and clinical diagnoses. We still don't have clear standards for diagnosing things like depression and other mental health issues, so creating these standards would help us compare new methods like machine learning more reliably. Even though machine learning models are often accurate, they can still be hard to understand because they don't explain why they make certain decisions. They just take in data and give an answer without showing how they pick certain features or patterns. It's important for doctors to trust the predictions made by machine learning tools, especially when diagnosing patients. Future research should involve both computer scientists and psychologists to make machine learning models easier to understand and trust. After reviewing all the literature and analyzing the results, several points have been highlighted. Various methodologies, outcomes, benefits, and challenges have been discerned in the examination of mental illness, encompassing the utilization of datasets. Machine learning methodologies can be effectively applied in conjunction with social networking platforms for the identification of users exhibiting social behaviors. Notably, algorithms such as Support Vector Machines (SVM), multivariate regression, and Radial Basis Function Networks (RBFN) demonstrate superior predictive capabilities for depression when compared to features based on volume. Furthermore, hybrid models that integrate Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) exhibit enhanced efficiency in the detection of depression. Additionally, the employment of machine learning and boosting algorithms proves instrumental in extracting sociodemographic and psychological factors contributing to the onset of depression. Voice change studies may aid in depression detection. Personalized evaluations and interventions show promise in clinical applications. Adjusting tuning parameters can improve model effectiveness. Some studies only focus on sadness, limiting comparisons with other depression symptoms. Causality is hard to determine in cross-sectional studies. Variation in participant demographics across studies affects generalizability. Smartphones assessments are more reliable than self-reporting. Some studies lack consideration of depression history. Integration of various forecasting capabilities remains limited. These findings also indicate limitations. Many studies use default parameters and a limited number of cases, possibly leading to biased results. Some studies collect data from a single institution or country, limiting generalizability. Parent-reported data may be biased, and interpersonal relationships of respondents are often overlooked. Some studies focus solely on self-reporting measures, neglecting interviews. It's challenging to determine causality in cross-sectional studies. Additional research conducted by medical professionals is required, coupled with the identification and assessment of internal and external factors of influence, in order to augment the accuracy of diagnoses.

8. FUTURE DIRECTIONS

In this section, we discuss potential future research directions based on previous studies we reviewed. Most studies on detecting depression have used small sample sizes. While small samples help build prediction models, larger samples are crucial for creating accurate models applicable to the wider population. Training models with large samples allows for more diverse representation of depressed patients, potentially resulting in models with therapeutic value. As studies incorporate larger datasets, methods are likely to evolve, showing improved validation metrics. Techniques like k-fold cross-validation with higher k-values can be employed to enhance model testing and generalizability. Selecting appropriate learning techniques is essential, with unlabeled data sometimes aiding in developing prediction models for large datasets with sparse data. Identifying the nature of incoming data (labeled, unlabeled, or a mix) guides the choice between supervised, unsupervised, or semi-supervised learning techniques. Considering the dataset's linearity helps prevent overfitting with small datasets and ensures effective handling of nonlinear data in larger datasets. Long-term objectives focus on improving predictive model accuracy. While current methods like Support Vector Machines (SVM) are reliable, future research aims to enhance accuracy and clinical relevance. Collaborations across disciplines such as psychology, physiology, computer science, and machine learning are vital for advancing Affective Disorder Estimation (ADE). Researchers should leverage diverse strengths to develop multimodal deep learning approaches for clinical applications. Challenges include limited availability of depression data and databases with varied modalities (audio, video, text, physiological signals). Existing databases like AVEC2013 and AVEC2014 are limited in size and modality coverage. Standardizing data collection criteria is crucial for ensuring consistency across studies. Additionally, methods for augmenting limited annotated data are needed, especially for deep learning applications. Standardizing data collection protocols and increasing dataset sizes are essential for advancing depression prediction research.

9. CONCLUSION

This research paper undertakes a comprehensive exploration of mental health issues through the application of machine intelligence methods, employing a systematic review and meta-analysis. The study delves deeply into the intricacies of machine intelligence, elucidating its applications, advantages, and limitations. Notably, machine learning (ML) techniques play a pivotal role in diagnosing a spectrum of mental health conditions, including PTSD, schizophrenia, depression, ASD, and bipolar disorders. Diverse data sources such as social media, clinical records, and mobile sensor data are investigated for mood disorder detection. The primary focus of our research centers on surveying contemporary investigations into ML-driven approaches for diagnosing depression. The goal is to elucidate the fundamental ML principles relevant to mental health, particularly depression, and discuss their practical implementation. The findings underscore the significant potential of ML in enhancing mental healthcare, facilitating predictive modeling, and enabling tailored interventions, with a specific emphasis on depression. Our examination extends to publication trends, mapping aspects, current scenarios, and computational considerations. The synthesis of literature suggests that machine intelligence holds promise for the timely detection of mental health issues and depression. However, we acknowledge substantial weaknesses in resource availability and treatment gaps, particularly in developing countries, where many patients face challenges in accessing care. Contributing factors identified include the absence of mental health policies, insufficient awareness, and societal stigma. Highlighting the importance of understanding the association between symptoms and internal/external factors, we propose validating methods through simulation to efficiently handle real events. Additionally, we recommend analyzing the association between risk factors and mental health problems for early-stage prediction. Recognizing the macroscopic nature of mental disorders, we stress the significance of proper intervention for prevention and prediction. We advocate for refining our understanding by considering the addition or removal of symptoms affecting causes and associated risks. In conclusion, our research underscores the potential of machine intelligence in revolutionizing mental healthcare, while acknowledging existing challenges and proposing avenues for future research and intervention. By embracing technology and adopting comprehensive approaches, we can enhance the diagnosis, treatment, and care for individuals experiencing mental health issues, ultimately improving their well-being and quality of life..

REFERENCE

[1] G. Malhi and J. Mann, “Depression,” Lancet, vol. 392, Nov. 2018, doi: 10.1016/S0140-6736(18)31948-2.

[2] D. D. Ebert et al., “Prediction of major depressive disorder onset in college students,” Depress. Anxiety, vol. 36, no. 4, pp. 294–304, 2019, doi: 10.1002/da.22867.

[3] Unicef, “The State of the World’s Children 2021,”, [Online]. Available: https://www.unicef.org/reports/state-worlds-children-2021

[4] A. K. Ibrahim, S. J. Kelly, C. E. Adams, and C. Glazebrook, “A systematic review of studies of depression prevalence in university students,” J. Psychiatr. Res., vol. 47, no. 3, pp. 391–400, 2013, doi: https://doi.org/10.1016/j.jpsychires.2012.11.015.

[5] J. P. Read, M. D. Wood, O. J. Davidoff, J. McLacken, and J. F. Campbell, “Making the Transition from High School to College: The Role of Alcohol–Related Social Influence Factors in Students’ Drinking,” Subst. Abus., vol. 23, no. 1, pp. 53–65, Mar. 2002, doi: 10.1080/08897070209511474.

[6] D. Coyle, A. Thieme, C. Linehan, M. Balaam, J. Wallace, and S. Lindley, “Emotional Wellbeing,” Int. J. Hum. Comput. Stud., vol. 72, no. 8, pp. 627–628, 2014, doi: https://doi.org/10.1016/j.ijhcs.2014.05.008.

[7] D. M. Hutchinson, A. B. R. Shatte, and S. J. Teague, “Machine learning in mental health: a scoping review of methods and applications,” Psychol. Med., vol. 49, no. 9, pp. 1426–1448, 2019, doi: DOI: 10.1017/S0033291719000151.

[8] M. Adamou, V. Lagani, G. Antoniou, P. Charonyktakis, E. Greasidou, and I. Tsamardinos, “Mining free-text medical notes for suicide risk assessment,” ACM Int. Conf. Proceeding Ser., 2018, doi: 10.1145/3200947.3201020.

[9] J. Aguilar-Ruiz, R. Costa, and F. Divina, Knowledge discovery from doctor-patient relationship, vol. 1. 2004. doi: 10.1145/967900.967960.

[10] B. Cao et al., “DeepMood: Modeling mobile phone typing dynamics for mood detection,” Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., vol. Part F1296, pp. 747–755, 2017, doi: 10.1145/3097983.3098086.

[11] O. DeMasi and B. Recht, “A step towards quantifying when an algorithm can and cannot predict an individual’s wellbeing,” UbiComp/ISWC 2017 - Adjun. Proc. 2017 ACM Int. Jt. Conf. Pervasive Ubiquitous Comput. Proc. 2017 ACM Int. Symp. Wearable Comput., pp. 763–771, 2017, doi: 10.1145/3123024.3125609.

[12] D. Galiatsatos, G. Konstantopoulou, G. Anastassopoulos, M. Nerantzaki, K. Assimakopoulos, and D. Lymberopoulos, “Classification of the most Significant Psychological Symptoms in Mental Patients with Depression using Bayesian Network,” ACM Int. Conf. Proceeding Ser., vol. 2015-Janua, no. October 2016, 2015, doi: 10.1145/2797143.2797159.

[13] A. L. Nobles, J. J. Glenn, K. Kowsari, B. A. Teachman, and L. E. Barnes, “Identification of Imminent Suicide Risk Among Young Adults using Text Messages,” in Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, in CHI ’18. New York, NY, USA: Association for Computing Machinery, 2018, pp. 1–11. doi: 10.1145/3173574.3173987.

[14] M. Rabbi, S. Ali, T. Choudhury, and E. Berke, “Passive and in-situ assessment of mental and physical well-being using mobile sensors,” UbiComp’11 - Proc. 2011 ACM Conf. Ubiquitous Comput., pp. 385–394, 2011, doi: 10.1145/2030112.2030164.

[15] K. Saha and M. De Choudhury, “Modeling stress with social media around incidents of gun violence on college campuses,” Proc. ACM Human-Computer Interact., vol. 1, no. CSCW, 2017, doi: 10.1145/3134727.

[16] D. Spathis, S. Servia-Rodriguez, K. Farrahi, C. Mascolo, and J. Rentfrow, “Passive mobile sensing and psychological traits for large scale mood prediction,” PervasiveHealth Pervasive Comput. Technol. Healthc., pp. 272–281, 2019, doi: 10.1145/3329189.3329213.

[17] K. Vaishnavi, U. N. Kamath, B. A. Rao, and N. V. S. Reddy, “Predicting Mental Health Illness using Machine Learning Algorithms,” J. Phys. Conf. Ser., vol. 2161, no. 1, 2022, doi: 10.1088/1742-6596/2161/1/012021.

[18] T. Trany, D. Phung, W. Luo, R. Harvey, M. Berk, and S. Venkatesh, “An integrated framework for suicide risk prediction,” Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., vol. Part F1288, pp. 1410–1418, 2013, doi: 10.1145/2487575.2488196.

[19] R. Wang et al., “Studentlife: Assessing mental health, academic performance and behavioral trends of college students using smartphones,” UbiComp 2014 - Proc. 2014 ACM Int. Jt. Conf. Pervasive Ubiquitous Comput., pp. 3–14, 2014, doi: 10.1145/2632048.2632054.

[20] C. Zakaria, R. Balan, and Y. Lee, “Stressmon: Scalable detection of perceived stress and depression using passive sensing of changes in work routines and group interactions,” Proc. ACM Human-Computer Interact., vol. 3, no. CSCW, pp. 1–29, 2019, doi: 10.1145/3359139.

[21] X. Tao, O. Chi, P. J. Delaney, L. Li, and J. Huang, “Detecting depression using an ensemble classifier based on Quality of Life scales,” Brain Informatics, vol. 8, no. 1, 2021, doi: 10.1186/s40708-021-00125-5.

[22] M. Gil, S. S. Kim, and E. J. Min, “Machine learning models for predicting risk of depression in Korean college students: Identifying family and individual factors,” Front. Public Heal., vol. 10, no. November, pp. 1–10, 2022, doi: 10.3389/fpubh.2022.1023010.

[23] P. Jain, K. R. Srinivas, and A. Vichare, “Depression and Suicide Analysis Using Machine Learning and NLP,” J. Phys. Conf. Ser., vol. 2161, no. 1, 2022, doi: 10.1088/1742-6596/2161/1/012034.

[24] D. Su, X. Zhang, K. He, and Y. Chen, “Use of machine learning approach to predict depression in the elderly in China: A longitudinal study,” J. Affect. Disord., vol. 282, pp. 289–298, 2021, doi: 10.1016/j.jad.2020.12.160.

[25] M. S. Zulfiker, N. Kabir, A. A. Biswas, T. Nazneen, and M. S. Uddin, “An in-depth analysis of machine learning approaches to predict depression,” Curr. Res. Behav. Sci., vol. 2, no. April, 2021, doi: 10.1016/j.crbeha.2021.100044.

[26] P. Debra Fulghum Bruce, “Depression.” [Online]. Available: https://www.webmd.com/depression/what-is-depression

[27] S. and B. W. Watson, “‘Depression,’” Gale Health and Wellness Online Collection, Gale, 2022. [Online]. Available: https://www.gale.com/open-access/depression

[28] F. Benazzi, “Various forms of depression,” Dialogues Clin. Neurosci., vol. 8, no. 2, pp. 151–161, 2006, doi: 10.31887/dcns.2006.8.2/fbenazzi.

[29] “Depression Assessment Instruments.” [Online]. Available: https://www.apa.org/depression-guideline/assessment

[30] M. Awad and R. Khanna, “Efficient learning machines: Theories, concepts, and applications for engineers and system designers,” Effic. Learn. Mach. Theor. Concepts, Appl. Eng. Syst. Des., no. April 2015, pp. 1–248, 2015, doi: 10.1007/978-1-4302-5990-9.

[31] G. Guo, H. Wang, D. Bell, Y. Bi, and K. Greer, “KNN model-based approach in classification,” Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 2888, no. November 2012, pp. 986–996, 2003, doi: 10.1007/978-3-540-39964-3_62.

[32] A. Cutler, D. R. Cutler, and J. R. Stevens, “Ensemble Machine Learning,” Ensemble Mach. Learn., no. January, 2012, doi: 10.1007/978-1-4419-9326-7.

[33] M. Maalouf, “Logistic regression in data analysis: An overview,” Int. J. Data Anal. Tech. Strateg., vol. 3, no. 3, pp. 281–299, 2011, doi: 10.1504/IJDATS.2011.041335.

[34] S. Cycles, “Chapter 9 Chapter 9,” Cycle, vol. 1897, no. Figure 1, pp. 44–45, 1989, doi: 10.1007/0-387-25465-X.

[35] C. Bentéjac, A. Csörgő, and G. Martínez-Muñoz, “A Comparative Analysis of XGBoost,” no. November 2019, 2019, doi: 10.1007/s10462-020-09896-5.

[36] Rish I, “An empirical study of the naive bayes classifier,” IJCAI 2001 Work. Empir. methods Artif. Intell., no. January 2001, pp. 41–46, 2001, [Online]. Available: http://www.cc.gatech.edu/home/isbell/classes/reading/papers/Rish.pdf

[37] F. A. Nazira, S. R. Das, S. A. Shanto, and M. F. Mridha, “Depression Detection Using Convolutional Neural Networks,” Proc. 2021 IEEE Int. Conf. Signal Process. Information, Commun. Syst. SPICSCON 2021, no. February 2023, pp. 9–13, 2021, doi: 10.1109/SPICSCON54707.2021.9885517.

[38] J. Singh, M. Wazid, D. P. Singh, and S. Pundir, “An embedded LSTM based scheme for depression detection and analysis,” Procedia Comput. Sci., vol. 215, pp. 166–175, 2022, doi: 10.1016/j.procs.2022.12.019.

[39] A. Yassine, C. Mohamed, and A. Zinedine, “Feature selection based on pairwise evalution,” 2017 Intell. Syst. Comput.

Vis., pp. 1–6, 2017, [Online]. Available: https://api.semanticscholar.org/CorpusID:24459212

[40] Y. AKHIAT, M. CHAHHOU, and A. ZINEDINE, "Ensemble Feature Selection Algorithm," Int. J. Intell. Syst. Appl., vol. 11, no. 1, pp. 24–31, 2019, doi: 10.5815/ijisa.2019.01.03.

[41] Y. Akhiat, Y. Asnaoui, M. Chahhou, and A. Zinedine, "A new graph feature selection approach," 2020 6th IEEE Congr. Inf. Sci. Technol., pp. 156–161, 2020, [Online]. Available: https://api.semanticscholar.org/CorpusID:233136414

[42] G. H. John, R. Kohavi, and K. Pfleger, "Irrelevant Features and the Subset Selection Problem," W. W. Cohen and H. B. T.-M. L. P. 1994 Hirsh, Eds., San Francisco (CA): Morgan Kaufmann, 1994, pp. 121–129. doi: https://doi.org/10.1016/B978-1-55860-335-6.50023-4.

[43] M. Ghosh, R. Guha, R. Sarkar, and A. Abraham, "A wrapper-filter feature selection technique based on ant colony optimization," Neural Comput. Appl., vol. 32, no. 12, pp. 7839–7857, 2020, doi: 10.1007/s00521-019-04171-3.

[44] I. Guyon and A. Elisseeff, "An Introduction of Variable and Feature Selection," J. Mach. Learn. Res. Spec. Issue Var. Featur. Sel., vol. 3, pp. 1157–1182, Jan. 2003, doi: 10.1162/153244303322753616.

[45] S. Okser, T. Pahikkala, and T. Aittokallio, "Genetic variants and their interactions in disease risk prediction – machine learning and network perspectives," BioData Min., vol. 6, no. 1, p. 5, 2013, doi: 10.1186/1756-0381-6-5.

[46] A. Tharwat, T. Gaber, A. Ibrahim, and A. E. Hassanien, "Linear discriminant analysis: A detailed tutorial," AI Commun., vol. 30, no. 2, pp. 169–190, 2017, doi: 10.3233/AIC-170729.

[47] N. Chawla, K. Bowyer, L. Hall, and W. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," J. Artif. Intell. Res., vol. 16, pp. 321–357, Jun. 2002, doi: 10.1613/jair.953.

[48] J. Pennebaker, M. Francis, and R. Booth, "Linguistic inquiry and word count (LIWC)," Jan. 1999.

[49] M. Pietrzykowski and W. Sałabun, "Applications of Hidden Markov Model: state-of-the-art," Jul. 2014.

[50] M. Ayyanar, S. Jeganathan, S. Parthasarathy, V. Jayaraman, and A. R. Lakshminarayanan, "Predicting the Cardiac Diseases using SelectKBest Method Equipped Light Gradient Boosting Machine," in 2022 6th International Conference on Trends in Electronics and Informatics (ICOEI), 2022, pp. 117–122. doi: 10.1109/ICOEI53556.2022.9777224.

[51] P. Bugata and P. Drotar, "On some aspects of minimum redundancy maximum relevance feature selection," Sci. China Inf. Sci., vol. 63, no. 1, p. 112103, 2019, doi: 10.1007/s11432-019-2633-y.

[52] M. B. Kursa, A. Jankowski, and W. R. Rudnicki, "Boruta - A system for feature selection," Fundam. Informaticae, vol. 101, no. 4, pp. 271–285, 2010, doi: 10.3233/FI-2010-288.

[53] M. Robnik-Sikonja and I. Kononenko, "Theoretical and Empirical Analysis of ReliefF and RReliefF," Mach. Learn., vol. 53, pp. 23–69, Oct. 2003, doi: 10.1023/A:1025667309714.

[54] Q. Liu and Y. Wu, "Supervised Learning," Jan. 2012, doi: 10.1007/978-1-4419-1428-6_451.

[55] S. Naeem, A. Ali, S. Anam, and M. M. Ahmed, "An Unsupervised Machine Learning Algorithms: Comprehensive Review," Int. J. Comput. Digit. Syst., vol. 13, no. 1, pp. 911–921, 2023, doi: 10.12785/ijcds/130172.

[56] P. Sarang, "ANN-Based Applications," pp. 289–327, 2023, doi: 10.1007/978-3-031-02363-7_18.

[57] I. H. Sarker, "Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions," SN Comput. Sci., vol. 2, no. 6, pp. 1–20, 2021, doi: 10.1007/s42979-021-00815-1.

[58] A.-N. Sharkawy, "Principle of Neural Network and Its Main Types: Review," J. Adv. Appl. Comput. Math., vol. 7, pp. 8–19, Aug. 2020, doi: 10.15377/2409-5761.2020.07.2.

[59] R. Siddiqua, N. Islam, J. F. Bolaka, R. Khan, and S. Momen, "AIDA: Artificial intelligence based depression assessment applied to Bangladeshi students," Array, vol. 18, no. February, p. 100291, 2023, doi: 10.1016/j.array.2023.100291.

[60] Vandana, N. Marriwala, and D. Chaudhary, "A hybrid model for depression detection using deep learning," Meas. Sensors, vol. 25, no. November 2022, p. 100587, 2023, doi: 10.1016/j.measen.2022.100587.

[61] G. F. Zhao and L. F. Sun, "Depression Identification of Students Based on Campus Social Platform Data and Deep Learning," Sci. Program., vol. 2022, 2022, doi: 10.1155/2022/6532384.

[62] S. M. Et. al., "Mental Health Prediction Models Using Machine Learning in Higher Education Institution," Turkish J. Comput. Math. Educ., vol. 12, no. 5, pp. 1782–1792, 2021, doi: 10.17762/turcomat.v12i5.2181.

[63] P. Muzumdar, G. P. Basyal, and P. Vyas, "An Empirical Comparison of Machine Learning Models for Student's Mental Health Illness Assessment," Asian J. Comput. Inf. Syst., vol. 10, no. 1, pp. 1–10, 2022, doi: 10.24203/ajcis.v10i1.6882.

[64] G. Di Liu, Y. C. Li, W. Zhang, and L. Zhang, "A Brief Review of Artificial Intelligence Applications and Algorithms for Psychiatric Disorders," Engineering, vol. 6, no. 4, pp. 462–467, 2020, doi: 10.1016/j.eng.2019.06.008.

[65] M. J. Patel, A. Khalaf, and H. J. Aizenstein, "Studying depression using imaging and machine learning methods," NeuroImage Clin., vol. 10, pp. 115–123, 2016, doi: 10.1016/j.nicl.2015.11.003.

[66] K. Pugazharasi, P. Kalaivani, J. Jayapriya, S. Kadhambari, and S. Mailvizhi, "Machine Learning Algorithms for Predicting Depression, Anxiety and Stress in Modern Life," AIP Conf. Proc., vol. 2587, no. 1, pp. 1258–1267, 2023, doi: 10.1063/5.0150604.

[67] A. Sharma and W. J. M. I. Verbeke, "Improving Diagnosis of Depression With XGBOOST Machine Learning Model and a Large Biomarkers Dutch Dataset (n = 11,081)," Front. Big Data, vol. 3, no. April, pp. 1–11, 2020, doi: 10.3389/fdata.2020.00015.

[68] L. He and C. Cao, "Automated depression analysis using convolutional neural networks from speech," J. Biomed. Inform., vol. 83, no. May, pp. 103–111, 2018, doi: 10.1016/j.jbi.2018.05.007.

[69] L. G. Tennenhouse, R. A. Marrie, C. N. Bernstein, and L. M. Lix, "Machine-learning models for depression and anxiety in individuals with immune-mediated inflammatory disease," J. Psychosom. Res., vol. 134, no. September 2019, p. 110126, 2020, doi: 10.1016/j.jpsychores.2020.110126.

[70] M. S. Zulfiker, N. Kabir, A. A. Biswas, T. Nazneen, and M. S. Uddin, “An in-depth analysis of machine learning approaches to predict depression,” Curr. Res. Behav. Sci., vol. 2, no. February, p. 100044, 2021, doi: 10.1016/j.crbeha.2021.100044.

[71] S. T. A. Shatte, D. Hutchinson, “Machine learning in mental health: A systematic scoping review of methods and applications Adrian B. R. Shatte*,” 2021.

[72] S. Aleem, N. Huda, R. Amin, S. Khalid, S. S. Alshamrani, and A. Alshehri, “Machine Learning Algorithms for Depression : Diagnosis ,” Electronics, 2022.

[73] M. R. Islam, M. A. Kabir, A. Ahmed, A. R. M. Kamal, H. Wang, and A. Ulhaq, “Depression detection from social network data using machine learning techniques,” Heal. Inf. Sci. Syst., vol. 6, no. 1, pp. 1–12, 2018, doi: 10.1007/s13755-018-0046-0.

[74] J. Fang et al., “Depression Prevalence in Postgraduate Students and Its Association with Gait Abnormality,” IEEE Access, vol. 7, pp. 174425–174437, 2019, doi: 10.1109/ACCESS.2019.2957179.

[75] M. M. Tadesse, H. Lin, B. Xu, and L. Yang, “Detection of depression-related posts in reddit social media forum,” IEEE Access, vol. 7, no. c, pp. 44883–44893, 2019, doi: 10.1109/ACCESS.2019.2909180.