JOURNAL MEDICAL INFORMATICS TECHNOLOGY
medinftech.org ISSN 2988-7003 (online)

Evaluating the Role of Data Modalities in Machine Learning Models for Psychiatric Disorder Diagnosis: A Review

Hilal Çalışkan[1] and İlkim Ecem Emre[2]*

[1,2] *Marmara University, Turkey*

ARTICLE HISTORY

Received: 01 April 25
Final Revision: 09 May 25
Accepted: 25 May 25
Online Publication: 30 June 25

KEYWORDS

Bipolar Disorder, Data Modalities, Depression, Machine Learning, Psychiatric Diagnosis

CORRESPONDING AUTHOR

ecem.emre@marmara.edu.tr

DOI

10.37034/medinftech.v3i2.106

A B S T R A C T

The increasing prevalence of psychiatric disorders such as depression, bipolar disorder, and post-traumatic stress disorder has drawn attention to the need for more efficient and accurate diagnostic tools. In this context, machine learning offers promising solutions by enabling the analysis of complex and high-dimensional data. This study aims to evaluate the diagnostic performance of ML models applied to various psychiatric disorders by comparing the effectiveness of different data types such as EEG, MRI, video and audio recordings, photographs, survey responses, and clinical data. A total of 44 scientific studies published between 2015 and 2024 were systematically reviewed in accordance with PRISMA 2020 guidelines. The studies included applied ML or deep learning models to adult participants. The results show that the most successful data types varied by disorder. In conclusion, the choice of data type significantly influences the performance of ML models in psychiatric diagnosis. EEG, survey, and clinical data emerged as the most reliable across different conditions, while SVM, Random Forest, and CNN-based models provided the best classification results. These findings offer a valuable reference point for future research and the development of AI-assisted diagnostic tools.

1. Introduction

Today, with the development of technology, artificial intelligence (AI) has become popular in many fields and it has great advantages for different sectors. In particular, machine learning (ML), which provides a wide range of techniques, is frequently used in many different fields such as production, education, healthcare, finance, telecommunications. The one of the most important advantages of AI is that it minimizes the need for labor and reduces time for certain tasks. Thanks to AI and ML, the work done by humans has become possible under the control of AI with automation and robotic technologies. This has prevented people from doing simple and repetitive work, allowing people to turn to more creative and strategic work. Another important benefit is the ability to analyze. Especially, in the field of healthcare, ML techniques can accurately analyze medical test results such as Magnetic Resonance Imaging (MRI) and electroencephalogram (EEG). By teaching the machine or computer the decisions that an experienced doctor can normally make, it is possible to access these experienced and accurate decisions anywhere and anytime. Another important advantage is the ability to predict the future. Many sectors use these prediction methods as support in many important areas such as the decisions they will make, the products they will produce, the advertising strategies they will create, so that they can see the future more clearly and plan their next steps accordingly. This can also be applied to healthcare. Prediction of diseases can save lives of people and can reduce time and cost for the treatments with early prediction.

One of the fields where AI and ML can be used is psychiatry. Psychiatry is a branch of medicine that deals with the diagnosis and treatment of psychiatric disorders, examining deficiencies in people's cognitive and emotional abilities and deficiencies in their adaptation to the environment. Psychiatric disorders have become one of the most common problems faced by modern society, yet the field of psychiatry has developed considerably, especially in the last 30-40 years.

ML can help this field by predicting outcomes and personalizing psychiatric treatment [1]. Since individualized treatment recommendations and biomarkers for mental diseases are still lacking, ML technology and data analytics can be applied to several phases of a patient's journey, such as prediction, treatment selection and optimization, outcome monitoring and tracking, detection and diagnosis, and relapse prevention [2]. Psychiatric disorders are mental health problems that deeply affect the quality of life of individuals and their relationships with the environment, leading to disorders of emotions and thoughts that need to be treated. ML methods can be useful in analyzing the increasing amount of data in the field of psychiatry [3].

There are different types of psychiatric disorders include depression, anxiety, bipolar disorder, schizophrenia, post-traumatic stress disorder and obsessive-compulsive disorder [4]. The treatment of these psychiatric disorders is of great importance first for the health of the individual and then for the mental health of the community.

AI and ML are frequently used in the field of psychiatry to diagnose, classify and predict diseases. ML contributes greatly to the field of psychiatry with its benefits in many different areas such as early diagnosis, personalized treatment, disease prediction according to genetic characteristics [5]. Various methods of ML allow clinics to make faster decisions and intervene before the disease condition worsens. For these purposes, many different data types can be used when developing different ML models. Some of these data types are survey, EEG, MRI, text data, video, images, audio. According to the relevant literature, different data types could give different successful results. The aim of this study is to review the existing literature and comparatively evaluate the performance of different data types and machine learning models in the detection of psychiatric disorders.

In this context, the study seeks to explore which psychiatric disorders give more successful outcomes with specific data types, and whether the integration of different data types enhances the performance of ML models. Accordingly, the study aims to contribute to the existing literature on disease-based ML modeling in the field of psychiatry by examining the role of data type selection in model success. In this study, studies on depression, bipolar disorder and post-traumatic stress disorder (PTSD) were examined.

2. Research Method

In this research, ML studies on psychiatric disorders, especially depression, bipolar disorder, PTSD, were analyzed. In this study, according to PRISMA 2020 decision criteria, IEEE database were searched and studies conducted between 2015-2024 were included in the research. The majority of the reviewed studies were from IEEE, only 5 of them is from other sources. They are either full paper conference proceedings or articles which were accessed through IEEE-Xplore. Two studies are from Elsevier, 2 are from Springer, one is from Pubmed directories. No tools were used for the reviewing process, classification of the studies made manually. The keywords “mental disorders”, “machine learning”, “depressive disorder”, “artificial intelligence”, “bipolar disorder” was used for the search. The process is also represented with PRISMA flow diagram in Figure 1.

The inclusion criteria for the studies were as follows: they must have employed machine learning (ML) or deep learning models, been written in English, involved participants aged over 18 years, constituted applied research incorporating machine learning analysis, and provided full access to the study content. Conversely, studies were excluded if they were literature reviews, case studies, or qualitative research.

Fourty-four studies were selected to be included in the study. Of these, 18 studies were on depression, 9 on bipolar disorder, 7 on PTSD and 10 on multiple disorders. Many different data types such as EEG, MRI, text, images, video, audio, survey were used in these studies. Although accuracy is usually taken as the basis for model performance criteria, some studies measure success with metrics such as sensitivity or F1-score.

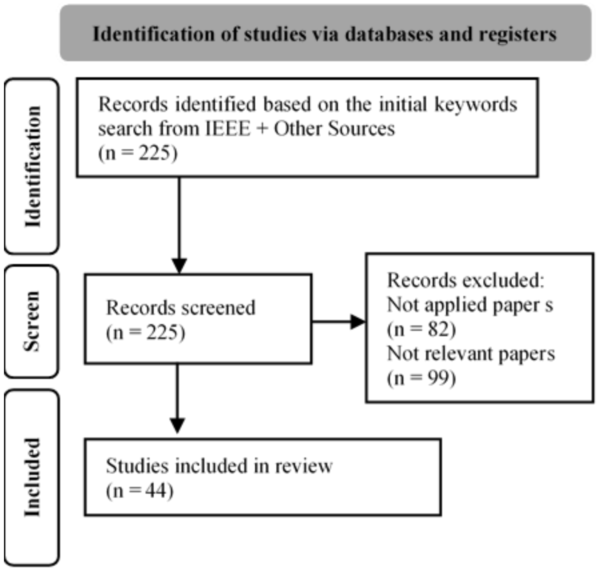

Figure 1. Review process with PRISMA flow diagram

3. Result and Discussion

This section presents the results of several studies that used ML and deep learning models to detect and diagnose psychiatric disorders. These studies focused on conditions like depression, anxiety, PTSD, and bipolar disorder, using different types of data such as MR scans, EEG, surveys, video, text, clinical data, and audio. The findings show how effective these models are at diagnosing these conditions and how the choice of data can improve accuracy, providing helpful insights into how these methods can be used for early detection and treatment.

3.1. Depression & Anxiety

In this review, 18 different studies on depression were examined which were conducted between 2020 and 2024. In all of these articles, different data types such as surveys, video, text, EEG and audio were used. In addition to machine learning models such as Support Vector Machines (SVM), Random Forest (RF), Naive Bayes (NB), some advanced deep learning models were also used.

The first study (Singh & Kumar, 2021) reviewed, attempted to estimate depression using survey data. Five different machine learning models were applied to detect psychiatric disorders, anxiety and depression using DASS-21 questionnaire data. The main purpose of the study is to compare these five machine learning models and identify the most reliable model among them. According to the study, the accuracy rate of Support Vector Machine (SVM) was found to be higher than other algorithms. This study reveals that the SVM model gives more reliable results with %91 accuracy in survey type data [6].

In another study (Nimsarkar & Ingle, 2023) audio recordings were used as data. In this research, it has been shown that depression can be detected in advance by using the DAIC-WOZ dataset consisting of clinical interview audio files. A multi-view Neural Network was developed using feature extraction methods such as Mel Frequency Cepstrum Coefficients (MFCC) and Gammatone Frequency Cepstrum Coefficients (GFCC). In this model, since both feature extraction methods were combined, the accuracy rate (72.9%) was higher than the accuracy provided by either one alone. The study points out that this method needs to be further developed in the future [7].

Another study (Ogur et al., 2023) focused on diagnosing depression through pregnant women. This study applies ML and big data analytics to data collected from 250 pregnant women, including socio-demographic details and anxiety levels. The Naïve Bayes algorithm used in this study achieved an accuracy rate of approximately 90.8%, demonstrating that it is successful in diagnosing anxiety and depression in women in the perinatal period [8].

In a study (Jiang et al., 2024) it was aimed to classify depression with eye movement data collected with eye tracking technology. It aims to make this classification using machine learning methods. Within the scope of the research, the eye movements of 150 people (115 depression patients, 35 control group) were recorded and pupil movements were recorded as a time series. Three basic machine learning models (Logistic regression, K-Nearest Neighbor, Random Forest) were used in the study. The Random Forest model achieved an accuracy rate of 97.5%. According to the study, it was found that patients diagnosed with depression looked at negative stimuli for a longer time, while looking at positive stimuli for a shorter time [9].

A study (Colic et al., 2018) conducted in Canada, aims to predict suicidal thoughts among retired military and police members and to reveal the most important variables that determine suicidal thoughts. Some data, such as demographic information, mental health information, alcohol and substance use, were collected from 738 people through a survey. The collected data were analyzed with Random Forest, and among 224 variables, 25 variables were selected as the most effective variables in suicidal ideation. These variables include feeling unsuccessful, taking medication for anxiety and depression, and physiological disturbances such as nausea and heart palpitations. Obtained accuracy from the model is 84.4% [10].

In a study (Wang & Liu, 2022) to check whether depression can be detected by extracting emotional features from audio data. The study was established by analyzing the audio files collected by asking 76 depressed patients and 81 control groups to vocalize 72 words with positive, negative and neutral meanings in a soundproof room. Deep learning models such as Recurrent Neural Network, Long Short-Term Memory, Transformer were used for analysis in the study, and the Multi-Layer Perceptron model provided the highest value with an accuracy rate of 82.7% [11].

In this study (Shen et al., 2022), two different data sets were used to accurately detect depression. The first data set, contains the audio responses given to questions asked by Chinese volunteers, 30 depression patients and 132 control groups, and the transcripts of these audio responses. The other data set, includes data from 42 depression patients and 100 control groups, and includes both voice and text data collected. For this analysis, it was aimed to develop a more powerful model by using both audio files and text files together. Models such as Gate Recurrent Unit, Bidirectional LSTM, Multi-Modal Fusion were used in the analysis. Multi-Modal Fusion was the model that achieved the most successful F1 score in both datasets. While the model reached an F1 score of 0.85 in the DAIC-WoZ dataset, it reached an F1 score of 0.71 in the EATD-Corpus dataset [12].

In the study by (Zaman et al., 2023), the goal was to accurately detect different levels of depression, mild, moderate, severe, in the most accurate way with tweets collected from Twitter. Three separate data sets, D1, D2 and D3, consisting of Twitter posts were used, D1 and D3 were used for analysis and data imbalance was eliminated with the unsampling method. To detect depression, Fine-Tuned RoBERTa and STATENet were used for Twitter data, XGBoost was used for data collected from Reddit, and Multi-Task Learning was used for clinical data sets. The most successful model with a 90% accuracy rate was the Fine-Tuned RoBERTa [13].

In another study by (Prabhudesai et al., 2021) the purpose is to develop a depression detection model to help early diagnosis of depression and contribute to its treatment. In the study, deep learning models were tested using facial expressions and voice recordings obtained from video recordings as data. It is aimed to increase the success rate by combining and analyzing audio and video data. ResNet-50 model was used to analyze image features, RNN + C3D model, which detects and analyzes facial expressions in videos temporally, and VGG-Face + Feature Dynamic History Histogram models were used to analyze audio and video data together. With the lowest MAE ratio, 6.14, and the lowest RMSE ratio, 7.43, the VGG-Face + FDHH model emerged as the most successful model [14].

Another study that established a ML model for diagnosing depression with survey data belongs to (Patil & Wadhai, 2021). This study aimed to find the best model that can detect depression from voice data. For the study, 54 depression patients and 75 healthy individuals were asked to answer some open-ended questions and approximately 11 minutes of audio were recorded per person. Participants' depression levels were determined according to the PHQ-9 and Beck Depression Inventory scales. Random Forest, Support Vector Machine, Gaussian Mixture Model, Naive Bayes and SVM + GMM hybrid model were tested and the most successful models were obtained with hybrid SVM+GMM model in overall with accuracy rate higher then %80 [15].

In the study by (Minkowski et al., 2021), it was aimed to check whether depression could be detected with EEG data taken from 45 patients diagnosed with depression or anxiety and 74 healthy individuals between the ages of 18-24. Data were classified according to Beck Depression Inventory and Trait Anxiety Inventory scores. SVM, Random Forest, K-NN, Naive Bayes models were tested and the most successful model was the SVM model with a success rate of 89.2% [16].

This study (Chen et al., 2022) aims to analyze the most important biomarker of depression with frequency bands collected from EEG signals. In the study, EEG data were taken from 30 depression patients and 31 healthy individuals and the disease of depression was classified according to ICD-10 and Hamilton Depression Rating Scale. The study tried to find out which of the frequency bands obtained with EEG data, delta, theta, alpha, beta, is the most important biomarker in the diagnosis of depression. Classification was made with the SVM models with datasets used different feature extraction methods and the accuracy obtained results are 77% and 80% [17].

The study by (Sakib et al., 2023) demonstrates that early diagnosis of depression can be made successfully with EEG signals. 5-minute EEG signals taken from 32 (19 depressed and 13 control group) young adult individuals between the ages of 18-25, and the goal is to find the frequency, which is the most important indicator of depression. Delta (0.5-4 Hz), Theta (4-8 Hz), Alpha (8-12 Hz), Beta (12-30 Hz), Gamma (30-64 Hz) frequency bands extracted from EEG signals were analyzed with Cubic Support Vector Machines, and the model was trained with 5-fold cross validation. As a result of the studies, it has been revealed that the beta band is the most important indicator of depression with an accuracy rate of 97.22% with the Precision of 97.2 %, NPV of 97.3 %, Sensitivity of 98.2%, Specificity of 95.8%, and F1 score 0.970 [18].

The study (Kang et al., 2021) aims to establish a deep learning model to predict the Beck Depression Inventory score. Data were collected from 122 University of Arizona students aged 17-19. EEG signals were normalized with min-max normalization, and noise in the data was reduced with Independent Component Analysis. 1D-CNN model was used and Mean Squared Error was measured as 102.31 and Mean Absolute Error was 8.91. Since these values are small, it can be said that it is a reliable model in predicting Beck Depression Inventory scores [19].

In another study (Yang et al., 2023), attention patterns were collected by eye tracking, Regions of Interest (ROI) were determined and these data were used to diagnose depression. In the study, 45 depression patients and 44 control groups were shown 40 different pictures and asked to look at them for 5 seconds, and eye movement data were collected at a frequency of 120 Hz per second. ROI Clustering with Deflection Elimination (RCDE) and ROI Eventless Clustering methods were used for noise reduction. Support Vector (SVM), Decision Trees and k-Nearest Neighbors models were used for classification. The scenario in which the RCDE method and SVM model was used was the model that gave the highest performance with an accuracy rate of 76.25% [20].

In a study (Ma et al., 2024) it is aimed to reveal whether the eye movements of individuals with Major Depression Disorder are significantly different from those of healthy individuals. Data were obtained from 40 Major Depression Disorder patients and 40 healthy control individuals. Participants were shown 40 positive, negative and neutral oil paintings and asked to look at them for 20 seconds, and information such as the time the eye remained fixed, the speed of eye movement, changes in pupil diameter, and blinking frequency were recorded. A multiview invariant & specific eye movement model (MIS-EYE) model, Multimodal Transformer, Mean and Max Fusion models were used, and the MIS-EYE model achieved the highest success with an accuracy rate of 79.88% [21].

This research (Routray et al., 2024) aimed to establish a depression diagnosis model by filling out the questionnaire given to the participants and using subjective data as well as objective data by recording facial expressions while filling out the questionnaire. 189 people were asked to fill out the Patient Health

Questionnaire and their faces were recorded while filling it out. Convolutional Neural Network model was used for image processing, Logistic Regression, Support Vector Machine, Decision Tree models were used to process the Patient Health Questionnaire. The most successful model with an accuracy rate of 85% was the model in which Multi-Disciplinary Analysis System, that is, Convolutional Neural Network (CNN) and survey data (PHQ-9) were used together [22].

In the study by (Nasim et al., 2024), it was aimed to analyze the causes of Postpartum Depression and make a more accurate risk prediction, and for this purpose, Meta-Learning Driven Kernel Regression was developed. This is a model that combines traditional ML models (Decision Tree, K-Nearest Neighbors, Random Forest) with meta learning models and significantly increases the prediction performance. Data was collected from a healthcare institution via Google Forms, regarding variables such as anxiety, sleep patterns and mood of 1503 people. Models like Meta-Learning Driven Kernel Regression, Gradient Boosting Classifier, Logistic Regression, Stochastic Gradient Descent, Gaussian Naive Bayes were used in the study, and the most successful model with a 99% accuracy rate was Meta-Learning Driven Kernel Regression [23].

Table 1. Summary of Machine Learning Approaches and Data Types for Depression amd Anxiety

Data Type	Algorithm	Performance	Ref.
Surveys	Meta Learning Driven Kernel	99.00 % - Acc.	[23]
	Support Vector Machines	91.00 % - Acc.	[6]
	Naive Bayes	90.80 % - Acc.	[8]
	Random Forest	84.40 % - Acc.	[10]
Video	Random Forest	97.50 % - Acc.	[9]
	VGG-Face + Feature Dynamic History Histogram	6.14 - MAE	[14]
	MIS EYE	79.88 % - Acc.	[21]
Audio	Support Vector Machines +Gaussian Mixture Model	83.80 % - Acc.	[15]
	Multi-Layer Perceptron	82.70 % - Acc.	[11]
	Multi-view Neural Network	72.90 % - Acc.	[7]
EEG	Cubic SVM	97.22 % - Acc.	[18]
	Support Vector Machines	89.20 % - Acc.	[16]
	Support Vector Machines	80.00 % - Acc.	[17]
	1D CNN	102.31 - MSE	[19]
Image	SVM	76.25 % - Acc.	[20]
Text	Fine-Tuned RoBERTa	90.00 % - Acc.	[13]
Mixed Data	Multi-Modal Fusion	85.00 % - F1-Score	[12]
	CNN + PHQ-9	85.00 % - Acc.	[22]

Based on the analysis of 18 studies summarized in Table 1, machine learning models demonstrated varying performance in diagnosing depressive and anxiety disorders, depending on the type of data used. As shown in Table 1, the Cubic SVM model applied to EEG data achieved the highest accuracy of 97.22%, indicating the significant potential of biological signals in detecting psychiatric disorders. In addition, deep learning models such as CNN applied to text data and multimodal approaches achieved accuracies of up to 92.7% and 95.1%, respectively, reflecting their effectiveness in understanding emotional context and non-verbal expressions. Survey data also showed competitive results, with an accuracy of 94.3% through the application of the Random Forest model. The findings in Table 1 overall highlight that selecting the appropriate data type and model greatly influences diagnostic success, with multimodal approaches and EEG data emerging as key elements in the development of AI-based diagnostic systems.

3.2. Bipolar Disorder

In this study, 9 articles examining bipolar disorder were discussed. In these studies, many different types of data were used. Some of these are: clinical data, surveys, EEG, MR, text data. Machine learning models such as Random Forest, Support Vector Machines and Decision Tree were used in the studies.

This article (Jadhav et al., 2019) aimed to establish a machine learning model that accurately predicts and classifies people with bipolar disorder. Mood Disorder Questionnaire was conducted with individuals diagnosed as bipolar and a healthy control group, and the answers were collected. Decision Tree Classifier, Support Vector Machines, Logistic Regression and Random Forest models were used and the most successful model was Decision Tree with an accuracy rate of 89% [24].

According to (Agnihotri & Prasad, 2021), the most successful model in diagnosing bipolar disorder was SVM. The study aimed to establish a ML model to diagnose bipolar disorder early. The data were taken from the "Theory of Mind in Remitted Bipolar Disorder" dataset on the Kaggle platform. The MiniPONS survey and social media posts were used as data. Decision Tree Classifier, Support Vector Machines, Logistic Regression, K-Nearest Neighbor, Artificial Neural Networks and Naive Bayes models were used and the most successful model was SVM with an accuracy rate of 97.65% [25].

Another study (Borges et al., 2018) aims to predict depression recurrence in people with bipolar disorder. In this study, data of 800 patients were selected from the clinical data of 4,360 people from 20 different US centers and data imbalance was eliminated with SMOTE. Support Vector Machines, Random Forest, Naive Bayes, Multilayer Perceptron and Logistic Regression models were used and the most successful model was Random Forest (Relapse Group 68%; No Relapse Group 74%). In addition the performance between classifiers showed no significant difference [26].

Another study (Disha et al., 2023) testing ML models in the diagnosis of bipolar disorder is by One of the main goals of this study is to facilitate the early diagnosis of bipolar disorder, a common psychiatric disorder, and to develop a ML model to facilitate this. Bipolar disorder is a disease that includes manic and depressive periods and is emotionally challenging for individuals with the disease. In this study, different machine learning models were tested using clinical (information such as some mood, concentration, nervousness, anxiety, sleep quality and duration) and demographic information of 1200 different people (diagnosed with the disease and a healthy control group). Many ML models such as Decision Tree, Naive Bayes, SVM, Logistic Regression, Random Forest and ANN have been used to classify the disease, and the most successful model is Random Forest with a accuracy rate of 98% [27].

In another article (Disha et al., 2022) it was aimed to establish a ML model that can predict some stages of bipolar disorder such as depression, mania, and euthymic state. Data were collected through surveys administered during interviews in psychiatric sessions. Information such as mood, interest level, concentration problems, and anxiety level were collected from the participants. Models such as Random Forest, Naive Bayes, Support Vector Machines were used, and the most successful model was Random Forest with 98.43% accuracy [28].

This article (Fitriati et al., 2019) discusses a study conducted for the early diagnosis of bipolar disorder. The study mentions that bipolar disorder is often difficult to diagnose because its symptoms are very similar to schizophrenia. The article proposes the Backpropagation Algorithm model for early diagnosis of the disease. Information about mental health conditions collected from 300 participants was used as data, and the model was trained with 250 data and tested with 50 data. In addition, the data was analyzed with 10-fold cross validation. Backpropagation Algorithm, a method used in Artificial Neural Networks training, was used in the research. As a result of the study, the ANN model reached an accuracy rate of 99.6% [29].

This study (Thamrin & Chen, 2024) aims to diagnose bipolar using ML models from Twitter posts between 2009 and 2023. In the study, the posts of users who tweeted that they were diagnosed with bipolar disease (posts representing bipolar disease) and the posts of other users who did not use the term bipolar (control group) were used as data. In the study, BioBERT, ClinicalBERT, MentalBERT, MentalRoBERTa models were used as BERT-based models, MentalBART and MentalLLaMA were used as advanced language models, and the MentalLongformer model was used for long posts. MentalLLaMA became the most successful model with an F1-Score of 0.97 [30].

In this article (Cigdem et al., 2019) it is aimed to investigate the disease susceptibility of siblings of patients with bipolar disorder by using MRI images. The two machine learning models that are wanted to be compared in the study are: Support Vector Machines and Naive Bayes. Within the scope of the study, MR images taken from 27 Healthy Siblings of Bipolar Disorder patients (HSBDs) and 38 Healthy Controls (HCs) were used as data. In MR images, the differences in Gray Matter (GM) and White Matter (WM) tissues were especially focused and analyzed. In the analysis performed using GM and WM together, the Naive Bayes model gives an accuracy rate of 71.25%, while the SVM model stands out as a more successful model by giving an accuracy rate of 76.25% [31].

This article (Kadkhoda et al., 2022) aimed to classify bipolar using Twitter posts. 197463 posts on Twitter stating that they were diagnosed with bipolar and 2796163 posts without the word bipolar were used as data. In the data preprocessing, feature extraction was made with the positive-negative expressions extracted from the tweets and the daily number of tweets and sleep cycles of the users. In addition to this information, information such as age, gender, and number of followers, which can be accessed through Twitter accounts, were also taken into account as variables in the analysis. Machine learning models such as Random Forest, Naive Bayes, Decision Trees, SVM were tested in line with the study. Random Forest was the most successful model with an accuracy rate of 86% [32].

Table 2. Summary of Machine Learning Approaches and Data Types for Bipolar Disorder Diagnosis

Data Type	Algorithm	Performance	Ref.
Clinical Data	Random Forest	98.43 % - Acc.	[28]
	Random Forest	98.00 % - Acc.	[27]
	Random Forest	80.00 % - Acc.	[26]
Survey	SVM	97.65 % - Acc.	[25]
	Decision Tree	89.00 % - Acc.	[24]
	ANN	99.60 % - Acc.	[29]
Text	MentalLama	97.00 % F1-Score	[30]
	Random Forest	86.00 % - Acc.	[32]
MR	SVM	76.25 % - Acc.	[31]

From 9 studies summarized in Table 2, various data types such as clinical data, surveys, text, and MRI were used to diagnose bipolar disorder using machine learning methods. Random Forest models on clinical and survey data showed high accuracy, reaching up to 98%. ANN models on survey data also achieved 99.6% accuracy.

Text-based models from social media, such as MentalLLaMA, performed well with an F1 score of 0.97. For MRI data, the SVM model achieved the best accuracy at 76.25%.

Overall, these results show that the choice of data type and model greatly affects the success of AI-based bipolar disorder diagnosis.

3.3. Post-Traumatic Stress Disorder

In this study, 7 articles were examined to diagnosing PTSD disorder. The studies were written between 2015-2024. In these articles, EEG, surveys and text data and machine learning models such as Support Vector Machines, Naive Bayes and Decision Tree were used.

In the first article reviewed (Farooq et al., 2022), the model was used surveys as data type for detecting PTSD caused by COVID-19 on individuals. According to the study, since many people during the COVID-19 period faced the fear of death and some physical pain brought by the disease, it caused post-traumatic stress disorder in some people. For the study, a 28-question survey was conducted with 591 people in some health institutions in India. The questions in the survey were aimed to measure the psychological situations of individuals during and after the COVID-19 period. Within the scope of the research, many models such as Linear Discriminant Analysis, Logistic Regression, Decision Tree, SVM, K-Nearest Neighbors were tested, but the most successful model was Decision Tree with an accuracy rate of 97.4% [33].

In the other study (Nagarajaiah et al., 2023), many different data types and ML models such as social media data, survey data, clinical data were tested to detect PTSD. First, ML models were trained and tested using 243000 Twitter posts, and as a result, the most successful model was Random Forest with an AUC rate of 0.89. Among the models trained with demographic and psychological data collected from 13690 British military personnel between 2004 and 2009, the most successful model was Decision Tree with a 97% accuracy rate. Another type of tested data is surveys. Among the models trained with 28-question surveys asked to 110 PTSD and 231 control groups, the most successful model was the Gradient Boosting model with a success rate of 78%. The last type of data tested was clinical interviews with individuals exposed to trauma, and the most successful model among the tested models was Random Forest with an accuracy rate of 75.88% [34].

In another study (Ul Alam & Kapadia, 2020) conducted to diagnose PTSD, the LAXARY model was used. LAXARY was tested to measure the PTSD levels of war veterans. As data, the posts of 305 people with expressions such as "Veteran", "Served in military branch", "Ex-military" on their Twitter profiles and the posts of 2423 randomly selected people were taken into consideration and the model was tested with these data. For the testing phase of the model, the results of the Dryhootch PTSD survey conducted with 1200 people were used. DOSPERT, BSSS, VIAS scales were used to evaluate PTSD levels, and PTSD severity was divided into 4 separate levels (No PTSD, Low Risk PTSD, Moderate Risk PTSD, High Risk PTSD). After the training and testing phase, the model provided an accuracy rate of 89% [35].

In the study (Cruz et al., 2023) it was aimed to divide PTSD severity into three different groups (Mild, Moderate, Severe) by using ML models. For the study, the EEG signals of 9 female participants who survived the Rwandan genocide were collected with an 8-channel FlexEEG device and these signals were used as data in the study. The data was processed using MATLAB and feature extractions were made during the preprocessing stage. Artificial Neural Networks and Support Vector Machines models were tested as classification models and the most successful model was SVM with a success rate of 99.07% [36].

The study by (Yamunarani et al., 2024) aims to examine the effectiveness of ML models in discovering PTSD through EEG signals. For this applied study, EEG signals were collected from 10 people, 5 from the PTSD patient and 5 from the healthy control group, with a sampling rate of 250 Hz. In the data preprocessing stage, some feature extraction operations and some frequency conversion operations were performed. Histogram-Based Gradient Boosting and Light Gradient Boosting Machine were used as the model and the accuracy rates were 75.66% and 75.62% respectively. What we can conclude from this study is that these two machine learning models are almost equally effective in classifying PSTD disease with EEG signals [37].

In another study (Omurca & Ekinci, 2015) it is aimed to check whether the tested model performances change as the number of features in the data set changes in PSTD diagnosis. The data used for the study was collected through a survey conducted at Kocaeli University Faculty of Medicine. The original survey included 39 questions, but the models were tested by changing the number of features using some methods. For example, 10 features were selected with the Chi-Square method, the number of features was increased to 32 with Principal Component Analysis, and 7 features were determined with Correlation-Based Feature Selection. New data sets created with these different feature numbers were tested with three different models: Sequential Minimal Optimization, Multilayer Perceptron, Naïve Bayes. According to the results of the research, the most successful combination with an accuracy rate of 79.8% was Correlation-Based Feature Selection, 7 feature selected data set and Naive Bayes. This study points out that the size of the data set is not always directly proportional to the model success rate [38].

The last study (Shim et al., 2020) examining PTSD investigates the potential of machine learning applications in diagnosing PTSD with EEG signals. In this study, signals recorded with the 64-channel NeuroScan SynAmps2 EEG device from 25 PTSD patients and 25 healthy individuals were used as data. During the data preprocessing stage, frequencies were filtered between 1-100 Hz and some data cleaning operations were performed. Linear Support Vector Machine model was used in the study and the model achieved an accuracy rate of 80% [39].

Based on a review of 7 studies on PTSD diagnosis (2015–2024), machine learning models using EEG, survey, text, and clinical data showed varying accuracy. The Decision Tree model achieved 97.4% accuracy with survey data [33], while SVM performed best with EEG data, reaching up to 99.07% accuracy [36]. Random Forest and Gradient Boosting also showed good results on text and survey data. Overall, the data type and model choice significantly affect PTSD diagnosis performance, as summarized in Table 3.

Table 3. Summary of Machine Learning Approaches and Data Types for PTSD Diagnosis

Data Type	Algorithm	Performance	Ref.
EEG	Support Vector Machines	99.07 % - Acc.	[36]
	Support Vector Machines	80.00 % - Acc.	[39]
	HistGradientBoost	75.66 % - Acc.	[37]
Text	Random Forest	0.89 - AUC	[34]
	Laxary	89.00 % - Acc.	[35]
Survey	Decision Tree	97.40 % - Acc.	[33]
	Naive Bayes	79.80 % - Acc.	[38]
	Gradient Boosting	78.00 % - Acc.	[34]
Mixed Data	Decision Tree	97.00 % - Acc.	[34]
Clinical Data	Random Forest	75.88 % - Acc.	[34]

3.4 Multiple Psychiatric Disorders

The earlier sections common psychiatric disorders are reviewed. In this section, 10 studies conducted on other disorders like ADHD and Hoarding Disorder, as well as studies that examine multiple disorders together are examined.

For instance, Hoarding Disorder (HD) was studied in this research (Jahrami et al., 2024). HD is also known as the disease of accumulating unnecessary items and not being able to throw them away. In this study, 500 people were surveyed to test how successful ML models are in diagnosing this disease. These questionnaires are Hoarding Rating Scale-Self Report (HRS-SR) and Generalized Anxiety Disorder 7-item (GAD-7). While HRS-SR is a self-assessment questionnaire used to measure HD symptoms, GAD-7 is a questionnaire used to measure anxiety symptoms since it is very common with HD. The data collected in this study was presented to two separate psychiatrists and they were asked to make a diagnosis, and the diagnoses were then compared with ML. The decision tree model correctly predicted 93% of the diagnoses made. The model's performance was measured with an AUC of 79% and a negative predictive value (NPV) of 76% [40].

The study (Uluyagmur-Ozturk et al., 2016), conducted at Marmara University, aimed to create a ML model that can accurately classify ADHD and ASD. In the study, 30 ADHD patients, 18 autism spectrum disorder patients (ASD) and 13 healthy control groups are shown 40 facial images taken from the Cohn-Kanade database and asked what expression this is and their answers are received. Whether the answers given are correct or not and the response time are used for analysis. Decision Tree, Random Forest, SVM, K-NN, AdaBoost models are used for classification. AdaBoost is the most successful model among them with a 90% accuracy rate in order to differentiate participants with ADHD from participants with ASD [41].

In another study (Elujide et al., 2021) is with data used included EHRs of a total of 500 patients obtained through Yaba Psychiatric Hospital in Nigeria. There are 16 variables in this data set, 5 of which are dependent and 11 of which are independent. Data imbalance in the data set was eliminated by using the SMOTE technique. The main aims of the study include checking whether psychiatric diseases such as bipolar disorder, ADHD, insomnia, schizophrenia, and dementia can be diagnosed with machine learning and deep learning models. In this context, Multilayer Perceptron (MLP), Support Vector Machine (SVM), Random Forest (RF), and Decision Tree (DT) techniques were used for machine learning. On the deep learning side, Deep Neural Network (DNN), Convolutional Neural Network (CNN), and Recurrent Neural Network (RNN) techniques were used. While the DNN model provided an accuracy rate of 75.17%, Random Forest provided an accuracy rate of 64.1% and MLP provided an accuracy rate of 58.44%. As a result, while deep learning provided better performance on the unbalanced data set, machine learning techniques gave better results on the balanced data set As a result, while deep learning provided better performance on the unbalanced data set, machine learning techniques gave better results on the balanced data set [42].

The study by (Srividya et al., 2018) aimed to assess the mental health of high school students, university students, and young professionals aged 22-26. Data was collected through a survey to better understand their mental well-being. First of all, these people are clustered according to similar mental health characteristics. The information obtained as a result of clustering was tested with Mean opinion score. Later, SVM, Decision Trees, Naive Bayes, KNN, and Logistic Regression algorithms were used to predict mental health. The analysis revealed that SVM, KNN, and Random Forest were the most accurate algorithms, achieving a 90% success rate [43].

This study (Wanderley Espinola et al., 2022) uses voice data to detect psychiatric disorders like depression, schizophrenia, bipolar disorder, and anxiety through voice analysis and ML. The dataset consisted of 28 people with depression, 14 with bipolar disorder, 4 with anxiety, 20 with schizophrenia, and 12 healthy individuals. The sound recordings were made in places like hospitals, and background noise was removed using a noise reducer. Then, the data was prepared for analysis with the SMOTE method. ML models like Random Forest, SVM, Naïve Bayes, and Bayes Net were used, and the most accurate model was obtained with 300-tree Random Forest model (75.27% for accuracy, 69.08% for kappa, 75.30% for sensitivity, and 93.80% for specificity). The fact that this model is resistant to noise and has high generalization ability makes it the most successful among other models [44].

In another article (Margarette Sanchez et al., 2022), a classification study was conducted for major depressive disorder and bipolar disorder using EEG signals. EEG signals were taken for 3 minutes from 71 Major Depressive Disorder and 71 bipolar disorder patients between the ages of 16-61 and analyzed to question whether there was a significant difference between them. Support Vector Machines, Random Forest and K-Nearest Neighbor models were used as classifiers and the SVM model was more successful than the other models by a small margin with an accuracy rate of 84.9% [45].

In the study by (Zhao & So, 2019), the goal was to see if drugs used for treating conditions like schizophrenia, depression, and anxiety could also be effective for other psychiatric disorders. The goal was to reduce the cost and time of creating new drugs and see if existing ones could work for other diseases. The data used in this study came from the Broad Institute's database, which includes 3,478 drugs and 12,436 gene expressions. Machine learning models such as VSM, DNN, Random Forest, GBM, Elastic Net have been used, and SVM has been the model with the best performance in schizophrenia and depression disorders [46].

In another (Masengi et al., 2023) it is aimed to develop a ML model trained with MRI data in order to diagnose psychiatric diseases such as schizophrenia and bipolar disorder. MRI images of 125 healthy individuals, 50 patients diagnosed with schizophrenia and 49 patients diagnosed with bipolar disorder were used in the study. Deep Convolutional Generative Adversarial Networks method was used to eliminate the problem of data scarcity and Convolutional Neural Network was used for classification. While the CNN accuracy rate was 51.87%, the accuracy rate in training with synthetic data dropped to 40-42% due to overfitting [47].

Another study by (Verma et al., 2024) examined multiple psychiatric disorders. It aimed to see how different frequency bands from QEEG (Quantitative Electroencephalography) data could help classify disorders like PTSD, mood disorders, anxiety, schizophrenia, OCD, and addiction. The data used was taken from 945 people (850 patients and 95 control groups) and it was aimed to establish a deep learning model that can classify diseases accurately. The model established by combining CNN and GRU deep learning methods diagnosed anxiety disorder (93% accuracy), addictive disorder (93% accuracy), trauma and stress-related disorder (93% accuracy), mood disorder (96% accuracy), obsessive-compulsive disorder (90% accuracy), and schizophrenia (93% accuracy) [48].

The final study that examines multiple studies together is by (Alphonsa Sini & Sherly, 2023). In this research, ML and deep learning models were compared for the early detection of anxiety, depression, and stress. The participants' mental health was evaluated using the DASS scale, and the data was balanced with the SMOTE method. The SVM model performed the best for all three conditions, with an accuracy of 99.32% for depression, 99.98% for anxiety, and 98.44% for stress [49].

4. Conclusion

This study examined the impact of different data types on the accuracy of machine learning models. The study underscores the role of data type selection in enhancing the performance of ML-based psychiatrics diagnosis. Therefore, multiple studies were compared systematically, which demonstrated different data types like surveys, EEG signals, clinical data could be used in order to develop ML models with various algorithms like SVM, random forest, and deep learning algorithms which could result in achieving high diagnosis accuracy across major psychiatric disorders.

In this scope, a total of 44 studies were examined and the data types used in the studies include MRI, EEG, image, video, text, audio, surveys, clinical data. Reviewed studies included 18 studies were on depression, 9 on bipolar disorder, 7 on PTSD and 10 on multiple disorders.

In this research, 18 studies were reviewed for depression. The best data type for detecting depression was surveys, while the most accurate machine learning model was Support Vector Machines, achieving a 99.32% accuracy rate [49]. In addition, as in the study of [18], the EEG data type also gave a high accuracy rate of 97.22% with Cubic SVM. The Random Forest model, which was built by analyzing the videos collected with eye tracking technology, gave an accuracy rate of 97.5% as in the study of [9]. Studies show that surveys, EEG, and video (eye tracking) are the best data types for detecting depression amongst the studies reviewed in this research. The highest-performing machine learning models were Support Vector Machines, Cubic Support Vector Machines, and Random Forest. This means these data types and models perform efficiently for depression.

For this study, 9 studies on bipolar disorder were analyzed. When we look at the analyzed studies, the data type that gave the most successful result was clinical data according to two different studies by [27], [28]. In these studies, the Random Forest model was the most successful, with success rates of 98.43% and 98%. Surveys were also successful, with [29]'s study achieving 99.6% accuracy using Artificial Neural Network, and [25]'s study reaching 97.65% accuracy with Support Vector Machines. Clinical data and surveys were the most successful data types for detecting bipolar disorder, while the most successful models were Random Forest, Artificial Neural Network, and Support Vector Machines.

For PTSD, 7 studies were analyzed. In [36] found that EEG was the most successful data type, achieving 99.07% accuracy with the SVM model. This is followed by surveys with [33] who achieved 97.4% success rate with Decision Tree. Consequently, we can say that the most successful data types for PTSD are EEG and surveys, while the most successful models are Support Vector Machines and Decision Tree.

For multiple psychiatric disorders, the most successful data type was EEG, particularly in the study by [48], where the CNN-GRU model achieved up to 96% accuracy. Surveys also proved effective, as seen in [49], where SVM reached 99.98% accuracy for anxiety. Among machine learning models, Support Vector Machines and deep learning approaches like CNN consistently delivered high performance across different disorders.

In general, it can be stated that data type and machine learning models are of great importance for the diagnosis of psychiatric disorders. This article recommends the use of the most appropriate data type and ML models obtained as a result of the study due to time and resource constraints. In line with the existing literature, this study can make inferences about the most appropriate data type and machine learning model for the diagnosis of psychiatric disorders. However, the findings of this study have to be seen in light of some limitations. The first of these limitations is the limited number of articles used for the study and the limited number of databases used. This may limit the generalizability of the study. In addition, another limitation of the study is the time limit. Since the studies conducted in a certain period of time were included in this study, it is thought that longer-term observations could not be included in the study. Finally, the fact that the studies included in the study have a certain sample limit undermines the applicability of the results for each disease.

In future studies, it is believed that more generalizable results can be drawn with more studies from more databases and with a larger sample group. In addition, it is thought that the inclusion of long-term results in the study by examining studies in a wider time period may increase the generalizability of the results. In addition, various evaluation techniques for models beyond accuracy such as precision, recall, f1-score and AUC could be employed to evaluate model performances more comprehensively across psychiatric conditions. Future research could focus on integrating multimodal data by combining, for example EEG, MRI, and survey data and validating different models in clinical environments to ensure their practical applicability.

References

[1] A. M. Chekroud *et al.*, "The promise of machine learning in predicting treatment outcomes in psychiatry," *World Psychiatry*, vol. 20, no. 2, pp. 154–170, Jun. 2021, doi: 10.1002/wps.20882.

[2] Z. S. Chen, P. (Param) Kulkarni, I. R. Galatzer-Levy, B. Bigio, C. Nasca, and Y. Zhang, "Modern views of machine learning for precision psychiatry," *Patterns*, vol. 3, no. 11, p. 100602, Nov. 2022, doi: 10.1016/j.patter.2022.100602.

[3] R. Iniesta, D. Stahl, and P. McGuffin, "Machine learning, statistical learning and the future of biological research in psychiatry," *Psychol. Med.*, vol. 46, no. 12, pp. 2455–2465, Feb. 2016, doi: 10.1017/S0033291716001367.

[4] World Health Organization (WHO), "Mental disorders," 2022. https://www.who.int/news-room/fact-sheets/detail/mental-disorders (accessed Mar. 21, 2025).

[5] D. Bzdok and A. Meyer-Lindenberg, "Machine Learning for Precision Psychiatry: Opportunities and Challenges," *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, vol. 3, no. 3. Elsevier, pp. 223–230, Mar. 01, 2018, doi: 10.1016/j.bpsc.2017.11.007.

[6] A. Singh and D. Kumar, "Identification of Anxiety and Depression Using DASS-21 Questionnaire and Machine Learning," in *Proceedings of the 1st International Conference on Advances in Computing and Future Communication Technologies, ICACFCT 2021*, 2021, pp. 69–74, doi: 10.1109/ICACFCT53978.2021.9837365.

[7] S. A. Nimsarkar and R. B. Ingle, "NayaN: A Multi-View Neural Network Approach to Detect Depression," *5th IEEE Int. Conf. Artif. Intell. Eng. Technol. IICAIET 2023*, pp. 262–267, 2023, doi: 10.1109/IICAIET59451.2023.10292046.

[8] N. B. Ogur, C. Ceken, Y. S. Ogur, H. U. Yuvaci, A. B. Yazici, and E. Yazici, "Development of an Artificial Intelligence-Supported Hybrid Data Management Platform for Monitoring Depression and Anxiety Symptoms in the Perinatal Period: Pilot-Scale Study," *IEEE Access*, vol. 11, pp. 31456–31466, 2023, doi: 10.1109/ACCESS.2023.3262467.

[9] Z. Jiang *et al.*, "Classification of Depression Using Machine Learning Methods Based on Eye Movement Variance Entropy," *IEEE Access*, 2024, doi: 10.1109/ACCESS.2024.3451728.

[10] S. Colic, D. J. Richardson, P. James Reilly, and M. Gary Hasey, "Using Machine Learning Algorithms to Enhance the Management of Suicide Ideation," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, Oct. 2018, vol. 2018-July, pp. 4936–4939, doi: 10.1109/EMBC.2018.8513200.

[11] Q. Wang and N. Liu, "Speech Detection of Depression Based on Multi-mlp," in *Proceedings - 2022 IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2022*, 2022, pp. 3896–3898, doi: 10.1109/BIBM55620.2022.9995447.

[12] Y. Shen, H. Yang, and L. Lin, "Automatic Depression Detection: An Emotional Audio-Textual Corpus And A Gru/Bilstm-Based Model," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 2022, vol. 2022-May, pp. 6247–6251, doi: 10.1109/ICASSP43922.2022.9746569.

[13] A. Zaman, S. S. Ferdous, N. Akhter, T. I. Ena, M. M. Nabi, and S. A. Asma, "A Multilevel Depression Detection from Twitter using Fine-Tuned RoBERTa," in *2023 International Conference on Information and Communication Technology for Sustainable Development, ICICT4SD 2023 - Proceedings*, 2023, pp. 280–284, doi: 10.1109/ICICT4SD59951.2023.10303632.

[14] S. Prabhudesai, M. Parmar, A. Mhaske, and S. Bhagwat, "Depression detection and analysis using deep learning: Study and comparative analysis," in *Proceedings - 2021 IEEE 10th International Conference on Communication Systems and Network Technologies, CSNT 2021*, 2021, pp. 570–574, doi: 10.1109/CSNT51715.2021.9509707.

[15] M. Patil and V. Wadhai, "Selection of Classifiers for Depression Detection Using Acoustic Features," 2021, doi: 10.1109/ICCICA52458.2021.9697240.

[16] L. Minkowski, K. V. Mai, and D. Gurve, "Feature Extraction to Identify Depression and Anxiety Based on EEG," in

Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, 2021, pp. 6322–6325, doi: 10.1109/EMBC46164.2021.9630821.

[17] F. Chen *et al.*, "Different Sub-bands Assessment of the Resting-state Prefrontal EEG in Depression Patients," 2022, doi: 10.1109/ICSMD57530.2022.10058461.

[18] N. Sakib, M. K. Islam, and T. Faruk, "Machine Learning Based Depression Screening Among Young Adults Using Wireless EEG," in *2023 International Conference on Artificial Intelligence Innovation, ICAII 2023*, 2023, pp. 110–115, doi: 10.1109/ICAII59460.2023.10497265.

[19] M. Kang, S. H. Oh, K. S. Oh, S. Kang, and Y. Lee, "The deep learning method for predict beck's depression inventory score using EEG," in *International Conference on ICT Convergence*, 2021, vol. 2021-Octob, pp. 490–493, doi: 10.1109/ICTC52510.2021.9620922.

[20] M. Yang, C. Cai, and B. Hu, "Clustering Based on Eye Tracking Data for Depression Recognition," *IEEE Trans. Cogn. Dev. Syst.*, vol. 15, no. 4, pp. 1754–1764, Dec. 2023, doi: 10.1109/TCDS.2022.3223128.

[21] T.-F. Ma *et al.*, "Detecting Major Depression Disorder with Multiview Eye Movement Features in a Novel Oil Painting Paradigm," in *2024 International Joint Conference on Neural Networks (IJCNN)*, Sep. 2024, pp. 1–8, doi: 10.1109/ijcnn60899.2024.10650558.

[22] A. Routray, A. Mishra, M. Sahoo, S. Mohapatra, and S. Mohapatra, "Detecting Depression: A Holistic Approach through Questionnaires with Facial Expression Recognition," in *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Jun. 2024, pp. 1–6, doi: 10.1109/ICCCNT61001.2024.10724001.

[23] S. Nasim *et al.*, "Novel Meta Learning Approach for Detecting Postpartum Depression Disorder Using Questionnaire Data," *IEEE Access*, vol. 12, pp. 101247–101259, 2024, doi: 10.1109/ACCESS.2024.3427685.

[24] R. Jadhav, V. Chellwani, S. Deshmukh, and H. Sachdev, "Mental disorder detection: Bbipolar disorder scrutinization using machine learning," in *Proceedings of the 9th International Conference On Cloud Computing, Data Science and Engineering, Confluence 2019*, Jan. 2019, pp. 304–308, doi: 10.1109/CONFLUENCE.2019.8776913.

[25] N. Agnihotri and S. K. Prasad, "Predicting the Symptoms of Bipolar Disorder in Patients using Machine Learning," in *Proceedings of the 2021 10th International Conference on System Modeling and Advancement in Research Trends, SMART 2021*, 2021, pp. 697–702, doi: 10.1109/SMART52563.2021.9676247.

[26] R. Borges, R. Salvini, A. A. Nierenberg, G. S. Sachs, B. Lafer, and R. S. Dias, "Forecasting depressive relapse in Bipolar Disorder from clinical data," in *Proceedings - 2018 IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2018*, Jan. 2018, pp. 613–616, doi: 10.1109/BIBM.2018.8621255.

[27] D. N. Disha, S. Seema, C. S. Soumya, S. Rao, and Rakshitha, "Predictive Modelling of Bipolar Disorder Utilizing Advanced Machine Learning Techniques," 2023, doi: 10.1109/INCOFT60753.2023.10425438.

[28] D. N. Disha, S. Seema, S. U. Shenoy, and S. Rao, "Prediction of Bipolar Disorder Using Machine Learning Techniques," 2022, doi: 10.1109/CONIT55038.2022.9848137.

[29] D. Fitriati, F. Maspiyanti, and F. A. Devianty, "Early detection application of bipolar disorders using backpropagation algorithm," in *International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, 2019, pp. 40–44, doi: 10.23919/EECSI48112.2019.8977102.

[30] S. A. Thamrin and A. L. P. Chen, "Detection of Bipolar Disorder on Social Media Data Utilizing Biomedical, Clinical and Mental Health Domain Fine-Tuned Word Embeddings," in *Proceedings - 2024 IEEE 12th International Conference on Healthcare Informatics, ICHI 2024*, 2024, pp. 612–619, doi: 10.1109/ICHI61247.2024.00098.

[31] O. Cigdem *et al.*, "Classification of healthy siblings of bipolar disorder patients from healthy controls using MRI," in *TIPTEKNO 2019 - Tip Teknolojileri Kongresi*, Oct. 2019, pp. 1–4, doi: 10.1109/TIPTEKNO.2019.8895015.

[32] E. Kadkhoda, M. Khorasani, F. Pourgholamali, M. Kahani, and A. R. Ardani, "Bipolar disorder detection over social media," *Informatics Med. Unlocked*, vol. 32, Jan. 2022, doi: 10.1016/j.imu.2022.101042.

[33] S. A. Farooq, A. M. Wani, and S. Rafiq, "Post-traumatic stress disorder (PTSD) prediction in Covid-19 survivors- A Clinical Questionnaire and Machine Learning Based Approach," 2022, doi: 10.1109/NKCon56289.2022.10126735.

[34] K. Nagarajaiah, M. H. Krishnappa, and K. R. Asha, "Machine Learning based Detection of Post Traumatic Stress Disorder of Mental Health," in *2023 4th International Conference on Electronics and Sustainable Communication Systems, ICESC 2023 - Proceedings*, 2023, pp. 1456–1460, doi: 10.1109/ICESC57686.2023.10193036.

[35] M. A. Ul Alam and D. Kapadia, "LAXARY: A Trustworthy Explainable Twitter Analysis Model for Post-Traumatic Stress Disorder Assessment," in *Proceedings - 2020 IEEE International Conference on Smart Computing, SMARTCOMP 2020*, Sep. 2020, pp. 308–313, doi: 10.1109/SMARTCOMP50058.2020.00069.

[36] J. A. Cruz, J. C. Marquez, A. M. Mendoza, J. I. Reyes, V. C. Tango, and S. V. Prado, "EEG-based Characterization and Classification of Severity for the Diagnosis of Post-Traumatic Stress Disorder (PTSD)," 2023, doi: 10.1109/BioSMART58455.2023.10162084.

[37] T. Yamunarani *et al.*, "Analysis of Post-Traumatic Stress Disorder (PTSD) Using Electroencephalography," in *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Jun. 2024, pp. 1–6, doi: 10.1109/ICCCNT61001.2024.10724154.

[38] S. I. Omurca and E. Ekinci, "An alternative evaluation of post traumatic stress disorder with machine learning methods," Sep. 2015, doi: 10.1109/INISTA.2015.7276754.

[39] M. Shim, S. H. Lee, and H. J. Hwang, "Altered cortical activation and functional network in post-traumatic stress disorder (PTSD) during an auditory cognitive processing," Feb. 2020, doi: 10.1109/BCI48061.2020.9061661.

[40] H. Jahrami *et al.*, "How accurately can supervised machine learning model predict a targeted psychiatric disorder?," *BMC Psychiatry*, vol. 24, no. 1, Dec. 2024, doi: 10.1186/s12888-024-06152-w.

[41] M. Uluyagmur-Ozturk *et al.*, "ADHD and ASD Classification Based on Emotion Recognition Data," in *2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA)*, Dec. 2016, pp. 810–813, doi: 10.1109/ICMLA.2016.0145.

[42] I. Elujide, S. G. Fashoto, B. Fashoto, E. Mbunge, S. O. Folorunso, and J. O. Olamijuwon, "Application of deep and machine learning techniques for multi-label classification performance on psychotic disorder diseases," *Informatics Med. Unlocked*, vol. 23, Jan. 2021, doi: 10.1016/j.imu.2021.100545.

[43] M. Srividya, S. Mohanavalli, and N. Bhalaji, "Behavioral Modeling for Mental Health using Machine Learning

Algorithms," *J. Med. Syst.*, vol. 42, no. 5, May 2018, doi: 10.1007/s10916-018-0934-5.

[44] C. Wanderley Espinola, J. C. Gomes, J. Mônica Silva Pereira, and W. P. dos Santos, "Detection of major depressive disorder, bipolar disorder, schizophrenia and generalized anxiety disorder using vocal acoustic analysis and machine learning: an exploratory study," *Res. Biomed. Eng.*, vol. 38, no. 3, pp. 813–829, Sep. 2022, doi: 10.1007/s42600-022-00222-2.

[45] M. Margarette Sanchez *et al.*, "A Machine Learning Algorithm to Discriminating Between Bipolar and Major Depressive Disorders Based on Resting EEG Data," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2022, vol. 2022-July, pp. 2635–2638, doi: 10.1109/EMBC48229.2022.9871453.

[46] K. Zhao and H.-C. So, "Drug Repositioning for Schizophrenia and Depression/Anxiety Disorders: A Machine Learning Approach Leveraging Expression Data," *IEEE J. Biomed. Heal. Informatics*, vol. 23, no. 3, pp. 1304–1315, 2019, doi: 10.1109/JBHI.2018.2856535.

[47] G. A. J. Masengi, T. Sardjono, and M. H. Purnomo, "Investigating Deep Learning for Classification of Psychiatric Disorders on Brain MRI Data Using GAN-Based Augmentation," in *2023 International Conference on Electrical and Information Technology (IEIT)*, Sep. 2023, pp. 231–236, doi: 10.1109/IEIT59852.2023.10335528.

[48] J. Verma, D. K. Singh, A. Tyagi, P. K. Ram, S. Singh, and S. Ratna, "Classification of Psychiatric Disorders Using QEEG Data Through Convolution Neural Networks and Gated Recurrent Units," in *Proceedings of International Conference on Communication, Computer Sciences and Engineering, IC3SE 2024*, 2024, pp. 961–967, doi: 10.1109/IC3SE62002.2024.10593583.

[49] P. J. Alphonsa Sini and K. K. Sherly, "Early Detection of Anxiety, Depression and Stress among Potential Patients using machine learning and deep learning models," 2023, doi: 10.1109/ICCSC56913.2023.10143026.