Iran J Psychiatry 2023; 18: 2: 237-247

Short Communication

An Overview of Bipolar Disorder Diagnosis Using Machine Learning Approaches: Clinical Opportunities and Challenges

Walter Antonio Campos-Ugaz[1*], Jessica Paola Palacios Garay[2], Oriana Rivera-Lozada[2], Mitchell Alberto Alarcón Diaz[3], Doris Fuster-Guillén[3], Arístides Alfonso Tejada Arana[2]

Abstract

Objective: Automatic diagnosis of psychiatric disorders such as bipolar disorder (BD) through machine learning techniques has attracted substantial attention from psychiatric and artificial intelligence communities. These approaches mostly rely on various biomarkers extracted from electroencephalogram (EEG) or magnetic resonance imaging (MRI)/functional MRI (fMRI) data. In this paper, we provide an updated overview of existing machine learning-based methods for bipolar disorder (BD) diagnosis using MRI and EEG data.
Method: This study is a short non-systematic review with the aim of describing the current situation in automatic diagnosis of BD using machine learning methods. Therefore, an appropriate literature search was conducted via relevant keywords for original EEG/MRI studies on distinguishing BD from other conditions, particularly from healthy peers, in PubMed, Web of Science, and Google Scholar databases.
Results: We reviewed 26 studies, including 10 EEG studies and 16 MRI studies (including structural and functional MRI), that used traditional machine learning methods and deep learning algorithms to automatically detect BD. The reported accuracies for EEG studies is about 90%, while the reported accuracies for MRI studies remains below the minimum level for clinical relevance, i.e. about 80% of the classification outcome for traditional machine learning methods. However, deep learning techniques have generally achieved accuracies higher than 95%.
Conclusion: Research utilizing machine learning applied to EEG signals and brain images has provided proof of concept for how this innovative technique can help psychiatrists distinguish BD patients from healthy people. However, the results have been somewhat contradictory and we must keep away from excessive optimistic interpretations of the findings. Much progress is still needed to reach the level of clinical practice in this field.

Key words: *Bipolar Disorder; Electroencephalogram; Magnetic Resonance Imaging; Machine Learning*

1. Universidad Nacional Pedro Ruiz Gallo, Lambayeque, Perú.
2. Universidad Privada Norbert Wiener, Lima, Perú.
3. Universidad Nacional Mayor de San Marcos, Lima, Perú.

***Corresponding Author:**
Address: Pinos de la Plata Block "A" Dpto 302, Postal Code: 14000.
Tel: 51 949519981, Fax: 51 074283146, Email: naneniwalter@gmail.com

Article Information:
Received Date: 2022/10/21, Revised Date: 2022/11/13, Accepted Date: 2022/11/14

Bipolar disorder (BD) is a chronic psychiatric illness that affects approximately 1% of the general population (1-3). BD causes a large socio-economic burden. Some challenges still remain for the management of BD, including its under-diagnosis in clinical settings (4). At this time, diagnosis is based solely on clinical evaluations, which can be influenced by subjective biases of patients and informants in reporting symptoms. This results in common misdiagnosis and a relatively long delay between onset and diagnosis of BD (5). Delay in treatment may lead to episode recurrence and poor outcomes (6). Therefore, early diagnosis and timely treatment in a personalized manner can have important effects on prognosis (7).

Growing evidence supports neurological changes in BD (8-10). Different neuroimaging and electrophysiological researches (mostly in the form of cross-sectional case-control studies) reported various alterations in the brain structure and function of patients with BD compared to healthy individuals (11, 12). For example, we can mention abnormal alpha fluctuations and deficits in visual and auditory steady-state potentials in EEG studies on BD (13, 14), as well as abnormal volumes of grey matter and dysfunction of frontal areas of the brain and the limbic system in neuroimaging studies on BD (15, 16). In recent years, automatic diagnosis of psychiatric disorders such as BD through machine learning techniques has attracted substantial attention from psychiatric and artificial intelligence communities (17). These approaches mostly rely on various biomarkers extracted from EEG or MRI/fMRI data (18-20).

Machine learning is a branch of artificial intelligence that may be used for big data analysis purposes. It utilizes complicated mathematical algorithms to implement learning models. Machine learning algorithms recognize various patterns in datasets for extracting knowledge from and making predictions on unlabeled data (Figure 1) (21). In machine learning approaches, prediction and classification models are trained to a research sample and produce outcomes at the individual level (supervised learning) or at the group level (unsupervised learning) (22). These models have the potential to address the complexity of BD neuropathology and become promising predictive tools in clinical settings in the near future. Therefore, a growing number of MRI and EEG studies have tried to utilize machine learning to differentiate BD from healthy individuals and other mental illnesses (18, 23). Machine learning can help distinguish BD from schizophrenia or major depressive disorder in the first episode of psychosis or depression. Furthermore, it may yield outcome biomarkers, directing the intensity and type of treatment at the individual level. Machine learning-based methods can also anticipate treatment outcomes according to neuroimaging or electrophysiological data (24, 25).

Figure 1. General Stages of Bipolar Disorder Diagnosis Using Machine Learning Approach

Previous review articles have discussed some important aspects of the field, however, without highlighting the clinical opportunities and challenges arising from the use of machine learning techniques for the management of BD (26, 27). In this paper, we provide an updated overview of existing machine learning-based methods for BD diagnosis using MRI and EEG data. In this regard, important methodological challenges of the original papers were assessed to determine the reliability of their results and observations. In addition, we attempt to provide practical suggestions for future machine learning studies.

EEG studies

We reviewed 10 articles (28-37) that attempted to classify patients with BD versus healthy subjects using machine learning techniques and EEG signals. Sample sizes in EEG studies ranged from 36 to 89 subjects. All studies adopted the 10-20 international standard for electrode placement. They also used Ag/AgCl electrode type to record the signal. Different preprocessing approaches were adopted in studies to prepare EEG data for further analysis. Most studies utilized the resting state with eyes open and closed for EEG recording, except for Nazhvani *et al.* (30) who used visual flash stimulation during signal recording to elicit visual evoked potentials reaching 92.85% accuracy. Most studies utilized band-pass filters, notch filters and independent component analysis to reduce noise, various artifacts and other interference of the EEG signals, except for two studies (31, 34) in which visual inspection was performed to remove artifacts from the data. These studies utilized both linear (e.g., frequency band power) and nonlinear (e.g., fractal dimension and entropy) features as inputs to machine learning methods. Leave-one-out is the most used cross-validation method in these studies. Hold-out and K-fold cross-validation are other methods used for validation purposes. Multilayer perceptron neural network and K-nearest neighbor are among the machine learning methods for classifying BD from healthy subjects using EEG signals. Two recent studies used deep learning methods for the EEG distinction of BD from healthy controls and reported relatively high accuracy compared to other techniques (36, 37). The reported classification accuracy is in the range of 76-96.88% (see Table 1).

Table 1. Summary of Electroencephalogram Studies on Bipolar Disorder Diagnosis through Machine Learning Techniques

Author (date)	Research sample	Recording protocol	Extracted features	Machine learning technique	Validation method	Outcomes
Sadatnezhad et al. (2011) (28)	22 BD patients, 21 ADHD patients (age range: 10-22 years)	Two open-eyes and closed-eyes resting states through 22 electrodes	Fractal dimensions, autoregressive model coefficients, band power, and wavelet coefficients	a combinatorial classifier based on extended classifier system for function approximation (XCSF) along with linear discriminant analysis (LDA)	Leave-one-out	Best accuracy = 86.44%
Alimardani et al. (2013) (29)	27 BD patients (age: 17.85±3.68 years), 26 schizophrenic patients (age: 20.92±4.29 years)	Closed-eyes resting state through 22 electrodes	Synchronization likelihood, robust synchronization and phase-locking value	Support vector machine	Leave-one-out	Best accuracy = 92.45%
Nazhvani et al. (2013) (30)	12 BD patients, 12 ADHD patients, 12 healthy subjects (age range: 10-22 years)	During 1-Hz visual stimuli in the form of a flash excitement through 22 electrodes	Amplitude and latency of P100 (a visual-evoked potential component)	K-Nearest Neighbor	Leave-one-out	Accuracy = 92.85%
Khaleghi et al. (2015) (31)	18 BD type I patients (age: 15.7±1.5 years), 20 BD type II patients (age: 16.1±1.5 years)	Open-eyes resting state through 22 electrodes	Morphological, time, frequency and time-frequency features as linear features and approximate entropy as a nonlinear feature	Multilayer perceptron neural network	Hold-out	Best accuracy = 91.83%
Erguzel et al. (2016) (32)	31 BD depressive episode patients, 58 unipolar depressive episode patients	closed-eyes resting state through 22 electrodes	Cardance values in delta, theta and alpha frequency bands	Back-propagation neural network	Six-fold cross-validation	Accuracy = 89.89%, sensitivity = 83.87% for BD and 93.1% for unipolar patients
Metin et al. (2018) (33)	18 frontotemporal dementia patients, 20 BD patients (age range: 52-77 years)	closed-eyes resting state through 22 electrodes	Relative frequency band power from F3, F4, T3, T5, T4, T6 electrodes	Multilayer perceptron neural network	Three-fold cross-validation	Accuracy = 76%
Khaleghi et al. (2019) (34)	21 BD type II patients (age: 16.1±1.51 years), 18 healthy subjects (age: 16.3±1.32 years)	Open-eyes resting state through 22 electrodes	Alpha power and alpha entropy	K-Nearest Neighbor	Hold-out	Accuracy = 95.8%, sensitivity = 95.1%,

						specificity = 97.3%
Mateo-Sotos et al. (2022) (35)	105 BD patients, 205 comparison subjects	Resting state through 32 electrodes	Band powers, approximate entropy, fractal dimension, detrended fluctuation analysis, hurst exponent	Extreme gradient boosting	10-fold cross-validation	Accuracy = 94%
Metin et al. (2022) (36)	169 BD type I patients (age: 44.51±14.37 years), 45 healthy subjects (age: 40.88±10.76 years)	closed-eyes resting state through 22 electrodes	Deep learning based features	Convolutional neural network	Hold-out	Accuracy = 96%, sensitivity = 98.45%, specificity = 85.62%
Lei et al. (2022) (37)	82 BD patients (age: 28.94±7.74 years), 101 MDD patients (age: 28.85±5.26 years), 81 healthy subjects (age: 26.26±5.62 years)	Open-eyes resting state through 64 electrodes	Deep learning based features	Convolutional neural network	10-fold cross-validation	Accuracy = 96.88%

Neuroimaging studies

We reviewed 16 articles (38-53) that attempted to classify patients with BD versus healthy subjects using machine learning techniques and brain images. The sample sizes (ranged from 27 to 3020 individuals) in these studies are considerably larger than EEG studies, which greatly contributes to the validity of the findings. The classification accuracy is in the range of 55.9-100%. Diffusion tensor imaging, structural MRI and functional MRI modalities were utilized in these studies. Both grey matter and white matter features were extracted, and most studies utilized vortex-wise features to classify BD from healthy individuals using brain images. K-fold cross-validation is the most widely used method in MRI studies. Support vector machine is the most popular classifier in neuroimaging studies for BD diagnosis. Structural MRI studies achieve 55.9-88.1% accuracies (42-44, 47, 50-52), resting-state functional MRI studies achieve 61.7-82.22% accuracies (39, 45, 49), task-based functional MRI studies achieve 59.72% and 83.5% accuracies (46, 48), and DTI studies achieve 78.12% and 100% accuracies (38, 41). Furthermore, one study used deep learning methods to distinguish BD from healthy controls through structural MRI data and reported a high accuracy of 99.72% (53) (see Table 2).

Table 2. Summary of Magnetic Resonance Imaging Studies on Bipolar Disorder Diagnosis through Machine Learning Techniques

Author (date)	Research sample	Modality	Extracted features	Machine learning technique	Validation method	Outcomes
Besga et al. (2012) (38)	12 BD type I patients (age: 69.55±7.58 years), 25 healthy subjects (age: 71.65±8.55 years)	Diffusion tensor imaging	Fractional anisotropy	Support vector machine	Leave-one-out	Accuracy = 100%, sensitivity = 100%, specificity = 100%
Anticevic et al. (2014) (39)	67 BD adult patients, 47 healthy adults	Resting state functional MRI	Thalamic connectivity map	Support vector machine	Leave-one-out	Accuracy = 61.7%, sensitivity = 75.5%, specificity = 72.2%
Jie et al. (2015) (40)	21 BD patients (age: 21.6±2.9 years), 25 MDD patients (age: 20.1±2.8 years), 23 healthy subjects (age: 20.5±1.9 years)	Resting state structural and functional MRI	Fractional magnitude of low frequency fluctuation, voxel-wise grey matter volume	Support vector machine-based adaptive forward-backward greedy algorithm	Leave-one-out	Accuracy = 80.78%
Mwangi et al. (2015) (41)	16 BD patients (age: 12.24±3.31 years), 16 healthy subjects (age: 12.42±3.06 years)	Diffusion tensor imaging	Radial diffusivity, fractional anisotropy, axial diffusivity	Support vector machine	Leave-one-out	Accuracy = 78.12%, sensitivity = 68.75%, specificity = 87.5%
Sacchet et al. (2015) (42)	40 BD type I patients (age: 37.8±9.6 years), 57 MDD patients (age: 37.1±10.1 years), 35 remitted MDD (age: 42.9± 8.6 years), 61 healthy	Structural MRI	Right and left ventral and caudate diencephalon volume residuals	Support vector machine	10-fold cross-validation	Accuracy = 55.9%

	subjects (age: 37.2±10.4 years)					
Mwangi et al. (2016) (43)	128 BD patients (age: 37.56±11.6 years), 128 healthy subjects (age: 36.33±12.25 years)	Structural MRI	Densities of grey and white matter, voxel-based morphometry	Relevance vector machine	Leave-one-out	Accuracy = 70.3%, sensitivity = 66.4%, specificity = 74.2%
Salvador et al. (2017) (44)	128 BD type I patients (age: 41.4±10.4 years), 128 schizophrenic patients (age: 41.5±10.3 years), 127 healthy subjects (age: 39.8±10.3 years)	Structural MRI	Cortical volume and thickness, white and grey matter voxel-based morphometry images in time and wavelet domains, region of interest based brain volumes	Random forest, support vector machine	10-fold cross-validation	Accuracy = 63%
Li et al. (2017) (45)	22 BD depressive episode patients (age: 28.73±10.11 years), 22 MDD patients (age: 27.68±8.65 years), 22 healthy subjects (age: 28.27±9.55 years)	Resting state functional MRI	Voxel-wise degree centrality maps	Support vector machine	Permutation tests	Accuracy = 81%, sensitivity = 77.27%, specificity = 72.73%
Frangou et al. (2017) (46)	30 BD type I patients (age: 34.7±7.7 years), 30 healthy subjects (age: 35.3±5.6 years)	Functional MRI during n-back test	General linear model coefficients, whole-brain beta maps	Gaussian process classifier	Leave-two-out	Accuracy = 83.5%, sensitivity = 84.6%, specificity = 92.3%
Doan et al. (2017) (47)	190 BD type I and II patients (age: 35.0±11.3 years), 223 schizophrenic patients (32.1±9.3 years), 284 healthy subjects (age: 35.2±9.6 years)	Structural MRI	Thickness maps, grey matter density maps	Random forest classifier	Leave-one-out	Accuracy = 66%, sensitivity = 58%, specificity = 72%
Bürger et al. (2017) (48)	36 BD patients (age: 38.56±12.3 years), 36 MDD patients (age: 40.72±11.58	Functional MRI during a task of emotional faces presentation	Three contrast images generated by an event-related analysis from whole brain and	Support vector machine, Gaussian process classifier	Leave-one-out	Accuracy = 59.72%

	years), 36 healthy subjects (age: 41.33±6.05 years)		regions of interest			
Jie et al. (2018) (49)	22 BD type I patients, 28 MDD patients, 23 healthy subjects (age range: 17-24 years)	Resting state functional MRI	Averaged regional time series, functional connectivity	Support vector machine-based adaptive forward-backward greedy algorithm	Leave-one-out	Accuracy = 82.22%, sensitivity = 81.82%, specificity = 82.61%
Matsuo et al. (2019) (50)	158 BD patients, 596 MDD patients, 777 healthy subjects from Japan; 36 BD patients, 43 MDD patients, 132 healthy subjects from united states; all adult age	Structural MRI	Voxel-based morphometry, regional grey matter volumes	Support vector machine	10-fold cross-validation and leave-one-out	Accuracy = 88.1%, sensitivity = 92.1%, specificity = 73.4% for Japanese sample; accuracy = 58.3%, sensitivity = 50.0%, specificity = 60.6% for united state sample
Schwarz et al. (2019) (51)	222 BD patients, 375 schizophrenic patients, 342 ADHD patients, 1729 healthy subjects; all adult age	Structural MRI	Voxel-based morphometry-based features	Random forest, support vector machine	Bootstrapping	Area under curve = 0.63
Nunes et al. (2020) (52)	853 BD type I and II patients (age: 37.43±11.64 years), 2167 healthy subjects (age: 34.89±12.41 years)	Structural MRI	Subcortical volumes, surface area, cortical thickness	Support vector machine	K-fold cross-validation (K = 3±1)	Accuracy = 65.23%, sensitivity = 66.02%, specificity = 64.9%
Li et al. (2021) (53)	40 BD patients (age: 25.15±6.59 years), 89 first-episode psychosis patients (age: 24.12±6.95 years), 83 healthy subjects (age: 21.45±7.37 years)	Structural MRI	Regional grey matter volumes	Convolutional neural network	10-fold cross-validation	Accuracy = 99.72%

Discussion

We investigated machine learning tasks applied to MRI and EEG data recorded from BD patients. These works present a proof of concept of anticipating BD from brain function above chance level. The reported accuracies for EEG studies is about 90%, while the reported accuracies for MRI studies remains below the minimum level for clinical relevance, i.e. about 80% of classification outcome. However, as will be noted below, most EEG studies suffer from more serious limitations than MRI studies. Comparing different researches in this field is a very challenging task due to heterogeneous methodology. Furthermore, the ways of reporting the findings are very variable and diverse. There is an increasing debate in the literature about the partial reporting of the results of predictive models in machine learning studies. The TRIPOD declaration was recently published to standardize reporting and direct researchers to better implement predictive models (54). None of the reviewed works, however, were written according to the standard reporting method. This led to heterogeneous and poorly comparable reports of findings. In the following, we explain the limitations observed in the literature and provide practical suggestions for future research in this field.

Methodological limitations and practical recommendations

Most of the reviewed studies suffer from several important limitations, including small sample sizes, heterogeneity of the patient group in terms of age, gender, drugs used, type of disorder, disease episode, and cognitive capacities. Almost all EEG studies and some MRI studies have small sample sizes, which considerably reduces the generalizability of their findings. Previous studies have shown that small sample sizes lead to falsely increased accuracies reported in machine learning models (55). In most studies, it is not clear whether patients with BD are in the depressive phase, manic/hypomanic phase (type I or type II), or out of the episode, or whether they are euthymic patients (28-30, 33, 35, 37, 39-41, 43, 48, 50, 51, 53). Meanwhile, previous studies have shown that there are different patterns of brain abnormalities between different episodes of this disorder (10). As a result, this issue can have a substantial effect on the reported findings and, thus, can affect their reliability and generalizability. Therefore, it seems necessary to conduct multicenter studies with large patient populations under fully controlled conditions (i.e. first episode drug-naïve patients, to eliminate possible confounding factors due to disease evolution, BD episode, or medications).

An important point to be noted is the validity and reliability of the findings and the accuracy reported in the literature. It is very important to incorporate the feature selection process in the cross-validation step to achieve a reliable result from the estimation process of classifier performance. In fact, by applying feature selection on the whole dataset before cross-validation, the test dataset will also consist of the most relevant features. As a result, the classifier artificially performs better on such test datasets compared to unseen data. Such an approach causes data leakage, which raises the possibility of overfitting and biases the obtained findings (56). A similar problem occurs when data standardization (i.e., statistical harmonization) is performed on the entire dataset before cross-validation (57). All of the EEG and MRI studies reviewed here suffer from this problem in their machine learning models for BD diagnosis. Therefore, it is strongly recommended that the feature selection process be performed during cross-validation to prevent overfitting to the dataset, hence, allowing for better generalizability and reproducibility of the classification outcomes.

Another issue for statistical control of machine learning performance is to assure that the performance of the classifier is not random. A permutation test is required to evaluate whether the obtained result is at the chance level or not. However, such tests are frequently neglected. Furthermore, balanced accuracy, area under the curve and confusion matrix can be more complete indicators for performance estimation than just accuracy. However, half of the reviewed studies were limited to accuracy index only and the other half reported sensitivity and specificity indices as well. These indicators must be reported systematically by researchers, particularly for unbalanced sample distributions.

Many psychiatric illnesses, including BD, are time-varying in nature. However, all reviewed studies have a cross-sectional design and, to date, no longitudinal research has been published on the diagnosis of BD through machine learning. Such research designs are necessary for the development of prognostic, diagnostic or response to treatment models of machine learning. Furthermore, most publications in this field have focused on the problem of two-group classification, whereas we generally deal with much more complex conditions in psychiatry and psychology. Therefore, we strongly recommend that future works develop new statistical strategies in machine learning models to address comorbidity and multiclass issues.

Many context-based ethical issues may emerge in applied machine learning in BD diagnosis and treatment processes. Currently, machine learning remains a research field and must be accompanied by ethical considerations, since no standard rules have been established to specify when a computer-based program is ethically authorized, given the highly complex circumstances of each individual. In addition, many machine learning models in BD are currently trained based on the judgment of psychiatrists, which can introduce human error into these models and bias the results. More use of unsupervised clustering and classification methods or the use of enhanced models with the help of labeled and unlabeled data (i.e. dynamic decision making) can help solve this problem.

In recent years, promising classification models have been developed based on the concept of deep learning, which can automate the learning process through their abstract representations of raw EEG or MRI data. Here, the reviewed studies using these algorithms seem to achieve better classification performance compared to traditional machine learning methods (36, 37, 53). Long short-term memory (LSTM) and convolutional neural network (CNN) are widely utilized deep learning techniques in these studies to diagnose BD. Deep learning algorithms are capable of automating the feature extraction process and increasing classification accuracy in order to objectively diagnose mental disorders. However, few studies to date have used these techniques for BD diagnosis, and future studies should focus more on deep learning methods for EEG/MRI data classification and automatic BD diagnosis.

Limitation

Although we tried to review many related studies and provide a fair perspective on the research in this field, failure to conduct a systematic review can lead to failure in reviewing all articles and related aspects in this field. However, the authors' experience suggests that the issues raised here exist in most machine learning studies.

Conclusion

Machine learning provides new opportunities for BD diagnostic purposes. Research utilizing machine learning applied to EEG signals and brain images has provided proof of concept for how this innovative technique can help psychiatrists distinguish BD from healthy people. In fact, considering the changes in brain volumes and electrophysiological activity of the brain, multiple researchers have demonstrated that machine learning can differentiate patients with BD from healthy individuals through EEG and MRI data. However, results have been somewhat contradictory and we must keep away from excessive optimistic interpretations of the findings. Much progress is still needed to reach the level of clinical practice in this field.

Acknowledgment

The authors thank the institutions where they work for encouraging the development of research.

Conflict of Interest

None.

References

1. Mohammadi MR, Ahmadi N, Khaleghi A, Mostafavi SA, Kamali K, Rahgozar M, et al. Prevalence and Correlates of Psychiatric Disorders in a National Survey of Iranian Children and Adolescents. Iran J Psychiatry. 2019;14(1):1-15.
2. Khaleghi A, Mohammadi MR, Zandifar A, Ahmadi N, Alavi SS, Ahmadi A, et al. Epidemiology of psychiatric disorders in children and adolescents; in Tehran, 2017. Asian J Psychiatr. 2018;37:146-53.
3. Talepasand S, Mohammadi MR, Alavi SS, Khaleghi A, Sajedi Z, Akbari P, et al. Psychiatric disorders in children and adolescents: Prevalence and sociodemographic correlates in Semnan Province in Iran. Asian J Psychiatr. 2019;40:9-14.
4. Rosa AR, González-Ortega I, González-Pinto A, Echeburúa E, Comes M, Martínez-Àran A, et al. One-year psychosocial functioning in patients in the early vs. late stage of bipolar disorder. Acta Psychiatr Scand. 2012;125(4):335-41.
5. Berk M, Dodd S, Callaly P, Berk L, Fitzgerald P, de Castella AR, et al. History of illness prior to a diagnosis of bipolar disorder or schizoaffective disorder. J Affect Disord. 2007;103(1-3):181-6.
6. Goi PD, Vianna-Sulzbach M, Silveira L, Grande I, Chendo I, Sodré LA, et al. Treatment delay is associated with more episodes and more severe illness staging progression in patients with bipolar disorder. Psychiatry Res. 2015;227(2-3):372-3.
7. Alda M, Manchia M. Personalized management of bipolar disorder. Neurosci Lett. 2018;669:3-9.
8. Giridharan VV, Sayana P, Pinjari OF, Ahmad N, da Rosa MI, Quevedo J, et al. Postmortem evidence of brain inflammatory markers in bipolar disorder: a systematic review. Mol. Psychiatry. 2020;25(1):94-113.
9. Clark L, Sahakian BJ. Cognitive neuroscience and brain imaging in bipolar disorder. Dialogues in clinical neuroscience. 2022.
10. Khaleghi A, Sheikhani A, Mohammadi MR, Moti Nasrabadi A. Evaluation of Cerebral Cortex Function in Clients with Bipolar Mood Disorder I (BMD I) Compared With BMD II Using QEEG Analysis. Iran J Psychiatry. 2015;10(2):93-9.
11. Scaini G, Valvassori SS, Diaz AP, Lima CN, Benevenuto D, Fries GR, et al. Neurobiology of bipolar disorders: a review of genetic components, signaling pathways, biochemical changes, and neuroimaging findings. Braz J Psychiatry. 2020;42(5):536-51.
12. Wang X, Luo Q, Tian F, Cheng B, Qiu L, Wang S, et al. Brain grey-matter volume alteration in adult patients with bipolar disorder under different conditions: a voxel-based meta-analysis. J Psychiatry Neurosci. 2019;44(2):89-101.
13. Xiao W, Manyi G, Khaleghi A. Deficits in auditory and visual steady-state responses in adolescents with bipolar disorder. J Psychiatr Res. 2022;151:368-76.
14. Moeini M, Khaleghi A, Mohammadi MR. Characteristics of Alpha Band Frequency in Adolescents with Bipolar II Disorder: A Resting-State QEEG Study. Iran J Psychiatry. 2015;10(1):8-12.
15. Hibar DP, Westlye LT, Doan NT, Jahanshad N, Cheung JW, Ching CRK, et al. Cortical abnormalities in bipolar disorder: an MRI analysis of 6503 individuals from the ENIGMA Bipolar

Disorder Working Group. Mol Psychiatry. 2018;23(4):932-42.
16. Hibar DP, Westlye LT, van Erp TG, Rasmussen J, Leonardo CD, Faskowitz J, et al. Subcortical volumetric abnormalities in bipolar disorder. Mol Psychiatry. 2016;21(12):1710-6.
17. Khaleghi A, Mohammadi MR, Shahi K, Nasrabadi AM. Computational Neuroscience Approach to Psychiatry: A Review on Theory-driven Approaches. Clin Psychopharmacol Neurosci. 2022;20(1):26-36.
18. Yasin S, Hussain SA, Aslan S, Raza I, Muzammel M, Othmani A. EEG based Major Depressive disorder and Bipolar disorder detection using Neural Networks:A review. Comput Methods Programs Biomed. 2021;202:106007.
19. Mohammadi MR, Khaleghi A, Nasrabadi AM, Rafieivand S, Begol M, Zarafshan H. EEG classification of ADHD and normal children using non-linear features and neural network. Biomed Eng Lett. 2016;6(2):66-73.
20. Zarafshan H, Khaleghi A, Mohammadi MR, Moeini M, Malmir N. Electroencephalogram complexity analysis in children with attention-deficit/hyperactivity disorder during a visual cognitive task. J Clin Exp Neuropsychol. 2016;38(3):361-9.
21. Murphy KP. Probabilistic machine learning: an introduction: MIT press; 2022.
22. Ayodele TO. Machine learning overview. Mach Lear. 2010;2:9-18.
23. Claude LA, Houenou J, Duchesnay E, Favre P. Will machine learning applied to neuroimaging in bipolar disorder help the clinician? A critical review and methodological suggestions. Bipolar Disord. 2020;22(4):334-55.
24. Passos IC, Mwangi B, Cao B, Hamilton JE, Wu MJ, Zhang XY, et al. Identifying a clinical signature of suicidality among patients with mood disorders: A pilot study using a machine learning approach. J Affect Disord. 2016;193:109-16.
25. Redlich R, Opel N, Grotegerd D, Dohm K, Zaremba D, Bürger C, et al. Prediction of Individual Response to Electroconvulsive Therapy via Machine Learning on Structural Magnetic Resonance Imaging Data. JAMA Psychiatry. 2016;73(6):557-64.
26. Librenza-Garcia D, Kotzian BJ, Yang J, Mwangi B, Cao B, Pereira Lima LN, et al. The impact of machine learning techniques in the study of bipolar disorder: A systematic review. Neurosci Biobehav Rev. 2017;80:538-54.
27. Passos IC, Ballester PL, Barros RC, Librenza-Garcia D, Mwangi B, Birmaher B, et al. Machine learning and big data analytics in bipolar disorder: A position paper from the International Society for Bipolar Disorders Big Data Task Force. Bipolar Disord. 2019;21(7):582-94.
28. Sadatnezhad K, Boostani R, Ghanizadeh A. Classification of BMD and ADHD patients using their EEG signals. Expert Systems with Applications. 2011;38(3):1956-63.
29. Alimardani F, Boostani R, Azadehdel M, Ghanizadeh A, Rastegar K. Presenting a new search strategy to select synchronization values for classifying bipolar mood disorders from schizophrenic patients. Eng Appl Artif Intell. 2013;26(2):913-23.
30. Nazhvani AD, Boostani R, Afrasiabi S, Sadatnezhad K. Classification of ADHD and BMD patients using visual evoked potential. Clin Neurol Neurosurg. 2013;115(11):2329-35.
31. Khaleghi A, Sheikhani A, Mohammadi MR, Nasrabadi AM, Vand SR, Zarafshan H, et al. EEG classification of adolescents with type I and type II of bipolar disorder. Australas Phys Eng Sci Med. 2015;38(4):551-9.
32. Erguzel TT, Sayar GH, Tarhan N. Artificial intelligence approach to classify unipolar and bipolar depressive disorders. Neural Comput Appl. 2016;27(6):1607-16.
33. Metin SZ, Erguzel TT, Ertan G, Salcini C, Kocarslan B, Cebi M, et al. The Use of Quantitative EEG for Differentiating Frontotemporal Dementia From Late-Onset Bipolar Disorder. Clin EEG Neurosci. 2018;49(3):171-6.
34. Khaleghi A, Mohammadi MR, Moeini M, Zarafshan H, Fadaei Fooladi M. Abnormalities of Alpha Activity in Frontocentral Region of the Brain as a Biomarker to Diagnose Adolescents With Bipolar Disorder. Clin EEG Neurosci. 2019;50(5):311-8.
35. Mateo-Sotos J, Torres A, Santos J, Quevedo O, Basar C. A Machine Learning-Based Method to Identify Bipolar Disorder Patients. Analog Integr Circuits Signal Process. 2022;41(4):2244-65.
36. Metin B, Uyulan Ç, Ergüzel TT, Farhad S, Çifçi E, Türk Ö, et al. The Deep Learning Method Differentiates Patients with Bipolar Disorder from Controls with High Accuracy Using EEG Data. Clin EEG Neurosci. 2022:15500594221137234.
37. Lei Y, Belkacem AN, Wang X, Sha S, Wang C, Chen C. A convolutional neural network-based diagnostic method using resting-state electroencephalograph signals for major depressive and bipolar disorders. Biomed Signal Process Control. 2022;72:103370.
38. Besga A, Termenon M, Graña M, Echeveste J, Pérez JM, Gonzalez-Pinto A. Discovering Alzheimer's disease and bipolar disorder white matter effects building computer aided diagnostic systems on brain diffusion tensor imaging features. Neurosci Lett. 2012;520(1):71-6.
39. Anticevic A, Cole MW, Repovs G, Murray JD, Brumbaugh MS, Winkler AM, et al. Characterizing thalamo-cortical disturbances in schizophrenia and bipolar illness. Cereb Cortex. 2014;24(12):3116-30.
40. Jie NF, Zhu MH, Ma XY, Osuch EA, Wammes M, Théberge J, et al. Discriminating Bipolar Disorder From Major Depression Based on SVM-FoBa: Efficient Feature Selection With Multimodal Brain Imaging Data. IEEE Trans Auton Ment Dev. 2015;7(4):320-31.
41. Mwangi B, Wu MJ, Bauer IE, Modi H, Zeni CP, Zunta-Soares GB, et al. Predictive classification of pediatric bipolar disorder using atlas-based

diffusion weighted imaging and support vector machines. Psychiatry Res. 2015;234(2):265-71.
42. Sacchet MD, Livermore EE, Iglesias JE, Glover GH, Gotlib IH. Subcortical volumes differentiate Major Depressive Disorder, Bipolar Disorder, and remitted Major Depressive Disorder. J Psychiatr Res. 2015;68:91-8.
43. Mwangi B, Wu MJ, Cao B, Passos IC, Lavagnino L, Keser Z, et al. Individualized Prediction and Clinical Staging of Bipolar Disorders using Neuroanatomical Biomarkers. Biol Psychiatry Cogn Neurosci Neuroimaging. 2016;1(2):186-94.
44. Salvador R, Radua J, Canales-Rodríguez EJ, Solanes A, Sarró S, Goikolea JM, et al. Evaluation of machine learning algorithms and structural features for optimal MRI-based diagnostic prediction in psychosis. PLoS One. 2017;12(4):e0175683.
45. Li M, Das T, Deng W, Wang Q, Li Y, Zhao L, et al. Clinical utility of a short resting-state MRI scan in differentiating bipolar from unipolar depression. Acta Psychiatr Scand. 2017;136(3):288-99.
46. Frangou S, Dima D, Jogia J. Towards person-centered neuroimaging markers for resilience and vulnerability in Bipolar Disorder. Neuroimage. 2017;145(Pt B):230-7.
47. Doan NT, Kaufmann T, Bettella F, Jørgensen KN, Brandt CL, Moberget T, et al. Distinct multivariate brain morphological patterns and their added predictive value with cognitive and polygenic risk scores in mental disorders. Neuroimage Clin. 2017;15:719-31.
48. Bürger C, Redlich R, Grotegerd D, Meinert S, Dohm K, Schneider I, et al. Differential Abnormal Pattern of Anterior Cingulate Gyrus Activation in Unipolar and Bipolar Depression: an fMRI and Pattern Classification Approach. Neuropsychopharmacology. 2017;42(7):1399-408.
49. Jie N-F, Osuch EA, Zhu M-H, Wammes M, Ma X-Y, Jiang T-Z, et al. Discriminating bipolar disorder from major depression using whole-brain functional connectivity: a feature selection analysis with SVM-FoBa algorithm. J Signal Process Syst. 2018;90(2):259-71.
50. Matsuo K, Harada K, Fujita Y, Okamoto Y, Ota M, Narita H, et al. Distinctive Neuroanatomical Substrates for Depression in Bipolar Disorder versus Major Depressive Disorder. Cereb Cortex. 2019;29(1):202-14.
51. Schwarz E, Doan NT, Pergola G, Westlye LT, Kaufmann T, Wolfers T, et al. Reproducible grey matter patterns index a multivariate, global alteration of brain structure in schizophrenia and bipolar disorder. Transl Psychiatry. 2019;9(1):12.
52. Nunes A, Schnack HG, Ching CRK, Agartz I, Akudjedu TN, Alda M, et al. Using structural MRI to identify bipolar disorders - 13 site machine learning study in 3020 individuals from the ENIGMA Bipolar Disorders Working Group. Mol Psychiatry. 2020;25(9):2130-43.
53. Li Z, Li W, Wei Y, Gui G, Zhang R, Liu H, et al. Deep learning based automatic diagnosis of first-episode psychosis, bipolar disorder and healthy controls. Comput Med Imaging Graph. 2021;89:101882.
54. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. Bmj. 2015;350:g7594.
55. Varoquaux G. Cross-validation failure: Small sample sizes lead to large error bars. Neuroimage. 2018;180(Pt A):68-77.
56. Shim M, Lee SH, Hwang HJ. Inflated prediction accuracy of neuropsychiatric biomarkers caused by data leakage in feature selection. Sci Rep. 2021;11(1):7980.
57. Kernbach JM, Staartjes VE. Foundations of Machine Learning-Based Clinical Prediction Modeling: Part II-Generalization and Overfitting. Acta Neurochir Suppl. 2022;134:15-21.