EEG Signal Analysis for Automatic Detection of Psychiatric Diseases

Saidi Meryem

January 2, 2024

EEG Signal Analysis for Automatic Detection of psychiatric diseases

1st Meryem Saidi
Higher School of Management of Tlemcen
Biomedical Engineering Laboratory, Tlemcen University
Tlemcen, Algeria
email address or ORCID

***Abstract*—Neurological diseases such as epilepsy, Parkinson's disease, and Alzheimer, or psychiatric diseases like depression, personality disorders, schizophrenia, or addictive behavior, among many others, affect numerous people around the world. Diagnosing these diseases is a challenge in medicine. The symptoms of neurological and psychiatric diseases can vary greatly, making it difficult for healthcare professionals to accurately diagnose and treat patients. Hence, it is based on the interpretation of symptoms by doctors and the analysis of EEG signals (the electroencephalogram).**
There is therefore an important need for diagnostic support systems to assist doctors in their decision-making. Particularly in Algeria, where the health system suffers from a shortage of doctors specializing in neurology and psychiatry.
In this work, we compare the performance of different methods used in creating systems to assist in the medical diagnosis of an EEG signal in the case of a mental illness. This system will be designed using intelligent algorithms on electroencephalogram signals (EEGs), which are generally non-stable and complex and whose interpretation is long and laborious. Hence, the application of artificial learning algorithms such as random forest (RF) or deep neuronal networks.

This study brings light to the importance of machine learning algorithms in significantly reducing the time and effort required for interpreting EEG signals. We also raised the critical short-come of these same algorithms under some conditions in real-world problems, such as imbalanced datasets.
***Index Terms*—EEG analysis, neurological diseases, psychiatric disorder, machine learning, automatic diagnosis.**

I. Introduction

Neurological and psychiatric diseases can have a profound impact on individuals and their families, affecting their quality of life and overall well-being. Early detection and diagnosis of these conditions play a crucial role in ensuring timely intervention and appropriate treatment. Some common neurological and psychiatric diseases include Alzheimer's disease, Parkinson's disease, multiple sclerosis, epilepsy, depression, personality disorders, and schizophrenia. Each of these diseases has its own unique characteristics and can present a variety of symptoms.
By identifying symptoms and risk factors at an early stage, healthcare professionals can implement targeted interventions and support systems, ultimately improving outcomes for patients.
EEG signal analysis refers to the process of extracting meaningful information from electroencephalogram (EEG) signals, which are electrical brain wave recordings. This analysis plays a crucial role in various fields, including neuroscience, clinical medicine, neuromarketing, and brain-computer interfaces. By analyzing EEG signals, researchers, and clinicians can gain insights into brain activity patterns, identify abnormalities or anomalies, and even decode specific mental states or intentions.
This analysis plays a crucial role in understanding brain functioning, diagnosing neurological and psychiatric disorders, and monitoring the effectiveness of treatments. By deciphering the patterns and abnormalities in EEG signals, researchers and medical professionals gain valuable insights into brain activity, cognitive processes, and overall brain health.

Diagnoses in the field of psychology are made on a phenomenological and categorical basis. Clinicians assess explicit and observable signs and symptoms and offer categorical diagnoses based on which those symptoms fit into, in accordance with the "International Classification of Disorders (ICD)" and the "Diagnostic and Statistical Manual for Mental Disorders (DSM)" [1], [2]. This descriptive nosology improves communication clarity, but it has limitations due to its reliance on clinician observation and/or presenting problems recorded by patients or informants, which may not be sufficiently objective [3].
By analyzing the patterns and abnormalities in EEG signals through machine learning algorithms, healthcare professionals can save valuable time and resources while ensuring a more objective evaluation of patients. This can greatly help the detection and monitoring of the progression of the disease and evaluate the effectiveness of different treatment strategies, enabling healthcare professionals to make informed decisions and adjustments to the treatment plan.

In this work, we focus on the use of advanced algorithms and machine learning techniques to analyze the EEG signal to identify patterns and markers of different psychiatric conditions. We briefly reviews existing methods and presents results of applying these classification algorithms to a psychiatric disorder dataset. The rest of this article is organized into six sections. In the second one we present a review of some existing methods of EEG signal analysis techniques for the automatic detection of neurological and psychiatric diseases . In third one, we introduce the used dataset. In the fourth section, we present the results of applying some of the well-known classification algorithms on the psychiatric dataset and we discuss the limitations and challenges faced during the analysis process. Finally, we present a conclusion that summarized the observations and the possible perspectives for future work.

II. EEG Signal Analysis Techniques

The brain is an incredibly complex and interesting organ that is responsible for everything from our thoughts and emotions to our physical movements and sensations. Cerebral electromagnetic activity is what scientists use to refer to the electrical and magnetic fields that are generated by brain neural activity. These fields are due to the flow of ions across the membranes of neurons, which creates an electrical potential that can be measured outside the brain. These electrical and magnetic fields can be detected using techniques such as electroencephalography (EEG).

Advanced machine learning algorithms can be applied to EEG signals to classify different brain states or detect specific abnormalities. These algorithms can be trained to recognize patterns in the EEG signals that are characteristic of different brain states, such as sleep, wakefulness, or certain mental disorders. By analyzing the EEG data in real-time, these algorithms can provide valuable insights into the brain's functioning and aid in the diagnosis and treatment of various neurological conditions. Furthermore, the integration of EEG with other neuro-imaging techniques, such as fMRI or PET scans, can provide a more comprehensive understanding of brain activity and its relation to different states or abnormalities.

A. Automatic Detection of Neurological and Psychiatric Diseases

In recent years, there has been significant progress in the development of automatic detection systems based on machine learning to analyze EEG data and identify potential signs of mental disorders.

Some research was conducted to detect abnormal EEG signals using machine learning classifiers as SVM [4], ensemble learning [5] or deep learning [6], [7], [8], [9]. Other research aims to identify a specific disease, such as epileptic seizures, using deep learning [10], SVM [11], [12] or linear programming boosting [13]. Saminu & al. [14] present a rigorous review of the existing methods.

These studies have shown promising results in accurately identifying mental disorders by detecting patterns and abnormalities in EEG signals. However, further research and validation are still needed to ensure the reliability and effectiveness of these systems in real-world clinical settings.

New intelligent approaches are presented to enable an earlier and more accurate diagnosis of Alzheimer disease; we cite SVM classifier [15], deep learning [16], [17], [18] or decision tree and K-nearest neighbor [19]. Parkinson disease detection is also an important research topic; Loh & al. have provided a comprehensive review of the existing literature on Parkinson's disease detection, shedding light on various approaches and techniques that have been explored in this field. This review serves as a valuable resource for researchers and clinicians seeking to enhance the accuracy and timeliness of Parkinson's disease diagnosis. [20].

Another important topic is the development of decision support systems for automatic detection of psychiatric diseases. These decision support systems leverage various techniques to identify such as Schizophrenia [18], [21], [22], autism [23], [24], [25], depression [26], [27], [28] or addiction [29], [30]. EEG-based approaches have shown promising results in detecting disorders, highlighting the potential for using these systems in a wide range of psychiatric diagnoses.

The aforementioned works, despite their promising results are limited to a single specific disorder. However, in [3] Park & al. intended to create new classification system for identifying patients with severe psychiatric illnesses from healthy controls. They gathered EEG data from patients with schizophrenia, mood disorders, anxiety disorders, obsessive-compulsive disorders, addictive disorders, trauma or stress-related disorders using RF and SVM algorithms.

This study aims to evaluate the performances of various classifiers including classification tree, random forests, and deep neural networks in classifying psychiatric disorders based on EEG data. By comparing the performances of these classifiers, we intend to identify the most effective approach for accurately diagnosing conditions such as depression, schizophrenia, and bipolar disorder using EEG signals. This research has the potential to contribute to the development of more reliable and efficient diagnostic tools for psychiatric disorders.

III. Method

ITo study the performance of different machine learning algorithms to identify and compare multiple psychiatric disorders using electroencephalography (EEG). We employ :

Classification And Regression Tree (CART): Proposed by Breiman et al. in (1984) [31], it is a supervised learning algorithm that uses the Gini index to find the best possible variable to split the node into two child nodes. The tree is grown to their maximum size until no splits are possible.

The K-nearest neighbors (K-nn): It is a non-parametric method based on a a calculation of distances between the characteristic vector of the instance to be classified and the vectors of the instances of the learning base. Then the instance to be classified is assigned the majority class among the classes of the k closest instances [32].

the Random Forest (RF) : Introduced by Breiman in 2001 [33], it is an ensemble method formed by a collection of trees, each tree contributes to the final decision made by majority vote.

Long Short-Term Memory (LSTM): is a variation of Recurrent Neural Network (RNN) that is intended to handle sequential data. This algorithm has a memory cell that is capable of long-term information storage. LSTM networks are capable of learning long-term dependencies in sequential data. They have the capability of identifying patterns in data that deviate from the norm [34].

We compare the results based on accuracy, precision, recall and F-score.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

$$Precision = \frac{TP}{TP+FP}$$

$$Recall = \frac{TP}{TP+FN}$$

$$F-score = 2 * \frac{Precision*Recall}{Precision+Recall}$$

IV. Dataset

Available EEG psychiatric dataset are rare making it difficult for researchers and clinicians to access sufficient data for their studies and diagnostic purposes. This scarcity of available EEG datasets hinders progress in understanding and treating psychiatric disorders, as EEG data plays a crucial role in understanding brain activity and identifying biomarkers. The quantitative EEG (QEEG) at resting-state assessments, medical records, psychological test batteries, and other data were gathered retrospectively from the Seoul Metropolitan Government-Seoul National University (SMG-SNU) Boramae Medical Center in Seoul, South Korea, from January 2011 to December 2018. The dataset included 945 subjects and the diagnostic decision for the patients was established from March 2019 to August 2019, by two psychiatrists and two psychologists. The EEG data included 5 min eyes-closed resting-state with 19 or 64 channels acquired with 500–1,000 Hz sampling rate and 0.1–100 on-line filters via Neuroscan. The data were down-sampled to 128 Hz, and 19 channels were selected based on the international 10–20 system. [3]. Figure 1 and table I represent the different disorder categories and their distribution.

V. Results and discussion

This section goal is to determine the best framework to perform an automatic diagnosis. This study employs different classifier CART, RF, K-nn, and DeepL.

The CART, RF, K-nn, and DeepL classifiers were chosen due to their proven effectiveness in similar studies. Each classifier was trained and tested on a dataset consisting of labeled samples belonging to different classes. The accuracy of each classifier was measured by comparing its predictions to the actual class labels.

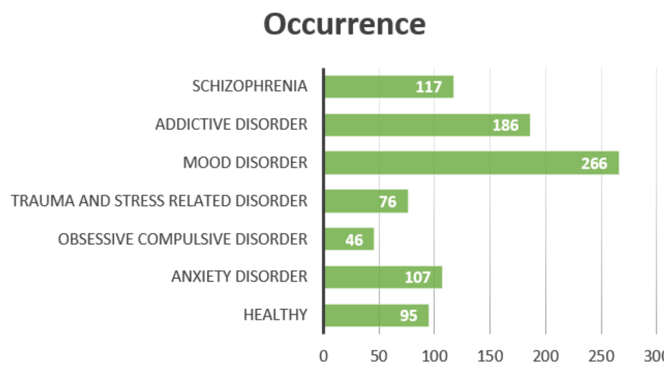

Fig. 1. Disorder categories distribution

TABLE I
Disorder categories frequencies

Main disorder	Occurrence
Healthy	95
Anxiety disorder	107
Obsessive compulsive disorder	46
Trauma and stress related disorder	76
Mood disorder	266
Addictive disorder	186
Schizophrenia	117

Each classifier is tested under various conditions such as number of classes. The performance of each classifier is evaluated based on metrics such as accuracy, precision, recall, and F-score.

Precision refers to the ability of a classifier to correctly identify positive cases, while recall measures its ability to find all positive cases. The F-score takes into account both precision and recall, providing a more balanced assessment of a classifier's performance. By comparing the performance of each classifier using these metrics, we can determine which framework is the most suitable for an automatic diagnosis system in this specific context.

The extremely bad scores above (table II) indicate that the dataset used in this study is not suitable for multi-class classification.
Additionally, the poor results may also be attributed to the choice of classification algorithm used. Considering alternative algorithms and tuning their parameters may lead to better performance.

TABLE II
MULTI-CLASSIFICATION RESULTS

Multi-classification	Accuracy	Precision	Recall	F-score
CART	24,55	19,33	19,43	19,15
K-nn	21,48	16,79	16,07	14,97
RF	29,20	21,66	18,19	14,99
DeepL	28,16	0,08	0,28	0,12

Moreover, it is important to consider other factors such as data quality, the complexity of the problem, or the distribution of the classes. Ultimately, a comprehensive evaluation of all these factors is necessary to determine the most suitable approach for multi-class classification in this study.
After further investigation and experimentation we adjust the framework to improve the classification performance by performing a binary classification.

TABLE III
BINARY CLASSIFICATION RESULTS

Binary classification	Accuracy	Precision	Recall	F-score
CART	84,65	50,76	52,71	51,33
K-nn	88,67	63,22	52,68	53,39
RF	89,52	66,96	52,74	52,90
DeepL	90.70	61.21	51.07	49.59

Therefore, in a second experimentation, we implement a binary classification of healthy EEG signal vs. abnormal EEG signal (psychiatric disorder) to gain insights into the distinguishability of healthy controls from specific disorders and pave the way for more refined multi-class classification models in the future.
This binary classification approach allows us to accurately identify the presence of psychiatric disorders based on distinct EEG signal characteristics, which can greatly improve diagnostic accuracy. With the knowledge gained from this experimentation, we can develop more advanced and precise multi-class classification models that can differentiate between various psychiatric disorders with greater accuracy and efficiency.
Table III shows the obtained results of the binary classification task. The binary classification lead to significant improvements in the classification performance. Based on these results, the Deep learning model (90.70%) reaches the first place in term of classification accuracy followed by the RF (89.52%), k-nn (88.67%) algorithm and CART (84.64%) respectively. However, when it comes to precision and F-score, the RF model outperforms the Deep learning model. This suggests that while DeepL may be better at overall classification accuracy, the RF model is more reliable in correctly identifying positive instances within the categories.

However, the precision and f-score remain insufficient, 51,33% for CART, 53,39% for K-nn, 52,90% for RF and 49.59 % for the Deep learning algorithm.

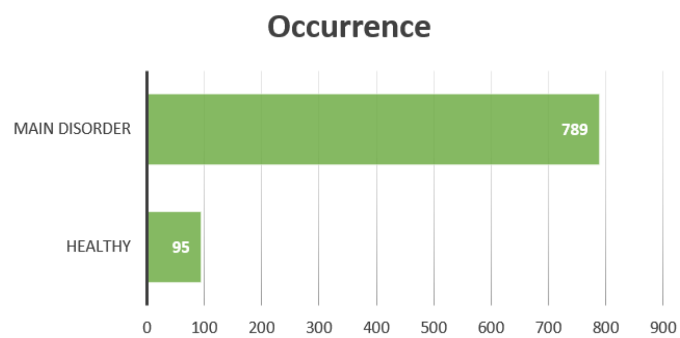

Fig. 2. healthy EEG signal vs. abnormal EEG signal distribution

TABLE IV
DISORDER CATEGORIES FREQUENCIES

Diagnosis	Occurrence
Healthy	95
Abnormal	789

Further analysis to understand the reasons behind these differences suggests that the lack of data for certain classes (see figure 2, table IV), leads to imbalanced training sets and affects the performance of the model.

Additionally, the chosen algorithm may not be the most suitable for the given problem. Indeed, traditional approaches presuppose that all errors are equal and aim to reduce misclassification errors.
However, the dataset being used is very unbalanced, and traditional classifiers frequently misclassify instances belonging to minority classes. The solution is the use of resampling techniques and cost-sensitive algorithms.

These techniques aim to address the issue of imbalanced training sets by either oversampling the minority class or undersampling the majority class. Oversampling involves duplicating instances from the minority class to balance the dataset, while undersampling involves removing instances from the majority class.

Cost-sensitive algorithms, on the other hand, assign different misclassification costs to different classes, prioritizing the minority class and reducing the bias towards the majority class. By utilizing these approaches, the performance of the model can be significantly improved in scenarios with imbalanced datasets.

VI. Conclusion

Applications of machine learning algorithms in EEG signal analysis have shown promising results in various medical fields, including the detection and diagnosis of neurological or psychiatric disorders such as epilepsy, sleep disorders, or schizophrenia.

By leveraging the power of machine learning algorithms, researchers have been able to extract meaningful information from EEG signals, enabling accurate and efficient detection of abnormalities. These algorithms cannot only identify specific patterns associated with different neurological conditions, but, in some cases, they allow the differentiation between various stages of a disease.

In this work, we have compared the performance of different well-known machine learning algorithms. The results have shown that the classical methods perform poorly in terms of handling imbalanced datasets. Since they tend to misclassify the minority class. This limitation can have serious consequences in the context of neurological disorder diagnosis, as early detection is crucial for effective treatment and most EEG signals datasets tend to be imbalanced.

In future works, we seek to use resampling techniques and cost-sensitive algorithms to improve performance. Since these techniques help in mitigating the bias towards the majority class and ensuring a more balanced classification.

VII. Acknowledgements

Park, S. M. (2021, August 16). EEG machine learning. Retrieved from osf.io/8bsvr

References

[1] W. H. Organization, *The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines*. World Health Organization, 1992, vol. 1.

[2] A. American Psychiatric Association, A. P. Association *et al.*, *Diagnostic and statistical manual of mental disorders: DSM-IV*. American psychiatric association Washington, DC, 1994, vol. 4.

[3] S. M. Park, B. Jeong, D. Y. Oh, C.-H. Choi, H. Y. Jung, J.-Y. Lee, D. Lee, and J.-S. Choi, "Identification of major psychiatric disorders from resting-state electroencephalography using a machine learning approach," *Frontiers in Psychiatry*, vol. 12, p. 707581, 2021.

[4] F. A. Alturki, K. AlSharabi, A. M. Abdurraqeeb, and M. Aljalal, "Eeg signal analysis for diagnosing neurological disorders using discrete wavelet transform and intelligent techniques," *Sensors*, vol. 20, no. 9, p. 2505, 2020.

[5] T. Wu, X. Kong, Y. Zhong, and L. Chen, "Automatic detection of abnormal eeg signals using multiscale features with ensemble learning," *Frontiers in Human Neuroscience*, vol. 16, p. 943258, 2022.

[6] H. Albaqami, G. M. Hassan, and A. Datta, "Automatic detection of abnormal eeg signals using wavenet and lstm," *Sensors*, vol. 23, no. 13, p. 5960, 2023.

[7] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggensperger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball, "Deep learning with convolutional neural networks for eeg decoding and visualization," *Human brain mapping*, vol. 38, no. 11, pp. 5391–5420, 2017.

[8] S. Roy, I. Kiral-Kornek, and S. Harrer, "Chrononet: A deep recurrent neural network for abnormal eeg identification," in *Artificial Intelligence in Medicine: 17th Conference on Artificial Intelligence in Medicine, AIME 2019, Poznan, Poland, June 26–29, 2019, Proceedings 17*. Springer, 2019, pp. 47–56.

[9] H. Albaqami, G. M. Hassan, A. Subasi, and A. Datta, "Automatic detection of abnormal eeg signals using wavelet feature extraction and gradient boosting decision tree," *Biomedical Signal Processing and Control*, vol. 70, p. 102957, 2021.

[10] M. Varlı and H. Yılmaz, "Multiple classification of eeg signals and epileptic seizure diagnosis with combined deep learning," *Journal of Computational Science*, vol. 67, p. 101943, 2023.

[11] A. K. Jaiswal and H. Banka, "Epileptic seizure detection in eeg signal using machine learning techniques," *Australasian physical & engineering sciences in medicine*, vol. 41, pp. 81–94, 2018.

[12] O. Faust, U. R. Acharya, L. C. Min, and B. H. Sputh, "Automatic identification of epileptic and background eeg signals using frequency domain parameters," *International journal of neural systems*, vol. 20, no. 02, pp. 159–176, 2010.

[13] A. R. Hassan and A. Subasi, "Automatic identification of epileptic seizures from eeg signals using linear programming boosting," *computer methods and programs in biomedicine*, vol. 136, pp. 65–77, 2016.

[14] S. Saminu, G. Xu, S. Zhang, I. Abd El Kader, H. A. Aliyu, A. H. Jabire, Y. K. Ahmed, and M. J. Adamu, "Applications of artificial intelligence in automatic detection of epileptic seizures using eeg signals: A review," 2022.

[15] L. R. Trambaiolli, A. C. Lorena, F. J. Fraga, P. A. Kanda, R. Anghinah, and R. Nitrini, "Improving alzheimer's disease diagnosis with machine learning techniques," *Clinical EEG and neuroscience*, vol. 42, no. 3, pp. 160–165, 2011.

[16] C. L. Alves, A. M. Pineda, K. Roster, C. Thielemann, and F. A. Rodrigues, "Eeg functional connectivity and deep learning for automatic diagnosis of brain disorders: Alzheimer's disease and schizophrenia," *Journal of Physics: complexity*, vol. 3, no. 2, p. 025001, 2022.

[17] Y. Zhao and L. He, "Deep learning in the eeg diagnosis of alzheimer's disease," in *Computer Vision-ACCV 2014 Workshops: Singapore, Singapore, November 1-2, 2014, Revised Selected Papers, Part I 12*. Springer, 2015, pp. 340–353.

[18] Y. Zhao, "Deep learning in the eeg diagnosis of alzheimer's disease 2014 asian conf," *Computer Visison*, pp. 1–15, 2014.

[19] D. Pirrone, E. Weitschek, P. Di Paolo, S. De Salvo, and M. C. De Cola, "Eeg signal processing and supervised machine learning to early diagnose alzheimer's disease," *Applied sciences*, vol. 12, no. 11, p. 5413, 2022.

[20] H. W. Loh, W. Hong, C. P. Ooi, S. Chakraborty, P. D. Barua, R. C. Deo, J. Soar, E. E. Palmer, and U. R. Acharya, "Application of deep learning models for automated identification of parkinson's disease: a review (2011–2021)," *Sensors*, vol. 21, no. 21, p. 7034, 2021.

[21] S. K. Khare and V. Bajaj, "A hybrid decision support system for automatic detection of schizophrenia using eeg signals," *Computers in Biology and Medicine*, vol. 141, p. 105028, 2022.

[22] Z. Aslan and M. Akin, "A deep learning approach in automated detection of schizophrenia using scalogram images of eeg signals," *Physical and Engineering Sciences in Medicine*, vol. 45, no. 1, pp. 83–96, 2022.

[23] M. N. A. Tawhid, S. Siuly, H. Wang, F. Whittaker, K. Wang, and Y. Zhang, "A spectrogram image based intelligent technique for automatic detection of autism spectrum disorder from eeg," *Plos one*, vol. 16, no. 6, p. e0253094, 2021.

[24] M. Radhakrishnan, K. Ramamurthy, K. K. Choudhury, D. Won, and T. A. Manoharan, "Performance analysis of deep learning models for detection of autism spectrum disorder from eeg signals." *Traitement du Signal*, vol. 38, no. 3, 2021.

[25] M. Liao, H. Duan, G. Wang *et al.*, "Application of machine learning techniques to detect the children with autism spectrum disorder," *Journal of Healthcare Engineering*, vol. 2022, 2022.

[26] A. Rafiei, R. Zahedifar, C. Sitaula, and F. Marzbanrad, "Automated detection of major depressive disorder with eeg signals: a time series classification using deep learning," *IEEE Access*, vol. 10, pp. 73 804–73 817, 2022.

[27] P. P. Thoduparambil, A. Dominic, and S. M. Varghese, "Eeg-based deep learning model for the automatic detection of clinical depression," *Physical and Engineering Sciences in Medicine*, vol. 43, pp. 1349–1360, 2020.

[28] G. Sharma, A. Parashar, and A. M. Joshi, "Dephnn: a novel hybrid neural network for electroencephalogram (eeg)-based screening of depression," *Biomedical signal processing and control*, vol. 66, p. 102393, 2021.

[29] W. Mumtaz, N. Kamel, S. S. A. Ali, A. S. Malik *et al.*, "An eeg-based functional connectivity measure for automatic detection of alcohol use disorder," *Artificial intelligence in medicine*, vol. 84, pp. 79–89, 2018.

[30] W. Mumtaz, P. L. Vuong, L. Xia, A. S. Malik, and R. B. A. Rashid, "An eeg-based machine learning method to screen alcohol use disorder," *Cognitive neurodynamics*, vol. 11, pp. 161–171, 2017.

[31] L. Breiman, J. Friedman, R. Olshen, and C. Stone, "Classification and regression trees. wadsworth int," *Group*, vol. 37, no. 15, pp. 237–251, 1984.

[32] N. S. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression," *The American Statistician*, vol. 46, no. 3, pp. 175–185, 1992.

[33] L. Breiman and R. Cutler, "Random forests machine learning [j]," *journal of clinical microbiology*, vol. 2, pp. 199–228, 2001.

[34] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.