



Conference Article

# Prediction of Schizophrenia Using Feature Extraction Methods with EEG Data

Osman KÜÇÜK<sup>1\*</sup>, İsmail CANTÜRK<sup>2</sup>

<sup>1</sup> Yildiz Technical University, Orcid ID: <https://orcid.org/0009-0002-1745-5952>, e-mail: osman.kucuk@yildiz.edu.tr

<sup>2</sup> Yildiz Technical University, Orcid ID: <https://orcid.org/0000-0003-0690-1873>, e-mail: icanturk@yildiz.edu.tr

\* Correspondence: osman.kucuk@yildiz.edu.tr

Received 16 October 2024

Received in revised form 06 December 2024

In final form 27 December 2024

4th International Conference on Design, Research and Development  
(RDCONF 2024)  
December 19 - 20, 2024

**Reference:** Küçük, O., & CANTÜRK, İ. (2024). Prediction of schizophrenia using feature extraction methods with EEG data. *Orclever Proceedings of Research and Development*, 5(1), 210-214.

## Abstract

*Schizophrenia is a mental disorder that causes some motor dysfunctions in individuals and causes psychotic symptoms. It is believed that machine learning algorithms offer support in the detection and treatment process of the disease. In this study, a system that predicts schizophrenia disease with machine learning algorithms is proposed using resting EEG data. Filtering process, feature extraction methods and cross-validation were performed before machine learning.*

**Keywords:** Biomedical Engineering, Machine Learning, Schizophrenia, EEG

## 1. Introduction

The widespread use of machine learning (ML) algorithms in disease detection demonstrates the potential benefits of this technology in the medical field [1-6] and it holds promise for aiding in the detection of schizophrenia [7, 8]. It is a psychological disorder that impacts roughly 1% of the global population. Hallucinations and delusions



are among the non-motor symptoms of the disease. Additionally, motor function disorders can also be observed in patients with schizophrenia[9]. Since schizophrenia diagnosis is revealed in line with clinical symptoms, a more objective and reliable measurement that can help psychiatrists in this process can provide convenience. This can provide early diagnosis and treatment of the disease [10].

De Miras et al. [9] investigated the effectiveness of machine learning methods in detecting schizophrenia. Linear and nonlinear measurements were calculated from resting EEG data, which demonstrated the capability to distinguish schizophrenia patients. The extracted features resulted in high classification accuracy across various machine learning algorithms.

Sun et al. [11] in their study, EEG time series were converted to RGB images after passing the pre-processing stage, time and frequency domain features and classification was performed with these images using hybrid deep neural networks (DNN). The research results revealed that the fuzzy entropy (FuzzyEn) feature was superior to fast Fourier transform (FFT) in schizophrenia classification.

Ahmedt-Aristizabal et al. [12] used EEG data and artificial intelligence methods in their study to detect children prone to schizophrenia at an early stage. They demonstrated that recurrent neural networks outperform ML techniques for sequence modeling. The results were validated using cross-validation.

In this study, resting-state EEG data recorded from healthy control individuals and schizophrenia patients were used to predict schizophrenia. To achieve this, a band-pass filter was applied, followed by three different feature extraction methods and cross-validation processes, which were tested using machine learning algorithms.

## **2. Materials and Methods**

The Schizophrenia disease diagnosis dataset utilized in this study was collected and analyzed in [9]. The healthy participants were selected from university students, university hospital staff, and an adult education school. The subjects in the patient group were selected from a university hospital. 31 subjects were used in this dataset. The group consisted of 20 healthy control subjects (13 males, 7 females) and 11 individuals with schizophrenia (9 males, 2 females). Demographic information is given in Table 1. EEG signals were recorded with 500 Hz frequency signals for 3 minutes while the subjects were at rest. The dataset consists of 31 channels. Eye blinks and other artifacts were already removed using infomax Independent Component Analysis (ICA) in [9].



Table 1: Demographic information of the subjects

Subjects	Male	Female	Age
Healthy Controls	13	7	40.7 ± 11.9
Schizophrenia Patients	9	2	36.2 ± 10.2

In this study, the dataset was tested in Matlab with 6 machine learning algorithms using 3 different feature extraction methods. First, EEG signals were filtered with a band pass filter between 0.5-45 Hz. Among the 3 different feature extraction methods, Streamline feature extraction in the signal processing toolbox in Matlab was first performed. This method consists of total 18 features including time and frequency domain features. The Automatic EEG Signal Preprocessing Toolbox in Matlab was used as the second feature extraction method. Here, Delta Mean Band Power, Theta Mean Band Power, Alpha Mean Band Power, Beta Mean Band Power, Gamma Mean Band Power and Theta-Beta Ratio were obtained from linear features. The EEG Feature Extraction Toolbox (EFETB) in Matlab was used as the third feature extraction method. Here, 12 features were obtained, including Band Power based, Hjorth based, statistical features. Feature extraction operations were performed using 3 different methods.

### 3. Results and Discussion

After feature extraction, normalization process was applied to the data, and all the features were scaled between -1 and 1. 3-fold cross validation method was applied to validate the scores. In this study, a data augmentation process was applied to the schizophrenia vectors in the training group to balance their number with that of the healthy vectors. No augmentation process was performed on the test set. ML algorithms K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Decision Tree (DT), Naive Bayes (NB), Logistic Regression (LR), And Random Forest (RF) were applied. Average accuracy was calculated for each algorithm. The results obtained are shown in Figure 1.

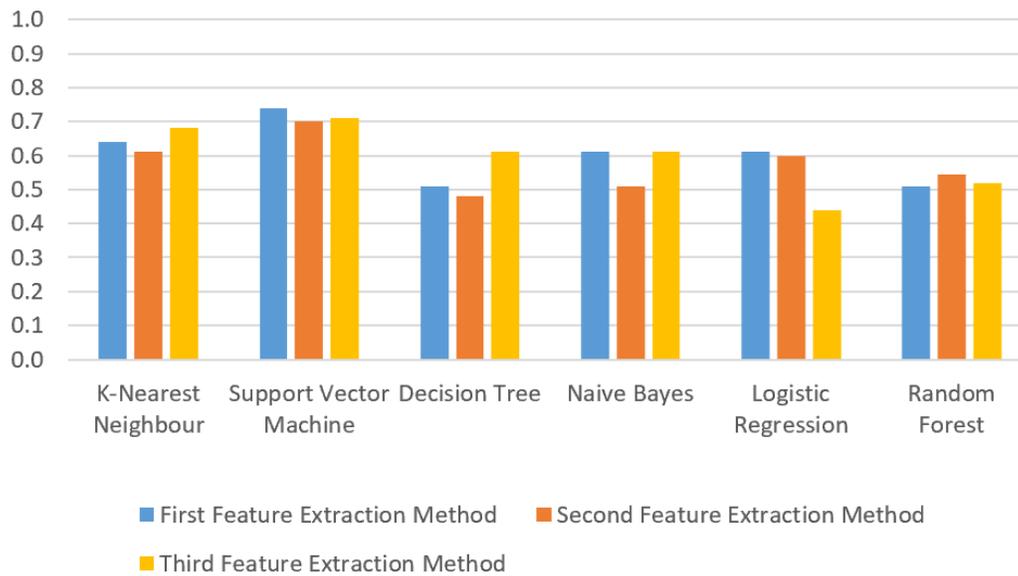


Figure 1: Average accuracies obtained with 6 different machine learning algorithms for 3 different feature methods

The highest average accuracy values across three different feature extraction methods were achieved using the SVM algorithm. These values were 0.75, 0.71, and 0.72 for the three feature extraction methods, respectively. The first feature extraction method performed best with 3 out of 6 ML models: SVM, NB, and LR. The second feature extraction method, in comparison, showed the best performance in the RF experiment. In the KNN and DT experiments, the third feature extraction method performed better.

## References

- [1] İ. Cantürk, "A computerized method to assess Parkinson's disease severity from gait variability based on gender," *Biomedical Signal Processing and Control*, vol. 66, p. 102497, 2021.
- [2] İ. Cantürk and O. Günay, "Investigation of Scalograms with a Deep Feature Fusion Approach for Detection of Parkinson's Disease," *Cognitive Computation*, pp. 1-12, 2024.
- [3] İ. Cantürk and L. Özyılmaz, "A Deep Feature Driven Expert System to Estimate the Postmortem Interval From Corneal Opacity Development," *Expert Systems*, p. e13757, 2024.
- [4] M. M. Ahsan, S. A. Luna, and Z. Siddique, "Machine-learning-based disease diagnosis: A comprehensive review," in *Healthcare*, 2022, vol. 10, no. 3: MDPI, p. 541.
- [5] İ. Cantürk, "Parkinson hastalığının derecesi ile yürüyüş değişkenliği arasındaki ilişkinin bulanık tekrarlılık grafiğine göre araştırılması," *Avrupa Bilim ve Teknoloji Dergisi*, no. 19, pp. 410-419, 2020.



- [6] İ. Cantürk, "A Feature Driven Intelligent System for Neurodegenerative Disorder Detection: An Application on Speech Dataset for Diagnosis of Parkinson's Disease," *International Journal on Artificial Intelligence Tools*, vol. 30, no. 03, p. 2150011, 2021.
- [7] J. Li *et al.*, "A Lightweight Multi-Mental Disorders Detection Method Using Entropy-Based Matrix from Single-Channel EEG Signals," *Brain Sciences*, vol. 14, no. 10, p. 987, 2024.
- [8] S. Bagherzadeh and A. Shalbaf, "EEG-based schizophrenia detection using fusion of effective connectivity maps and convolutional neural networks with transfer learning," *Cognitive Neurodynamics*, pp. 1-12, 2024.
- [9] J. R. De Miras, A. J. Ibáñez-Molina, M. F. Soriano, and S. Iglesias-Parro, "Schizophrenia classification using machine learning on resting state EEG signal," *Biomedical Signal Processing and Control*, vol. 79, p. 104233, 2023.
- [10] M. Nieuwenhuis, N. E. van Haren, H. E. H. Pol, W. Cahn, R. S. Kahn, and H. G. Schnack, "Classification of schizophrenia patients and healthy controls from structural MRI scans in two large independent samples," *Neuroimage*, vol. 61, no. 3, pp. 606-612, 2012.
- [11] J. Sun *et al.*, "A hybrid deep neural network for classification of schizophrenia using EEG Data," *Scientific Reports*, vol. 11, no. 1, p. 4706, 2021.
- [12] D. Ahmedt-Aristizabal *et al.*, "Identification of children at risk of schizophrenia via deep learning and EEG responses," *IEEE Journal of biomedical and health informatics*, vol. 25, no. 1, pp. 69-76, 2020.