Scientific Research Publishing

Advances in Artificial Intelligence and Robotics Research, 2026, 2(3), 162-189
https://www.scirp.org/journal/airr
ISSN Online: 3143-3995
ISSN Print: 3143-3987

Contribution of the Chebyshev Polynomial to Automatic Diagnosis of Alzheimer's Disease Based on Machine Learning

Stephanie-Claude Pelap Ngassa[1*], Edith Belise Kenmogne[1], Clementin Tayou Djamegni[1,2]

[1]Department of Mathematics and Computer Science, FS, University of Dschang, Dschang, Cameroon
[2]Department of Computer Engineering, UIT-FV, University of Dschang, Dschang, Cameroon
Email: *scpelap@yahoo.fr, ebkenmogne@gmail.com, dtayou@gmail.com

How to cite this paper: Ngassa, S.-C.P., Kenmogne, E.B. and Djamegni, C.T. (2026) Contribution of the Chebyshev Polynomial to Automatic Diagnosis of Alzheimer's Disease Based on Machine Learning. *Advances in Artificial Intelligence and Robotics Research*, **2**, 162-189.
https://doi.org/10.4236/airr.2026.23010

Received: July 23, 2026
Accepted: September 13, 2026
Published: September 16, 2026

Abstract

Alzheimer's disease (AD) is a neurodegenerative disorder affecting brain lobe cells. Several techniques have been developed to diagnose AD. This study proposes a new technique for the automatic diagnosis of AD based on the analysis of EEG signals. This involves preprocessing, spectral analysis based on the Discrete Chebyshev Transform (DChT), extraction of 85 features, dimensionality reduction, and classification. Classification is performed using k-nearest neighbors, support vector machines, XGBoosting, AdaBoosting, GradientBoosting, Random Forest and Multi-Layer Perceptron. In addition, the classification accuracies obtained without the proposed model were examined. Furthermore, Boruta's feature selection algorithm and the Mann-Whitney test confirm the benefits of using the proposed model, especially DChT as a spectral analysis tool. The impact of training set size on classification accuracy was analyzed. For classification using dimensionality reduction, the best accuracy is 88.64% and was achieved by GradientBoosting. Without dimensionality reduction, the best accuracy is 93.24% and was achieved by XGBoost. These performances are as high as those of other methods proposed in the literature. Therefore, DChT is an effective tool for improving AD diagnosis. However, using DChT is beneficial for classification and the results are the best; this work found a new technique for diagnosing AD using artificial intelligence.

Keywords

Discrete Chebyshev Transform (DChT), EEG Signal, Dimension Reduction, Classification, Feature Extraction

1. Introduction

Alzheimer's disease (AD), first described in 1906 by Alois Alzheimer, is the leading cause of dementia worldwide. It is a progressive and irreversible neurodegenerative disease that affects cognitive functions and memory [1]. AD is characterized by the abnormal accumulation of Tubulin-Associated Unit (TAU) proteins in neurons, and of beta-amyloid proteins, which form amyloid plaques, causing neuronal degeneration as a result of these lesions: firstly, the abnormal accumulation on the outside of nerve cells of a protein called β-amyloid peptide (or A-beta peptide or A β peptide), leading to the formation of amyloid plaques, also known as "senile plaques", and secondly, the abnormal accumulation of the TAU protein in neurons, leading to their degeneration. Several tools are available for the diagnosis of AD [2]. However, these methods have a number of limitations, as they require the presence of a neurologist and long waiting times for results. Several automatic diagnostic methods have been developed in recent years to overcome these problems. The work presented in this paper follows this trend. The aim is to focus on machine learning such as Support Vector Machine (SVM), k-nearest neighbors (kNN), Ensemble Learning (EL) and deep learning such as Multi-Layer Perceptron (MLP) in EEG signals, specifically spectral analysis, feature extraction, dimension reduction, classification, metrics evaluation and performance evaluation.

The remainder of this paper is organized as follows. Section 2 presents the materials and methods used. First, it presents the dataset used for the tests. Second, it presents the methodology used in detail, specifically preprocessing, expansion of the EEG signal in the orthogonal Chebyshev basis for spectral analysis, feature extraction, dimension reduction, classification and metric evaluation. Section 3 applies the methodology presented in Section 2 and presents the results obtained. Section 4 discusses the results obtained from the state-of-the-art. Finally, Section 5 concludes the study.

2. Materials and Methods

2.1. Data Acquisition

The database was designed by researchers at Florida State University and recorded from 19 scalp loci (Fp1, Fp2, Fz, F3, F4, F7, F8, Cz, C3, C4, T3, T4, Pz, P3, P4, T5, T6, O1 and O2) of the international 10 - 20 system using a Biologic Systems Brain Atlas III Plus workstation. The letters F, C, P, O and T represent the brain lobes (F: frontal, C: central, P: parietal, O: occipital and T: temporal). The recordings, which comprised four groups (A, B, C and D), were made in two resting states: eyes open (groups A and C) by visual fixation, and eyes closed (groups B and D) using a mandible-bound frontal reference ground. Groups A and B represent healthy controls and consisted of 24 healthy elderly people, with no neurological or psychiatric disorders. The age of healthy people ranged from 61 to 83 years, with an average age of 72 years and the age of AD patients ranged from 53 to 85 years, with an average age of 69 years. Groups C and D consisted of 24 patients

with AD diagnosed according to the criteria of the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS), AD and Related Disorders Association (ADRDA) and Diagnostic and Statistical Manual of Mental Disorders (DSM)-III-R, as described above. EEG segments had a recording time of 8 s, which has limited to a bandwidth of 1 - 30 Hz. EEG segments were recorded at a sampling rate of 128 Hz (without eye movements, blinks or myogenic artifacts) and extracted from EEG recordings. An EEG technician was present with each patient during the recordings to monitor the patients' state of alertness [3]. The database was created on January 21, 2020, and was provided by Dr. Dennis Duke of Florida State University. It is available at https://doi.org/10.17605/OSF.IO/S74QF.

2.2. Methodology

In this study, a model was proposed for the automatic diagnosis of AD based on EEG signals. The steps of the proposed learning model are illustrated in **Figure 1**. These steps are described in detail.

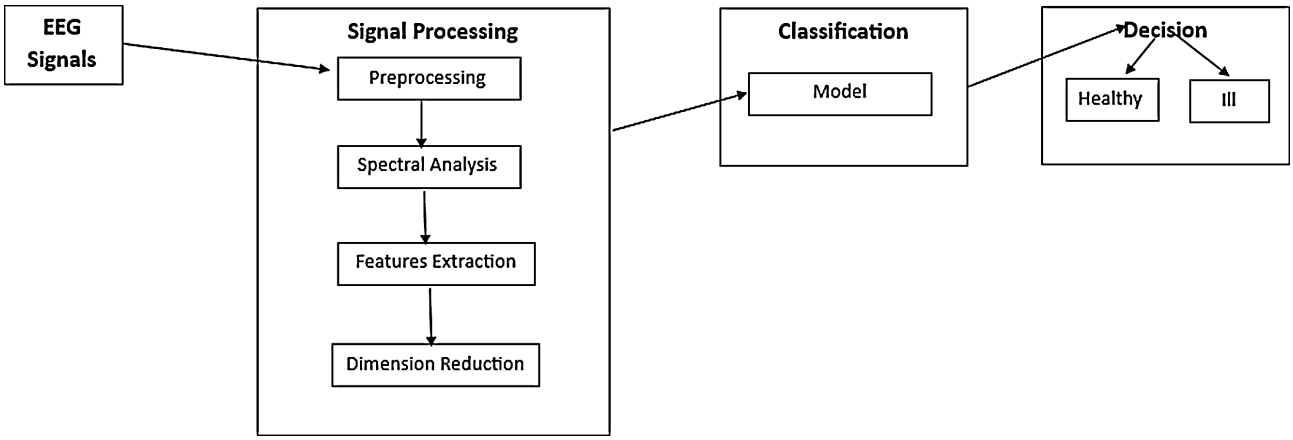

Figure 1. Architecture of the proposed learning model.

Generally, in preprocessing, the recorded data are cleaned and denoised, the artifacts are eliminated, the redundancies are eliminated, and the missing data and the bottlenecks are adjusted [4].

For spectral analysis, the signal is represented as a mathematical function of one or more variables. The 1st species of the Chebyshev polynomial, noted $T_k(x)$, is a special case of Jacobi polynomials with interesting properties that make them attractive for filter design and optimal polynomial interpolation. Chebyshev polynomials are orthogonal in the interval $[-1; 1]$ with respect to the weight function

$$\omega(x) = (1-x)^{\frac{-1}{2}} \tag{1}$$

The explicit relationship describing Chebyshev polynomials $T_k(x)$ is the following trigonometric form:

$$T_k(x) = \cos\left[k\cos^{-1}(x)\right], k \geq 0 \tag{2}$$

The expansion of a signal $s(x)$ in the "L" order in the orthogonal Chebyshev basis is given by:

$$s(x)=\sum_{k=0}^{L}\alpha_k T_k(x) \tag{3}$$

The decomposition coefficients α_k are then evaluated using the Gauss-Lobatto quadrature method with a quadrature of order $L+1$. The Gauss-Lobatto method states that for a quadrature of order L, and a family of orthogonal Chebyshev polynomials $T_k(x)$ in the interval $[-1; 1]$, with respect to the weight function $w(x)$, the following approximation was made by [5]:

$$\int_{-1}^{1} f(x)\omega(x)\mathrm{d}x \approx \sum_{j=0}^{L} G_j f(x_j) \tag{4}$$

where $f(x)\in L^2[a,b]$, x_j, are the L nodes of the Gauss quadrature formed by the L roots or zeros of $T_L(x)$, and $\{G_j\}_{j=1}^{L}$ called Christoffel numbers, is given by

$$G_j=\frac{\pi}{L+1},\ j=1,2,\cdots,L \tag{5}$$

Applying the Gauss-Lobatto quadrature method gives:

$$\int_{-1}^{1}\frac{s(x)T_k(x)}{\sqrt{1-x^2}}\mathrm{d}x=\frac{\pi}{M+1}\sum_{j=0}^{M}s(x_j)T_k(x_j) \tag{6}$$

The EEG signal is decomposed into a set of Chebyshev spectral coefficients using the DChT Equation (7) [6]:

$$\begin{cases}\alpha_0=\dfrac{1}{M+1}\displaystyle\sum_{j=1}^{M+1}s(x_j)=\dfrac{1}{M+1}\sum_{j=1}^{M+1}s\left\{\cos\left[\dfrac{2j-1}{2(M+1)}\pi\right]\right\}\\ \alpha_k=\dfrac{2}{M+1}\displaystyle\sum_{j=1}^{M+1}s(x_j)\cos\left[\dfrac{k(2j-1)}{2(M+1)}\pi\right]\\ \quad=\dfrac{2}{M+1}\displaystyle\sum_{j=1}^{M+1}s\left\{\cos\left[\dfrac{2j-1}{2(M+1)}\pi\right]\right\}*\left\{\cos\left[\dfrac{k(2j-1)}{2(M+1)}\pi\right]\right\}\\ \quad k=1,2,3,\cdots,M\end{cases} \tag{7}$$

The main aim of feature extraction is to obtain additional information from the raw signal by transforming voluminous data into a smaller feature vector. After analyzing the data using DChT, the feature extraction methodology analyzes the signals to extract the most important features that are representative of the different signal classes [4]. To identify and quantify the most relevant properties of a signal and facilitate its analysis or classification, 85 parameters were extracted. 31 parameters were extracted using the spectral decomposition coefficients, 22 parameters were extracted using the absolute values of these decomposition coefficients, 30 parameters were extracted using the original signal and 2 parameters were extracted using the reconstructed signal. These two parameters are the Percent Root Square Difference (PRD) and Energy of the approximated signal. 85 parameters were extracted by applying the formulas described below. However,

scaling, hyperparameter tuning, and Boruta selection were performed only on each training fold prior to classification. Some of these informative, descriptive, statistical, energetic and entropic parameters are presented below:

- Maximum
- Minimum
- Mean absolute deviation:

$$\text{Mad} = \text{Me}\left(\left|X_i - \mu\right|\right) \tag{8}$$

- PRD (Percent Root Square Difference): for a signal Δ_n and the approximate signal δ_n $n = 1, 2, \cdots$

$$\text{PRD}(\%) = 100\sqrt{\sum_n \left(\Delta_n - \delta_n\right)^2} \tag{9}$$

- Coefficient of asymmetry or skewness coefficient ($\alpha 1$): μ and σ respectively represent the mean and standard deviation of the set X.

$$\alpha 1 = \frac{\sum_i^N \left(X_i - \mu\right)^3}{\sigma^3} \tag{10}$$

- Kurtosis ($\beta 1$): μ and σ respectively represent the mean and standard deviation of the set X

$$\beta 1 = \frac{\sum_i^N \left(X_i - \mu\right)^4}{\sigma^2} \tag{11}$$

- Number of Zero Crossings (NZC)
- Magnitude Peak (PM1):

$$\text{PM1} = \frac{\text{Max}\left(\text{Abs}\left(X\right)\right)}{\sqrt{\frac{\sum_{i=1}^N \left(X_i - X_i'\right)^2}{N}}} \tag{12}$$

- Latitude factor (LF):

$$\text{LF} = \frac{N * \text{Max}\left(\text{Abs}\left(X\right)\right)}{\frac{\sqrt{\sum_{i=1}^N X_i}}{N}} \tag{13}$$

- Shape Factor (SF):

$$\text{SF} = \frac{\sqrt{\frac{\sum_{i=1}^N X_i^2}{N}}}{\left|\frac{\sum_{i=1}^N X_i}{N}\right|} \tag{14}$$

- Coefficient of Variation (CV):

$$\text{CV} = \frac{\sqrt{\frac{\sum_{i=1}^N \left(X_i - X_i'\right)^2}{N}}}{\frac{1}{N}\sum_{i=0}^N X_i} \times N \tag{15}$$

Dimension reduction techniques such as Principal Component Analysis (PCA) can be applied to reduce the dimensions of the original data and remove irrelevant and redundant information. This reduces the computational costs and reduces noise depending on how many principal components are kept [4].

Machine Learning (ML) and Deep Learning (DL) algorithms were used to perform classification. These include Support Vector Machine (SVM), k-nearest neighbors (kNN), Ensemble Learning (EL) algorithms such as XGBoost, GradientBoosting and AdaBoosting, Random Forest (RF) and Multi Layer Perceptron (MLP).

Some metrics are evaluated after classification:

- Confusion matrix

It depicts the structure of the confusion matrix for a binary classifier. The estimates of the possibilities of classification models are derived from the expressions True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN), which exist in the confusion matrix [7].

- Basic indicators

The performance evaluation criteria in our study consist of measuring the performance of a classification according to the standard criteria described below [7] [8]:

1. Accuracy (acc):

$$\mathrm{ACC} = \frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{FP}+\mathrm{TN}+\mathrm{FN}} \times 100 \tag{16}$$

2. Sensitivity (Sen) or True Positive Rate (TPR) or Recall:

$$\mathrm{Sen} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} \times 100 \tag{17}$$

3. Specificity or True Negative Rate (TNR ou spe):

$$\mathrm{Spe} = \frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FP}} \times 100 \tag{18}$$

4. False Positive Rate (FPR):

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{TN}+\mathrm{FP}} \times 100 = 1 - \mathrm{Spe} \tag{19}$$

5. Precision (Pre):

$$\mathrm{Pre} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}} \times 100 \tag{20}$$

6. F-Measure or F-score:

$$\text{F1-score} = 2 \times \frac{\mathrm{Pre} \times \mathrm{Recall}}{\mathrm{Pre}+\mathrm{Recall}} \tag{21}$$

- Receiver Operating Characteristic (ROC) area and Area Under the Curve (AUC):

The ROC curve visualizes the trade-off between TPR and FPR. For each threshold, TPR and FPR were calculated and plotted on one graph. The higher the TPR

and the lower the FPR for each threshold, the better. The area below the ROC curve is called the ROC-AUC score, a number that determines how good the ROC curve is [7].

- K-fold cross-Validation (KCV):

The KCV technique is one of the most commonly used approaches by practitioners for model selection and error estimation of classifiers. It consists of splitting a dataset into k subsets; then, iteratively, some of them are used to learn the model, whereas the others are exploited to assess its performance [9].

3. New Results

This section applies the proposed learning model and presents the results obtained at each step of the learning model.

3.1. Signal Plot

In this database, 19 columns represent each electrode and 1024 rows represent the total duration of the recordings, that is 8 s. This means that one line recorded the amplitude of the electrode for a period of 7.8125 ms. Examples of the data from channel F7 are shown in **Figure 2**. This figure shows a signal trace for each F7 channel from each group A to D. Group A refers to healthy subjects with open eyes, group B to healthy subjects with closed eyes, group C to AD subjects with open eyes and group D to AD subjects with closed eyes.

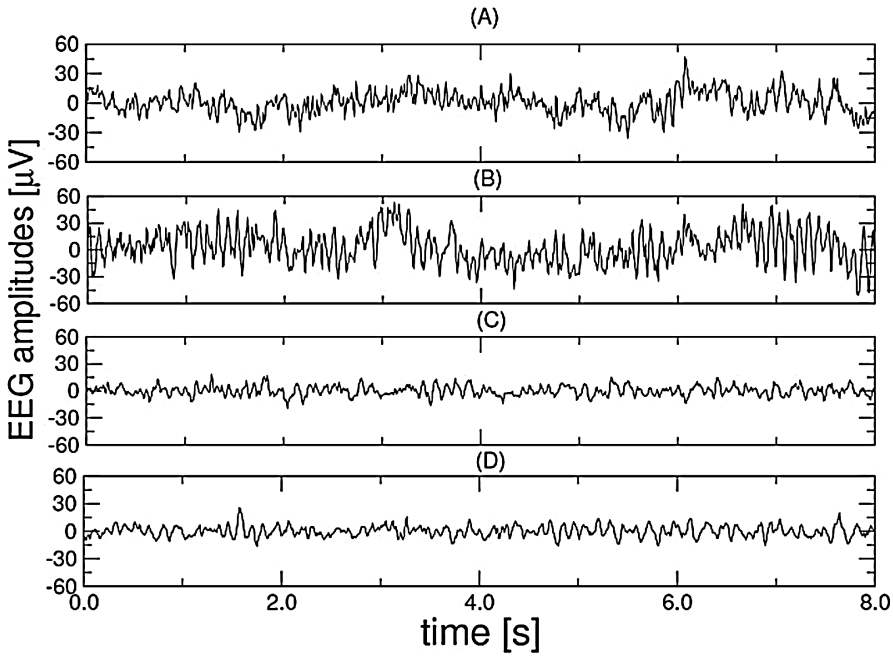

Figure 2. EEG signal presentation at electrode F7 for each group (A, B, C and D) [10].

Figure 3 shows the signal obtained at electrode F7 for each group after preprocessing. In this figure, group A is labeled HC_oe, group B is labeled HC_ce, group C is labeled AD_oe, and group D is labeled AD_ce.

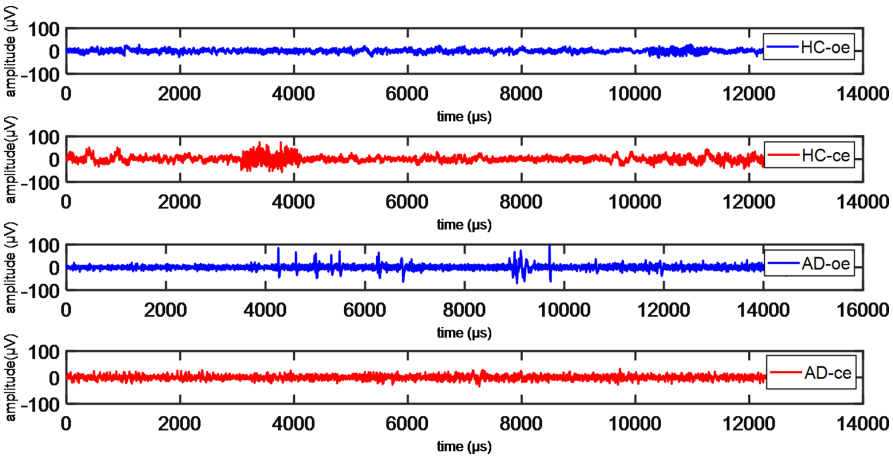

Figure 3. EEG signal at electrode F7 after preprocessing.

The dataset consisted of four data groups, each containing the recorded data from 12 patients. Some recordings lasted longer than 8 seconds and, therefore, contained more than 1024 rows for a single patient. Thus, for each data group, there were between 12,288 and 15,360 rows and 19 columns, each representing an electrode. Preprocessing verifies whether patient data is valid (no bottlenecks, missing values, outliers,..) and divides the data by sampling it into rows or data units. The records were fully preserved during preprocessing to prevent subject-level information leaks. The information is categorized based on ocular status, and each patient's data is added progressively and processed independently during testing.

3.2. Spectral Analysis

Spectral analysis involves approximating or interpolating the signals to extract better information. It breaks down a signal into smaller signals to extract characteristics; This is performed using a transfer function. In our case, it is the Jacobi polynomial, in particular the Chebyshev polynomials. **Figure 4** shows the interpolation coefficients on the orthogonal Chebyshev basis for each of the HC_oe, HC_ce, AD_oe and AD_ce groups. The graphs in **Figure 4** represent the first and second spectral coefficients, respectively.

Using Chebyshev interpolation, the number of columns increases from 19 to 39 as given by Equation (7) to evaluate the spectral coefficients α_k.

However, the spectral coefficients gradually decrease until they cancel each other out. This is explained by the Runge phenomenon, which appears to be a consequence of interpolation using the Jacobi polynomial.

3.3. Feature Extraction

After spectral analysis, the dataset has a dimension of 12,288 × 39; and the coefficients obtained after interpolation of the signal in the orthogonal Chebyshev basis will make it possible to extract 85 parameters represented by 102,425 figures in dimensions 1, 2 and 3, as shown respectively in **Figures 5-7** for a dataset of size 12,288 × 85.

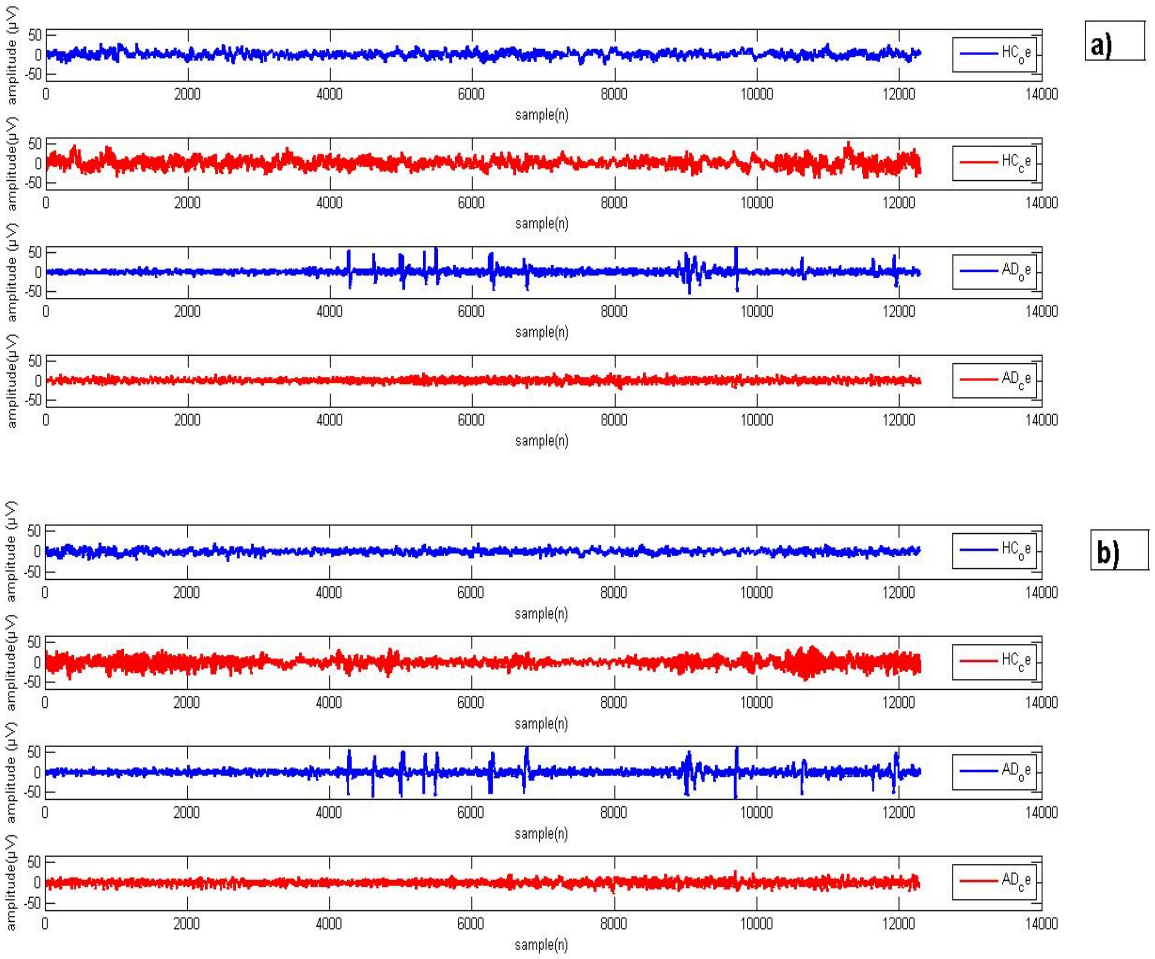

Figure 4. Spectral coefficients after interpolation by the Chebyshev polynomial.

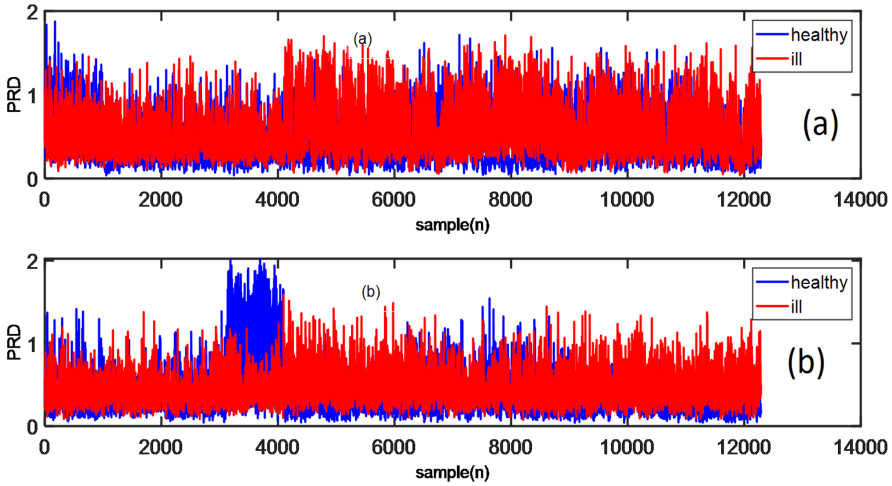

Figure 5. 1D parameter features: PRD.

Figure 5 presents the 1D parameter features. It contains two graphs. Graph (a) represents the "eyes opened" parameter and graph (b) the "eyes closed" parame-

ter. Figure 5 presents the characteristic PRD in relation to the sample in dimension 1. As defined by formula (9), the PRD represents the difference between the original signal and the reconstructed signal for each of the parameters "eyes opened" and "eyes closed" for sick and healthy individuals. This difference is minimal. This means that signal interpolation on the orthogonal Chebyshev basis was carried out with minimal error.

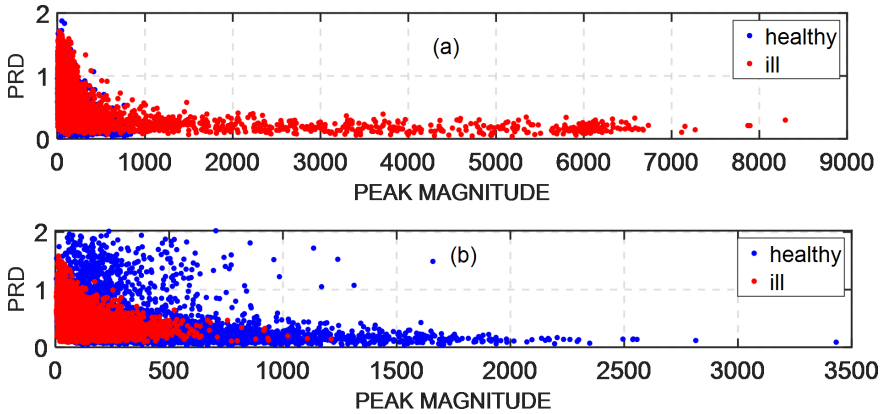

Figure 6. 2D parameter features: PM1 and PRD.

Figure 6 presents 2D parameter features. It contains two graphs. Graph (a) represents the "eyes opened" parameter and graph (b) the "eyes closed" parameter. For each of the "eyes opened" and "eyes closed" parameters, Figure 6 shows the relationship between the PRD and PM1 characteristics for healthy and sick individuals. The PM1 parameter is defined using formula (12). Note that the higher the signal intensity, the lower the signal reconstruction error.

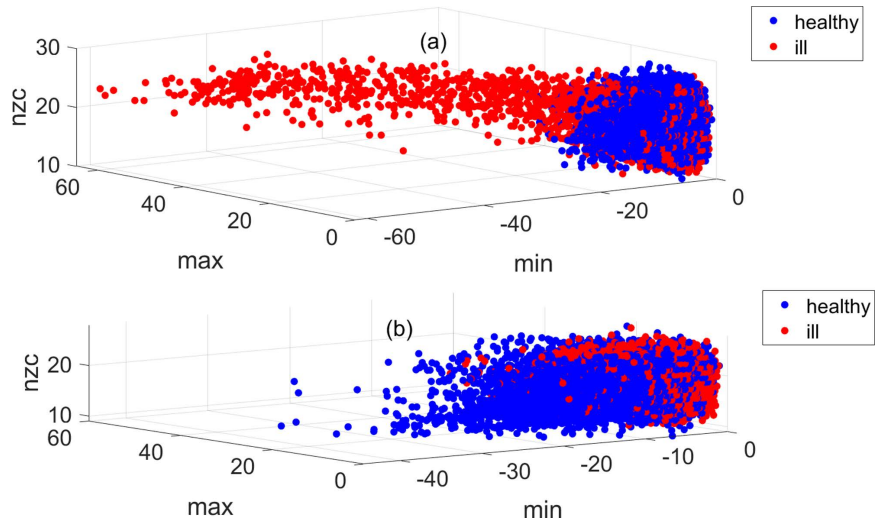

Figure 7. 3D parameter features: Min, Max and NZC.

Figure 7 presents 3D parameter features. The Min, Max and NZC characteristics are observed for the parameters "eyes opened" in graph (a) and "eyes closed" in graph (b) for both sick and healthy individuals. In this figure, the values of the

metrics are much more concentrated towards the center, which means that the minimum and maximum values fall within a relatively narrow range and the distribution is random.

3.4. Dimension Reduction

After application of the PCA, the result of the dimension reduction is a matrix of size 85 × 85, and the datasets are represented in dimensions 1, 2 and 3 as shown in Figures 8-10.

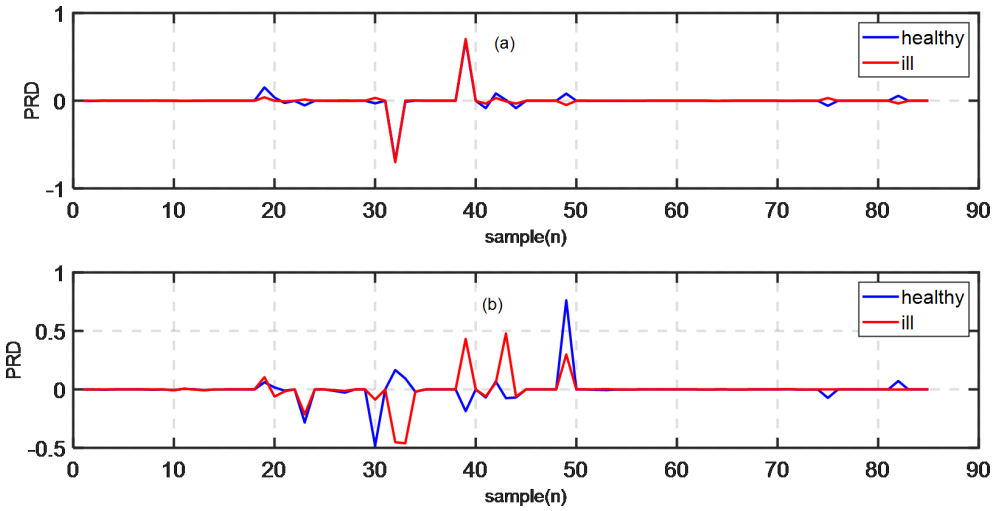

Figure 8. PRD parameter as a function of 1D sample size after dimension reduction.

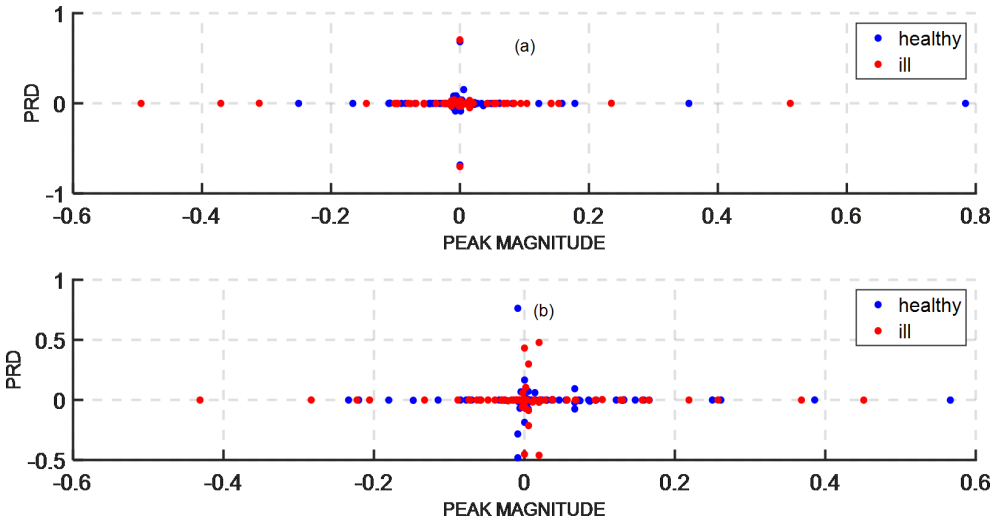

Figure 9. 2D parameter after dimension reduction.

Figure 8 shows two PRD plots, in graphs (a) and (b), which represent the signal reconstruction error in the Chebyshev orthogonal basis for the "eyes opened" and "eyes closed" parameters respectively. After dimension reduction using PCA, the error becomes increasingly smaller.

Figure 9 shows the peak signal magnitude as a function of the reconstruction

error for parameters "eyes opened" and "eyes closed" in graphs (a) and (b) respectively. The metrics were increasingly concentrated towards the center and of low intensity for both healthy and sick individuals.

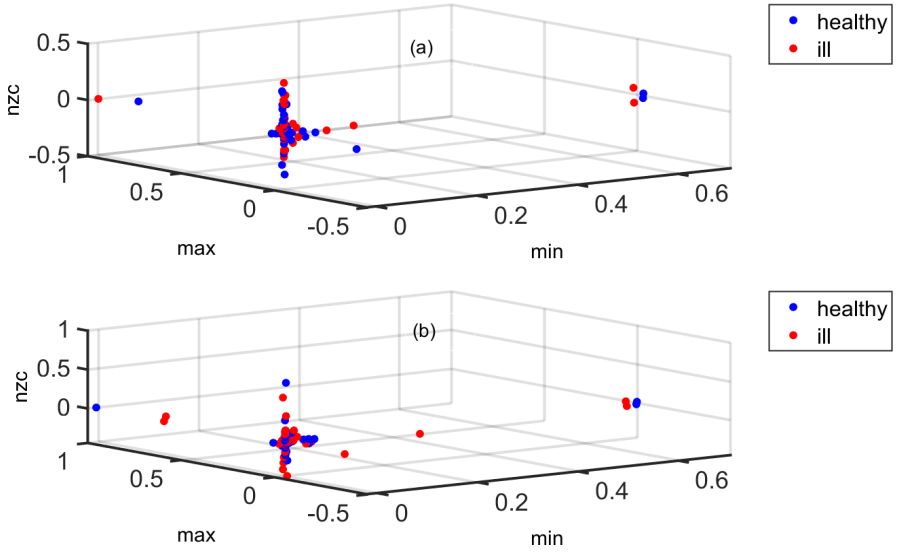

Figure 10. Min, Max and NZC parameters after dimension reduction.

Figure 10 shows the minimum, maximum and number of positive values in the signal for parameters "eyes opened" and "eyes closed" in graphs (a) and (b) respectively. After dimension reduction with PCA, the metrics were increasingly concentrated towards the center and of lower intensity for both healthy and sick individuals.

3.5. Classification

In this subsection, three categories of results are presented. The first category of results was obtained using dimension reduction, the second category was obtained without dimension reduction and the third category of results was obtained without spectral analysis and without dimension reduction.

3.5.1. Classification Results Using Dimension Reduction

After parameter extraction and dimension reduction by signal interpolation in the orthogonal Chebyshev basis, the very low-dimensional parameter vectors were used directly as classifier inputs.

Using the SVM algorithm, on a dataset of which 80% is used for training and 20% for testing, the ROC curve obtained in Figure 11 shows that the classifier is not very discriminative, with an AUC = 0.426.

The associated confusion matrix is given by

$$\begin{pmatrix} 17 & 17 \\ 22 & 12 \end{pmatrix}$$

and leads to the following metric values: Acc = 42.65%, Sen = 43.59%, Spe = 41.38%, Pre = 50% and F-Score = 46.58%.

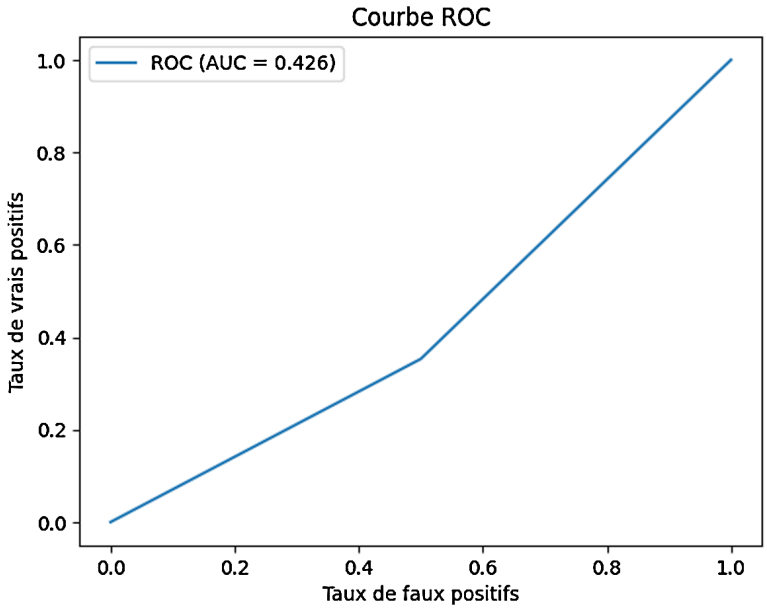

Figure 11. ROC curve obtained for SVM classification with dimension reduction.

Using the kNN algorithm on a dataset in which 80% was used for training and 20% for testing, the best result was obtained with k = 1. The corresponding ROC curve is shown in **Figure 12**.

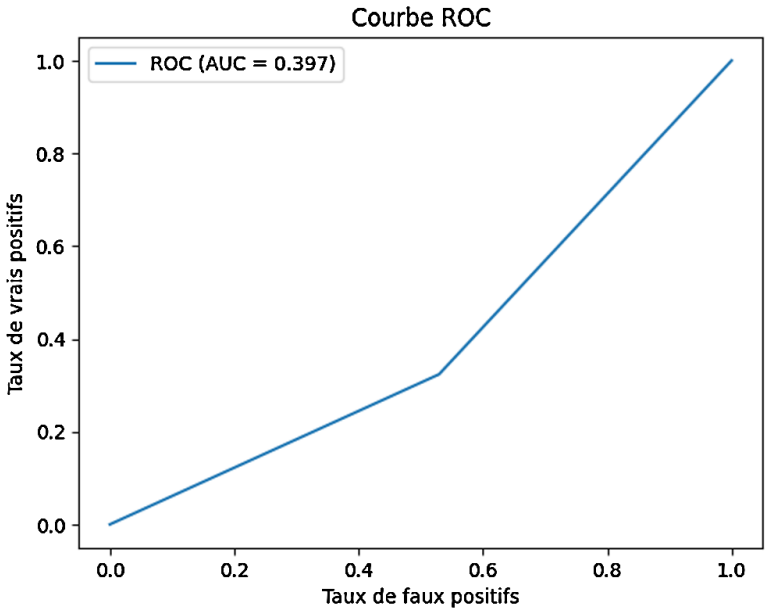

Figure 12. ROC curve obtained for kNN classification with dimension reduction.

The associated confusion matrix is given by

$$\begin{pmatrix} 16 & 18 \\ 23 & 11 \end{pmatrix}$$

This leads to the following metrics deduced: Acc = 39.71%, Sen = 41.03%, Spe

= 37.93%, Pre = 47.06%, F-Score = 43.84%. In this case, these metric values show that the kNN classifier performed less well than the SVM classifier.

Using the XGBoost algorithm on a dataset in which 80% was used for training and 20% for testing, the following metric values were obtained: Acc = 84.38%; Pre = 76%; Sen = 100%; F-score = 86% and AUC = 0.844. This shows that the XGBoost classifier used here is highly discriminative and delivers a precise result.

Using the AdaBoost algorithm, on a dataset in which 70% was used for training and 30% for testing, the following results were obtained: Acc = 80.77%; Pre = 82%; Sen = 81%; F-score = 81%. The corresponding ROC curve is shown in Figure 13. The AdaBoost classifier performed well on this dataset. This makes it a good tool for distinguishing sick from healthy subjects.

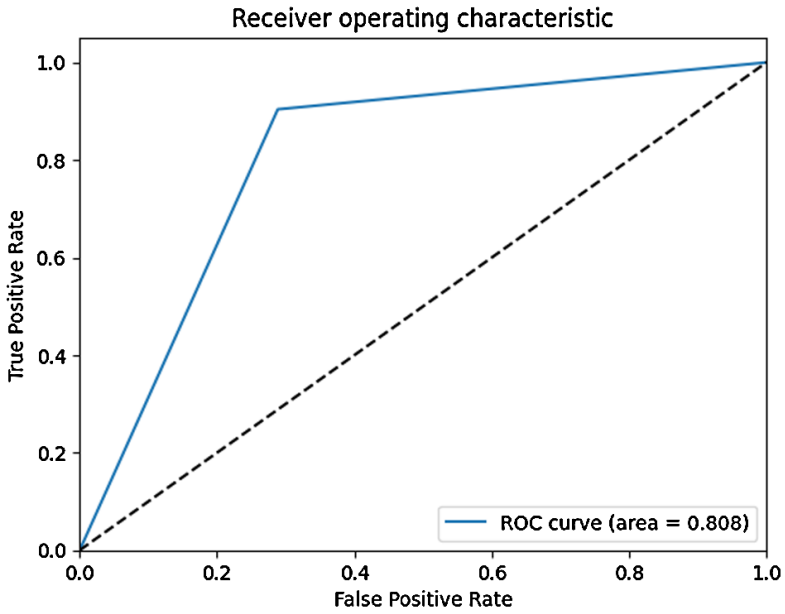

Figure 13. ROC curve obtained for Adaboost classification with dimension reduction.

Using the GradientBoosting algorithm, on a dataset in which 80% was used for training and 20% for testing, the ROC curve in Figure 14 was obtained.

The associated confusion matrix is given by

$$\begin{pmatrix} 29 & 5 \\ 6 & 28 \end{pmatrix}$$

This led to the following metric values: Acc = 83.82%; F-score = 83.58%; Sen = 82.35% and Pre = 84.84%. According to the results of the experiments, the GradientBoosting classifier is one of the best tool for diagnosing patients.

Using the GradientBoosting algorithm by varying the percentages of the dataset used in the training and testing phases to achieve learning accuracy based on the percentage of the dataset used in the training phase, Figure 15 was obtained. In this figure, the x-axis represents the training set percentage and the y-axis represents the accuracy after classification. The experiment began with a value of 5% for the training set and 95% for the testing set. At this value of x, the accuracy was

58.33%. The abscissas evolved in steps of 10%. Accuracies evolved from 56.82% to 88.64% and were calculated for different values of the abscissas from 5% to 95%. The best accuracy is obtained with a training set percentage of 88% and a testing set percentage of 12%. However, the accuracy curve increases with the training set.

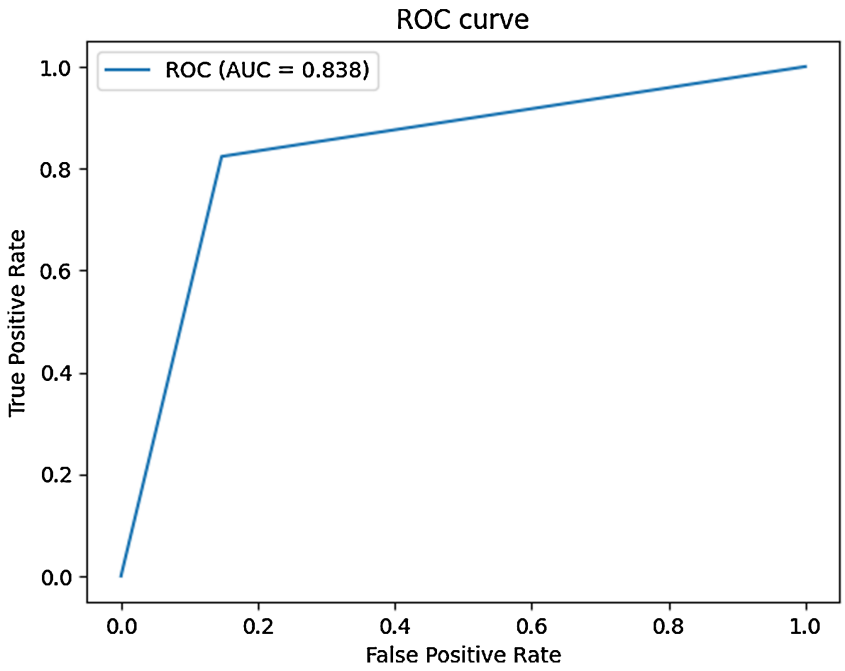

Figure 14. ROC curve obtained for GradientBoosting classification with dimension reduction.

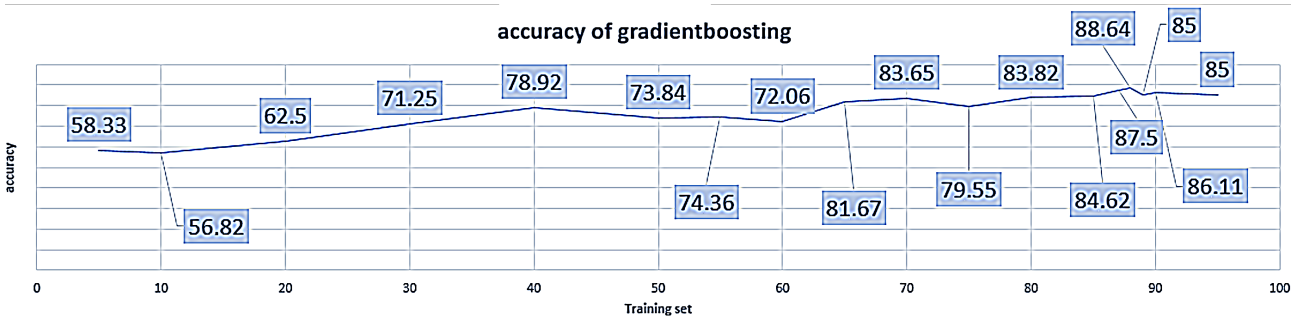

Figure 15. Evolution of accuracy in GradientBoosting classification with dimension reduction.

Using the Random Forest algorithm on a dataset in which 80% was used for training and 20% for testing, results in an accuracy of 73.53%, Pre = 72.22%, Recall = 76.71%, F1-score = 74.29% and AUC= 0.735. The associated confusion matrix is given by

$$\begin{pmatrix} 24 & 10 \\ 8 & 26 \end{pmatrix}$$

The corresponding ROC curve is shown in **Figure 16**. RF classifier is one of the best tool for diagnosing patients.

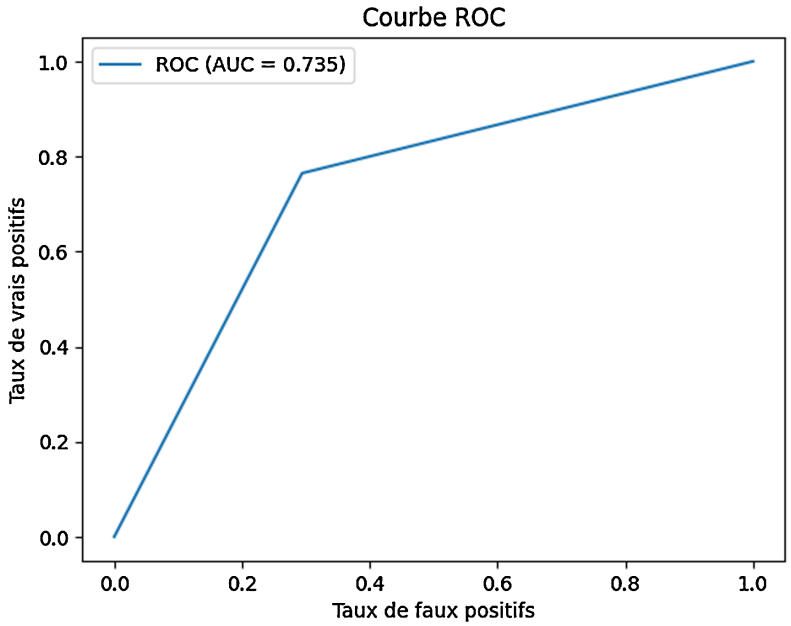

Figure 16. ROC curve obtained for Random Forest classification with dimension reduction.

The data used for classification using artificial neural networks such as the MLP were divided into three blocks; 70% of the data were used for training, 15% for validation and the remaining 15% for data prediction. The data at the input of the classifier correspond to the various features extracted in the previous phase. RandomizedSearchCV is a powerful technique for optimizing the hyperparameters of a machine learning model.

Table 1. Results for MLP classification with RandomisedSearchCV.

Hyperparameters	Architecture	Metrics
Optimiser: Adam; Activation function: Logistic; Learning rate: 0.01; Cross-validation: 5.	85-43-43-1	Accuracy = 53.79%; Prediction = 50%; Auc = 0.5.
Optimiser: SGD; Activation function: ReLU; Learning rate: 0.001; Cross-validation: 5.	85-43- ~ -1	Accuracy = 46.68%; Prediction = 46.15%; Auc = 0.558.

The results obtained using the RandomizedSearchCV algorithm to tune the hyperparameters are listed in Table 1. The accuracy was observed by varying the hyperparameters used in the MLP architecture. This makes it possible to search for hyperparameters and architectures that maximize the precision.

Table 2 lists the results obtained without using the RandomizedSearchCV algorithm to tune the hyperparameters. In order to avoid overlearning, the SGD

Table 2. Results for MLP Classification without RandomisedSearchCV.

Hyperparameters	Architecture	Metrics
Optimiser: SGD; Loss: binarycrossentropy; Learning rate: 0.001; Epoch: 25; Batch size: 32.	85-43-43-1	Accuracy_train = 46.72%; Auc = 0.491; Accuracy_test = 48.07%.
	85-85-1	Accuracy_train = 44.79%; Auc = 0.436; Accuracy_test = 36.59%.
	85-85-43-1	Accuracy_train = 51.44%; Accuracy_test = 50%; Auc = 0.491.
	85-43-85-1	Accuracy_train = 49.16%; Accuracy_test = 50%; Auc = 0.472.
	85-43-21-1	Accuracy_train = 48.98%; Accuracy_test = 38.46%; Auc = 0.342.
	85-43-100-1	Accuracy_train = 54.31%; Accuracy_test = 46.15%; Auc = 0.487.
	85-43-1-1	Accuracy_train = 52.11%; Accuracy_test = 55.76%; Auc = 0.581.
	85-43-500-1	Accuracy_train = 49.74%; Accuracy_test = 50%; Auc = 0.392.
	85-43-50-1	Accuracy_train = 48.76%; Accuracy_test = 42.31%; Auc = 0.433.
	85-43-1	Accuracy_train = 43.65%; Accuracy_test = 55.77%; Auc = 0.584.
	85-43-21-1	Accuracy_train = 44.02%; Accuracy_test = 40.38%; Auc = 0.414.
	85-85-85-1	Accuracy_train = 52.55%;

Continued

		Accuracy_test = 51.19%; Auc = 0.496.
Optimiser: SGD; Loss: binarycrossentropy; Activation function: ReLU; Learning rate: 0.001; Epoch: 30; Batch size: 32; DropOut: 0,3.	85-43-1	Accuracy_train = 48.07%; Accuracy_test = 47.62%; Auc = 0.402.
Optimiser: SGD; Loss: MSE; Activation function: ReLU; Learning rate: 0.001; Epoch: 30; Batch size: 32.	85-85-43-1	Accuracy_train = 52.78%; Accuracy_test = 48.07%; Auc = 0.438.
Optimiser: SGD; Loss: Binarycrossentropy; Learning rate: 0.001; Batch size: 32; DropOut: 0.3; Epoch: 100; Validation split = 20%.	85-1000-1	Accuracy_test = 59.38%; Auc = 0.488.

optimizer was used. The values of the hyperparameters were taken at random and the regularization techniques used for classification were designed to improve the test accuracy by enhancing learning. After this series of experiments with the MLP, the best accuracy obtained was 59.38% with the architecture 85-1000-1 with a dataset divided into three parts: 80% for the training set (20% of the training set was used for the validation), 10% for the testing set and the remaining part for the prediction. The ROC curve obtained after this classification is shown in **Figure 17** with AUC= 0.488. This value shows that the classifier is discriminative, and with this average accuracy, it can briefly distinguish sick from healthy individuals.

3.5.2. Classification Results without Dimension Reduction

Using the kNN algorithm on a dataset of which 80% is used for training and 20% for testing, the best result was obtained when k=20. The corresponding ROC curve is shown in **Figure 18**.

The associated confusion matrix is given by

$$\begin{pmatrix} 3659 & 1257 \\ 1516 & 3400 \end{pmatrix}$$

This leads to deducing the values of the following metrics: Acc = 71.80%, Recall = 69.16%, Pre = 73.01% and F-Score = 71.03%. It can be observed that the kNN classifier in this case performs better without dimension reduction.

Using the XGBoost algorithm on a dataset of which 70% is used for training and 30% for testing, the following metric values were obtained: Acc = 92.72%, Pre

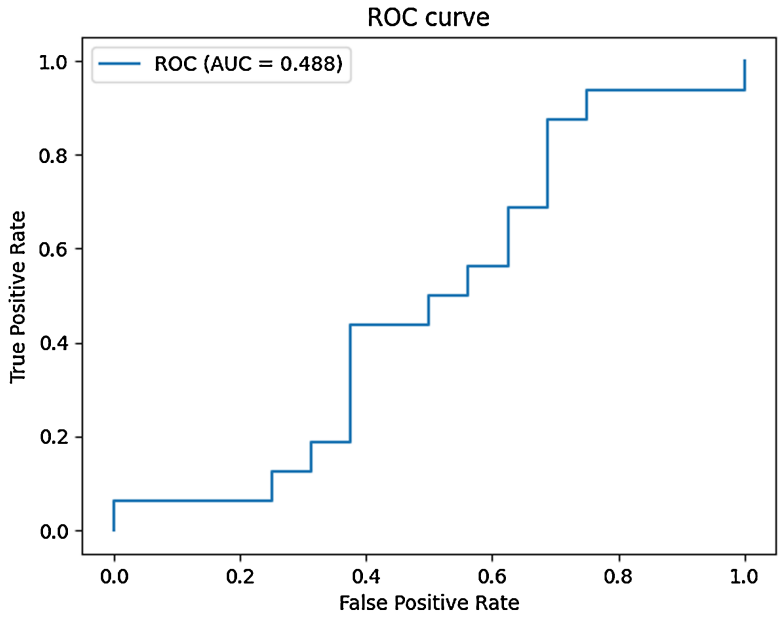

Figure 17. ROC curve obtained for MLP classification with dimension reduction.

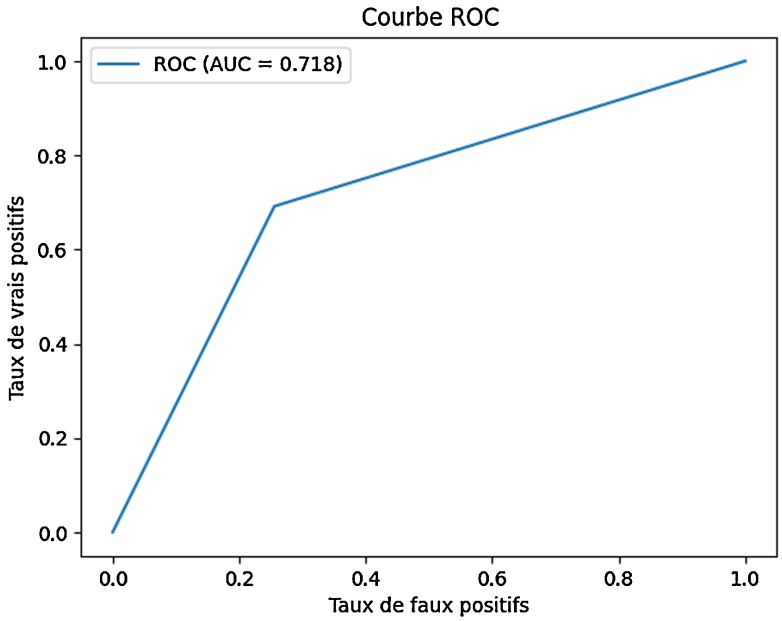

Figure 18. ROC curve obtained for kNN classification without dimension reduction.

= 95%, Recall = 90%, F1-score = 93% and Auc = 0.927. The XGBoost classifier is more discriminating and can easily distinguish between sick and healthy patients with precision. A closer look at the accuracies obtained after classification is shown in **Figure 19**. The best accuracy was 93.24%, obtained with a training set of 90% and a testing set of 10%.

Using the AdaBoost algorithm, on a dataset of which 70% was used for training and 30% for testing, the following results were obtained: Acc = 78.03%; Pre = 78%; Sen = 78%; F1-score = 78%. The corresponding ROC curve is shown in **Figure 20**. The AdaBoost classifier is a good tool for distinguishing sick from healthy subjects.

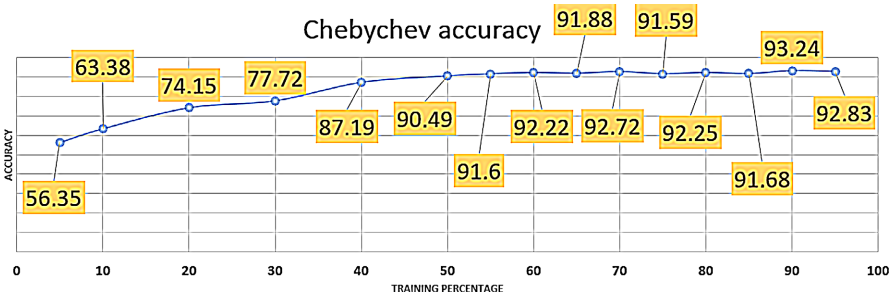

Figure 19. Evolution of the accuracy in the XGBoost classification.

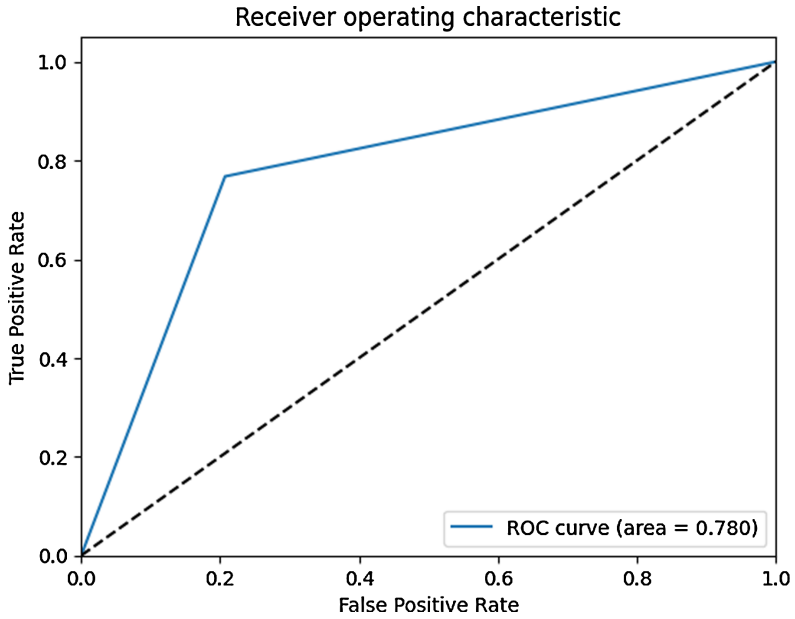

Figure 20. ROC curve obtained for Adaboost classification without dimension reduction.

Using the GradientBoosting algorithm on a dataset of which 70% was used for training and 30% for testing, results in an accuracy of 84.01%. Using dimension reduction, the best accuracy is achieved using 88% of the dataset for training. Therefore, using this percentage for the training set, the best accuracy obtained was 84.59%, Pre = 85%, Recall = 85%, F1-score = 85% and AUC= 0.85. The associated confusion matrix is given by

$$\begin{pmatrix} 2567 & 383 \\ 526 & 2424 \end{pmatrix}$$

The GradientBoosting classifier emerges as one of the best tool for diagnosing patients.

Using the Random Forest algorithm on a dataset of which 80% was used for training and 20% for testing, results in an accuracy of 80.73%, Pre = 83.47%, Recall = 76.63%, F1-score = 79.90% and AUC= 0.807. The associated confusion matrix is given by

$$\begin{pmatrix} 4170 & 746 \\ 1149 & 3767 \end{pmatrix}$$

The corresponding ROC curve is shown in Figure 21. The Random Forest classifier emerges as one of the best tool for diagnosing patients.

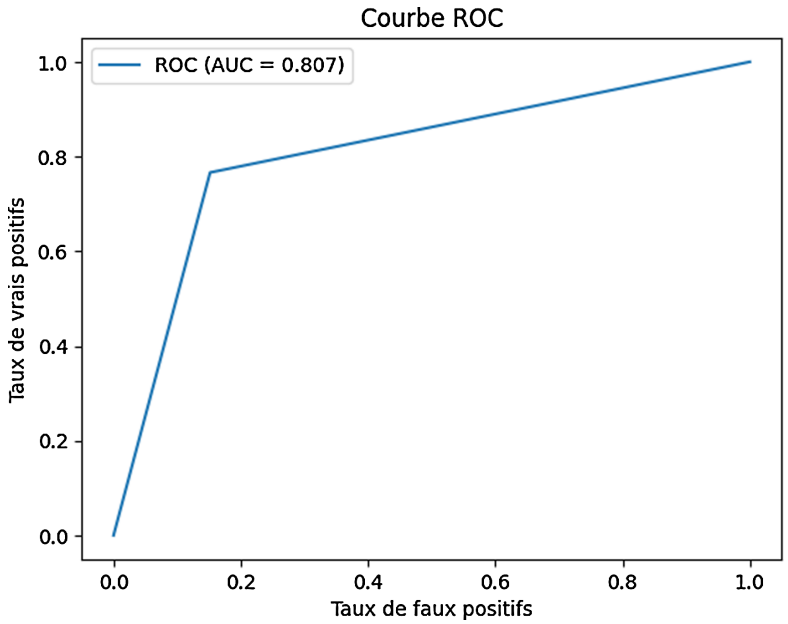

Figure 21. ROC curve obtained for Random Forest classification without dimension reduction.

The data used for classification with artificial neural networks such as the MLP were divided into three blocks; 80% of the data were used for training (20% of the training set is used for validation), 10% for the test and the remaining 10% for data prediction. The data taken as the input of the classifier correspond to the various features extracted in the previous phase. Therefore, using the architecture 39-1000-1, the hyperparameters are: epochs = 100, batch-size = 32, learning_rate = 0.001, optimizer = SGD, loss = binarycrossentropy, dropout = 0.3 and the best accuracy obtained was 50% with AUC= 0.500. This shows that the classifier is discriminative and, therefore, able to distinguish between sick and healthy people.

3.5.3. Classification Results without Spectral Analysis and without Reduction of Dimension

In this subsection, classification is performed using the original signal without the proposed methodology.

For the AdaBoost algorithm, on a dataset in which 70% was used for training and 30% for testing, the following results are obtained: Acc = 75.91%; Pre = 76%; Sen = 76%; F1-score = 76%. The corresponding ROC curve is shown in Figure 22. The AdaBoost classifier is a good tool for distinguishing sick from healthy subjects.

3.6. Tests to Validate the Use of Spectral Analysis

In this subsection, tests are done. First, the original signal was divided into 96 segments, each 1s in length. Temporal features were then extracted to construct a matrix. These include:

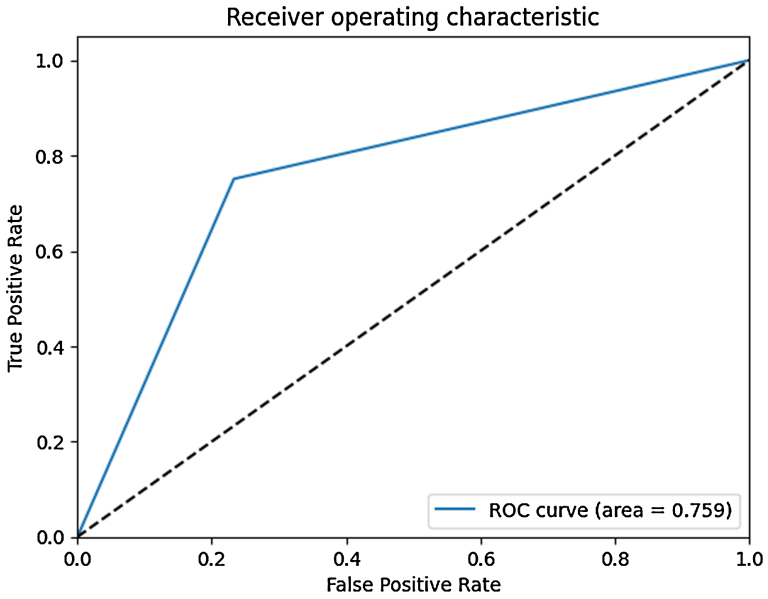

Figure 22. ROC curve obtained for Adaboost classification without spectral analysis and without dimension reduction.

- Mean and Standard deviation to see the overall range
- Skewness is given by formula (10) and kurtosis is given by formula (11) to detect whether the original signal contains artifacts or spikes
- Root Mean square to determine the total energy of the signal before filtering
- Peak-to-Peak to detect sudden changes in the signal
- Number of zero crossings

The spectral coefficients were divided into 96 segments. Spectral features were extracted from the spectral coefficients obtained after interpolation using the Chebyshev polynomial. These features are used to construct a matrix, which includes the following:

- total spectral power: measures the total energy of the signal after filtering
- spectral centroid: indicates where the main concentration of energy lies in the spectrum
- spectral bandwidth: measures the spread of energy
- spectral entropy: measures the disorder or complexity of the spectrum
- maximum amplitude and its position to determine the dominant filtered frequency and its strength

These two feature matrices are combined and merged into a single dataset. The Boruta algorithm is used to select the most important features. After applying this algorithm, all temporal and spectral features were selected, demonstrating that spectral analysis using Chebyshev polynomial interpolation had a significant impact.

With these two sets of data—where Set 1 corresponds to the temporal characteristics of the original signal and Set 2 corresponds to the characteristics selected by Boruta; the Shapiro test is performed to verify the normality of the data sets. Since the Shapiro test yielded a negative result, the Mann-Whitney U test was ap-

plied.

In this test, the null hypothesis (H0) states that the accuracy scores obtained by applying the XGBoost classification algorithm to each group are equal. The alternative hypothesis (H1), on the other hand, posits that accuracy improves when using the selected features, which include Chebyshev coefficients.

At the end of this test, the p-value = 0. This led us to conclude that spectral analysis using the Chebyshev polynomial is significant and beneficial for classification.

4. Discussions

The initial dataset consists of two subsets, namely the subset of healthy individuals and the subset of sick individuals, and takes into account an eye state parameter having two values, namely "eyes opened" and "eyes closed", which influence the EEG signal of the individuals involved in the signal recording. The new technique for the automatic diagnosis of AD proposed in this paper yields performances as high as those of the methods proposed in the literature. This technique is based on the analysis of EEG signals using orthogonal Chebyshev and has not yet been used to diagnose this pathology.

One of the strengths of this study is in the use of Chebyshev polynomials for signal processing, specifically spectral analysis. In methods used in the literature, such as VMD, Discrete Wavelet Transform, and Multitaper, the signal is divided into segments to be approximated for each segment. Unlike these methods, which result in significant data loss, estimation errors, and discontinuities owing to segmentation, Chebyshev polynomials provide optimal signal interpolation. Indeed, they allow for the creation of a continuous curve on which sampling can be performed without discontinuities, without the risk of losing a peak or the original coherent structure that acts as a natural low-pass filter, thus preserving pathological details while filtering out noise. These results suggest that the DChT is highly accurate in detecting low frequencies that are often associated with neurodegeneration. Using the DCHT, spectral coefficients were calculated, which reflect changes in brain activity in AD.

Feature extraction and selection are critical aspects of AD diagnosis. To further demonstrate the importance of this spectral analysis method, Boruta's algorithm was used to select relevant features from the two sets of data. Set1 comprises temporal features extracted from the original signal, whereas Set2 comprises spectral features extracted from the spectral coefficients after interpolation into the Chebyshev orthogonal basis. After applying Boruta, which removes noisy features by comparing the magnitude of the actual variables to shadow features, all features from both sets were selected, which supports DCHT as a useful spectral analysis tool.

Furthermore, Shapiro's test was applied to the two Sets to verify their normality. The first Set is Set1, and the second Set includes all features selected by Boruta. Because the Shapiro test yielded a negative result, the Mann-Whitney test was in-

tegrated after feature selection. The use of this nonparametric test confirmed that the characteristics identified by Boruta exhibit statistically distinct distributions between the two sets. The p-value is $0.0 < 0.5$. This second validation underscores the importance of using DCHT as a spectral analysis tool.

After the spectral analysis, feature extraction was performed. Each extracted feature provides unique information to maximize dimensionality reduction. This process is performed to condense complex information without significant loss. The use of PCA allowed us to reduce the dataset from $12{,}288 \times 85$ to 85×85. This reduction in size minimized the risk of overfitting. It can be noticed that dimensionality reduction has a significant impact on the classification results depending on the classification algorithm. Next, to ensure that as much information as possible was retained during dimension reduction, two types of features were extracted: 85 features (based on the spectral coefficients, their absolute values, the original signal, and the reconstructed signal) to capture as much signal information as possible, and seven spectral features (based only on each spectral coefficient) to test the robustness of Chebyshev using Boruta's algorithm and the Mann-Whitney test. On the one hand, the accuracy without dimension reduction is better than that with dimension reduction for the kNN and XGBoost algorithms. On the other hand, the accuracy with dimension reduction is better than that without dimension reduction for the AdaBoost, GradientBoosting and MLP algorithms. Regarding dimension reduction, the best accuracy is achieved by the GradientBoosting algorithm, with a value of 88.64%. Without dimensionality reduction, the best accuracy of 93.24% was achieved by the XGBoosting algorithm. The impact of dimensionality reduction on classification is examined to ensure that rare information is not lost, which enables the diagnosis of rare cases of AD. Our results are better than those reported in the literature.

A ROC plots the possible TPR against the FPR to graphically represent the performance of a binary classifier at different classification thresholds. The corresponding ROC-AUC score shows how well the classifier distinguishes between the positive and negative classes. Knowing that a high ROC-AUC score indicates better performance and that this score is 0.932 for the GradientBoosting algorithm and 0.808 for the XGBoosting algorithm, we can say that these two algorithms make a relatively clear distinction between sick and healthy individuals. As shown in **Table 3**, the classification accuracies observed for AD detection using the EEG signals varied considerably. The accuracies range from 39.71% to 93.24% for the proposed technique. While maintaining the recording criterion "eyes opened/eyes closed", the model distinguished healthy individuals from sick individuals.

The curves of the evolution of the accuracy of the Gradient Boosting and XGBoost algorithms as a function of the percentage of data allocated to training in **Figure 15** and **Figure 19** show that the accuracy increases overall with this percentage. These figures also show interesting accuracies for low percentages of the dataset allocated to training. For instance, when the training set represents 5% of the dataset, and the testing set represents 95% of the dataset, the accuracies

Table 3. Summary of work on the automatic diagnosis of AD.

Author	Input data	Number of extracted features	Signal processing	Classification algorithm	Dimension Reduction	Accuracy and train-test percentage
[11]	MRI (382 subjects)	/	/	CNN (DL)	No	75%
[12]	MRI (416 subjects)	/	/	CNN (DL)	No	94%
[13]	MRI (9430 subjects)	/	/	MCENN (ML)	No	90%
[14]	EEG (30 subjects)	/	/	SVM (ML)	No	92%
[15]	EEG (88 subjects)	/	ICA	MLP (DL)	No	73.12%
[16]	EEG (48 subjects)	49	Multitaper	EL (ML)	No	93.04% (70-30)
[17]	EEG (48 subjects)	/	DWT	kNN (ML)	No	91.12%
[18]	EEG (48 subjects)	/	MSPCA VMD	RF (ML)	No	98.42%
[19]	MRI	/	/	Neural Network (DL)	No	98.36%
[20]	EEG (100 subjects)	/	Spectral and Wavelet Transform	SVM (ML)	No	94%
[21]	EEG (20 subjects)	1	/	Deep Neural Network (DL)	No	75%
Our Work	EEG (48 subjects)	85	Preprocessing	SVM (ML)	Yes	42.65% (80-20)
			DChT PCA	kNN (ML)	Yes	39.71% (80-20)
				kNN (ML)	No	71.80% (80-20)
				XGBoost (ML)	Yes	84.38% (80-20)
				XGBoost (ML)	No	93.24% (90-10)
				AdaBoost (ML)	Yes	80.77% (70-30)
				AdaBoost (ML)	No	78.03% (70-30)
				GradientBoosting (ML)	Yes	88.64% (88-12)
				GradientBoosting (ML)	No	84.59% (88-12)
				RF (ML)	Yes	73.53% (80-20)
				RF (ML)	No	80.73% (80-20)
				MLP (DL)	Yes	59.38% (80-10-10)
				MLP (DL)	No	50% (80-10-10)
			No	AdaBoost (ML)	No	75.91% (70-30)

are greater than 56%. When combined with the proposed model, these classifiers suggest that the model has a good generalization ability.

Classification was also performed on a model that contained only the original features to study the impact of the proposed model on classification. The results obtained using DChT for the AdaBoost classifier are quite interesting: 80.77% with dimensionality reduction and 78.03% without dimensionality reduction,

compared to 75.91% without spectral analysis and without dimensionality reduction. This clearly demonstrates that the DChT is a very effective spectral analysis tool for diagnosing AD.

Compared to recent studies, our model distinguishes between sick and healthy individuals with 100% sensitivity. Each feature can be linked to the signal and is thus interpreted. These capabilities suggest that this model can be integrated into decision-support solutions for the diagnosis of AD.

Despite these promising results, this study had several limitations. First, we did not have access to the programs of the algorithms previously proposed by the authors in the literature, nor to their datasets to test their algorithms on the same dataset as the technique proposed in this article. Second, although the sample size was statistically significant, it was limited. Testing on a much larger dataset would provide greater assurance of the generalizability of the proposed Chebyshev-based model. Furthermore, the spectral efficiency depends on the choice of the order n of the Chebyshev polynomial. We choose a variable value of (2 × number of variables). However, a more in-depth analysis of the impact of this order on the classification is required. Therefore, considering the results obtained by this new technique, we can say that Chebyshev polynomials constitute an effective tool and are a completely new technique for improving AD diagnosis.

5. Conclusion

In this study, we investigated the contribution of the Chebyshev polynomial to automatic diagnosis of AD using machine learning. This study led to a new technique for the automatic diagnosis of AD for which EEG signals are preprocessed and interpolated according to the Chebyshev orthogonal basis; spectral analysis was performed on this basis, 85 features were extracted, dimension reduction was performed to reduce observed redundancies and persistent noises, and classification was performed based on both machine learning and deep learning algorithms. The evaluation of the new technique involved the evaluation of accuracy, ROC-AUC, the impact of dimension reduction on accuracy, and the change in accuracy as a function of the percentage of the dataset used for training. This shows that the classification performance obtained by the proposed technique is as high as that of the methods proposed in the literature. The GradientBoosting algorithm provided the best accuracy, with a value of 88.64%, when dimensionality reduction was performed. In addition, without dimensionality reduction, the XGBoost algorithm provided the best accuracy, with a value of 93.24%. Furthermore, using the AdaBoost classifier without spectral analysis, the accuracy was 75.91%; with spectral analysis and dimension reduction, it achieved an accuracy of 80.77%, and without dimension reduction, the accuracy was 78.03%. This study highlights that Chebyshev polynomials constitute an effective tool for improving AD diagnosis. Moreover, the Mann-Whitney test was used to prove that the Chebyshev polynomial has an effect on the classification using the null hypothesis H0, which posits the equality of all the accuracies after classification of the selected

features and the temporal features using XGBoost. Therefore, H0 is rejected in favor of the superiority of the classification accuracy with all selected features. In conclusion, spectral analysis with DChT is an effective tool for improving classification, and the proposed model is one of the best for diagnosing AD. In the future, it will be interesting to explore new lines of research. For example, the early diagnosis of AD, as well as the diagnosis of AD at different stages of the disease.

Data Availability Statement

The authors confirm that the data supporting the results of this study are available at the following address: https://doi.org/10.17605/OSF.IO/S74QF.

Author Contributions

Conceptualization, Stéphanie-Claude Pelap Ngassa, Edith Belise Kenmogne and Clémentin Tayou Djamegni; methodology, Stéphanie-Claude Pelap Ngassa, Edith Belise Kenmogne and Clémentin Tayou Djamegni; writing—original draft preparation, Stéphanie-Claude Pelap Ngassa; writing—review and editing, Stéphanie-Claude Pelap Ngassa, Edith Belise Kenmogne and Clémentin Tayou Djamegni. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Meyer, F. (2021) La maladie d'alzheimer à l'ère des biomarqueurs: Le point de vue du clinicien. In: *La maladie de Alzheimer et son dépistage au laboratoire*, Université de Liège/CHU de Liège, Belgique, 4

[2] Walhovd, K.B., Fjell, A.M., Brewer, J., McEvoy, L.K., *et al.* (2010) Combining MR Imaging, Positron-Emission Tomography, and CSF Biomarkers in the Diagnosis and Prognosis of Alzheimer Disease. *American Journal of Neuroradiology*, **31**, 347-354.

[3] Alves, C.L., Pineda, A.M., Roster, K., Thielemann, C. and Rodrigues, F.A. (2022) EEG Functional Connectivity and Deep Learning for Automatic Diagnosis of Brain Disorders: Alzheimer's Disease and Schizophrenia. *Journal of Physics*: *Complexity*, **3**, Article 025001. https://doi.org/10.1088/2632-072x/ac5f8d

[4] Emigdio, Z. (2017) Human Mental States Classification Using EEG by Means of Genetic Programming. PhD Thesis, Instituto Tecnologico de Tijuana.

[5] Eisinberg, A. and Fedele, G. (2007) Discrete Orthogonal Polynomials on Gauss-Lobatto Chebyshev Nodes. *Journal of Approximation Theory*, **144**, 238-246.

[6] Nkengfack, L.C.D., Tchiotsop, D., Atangana, R., Louis-Door, V. and Wolf, D. (2021) Classification of EEG Signals for Epileptic Seizures Detection and Eye States Identification Using Jacobi Polynomial Transforms-Based Measures of Complexity and Least-Square Support Vector Machine. *Informatics in Medicine Unlocked*, **23**, Article 100536. https://doi.org/10.1016/j.imu.2021.100536

[7] Vujovic, Ž.Đ. (2021) Classification Model Evaluation Metrics. *International Journal of Advanced Computer Science and Applications*, **12**, 599-606. https://doi.org/10.14569/ijacsa.2021.0120670

[8] Hakim, B. and Brahim, S. (2018) Classification des images avec les réseaux de neurones convolutionnels. PhD Thesis, Université Mouloud Mammeri.

[9] Anguita, D., Ghelardoni, L., Ghio, A., *et al.* (2012) The 'K' in K-Fold Cross Validation. *The* 20*th European Symposium on Artificial Neural Networks*, *Computational Intelligence and Machine Learning*, Bruges, 25-27 April 2012, 441-446.

[10] Pineda, A.M., Ramos, F.M., Betting, L.E. and Campanharo, A.S.L.O. (2020) Quantile Graphs for EEG-Based Diagnosis of Alzheimer's Disease. *PLOS ONE*, **15**, e0231169. https://doi.org/10.1371/journal.pone.0231169

[11] Boudali, A. (2022) Détection précoce de la maladie d'Alzheimer sur les images IRM. PhD Thesis, Université Ibn Khaldoun-Tiaret.

[12] Islam, J. and Zhang, Y.Q. (2018) Brain MRI Analysis for Alzheimer's Disease Diagnosis Using an Ensemble System of Deep Convolutional Neural Networks. *Brain Informatics*, **5**, Article No. 2. https://doi.org/10.1186/s40708-018-0080-3

[13] Liu, C.Q., Huang, F. and Qiu, A.Q. (2023) Monte Carlo Ensemble Neural Network for the Diagnosis of Alzheimer's Disease. *Neural Networks*, **159**, 14-24. https://doi.org/10.1016/j.neunet.2022.10.032

[14] Zhao, Y.L. and He, L.H. (2014) Deep Learning in the EEG Diagnosis of Alzheimer's Disease. In: *Lecture Notes in Computer Science*, Springer, 340-353. https://doi.org/10.1007/978-3-319-16628-5_25

[15] Miltiadous, A., Tzimourta, K.D., Afrantou, T., Ioannidis, P., Grigoriadis, N., Tsalikakis, D.G., *et al.* (2023) A Dataset of Scalp EEG Recordings of Alzheimer's Disease, Frontotemporal Dementia and Healthy Subjects from Routine EEG. *Data*, **8**, Article 95. https://doi.org/10.3390/data8060095

[16] Göker, H. (2023) Detection of Alzheimer's Disease from Electroencephalography (EEG) Signals Using Multitaper and Ensemble Learning Methods. *Uludağ Üniversitesi Mühendislik Fakültesi Dergisi*, **28**, 141-152. https://doi.org/10.17482/uumfd.1142345

[17] Aslan, Z. (2022) EEG sinyallerini kullanarak alzheimer hastalığının otomatik tespiti için bilgisayar destekli tanı sistemi. *Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi*, **13**, 213-220. https://doi.org/10.24012/dumf.1092569

[18] Aslan, Z. (2024) A Novel Computer-Aided Diagnostic System for Alzheimer's Diagnosis Using Variational Mode Decomposition Method. *Circuits*, *Systems*, *and Signal Processing*, **43**, 615-633. https://doi.org/10.1007/s00034-023-02496-y

[19] Lodha, P., Talele, A. and Degaonkar, K. (2018) Diagnosis of Alzheimer's Disease Using Machine Learning. 2018 4*th International Conference on Computing Communication Control and Automation* (*ICCUBEA*), Pune, 16-18 August 2018, 1-4. https://doi.org/10.1109/iccubea.2018.8697386

[20] Bairagi, V. (2018) EEG Signal Analysis for Early Diagnosis of Alzheimer Disease Using Spectral and Wavelet Based Features. *International Journal of Information Technology*, **10**, 403-412. https://doi.org/10.1007/s41870-018-0165-5

[21] Kim, D. and Kim, K. (2018) Detection of Early Stage Alzheimer's Disease Using EEG Relative Power with Deep Neural Network. 2018 40*th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (*EMBC*), Honolulu, 17-21 July 2018, 352-355. https://doi.org/10.1109/embc.2018.8512231