Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Schizophrenia classification using machine learning on resting state EEG signal

J. Ruiz de Miras [a,*], A.J. Ibáñez-Molina [b], M.F. Soriano [c], S. Iglesias-Parro [b]

[a] *Software Engineering Department, University of Granada, Granada, Spain*
[b] *Department of Psychology, University of Jaén, Jaén, Spain*
[c] *St. Agustín University Hospital, Linares, Jaen, Spain*

ARTICLE INFO

ABSTRACT

Schizophrenia is a severe mental disorder associated with a wide spectrum of cognitive and neurophysiological dysfunctions. Early diagnosis is still difficult and based on the manifestation of the disorder. In this study, we have evaluated whether machine learning techniques can help in the diagnosis of schizophrenia, and proposed a processing pipeline in order to obtain machine learning classifiers of schizophrenia based on resting state EEG data. We have computed well-known linear and non-linear measures on sliding windows of the EEG data, selected those measures which better differentiate between patients and healthy controls, and combined them through principal component analysis. These components were finally used as features in five standard machine learning algorithms: k-nearest neighbours (kNN), logistic regression (LR), decision trees (DT), random forest (RF) and support vector machines (SVM). Complexity measures showed a high level of ability in differentiating schizophrenia patients from healthy controls. These differences between groups were mainly located in a delimited zone of the right brain hemisphere, corresponding to the opercular area and the temporal pole. Based on the area under the curve parameter in receiver operating characteristic curve analysis, we obtained high classification power in almost all of the machine learning algorithms tested: SVM (0.89), RF (0.87), LR (0.86), kNN (0.86) and DT (0.68). Our results suggest that the proposed processing pipeline on resting state EEG data is able to easily compute and select a set of features which allow standard machine learning algorithms to perform very efficiently in differentiating schizophrenia patients from healthy subjects.

1. Introduction and related work

Schizophrenia is a severe mental disorder which affects approximately 1 % of the world's population. Patients with schizophrenia suffer psychotic symptoms such as hallucinations and delusions [1], along with a wide variety of cognitive and motor dysfunctions. The symptoms' onset is usually during adolescence and early adulthood (between 14 and 30), and it has been consistently found that the time between the onset of symptoms and diagnosis and treatment is one of the best predictors of later prognosis [2]. Thus, early diagnosis is of great relevance in the course of the disorder. Currently, there are no reliable biomarkers for schizophrenia and, therefore, its diagnosis relies on the subjective evaluation of the clinicians [3].

Machine learning is a set of artificial intelligence tools used to construct data models representing the knowledge hidden in huge data sets [4]. These data models are simplified versions of the original data

constructed by using a set of features extracted from them. The machine learning tools consist of algorithms that construct these data models automatically from the set of features extracted from the original data (training data). The data models are then used to predict or classify new data (test data). In this study, we have applied five of the most widely used machine learning classification algorithms on an EEG dataset from patients with schizophrenia and healthy controls at rest: k-nearest neighbours (kNN), logistic regression (LR), decision trees (DT), support vector machines (SVM) and random forest (RF). These techniques have been used in many successful applications of machine learning in biomedicine such as identification of anticancer peptides (iACP-GAEnsC model) [5], classification of DNA-bindings proteins (DP-BINDER method) [6] and prediction of antifungal peptides (iAFPs-EnC-GA predictor) [7], among many others.

Several recent studies have shown that machine learning can also be a very useful tool in classifying schizophrenia. In [8], an ensemble

---

Table 1

Demographic data. Values expressed as mean $\pm$ standard deviation. SCIP-S: Spanish version of Screen for Cognitive Disability in Psychiatry; VLi: immediate verbal learning; VLd: delayed verbal learning; VF: verbal fluency; WM: working memory and PS: processing speed. [a] Mann-Whitney test; [b] χ^2 test; [c] t-student test.

	HC	SCZ	Test and p-value
N	20	11	
Age (y)	40.7 $\pm$ 11.9	36.2 $\pm$ 10.2	$U = 85, p = 0.31$ [a]
Gender (M:F)	13:7	9:2	$\chi^2 = 0.97, p = 0.32$ [b]
Education (Primary:Secondary: High)	1:12:7	2:8:1	$\chi^2(2) = 3.29, p = 0.19$ [b]
SCIP-S – VLi	21.1 $\pm$ 4.4	16.2 $\pm$ 3.9	$t = -3.1, p = 0.004$ [c]
SCIP-S – VLd	6.8 $\pm$ 2.4	5.8 $\pm$ 2.1	$t = -1.07, p = 0.28$ [c]
SCIP-S – VF	17.9 $\pm$ 4.1	13.6 $\pm$ 4.5	$t = -2.69, p = 0.01$ [c]
SCIP-S – WM	19.5 $\pm$ 3.0	16.7 $\pm$ 4.0	$t = -2.16, p = 0.03$ [c]
SCIP-S – PS	11.3 $\pm$ 2.9	7.6 $\pm$ 3.0	$t = -3.27, p = 0.003$ [c]

method was used to classify first-episode schizophrenia from functional magnetic resonance imaging (fMRI) data. A classification accuracy of 84.7 % was obtained with that method. Ensemble methods and several neuroanatomical measures extracted from MRI data were also used in [9] to achieve a classification accuracy of 87 %. An accuracy of 85 % in classifying patients with schizophrenia and autism by using kNN and cortical features extracted from MRI data was achieved in [10]. A SVM classifier based on several structural MRI features allowed Zhu et al. to achieve an accuracy of 76 % in classifying schizophrenia at early clinical stages [11]. SVM was also used to classify schizophrenia with an accuracy of 85 % by using brain-wise functional connectivity features from MRI data [12]. The capacity of generalization and replication of machine learning for classification of schizophrenia based on functional MRI data were assessed in [13]. An accuracy of 70 % was obtained when an external validation was performed on the model trained with the main dataset. Li et al. [14] proposed a deep canonically-correlated sparse autoencoder to classify schizophrenia patients from healthy controls. Their method was applied on both single nucleotide polymorphisms and fMRI. An accuracy of 95.65 % was achieved with the genomic data, while the accuracy was 80.53 % when using fMRI. Textual data have also been used to classify schizophrenia. In [15], several linguistic features were extracted from social media posts, and then a RF method was used to obtain a classification accuracy of 96 %. Demographic data and clinical assessment questionnaires were used in [16] to train a SVM classifier which achieved 90 % of accuracy in classifying schizophrenia. A review of machine learning techniques applied in classifying schizophrenia from several kinds of data such as structural and functional MRI, genomic information, interview audio records and textual clinical data can be found in [17].

There have been several applications of machine learning techniques in schizophrenia classification based on EEG data. These methods can be grouped together taking into account the machine learning algorithm used, the measures determining the features in the algorithm and how the input EEG data was obtained. We briefly summarize now the most recent methods. For the case of event-related potential EEG signals, Zhang [18] developed a random forest algorithm for the differentiation between schizophrenia and healthy controls, using N1 and *P*3 features extracted when basic sensory tasks involving button press and auditory tone were performed. In this study the best classification accuracy obtained was 81 %. In [19], kNN, decision tree and SVM classifiers were used to identify children at risk of developing schizophrenia. In this study the mean amplitude of early and mid-latency event related potential were used as features during a passive auditory task. The best accuracy obtained was 44 %, improved to 72 % by using a deep learning approach based on convolutional neural networks. Several recent studies have analyzed the EEG signal obtained in resting state. Buettner et al. [20] achieved 71 % of accuracy in classifying schizophrenia using the random forest classifier based on power features extracted from

several frequency bands. Twelve statistical features on five bands extracted from 5-minute resting-state EEG signal were used in a SVM model allowing Tikka et al. to achieve an accuracy of 79 % in classifying schizophrenia patients from healthy controls [21]. Non-linear measures such as entropy, largest Lyapunov exponent, Hurst exponent, and recurrence quantification analysis have also been employed in SVM algorithms, achieving an accuracy of approximately 92 % in the classification of patients [22,23]. Finally, Sun et al. [24] developed a method that, from resting state EEG data, constructed an RGB image from which a feature based on the fuzzy entropy was extracted. This feature was used in a hybrid model based on convolutional neural networks and long-short-term memory networks to obtain a schizophrenia classifier with a high accuracy of 99 %.

In the present study, we further explore the capabilities of standard machine learning algorithms in obtaining classifiers of schizophrenia based on resting state EEG data. Our approach consists of a processing pipeline whereby the EEG data is split into sliding windows over which a set of linear and non-linear measures are computed. Those measures which better differentiate between patients and controls are selected, and a principal component analysis is then performed on them. The results of the principal component analysis are finally used as the features in the machine learning algorithms. This processing pipeline allows us to select a set of features which works efficiently regardless of the machine learning algorithm used. This characteristic is not present in most of the previous studies, where good accuracy values were obtained for only one technique or a very limited number of machine learning methods. Our method is based on well-known and easily computable EEG measures which are used in standard machine learning algorithms available in common statistical software tools. This fact allows the scientific community to easily reproduce and use our processing pipeline. This is not a common characteristic in those previous studies which obtained very good accuracy results but at the cost of using complex and custom algorithms in order to extract the features and design de machine learning models.

In the following sections, we first describe the dataset of the study and how it was obtained. Then, we briefly review the EEG measures used and the machine learning algorithms tested. Next, we show the process implemented to extract the features from the EEG data. And finally, we present and discuss the classification results obtained by the machine learning algorithms.

2. Methods

2.1. Subjects

We included 31 subjects in this study (see Table 1 for demographics): 20 healthy control (HC) subjects (13 males, 7 females, mean age: 40.7 $\pm$ 11.9) and 11 subjects suffering from schizophrenia (SCZ) (9 males, 2 females, mean age: 36.2 $\pm$ 10.2). Participants in the control group were recruited from the students of the University of Jaén, the staff of the University Hospital of San Agustin Linares (Jaén) and from an Adult School. The patient group was recruited at the Hospital Universitario San Agustín (Linares, Jaén).

The inclusion criteria for participation were an ICD-10 diagnosis of schizophrenia (F20), psychotic disorder (F23) or schizophrenic disorder (F25). The diagnosis was made by the clinician in charge of the patient. The duration of the disease was, on average, 15.72 $\pm$ 10.19 years. Regarding medication, all participants were receiving antipsychotic medication (all atypical). In addition to antipsychotics, one of the participants was receiving antidepressants. Due to differences in active principles, doses and administration methods, we converted the doses of all antipsychotics into chlorpromazine equivalents (818.18 $\pm$ 407.75 mg).

For both groups, the exclusion criteria were any of the following: a concurrent diagnosis of neurological disorder, a concurrent diagnosis of substance abuse, a history of developmental disability, an inability to
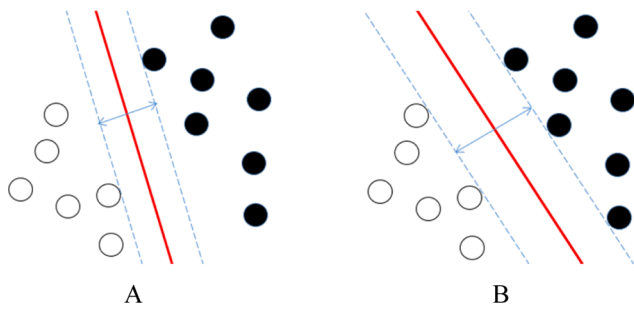
Fig. 1. Margins and support vectors in SVM. A) Hyperplane separating the classes but with a small margin; the support vector contains two instances. B) Hyperplane separating the classes with the maximum margin; the support vector contains five instances.

sign informed consent, vision disorders (vision disorders that, although corrected with surgery, glasses or contact lenses, result in a loss of visual acuity, e.g., cataracts), or hearing disorders (unless corrected with hearing aids or surgery). In addition, an exclusion criterion for the control group was the diagnosis of a mental disorder (as reported verbally by the participants). All participants gave written informed consent in accordance with the Declaration of Helsinki, and the Jaén Research Ethics Committee approved the study.

Finally, in order to measure the cognitive functioning in both groups, we used a Spanish adaptation of the Screen for Cognitive Disability in Psychiatry (SCIP-S, [25]), which has shown satisfactory reliability and good concurrent validity with neuropsychological batteries [26]. SCIP-S has been specifically developed to detect the main cognitive deficits of people with some type of mental illness, although it can also be used to assess the cognitive status of adults without mental illness. SCIP-S is a short paper-and-pencil test that can be completed in about 15 min. It provides a score for each of the following subscales: immediate and delayed verbal learning, verbal fluency, working memory and processing speed. Descriptive statistics for each group can be found in Table 1. We performed t-tests for the scores on the five subtests of the SCIP-S. As

can be seen in Table 1, the group of healthy controls scored significantly better than the patients on all dimensions of the SCIP-S except for delayed verbal learning.

2.2. EEG acquisition and processing

EEG data was acquired in a session where participants sat in a laboratory room at the hospital in a comfortable chair with a laptop computer placed on the desk in front of them. The laptop screen was located about 70 cm from the participant's eyes. Then, the experimenter proceeded to place the cap montage of 31 active electrodes in the 10–20 system with positions FP1, FP2, F7, F3, Fz, F4, F8, FT9, FC5, FC1, FC2, FC6, FT10, T7, C3, C4, T8, TP9, CP5, CP1, CP2, CP6, TP10, P7, P3, Pz, P4, P8, O1, Oz and O2. All electrodes were referenced to both mastoids, and impedances were kept below 5 kOhm. Signals were recorded at a frequency of 500 Hz. During the resting state task, participants were instructed to look at a light grey fixation cross placed in the centre of a black background for 3 min. They were asked not to move and were told that they could think about whatever they wanted, in silence. The experimenter was seated behind them, out of view of the participants, in the same room. In a second session, participants underwent cognitive evaluation (SCIP-S), and the remainder of demographic data (age, gender and educational level) were collected.

Data processing was performed with EEGLAB and custom MATLAB functions. For each participant, we selected 3 min of continuous data. Blinks and other artefacts were extracted using infomax ICA. ICA components with artefacts were selected by visual inspection of the scalp topography, power spectra and raw activity from all components. The resulting EEGs, after denoising, were used as inputs for a custom MATLAB script developed to obtain the measures described in the following section.

2.3. Extracting features from EEG

Seventeen different linear and non-linear measures were extracted from EEG data, namely Higuchi fractal dimension (HFD), approximate
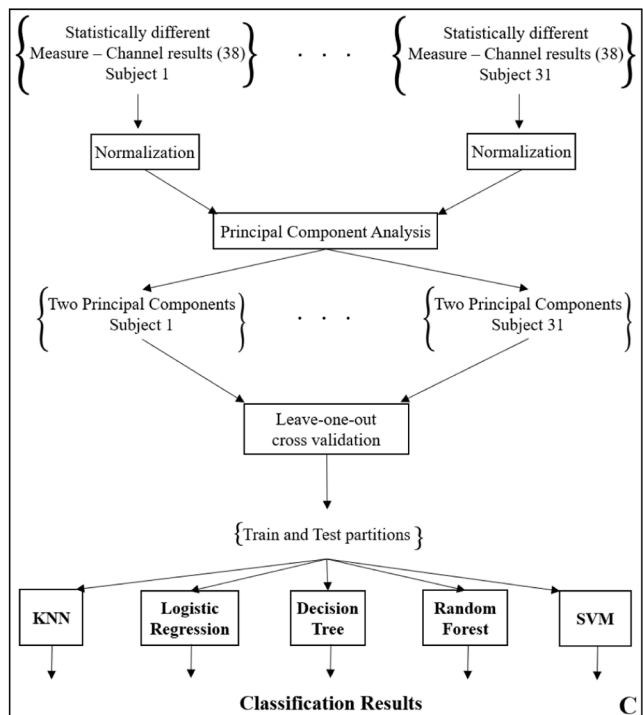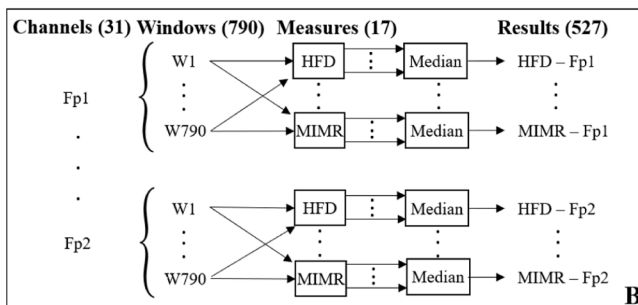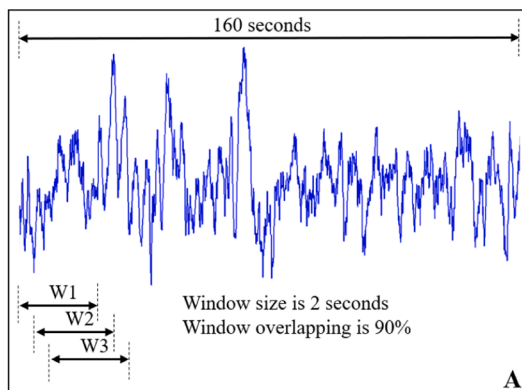

Fig. 2. Sliding window configuration (A). Computation of all measures for each channel of one subject (B). Classification pipeline: selection of pairs measure-channel showing significant differences, normalization, principal components analysis, leave-one-out cross validation and machine learning classification (C).

Table 2

Measures and channels for which HC vs SCZ comparison is statistically different. Channels Fp1, Fz, F3, FT9, C3, T7, TP9, CP1, Pz, P3, P7, O1, Oz, P8, F4 and Fp2, and measures detrended fluctuation analysis, Hjorth activity, total energy, power band alpha, power band beta and power band theta did not show significant differences for any measure nor in any channel, respectively, between HC and SCZ groups. *: *t*-test p-value < 0.05, **: *t*-test p-value < 0.01. HFD: Higuchi fractal dimension, ApEn: approximate entropy, SamEn: sample entropy, CD: correlation dimension, LZC: Lempel-Ziv complexity, MIMR: mutual information of multiple rhythms, HM: Hjorth mobility, HC: Hjorth complexity, DP: delta power band, LLE: largest Lyapunov exponent and DET: determinism.

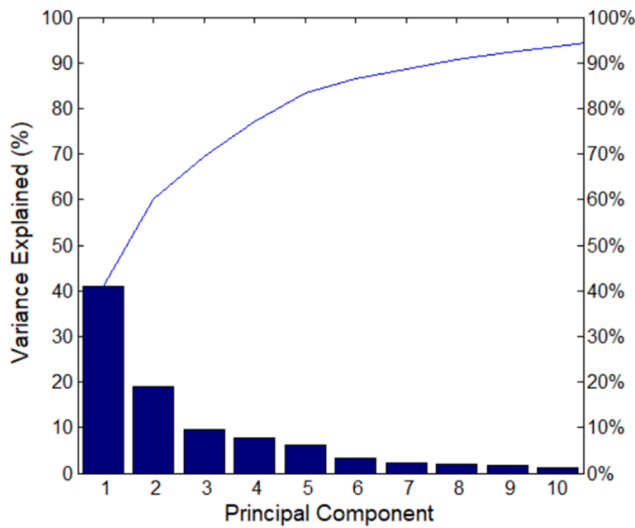|	HFD	ApEn	SamEn	CD	LZC	MIMR	HM	HC	DP	LLE	DET
F7									*		
FC5	*								*		
FC1									*		
CP5										*	
O2	*									*	
P4										*	
TP10										*	
CP6										**	
CP2										*	
C4	*									*	
T8						*				**	
FT10				*	*					*	
FC6	*	*	*	*	*	*	*	*		*	*
FC2						*					
F8	*	*	*	*	*	*	*	*	*		*

Fig. 3. Percentage of variance explained by the first ten principal components.

entropy (ApEn), sample entropy (SamEn), correlation dimension (CD), scaling coefficient in detrended fluctuation analysis (DFA), Hjorth activity (HA), Hjorth mobility (HM), Hjorth complexity (HC), total energy (TE), power of alpha, beta, delta and theta bands (AP, BP, DP and TP), Lempel-Ziv complexity (LZC), largest Lyapunov exponent (LLE), determinism measure in recurrence quantification analysis (DET) and mutual information of multiple rhythms (MIMR).

This range of measures was selected because we wanted to cover widely used metrics in both time and frequency domains as well as from non-linear analysis. Next, we briefly describe the theory behind each measure and the experimental configurations used for its computation in this study.

2.3.1. Higuchi fractal dimension

The fractal dimension is a metric of the complexity of a signal which measures how the signal occupies its geometrical target space [27]. Higuchi's algorithm calculates the fractal dimension of a temporal signal (HFD) as follows [28]:

From a given signal $s = \{s(1), s(2), \ldots, s(N)\}$, k curves s_m^k are constructed as:

Table 3

Classification results. Bold results are misclassifications.

Subject	Group	KNN	Logistic Regression	Decision Tree	Random Forest	SVM
1	SCZ	SCZ	SCZ	SCZ	SCZ	SCZ
2		SCZ	SCZ	SCZ	**HC**	SCZ
3		SCZ	SCZ	SCZ	SCZ	SCZ
4		SCZ	**HC**	SCZ	**HC**	**HC**
5		**HC**	**HC**	**HC**	**HC**	**HC**
6		SCZ	SCZ	SCZ	SCZ	SCZ
7		SCZ	SCZ	SCZ	SCZ	SCZ
8		SCZ	SCZ	SCZ	SCZ	SCZ
9		SCZ	SCZ	SCZ	SCZ	SCZ
10		**HC**	**HC**	**HC**	**HC**	**HC**
11		SCZ	**HC**	**HC**	**HC**	**HC**
12	HC	HC	HC	HC	HC	HC
13		HC	HC	HC	HC	HC
14		HC	HC	HC	HC	HC
15		HC	HC	HC	HC	HC
16		**SCZ**	HC	HC	**SCZ**	HC
17		HC	HC	HC	HC	HC
18		HC	HC	**SCZ**	**SCZ**	HC
19		HC	HC	**SCZ**	HC	HC
20		HC	HC	HC	HC	HC
21		HC	HC	HC	HC	HC
22		HC	HC	HC	HC	HC
23		HC	HC	HC	HC	HC
24		HC	HC	**SCZ**	HC	HC
25		**SCZ**	**SCZ**	**SCZ**	**SCZ**	**SCZ**
26		HC	HC	HC	HC	HC
27		HC	HC	HC	HC	HC
28		HC	HC	HC	HC	HC
29		HC	HC	HC	HC	HC
30		HC	**SCZ**	HC	HC	**SCZ**
31		HC	HC	HC	HC	HC

$$s_m^k = \left\{ s(m), s(m+k), \cdots, s\left(m + \left\lfloor \frac{N-m}{k} \right\rfloor \cdot k \right) \right\}, m = 1, 2, \cdots, k \qquad (2)$$

where m and k are integers and they indicate the initial time and the time interval, respectively. The length $L_m(k)$ of each curve s_m^k is calculated as:

$$L_m(k) = \frac{\left(\sum_{i=1}^{\lfloor \frac{N-m}{k} \rfloor} |s(m+ik) - s(m+(i-1)k)| \right) \cdot \frac{N-1}{\lfloor \frac{N-m}{k} \rfloor k}}{k} \qquad (3)$$

The length of the curve for time interval k, $L(k)$, is calculated as the average of the m curves $L_m(k)$ for $m = 1, \ldots, k$. If $L(k) \propto k^{-D}$, then the
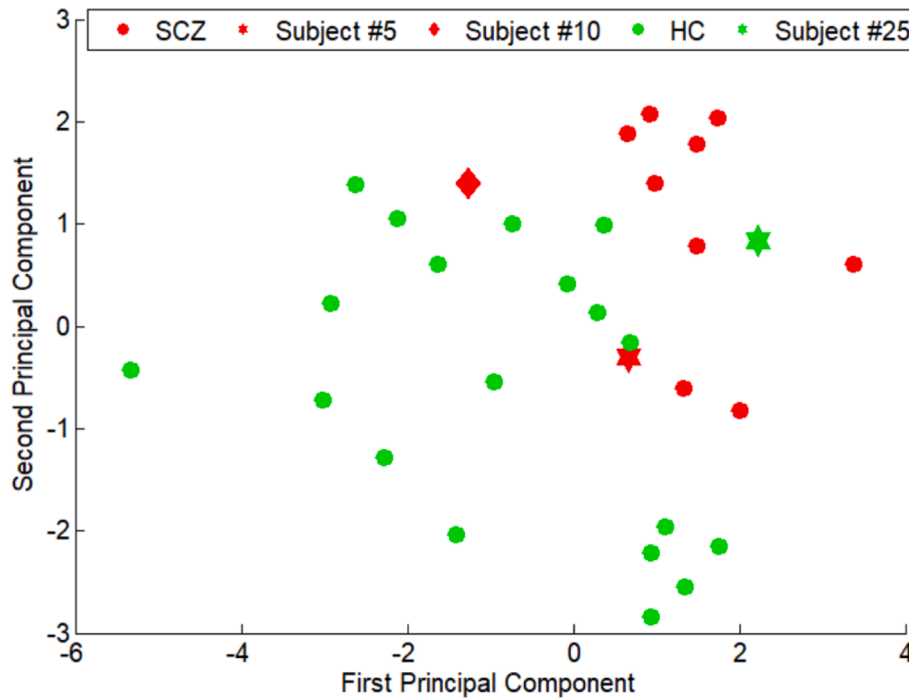
Fig. 4. Graphical representation of the dataset based on their two principal components.

Table 4
Performance scores for each classification algorithm.

	Sensitivity	Specificity	Accuracy	ROC AUC
KNN	0.82	0.90	0.87	0.86
Logistic Regression	0.64	0.90	0.81	0.86
Decision Tree	0.73	0.80	0.77	0.68
Random Forest	0.55	0.85	0.74	0.87
SVM	0.63	0.90	0.81	0.89

signal s is fractal-like with the dimension D.

HFD values of the EEG signals were computed in MATLAB R2013a using the software written by Selvam and Nadu [29]. Higuchi's algorithm needs a manual-set parameter, $kmax$, in order to establish the maximum size of the time intervals. We set the parameter $kmax$ with a value equal to half the number of time samples in the entry EEG. This value ensures that the curve obtained plotting HFD vs $kmax$, for $kmax$ between 1 and the number of time samples, plateaus at that value of $kmax$ [30].

2.3.2. Approximate and sample entropy (ApEn, SamEn)

Entropy can be defined as the rate of information production in a dynamical system [31]. There are methods for a direct estimation of the entropy of time series, but these methods are not adequate for analyzing EEG data, due to its length and noisy nature. Nevertheless, approximate entropy (ApEn) and sample entropy (SamEn) are two measures, closely related to entropy, which are easily applied to EEG data and other biological signals.

ApEn and SamEn measure the regularity of a signal by dividing it into epochs. The more frequent and more similar the epoch is, the lower the value of ApEn and the more regular the signal. Given three parameters: N, m and r, where N is the length of the EEG time series, m is the length of the epochs to be compared and r is a value of tolerance for accepting matches, ApEn(N, m, r) of the signal $s = \{s(1), s(2), …, s(N)\}$ can be defined as [32]:

$$\text{ApEn}(N, m, r) = \frac{\sum_{i=1}^{N-m+1} \ln\left(C_i^m(r)\right)}{N-m+1} - \frac{\sum_{i=1}^{N-m} \ln\left(C_i^{m+1}(r)\right)}{N-m}.$$

where $C_i^m(r) = \frac{B_i}{N-m+1}$, and B_i are the number of epochs $x_m(j)$ within r of $x_m(i)$, where $x_m(i) = \{s(i + k): 0 \leq k \leq m - 1\}$ is the epoch of m data points from $s(i)$ to $s(i + m - 1)$.

Sample entropy (SamEn) is an optimization of ApEn in order to avoid self-matching of the epochs, and it is defined as [31]:

ApEn$(N, m, r) = -ln\frac{A_i}{B_i}$, where A_i is the number of epochs $x_{m+1}(j)$ within r of $x_{m+1}(i)$.

In this study, ApEn and SamEn values of the EEG data were computed using the MATLAB code provided by Lee [33] and Martínez [34]. Parameters m and r were set to 2 and 0.1 respectively, as recommended for EEG data [31].

2.3.3. Correlation dimension (CD)

The correlation dimension (CD) is a type of fractal dimension that measures the dimensionality of the space occupied by a set of points [35]. CD behaves well for signals with a small number of points, and computing CD is easy and fast.

Given the signal $s = \{s(1), s(2), …, s(N)\}$, the correlation integral $C(l)$ is defined as [36]:

$C(l) = \lim_{N\to\infty}\frac{g}{N^2}$, where g is the number of pairs (i, j) whose distance $|s(i) - s(j)|$ is less than l. For small l's, $C(l)$ grows like a power $C(l) \sim l^{CD}$. CD can be estimated easily by using a log–log plot of correlation integral versus l.

In this study, we used the Gaussian kernel algorithm to estimate the values of CD of our EEG data using the MATLAB code provided in [37].

2.3.4. Lempel-Ziv complexity (LZC)

The Lempel-Ziv complexity (LZC) is a measure of the complexity of binary sequences that tests the randomness of the sequence by searching for repeated patterns in it [38]. The more irregular the sequence is, the greater the LZC. LZC can be computed easily following the algorithm proposed in [39]. The binary sequence is parsed into subsequences with different patterns. A complexity counter increases when a new subsequence appears and the following symbol is regarded as the beginning of the next subsequence. The scan continues until the end of the sequence, and finally LZC equals the complexity counter.
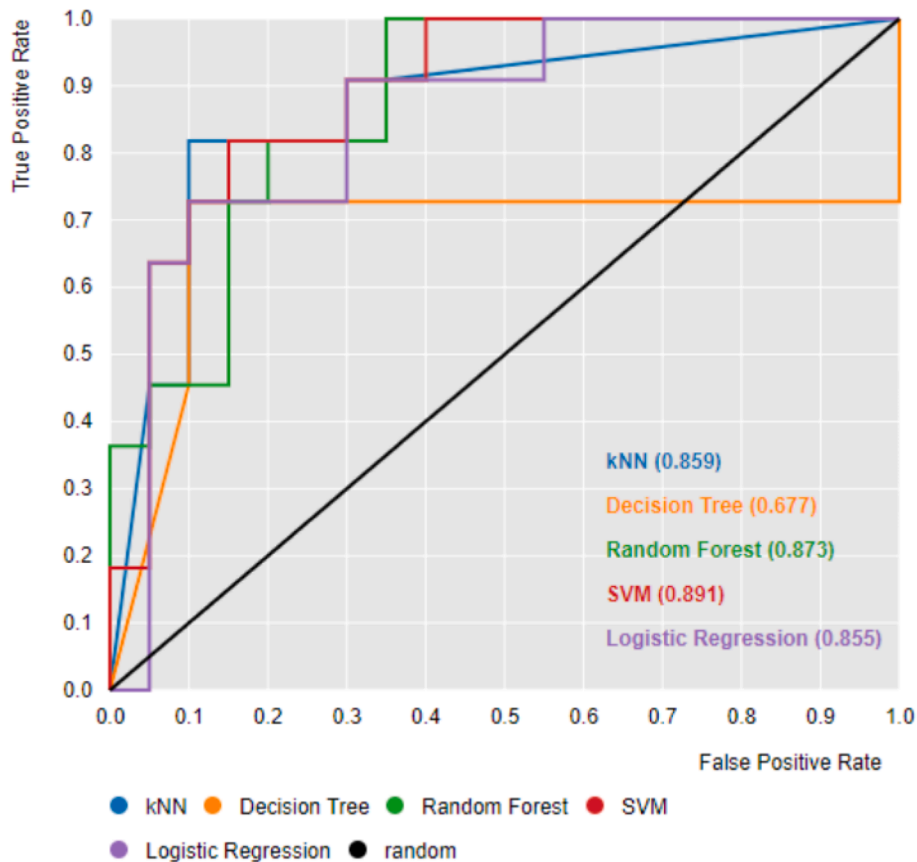
Fig. 5. ROC curve analysis evaluating the performance of classification algorithms. AUC values are shown inside parentheses.

Table 5
Performance scores for each classification algorithm using cognitive measures VLi, VF, WM and PS as input features.

	Sensitivity	Specificity	Accuracy	ROC AUC
KNN	0.55	0.70	0.65	0.66
Logistic Regression	0.64	0.80	0.74	0.75
Decision Tree	0.55	0.60	0.58	0.58
Random Forest	0.55	0.75	0.68	0.66
SVM	0.27	0.80	0.61	0.73

For the case of non-binary signal like EEG, a process of binarization is required. In this study we use the mean of the EEG signal as the threshold for its binarization. The LZC was then computed using the MATLAB code provided in [40].

2.3.5. Mutual information of multiple rhythms (MIMR)

The mutual information of multiple rhythms (MIMR) is a measure of the cross frequency coupling between multiple neural rhythms present in an EEG signal [41]. MIMR obtains a series of binary sequences extracted from the original EEG, where each sequence includes information from a particular rhythm (i.e., alpha, beta, gamma and theta oscillations). The information from all the series is combined into a new symbolic series, from which the delayed mutual information [42] is computed as an estimation of dependence between them. The more dependent the activity from the selected rhythms is, the more coupled the rhythms are and the higher the value of MIMR.

In order to calculate MIMR, the signal $s = \{s(1), s(2), ..., s(N)\}$ is smoothed using a median moving window with m window sizes $\{w_1, ..., w_m\}$. The binary sequences, of size $N - w_m - 1$, are obtained from the expression [41]:

$H_k[s_k(n)] = H[\text{median}(s_{k-1}(n - (w_{k-1}-1)/2), ..., s_{k-1}(n), ..., s_{k-1}(n + (w_{k-1}-1)/2)) - \text{median}(s_k(n - (w_k - 1)/2), ..., s_k(n), ..., s_k(n + (w_k - 1)/2))].$

where $(w_m - 1)/2$ are the total points from the sides that are not included in the smoothing, w_0 corresponds to the original signal $s(n)$; and $H[n]$ is the discrete Heaviside function:

$$H[n] = \begin{cases} 1 \, if \, n > 1 \\ 0 \, if \, n \leq 0 \end{cases}$$

All binary sequences H_k are transformed into the corresponding integer so that single values can identify all possible multiscale states. These integer values are used to construct the new signal $\{Y(n)\}$. Finally, MIMR is obtained as the delayed mutual information [42] of integers in $\{Y(n)\}$:

$$\text{MIMR} = I(Y(n - \tau), Y(n))$$

where τ is estimated as the first value that is closer to zero in the autocorrelation function of $Y(n)$.

In this study, we use a homemade MATLAB code to compute MIMR. More information about that implementation can be found in [41].

2.3.6. Detrended fluctuation analysis (DFA)

Detrended fluctuation analysis (DFA) measures the statistical self-affinity of a signal by detecting the long-range correlations embedded in it [43]. In order to calculate DFA, the signal $s = \{s(1), s(2), ..., s(N)\}$ is divided into N/l non-overlapping boxes s_l of size l. For each box, a cumulative sum S_l is calculated as $S_l = \sum_{i=1}^{l}(s_i - \bar{s})$, where $\bar{s}$ is the mean value of the signal s. Then the local trend Y_l of each box is computed as the ordinate of the least squares straight-line fit for the set of time windows of size n in S_l. Next, the fluctuation of each trend is calculated as:
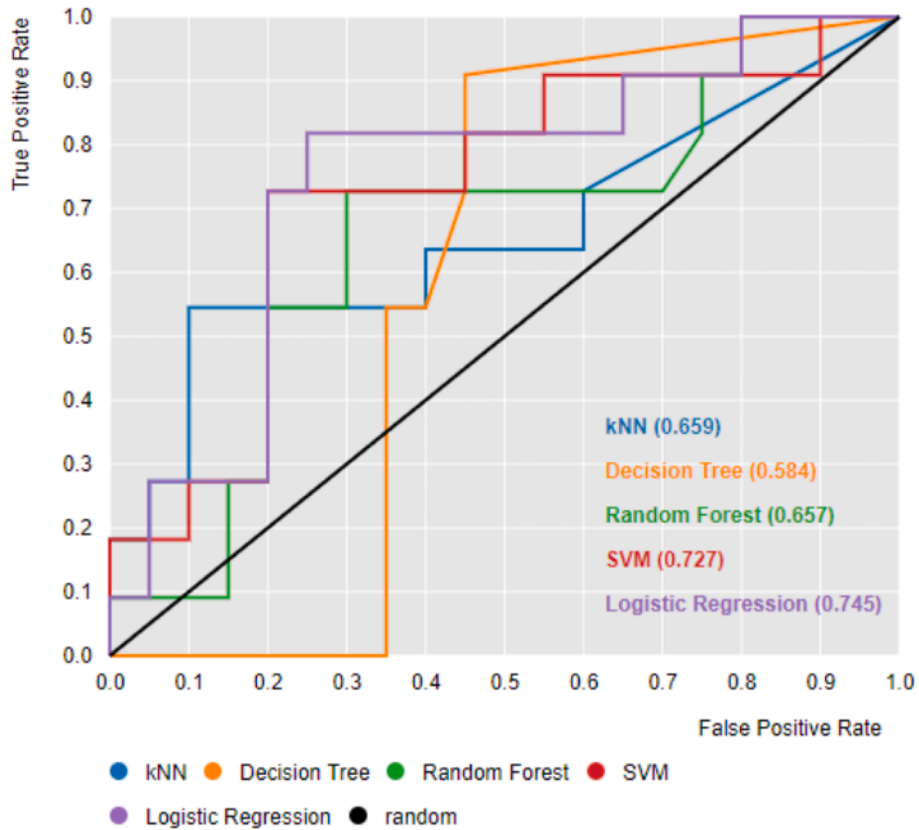
Fig. 6. ROC curve analysis evaluating the performance of classification algorithms using cognitive measures VLi, VF, WM and PS as input features. AUC values are shown inside parentheses.

Table 6
Performance scores of schizophrenia classification based on features extracted from EEG signals in recent studies. N/A: not available.

Reference	Method	Sensitivity	Specificity	Accuracy	ROC AUC
Present work	KNN	0.82	0.90	0.87	0.86
Present work	SVM	0.63	0.90	0.81	0.89
[24]	Deep Neural Networks	N/A	N/A	0.99	N/A
[23]	SVM	N/A	N/A	0.92	N/A
[22]	SVM	0.93	0.92	0.92	N/A
[18]	RF	N/A	N/A	0.81	N/A
[21]	SVM	0.85	0.67	0.79	0.76
[19]	Deep Neural Networks	0.62	0.75	0.73	0.77
[20]	RF	0.60	1.00	0.71	N/A

$$F(n) = \sqrt{\frac{\sum_{l=1}^{N}(S_l - Y_l)^2}{N}}$$

The previous process is repeated for a range of different window lengths *n*, and the slope of the log–log plot of $F(n)$ versus *n* is the scaling coefficient α, the result of the detrended fluctuation analysis.

We used the MATLAB code provided by Magris in [44] in order to perform the detrended fluctuation analysis of our EEG data. The window lengths ranged from 20 to 500 EEG samples.

2.3.7. Hjorth activity, mobility and complexity (HA, HM and HC)

Hjorth introduced three parameters, namely activity, mobility and complexity, to characterize EEG signals in the time domain [45].

Activity is computed as the standard deviation of the signal. The mobility parameter is defined as the square root of the standard deviation of the first derivative of the signal divided by the standard deviation of the signal. Finally, Hjorth complexity can be calculated as the ratio between the mobility of the first derivative of the signal and the mobility of the signal itself. These three parameters characterize the EEG signal in terms of amplitude, time scale and complexity, respectively.

In order to compute HA, HM and HC from our EEG data we used the MATLAB code provided by Too in [46].

2.3.8. Power of delta, theta, alpha and beta bands (DP, TP, AP and BP) and total energy (TE)

Power spectral analysis is a conventional method in EEG analyses. It is based on decomposing the signal, using Fourier transform, into functionally distinct frequency bands: delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–12 Hz) and beta (12–30 Hz). Then an estimation of the power spectral density of the signal is performed, from which the average band power is computed for each frequency band (DP, TP, AP and BP). The total energy (TE) is easily computed as the sum of the squared amplitude of the signal in each time sample.

In this study, DP, TP, AP, BP and TE measures for our EEG data were computed in MATLAB using the function *bandpower*.

2.3.9. Largest Lyapunov exponent (LLE)

The Lyapunov exponents of a dynamical system are the average exponential rates of divergence or convergence of infinitesimally close orbits in phase space [47]. A system with at least one positive Lyapunov exponent is chaotic. The value of the exponent indicates the time scale when the dynamical system becomes unpredictable. The larger the exponent, the more chaotic the system.

For a discrete time series, the largest Lyapunov exponent (LLE) can be computed from the constructed attractor obtained by means of phase
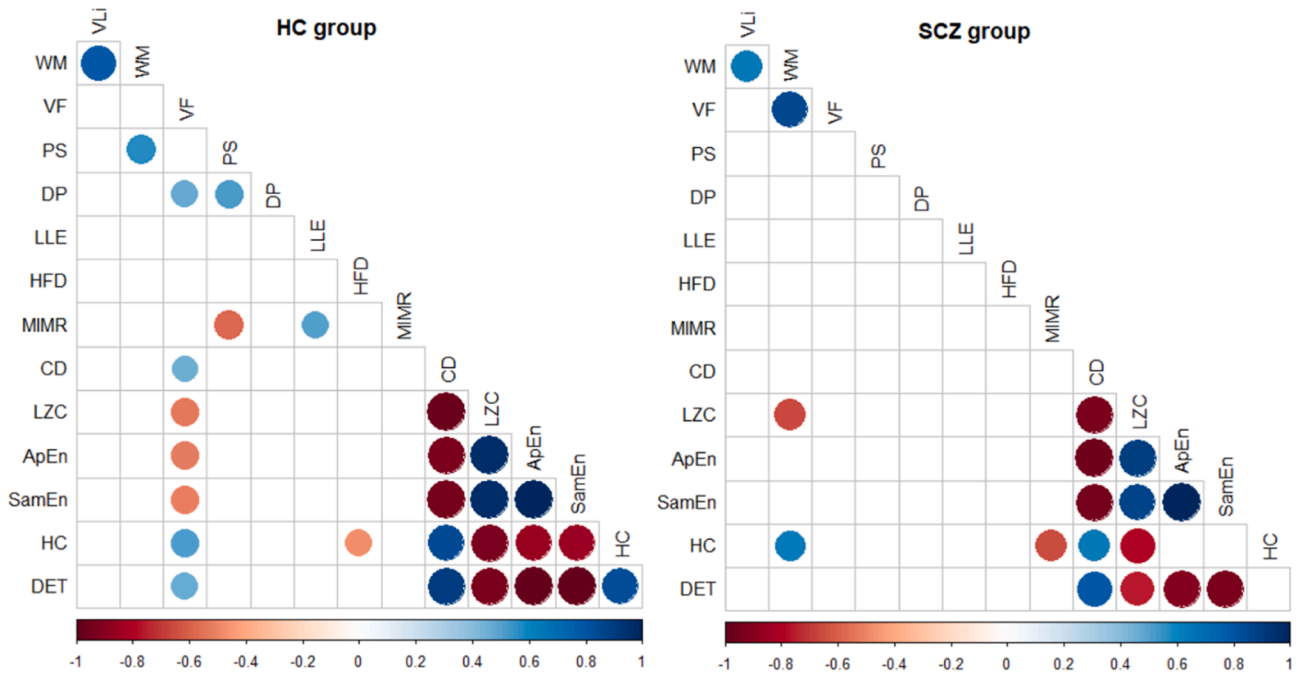
Fig. 7. Significant Spearman correlations between SCIP-S cognitive measures and EEG extracted features in the HC and SCZ groups. Results were corrected for multiple comparisons using the false discovery rate method. Only statistically significant correlations ($p < 0.05$) are shown. VLi: immediate verbal learning, WM: working memory, VF: verbal fluency, PS: the information processing speed, DP: Delta power, LLE: largest Lyapunov exponent, HFD: Higuchi fractal dimension, MIMR: mutual information of multiple rhythms, CD: correlation dimension, LZC: Lempel-Ziv complexity, ApEn: approximate entropy, SamEn: sample entropy, HC: Hjorth complexity, and DET: determinism in Recurrence quantification analysis.

space reconstruction with delay coordinates [47].

We used the MATLAB code provided by Wolf [48] to compute the LLE values for our EEG data.

2.3.10. Recurrence quantification analysis (DET)

Recurrence quantification analysis (RQA) is a useful technique for analyzing the non-linear dynamical behaviour of EEG data [49]. This technique is based on recurrence plots, which describe the recurrence property of a deterministic dynamical system by visualizing in a binary square matrix the time-dependent behaviour of orbits in phase space. Recurrence plots can reveal large-scale and small-scale patterns of a dynamical system, such as EEG [50]. One measure of RQA is the determinism measure (DET), based on the diagonal lines of the recurrence plot. DET is defined as the ratio of recurrence points on the diagonals of the recurrence plot to all recurrence points. This measure was introduced as a determinism measure of the system, where the higher the DET value the more predictable the system is.

We computed the recurrence plots of our EEG data and their DET values by using the MATLAB code provided by Ouyang in [51].

2.4. Machine learning classification

In this section, we briefly describe the main features of the machine learning algorithms that we used, and how we configured them using the software tool KNIME [52].

2.4.1. k-nearest neighbours (kNN)

kNN is a lazy learning algorithm where the data model is simply the training data. The basic idea in classifying the test data is to find similar instances inside the training data. Similarity between data instances is computed through distance functions such as the Euclidean distance. In order to classify a new test instance, the distance from it to all the instances in the training data must be computed. Then, the most frequent class of the k nearest training instances (neighbours) is selected as the class of the new test instance.

The parameter k is the key in kNN. In this study, we set k to a value of 3. Additionally, neighbors were weighted by distance, meaning that closer neighbors have greater influence on the resulting class than those farther away.

2.4.2. Logistic regression

Logistic regression is a classification method based on approximating the output function.

$$y(w, x) = \frac{1}{1 + e^{w_0 + x_1 w_1 + \cdots + x_n w_n}}$$

where $w = (w_1, \ldots, w_n)$ are the coefficients (w_0 is the independent term) to be fitted based on the values of each training instance x. In order to use logistic regression for classification, the output of the function y is restricted to the values 0 and 1. The w coefficients are computed using a solver. In this study, we configured KNIME to use the stochastic average gradient solver because it performs well regardless of both the number of instances in the training data and the number of features of each instance [53].

2.4.3. Decision trees

A decision tree is a classifier based on simple conditions or decision rules with the format IF-THEN-ELSE. These conditions are hierarchically chained within a tree-based structure in order to obtain the final decision [54]. Leaves of a decision tree are the output classes of the classifier. Each internal node of the tree contains a simple condition to check for a tuple < feature, value >. Children of a node are the possible results of the condition.

In the training step, the decision tree is constructed looking for a disjoint separation of the possible output classes in each node. The splitting criteria for each node is usually based on two measures: information gain and inequality (Gini index). For each node, the combination feature/value that minimizes the information gain or maximizes the Gini index is selected.

In this study, we configured KNIME by setting the Gini index as the

quality measure of the node. Additionally, we used a pruning method to trim the tree in a post-processing step by replacing each node with its most popular class if the prediction accuracy did not decrease.

2.4.4. Random forest

Random forest is a machine learning classifier that combines the results of several decision trees to produce an only output. Random forest is based on the concept of *bagging*, where several subsets are generated from the original training data and then used to train several decision trees independently. Each training data subset of size N is created selecting, randomly and with replacement, N instances from the original training data of size N. Each decision tree constructs a different model with an independent output class for each instance of the test data, and the final output of the random forest classifier corresponds to the most frequent class obtained.

In this study, the random forest model configured in KNIME consisted of 100 decision trees where the Gini index was also used as the quality measure in each tree node.

2.4.5. Support vector machines (SVM)

The goal of Support Vector Machines (SVM) is to find a hyperplane that separates the instances of two different classes. A hyperplane is a function of the form.

$$y(w, x) = w_0 + x_1 w_1 + \cdots + x_n w_n$$

where $w = (w_1, \ldots, w_n)$ are the coefficients (w_0 is the independent term) to be fitted based on the values of the features of each training instance x. From all possible hyperplanes dividing the training instances into two classes, SVM computes the one that obtains the maximum margin. The margin of a hyperplane is the maximum distance between border instances, and the training instances closest to the hyperplane are called the support vector (see Fig. 1).

Usually, a hyperplane is not a good solution when the two classes are not linearly separable. To solve this problem, SVM uses *kernel functions* that increment the dimensionality of the training data by adding more variables through non-linear transformations. Typical examples of kernel functions are the polynomial kernel (the hyperplane is transformed into a polynomial of arbitrary grade) and the radial base function kernel (the hyperplane is transformed into a circular-shape function).

In this study, we used in KNIME a polynomial kernel of the form $(\gamma < w, x > + r)^d$ with $d = 1$, $\gamma = 1.5$ and $r = 1$.

2.4.6. Classification pipeline

The pipeline proposed for computing and selecting the features in order to train the classifiers and test them using our schizophrenia EEG dataset is described graphically in Fig. 2.

Firstly, we computed all the measures for each channel of the EEG data. In order to accomplish that computation, each EEG epoch of 160 s was split into sliding windows of 2 s with 90 % of overlapping [55] (see Fig. 2.A). The measures were computed for each sliding window and then the median of all of them was considered as the representative value of each measure in each channel of the epoch. In this way, we obtained 527 values (17 measures × 31 channels) for each EEG epoch (see Fig. 2.B).

Using 527 different measures as features for each epoch in the machine learning models would not be a good strategy for a dataset of only 31 subjects. Moreover, such a large set of features could present additional problems, such as the appearance of confounding and correlated variables. So a dimensionality reduction process was performed in order to adjust the number of features to the size of our dataset. First, we selected those measures and channels where significant differences between groups (HC vs SCZ) were found in a *t*-test (see Table 2). In this way, the number of variables (pairs measure-channel) was reduced to 38. The scale of these variables is not the same, so a normalization

process to [-1, 1] was also applied in order to adequately combine and compare values from different measures (see Fig. 2.C).

Finally, we used principal component analysis (PCA) [56] to identify those variables which preserved the maximum variance in the data. Fig. 3 shows the percentage of variance explained by the first ten principal components. The first and second principal components explain 41 % and 19 % of the variance, respectively. However, the third principal component only explains 9 % of the variance, and this percentage decreases for subsequent principal components. Therefore, we finally selected principal components one and two as the features characterizing each subject.

The transformed data with 31 subjects (20 healthy controls and 11 schizophrenia patients) and two features (two principal components of 38 variables showing significant differences between groups) was then divided into train and test sets by using a leave-one-out cross validation [57] (see Fig. 2.C).

3. Results

Table 3 shows the classification results obtained for each subject using the five machine learning algorithms, configured in KNIME as previously described in Sections 2.4.1 to 2.4.5. The classification result (SCZ or HC) was assigned when the probability returned by the cross-validation process of the classification algorithm for that group was greater than 50 %. Three subjects (numbers five, ten and twenty-five) were misclassified with all algorithms. Fig. 4 shows a graphical representation of the dataset based on their two principal components. In this figure, subjects five, ten and twenty-five have been highlighted for a better understanding of their misclassification with all algorithms.

Table 4 shows the main scores obtained for each classification algorithm. The best accuracy was obtained for kNN. Nevertheless, SVM and Random Forest were the classification algorithms which obtained the best values for the AUC (area under curve) parameter of the ROC (Receiver Operating Characteristic) curve [58]. ROC curves of all classification algorithms are shown in Fig. 5.

As well as EEG data analysis, we performed an SCIP-S test on both schizophrenia patients and healthy controls (see Table 1), so we have four additional cognitive measures which also showed significant differences between groups. In order to assess the effectiveness of the proposed classification method based on complexity measures extracted from EEG data, we repeated the classification pipeline shown in Fig. 2.C, but using only those four cognitive measures as input features. The scores obtained for each classification algorithm are shown in Table 5 and Fig. 6.

Classification results using cognitive measures as features are clearly worse than those obtained for measures extracted from the EEG data (see Table 4 and Fig. 5). The best AUC value achieved was 0.75 LR classifier, while an AUC of up to 0.89 was obtained for SVM when EEG measures were used, as previously seen.

Table 6 shows, in a comparative way, the performance scores obtained in recent studies on machine learning classifiers for schizophrenia based on features extracted from EEG signals and those obtained with our methodology.

Finally, in order to analyze the relationship between cognitive and EEG measures we first calculated the average across the EEG channels for the variables in which we found significant differences between patients and healthy controls (see Table 2). We then calculated the Spearman correlation coefficient between these EEG variables and the four dimensions of the SCIP-S test (VLi, WM, VF and PS) which showed significant differences between groups (see Table 1). To control the type I error rate due to multiple comparisons, we used the false discovery rate approach [59]. After adjustment, only the correlations that were significant at a 95 % confidence level were considered. The results are presented in Fig. 7. No strong correlations were found between cognitive measures and EEG features in any group. VF showed weak correlations with CD, LZC, ApEn, SamEn, HC and DET in the HC group. For the SZC

group, WM correlated weakly with LZC and HC.

4. Discussion and conclusions

In the last few years, several studies have proposed using machine learning algorithms as a tool for distinguishing between patients with schizophrenia and healthy controls using EEG data. A wide range of measures have been used as features for those algorithms, including linear and non-linear measures. In the present study we have focused on providing a processing pipeline for computing and selecting those features which better differentiate schizophrenia patients from healthy controls. These features are then used as inputs in traditional machine learning techniques.

Our results show that complexity measures are the most effective measures in differentiating schizophrenia (see Table 2). HFD, MIMR, CD, LZC and especially LLE were the measures which showed significant differences in the highest number of channels. On the contrary, DFA, HA, TP, AP, BP and TE did not show significant differences in any channel. These findings are in line with recent studies, such as [22,23], which also showed the capability of complexity measures in distinguishing schizophrenia from healthy controls.

Channels FC6 and F8 are clearly the channels which present significant differences between subjects in a high number of measures, followed by channels FT10 and T8 (see Table 2). This finding implies that, according to our results, the main differences in brain activity between patients and healthy subjects are located in a clearly delimited area of the right hemisphere: the zone covered by the opercular area and the temporal pole. Some other studies [60] also reported those differences for complexity measures in the right brain hemisphere, but not delimiting the affected zone with the precision of the present study.

According to ROC curve analysis based on the AUC parameter (see Fig. 5), SVM (0.89), random forest (0.87), logistic regression (0.86) and kNN (0.86) performed very well. Only the decision tree algorithm performed a little worse, with an AUC of 0.68. This result suggests that our processing pipeline for computing and selecting the features works very efficiently regardless of the machine learning algorithm used. We believe that this fact is relevant, since it indicates that a good selection of the features would guarantee a good classification rate for use in other new or more sophisticated classification algorithms. This was not so common in several previous studies, where good accuracy values were obtained but only for a certain technique [8,18,20,60], performing poorly with the other algorithms tested [19,23].

Cognitive deficits in all cognitive domains are a central and persistent feature of schizophrenia [61–64] and have traditionally been used as an endophenotype to diagnose the underlying pathogenesis [65]. For this reason, we compared the discriminative capacity of EEG-based measures with respect to measures of cognitive functioning (VLi, VF, WM and PS). This comparison revealed that features extracted from the EEG data allowed classifiers to perform better, as shown in Table 5 and Fig. 6. Several authors defend the hypothesis that cognitive impairment is a consequence of schizophrenia and, therefore, cognitive measures would be reliable criteria for its diagnosis [65,66]. Nevertheless, our results suggest that neurophysiological measures extracted from EEG data provide better features for classification algorithms. These results would be in line with research with schizophrenia patients, which shows that neurocognition only accounts for a moderate association with variance in functional outcome [67,68]. Furthermore, the analysis of correlations between cognitive scores and complexity measures (see Fig. 7) pointed in the direction that these two groups of measures are not related to each other, and provide different information. Only verbal fluency (VF) was weakly correlated with several complexity measures (CD, LZC, ApEn, SamEn, HC and DET) for the HC group. For the case of the SCZ group, only working memory (WM) correlated with LZC and HC, but this correlation was weak as well. Although we do not have an explanation for these results at the moment, it is interesting to note that different metrics are associated with different aspects in both groups;

and this should be further explored in future studies.

The main limitation of our study is not being able to provide the best classification accuracy among the state-of-the-art EEG schizophrenia classification tools. To the best of our knowledge, this rank is led by Sun et al [24] with their hybrid model, based on convolutional neural networks and long-short-term memory networks, that achieved an accuracy of 99 % in classifying schizophrenia (see Table 6). Nevertheless, we believe that our method, even while not performing as well as Sun's, shows a valuable benefit for most scientists: it is much simpler to reproduce. No image construction is needed and no complex neural network has to be trained. Our method relies on well-known EEG measures easily computable with public source code. These measures are used in standard machine learning algorithms, which are available and easily configurable in common statistical software tools.

Another limitation of our study is the relatively small size of the dataset, so caution should be observed in generalizing our results to other datasets where patients underwent different treatment.

CRediT authorship contribution statement

J. Ruiz de Miras: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing – original draft, Funding acquisition. **A.J. Ibáñez-Molina:** Methodology, Software, Investigation, Resources, Data curation, Writing – review & editing. **M. F. Soriano:** Methodology, Investigation, Resources, Data curation, Writing – review & editing. **S. Iglesias-Parro:** Methodology, Software, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

References

[1] J. McGrath, S. Saha, D. Chant, J. Welham, Schizophrenia: A Concise Overview of Incidence, Prevalence, and Mortality, Epidemiol. Rev. 30 (2008) 67–76, https://doi.org/10.1093/epirev/mxn001.

[2] T.H. McGlashan, Duration of untreated psychosis in first-episode schizophrenia: marker or determinant of course? Biol. Psychiatry. 46 (1999) 899–907, https://doi.org/10.1016/S0006-3223(99)00084-0.

[3] M. Nieuwenhuis, N.E.M. van Haren, H.E. Hulshoff Pol, W. Cahn, R.S. Kahn, H.G. Schnack, Classification of schizophrenia patients and healthy controls from structural MRI scans in two large independent samples, Neuroimage. 61 (2012) 606–612, https://doi.org/10.1016/j.neuroimage.2012.03.079.

[4] E. Alpaydin, Introduction to machine learning, 4th ed., MIT Press, 2020.

[5] S. Akbar, M. Hayat, M. Iqbal, M.A. Jan, iACP-GAEnsC: Evolutionary genetic algorithm based ensemble classification of anticancer peptides by utilizing hybrid feature space, Artif. Intell. Med. 79 (2017) 62–70, https://doi.org/10.1016/j.artmed.2017.06.008.

[6] F. Ali, S. Ahmed, Z.N.K. Swati, S. Akbar, DP-BINDER: machine learning model for prediction of DNA-binding proteins by fusing evolutionary and physicochemical information, J. Comput. Aided. Mol. Des. 33 (2019) 645–658, https://doi.org/10.1007/s10822-019-00207-x.

[7] A. Ahmad, S. Akbar, M. Tahir, M. Hayat, F. Ali, iAFPs-EnC-GA: Identifying antifungal peptides using sequential and evolutionary descriptors based multi-information fusion and ensemble learning approach, Chemom. Intell. Lab. Syst. 222 (2022), 104516, https://doi.org/10.1016/j.chemolab.2022.104516.

[8] S. Han, Y. Wang, W. Liao, X. Duan, J. Guo, Y. Yu, L. Ye, J. Li, X. Chen, H. Chen, The distinguishing intrinsic brain circuitry in treatment-naïve first-episode

schizophrenia: Ensemble learning classification, Neurocomputing. 365 (2019) 44–53.

[9] G.S. Chilla, L.Y. Yeow, Q.H. Chew, K. Sim, K.N.B. Prakash, Machine learning classification of schizophrenia patients and healthy controls using diverse neuroanatomical markers and Ensemble methods, Sci. Rep. 12 (2022) 2755, https://doi.org/10.1038/s41598-022-06651-4.

[10] W. Yassin, H. Nakatani, Y. Zhu, M. Kojima, K. Owada, H. Kuwabara, W. Gonoi, Y. Aoki, H. Takao, T. Natsubori, N. Iwashiro, K. Kasai, Y. Kano, O. Abe, H. Yamasue, S. Koike, Machine-learning classification using neuroimaging data in schizophrenia, autism, ultra-high risk and first-episode psychosis, Transl. Psychiatry. 10 (2020) 278, https://doi.org/10.1038/s41398-020-00965-5.

[11] Y. Zhu, H. Nakatani, W. Yassin, N. Maikusa, N. Okada, A. Kunimatsu, O. Abe, H. Kuwabara, H. Yamasue, K. Kasai, K. Okanoya, S. Koike, Application of a Machine Learning Algorithm for Structural Brain Images in Chronic Schizophrenia to Earlier Clinical Stages of Psychosis and Autism Spectrum Disorder: A Multiprotocol Imaging Dataset Study, Schizophr. Bull. 48 (2022) 563–574, https://doi.org/10.1093/schbul/sbac030.

[12] L.-H. Lee, C.-H. Chen, W.-C. Chang, P.-L. Lee, K.-K. Shyu, M.-H. Chen, J.-W. Hsu, Y.-M. Bai, T.-P. Su, P.-C. Tu, Evaluating the performance of machine learning models for automatic diagnosis of patients with schizophrenia based on a single site dataset of 440 participants, Eur. Psychiatry. 65 (2022) e1. https://doi.org/DOI: 10.1192/j.eurpsy.2021.2248.

[13] X.-L. Cai, D.-J. Xie, K.H. Madsen, Y.-M. Wang, S.A. Bögemann, E.F.C. Cheung, A. Møller, R.C.K. Chan, Generalizability of machine learning for classification of schizophrenia based on resting-state functional MRI data, Hum. Brain Mapp. 41 (2020) 172–184, https://doi.org/10.1002/hbm.24797.

[14] G. Li, D. Han, C. Wang, W. Hu, V.D. Calhoun, Y.-P. Wang, Application of deep canonically correlated sparse autoencoder for the classification of schizophrenia, Comput. Methods Programs Biomed. 183 (2020), 105073, https://doi.org/10.1016/j.cmpb.2019.105073.

[15] Y.J. Bae, M. Shim, W.H. Lee, Schizophrenia Detection Using Machine Learning Approach from Social Media Content, Sensors. 21 (17) (2021) 5924.

[16] T.V. Rampisela, Z. Rustam, Classification of Schizophrenia Data Using Support Vector Machine (SVM), J. Phys. Conf. Ser. 1108 (2018) 12044, https://doi.org/10.1088/1742-6596/1108/1/012044.

[17] J.A. Cortes-Briones, N.I. Tapia-Rivas, D.C. D'Souza, P.A. Estevez, Going deep into schizophrenia with artificial intelligence, Schizophr. Res. 245 (2022) 122–140, https://doi.org/10.1016/j.schres.2021.05.018.

[18] L. Zhang, EEG Signals Classification Using Machine Learning for The Identification and Diagnosis of Schizophrenia, in: 2019 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., 2019: pp. 4521–4524. https://doi.org/10.1109/EMBC.2019.8857946.

[19] D. Ahmedt-Aristizabal, T. Fernando, S. Denman, J.E. Robinson, S. Sridharan, P. J. Johnston, K.R. Laurens, C. Fookes, Identification of Children at Risk of Schizophrenia via Deep Learning and EEG Responses, IEEE J. Biomed. Heal. Informatics. 25 (2021) 69–76, https://doi.org/10.1109/JBHI.2020.2984238.

[20] R. Buettner, M. Hirschmiller, K. Schlosser, M. Rossle, M. Fernandes, I.J. Timm, High-performance exclusion of schizophrenia using a novel machine learning method on EEG data, in: 2019 IEEE Int. Conf. E-Health Networking, Appl. Serv. Heal. 2019. (2019). https://doi.org/10.1109/HEALTHCOM46333.2019.9009437.

[21] S.K. Tikka, B.K. Singh, S.H. Nizamie, S. Garg, S. Mandal, K. Thakur, L.K. Singh, Artificial intelligence-based classification of schizophrenia: A high density electroencephalographic and support vector machine study, Indian J. Psychiatry. 62 (2020) 273–282, https://doi.org/10.4103/psychiatry.IndianJPsychiatry_91_20.

[22] V. Jahmunah, S. Lih Oh, V. Rajinikanth, E.J. Ciaccio, K. Hao Cheong, N. Arunkumar, U.R. Acharya, Automated detection of schizophrenia using nonlinear signal processing methods, Artif. Intell. Med. 100 (2019), 101698, https://doi.org/10.1016/J.ARTMED.2019.07.006.

[23] S.K. Prabhakar, H. Rajaguru, S.-H. Kim, A. Dourado, Schizophrenia EEG Signal Classification Based on Swarm Intelligence Computing, Comput. Intell. Neurosci. 2020 (2020) 1–14.

[24] J. Sun, R. Cao, M. Zhou, W. Hussain, B. Wang, J. Xue, J. Xiang, A hybrid deep neural network for classification of schizophrenia using EEG Data, Sci. Reports 2021 111. 11 (2021) 1–16. https://doi.org/10.1038/s41598-021-83350-6.

[25] Ó. Pino, G. Guilera, J.E. Rojo, J. Gómez, S.E. Purdon, SCIP-S, Screening del Deterioro Cognitivo en Psiquiatría, TEA Ediciones, 2014.

[26] M.J. Cuesta, O. Pino, G. Guilera, J.E. Rojo, J. Gómez-Benito, S.E. Purdon, M. Franco, A. Martínez-Arán, N. Segarra, R. Tabarés-Seisdedos, E. Vieta, M. Bernardo, B. Crespo-Facorro, F. Mesa, J. Rejas, Brief cognitive assessment instruments in schizophrenia and bipolar patients, and healthy control subjects: A comparison study between the Brief Cognitive Assessment Tool for Schizophrenia (B-CATS) and the Screen for Cognitive Impairment in Psychiatry (SCIP), Schizophr. Res. 130 (2011) 137–142, https://doi.org/10.1016/j.schres.2011.05.020.

[27] B.B. Mandelbrot, The Fractal Geometry of Nature, 1983. https://doi.org/10.1119/1.13295.

[28] T. Higuchi, Approach to an irregular time series on the basis of the fractal theory, Phys. D Nonlinear Phenom. 31 (1988) 277–283, https://doi.org/10.1016/0167-2789(88)90081-4.

[29] S. Selvam, T. Nadu, Complete Higuchi Fractal Dimension Algorithm, (2013). https://www.mathworks.com/matlabcentral/fileexchange/30119-complete-higuchi-fractal-dimension-algorithm.

[30] C. Gómez, Á. Mediavilla, R. Hornero, D. Abásolo, A. Fernández, Use of the Higuchi's fractal dimension for the analysis of MEG recordings from Alzheimer's disease patients, Med. Eng. Phys. 31 (2009) 306–313, https://doi.org/10.1016/j.medengphy.2008.06.010.

[31] J.S. Richman, J.R. Moorman, Physiological time-series analysis using approximate entropy and sample entropy, Am. J. Physiol. Circ. Physiol. 278 (2000) H2039–H2049, https://doi.org/10.1152/ajpheart.2000.278.6.H2039.

[32] S. Pincus, Approximate entropy (ApEn) as a complexity measure, Chaos An Interdiscip, J. Nonlinear Sci. 5 (1995) 110–117, https://doi.org/10.1063/1.166092.

[33] K. Lee, Fast Approximate Entropy, MATLAB Cent. (2012). https://es.mathworks.com/matlabcentral/fileexchange/32427-fast-approximate-entropy.

[34] V. Martínez, Sample Entropy, MATLAB Cent. (2018). https://es.mathworks.com/matlabcentral/fileexchange/69381-sample-entropy.

[35] P. Grassberger, Generalized dimensions of strange attractors, Phys. Lett. A. 97 (1983) 227–230, https://doi.org/10.1016/0375-9601(83)90753-3.

[36] P. Grassberger, I. Procaccia, Measuring the strangeness of strange attractors, Phys. D Nonlinear Phenom. 9 (1983) 189–208, https://doi.org/10.1016/0167-2789(83)90298-1.

[37] M. Small, Correlation dimension, 2002. https://github.com/m-small/dimension-matlab/tree/master/dimension.

[38] A. Lempel, J. Ziv, On the Complexity of Finite Sequences, IEEE Trans. Inf. Theory. 22 (1976) 75–81, https://doi.org/10.1109/TIT.1976.1055501.

[39] F. Kaspar, H. Schuster, Easily calculable measure for the complexity of spatiotemporal patterns, Phys. Rev. A, Gen. Phys. 36 (1987) 842–848, https://doi.org/10.1103/PHYSREVA.36.842.

[40] S. Faul, Kolmogorov algorithm for Lempel-Ziv complexity, MATLAB Cent. (2005). https://es.mathworks.com/matlabcentral/fileexchange/6886-kolmogorov-complexity.

[41] A.J. Ibáñez-Molina, M.F. Soriano, S. Iglesias-Parro, Mutual Information of Multiple Rhythms for EEG Signals, Front. Neurosci. (2020) 1312, https://doi.org/10.3389/FNINS.2020.574796.

[42] A.M. Fraser, H.L. Swinney, Independent coordinates for strange attractors from mutual information, Phys. Rev. A. 33 (2) (1986) 1134–1140.

[43] C.-K. Peng, S.V. Buldyrev, S. Havlin, M. Simons, H.E. Stanley, A.L. Goldberger, Mosaic organization of DNA nucleotides, Phys. Rev. E. 49 (2) (1994) 1685–1689.

[44] M. Magris, Detrended fluctuation analysis, MATLAB Cent. (2018). https://es.mathworks.com/matlabcentral/fileexchange/67889-detrended-fluctuation-analysis-dfa.

[45] B. Hjorth, EEG analysis based on time domain properties, Electroencephalogr. Clin. Neurophysiol. 29 (1970) 306–310, https://doi.org/10.1016/0013-4694(70)90143-4.

[46] J. Too, EEG Feature Extraction Toolbox, GitHub. (2020). https://github.com/JingweiToo/EEG-Feature-Extraction-Toolbox.

[47] A. Wolf, J.B. Swift, H.L. Swinney, J.A. Vastano, Determining Lyapunov exponents from a time series, Phys. D Nonlinear Phenom. 16 (1985) 285–317, https://doi.org/10.1016/0167-2789(85)90011-9.

[48] A. Wolf, Wolf Lyapunov exponent estimation from a time series, MATLAB Cent. (2019). https://es.mathworks.com/matlabcentral/fileexchange/48084-wolf-lyapunov-exponent-estimation-from-a-time-series.

[49] C.L. Webber, J.P. Zbilut, Dynamical assessment of physiological systems and states using recurrence plot strategies, J. Appl. Physiol. 76 (2) (1994) 965–973.

[50] G. Ouyang, X. Li, C. Dang, D. Richards, Using recurrence plot for determinism analysis of EEG recordings in genetic absence epilepsy rats, Clin. Neurophysiol. 119 (2008) 1747–1755, https://doi.org/10.1016/J.CLINPH.2008.04.005.

[51] G. Ouyang, Recurrence quantification analysis, MATLAB Cent. (2014). https://es.mathworks.com/matlabcentral/fileexchange/46765-recurrence-quantification-analysis-rqa.

[52] M.R. Berthold, N. Cebron, F. Dill, T.R. Gabriel, T. Kötter, T. Meinl, P. Ohl, C. Sieb, K. Thiel, B. Wiswedel, KNIME: The Konstanz Information Miner, Stud. Classif. Data Anal. Knowl. Organ. (2008) 319–326. https://doi.org/10.1007/978-3-540-78246-9_38.

[53] M. Schmidt, N. Le Roux, F. Bach, Minimizing finite sums with the stochastic average gradient, Math. Program. 2016 1621. 162 (2016) 83–112. https://doi.org/10.1007/S10107-016-1030-6.

[54] L. Rokach, O. Maimon, Data Mining with Decision Trees, 2nd ed., WORLD SCIENTIFIC, 2014. https://doi.org/10.1142/9097.

[55] Y.-J. Liu, M. Yu, G. Zhao, J. Song, Y. Ge, Y. Shi, Real-Time Movie-Induced Discrete Emotion Recognition from EEG Signals, IEEE Trans. Affect. Comput. 9 (2018) 550–562, https://doi.org/10.1109/TAFFC.2017.2660485.

[56] I.T. Jolliffe, Principal Component Analysis, 2nd ed., Springer-Verlag, New York, 2002. https://doi.org/10.1007/B98835.

[57] D.W. Gareth James Trevor Hastie, Robert Tibshirani, An introduction to statistical learning : with applications in R, 2nd ed., Springer, New York, 2021. https://search.library.wisc.edu/catalog/9910207152902121.

[58] J.A. Hanley, B.J. McNeil, The meaning and use of the area under a receiver operating characteristic (ROC) curve, Radiology. 143 (1982) 29–36, https://doi.org/10.1148/radiology.143.1.7063747.

[59] Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing, J. R. Stat. Soc. 57 (1995) 289–300, https://doi.org/10.2307/2346101.

[60] A. Goshvarpour, A. Goshvarpour, Schizophrenia diagnosis using innovative EEG feature-level fusion schemes, Phys. Eng. Sci. Med. 43 (2020) 227–238, https://doi.org/10.1007/s13246-019-00839-1.

[61] B.R. Rund, A Review of Longitudinal Studies of Cognitive Functions in Schizophrenia Patients, Schizophr. Bull. 24 (1998) 425–435, https://doi.org/10.1093/oxfordjournals.schbul.a033337.

[62] N.F. Forbes, L.A. Carrick, A.M. McIntosh, S.M. Lawrie, Working memory in schizophrenia: a meta-analysis, Psychol. Med. 39 (2009) 889–905, https://doi.org/10.1017/S0033291708004558.

[63] M.K. Manglam, A. Das, Verbal learning and memory and psychopathology in schizophrenia, Asian J. Psychiatr. 6 (2013) 417–420, https://doi.org/10.1016/j.ajp.2013.05.009.

[64] K.H. Nuechterlein, M.F. Green, M.E. Calkins, T.A. Greenwood, R.E. Gur, R.C. Gur, L.C. Lazzeroni, G.A. Light, A.D. Radant, L.J. Seidman, L.J. Siever, J.M. Silverman, J. Sprock, W.S. Stone, C.A. Sugar, N.R. Swerdlow, D.W. Tsuang, M.T. Tsuang, B. I. Turetsky, D.L. Braff, Attention/vigilance in schizophrenia: performance results from a large multi-site study of the Consortium on the Genetics of Schizophrenia (COGS), Schizophr. Res. 163 (2015) 38–46, https://doi.org/10.1016/j.schres.2015.01.017.

[65] S.K. Hill, J.L. Reilly, R.S.E. Keefe, J.M. Gold, J.R. Bishop, E.S. Gershon, C. A. Tamminga, G.D. Pearlson, M.S. Keshavan, J.A. Sweeney, Neuropsychological impairments in schizophrenia and psychotic bipolar disorder: findings from the Bipolar-Schizophrenia Network on Intermediate Phenotypes (B-SNIP) study, Am. J. Psychiatry. 170 (2013) 1275–1284, https://doi.org/10.1176/appi.ajp.2013.12101298.

[66] S.A. Chong, Inclusion of cognitive impairment in the DSM diagnosis of schizophrenia: if not now, when? World Psychiatry. 7 (2008) 37–38, https://doi.org/10.1002/J.2051-5545.2008.TB00150.X.

[67] M.F. Green, R.S. Kern, R.K. Heaton, Longitudinal studies of cognition and functional outcome in schizophrenia: implications for MATRICS, Schizophr. Res. 72 (2004) 41–51, https://doi.org/10.1016/j.schres.2004.09.009.

[68] S.J. Schmidt, D.R. Mueller, V. Roder, Social Cognition as a Mediator Variable Between Neurocognition and Functional Outcome in Schizophrenia: Empirical Review and New Results by Structural Equation Modeling, Schizophrenia Bull. 37 (suppl 2) (2011) S41–S54.