A new Machine Learning approach for epilepsy diagnostic based on Sample Entropy

Zayneb BRARI and Safya BELGHITH

RISC laboratory, École Nationale d'Ingénieurs de Tunis , Université de Tunis El Manar, BP 37, 1002, Le Bélvédère, 1002, Tunis, Tunisia (e-mail: brarizayneb@gmail.com and safya.belghith@enit.utm.tn).

Abstract: Irregularity is the main characteristic of electroencephalographic signals (EEG), which needs a specific analysis method for neurological disease diagnosis. An efficient tool for signal irregularity analysis is Sample Entropy (SampEn). In this context, our paper was elaborated. We used SampEn to design a Machine Learning model for brain state detection based on EEG signals, which allows to differentiate between healthy (H) subjects, epileptic subjects during seizures free intervals (E) and epileptic subjects during seizures (S). Two main novelties are presented in our paper. The first one is related to the outline of the designed machine learning model, signal derivatives are determined as preprocessing step, then extracted features are SampEn and Standard Deviation (STD) from EEG signals and its first and second derivatives. These features are firstly used to train a K-Nearest Neighbor classifier (KNN) and yield high accuracy. After that, we select the most relevant features and we design our proposed classifier that provides better accuracy. The second one is related to the performance of our model to overcome some crucial purposes. In addition to the highest achieved accuracy, 100% for seizure detection, 99.2% for epilepsy detection and 99.86% for three class classification cases, our model used few features and simple classifier which involves fast running time. That is why we can consider our model as a suitable tool for real time applications.

Keywords: Epilepsy, EEG, Features Selection, Machine Learning, Sample Entropy.

1. INTRODUCTION

Context : Epilepsy is a common brain disorder that affects about 50 million people around the world, it's characterized by abnormal activity of brain cells that provokes epileptic seizures, which can lead to dangerous situations. The main diagnostic test is EEG traces, thanks to its low cost, acquisition simplicity and significant results. These records are complicated traces that need an expert to investigate them to identify epilepsy, which is not possible in all cases and not to mention it is time consuming. In addition, epileptic seizures prevention can save patient lives, by protecting them before seizure attacks. To overcome these problems, automatic analysis of EEG signals is made in many published works, which compete to produce the most accurate machine learning model that detects patients brain states. In general, the machine learning model approach goes through four stages, Khosla et al. (2020). Firstly, *signal preprocessing*, is often used for noise removal, using signal decomposition or transformation of the signal into one or both time and frequency domain. Secondly, *features extraction*, is utilised for selecting the most relevant information from studied signals, using different approaches such as statistical, spectral, nonlinear, chaotic... Thirdly, *feature selection and dimensionality* reducing, are used for selecting the most relevant information and delete redundant ones and for reducing the dimension of study space by various methods like Fisher score, Kruskal Wallis test, Principal Component Analysis, Generalized Gauss Distribution... Finally, *classification* allows to detect brain state, using several of classifiers; K-Nearest Neighbours, Support Vector Machine, Decision Tree, Artificial Neural Network...

Problematic : The proposed models define two kinds of challenges. The first one is related to EEG signals nature, which is characterized by nonlinear, non stationary and irregular aspect. The second one is related to the real time applications of these models. These require, in addition to the high accuracy, fast run time, program simplicity and least possible storage memory.

Hypothesis : This paper has relied on two research findings. The first one exists in Richman and Moorman (2000) paper's, which gives the first definition of Sample Entropy and shows that it's a powerful tool for studying irregularity of physiological time series. The second one was defined in our previous work, Brari and Belghith (2020), where we have shown that EEG signals derivatives is a powerful tool for preprocessing step.

Methods : We propose a simple machine learning model composed of four stages. First, we determine EEG the first and the second derivatives in the preprocessing stage. Second, we extract Sample entropy from each one based on Richman and Moorman (2000) method, Sample entropy is calculated with threshold proportional to Standard Deviation of studied signal. Both extracted Sample entropy and Standard deviation are used to train K Nearest Neighbor classifier which gives an important accuracy. After that, we select the most relevant features to delete redundant and unused information. Then, selected features permit to classify signals by simple linear and Piecewise linear functions. In this paper, in addition to the efficiency of our model, we prove the importance of features selection in classification stage. It's basically efficient to reduce features space, so decrease used memory and run time. Moreover, in this work, it is put in use to increase classification accuracy .

Contributions: The main contributions of this paper are:

- A new machine learning model for epilepsy diagnosis with enhanced accuracy percentage in comparison to the state of the art works,
- Sample entropy extracted from EEG signals and its first and second derivatives are used in the features extraction stage which consist of an efficient way to characterize studied signals,
- Our model uses few features and a simple classification function which is beneficial in real applications.

Organization: This paper is organized as follows, section 2 describes the used database as well as other published work which uses the latter to develop an automatic model for epilepsy diagnosis. Our method is introduced in section 3. Experimental results are presented and discussed with other published works in section 4. Finally, conclusions and some perspectives are displayed in section 5.

2. PRELIMINARY

2.1 Used Database

In this paper, we used a publicly available database in the web site of Bonn University, described in Andrzejak et al. (2001). It contains five subsets A, B, C, D and E. A and B signals are respectively extracted from healthy subjects with open and closed eyes. C an D signals are respectively extracted from epileptic subjects during seizures free intervals from the hippocampal formation of brain hemisphere and the epileptogenic zone. E signals are extracted from an epileptic subjects during seizures. In our work we will use only four subsets A, B, C and E, to study different classification cases which are summarized in Table 1.

Table 1. EEG classification cases

cases	Description
case 1	H vs E
case 2	H vs S
case 3	E vs S
case 4	(H and E) vs S
case 5	H vs (E and S)
case 6	H vs E vs S

2.2 Literature review

Several published works have studied EEG signals using the Bonn database. In this paper, we focus on comparing our approach to other published ones that used Entropy, as features, with different methods. Kumar et al. (2010) divided EEG signal into 1 second, 2 second and 5 second samples, and extract three types of entropies, Wavelet Entropies (WE1, WE2, WE5), Sample Entropies (SampEn1, SampEn2, SampEn5) and spectral entropies (SpEn1, SpEn2, SpEn5) for each epoch. Then they classify signals using Recurrent Elman Network and Radial Basis network. Acharya et al. (2012) extract four entropies as features, Approximate Entropy, Sample Entropy, Phase Entropy 1 and Phase Entropy 2 and classify them by different classifiers, Fuzzy Sugeno Classifier, Support Vector Machine, K-Nearest Neighbour, Probabilistic Neural Network, Decision Tree, Gaussian Mixture Model and Naive Bayes Classifier. They proved that Fuzzy Classifier is the best classifier that gives the highest accuracy. Acharya et al. (2015) studied EEG signals with many types of Entropy, Approximate Entropy, Fuzzy Entropy, Sample Entropy, Renyi's Entropy, Spectral Entropy, Permutation Entropy, Wavelet Entropy, Tsallis entropy, Higher Order Spectra Entropies, Kolmogorov–Sinai entropy and Recurrence Quantification Analysis entropy. Nkengfack et al. (2021) extract, Alpha α and Gamma γ rhythms of EEG are using Jacobi Polynomial Transform, Discrete Legendre Transforms and discrete Chebychev Transform, then they extract different measures of complexity: Approximate Entropy, Fuzzy Entropy, Sample Entropy and Permutation Entropy. Extracted features are classified using Least-Square SVM. Previously our work, Brari and Belghith (2020), show with the variance extracted from EEG and its derivatives and transformed in redescription space, and using 3D log-log-log plot and a simple kernel trick, we classify signals by a linear classifier.

2.3 Sample Entropy

Sample entropy is a modified version of Approximate entropy. It was initially defined by Richman and Moorman (2000),based on Pincus and Ehrenkranz (1991) works, to study the irregularity of physiological signals. In fact, it examines the similarity of different signal epochs and attributes high value to irregular and random signals.

For time series, of length N, with a constant time interval τ given by :

$$X_N = (x_1, x_2, ...x_N) \tag{1}$$

We define m vectors $X_m(1)$, $X_m(2)$,... ,$X_m(N - m1)$:

$$X_m(j) = [x(j), x(j + 1), ..., x(j + m - 1)], \tag{2}$$

where $1 \leq j \leq (N - m + 1)$.
Sample Entropy is given by:

$$SampEn(\tau, r, m, N) = -log(\frac{A}{B}) \tag{3}$$

Where

- r is threshold measure,according to Richman and Moorman (2000) works, r is taken proportional to the standard deviation of the studied time series and m is the embedding dimension.
- B_i the probability that two sequences closer than r for m points, $B_i = d(X_m(i), X_m(j)) < r$. d is the measured distance. So $B_i^m = \frac{1}{N-m-1} B_i$.

$$B = \frac{1}{N - m} \sum_{j=1}^{N-m} B_i^m \tag{4}$$

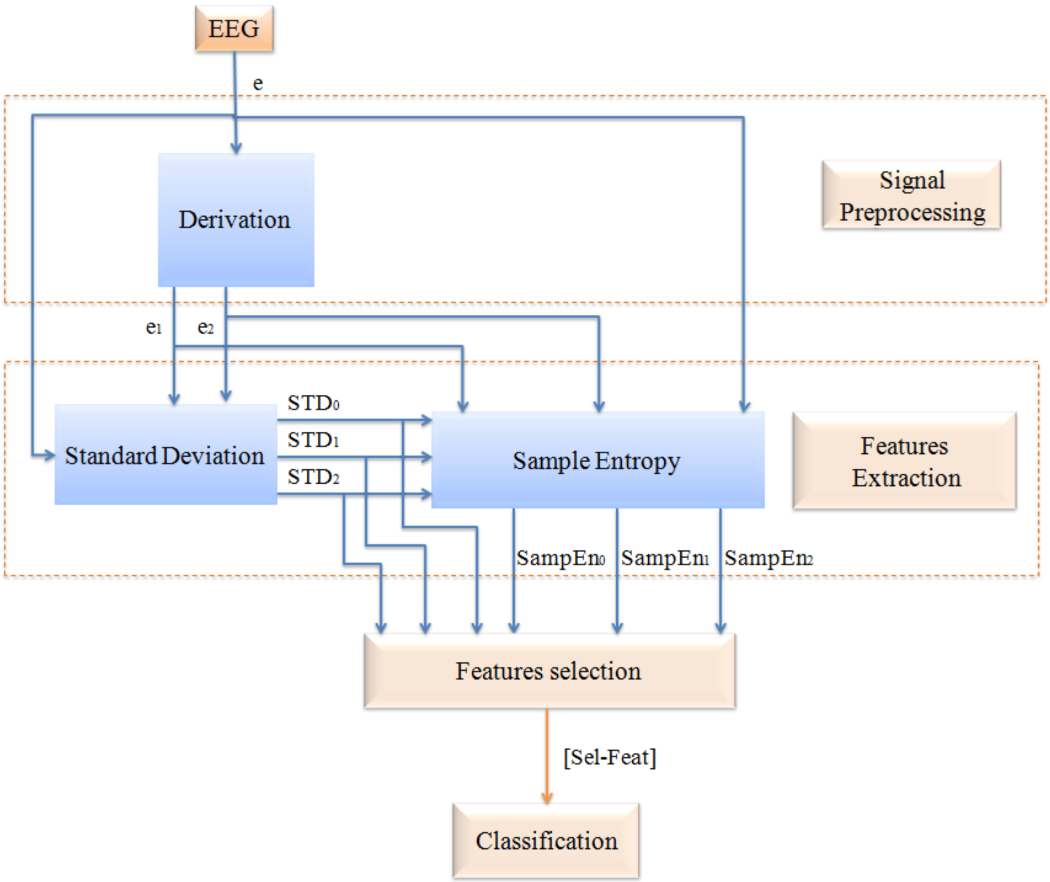

Fig. 1. Block diagram of the proposed model

- A_i the probability that two sequences closer than r for (m+1) points. In the same way, We calculate $A_i^m = \frac{1}{N-m-1} A_i$.

$$A = \frac{1}{N-m} \sum_{j=1}^{N-m} A_i^m \tag{5}$$

2.4 Features Selection

Feature selection is an interesting step that is employed to optimize learning. It consists in retrieving an optimal subset of features. So, it eliminates redundant and irrelevant information according to the fixed goals. Moreover, it improves classification accuracy and reduces the problem of overfitting. It also aims to reduce the dimension of the data space to simplify data representation, usually in 2D or 3 D space to improve data visualization. Furthermore, it is beneficial for the reduction of the storage space and the learning time, Stańczyk (2015). There are three types of selections:

- *Filter approaches*, feature selection is independent of used classification method.
- *Wrappers approaches*, use the learning algorithm as a criterion of performance evaluations.
- *Embedded approaches*, select features in parallel to the ranking process.

2.5 Classification

Machine learning models are often designed through two basic approaches: supervised learning and unsupervised learning. In our work, we are interested in supervised learning which can be done using different algorithms. Two types of classification algorithms will be used. The first one is K-nearest neighbor is a classifier based on the majority voting. The algorithm classifies the objects according to their nearest neighbors. The second one is Kernel classifier based on nonlinear kernel to linearize nonlinear problems and linear classifier for linear cases, during the learning phase the algorithm aims to look for the optimal hyperplane which separates the classes.

3. METHODOLOGY

In this paper, we propose a new machine learning model for the automatic detection of brain state, as shown in figure 1, our approach is mainly divided into four tasks.

- *Signal Preprocessing* : we calculate first, e1, and second, e2, derivatives from EEG signals,e.
- *Features Extraction* : From each signal, (e, e1, e2), we determine Sample Entropy based on Richman and Moorman (2000) algorithm. Sample Entropy is a logarithmic measure of the probability that two simulta-

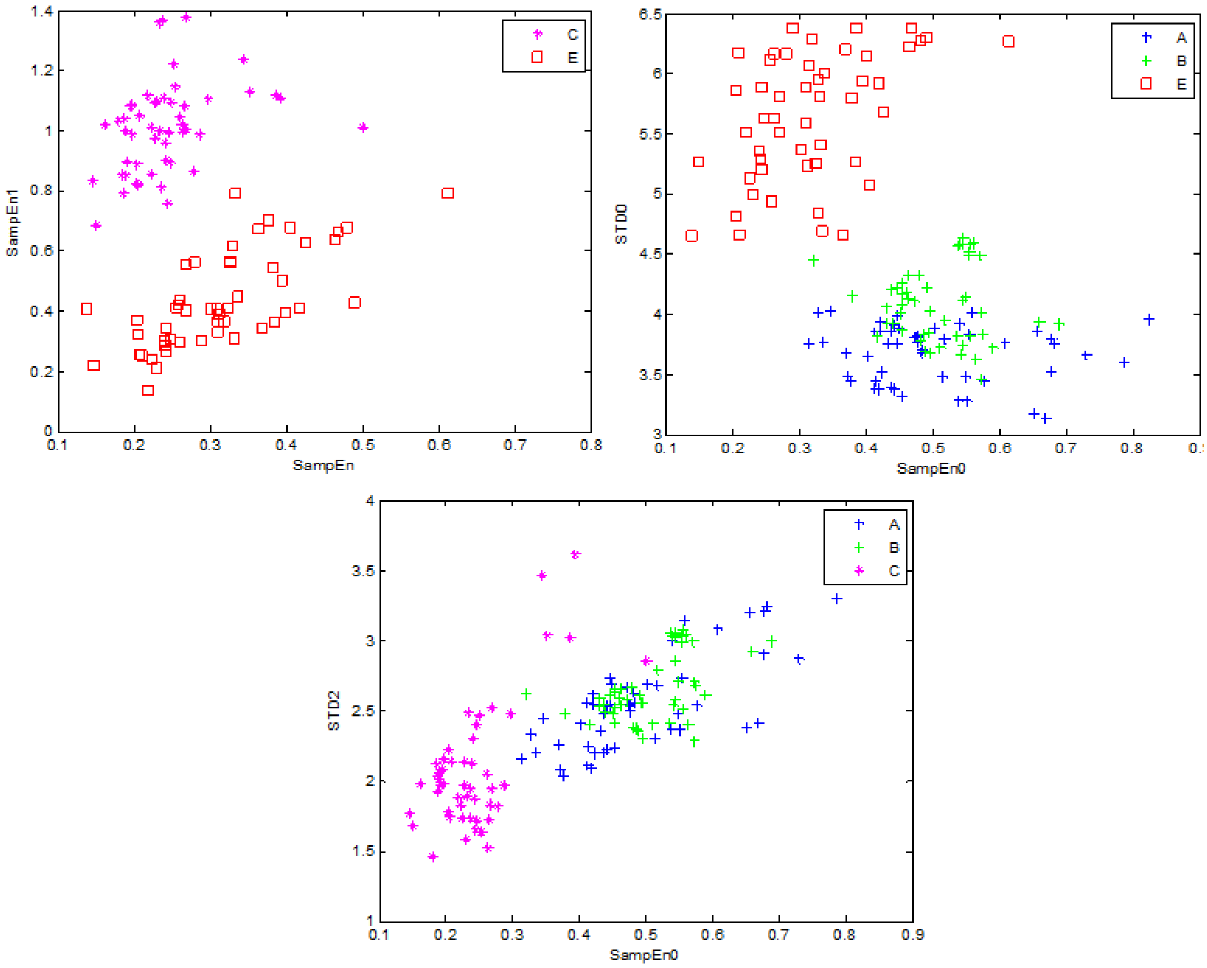

Fig. 2. Projections in 2D spaces of extracted features. (Healthy subjects:A/B, Epileptic subjects during seizures free intervals: C, Epileptic subjects during seizures: E)

neous states of length m are closer than a threshold r then simultaneous states of length m+1 are also closer than r. In Richman and Moorman (2000) work, in which the first definition of Sample entropy is given, r is taken proportional to the standard deviation of the studied signal.

- *Features Selection* : we select the useful features and delete the redundant and useless ones, using wrapper approach.
- *Classification* : we use selected features for the detection of different states and separate them by assigning a class label, based on training step.

4. EXPERIMENTAL RESULTS

In our work, we will take 50% of the features for the training step and the other 50% for the test step.

4.1 Classification using KNN

In the first stage, all extracted Features (all-feat), Sample Entropy and Standard Deviation extracted from (e, e1, e2), are tested by the K nearest neighbor classifier(KNN). In Table 2, we present the achieved accuracy by this method. KNN gives high accuracy, but it is slow because it is necessary to review all the examples of training each classification test,not to mention that it requires a large memory.
In the rest of this work, we will look for a new classifier that leads to overcome these problems and increase program accuracy.

4.2 Proposed Model (PM)

In the second stage, we will reduce features vector by selecting the most pertinent features, we start by studying the first, the second and the third classification cases, which are binary classifications using two types of signals. In Fig.2, we present a selection of projections in 2D space of extracted features, which is helpful to resolve binary classification problems using two types of signals.

- **Case 1 :** In subfigure 2(a), we show that case 1 is solved in 2D spaces composed of coordinates Sample entropy extracted from e and logarithm of Standard Deviation extracted from e2.
- **Case 2 :** In subfigure 2(b), case 2 is solved with linear function in 2D space with coordinates Sample Entropy and logarithm of standard deviation extracted from e.

Table 2. Classification Accuracy (%) for the different studies cases

Cases	Used sets	KNN(all-feat)	PM(sel-fet)
1	A-E	100	100
	B-E	98.96	100
	AB-E	99.02	100
2	A-C	98.58	99
	B-C	98.94	99.2
	AB-C	99	99.1
3	C-E	99.8	100
4	(A-C)-E	99.82	100
	(B-C)-E	98.82	100
	(AB-C)-E	99.32	100
5	A-(C-E)	99.02	99.33
	B-(C-E)	98.95	99.73
	AB-(C-E)	99.30	99.5
6	A-C-E	98.71	99.33
	B-C-E	98.69	99.86
	AB-C-E	99.30	99.5

- **Case 3 :** In subfigure 2(c), case 3 classification, we use 2D space with coordinates Sample entropy extracted from e and e1.

For binary classification using three types of signals and 3-classes classification cases, **case 4, case 5 and case6**, we form a new features space from selected features (Sel-Feat). Obtained results by our method (PM) are presented in Table 2.

The number of used features and computation time in seconds is given in Table 3. We show that our model uses a few features with fast run time.

Table 3. Performance of Proposed Model

Cases	Used sets	N-Feat	T(s)
1	AB-E	2	0.4
2	AB-C	2	0.36
3	C-E	2	0.75

5. DISCUSSION

In table 4, we summarize the achieved results by our method and other published works that used the same database and Sample Entropy. Two main issues are discussed in this section:
i) *State of the art methods :* Kumar et al. (2010) studied a binary classification problem, case 1, using two subset combinations (A-E, B-E), and 3-class classification problem (C-D-E). They prove the efficiency of the Recurrent Elman Network classifier compared to the Radial Basis network. In addition, they show that extract features employing Wavelet entropies is better than Sample Entropies and Spectral entropies, which yield to the best classification accuracies in most the cases. 99.35% in B-E case and 99.4 % in C-D-E case. Acharya et al. (2012) showed the efficiency of a Fuzzy Classifier applied to a feature vector based on four types of entropy (Approximate Entropy, Sample Entropy, Phase Entropy 1, Phase Entropy 2). An accuracy of 98.1% was achieved in case 6. Nkengfack et al. (2021) studied only binary classification problems. An accuracy of 100% was archived in detecting epileptic seizures using different subset combinations (A-E, B-E, C-E, D-E, AB-E, CD-E, ACD-E, ABCD-E). Besides, an accuracy of 97.5% for case 1 (AB-CD) and 95.00% for case5 (AB-CDE). They also studied eyes state (A-B) and achieve an accuracy of 88.75%.
ii) *Our methods :* In this work, Sample Entropy and Standard Deviation were extracted from EEG signals and their derivatives yield to resolve many classification cases using few features and by simple classifier. A highest accuracy value, 100%, was achieved in case 1 (A-E, B-E, AB-E), case 3 (C-E) and case 4 (AC-E, BC-E, AB-CE). Moreover, the best classification accuracy for case 2 is 99.2%, using B-E, for case 5 is 99.73%, using B-CE, and for case 6 is 99.86%, using B-C-E. Through our model, we exceed all published work in term of the achieved accuracy and the number of studied cases on one side and in term of program simplicity and used features on the other side.

6. CONCLUSION AND PERSPECTIVES

In this study, a new approach for machine learning model building is proposed. Our main contribution consists in extracting Sample entropy and Standard Deviation from EEG signals and its first and second derivatives, which are investigated to propose an automatic system for epilepsy diagnosis. The proposed approach exceeds other published work in terms of accuracy, a high value is achieved in all cases, and in terms of simplicity, we use an optimal feature vector with small size and simple classifier. We prove that Sample Entropy is an efficient tool for brain signals characterization in epilepsy monitoring applications. Our future direction will concern the application of the proposed approach in real time frameworks and its application to detect other brain pathology to value our contributions

REFERENCES

Acharya, U.R., Fujita, H., Sudarshan, V.K., Bhat, S., and Koh, J.E. (2015). Application of entropies for automated diagnosis of epilepsy using eeg signals: A review. *Knowledge-based systems*, 88, 85–96.

Acharya, U.R., Molinari, F., Sree, S.V., Chattopadhyay, S., Ng, K.H., and Suri, J.S. (2012). Automated diagnosis of epileptic eeg using entropies. *Biomedical Signal Processing and Control*, 7(4), 401–408.

Andrzejak, R.G., Lehnertz, K., Mormann, F., Rieke, C., David, P., and Elger, C.E. (2001). Indications of nonlinear deterministic and finite-dimensional structures in

Table 4. Comparison between achieved results by our method and others used the same database

Authers	Methods	Used sets	Accuracy(%)
Kumar et al. (2010)	Recurrent Elman Network	A-E	99.75
	Wavelet Entropy (WE1, WE2, WE5)	B-E	99.35
		C-D-E	99.4
	Sample Entropy (SampEn1, SampEn2, SampEn5)	A-E	98.65
		B-E	93.5
		C-D-E	88.3
	Spectral Entropy (SpEn1, SpEn2, SpEn5)	A-E	99.8
		B-E	98.8
		C-D-E	79
	Radial Basis network	A-E	97.5
	Wavelet Entropy(WE1, WE2, WE5)	B-E	96.25
		C-D-E	90
		A-E	96.25
	Sample Entropy (SampEn1, SampEn2, SampEn5)	B-E	90.5
		C-D-E	86.25
		A-E	96
	Spectral Entropy (SpEn1, SpEn2, SpEn5)	B-E	95.25
		C-D-E	88.5
Acharya et al. (2012)	Approximate Entropy, Sample Entropy, Phase Entropy 1 and Phase Entropy 2		
	Fuzzy Classifier	case 6	98.1
Nkengfack et al. (2021)	Jacobi polynomial transform	A-E	100
	Discrete Legendre transforms	B-E	100
	Chebychev transform	C-E	100
	Approximate Entropy	D-E	100
	Fuzzy Entropy	AB-E	100
	Sample Entropy	CD-E	100
	Permutation Entropy	ACD-E	100
	Least-Square-Support Vector Machine	ABCD-E	100
		AB-CDE	95
		AB-CD	97.50
		A-B	88.75
This work	Derivation	A-E	100
	Sample Entropy and Standard Deviation	B-E	100
	Linear and piecewise linear classifiction	AB-E	100
		A-C	99
		B-C	99.2
		AB-C	99.1
		C-E	100
		AC-E	100
		BC-E	100
		ABC-E	100
		A-CE	99.33
		B-CE	99.73
		AB-CE	99.5
		A-C-E	99.33
		B-C-E	99.86
		AB-C-E	99.5

time series of brain electrical activity: Dependence on recording region and brain state. *Physical Review E*, 64(6), 061907.

Brari, Z. and Belghith, S. (2020). A new method for the detection of epilepsy and epileptic seizures based on the variance of eeg signals and its derivatives with a simple kernel trick. In *2020 4th International Conference on Advanced Systems and Emergent Technologies (IC_ASET)*, 261–265. IEEE.

Khosla, A., Khandnor, P., and Chand, T. (2020). A comparative analysis of signal processing and classification methods for different applications based on eeg signals. *Biocybernetics and Biomedical Engineering*, 40(2), 649–690.

Kumar, S.P., Sriraam, N., Benakop, P., and Jinaga, B. (2010). Entropies based detection of epileptic seizures with artificial neural network classifiers. *Expert Systems with Applications*, 37(4), 3284–3291.

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, 100536.

Pincus, Steven M., I.M.G. and Ehrenkranz, R.A. (1991). A regularity statistic for medical data analysis. *Journal of clinical monitoring*, 7(4), 335–345.

Richman, J.S. and Moorman, J.R. (2000). Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology-Heart and Circulatory Physiology.*

Stańczyk, U. (2015). Feature evaluation by filter, wrapper, and embedded approaches. In *Feature Selection for Data and Pattern Recognition*, 29–44. Springer.