IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Traitement du Signal
Vol. 37, No. 2, April, 2020, pp. 279-289
Journal homepage: http://iieta.org/journals/ts

Comparison of Classification Models Using Entropy Based Features from Sub-bands of EEG

Arshpreet Kaur[1*], Karan Verma[1], Amol P. Bhondekar[2], Kumar Shashvat[1]

[1] National Institute of Technology, Delhi 110040, India
[2] Central Scientific Instruments Organization, Chandigarh 160030, India

Corresponding Author Email: arshpreet@nitdelhi.ac.in

https://doi.org/10.18280/ts.370214

Received: 10 September 2019
Accepted: 26 January 2020

Keywords:
EEG classification, approximate entropy, sample entropy, fuzzy approximate entropy, random forest, AdaBoost, gradient boosting, naïve Bayes, linear discriminant analysis, quadratic discriminant analysis

ABSTRACT

The purpose of this study is to distinguish between different epileptic states automatically in an EEG. The work focuses on distinguishing activity of a controlled patient from inter-ictal and ictal activity and also from each other. Publically available Bonn database is used in this study. Seven such cases are considered. For this study three entropy features: approximate entropy, sample entropy and fuzzy approximate entropy are extracted from frequency sub-bands and are used with six classification algorithms which are Naive Bayes, LDA (Linear Discriminant Analysis), QDA(Quadratic Discriminant Analysis) from the generative group and RF(Random Forest), GB(Gradient Boosting) and Ada Boost from the ensemble group. The performance is evaluated on basis ofClassification accuracy, Sensitivity and Specificity.The results obtained direct that LDA as a classifier from the generative class and Ada boost from the ensemble group has outperformed other classifiers achieving the highest classification accuracies for three cases each respectively. Evaluating the results from sub-bands, we find out that D2 (21.7-43.4 Hz) sub-band clearly outperformed all the bands. Among the entropies used as features from sub bands, sample entropy outperforms the other entropies. From the results obtained it is established that frequency features from higher sub-band such as D2 (21.7-43.4 Hz) contain substantial information which can be used for identification of epileptic discharges which are however missed during visual analysis. This shows the impact automated methods can make in the field of identification of ictal and inter-ictal activity.

1. INTRODUCTION

Epilepsy is a seizure disorder in which range of severity varies with the patient. 9.72% people of 7.2 billion people worldwide suffer from epilepsy; out of this 17.14% reside in India [1]. Approximately more than a quarter people with this disease fall into the category where medication has no effect. Epileptic patients are also prone to SUDEP, which is sudden unexpected death during epilepsy. The death can occur during or post a seizure without any anatomical cause. Thus, making timely diagnose of epilepsy is of utmost importance. The time a patient spends having a seizure is fringe and is known as the ictal interval. This period is usually marked by sensory disturbances, loss of consciousness, convulsions, associated with abnormal electrical activity in the brain [1]. Patients usually get examined in their inter-ictal interval. Epileptiform activity present in patients EEG (Electroencephalography) post the seizures is referred to as inter-ictal activity. It is the time period where there is no clinical sign of epilepsy. However, it is not important for the inter-ictal activity to be always present and detailed study of patient history is always required for neurologist to make diagnosis. EEG is a safe technique used for diagnoses and monitoring for the presence of epileptiform discharges. The visual interpretation of EEG data for identification of epileptiform activity is a difficult task and needs a high level of expertise. With an enormous number of cases, the time for evaluation by neurologist/epileptologist also increases. To identify inter-ictal and ictal activity in EEG of a patient through automation has been a topic of interest for more than a decade. The goal of current work is to automate the process of EEG interpretation and facilitate in labeling of inter-ictal and ictal activity in EEG amid of various artifacts. The brain waves captured through EEG can be divided into five types' Delta (0-4 Hz), Theta (4-8 Hz), and Alpha (8-13 Hz), Beta (13-30Hz) and Gamma range (30-60 Hz). The beta waves and gamma waves are difficult to interpret visually and hence are overlooked by neurologist during visual analysis. This works contributes by considering the frequency band of higher frequency range and analyzing their contribution. Figure 1 shows the dissimilarity between electrical activities of controlled, inter- ictal and ictal states of patients taken from Bonn database.

To identify inter-ictal and ictal activity different linear and nonlinear parameters have been used along with different classifiers [2]. Entropy based parameters are popular among researchers as a feature and have been used for this problem over time such as in the researches [3-9]. Discriminative classifiers too have been the popular choice of researchers among which SVM (Support Vector Machine), ANN (Artificial Neural network) and ELM (Extreme Learning Machine) have been used most commonly with linear as well as nonlinear methods such as in the researches [10-13]. Chandaka et al. [10] used correlation a non-linear technique as a feature with MLPNN (Multilayer perceptron neural network)

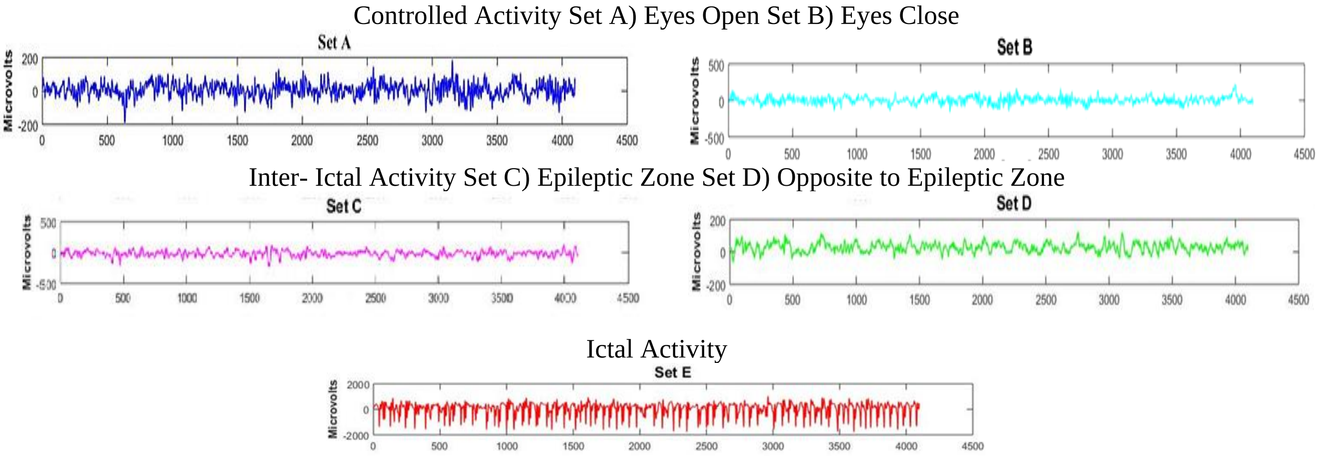

Figure 1. EEG of all groups

and SVM as classifier. They reported classification accuracy of 93.2% and 95.96% of for case A-E respectively. Rivero et al. [11] used relative wavelet energy from the five sub-bands decomposed using DWT (Discrete Wavelet Transform), and using ANN as a classifier. The classification accuracy of 95.52% was achieved for case A-E. Song et al. [12] explored Sample entropy as a feature changing the value of parameter '*m*' (embedding dimension) from one to three; and value of *'r'* (tolerance window) from 10%- 50% of standard deviation of data with increase of 10% at each step which is vector comparison distance. The value of N (Data Points) were also varied for which the chosen were 256, 512, 1024, 2048 and 4097. Extreme learning machines and BPNN were used as classifiers. Average learning accuracy of 95.67% with average learning time of only 0.25seconds was achieved using ELM (Extreme Learning Machines) with parameters *m*=3, *r=0.1 times standard deviation* and N=1024. The paper compared between the sets A, D and E (A-D-E).

In the study [13], Multilayer perceptron neural network based classification model was used to classify between five different cases. The cases considered were ABCD-E, A-E, AB-CDE, AB-CD-E and A-D-E. Discrete Wavelet Transform (DWT) was used to decompose the signal into five respective sub bands. By implementation of k means wavelet coefficients were clustered for each frequency sub-band and probability distributions were computed. The classification accuracies of 99.6%, 100%, 98.8%, 95.6% and 96.67% respectively were achieved for the cases specified above in order.

Chen et al. [14] compared ELM and SVM using three nonlinear features approximate entropy, Sample Entropy and RQA. These were extracted from the wavelet decomposed sub-bands. The preeminent performance was achieved using Sample entropy and ELM achieving utmost accuracy of 92.6%. The study by Kumar et al. [15] depicts the potential of sub-bands extracted using wavelet transform, (A1-A5) and (D1-D5). The study used approximate entropy with values of parameters *r* and *m* being 0.2 times standard deviation and 0.2 respectively. Performance of Artificial neural network and support vector machine was compared when fed with each sub band as a feature. A total of six cases were considered in this work; which were case 1(A-E), case 2(B-E), case 3(C-E), case 4(D-E), case 5 (ACD-E) and case 6 (BCD-E). The highest classification accuracy of 100% was achieved using approximate entropy as a feature from sub band D1 (43.4-86.8 Hz) and FBNN for the case (A-E) and case(C-E) respectively. Kaya et al. (2014) [16] implemented different classifiers such as SVM, ANN, Naive Bayes and others were used with the 1D-LBP approach for six cases. For case (A-E) and case (A-D) it achieved highest accuracy of 99.50% with FT and Naïve Bayes respectively. For all other cases which were D-E, E-CD, ABCD-E and A-D-E it achieved top classification accuracy of 95.5%, 97%, 93% and 95.67% respectively with Bayes Classifier. The results showed that Bayes classifier had the potential for classifying between different groups.

The proposed method by Xiang et al. [17] used fuzzy approximate entropy with dimension of phase space as two, i.e. (*m*=2) and similarity tolerance (*r*=0.25 times standard deviation). This is calculated from sub-bands using discrete wavelet transform and classified using SVM-RBF has shown 100% accuracy. Tawfik et al. [18] used weighted Permutation Entropy (WPE) from different sub bands of EEG signal extracted using DWT were fed into SVM for classification. For the two cases considered, A-B-C-D-E and A-D-E the author reported highest accuracies of 97.5% and 93.75% respectively. Supriya et al. [19] edge weight method using visibility graph in the complex network was implemented. Features including the average weighted degree of the complex network were inspected and fed into support vector machine (SVM). For the considered case (A-E), 100% of classification accuracy was achieved.

The aim of the work is to find the combination of entropy based feature and classifier which has the potential to distinguish between various considerations of inter ictal and ictal activity as well as inter-ictal and controlled activity. The group division and cases considered for this work are described in Table 2. For this work firstly, three entropies (approximate entropy, sample entropy and fuzzy approximate entropy) used in this work, which are explained in section 2.2 are extracted from five sub-bands specified in discussed in section 2.2. The work consists of two scenarios, for the first we take each entropy feature extracted from all five sub-bands considered as a feature set and feed into all six classifiers. For second scenario we used all three entropies extracted from five sub-bands; but each sub-band was used individually as a feature set for all six classifiers. This was done to find out if entropy extracted feature extracted from a single sub-band has potential to distinguish between different groups. The results obtained by first scenario are shown in Table 3-7 and for second scenario are shown in Table 8-10. Also, this work focuses to find if it is possible to successfully classify between different cases using a single sub band; considering this each entropy based feature which is extracted out of five sub-bands is used as only feature for all six classifiers. With the proposed method the D2 sub-band achieved highest classification accuracy of 100% for case 4(C-E). Also this sub band showed potential as important frequency range for feature extraction. The results obtained by using entropy extracted from each sub-band as only feature are discussed in Table 8-10. Comparison with existing methods is also established in Table 12.

2. CLINICAL DATA AND METHODOLOGY

2.1 EEG database and group division

Table 1. Details of data used

Total Data Folders	5
Controlled Group	2(A,B)
Inter-Ictal	2(C,D)
Ictal	1(E)
Time period of Signals	23.6 seconds
Sampling Frequency of Signal	173.6 Hz
Total no. of Data Points in each Signal	4097

Table 2. Division of groups for classification

Group 1: Healthy- Ictal	
Case 1	A-E
Case 2	B-E
Case 3	AB-E
Group 2: Inter- Ictal and Ictal	
Case 4	C-E
Case 5	D-E
Case 6	CD- E
Group 3: Healthy and Inter- Ictal	
Case 7	AB-CD

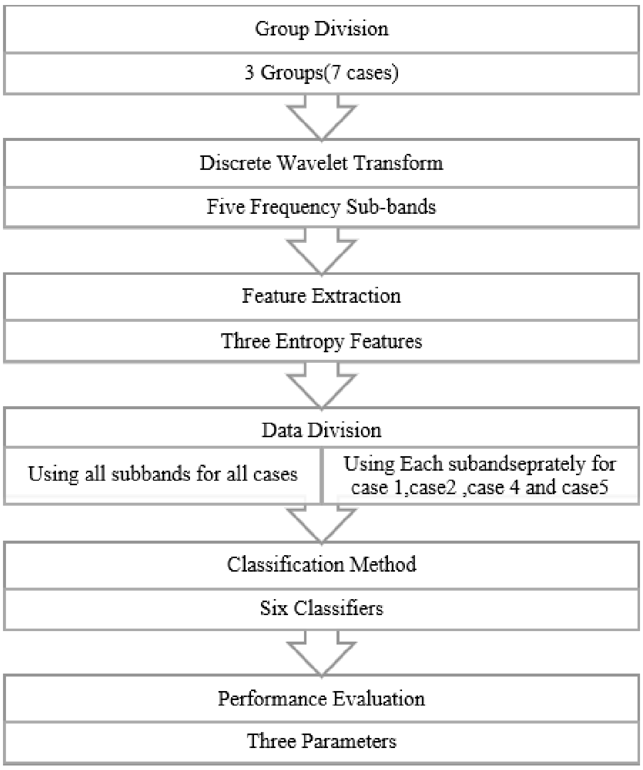

Figure 2. Methodology

Data is taken from the online source, University of Bonn [20]. The data available is already divided into five folders and three basic groups i.e. Controlled, Epileptic in ictal period and Epileptic in inter-ictal period. Each folder has 100 files, taken from 5 healthy and 5 epileptic subjects. Folder A and B contain data from healthy patients collected using surface electrodes. C and D have inter-ictal data and folder E has ictal data collected using the intracranial method. Table 1 hold detail about the data.

Figure 2 depicts the work flow for the current work.

The first step is to divide the data into groups as per the aim of the work. Table 2 holds the summary of the division.

2.2 Feature extraction

The brain waves are divided into five waves which are Delta (0-4 Hz), Theta (4-8 Hz), and Alpha (8-13 Hz), Beta (13-30Hz) and Gamma range (30-60 Hz). Different features linear and non-linear are extracted from these waves and are most generally used for seizure classification [21, 22]. In this work to understand the contribution of different frequency sub-bands and apprehend the role of higher frequency sub-bands in differentiating epileptiform discharges discrete wavelet transform was applied to the signal. The sampling frequency of the data is 173.6 Hz. DWT divides the complete frequency into different levels where each level of discrete wavelet transform corresponds to a specific sub-band. For this work level five decomposition is implemented using db4 wavelet. The level 5 was chosen as this decomposition allows to divide the sub-bands in frequency ranges closest to the required ranges of delta, theta, alpha, beta and gamma. Moreover, after the selected frequency range there is more possible of occurrence of the artefacts such as 50Hz (power line artefact).

The decomposition divides the data into following frequency: A5 (0-2.7125), D5 (2.71-5.4), D4 (5.4-10.8), D3 (10.85-21.7) and D2 (21.7-43.4). Figure 3 diagrammatically shows the process.

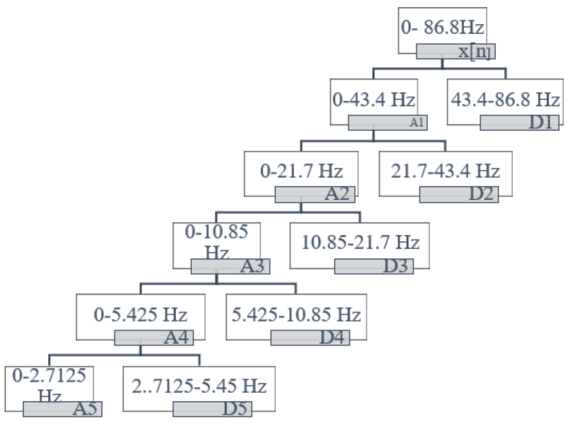

Figure 3. Level 5 decomposition using DWT

From these sub-bands three nonlinear entropy based features approximate entropy [23], sample entropy [24] and fuzzy approximate entropy [25] were extracted. Since the data is of 23.6seconds, we extract the features using complete data length. The entropy based features are Here N is the number of data points which is 4097 for all the three entropy parameters. In this work for all entropy features the value of *m* (embedding dimensions) is taken as 2 is and *r* (vector comparison distance) is 0.3 times the standard deviation of the signal.

DWT formula:

$$f(x) = \frac{1}{\sqrt{M}}\sum_{k} W_{\phi}(j_0,k)\phi_{j_0,k}(x) + \frac{1}{\sqrt{M}}\sum_{j-j_0}^{\infty}\sum_{k} W_{\varphi}(j,k)\varphi_{j,k}(x) \quad (1)$$

where, j_0 is an arbitrary starting scale

$$W_{\phi}(j_0,k) = \frac{1}{\sqrt{M}}\sum_{x=0}^{M-1} f(x)\,\ddot{\phi}_{j_0,k}(x) \quad (2)$$

$W_{\phi}(j_0,k)$ is called the approximation or scaling coefficients

$$W_{\psi} = \frac{1}{\sqrt{M}}\sum_{x=0}^{M-1} f(x)\,\widetilde{\psi_{J,k}}(x) \quad (3)$$

W_{ψ} is called the detail or wavelet coefficients

2.2.1 Approximate entropy

It is the likelihood that runs of patterns that are close remain close on next incremental comparisons. It is a measure of complexity that is applicable to noisy, medium-sized datasets. A high value of Approximate Entropy indicates random and unpredictable variation, whereas a low value of Approximate Entropy indicates regularity and predictability in a time series.

$$A_E(m,r,N) = \frac{1}{N-m}\sum_{i=1}^{N-m} ln\frac{n_i^m}{n_i^{m+1}} \quad (4)$$

2.2.2 Sample entropy

Sample Entropy does not amount a self-match, thus eradicating the prejudice in the direction of regularity. Sample Entropy has been suggested to be independent of data length and demonstrates relative consistency. It is less sensitive to noise.

$$S_E(m,r,N) = ln\frac{\sum_{i=1}^{N-m} n_i^m}{\sum_{i=1}^{N-m} n_i^{m+1}} \quad (5)$$

2.2.3 Fuzzy approximate entropy

In a real world scenario it is difficult to categorize an input to a specific class. Fuzzy approximate entropy works on this concept. With the concept of Lotfi, Zahed theory membership degree is introduced by fuzzy function u_z (x) having a real value between the range [0, 1].

For N data points u (i) =u (1), u (2), u (3)…..u (N) for finding the fuzzy approximate entropy

$$X_i^m = \{u(i), u(i+1), \ldots\ldots u(i+m-1\} - u_0(i) \quad \text{for i=1,2,3…N-m+1} \quad (6)$$

$u_0(i)$is baseline value:

$$u_0(i) = \frac{1}{m}\sum_{j=0}^{m-1} u(i+j) \quad (7)$$

Distance d_{ij}^m between two vectors X_i^mandX_j^m is defined as:

$$d_{ij}^m = d[X_i^m, X_j^m] = max_{k\in(0,m-1)}|u(i+k) - u_0(i) - (u(j+k) - u_0(j)|,\ j\neq I \quad (8)$$

For a given r, the similarity degree D_{ij}^m between X_i^m *and* X_j^m is determined by a fuzzy membership function u (d_{ij}^m, r).

$$D_{ij}^m = u(d_{ij}^m, r) \quad (9)$$

$$u(d_{ij}^m, r) = \exp\left({}^{-d_{ij}^2}/_{r}\right) \quad (10)$$

$$C_r^m(i) = \frac{1}{N-m+1}\sum_{j=1, j\neq I}^{N-m+1} D_{ij}^m \quad (11)$$

$$_{\varphi}{}^{m}(r) = \frac{1}{N-m+1}\sum_{i=1}^{N-m+1} \ln[C_r^m(i)] \quad (12)$$

$$\text{FAE}(m,r,N) = {}_{\varphi}{}^{m}(r) - {}_{\varphi}{}^{m+1}(r) \quad (13)$$

2.3 Classification

Two different categories of classifiers have been used in this work, generative class and ensemble class. The models are LDA, QDA, Naive byes, Random Forest, Ada boost and Gradient Boosting. For random forest, Ada boost and gradient boosting; 10 trees are used for this work. For gradient boosting algorithm the value of alpha which is the regularization coefficient is set to 1.

2.4 Data division

The data is divided into training (70%) and testing (30%). Performance parameters of the testing data are used for the comparison and evaluation.

2.5 Performance evaluation

Parameters which will evaluate the performance such as Accuracy, Sensitivity, Specificity, and Recall are calculated.

2.5.1 Accuracy

Proportion of people correctly identified into their actual groups i.e. in case1 (A-E) the accuracy will be high if all samples are allocated to their actual group.

2.5.2 Specificity

It is the measures the percentage of genuine positive cases that are suitably recognized as such. For example in case1 A-E, specificity will be high if number of signals correctly identified as ictal; i.e. correctly classified in group E will be high.

2.5.3 Sensitivity

Specificity events the percentage of actual negatives that are fittingly recognized as such for example in the case of B- E; the sensitivity will be high if the proportion of people identified as a controlled group i.e. with no epileptic discharge will be high.

3. RESULTS

EEG signals from different sets were decomposed into sub-bands A5 (0-2.71 Hz), D5 (2.71-5.4 Hz), D4 (5.4-10.8 Hz), D3 (10.85-21.7 Hz) and D2 (21.7-43.4 Hz). From these decomposed sub-bands approximate entropy, Sample entropy and fuzzy approximate entropy were computed. Six different classifiers LDA, QDA, Naive Bayes, Random Forest, Ada Boost, and Gradient Boosting are used in this study. The following table summarizes the results obtained when each type of entropy was extracted from five specified sub bands and used as set of features for a classifier; Table 3-7 hold the results for this scenario.

For Case 1 (A-E) the highest classification accuracy of 96.67% has been achieved by using a combination of Fuzzy Approximate Entropy and Naive Bayes as well as with Sample Entropy and Random Forest. In Case 2 the highest accuracy of 96.67% is achieved by using LDA as classifier and sample entropy or approximate Entropy as a feature. Also, for case 2 the ensemble models achieved the same result with respect to classification accuracy; when Random forest and Adaboost are used as classifiers with the combination of approximate entropy or sample Entropy as feature from all sub-bands.

Table 3. Results by generative models on case 1, case2, case 4 and case 5

					A-E				
	LDA			QDA			Naive Bayes		
	SE	AE	FAE	SE	AE	FAE	SE	AE	FAE
AC (%)	90	90	85	95	91.66	95	93.33	93.33	96.67
SP (%)	96.15	92.05	95.65	96.55	96.29	100	96.42	100	100
SN (%)	85.29	87.5	78.37	93.54	87.87	90.90	90.62	88.23	93.75
					B-E				
	LDA			QDA			Naïve Bayes		
	SE	AE	FAE	SE	AE	FAE	SE	AE	FAE
AC (%)	96.66	96.67	91.66	88.33	93.33	91.66	90	91.67	81.66
SP (%)	100	100	93.10	87.09	100	93.10	96.15	100	73.17
SN (%)	93.75	93.75	90.32	89.65	88.23	90.32	85.29	85.71	100
					C-E				
	LDA			QDA			Naïve Bayes		
	SE	AE	FAE	SE	AE	FAE	SE	AE	FAE
AC (%)	96.66	96.67	88.33	98.33	100	91.67	98.33	95	**91.67**
SP (%)	100	100	100	96.77	100	100	96.77	100	**100**
SN (%)	93.75	93.75	81.05	100	100	85.71	100	90.90	**85.71**
					D-E				
	LDA			QDA			Naïve Bayes		
	SE	AE	FAE	SE	AE	FAE	SE	AE	FAE
AC (%)	80	83.33	73.3	86.66	81.66	86.67	73.33	71.66	68.33
SP (%)	84.61	85.71	93.75	92.30	85.18	92.30	81.81	84.21	64.86
SN (%)	76.47	81.25	65.90	82.35	78.78	82.35	68.42	65.85	73.91

Table 4. Results by ensemble models on case 1, case2, case 4 and case 5

					A-E				
	Random Forest			Gradient Boosting			Ada Boost		
	SE	AE	FAE	SE	AE	FAE	SE	AE	FAE
AC (%)	96.6	95	95	95	95	95	91.679	95	93.33
SP (%)	93.75	93.54	93.54	96.55	93.54	96.55	87.87	96.5	96.42
SN (%)	100	93.54	96.52	93.54	96.55	93.54	96.29	93.54	90.62
					B-E				
	Random Forest			Gradient Boosting			Ada Boost		
	SE	AE	FAE	SE	AE	FAE	SE	AE	FAE
AC (%)	96.67	91.66	88.33	86.66	93.33	83.33	95	96.66	95
SP (%)	100	100	87.09	95.83	100	88.46	96.55	100	93.54
SN (%)	93.75	85.71	89.65	80.55	88.23	79.41	93.54	93.75	96.55
					C-E				
	Random Forest			Gradient Boosting			Ada Boost		
	SE	AE	FAE	SE	AE	FAE	SE	AE	FAE
AC (%)	98.33	95	88.33	96.66	100	85	100	98.33	90
SP (%)	100	100	96	100	100	88.88	100	100	96.154
SN (%)	96.77	90.90	82.57	93.75	100	81.81	100	96.77	85.29
								D-E	
	Random Forest			Gradient Boosting			Ada Boost		
	SE	AE	FAE	SE	AE	FAE	SE	AE	FAE
AC (%)	86.66	83.33	81.66	86.66	76.66	83.33	86.66	85	81.66
SP (%)	86.66	83.33	77.30	95.83	80.76	85.71	82.35	83.87	85.18
SN (%)	86.66	83.33	75.67	80.55	73.52	81.25	92.30	86.20	78.75

In Case 4, 100% accuracy, specificity and sensitivity were achieved through classification with approximate entropy (AE) and two classifiers (Naive Bayes and QDA). In Case 5, 86.67% accuracy was achieved through sample entropy (SE), AdaBoost, and QDA. In Case 3, the highest accuracy (93.3%) was achieved through the combination of fuzzy approximate entropy (FAE) and Naive Bayes. In Case 6, the highest accuracy (83.3%) was achieved through the combination of approximate entropy and linear discriminant analysis (LDA), and that of with sample approximate entropy (SAE) and LDA. In Case 7, the highest accuracy (80 %) was achieved using FAE+LDA. Furthermore, different types of entropies from each sub-band were taken as a singular feature of the six classifiers. Tables 8-10 list the results of the sub-bands with substantial performance. In Case 1 (A-E), the highest accuracy (91.66%) was achieved using FAE and random forest (RF) for D2, followed by 85% with AdaBoost+AE for D3, and 80% with SE+LDA for D4. In Case 2, the highest accuracy (93.33%) was achieved by AdaBoost+SAE and AE+RF for D2, followed by 73.3% with RF+AE for A5.

Cases 4 and 5 achieved similar accuracy in all sub-bands. In Case 4, 100% accuracy, specificity and sensitivity were achieved with RF+SE or RF+AE for D2; 85% accuracy was achieved with SE+LDA, SE+QDA, and SE+ Naïve Bayes for D3. In Case 5, 86.67% accuracy was achieved with AdaBoost+SE and with AdaBoost+AE for D2.

Table 5. Results by all models on case 3

Case 3 (AB-E)									
	Approximate Entropy			**Sample Entropy**			**Fuzzy Approximate Entropy**		
	AC (%)	SN (%)	SP (%)	AC (%)	SN (%)	SP (%)	AC (%)	SN (%)	SP (%)
LDA	93.33	90.91	100	92.22	89.55	100	88.89	85.71	100
QDA	93.33	93.55	92.86	93.33	93.55	92.86	93.33	93.55	92.86
Naive Bayes	90	91.80	86.21	90	91.80	86.21	96.67	96.72	96.55
Adaboost	91.11	91.94	89.29	92.22	94.92	87.10	90	89.23	92
RF	82.22	69.44	90.74	93.33	92.86	93.55	76.67	59.57	95.35
Gradient Boosting	83.33	92.45	70.27	93.33	93.55	92.86	77.78	93.48	61.36

Table 6. Results by all models on case 6

Case 6 (CD-E)									
	Approximate Entropy			**Sample Entropy**			**Fuzzy Approximate Entropy**		
	AC (%)	SN (%)	SP (%)	AC (%)	SN (%)	SP (%)	AC (%)	SN (%)	SP (%)
LDA	83.33	86.89	75.86	83.33	86.89	75.86	75.56	77.941	68.18
QDA	81.11	89.09	68.57	81.11	89.09	68.57	81.11	89.091	68.57
Naive Bayes	75.56	82.76	62.5	75.56	82.76	62.5	70	77.966	54.84
Adaboost	78.89	100	61.22	78.89	81.54	72	78.89	81.538	72
RF	78.89	61.70	97.67	80	63.04	97.73	80	66.667	88.89
Gradient Boosting	82.22	89.29	70.59	82.22	89.29	70.59	78.89	85.97	66.67

Table 7. Results by all models on case 7

Case 7 (AB-CD)									
	Apen			**Sample Entropy**			**Fuzzy Approximate Entropy**		
	AC (%)	SN (%)	SP (%)	AC (%)	SN (%)	SP (%)	AC (%)	SN (%)	SP (%)
LDA	55	54.69	55.36	60	58.82	61.54	80	92.86	73.08
QDA	69.17	64.94	76.74	73.33	70.59	76.92	64.17	70.73	60.76
Naive Bayes	55	53.26	60.71	55.83	54.12	60	70.83	73.59	68.66
Adaboost	69.17	69.49	68.85	70.83	70.49	71.19	60.83	65.85	58.23
RF	79.17	73.97	87.23	75	72.06	78.85	66.67	61.91	77.78
Gradient	67.5	69.09	66.15	76.67	79.63	74.24	64.17	68.09	61.64

Table 8. Results of generative models with sub bands by sample entropy and approximate entropy or case 1, 2, 4 and 5

A-E																		
	LDA						**QDA**						**Naive Bayes**					
	Sample Entropy			**Approximate Entropy**			**Sample Entropy**			**Approximate Entropy**			**Sample Entropy**			**Approximate Entropy**		
	CA	SN	SP	CA	SN	SP	CA	SN	SP	CA	SN	SP	CA	SN	SP	CA	SN	SP
D3	85	76.92	100	83.3	75	100	83.33	76.32	95.46	83.33	76.32	95.46	83.33	76.32	95.46	83.33	76.32	95.46
D4	80	75	87.5	73.3	69.4	79.1	80	76.5	84.6	70	66.7	75	80	76.5	84.6	70	66.7	75
C-E																		
	LDA						**QDA**						**Naive Bayes**					
	Sample Entropy			**Approximate Entropy**			**Sample Entropy**			**Approximate Entropy**			**Sample Entropy**			**Approximate Entropy**		
	CA	SN	SP	CA	SN	SP	CA	SN	SP	CA	SN	SP	CA	SN	SP	CA	SN	SP
D3	85	76.9	100	81.7	73.2	100	85	76.9	100	83.3	75	100	85	76.9	100	83.3	75	100
A5	81.7	77.1	88	66.7	66.7	66.7	76.7	71.1	86.4	66.7	61.9	77.8	76.7	71.1	86.4	66.7	61.9	77.8

Table 9. Results of ensemble models with sub bands by sample entropy and approximate entropy or case 1, 2, 4 and 5

				Random Forest					
				A-E					
	Approximate			**Sample**			**Fuzzy**		
	CA	SN	SP	CA	SN	SP	CA	SN	SP
D2	85	78.38	95.65	85	76.92	100	91.66	93.10	90.32
				Adaboost					
D2	85	100	76.92	86.67	95.83	80.56	85	76.92	100
D3	85	76.92	100	85	76.92	100	61.67	56.60	100
				B-E					
				Random Forest					
D2	93.33	88.23	100	86.66	82.35	92.30	81.66	75.67	91.30
				Adaboost					
D2	93.33	88.24	100	93.33	88.24	100	86.67	78.95	100
A5	73.33	70.59	76.92	66.67	63.16	72.73	68.33	64.10	76.19
				C-E					
				Random Forest					
D2	100	100	100	100	100	100	85	83.87	86.20
				Adaboost					
				D-E					
D2	86.67	100	78.95	86.67	100	78.95	73.33	93.75	65.91

Table 10. Results of ensemble models with sub bands by fuzzy approximate entropy or case 1, 2, 4 and 5

				Fuzzy Approximate Entropy					
				D-E					
	LDA			**QDA**			**Naive Bayes**		
	CA	SN	SP	CA	SN	SP	CA	SN	SP
A5	73.33	69.44	79.17	61.67	57.78	73.33	61.67	57.78	73.33

4. DISCUSSION

EEG sub bands, A5 (0-2.7 Hz), D5 (2.71-5.4 Hz), D4 (5.4-10.8 Hz), D3 (10.85-21.7 Hz) and D2 (21.7-43.4 Hz) were considered for this work. The data was divided into three groups and seven cases. Three entropy features which are approximate entropy, sample entropy and Fuzzy approximate entropy were calculated for all five sub-bands for each set for each sample. From analysis it has been observed that subband D2 plays important part in identification of inter-ictal and ictal discharges. Each set had a total of a hundred samples. The below given line plots with depict how approximate entropy and sample entropy for various samples of set C and set D vary for sub band D2.

In Figure 4, Figure 5, Figure 6 and Figure 7 the x-axis denotes the sample number and y-axis denotes the value of corresponding approximate entropy or sample entropy of the sub band D2 of specified set. From Figure 4 and Figure 5 it is evident that for sub-band D2 of set C has an evidently higher approximate entropy and sample entropy of sub-band E. Both the entropy features show a very similar trend when compared for set C and set D. In Figure 6 comparison of approximate entropy of sub-band D2 for set D and Set E is done while Figure 7 compares both the sets on sample entropy of samples obtained from sub band D2 .From these it is evaluated that both parameters i.e. approximate entropy and sample entropy of sub band set D lie in the lower spectrum compared from set E.

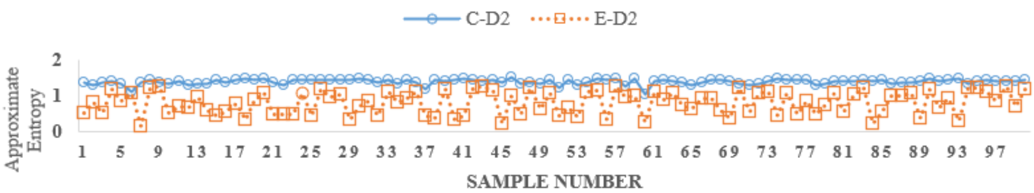

Figure 4. Line plots of Approximate entropy for case 4

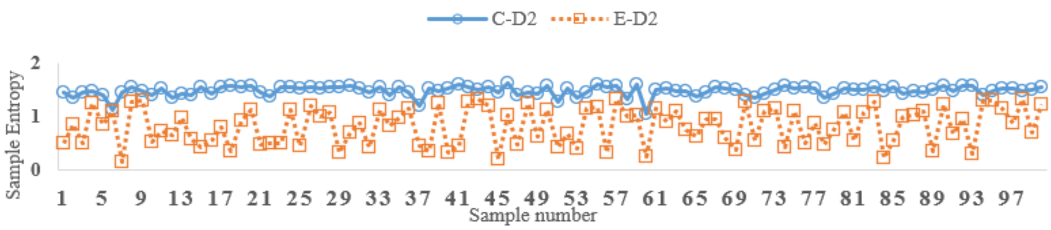

Figure 5. Line plots of Sample Entropy for case 4

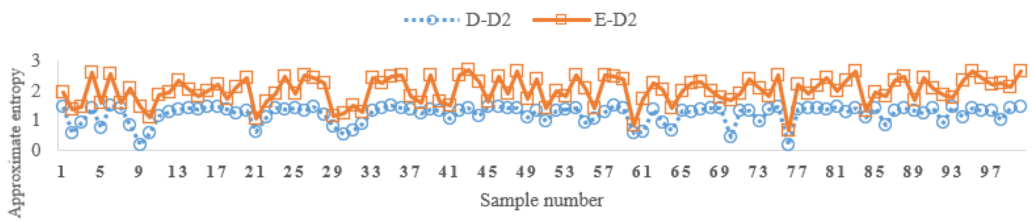

Figure 6. Line plots of Approximate Entropy for case 5

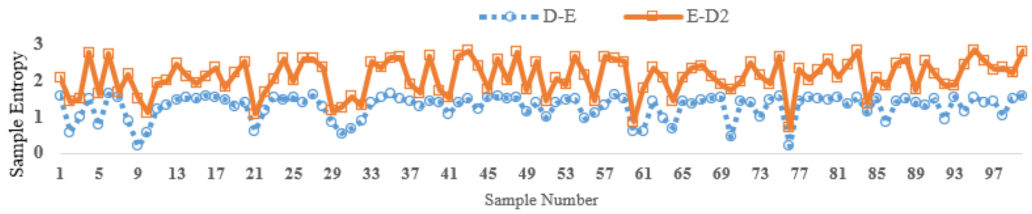

Figure 7. Line plots of sample entropy for case 5

The results obtained established that D2 sub-band has outperformed other bands as a singular feature with different classifiers. Table 11 summarizes the result obtained by D2 sub-band.

Table 11. Performance analysis of D2 sub band

D2 sub band			
Case	**CA (%)**	**Feature**	**Classifier**
A-E	91.66	FAE	RF
B-E	93.3	AE	RF
B-E	93.3	AE	ADABOOST
B-E	93.3	SE	ADABOOST
C-E	100	AE	RF
C-E	100	SE	RF
D-E	86.67	AE	ADABOOST
D-E	86.67	SE	ADABOOST

To establish a fair comparison this work is compared with other works done on similar lines. Table 12 holds comparison of current work with the previous work done by researchers. Kumar et al. [15] used approximate Entropy with artificial neural network and support vector machine; both of which are discriminative classifiers. For the work case 1, case 2, case 4 and case 5 considered in this study, were also evaluated along with some others. Comparing the results, we find that for case 2 using approximate entropy with LDA and Ada boost algorithm we achieved highest accuracy of 96.67% which is an improvement from 92.5% by the previous research. Xiang et al. [17] fuzzy approximate entropy and sample entropy were used with support vector machine for case 4 and case 5 considered in this study. Though, with Fuzzy approximate entropy and SVM 100% classification accuracy was achieved in both the cases; but with Sample entropy the accuracy yielded were 88.6% and 88.5%. In this study improvement with use of sample entropy has been achieved by its combination with all the other classifiers, where with Ada boost it achieved 100% accuracy and with LDA and Gradient Boosting it achieved 96.6% and with QDA and NB it achieved 98.33%. The improvement was also seen in parameters of specificity and sensitivity where the highest of 100% for both parameters was reached by the combination of this parameter with Ada Boost. Conventional features such a mean absolute value, standard deviation and others [26] were extracted from sub bands D3-D5 and A5 and classifiers SVM and Naïve Bayes was used in this work for classification. Case 4 used in this study was also considered by them; the researchers achieved 99.5% classification accuracy with 12 features; while, Kumar et al. [25] achieved highest accuracy of 99.6% for the same case with fuzzy entropy using all five sub bands. However, this study achieved 100% in all three statistical parameters with single sub band D2 with approximate entropy and random forest as well as with sample entropy and random forest.

Table 12. Comparison with existing work

Researcher and Year	Signal Used	Features Extracted	Classification	Data	Cases Considered	CA%
Kannathal et al. 2005 [7]	-	Entropy measures	Neuro-fuzzy inference system	Bonn Data	A-E	92.25
Umut Orhan et al. 2011 [13]	DWT	Clustered K means Coefficients of all sub bands	MLPNN	Bonn Data	ABCD-E	99.6
					A-E	100
					AB-CDE	98.8
					AB-CD-E	95.6
					A-D-E	96.67
Yatindra Kumar et al. 10 August 2014	Discrete wavelet transforms (DWT)	Approximate entropy (ApEn)·	Artificial neural network (ANN)·Support vector machine (SVM) 5-fold scheme	Bonn Data		

[15]			holdout method		A–E	100%	
					B–E	92.5	
					C–E	100	
					D–E	95	
					BCD–E	94	
					ABCD–E	94	
Yatindra Kumar et al. 20 January 2014 [25]	Discrete wavelet transforms (DWT)	Fuzzy approximate entropy	Support vector machine (SVM)	Bonn Data	A–E	100	
					B–E	100	
					C–E	99.6	
					D–E	95.85	
					ACD–E	98.15	
					BCD–E	98.15	
					ABCD–E	97.38	
Jie Xiang et al. 10 January 2015 [17]	Complete Signal	Fuzzy approximate entropy Sample approximate entropy	SVM	Bonn Data CHB-MIT		D-E	
						Fuzzy_E	Sample_E
					CA	100	87.6
					SP	100	90.79
					SN	100	87.5
						C-E	
						Fuzzy_E	Sampke_E
					CA	100	88.5
					SP	100	90.36
					SN	100	87.63
Current Work	DWT	Sample Entropy, Fuzzy Approximate Entropy, Sample Entropy	Naïve Bayes, LDA, QDA, Ada Boost, Gradient Boost, Random Forest	Bonn Data	Case	CA(%)	Technique
					A-E	*96.67*	*FAE+NB,SE +RF*
					B-E	*96.67*	*Apen+ LDA,Adaboost, Sample+LDA,RF*
					C-E	*100*	*Apen+QDA,GB*
					D-E	*86.66*	*Sampen+AdaBoos, QDA*
					AB-E	93.33	Apen+LDA,QDA
					CD-E	96.66	FAE+NB
					AB-CD	80	FAE+LDA

5. CONCLUSION

Among the entropies used as features from sub bands, sample entropy outperforms the other entropies. It achieved highest accuracy with combination with Random forest for case 1. With LDA and Random forest for case 2. For case 4 and case 5 with Ada boost as well as Gradient Boosting; and only with LDA for case 6. LDA has outperformed all the classifiers achieving the highest accuracy for five out of seven cases which are case 2, case 3, case 4, case 6 and case 7. For case 2, case 3 and case 6 it provided the highest accuracy with approximate entropy while for case 4 and case 7 it was the combination of sample and fuzzy approximate entropy. Naive Bayes achieved the highest accuracy in consideration for case1 with fuzzy approximate entropy. Among the ensemble methods Ada boost has achieved the highest accuracy for case 2 with approximate entropy and case 4, case 5 with Sample Entropy. Gradient boosting achieved the highest accuracy for case 4 with Approximate Entropy. The D2 sub band has outperformed all the other sub band; it achieved accuracy as high as achieved using all the sub bands together for case 3 and case 4 which was 86.66% and 100% respectively. Ada boost has achieved the highest accuracy for case 2 and case 5 with sample and approximate entropy. While Random forest has achieved with case 4 and case 2 with sample and approximate entropy.

ACKNOWLEDGMENT

The authors would like to acknowledge Dr R.G.Andrzejak of University of Bonn, Germany, for providing permission to use the EEG data available in the public domain.

REFERENCES

[1] Amudhan, S., Gururaj, G., Satishchandra, P. (2015). Epilepsy in India I: Epidemiology and public health. Annals of Indian Academy of Neurology, 18(3): 263-277. https://doi.org/doi:10.4103/0972-2327.160093

[2] Hsu, K.C., Yu, S.N. (2010). Detection of seizures in EEG using subband nonlinear parameters and genetic algorithm. Computers in Biology and Medicine, 40(10): 823-830. https://doi.org/doi:10.1016/j.compbiomed.2010.08.005

[3] Quiroga, R.Q., Arnhold, J., Lehnertz, K., Grassberger, P. (2000). Kulback-Leibler and renormalized entropies:

Applications to electroencephalograms of epilepsy patients. Physical Review E, 62(6): 8380-8386. https://doi.org/10.1103/physreve.62.8380
[4] Kopitzki, K., Warnke, P.C., Timmer, J. (1998). Quantitative analysis by renormalized entropy of invasive electroencephalograph recordings in focal epilepsy. Physical Review E, 58(4): 4859-4864. https://doi.org/10.1103/physreve.58.4859
[5] Ocak, H. (2009). Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy. Expert Systems with Applications, 36(2): 2027-2036. https://doi.org/10.1016/j.eswa.2007.12.065
[6] Acharya, U.R., Molinari, F., Sree, S.V., Chattopadhyay, S., Ng, K., Suri, J.S. (2012). Automated diagnosis of epileptic EEG using entropies. Biomedical Signal Processing and Control, 7(4): 401-408. https://doi.org/10.1016/j.bspc.2011.07.007
[7] Kannathal, N., Acharya, U.R., Lim, C.M., Sadasivan, P.K. (2005). Characterization of EEG—A comparative study. Computer Methods and Programs in Biomedicine, 80(1): 17-23. https://doi.org/10.1016/J.CMPB.2005.06.005
[8] Nicolaou, N., Georgiou, J. (2012). Detection of epileptic electroencephalogram based on permutation entropy and support vector machines. Expert Systems with Applications, 39(1): 202-209. https://doi.org/10.1016/j.eswa.2011.07.008
[9] Kannathal, N., Lim, M., Acharya, U.R., Sadasivan, P.K. (2005). Entropies for detection of epilepsy in EEG. Computer Methods and Programs in Biomedicine, 80(3): 187-94. https://doi.org/10.1016/j.cmpb.2005.06.012
[10] Chandaka, S., Chatterjee, A., Munshi, S. (2009). Cross-correlation aided support vector machine classifier for classification of EEG signals. Expert Systems with Applications, 36(2 PART 1): 1329-1336. https://doi.org/10.1016/j.eswa.2007.11.017
[11] Guo, L., Rivero, D., Seoane, J.A., Pazos, A. (2009). Classification of EEG signals using relative wavelet energy and artificial neural. In Proceedings of the first ACM/SIGEVO Summit on Genetic and Evolutionary Computation (GEC '09), pp. 177-184. https://doi.org/10.1145/1543834.1543860
[12] Song, Y., Liò, P. (2010). A new approach for epileptic seizure detection: sample entropy based feature extraction and extreme learning machine. Journal of Biomedical Science and Engineering, 3(6): 556-567. https://doi.org/10.4236/jbise.2010.36078
[13] Orhan, U., Hekim, M., Ozer, M. (2011). EEG signals classification using the K-means clustering and a multilayer perceptron neural network model. Expert Systems with Applications, 38(10): 13475-13481. https://doi.org/10.1016/j.eswa.2011.04.149
[14] Chen, L.L., Zhang, J., Zou, J.Z., Zhao, C.J., Wang, G.S. (2014). A framework on wavelet-based nonlinear features and extreme learning machine for epileptic seizure detection. Biomedical Signal Processing and Control, 10(1): 1-10. https://doi.org/10.1016/j.bspc.2013.11.010
[15] Kumar, Y., Dewal, M.L., Anand, R.S. (2012). Epileptic seizures detection in EEG using DWT-based ApEn and artificial neural network. Signal, Image and Video Processing, 8(7): 1323-1334. https://doi.org/10.1007/s11760-012-0362-9
[16] Kaya, Y., Uyar, M., Tekin, R., Yıldırım, S. (2014). 1D-local binary pattern based feature extraction for classification of epileptic EEG signals. Applied Mathematics and Computation, 243: 209-219. https://doi.org/10.1016/j.amc.2014.05.128
[17] Xiang, J., Li, C., Li, H., Cao, R., Wang, B., Han, X., Chen, J. (2015). The detection of epileptic seizure signals based on fuzzy entropy. Journal of Neuroscience Methods, 243: 18-25. https://doi.org/10.1016/j.jneumeth.2015.01.015
[18] Tawfik, N.S., Youssef, S.M., Kholief, M. (2016). A hybrid automated detection of epileptic seizures in EEG records. Computers and Electrical Engineering, 53: 177-190. https://doi.org/10.1016/j.compeleceng.2015.09.001
[19] Supriya, S., Siuly, S., Zhang, Y. (2016). Automatic epilepsy detection from EEG introducing a new edge weight method in the complex network. Electronics Letters, 52(17): 17-18. https://doi.org/10.1049/el.2016.1992
[20] Andrzejak, R.G., Lehnertz, K., Mormann, F., Rieke, C., David, P., Elger, C.E. (2001). Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. Physical Review E, 64(6): 061907. https://doi.org/10.1103/PhysRevE.64.061907
[21] Jahankhani, P., Kodogiannis, V., Revett, K. (2006). EEG signal classification using wavelet feature extraction and neural networks. Proceedings - IEEE John Vincent Atanasoff 2006 International Symposium on Modern Computing, Sofia, pp. 120-124. https://doi.org/10.1109/JVA.2006.17
[22] Guo, L., Rivero, D., Pazos, A. (2010). Epileptic seizure detection using multiwavelet transform based approximate entropy and artificial neural networks. Journal of Neuroscience Methods, 193(1): 156-163. https://doi.org/10.1016/j.jneumeth.2010.08.030
[23] Burioka, N., Miyata, M., Cornélissen, G., Halberg, F., Takeshima, T., Kaplan, D.T., Shimizu, E. (2005). Approximate entropy in the electroencephalogram during wake and sleep. Clinical EEG and Neuroscience, 36(1): 21-24. https://doi.org/10.1177/155005940503600106
[24] Al-Angari, H.M., Sahakian, A.V. (2007). Use of sample entropy approach to study heart rate variability in obstructive sleep apnea syndrome. IEEE Transactions on Biomedical Engineering, 54(10): 1900-1904. https://doi.org/10.1109/TBME.2006.889772
[25] Kumar, Y., Dewal, M.L., Anand, R.S. (2014). Epileptic seizure detection using DWT based fuzzy approximate entropy and support vector machine. Neurocomputing, 133: 271-279. https://doi.org/10.1016/j.neucom.2013.11.009
[26] Sharmila, A., Geethanjali, P. (2016). DWT based detection of epileptic seizure from EEG signals using naive Bayes and k-NN classifiers. IEEE Access, 4: 7716-7727. https://doi.org/10.1109/ACCESS.2016.2585661

NOMENCLATURE

SE	Sample Entropy
AE	Approximate Entropy
FAE	Fuzzy Approximate Entropy
LDA	Linear Discriminant Analysis

RF	Random forest	ELM	Extreme Learning Machine
QDA	Quadratic Discriminant Analysis	SP	Specificity
NB	Naive Bayes	SN	Sensitivity
SVM	Support vector Machine	AC	Accuracy