www.ijecs.in
International Journal Of Engineering And Computer Science ISSN:2319-7242
Volume – 5 Issue -04 April, 2016 Page No. 16073-16078

Feature Extraction and Selection of a Combination of Entropy Features for Real-time Epilepsy Detection

B. Abhinaya[1], D. Charanya[1], K. Palani Thanaraj[2*]

[1]Department of Electronics and Instrumentation Engineering,
St. Joseph's College of Engineering, Anna University, Chennai-600119, India
abhinayababu22@gmail.com

[1]Department of Electronics and Instrumentation Engineering,
St. Joseph's College of Engineering, Anna University, Chennai-600119, India
charanya.suga@gmail.com

[2]Department of Electronics and Instrumentation Engineering,
St. Joseph's College of Engineering, Anna University, Chennai-600119, India
palanithanaraj.k@gmail.com

Abstract*: Epilepsy is associated with the abnormal electrical activity in the brain which is detected by recording EEG (Electroencephalogram) signals. This signal is non-linear and chaotic and hence, it is very time-consuming and tedious to analyse them visually. In this work, we have extracted five entropy features such as Approximate Entropy, Sample Entropy, Fuzzy Entropy, Permutation Entropy and Multi-scale Entropy for characterizing the focal signals. We have used Sequential Forward Feature Selection (SFFS) algorithm to select two significant features for epilepsy classification. These two features are given as input to the Least Square Support Vector Machine (LS-SVM) classifier to differentiate normal and focal signal. The classification accuracy of our method is 82%. Moreover, the average computational time for the selected feature set is 47.94 seconds.*

Keywords: Epilepsy, EEG signal, Entropy, LS-SVM classifier

1. INTRODUCTION

Epilepsy is a nervous disorder that causes people to have repeated seizures. During epilepsy, the normal neuronal activity is disturbed which results in strange emotions and behaviour or sometimes convulsions, and loss of consciousness. Seizures results when the normal neural activity gets affected ranging from illness, brain damage to abnormal brain development (1).

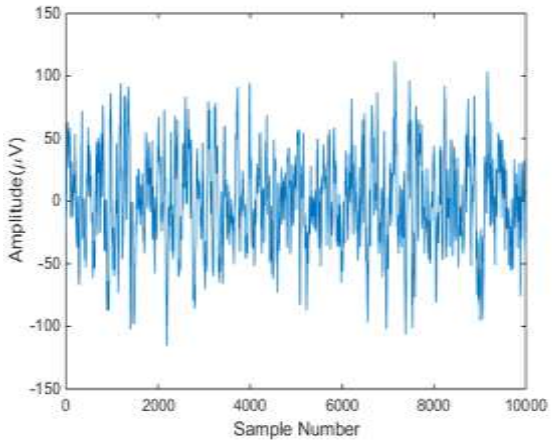
The EEG signal is widely used in the diagnosis of epilepsy(2). During epilepsy, only a small number of processes and functions are active in the brain regions (3). Epileptic seizures can be identified from non-epileptic waveforms by observing the EEG recordings. Fig.1 shows that the randomness of the normal signal is more when compared to the focal signal. Analysis by visual inspection of EEG to detect epilepsy is a tedious process and causes immense workload on the physician (2). Therefore, there is a need to develop automated epilepsy detection systems.

Recently, Rajeev Sharma *et. al.* decomposed the EEG signals using empirical mode decomposition (EMD) to extract intrinsic mode functions (IMF)(4). They used five entropy features and the LS-SVM classifier in their approach. Chia-Ping Shen *et. al.* used multichannel EEG signals and employed genetic algorithm along with a support vector machine for epilepsy classification(5). Xiang *et. al.,* proposed fuzzy entropy to classify seizures and validated their approach using two datasets (6).

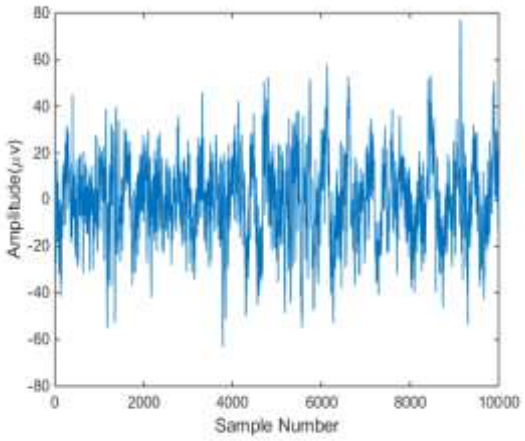
They used SVM classifier and reported an accuracy of 98.3% for CHB-MIT database. Acharya *et. al.* used different complexity and entropy measures for epilepsy classification. They selected significant features based on Analysis Of Variance (ANOVA) statistical test for epilepsy classification(7). However, much work on EEG classification using entropy features is done in the transform domain such as empirical mode decomposition or wavelet domain(8). Investigation of a combination of entropy features in the time domain has lacked behind. Moreover, the extraction of features in the transform domain requires more computational time, and limits the application of the autonomous epilepsy detection system in real time portable medical systems.

Thus, the main aim of our current study is to investigate a combination of entropy features in the time domain. We also analyse the computational time of the entropy features. We also propose to use Sequential Forward Feature Selection (SFFS) algorithm for feature selection(9). Moreover, we also examine the computation time and classification accuracy of a combination of entropy features and compare with the selected features from SFFS algorithm for real-time medical applications.

DOI: 10.18535/ijecs/v5i4.03



a)



b)

Fig. 1 EEG signal a) Normal b) Focal

2. MATERIAL AND METHOD

The EEG recordings used in our current work of epileptic seizure detection is collected from *Bern-Barcelona EEG database* (University Pompeu Fabra (UPF), Barcelona) (10).

2.1 The Bern-Barcelona EEG Database:

The datasets 'F' and 'N' are used to refer focal and non-focal signals in the Bern database(10). Focal refers to the epileptic patients and non-focal refers to normal subjects. All the EEG data is free from artifacts. Each data file consists of signal pair that includes x-signal and y-signal. The x-signal is present in the first column and the y-signal is present in the second column of the data file. The amplitude values for all these signal pairs are given in ASCII format in the database. We have used 50 focal and 50 non-focal signals from the open database and they are processed in MATLAB.

2.2 Feature Extraction:

In this paper, we use a range of entropy features that includes Sample Entropy, Fuzzy Entropy, Permutation entropy, Multi Scale Entropy and Approximate Entropy for EEG signal analysis and classification.

2.2.1 Entropy:

Entropy is a non-linear index that describes the disorder or randomness of a signal (2,3). The principle behind this feature is that the irregularity within an EEG signal decreases during epilepsy(2).

(a) Approximate Entropy:

Approximate Entropy (ApEn) is a technique that quantifies the degree of irregularity and the unpredictability of fluctuations over time-series data. The ApEn value is higher when the signal is less predictable and the ApEn value is less for a predictable time series (11).

$$\boldsymbol{ApEn(S_N, m, r) = ln\left[\frac{C_m(r)}{C_{m+1}(r)}\right]} \qquad \textbf{(1)}$$

Where S_N is the given sequence consisting of N instantaneous measurements. Here *'m'* specifies the pattern length, *'r'* defines the criterion of similarity, $C_{im}(r)$ is the fraction of patterns of length *'m'* that is similar to the pattern of the same length that starts at interval *'i'*.

(b) Sample Entropy:

Sample entropy (SampEn) is a modification of Approximate Entropy. SampEn does not include self-similar patterns as the ApEn. SampEn is a measure of complexity of a signal (12).

$$\textbf{SampEn}(\boldsymbol{m, r, N}) = \textbf{-ln}\left[\frac{A^{m+1}(r)}{A^m(r)}\right] \qquad \textbf{(2)}$$

Where $A^m(r) = \sum_{i=1}^{N-M+1} \frac{C_i^m(r)}{N-M+1}$

(c) Fuzzy Entropy:

Fuzzy entropy is a measure of relative degree of randomness and chaos in the signal(6). In equation 3, A is a fuzzy variable with a continuous membership function and its entropy is given by

$$\mathbf{H}[\boldsymbol{A}] = \int_{-\infty}^{\infty} \mathbf{S}\,(\boldsymbol{C_r}\{\boldsymbol{A \geq t}\})\,\mathbf{dt} \qquad \textbf{(3)}$$

Where

S(y) = -y ln y – (1-y) ln (1-y)

The main advantage of fuzzy is that it is sensitive to the information content rather than noise component in the signals.

(d) Permutation Entropy:

Permutation Entropy provides an alternate method to estimate complexity and randomness of a time series data(13). It is a measure based on analysis of permutation patterns in the

signal. Permutation Entropy (PE) can be applied to non-stationary signals such as EEG signals. PE is given by,

$$PE = - \sum_{j=1}^{n} p_j \log_2 p_j \qquad (4)$$

(e) Multiscale Entropy:

The Multiscale entropy (MSE) analysis is useful for investigating complexity of physiologic signals and other time series signals that has statistical relation at different time scales. The MSE is based on the evaluation of SampEn on multiple scales of the input signal (14).

2.3 Feature Selection:

We have used Sequential Forward Feature Selection (SFFS) algorithm to select significant entropy features(9). SFFS is also known as heuristic feature search. In this algorithm, the most significant feature is first selected, and then feature pairs are formed by combining the best feature with the remaining features and then the best pair is selected.

2.4 Feature Classification:

In our current work, we use Least Squares Support Vector Machines for feature classification. The least squares support vector machine (LS-SVM) method consists of a set of linear equations to create a decision boundary (hyper plane) between two different groups of patterns (here normal and focal signals)(15). The performance of the LS-SVM classifier with the selected entropy feature set is evaluated using *k*-fold (*k*=10) cross validation technique. The performance measures such as Sensitivity (SEN), Specificity (SPF), Accuracy (ACC), Positive Predictive value (PPV) and Negative Predictive value (NPV) are computed.

Sensitivity is the number of true positives divided by true positives and false negatives. **Specificity** is the number of true negatives divided by true negatives and false positives. **Accuracy** is the number of true positives and true negatives divided by total number of EEG dataset (true positive and true negative and false positive and false negative). **Positive Predictive Value** is the number of true positives divided by true and false positives.

Negative Predictive Value is the number of true negatives divided by true and false negatives. True positive (TP) signifies epileptic data correctly identified. False positive (FP) signifies normal data incorrectly identified as epileptic signal. True negative (TN) refers to normal data correctly identified. False negative (FP) signify the epileptic data incorrectly identified as normal.

3. RESULTS AND DISCUSSION

We have evaluated the combination of entropy features of EEG signals obtained from fifty normal and fifty epileptic patients from the Bern database. By processing these signals in MATLAB, we extracted five entropy features. Table 1 shows the entropy values of five normal and Table 2 shows the entropy measures of five epileptic patients. We can observe from the table that the entropy values for epileptic patients are less when compared to normal subjects. The entropy values are also plotted using boxplot. The X-axis in the boxplot represents the normal and focal EEG signals and Y-axis represents the entropy values .The red line inside the box indicates the median of all the values.

In fig. 2 the sample entropy values for normal signal ranges between 0.14 to 0.97 and focal signal ranges between 0.12 to 0.7. In fig. 3 the fuzzy entropy values for the normal signal ranges between 0.00007 to 0.0147 and focal signal ranges from 0.000095 to 0.0284. In fig. 4, the permutation entropy values for the normal signal ranges from 0.6912 to 0.6931 and focal signal ranges from 0.6923 to 0.6931. In fig. 5, the multiscale entropy values for the normal signal ranges from 0.0146 to 0.975 and focal signal ranges from 0.01285 to 0.7045. In fig. 6, the approximate entropy values for the normal signal ranges from 0.1926 to 1.044 and focal signal ranges from 0.1715 to 0.8642.

The *p* values are also obtained for all five entropies by using *student t-test* in MATLAB (refer Table 3). Although, the *p* value is less than 0.05 for all the entropy measures except PE, we have chosen two significant features (SampEn & ApEn) from five entropies based on SFFS algorithm .The selected two entropy features are given as input to the LS-SVM classifier. The LS-SVM classifies the signal as normal and focal signals. We have used Radial Basis Function (RBF) as kernel function in the LS-SVM subroutine.

In fig. 7, class 1 represented by star symbol corresponds to normal signal and class 2 represented by square symbol corresponds to focal signal. The X1 and X2 in fig. 7 represents the two features SampEn and ApEn. The Receiver Operating Characteristic (ROC) curve is also plotted as shown in fig. 8. In fig. 8, X-axis represents 1-Specificity and Y-axis represents the Sensitivity. The area of the ROC is 0.8804. The standard deviation of ROC is 0.033264. The overall maximum classification accuracy obtained in our analysis is 82% as shown in Table 4. Moreover, the other performance measures such as Sensitivity, Specificity, NPV, and PPV are shown in Table 4 for reference.

DOI: 10.18535/ijecs/v5i4.03

Table 1 Entropy values of Normal patients

Entropy Values of Normal Patients										
Patients	*Sample Entropy*		*Fuzzy Entropy*		*Permutation Entropy*		*Multiscale Entropy*		*Approximate Entropy*	
	X	*Y*	*X*	*Y*	*X*	*Y*	*X*	*Y*	*X*	*Y*
1	0.5208	0.5283	0.0021	0.0027	0.6930	0.6930	0.5208	0.5283	0.5771	0.5886
2	0.5529	0.5164	0.0022	0.0019	0.6931	0.6931	0.5530	0.5164	0.6259	0.5897
3	0.6938	0.6562	0.0051	0.0062	0.6931	0.6931	0.6936	0.6561	0.7528	0.7230
4	0.7350	0.6397	0.0128	0.0120	0.6931	0.6931	0.7350	0.6396	0.7889	0.7022
5	0.6172	0.6317	0.0072	0.0065	0.6931	0.6931	0.6171	0.6317	0.6855	0.6870

Entropy Values of Patients with Epilepsy										
Patients	*Sample Entropy*		*Fuzzy Entropy*		*Permutation Entropy*		*Multiscale Entropy*		*Approximate Entropy*	
	X	*Y*	*X*	*Y*	*X*	*Y*	*X*	*Y*	*X*	*Y*
1	0.5054	0.5341	0.0020	0.0025	0.6931	0.6931	0.5053	0.5339	0.5658	0.5926
2	0.4687	0.5686	0.0013	0.0042	0.6931	0.6931	0.4685	0.5687	0.5338	0.6726
3	0.2775	0.2806	0.0005	0.0007	0.6931	0.6931	0.2774	0.2805	0.3647	0.3936
4	0.4803	0.5910	0.0020	0.0031	0.6931	0.6931	0.4805	0.5911	0.5522	0.6545
5	0.3520	0.2824	0.0006	0.0004	0.6931	0.6931	0.3521	0.2824	0.4281	0.3544

Table 2 Entropy values of Epilepsy patients

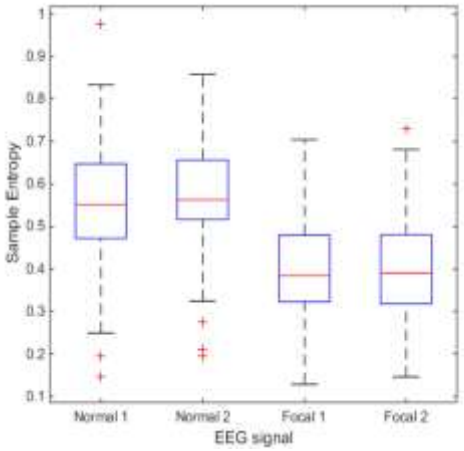


Fig. 2: Boxplot of Sample Entropy

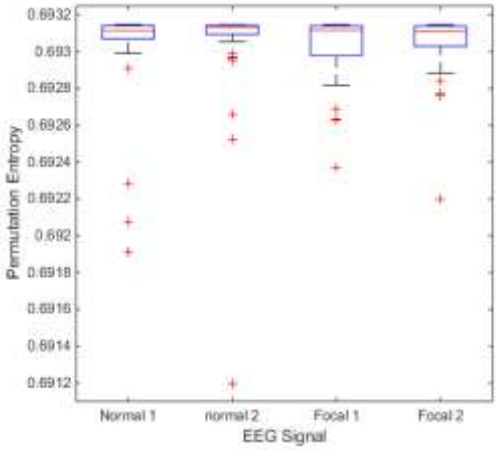
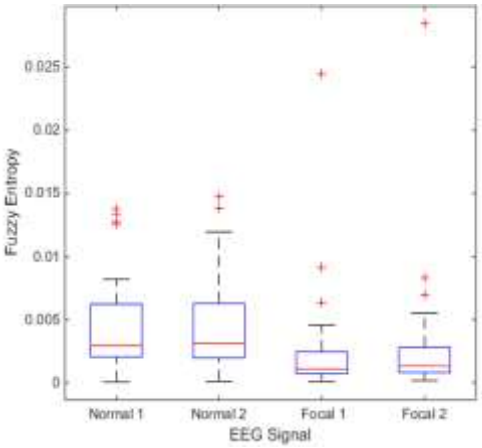
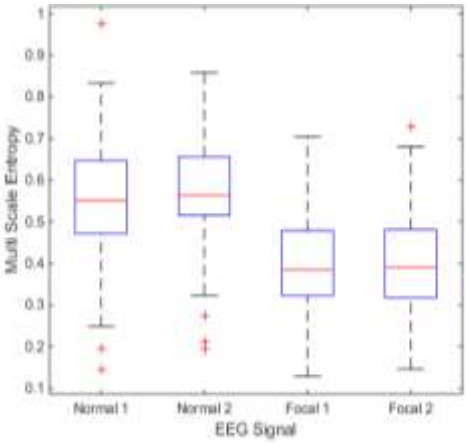


Fig. 4: Boxplot of Permutation Entropy





DOI: 10.18535/ijecs/v5i4.03

Fig. 3: Boxplot of Fuzzy Entropy

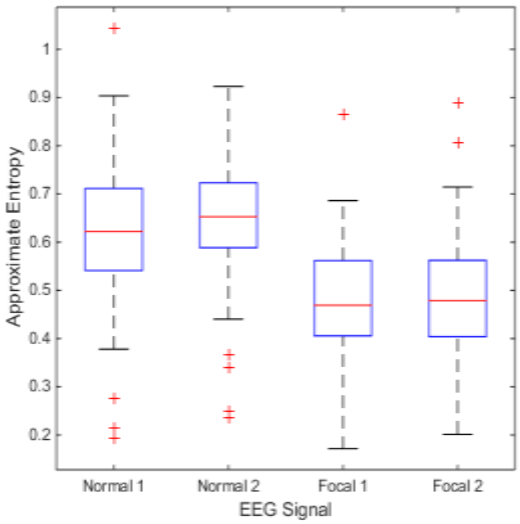


Fig. 6: Boxplot for Approximate entropy

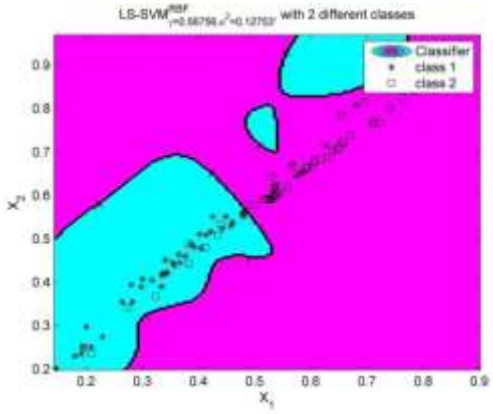


Fig. 7: LS-SVM classifier output showing the hyper planes

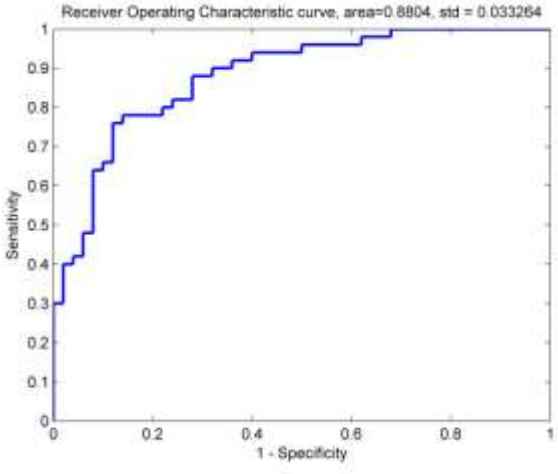


Fig. 8: Receiver Operating Characteristic curve

Fig. 5: Boxplot of Multiscale Entropy

Table 3: *p*-values of various entropy features

Entropy measure	*p*-value
MSE	0.0000000780
SampEn	0.0000000778
ApEn	0.0000021536
FEn	0.0047383791
PE	0.6844201270

Table 4: LS-SVM Classifier performance measures

Sensitivity (%)	***Specificity (%)***	***Maximum Accuracy (%)***	***PPV (%)***	***NPV (%)***	***AUC***
70.62	68.52	82	77.696	78.18	0.8804

The accuracy could be increased by considering more features but the computational time also increases accordingly. Our work is more focused on selection of limited features that can produce acceptable classification accuracy with less computation time. Table 5 shows that the computational time for extraction of two features is less when compared to extraction of five features. The experiments were performed in Intel based Mobile PC with 4 GB of RAM and i3 processor clocked at 2.4 GHz.

Table 5: Comparison of computation time of extracted features

No. of Entropy measures	**5**	**2**
Computation time for Normal signal (sec)	102.3	47.24
Computation time for Focal signal (sec)	103.04	48.65
Average Computation time (sec)	102.67	47.94
Classification Accuracy (%)	85	82

4. CONCLUSION

In this paper, we have analysed a combination of entropy features for classification of normal and focal EEG signals. Our method selects two significant entropy measures (SampEn and ApEn) as features, which are given as inputs to LS-SVM classifier in order to differentiate normal and focal signals. It achieved a classification accuracy of 82%. Moreover, the area under the curve is 0.8804 obtained by Receiver Operating characteristic Curve (ROC) plot. Even though combining the other features produced a higher accuracy, the proposed method achieved an average computation time of 47.94

seconds for feature extraction of ten normal and ten epileptic subjects, which is significantly less compared to combination of all the features. Thus, the current work suggests that combining SampEn and ApEn can provide acceptable classification accuracy in real-time implementation of epilepsy detection system.

Acknowledgements

None

Conflict of Interest

None

REFERENCES

1. Scheffer IE, Mullen S a. Epilepsy in 2012: Advances in epilepsy shed light on key questions. Nature reviews Neurology [Internet]. 2013;9(2):66–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23296338

2. Thanaraj KP, Chitra K. Multichannel Feature Extraction and Classification of Epileptic States Using Higher Order Statistics and Complexity Measures. International Journal of Engineering & Technology (0975-4024) [Internet]. 2014;6(1):102–9. Available from: http://www.enggjournals.com/ijet/abstract.html?file=14-06-01-163

3. Kannathal N, Choo ML, Acharya UR, Sadasivan PK. Entropies for detection of epilepsy in EEG. Computer Methods and Programs in Biomedicine. 2005;80:187–94.

4. Sharma R, Pachori RB, Gautam S. Empirical Mode Decomposition Based Classification of Focal and Non-focal EEG Signals. 2014 International Conference on Medical Biometrics [Internet]. 2014;135–40. Available from: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6845839

5. Shen CP, Liu ST, Zhou WZ, Lin FS, Lam AYY, Sung HY, et al. A Physiology-Based Seizure Detection System for Multichannel EEG. PLoS ONE. 2013;8(6).

6. Xiang J, Li C, Li H, Cao R, Wang B, Han X, et al. The detection of epileptic seizure signals based on fuzzy entropy. Journal of neuroscience methods [Internet]. 2015 Mar 30 [cited 2016 Jan 12];243:18–25. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25614384

7. Acharya UR, Molinari F, Sree SV, Chattopadhyay S, Ng KH, Suri JS. Automated diagnosis of epileptic EEG using entropies. Biomedical Signal Processing and Control. 2012;7(4):401–8.

8. Acharya UR, Vinitha Sree S, Swapna G, Martis RJ, Suri JS. Automated EEG analysis of epilepsy: A review. Knowledge-Based Systems [Internet]. 2013;45:147–65. Available from: http://dx.doi.org/10.1016/j.knosys.2013.02.014

9. Liu Y, Zheng YF. FS_SFS: A novel feature selection method for support vector machines. Pattern Recognition. 2006;39(7):1333–45.

10. Andrzejak RG, Schindler K, Rummel C. Nonrandomness, nonlinear dependence, and nonstationarity of electroencephalographic recordings from epilepsy patients. Physical Review E - Statistical, Nonlinear, and Soft Matter Physics. 2012;86(4).

11. Pincus SM. Approximate entropy as a measure of system complexity. Proceedings of the National Academy of Sciences of the United States of America [Internet]. 1991;88(6):2297–301. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11607165

12. Bai D, Qiu T, Li X. The sample entropy and its application in EEG based epilepsy detection. Sheng wu yi xue gong cheng xue za zhi = Journal of biomedical engineering = Shengwu yixue gongchengxue zazhi. 2007;24:200–5.

13. Nicolaou N, Georgiou J. Detection of epileptic electroencephalogram based on Permutation Entropy and Support Vector Machines. Expert Systems with Applications. 2012;39:202–9.

14. Costa M, Goldberger AL, Peng C-K. Multiscale entropy analysis of complex physiologic time series. Physical review letters. 2002;89(6):068102.

15. Vandewalle J. Least Squares Support Vector Machine Classifiers. Neural Processing Letters. 1999;9(3):293–300.