A Novel SVM and K-NN Classifier Based Machine Learning Technique for Epileptic Seizure Detection

https://doi.org/10.3991/ijoe.v19i07.37881

Gowrishankar K.[1(✉)], Muthukumar V.[2], Sudhakara Pandian R.[3],
Deivasigamani S.[4], Chun Kit Ang[4]
[1]Department of Marine Engineering, AMET University, Tamilnadu, India
[2]Faculty of Computing & Engineering, QUEST International University, Ipoh, Malaysia
[3]School of Mechanical Engineering, Vellore Institute of Technology, Tamilnadu, India
[4]Faculty of Engineering, Technology & Build Environment, UCSI University,
Kuala Lumpur, Malaysia
gowrishankark@ametuniv.ac.in

Abstract—An EEG signal is used for capturing the signals from the brain, which helps in localization of epileptogenic region, thereby which plays a vital role for a successful surgery. The focal and non-focal signals are obtained from the epileptogenic region and normal region respectively. The localization of epileptic seizure with the help of focal signal is necessary while detecting seizures. Hence, the present article provides detailed analysis of EEG signals. The Focal and Non-focal signals are decomposed using EMD-DWT. A combination of EMD-DWT decomposition method in accordance with log-energy entropy gives an efficient accuracy in comparison to other entropy in differentiating the Focal from Non-focal signals. The extracted features are subjected to SVM and KNN classifiers whose performance will be calculated and verified with respect to accuracy, sensitivity and specificity. At the end, it will be shown that KNN produces the highest accuracy when compared to SVM classifier.

Keywords—epileptogenic, EMD-DWT, focal, log-energy entropy, non-focal

1 Introduction

Epilepsy is the fourth common disorder among the various neurological disorder affecting almost people of all ages. A spectrum condition which involves seizure of several types that varies among peoples are referred to as epilepsy. A number of methods exists for detecting epilepsy with some limitations. Epilepsy is a neurological disorder [1]–[3] which comprises of recurrent seizures, that mostly affects the Central Nervous System (CNS) of our brain. Epilepsy is otherwise known as “seizure disorder”. It disrupts the activity of the brain cells called neurons, which transmits the electrical and chemical signals. In a person with normal condition, these electrical and chemical signals acts on other neurons, muscles and glands to produce feelings, thoughts and actions;

whereas in a seizure condition, irregular firing of neurons takes place which causes excessive electrical activity resulting in loss of consciousness [4]–[7]. Epilepsy are characterized by unpredictable seizures and other causes such as head injuries, stroke, brain damage etc., as shown in the Figure 1. During the period of seizure, a person experiences abnormal behaviour, symptoms and sensations which also includes loss of consciousness. Epilepsy differs from seizure in such a way that [8] [9] a seizure comprises of a single event; whereas an epilepsy is characterized by more than 2 unprovoked seizures.

The various methods available for detection of epilepsy are MRI, PET and SPECT etc. But, [10]–[15] the accuracy levels of those methods were not up to the level; such that it could be further continued for the purpose of detection. Hence, this paved a way to a new method of detection using EEG electrodes.

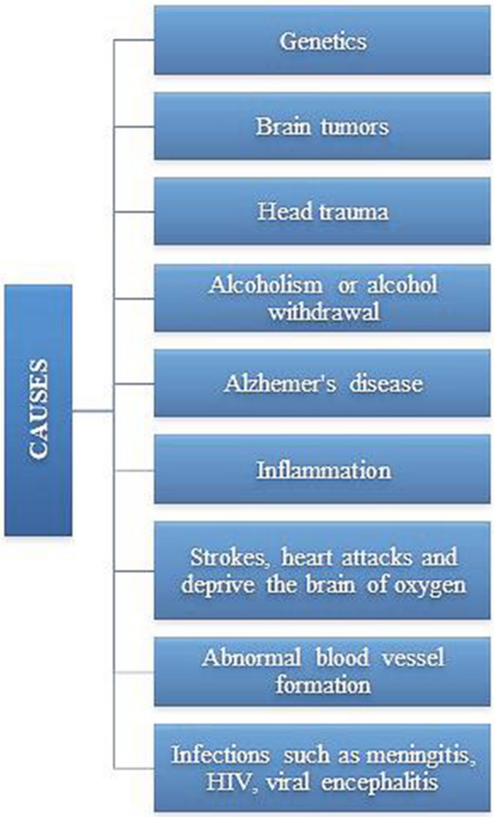


Fig. 1. Various causes of epilepsy

An electroencephalogram (EEG) electrode is a 10–20 electrode lead system which is placed on the scalp for capturing the signals from brain. The patients with epilepsy are recommended surgery as a part of curing the disorder, which are not cured by medicine. For the successful epilepsy surgery, "localization of epileptogenic region," a region that initiates the seizure; plays a vital role.The epileptogenic region is detected through analysis of EEG signals. Signals acquired from the epileptogenic region in the brain are called as the focal signal and that of obtained from the normal region of the brain are known as non-focal signal (Figure 2). The epileptogenic region in the brain, [16]–[18] found with help of focal signal is necessary for the purpose of detection of epileptic seizures. Hence, by differentiating the focal and non-focal signal, the process of detection is made serene. The Focal and Non-focal signals are decomposed using EMD-DWT. A combination of EMD-DWT decomposition method in accordance with log-energy entropy gives an efficient accuracy in comparison to other entropy in differentiating the Focal from Non-focal signals. The extracted features are subjected to SVM and KNN classifiers whose performance will be calculated and verified with respect to accuracy, sensitivity and specificity. At the end, it will be shown that KNN produces the highest accuracy when compared to SVM classifier.

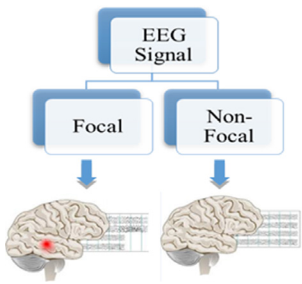


Fig. 2. Discrimination between focal and non-focal signal

2 Literature survey

Non-invasive EEG signal [19], [32] were made used in detecting epileptic seizure. It utilized wavelet decomposition for capturing morphology and spatial distribution to differentiate epileptic from non-epileptic seizure through Support Vector Machine algorithms. They tested 36 pediatric subjects which detected almost 131 of 139 seizure events and declared 15 false detection for a duration of 60 hours in clinical testing. Average sample entropy and Average variance [20] were used for classifying focal and non-focal signals. These features were made used as an input to LS-SVM classifier. In [21] Fuzzy Entropy were employed for detecting epileptic seizure. It obtained EEG signals from various epileptic states and obtained classification features which were trained and helped in classification using SVM.

In [22] EEG signals such as Focal and non-focal were classified based on Entropies. IMFs were extracted by decomposing EEG signals. Entropy features obtained from IMFs were fed to LS-SVM for classifying EEG signals. Discrimination of focal from non-focal signal [23], [31] with the help of Focal and Non-Focal Index (FNFI).

They decomposed EEG signals into 6 levels using FNFI and measured the entropy features using different classifiers such as KNN, PNN and LS-SVM. In [24] analysis of the two classes of EEG signals was done in EMD-DWT domain. Accordingly it obtained spectral entropy features which were fed as input to the KNN, primarily consists of log-energy entropy features. In [25] filters based on Discrete Fourier Transform for classifying two classes of EEG signals by computing the rhythms of EEG signals were proposed. They derived two features namely RMS and MF which acted as an input for LS-SVM for carrying out the classification of EEG signals into focal and non-focal signals.

In [26], [30] EEG signals are classified into focal and non-focal signals using IMFs which are obtained by disintegrating Electroencephalography signals through EMD were carried out. Polynomials and RBF Kernel as classifier for the purpose of classification were used.

3 Methodology

The methodology involved in localization of epileptic seizure is carried out through a series of process (Figure 3). They are carried out in 3 main process- decomposition, feature extraction and classification. The process is initiated by collecting the dataset and the one being used here is accessed through an open source, [27] "Bern-Barcelona". It consists of a collection of focal and non-focal dataset which individually consists of x and y signals. They are categorized as x, y and x-y signal of Focal and Non-focal signals. These initial processes are followed by the decomposition using Empirical Mode Decomposition (EMD) and Discrete Wavelet Transform (DWT). Further, dataset are decomposed using Empirical Mode Decomposition [28–29] so as to reduce the noise, if any present in the signals. A combination of EMD-DWT is being employed in this method as they produce a good performance in analysing non-stationary signal. By using EMD, the input signals are transformed into IMFs (Intrinsic Mode Functions). These IMFs are exported to decompose into sub-bands using DWT. The features of sub-bands are extracted into 5 levels as d1–d4 and a4 in the process of feature extraction. It involves Log-energy entropy as a main feature and outcomes are computed for every level of the sub-bands. And, once done with all these processes, the extracted features are subjected to classification where the signals are classified into focal and non-focal signals using an appropriate algorithm.

The further steps involved in classifying the EEG signals will be discussed in the following section.

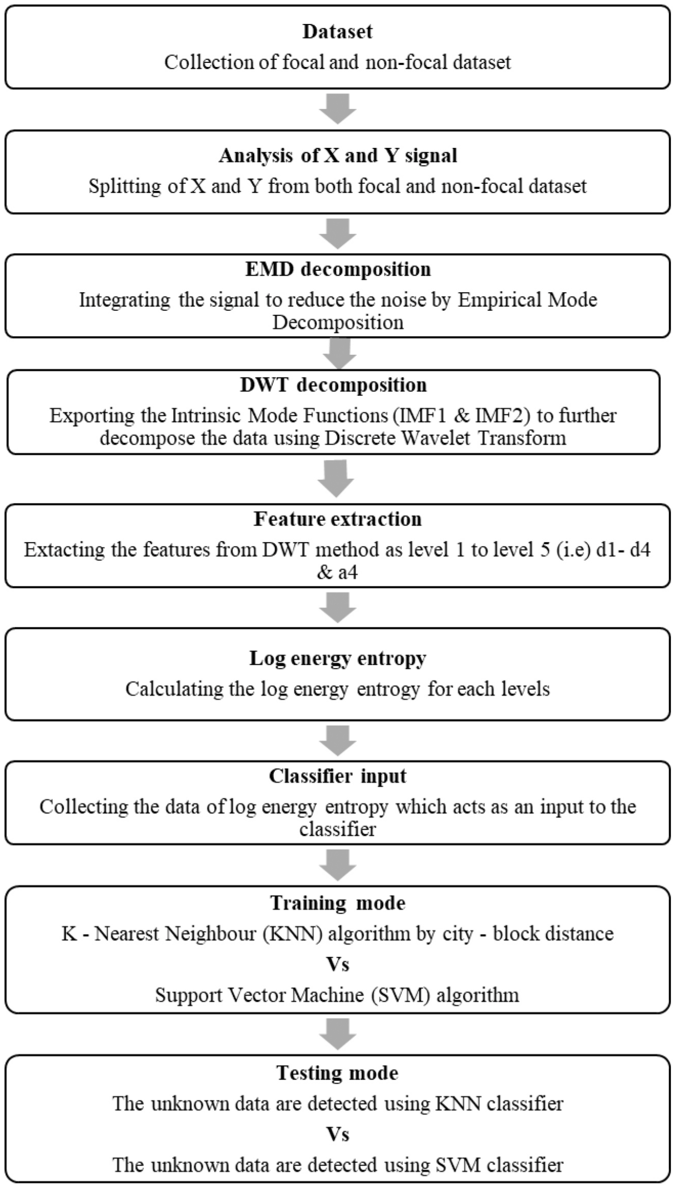


Fig. 3. Flowchart depicting the stages involved in classification

The extracted features are processed for classification which is carried out in two modes namely, training mode and testing mode. For a comparison purpose, presently 2 algorithms are made used for classifying the EEG signals. In training mode, 60 dataset comprising of focal and non-focal are trained via K-Nearest Neighbor (KNN) algorithm as well as Support Vector Machine. Later, testing dataset with the trained dataset as input are tested by KNN and SVM. Finally, their performance in terms of accuracy are compared and verified.

3.1 Steps involved in methodology

The input EEG signals are recorded using EEG electrodes as shown in Figure 4.

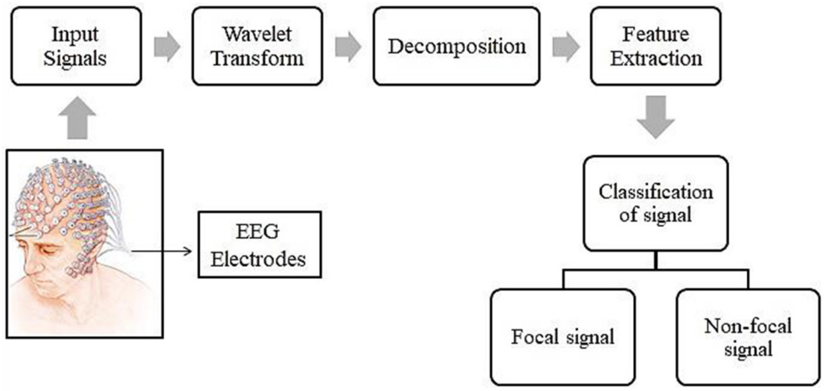


Fig. 4. Steps involved in classification of EEG signal

Dataset. The dataset from the EEG signal are collected from the open source website known as Bern Barcelona. The dataset consists of two types of signals as discussed in the above section. Individually, the focal and non-focal consist of 3750 pairs (x & y) of recorded signal. Each of 3750 pairs consists of 10,240 samples (Figure 5) recorded for a duration of 20s. Initially, 60 pairs of focal and non-focal EEG signals are utilized for feature extraction &classification.

	1	2	3	4	5
1	-23.5845	22.6215			
2	-20.1802	26.5471			
3	-16.6318	28.4316			
4	-12.7621	30.0987			
5	-10.4520	30.0309			
6	-9.7885	27.8641			
7	-9.0921	27.2800			
8	-7.9198	30.0054			

Fig. 5. Display of 10,240 samples in the workspace

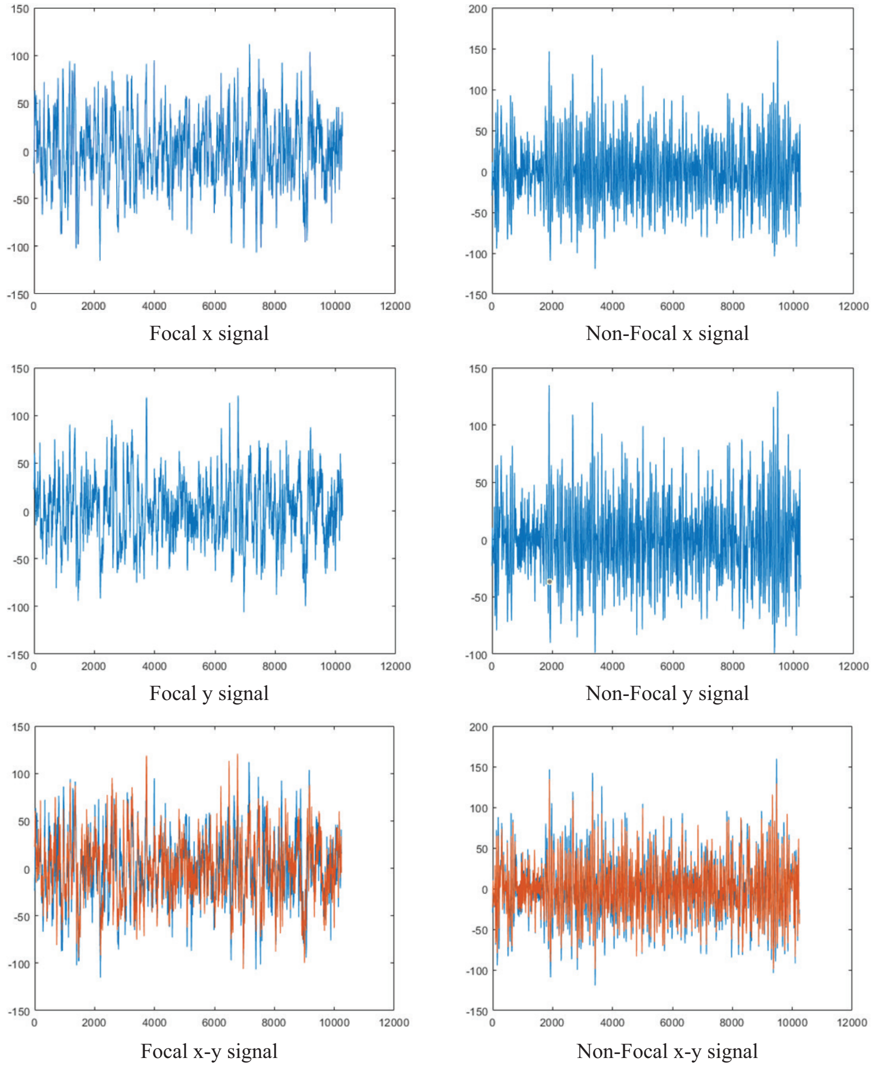


Fig. 6. Analysis of focal and non-focal signals

The x & y signals from the focal and non-focal EEG signals are separated and analysed. Later, these x & y signals are combined for analysing the EEG signals. Figure 5 depicts the x, y and x-y focal signals in the first row, whereas the second row consists of x, y and x-y non-focal signals. The difference between focal and non-focal signals can be clearly observed on seeing the Figure 6.These signals after analysing are subjected to EMD-DWT based Decomposition and then feature extraction. From thereon, the features extracted are given as input to the classifier; to classify the EEG signals into focal and non-focal signals.

EMD-DWT based decomposition. Decomposition is the main step in signal processing. It helps to integrate the high frequency signal into clear noiseless signal. It is processed by using Empirical Mode Decomposition (EMD) method. It is a method of fragmenting the signals without leaving the time domain. This method is highly suitable for the non-linear and non-randomness signals like EEG. The Fourier transform and wavelet are the other decomposition methods which are compared to the EMD decomposition.

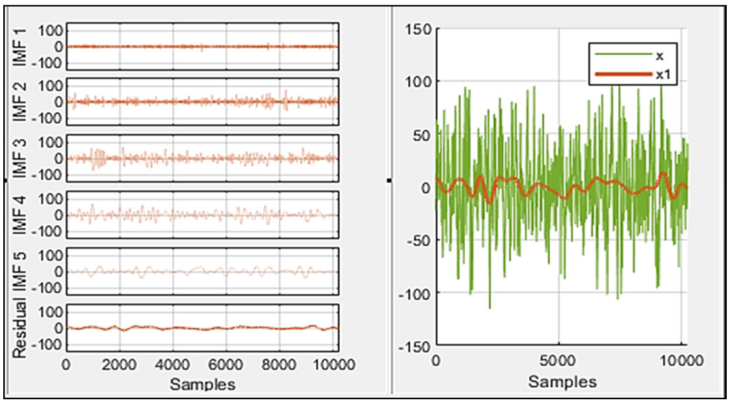


Fig. 7. Decomposition using EMD

This process is mainly used for analyzing natural signals. From the original signal, it filters out function which form a complete and nearly orthogonal basis. This way of decomposing the signal shows that EMD is a method of completeness. EMD (Empirical Mode Decomposition) transforms the input signal x(t) into a set of IMFs (Figure 7). Intrinsic Mode Function is the main function that describes the signal, even though they are not necessarily orthogonal in nature. The Intrinsic Mode Functions (IMFs) are amplitude & frequency modulated oscillatory patterns. After the decomposing both focal and non-focal signal, the IMFs of five layer are taken into matrix and further the decomposition proceeds with Discrete Wavelet Transform (DWT) method.

The extracted IMFs (IMF 1 – IMF 5) are decomposed into sub-bands (Figure 8) using DWT. The motivation behind EMD-DWT method is that it analyses the non-stationary signal in an effective manner. The Discrete Wavelet Transform is the method which also process the signal in orthogonal nature, time – frequency based method which is highly needed for the decomposition process to provide a clear signal. Discrete wavelet transform are defined by Daubechies (DB) wavelets. An analysis of orthogonal multiresolution is generated by scaling function under DWT. At last, the four level of matrix are obtained from the EMD-DWT based decomposition. Hence from this, the signal multiresolution application is named which has the both EMD and DWT as decomposition method.

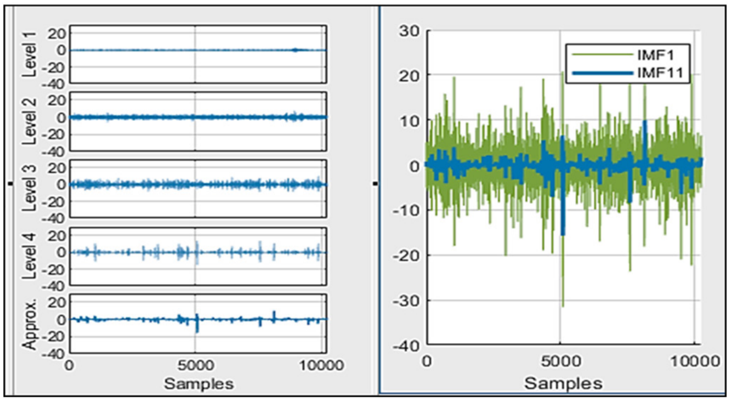


Fig. 8. Decomposition using DWT

Feature extraction. After the EMD-DWT based decomposition, the process of feature extraction are performed. The feature extraction is mainly used to reduce the number of resources which are needed for processing without losing the relevant information. It can also reduce the amount of redundant data for a given analysis.

It is the process of extracting the main features or parameters from the decomposed signal, which gives the differential values that helps in differentiating focal from the normal signal. Here, the features based on entropy are extracted. Generally, entropy is a measure of randomness which helps in understanding the dynamics of EEG signals. Basically, it measures the complexity of the signals. The log-energy entropy method of feature extraction provides the most reliable features in classifying EEG of absolute minimal error as 0.01. In particular, the machine learning algorithm is implemented to distinguish between seizure and non-seizure activity where the Log-En (log-energy entropy) values known as signal features [23],[30] by which the EEG complexity is characterized. The matrix from the decomposition process acts as an input from which the features are extracted as x and y signal. The log energy entropy value for every five level of DB are calculated using the formula as,

$$e = wentropy\,(a, \text{‘}logenergy\text{’}) \tag{1}$$

The calculated Log-energy entropy values for each and every level (d1–d4, a4) of 60 dataset are trained and tested for classifying the EEG signals. Table 1 to 6 shows the values of Log-Energy entropy for a particular set of dataset. These values are subjected to classification using SVM and KNN algorithm.

Table 1. Entropy values for focal x-signals

Sub-Bands	IMF 1	IMF 2	IMF 3	IMF 4	IMF 5
d1	–3.3124e+04	–1.1669e+05	–2.3685e+05	–4.4025e+05	–5.8935e+05
d2	–6.9048e+03	–5.8909e+04	–1.3843e+05	–2.5773e+05	–4.7574e+05
d3	–4.5030e+03	–5.4565e+03	–8.0436e+04	–1.5847e+05	–3.1935e+05
d4	–1.7489e+04	1.7361e+04	–2.2955e+04	–9.6008e+04	–1.9243e+05
a4	–2.4498e+04	1.8992e+04	4.2151e+04	4.7363e+04	4.0646e+04

Table 2. Entropy values for focal y-signals

Sub-Bands	IMF 1	IMF 2	IMF 3	IMF 4	IMF 5
d1	–3.1158e+04	–1.2088e+05	–2.3225e+05	–4.4177e+05	–5.8367e+05
d2	–6.1630e+03	–6.3243e+04	–1.3686e+05	–2.6371e+05	–4.6538e+05
d3	941.8583	–8.3406e+03	–8.0670e+04	–1.6250e+05	–3.0950e+05
d4	–1.0276e+04	1.6161e+04	–2.2693e+04	–9.8993e+04	–1.9195e+05
a4	–1.9668e+04	1.5324e+04	4.1070e+04	4.4918e+04	4.3687e+04

Table 3. Entropy values for focal x-y signals

Sub-Bands	IMF 1	IMF 2	IMF 3	IMF 4	IMF 5
d1	–6.4282e+04	–2.3758e+05	–4.6910e+05	–8.8202e+05	–1.1730e+06
d2	–1.3068e+04	–1.2215e+05	–2.7528e+05	–5.2143e+05	–9.4112e+05
d3	–3.5612e+03	–1.3797e+04	–1.6111e+05	–3.2097e+05	–6.2885e+05
d4	–2.7766e+04	3.3522e+04	–4.5648e+04	–1.9500e+05	–3.8438e+05
a4	–4.4166e+04	3.4316e+04	8.3221e+04	9.2281e+04	8.4333e+04

It can be inferred from the Tables 1, 2 and 3 that the Log-energy entropy values are quite small for focal signals. In general, IMFs 3, 4 or 5 have greater values when compared to 1 and 2. However, the values of x – y signal are smaller than that of computed individually (i.e.) x and y signal.

Table 4. Entropy values for non-focal x-signals

Sub-Bands	IMF 1	IMF 2	IMF 3	IMF 4	IMF 5
d1	–4.1127e+04	–1.1528e+05	–3.5683e+05	–5.0125e+05	–6.0078e+05
d2	–1.7153e+04	–5.5112e+04	–1.8190e+0	–3.1708e+05	–4.9587e+05
d3	–2.4201e+04	–1.0874e+04	–1.1049e+05	–2.0036e+05	–3.4230e+05
d4	–3.8016e+04	1.8257e+04	–5.5476e+04	–1.3082e+05	–2.1672e+05
a4	–4.2690e+04	5.1831e+04	4.2898e+04	2.8726e+04	2.7835e+04

Table 5. Entropy values for non-focal y-signals

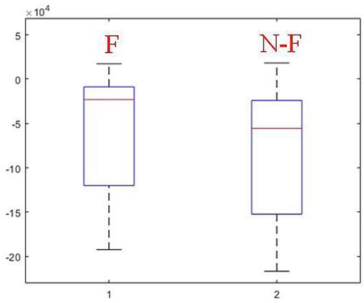
Sub-Bands	IMF 1	IMF 2	IMF 3	IMF 4	IMF 5
d1	–4.1563e+04	–1.4872e+05	–3.6380e+05	–5.1239e+05	–6.3228e+05
d2	–1.7032e+04	–7.6540e+04	–1.8857e+05	–3.3150e+05	–5.3873e+05
d3	–1.4498e+04	–2.5142e+04	–1.1606e+05	–2.0718e+05	–3.8381e+05
d4	–1.8106e+04	1.1865e+04	–6.0979e+04	–1.3453e+05	–2.4484e+05
a4	–1.8692e+04	4.8002e+04	3.6183e+04	2.9122e+04	1.7301e+04

Table 6. Entropy values for non-focal x-y signals

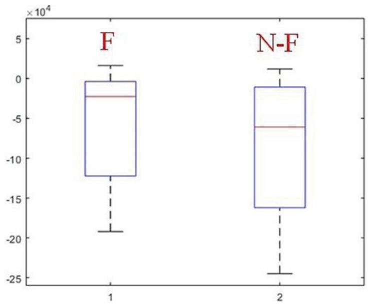
Sub-Bands	IMF 1	IMF 2	IMF 3	IMF 4	IMF 5
d1	–8.2690e+04	–2.6400e+05	–7.2062e+05	–1.0136e+06	–1.2331e+06
d2	–3.4186e+04	–1.3165e+05	–3.7047e+05	–6.4858e+05	–1.0346e+06
d3	–3.8699e+04	–3.6016e+04	–2.2654e+05	–4.0754e+05	–7.2611e+05
d4	–5.6122e+04	3.0122e+04	–1.1645e+05	–2.6535e+05	–4.6156e+05
a4	–6.1383e+04	9.9833e+04	7.9081e+04	5.7848e+04	4.5136e+04

As in the case of focal signal, the same is followed in non-focal signal. It can be inferred from the Tables 4, 5 and 6 that the Log-energy entropy values are quite small for x, y and x – y non-focal signals. In general, the values of IMFs 3, 4 or 5 have greater values compared to 1 and 2 [24]. However, the values of x – y signal are smaller than that of computed individually (i.e.) x and y signal.

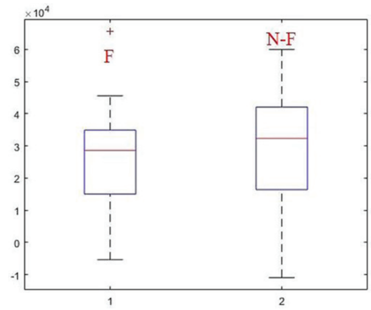
The Figure 9(a, b, c) indicates the box-plot representation of d4 sub-bands extracted from the EMD-DWT domain. Box plot represents the discriminating ability of entropy features. Here, the interquartile range, the red line indicates the median value for focal x and y signals of d4 are between 0 and –5 whereas for non-focal, the values are between –5 to –10. The log-energy values act as a feature vector to perform the classification process.



a)



b)



c)

Fig. 9. Box-plot representation of log-energy entropy (a) d4 sub-bands for x-signal (b) d4 sub-bands for y-signal (c) d4 sub-bands for x-y signal

4 Results and discussion

As discussed earlier, the process of classifying an EEG signal into focal and non-focal signal is carried out in 2 modes such as training and testing mode. Here in this paper, two classifiers have been used namely SVM classifier and KNN classifier.

4.1 SVM classifier

Figure 10 shows the Scatter Plot of SVM Classifier which represents the scattered signals on the plot. Basically, it consists of 5 predictors (d1–d4, a4) and 2 response classes (focal and non-focal). Usually they are plotted against 2 predictors and here it is plotted between d1 and d2 that can be changed according to our convenient. The incorrect values are mostly scattered out of the focus which is clearly evident from Figure 10.

Figure 11 provides a comparison with respect to accuracy among various types of SVM classifiers which includes Medium Gaussian SVM, Fine Gaussian SVM, Cubic SVM, Quadratic SVM and Coarse Gaussian SVM. It can be noted that the Fine Gaussian SVM trains the data with the highest accuracy of 62.3% when compared to other SVM algorithms.

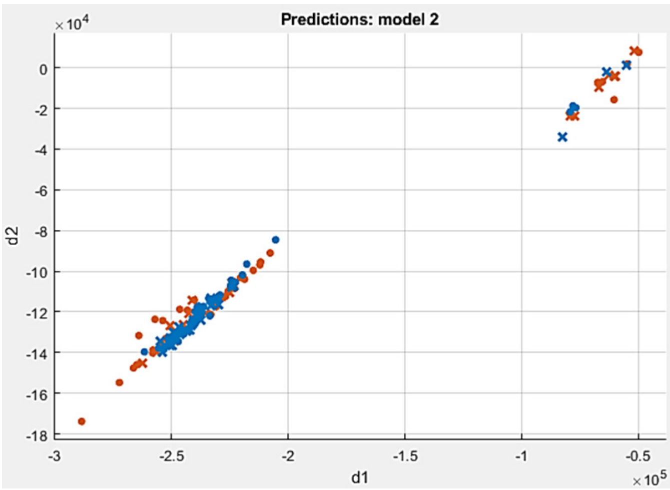


Fig. 10. Scatter plot of SVM classifier

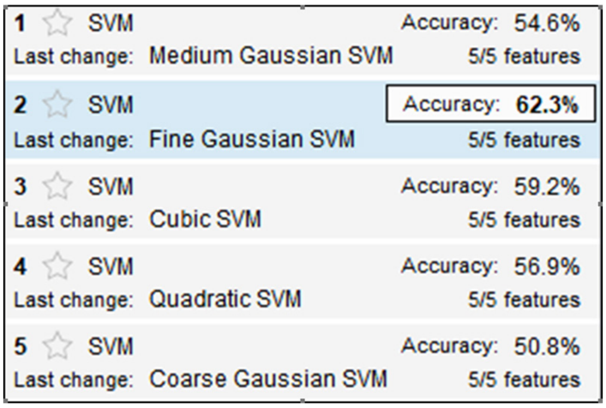


Fig. 11. Comparison of accuracy among various SVM classifiers

Fig. 12. Confusion matrix of Fine Gaussian SVM

Figure 12 shows Fine Gaussian SVM classifier values plotted between true class and predicted class of both predictors (i.e) focal and non-focal signal. It also provides the True Positive (TP) rate which defines the correct predicted rate of focal as 68% & non-focal as 57% and the False Negative (FN) rate which defines the wrong values of focal & non-focal as 32% and 43% respectively. Figure 13 shows the Region of Convergence (ROC) Curve of Fine Gaussian SVM classifier which represents the performance of a classification model. The red spot on the curve denotes the position of current classifier with False Positive rate as 0.43 towards x-axis and True Positive rate as 0.68 towards y axis.

The Area under the Curve (AUC) summarizes the ROC curve which measures the ability of a classifier to differentiate between various classes. The higher the AUC, the better will be the performance of the model at distinguishing between the positive and negative classes.

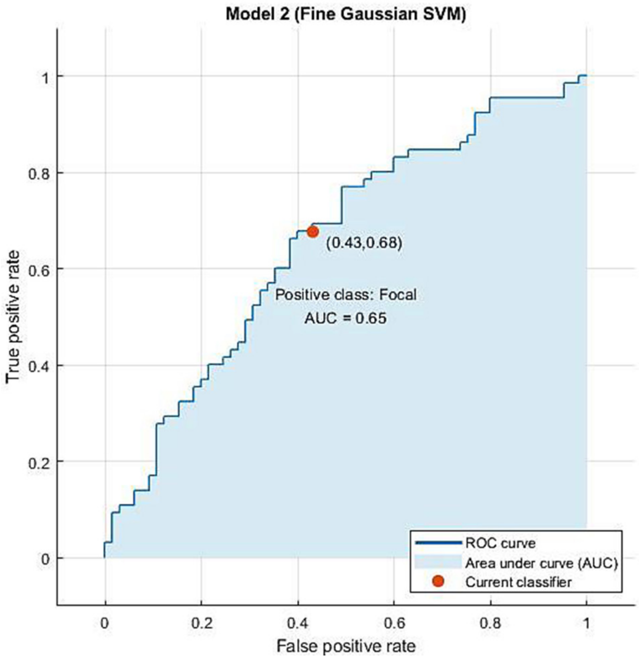


Fig. 13. ROC curve of Fine Gaussian SVM

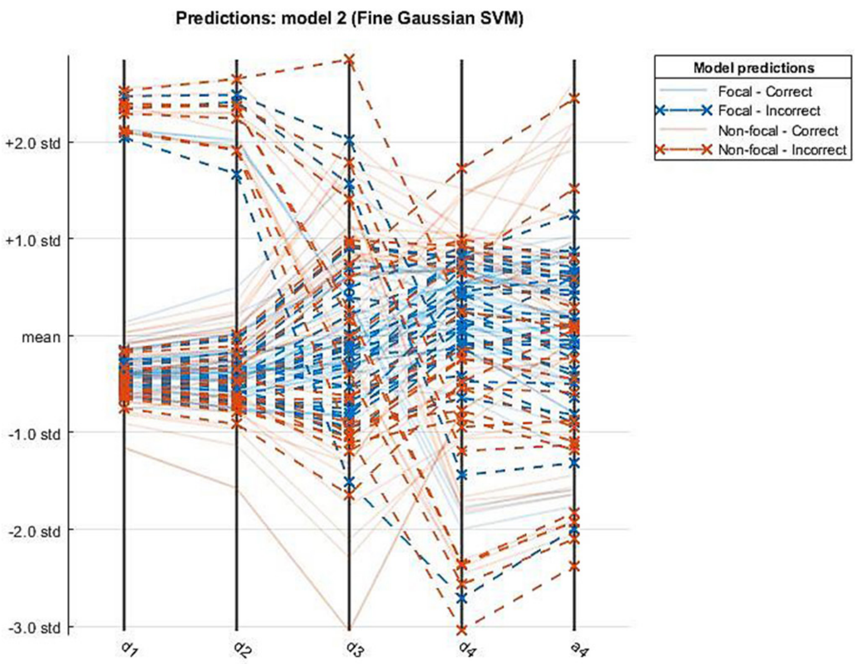


Fig. 14. Parallel coordinates plot of Fine Gaussian SVM classifier

From the Figure 14 of Parallel Coordinates plot, it is analyzed that the sub-band of d1 has true values whose values lies between +1.0 std and –1.5 std and the corrected values gradually increases at d3 sub-band. Finally, at approximation band it is shown that the corrected values are equally distributed at each range of std between +2.0 std and –2.0 std.

	1 d1	2 d2	3 d3	4 d4	5 a4
1	-64282	-13068	-3.5612e+03	-27766	-44166
2	-51706	8.1113e+03	24799	25091	17988
3	-62649	-3.8275e+03	9340	-7.5593e+03	-18455
4	-78091	-18717	-11479	-32226	-47814
5	-60080	-4.0882e+03	-14376	-46436	-55001
6	-60299	-4.3504e+03	-12875	-46481	-58666
7	-76883	-19667	-11411	-36448	-52524
8	-79259	-21619	-9054	-30293	-45828
9	-66794	-9.5612e+03	3.8337e+03	-14360	-28200
10	-77567	-23578	-22382	-52103	-65483

Fig. 15. Outcome of trained model using Fine Gaussian SVM

Figure 15 shows the output of Fine Gaussian SVM which consists of classification of focal and non-focal data using the trained model. Depending upon the outcome, their performance will be evaluated in terms of accuracy.

Table 7. Analysis of signals with respect to increasing data for SVM

No. of Testing Data	20 Data		40 Data		60 Data	
Signal	**Focal**	**Non-Focal**	**Focal**	**Non-Focal**	**Focal**	**Non-Focal**
Total	10	10	20	20	30	30
True value	2	4	7	12	14	21
False value	8	6	13	8	16	9

For testing, the data were gradually increased and with an increase in data at every stage, the capability of recognizing the true values increased. Initially, 20 data were taken and subsequently were increased in steps of 20. At a certain stage, when 60 data were taken, the values got saturated and became constant. Table 7 tabulates the analysis made while increasing data.

Figure 16 provides a graphical representation that gives a clear idea how their performance got enhanced with an increase in data. Their performance are calculated and tabulated in Table 9. By using SVM classifier, it obtained an accuracy of 58.33%.

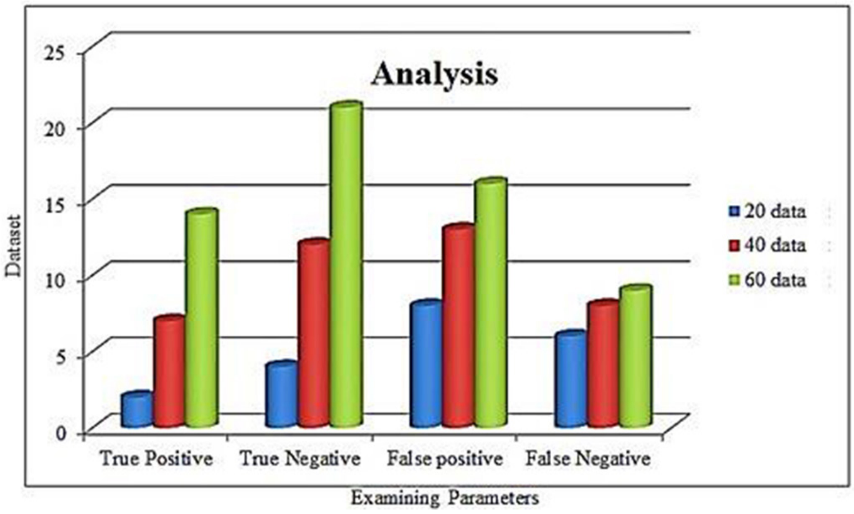


Fig. 16. Depicting the analysis of signals with an increase in data for SVM

4.2 KNN classifier

Figure 17 shows the Scatter Plot of KNN Classifier which represents the scattered signals on the plot. Basically, it consists of 5 predictors (d1–d4, a4) and 2 response classes (focal and non-focal). Usually they are plotted against 2 predictors and here it is plotted between d1 and d2 that can be changed according to our convenient. The incorrect values are mostly scattered out of the focus which is clearly evident from Figure 17.

Figure 18 provides a comparison with respect to accuracy among various KNN classifiers which includes Medium KNN, Fine KNN and Cubic KNN. It can be noted that the Medium KNN trains the data with the highest accuracy of 54.6% when compared to other KNN algorithms.

Figure 19 shows Medium KNN classifier values plotted between true class and predicted class of both predictors (i.e) focal and non-focal signal. It also provides the True Positive (TP) rate which defines the correct predicted rate of focal as 72% & non-focal as 37% and the False Negative (FN) rate which defines the wrong values of focal & non-focal as 28% and 63% respectively. Figure 20 shows the Region of Convergence (ROC) Curve of Medium KNN classifier which represents the performance of a classification model. The red spot on the curve denotes the position of current classifier with False Positive rate as 0.63 towards x-axis and True Positive rate as 0.72 towards y axis.

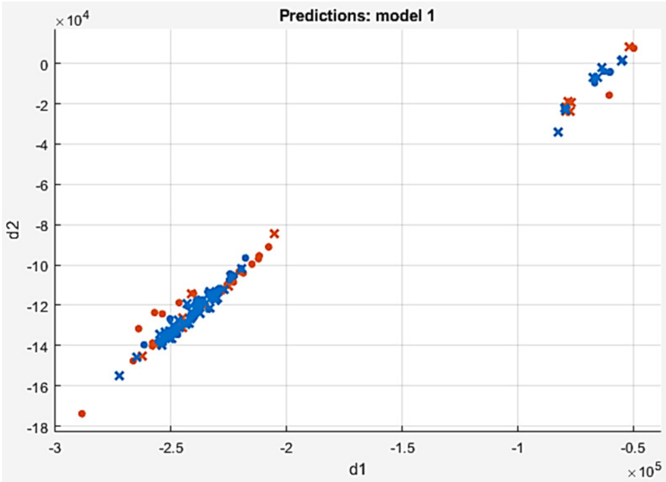


Fig. 17. Scatter plot of KNN classifier

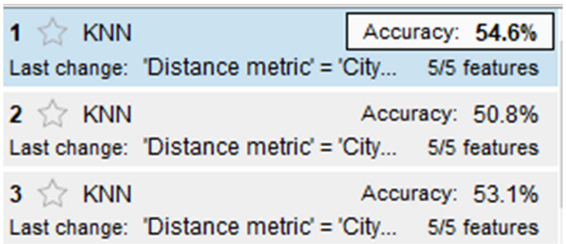


Fig. 18. Comparison of accuracy among various KNN classifier

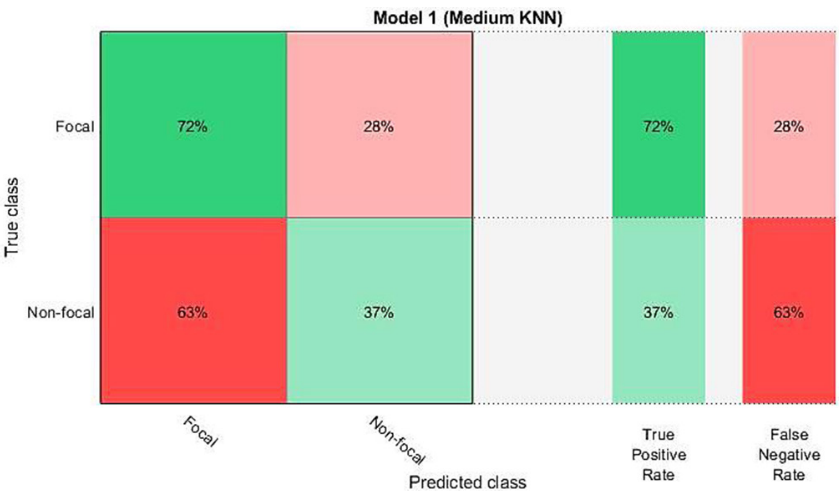


Fig. 19. Confusion matrix of medium KNN

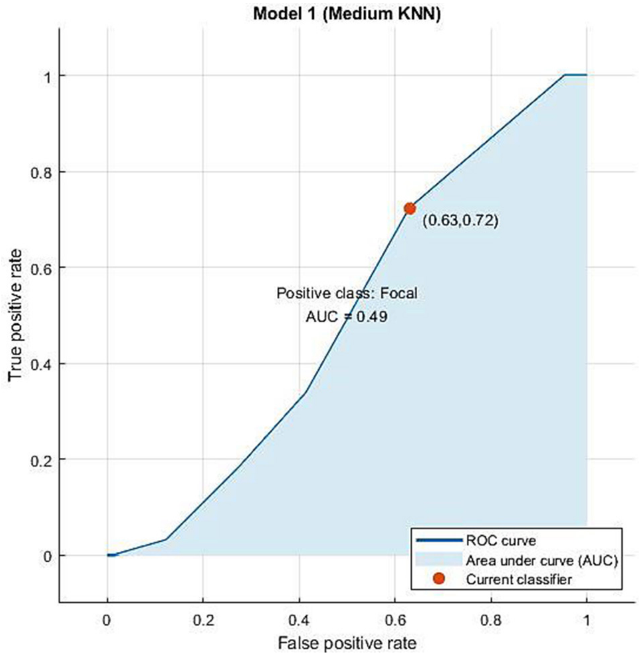


Fig. 20. ROC curve of medium KNN

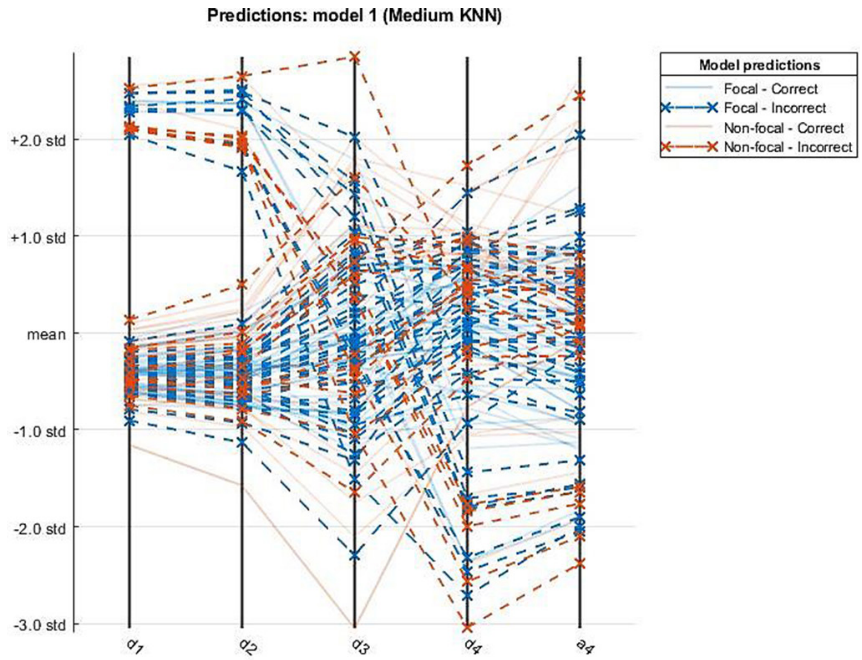


Fig. 21. Parallel coordinates plot of medium KNN classifier

From the Figure 21 of Parallel Coordinates plot, it is analyzed that the sub-band of d1 has true values whose values lies between +0.2 std and –1.0 std and the corrected values gradually increases at d3 sub-band. Finally, at approximation band it is shown that the corrected values are equally distributed at each range of std between +1.2 std and –2.0 std.

For testing, the data were gradually increased and with an increase in data at every stage, the capability of recognizing the true values increased. Initially, 20 data were taken and subsequently were increased in steps of 20. At a certain stage, when 60 data were taken, the values got saturated and became constant. Table 8 tabulates the analysis made while increasing data. Figure 22 provides a graphical representation that gives a clear idea how their performance got enhanced with an increase in data. Their performance are calculated and tabulated in Table 9. By using KNN classifier, it obtained an accuracy of 75%.

Table 8. Analysis of signals with respect to increasing data for KNN

No. of Testing Data	20 Data		40 Data		60 Data	
Signal	**Focal**	**Non-Focal**	**Focal**	**Non-Focal**	**Focal**	**Non-Focal**
Total	10	10	20	20	30	30
True value	3	6	10	13	21	24
False value	7	4	10	7	9	6

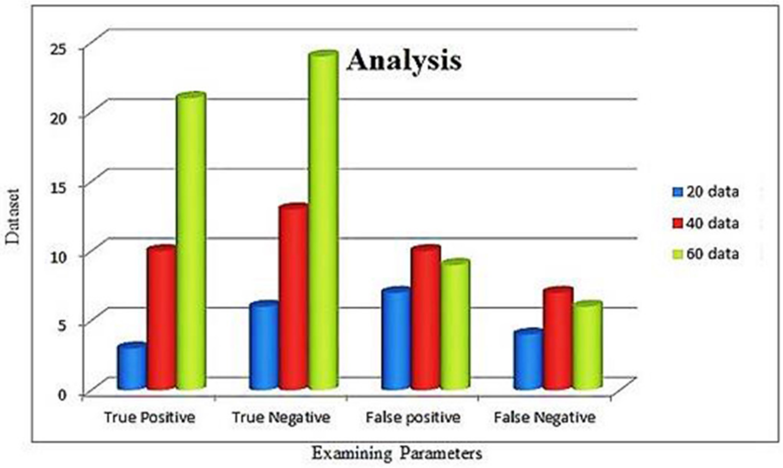


Fig. 22. Depicting the analysis of signals with an increase in data for KNN

Comparison between SVM and KNN classifier

Case 1: Comparison between SVM and KNN for 20 data

Initially, when 20 data were taken, it correctly predicted the focal (i.e) True Positive (TP) as 2 in the case of SVM whereas 3 in the case of KNN (Figure 23).

Case 2: Comparison between SVM and KNN for 40 data

Later, when 40 data were taken, it correctly predicted the focal (i.e) True Positive (TP) as 7 in the case of SVM whereas 10 in the case of KNN (Figure 24).

Case 3: Comparison between SVM and KNN for 60 data

Finally, when 60 data were taken, it correctly predicted the focal (i.e) True Positive (TP) as 14 in the case of SVM whereas 21 in the case of KNN (Figure 25).

Hence, from all the 3 cases mentioned above, it is understood that the number of data plays a crucial role in this process of classification. It is also inferred that, in all the 3 cases, KNN predicts the highest number of TP value.

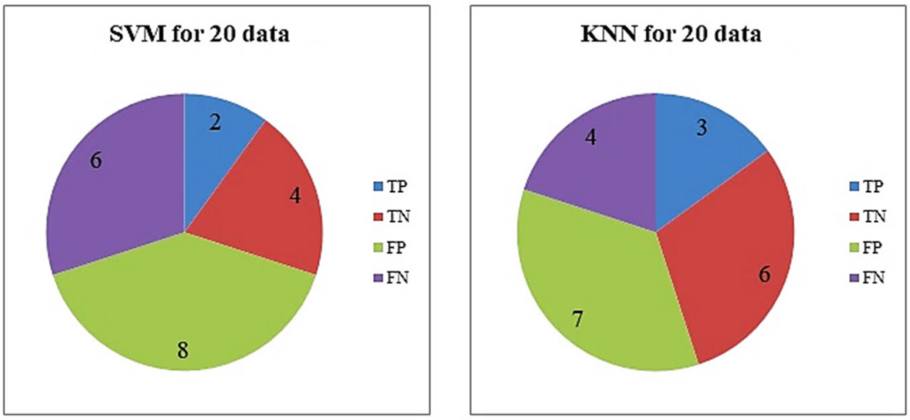


Fig. 23. Comparison between SVM and KNN for 20 data

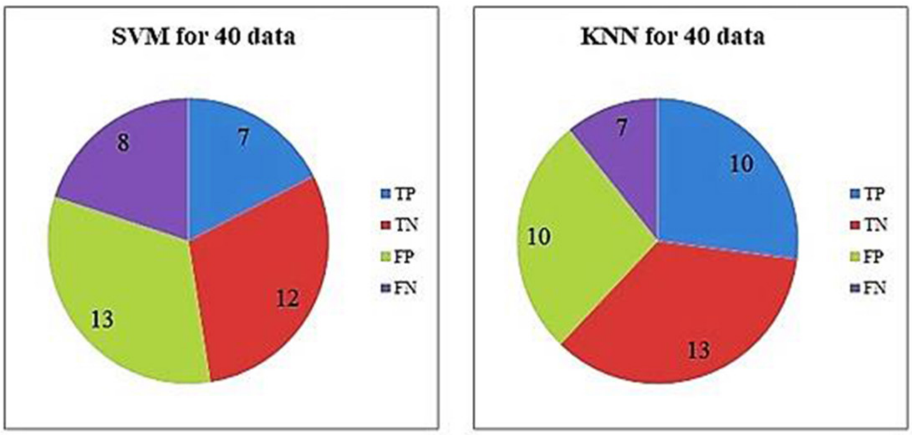


Fig. 24. Comparison between SVM and KNN for 40 data

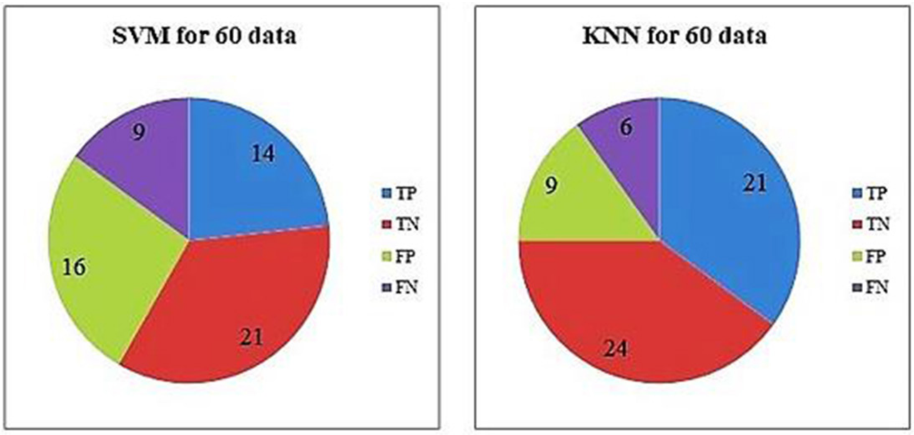


Fig. 25. Comparison between SVM and KNN for 60 data

The obtained values of True Positive (*TP*), True Negative (*TN*), False Positive (*FP*) and False Negative (*FN*) for both SVM and KNN can be noted from Table 7 and 8 respectively. Based upon *TP*, *TN*, *FP* and *FN*, their performance in terms of accuracy, sensitivity and specificity are calculated using the below formulae and are tabulated in Table 9.

ACCURACY:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%$$

For SVM,

$$\text{Accuracy} = \frac{14 + 21}{14 + 21 + 16 + 9} \times 100 = 58.33\%$$

For KNN,

$$\text{Accuracy} = \frac{21 + 24}{21 + 24 + 9 + 6} \times 100 = 75\%$$

SENSITIVITY:

$$Sensitivity = \frac{TP}{TP + FN} \times 100\%$$

For SVM,

$$\text{Sensitivity} = \frac{14}{14 + 9} \times 100 = 60.87\%$$

For KNN,

$$\text{Sensitivity} = \frac{21}{21+26} \times 100 = 77.77\%$$

SPECIFICITY:

$$Specificity = \frac{TN}{TN+FP} \times 100\%$$

For SVM,

$$\text{Specificity} = \frac{21}{21+16} \times 100 = 56.76\%$$

For KNN,

$$\text{Specificity} = \frac{24}{24+9} \times 100 = 72.73\%$$

Table 9. Performance of SVM and KNN classifier

Parameters	Classifier	
	SVM	KNN
Accuracy	58.33%	75%
Sensitivity	60.87%	77.78%
Specificity	56.76%	72.73%

The above tabulated performance of SVM and KNN classifier is graphically represented in Figure 26 for an easy understanding. Thus from Table 9 and Figure 26, it is affirmed that KNN provides the highest accuracy of 75%, sensitivity of 77.78% and specificity of 72.73%.

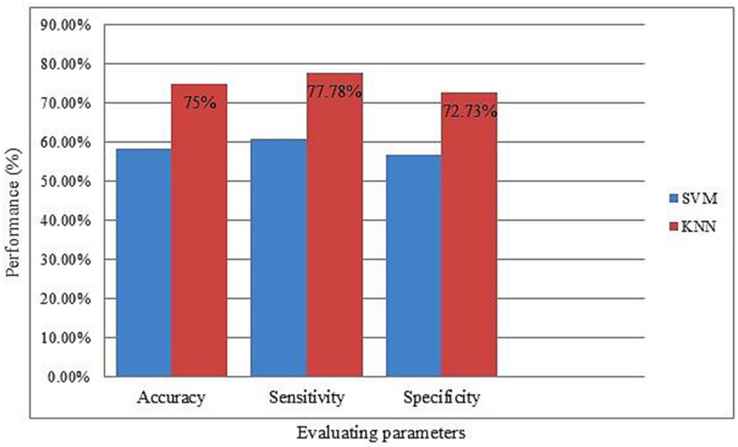


Fig. 26. Depicting the comparison between SVM and KNN in terms of evaluating parameters

5 Conclusion

In this paper, the classification of EEG signals into focal and non-focal signal were carried out through KNN method whose performance was calculated and verified in terms of their accuracy, sensitivity and specificity. It was also inferred that with an increase in data at each and every point, the performance parameters were enhanced and later got saturated after a certain specific point. Further, KNN classifier obtained the highest accuracy of 75%, sensitivity of 77.78% and specificity of 72.73% while SVM classifier obtained accuracy of 58.33%, sensitivity of 60.87%, and specificity of 56.76%. Thus it can be stated that KNN classifier provided the highest accuracy when compared to SVM classifier.

6 References

[1] R.B. Pachori (2008), Discrimination between ictal and seizure-free EEG signals using empirical mode decomposition – Research Letter Signal Processing, 2008, 1–5. https://doi.org/10.1155/2008/293056
[2] A.B. Das, M.I.H. Bhuiyan (2014), A sub-band correlation-based method for the automatic detection of epilepsy and seizure in the dual tree complex wavelet transform domain. IEEE Conference on Biomedical Engineering and Sciences, Sarawak, Malaysia, December 8–10, 2014, pp. 811–816.
[3] R.B. Pachori, V. Bajaj (2013), Epileptic seizure detection based on the instantaneous area of analytic intrinsic mode functions of EEG signals. Biomedical Engineering Letters, 3, 17–21. https://doi.org/10.1007/s13534-013-0084-0
[4] A.G. Correa, L. Orosco, A. Torres, J.P. Graffigna, E. Laciar (2009), An epileptic seizures detection algorithm based on the empirical mode decomposition of EEG. IEEE Conference on Engineering in Medicine and Biology Society (EMBS), Minnesota, USA, September 2–6, 2009, pp. 2651–2654.
[5] R.B. Pachori, V. Bajaj (2011), Application of the sample entropy for discrimination between seizure and seizure free EEG Signals. Indian International Conference on Artificial Intelligence, Bangalore, India, December 14–16, 2011, pp. 1232–1247.
[6] V. Bajaj, R.B. Pachori (2011), Analysis of normal and epileptic seizure EEG signals using empirical mode decomposition. Computer Methods and Programs in Biomedicine, 104, 373–381. https://doi.org/10.1016/j.cmpb.2011.03.009
[7] D.P. Mandic, Y. Xia (2010), Application of multivaritate empirical mode decomposition for seizure detection in EEG signals. Annual International Conference of IEEE Engineering in Medicine and Biology Society, Buenos Aires, Argentina, August 31–September 04, 2010, pp. 1650–1653.
[8] M.I.H. Bhuiyan, S.M. Shafiul Alam (2011), Detection of epileptic seizures using chaotic and statistical features in the EMD domain. Annual IEEE Conference (INDICON), Hyderabad, India, December 16–18, 2011, pp. 1–4.
[9] M.I.H. Bhuiyan, S.M. Shafiul Alam (2013), Detection of seizure and epilepsy using higher order statistics in the EMD domain. IEEE Transactions on Information Technology in Biomedicine, 17, 312–318. https://doi.org/10.1109/JBHI.2012.2237409
[10] F. Bahari, A. Janghorbani (2013), EEG-based emotion recognition using recurrence plot analysis and K nearest neighbor classifier. 20th Iranian Conference on Biomedical Engineering (ICBME), Tehran, Iran, December18–20, 2013, pp. 228–233. https://doi.org/10.1109/ICBME.2013.6782224

[11] Md. Ashfanoor Kabir, C. Shahnaz (2012), Denoising of ECG signals based on noisereduction algorithms in EMD and wavelet domains. Biomedical Signal Processing and Control, 7, 481–489. https://doi.org/10.1016/j.bspc.2011.11.003
[12] S. Chen, Y. Liu, Q. Yuan, W. Zhou (2012), Automatic seizure detection using wavelet transformation and SVM in long term intracranial EEG. IEEE Transactions on Neural System and Rehabilitation Engineering, 6, 749–754. https://doi.org/10.1109/TNSRE.2012.2206054
[13] R.B. Pachori, Bajaj (2012), Classification of seizure and non-seizure EEG signals using empirical mode decomposition. IEEE Transactions on Information Technology in Biomedicine, 16, 1135–1142. https://doi.org/10.1109/TITB.2011.2181403
[14] Shou-Zen Fan, Jeng-Rung Huang, Kuo-Kuang Jen, Maysam F. Abbo, Jiann-Shing Shieh Jeng-Fu Wu (2013), Application of multivariate empirical mode decomposition and sample entropy in EEG signals via artificial neural networks for interpreting depth of anesthesia. Entropy, 15, 3325–3339. https://doi.org/10.3390/e15093325
[15] P.S. Khobragade, R. Panda, P.D. Jambhule, P.R. Pal, S.N. Jengthe, T.K. Gandhi (2010), Classification of EEG signal using wavelet transform and support vector machine for epileptic seizure prediction. IEEE Systems in Medicine and Biology (ICSMB), 405–408.
[16] R.B. Pachori, V. Bajaj (2011), EEG signal classification using empirical mode decomposition and support vector machine. International Conference on Soft Computing for Problem Solving, Roorkee, India, December 20–22, 2011, pp. 623–635. https://doi.org/10.1007/978-81-322-0491-6_57
[17] S. Patidar, R.B. Pachori (2014), Epileptic seizure classification in EEG signal using second-order difference plot of intrinsic mode function. Computer Methods and Programs in Biomedicine, 113, 494–502. https://doi.org/10.1016/j.cmpb.2013.11.014
[18] G.F. Boudreaux Bartels, W. Besio, Y. Boudria, A. Feltane (2013), Seizure detection using empirical mode decomposition and time-frequency energy concentration. IEEE Signal Processing in Medicine and Biology Symposium (SPMB), New York, USA, December 07, 2013, pp. 1–6.
[19] H. Edwards, J. Connolly, H. Shoeb, B. Bourgeois, J. Guttag, S.T. Treves (2004), Patient-specific seizure onset detection. Epilepsy and Behavior, 5, 483–498. https://doi.org/10.1016/j.yebeh.2004.05.005
[20] S. Gautam, R. Sharma, R.B. Pachori (2014), Empirical mode decomposition-based classification of focal and non–focal EEG signals. International Conference on Medical Biometrics, Shenzhen, China, May 30–June 01, 2014, pp. 135–140.
[21] X. Han, J. Xiang, H. Li, C. Li, C.R.B. Wang, J. Chen (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
[22] U.R. Acharya, R.B. Pachori, R. Sharma (2015), Application of entropy measures on intrinsic mode functions for the automated identification of focal electroencephalogram signals. Entropy, 17, 669–691. https://doi.org/10.3390/e17020669
[23] M.I.H. Bhuiyan, A.B. Das (2016), Discrimination and classification of focal and non-focal EEG signals using entropy-based features in the EMD-DWT domain. Biomedical Signal Processing and Control, 29, 11–21. https://doi.org/10.1016/j.bspc.2016.05.004
[24] R.B. Pachori, P. Singh (2017), Classification of focal and non-focal EEG signals using features derived from Fourier-based rhythm. Journal of Mechanics in Medicine and Biology, 17, 1–16. https://doi.org/10.1142/S0219519417400024
[25] V. VijayaBaskar, R. Krishnaprasanna (2017), Focal and non-focal EEG signal classification by computing area of 2D-PSR obtained for IMF. Journal of ICT Standardization, 5, 171–186. https://doi.org/10.13052/jicts2245-800X.523

[26] K. Schindler, C. Rummel, R.G. Andrzejak (2012), Nonrandomness, nonlinear dependence, and non-stationarity of electroencephalographic recordings from epilepsy patients. Physical Review E, 86, 1–7. https://doi.org/10.1103/PhysRevE.86.046206
[27] C. Kiran Kumari, P. Preethi, M. Revathi, A. Shakiya Farhana, K. Gowrishankar (2020), Detection of epileptic seizure: A survey. Journal of Critical Reviews, 7, 4370–4379.
[28] S. Deivasigamani, C. Senthilpari, W.H. Yong (2021), Machine learning method based detection and diagnosis for epilepsy in EEG signal. Journal of Ambient Intelligent Human Computing, 12, 4215–4221. https://doi.org/10.1007/s12652-020-01816-3
[29] S. Deivasigamani, C. Senthilpari, W.H. Yong (2016), Classification of EEG focal and nonfocal EEG signals using ANFIS classifier for epilepsy detection, International Journal Imaging System and Technology, 26, 277–283. https://doi.org/10.1002/ima.22199
[30] A. Rizal, W. Priharti, S. Hadiyoso (2021), Seizure detection in epileptic EEG using short-time fourier transform and support vector machine. International Journal of Online and Biomedical Engineering (iJOE), 17, pp. 65–78. https://doi.org/10.3991/ijoe.v17i14.25889
[31] J. Katona, A. Kovari (2015), EEG-based computer control interface for brain-machine interaction. International Journal of Online and Biomedical Engineering (iJOE), 11, pp. 43–48. https://doi.org/10.3991/ijoe.v11i6.5119
[32] T. Najafi, R. Jaafar, R. Remli, W.A. Wan Zaidi, K. Chellappan (2022). Brain dynamics in response to intermittent photic stimulation in epilepsy. International Journal of Online and Biomedical Engineering (iJOE), 18, pp. 80–95. https://doi.org/10.3991/ijoe.v18i05.27647

7 Authors

Dr. K. Gowrishankar working as a Professor in the Department of Marine Engineering, AMET deemed to be University, Chennai, Tamilnadu, India. He completed his doctoral degree from University Tenaga Nasional, Malaysia. And master's degree in the field of Control Systems and Instrumentation Engineering from SASTRA University, Tamilnadu, India. He is a professional member of various bodies IEEE, IE (India), ISTE, and IAENG. He published around 50 articles in reputed journals and conferences. He filed 9 patents. His area of interest includes biomedical engineering, industrial automation, control system, optimization algorithms.

Mr. V. Muthukumar received the master's in engineering specialized in Electronics Engineering degree from the College of Engineering-Guindy, Anna University, Chennai. As an academia for more than 2 decades, he involved in community-related engineering projects and other research projects as well. He is currently pursuing his PhD in the area of Machine learning. His research interests include Bio-MEMS, Microelectronics, Biosignal processing, and devices for healthcare applications.

Dr. R. Sudhakara Pandian is working as a Professor of School of Mechanical Engineering, Vellore Institute of Technology, Vellore. He has done his PhD in NIT Rourkela, India. He has 22 years of Teaching and 12 years of Research Experiences. He has published 30 International Journals and 28 International Conferences. He is a Co-Editor of a Book published by IGI-Global USA. He has done his post-doctoral work at Technical university of Kosice, Slovakia in 2009-10 and 2018. He is working on a SME 4.0 project funded by EU H2020. He has acted as session chairs for 10 International Conferences. He is an active Reviewer of Elsevier, MDPI and Hindawi Journals.

Dr. S. Deivasigamani received his B.Eng. in Electrical & Electronics Engineering, Thiagarajar College of Engineering, M.Eng in Applied Electronics Engineering, Thanthai Periyar Government Institute of Technology from the University of Anna, India, Ph.D. in Engineering (Medical Signal Processing using Machine Learning Methods), Multimedia University, Malaysia. Currently working as an Assistant Professor at the Faculty of Engineering, Technology & Built Environment, UCSI University, Malaysia. He served in various academic positions such as Deputy Dean, HOD, and Programme Co-ordinator. He has to date, published over 35 scientific articles in international journals and conferences. His current research expertise and interest areas include Medical Signal and Image Processing and VLSI applications.

Ir. Ts. Dr. Chun Kit Ang received Bachelor of Engineering Degree with First Class Honors in Mechatronic Engineering from UCSI University, Malaysia in 2010. In the end of 2010, he pursued his PhD in Mechanical Engineering (research in robotics and Artificial Intelligence) in Universiti Putra Malaysia (UPM) and graduated in 2014. Currently, he is the Dean for Faculty of Engineering and Built Environment, UCSI University. He has published more than 30 indexed journal articles within these 5 years and secured grants with more than RM 250, 000 in total. He was also appointed as editorial board members and committee members for more than 5 conferences in the past 2 years. Lastly, one of his latest best achievements in year 2020 was that managed to secure a Fundamental Research Grant Scheme (FRGS) and Post Covid-19 Grant from Ministry of Education (MoE).

Article submitted 2023-01-23. Resubmitted 2023-04-07. Final acceptance 2023-04-09. Final version published as submitted by the authors.