4

Predicting Epilepsy Seizures Using Machine Learning and IOT

Bahubali Shiragapur, Tanuja S. Dhope (Shendkar), Dina Simunic, Vijayalaxmi Jain and Nishikant Surwade

CONTENTS

4.1 Introduction

As per the World Health Organization (WHO) survey, around 60 million people are affected by epilepsy diseases [1]. This is a brain disorder which needs to predict occurrences to prevent life-threatening situations. Epileptic seizures are a result of sudden changes in electroencephalogram (EEG) signals reflected as transient high peaks in EEGs [2]. The EEG signal is widely used for clinical assessments for measuring brain activity and detection of seizures. Traditional, visual scanning of a patient's EEG data is a tedious and time-consuming process. Thus, there is a need for a reliable and automated system to predict, classify and detect epilepsy, for better treatment. Furthermore, this kind of system can reduce the effects of long-term treatment with antiepileptic drugs which are harmful to the human neurological system. Thus, there is a need identify a better predictive system that will help to doctors to make decisions and will reduce clinical observation errors.

A seizure represents a single occurrence event in EEG. However, epilepsy is defined as a neurological state characterized by two or more unprovoked seizures. Apart from various types of seizures, most common seizures can be classified as generalized, or in the form known as focal [3]. During focal (or partial) seizures, the seizure activity is restricted to a portion of one brain hemisphere. The seizure begins in a part of the brain. Focal seizures are of two types: A simple partial seizure is a focal seizure with retained awareness. A complex partial seizure or focal dyscognitive seizure is a seizure with loss of awareness.

The EEG is a biomedical clinical tool used to predict human brain abnormalities. Furthermore, the use of better state-of-the-art techniques, such as the 10–20 international system, are used to record brain activity [11,12]. This multichannel time variant signal can be further analyzed by using a signal processing tool to extract different signal features. It is possible to detect and predict epilepsy. This study's primary goal is to detect and predict by the use of a machine learning approach and to process the data by Internet of Things (IoT) devices. Furthermore, this chapter may motivate the research group to solve the societal problem.

We have processed the publicly available EEG signals of normal as well as epileptic disorder patients using Chebyshev filter wavelet analysis, and extracted the features using wavelet decomposition that captured the frequency of the dataset. The seven different techniques, including LPC (linear predictive coding), kurtosis, mean, auto-correlation, skewness, spectral energy and feature extraction is done by using PCA (Principal Component Analysis). The state-of-the-art methods such as the Artificial Neural Network (ANN) are used to make decisions on medical events whether the EEG signal is free from epileptic seizure or not.

4.2 Literature Survey

A set of several unique features can be used for predicting the preictal state of epileptic seizures. Rasekhi et al. [5] have proposed univariate linear feature detection for seizure prediction; 132-dimensional feature space has been used by utilizing six EEG channels and extracting various univariate linear properties. The work suggests that the "preictal time" begins 10 to 40 minutes ahead of the "ictal" state, with a difference of 10 minutes. "Preictal" and "ictal" states are considered as a binary classification tests to predict an epileptic disorder. They reported that prediction sensitivity is 73.9%. They used a Support Vector Machine (SVM) for the classification of EEG signals "preictal" and "ictal." The authors suggested the use of univariate linear features with a fixed window size, with certain regularization decisions on EEG signals. They have considered 5-second segmented data for further processing and filtering to reduce noise.

Teixeira et al. [6] have developed a technique for predicting the epileptic seizure in real time. The medical event prediction is based on machine learning classification methods, viz. SVM, radial basis functions neural networks (RBF) and Multilayer Perception Neural Networks (MLP). The authors [6] have suggested filtering techniques as a preprocessing step, which can be used for removal of artifacts. They suggested the SVM classifier can achieve seizure detection sensitivity of 75.8% [8]. The use of the wavelet signal processing method for prediction of seizures has been suggested in Gadhoumi et al. [9]. The wavelet energy and wavelet entropy are considered as features to train the neural network. For testing, six patients' data was selected from two or three channels. The reported sensitivity is 88%.

Zandi et al. [10] have also suggested a model for prediction of seizures based on zero crossing by scalp EEG signals. The computation of histogram for all intervals in moving the average window for observations at different sets of points (preictal and interictal). The work suggests an alarm is created at the start or beginning of the "preictal" state of the seizure, which predicts seizures. The suggested model provides a sensitivity of 88.3%, with a predicted time of 22.5 minutes.

The authors [12] suggested that statistical moments, the first four as features, are extracted. These features are used to measure the variance, similarity and symmetry of successive EEG signal samples. As the EEG is a non-stationary signal, it is necessary to eliminate noise smoothing of EEG signal for better sensitivity and performance analysis of epilepsy EEG dataset signals.

The authors [13] have shown the experiment for detection of epileptic seizure by using filtering and wavelet transformation techniques. They [14] have presented a classification of sleep stages operating on wavelet transformation and neural networks. The detailed and db-4 approximation coefficients and back propagation algorithm EEG are used to train the neural network model for classifying the stages of sleep signals. The authors [16] have proposed the use of lower devices MSP-432 for seizure detection. Another research group [17] discussed accurate detection of epilepsy with a 10-second prediction time well before the occurrence of the medical event.

4.3 Proposed Approach

Figure 4.1 shows the block diagram of the proposed approach.

4.3.1 EEG Dataset

A public dataset [15] containing five EEG sets are used for experimentation. The details of the dataset are as follows:

- Signal recording segments: 100
- Sampling frequency: 1,000 Hz
- Duration: 23.6 seconds of every segment

For experimentation we have used only two sets: A, E.

Healthy patients' EEG signals are in Set-A, whereas epileptic patients' EEG signals during epilepsy seizures are in Set-E.

4.3.2 Preprocessing

The Chebyshev filtering method is used to remove the artifacts of EEG signal as it is a non-stationary signal. This will enhance the prediction and detection of the epilepsy seizure event.

4.3.3 Decomposition

Decomposition is a sub-band coding that uses Daubechies wavelet families to decompose the signal for analysis of low frequency band signal. For experimentation, the db-4 method is used for analysis of the EEG dataset.

4.3.4 Feature Extraction

Various statistical methods are considered, such as LPC, kurtosis, mean, auto-correlation and PCA. These features are considered to train the network for further classification.

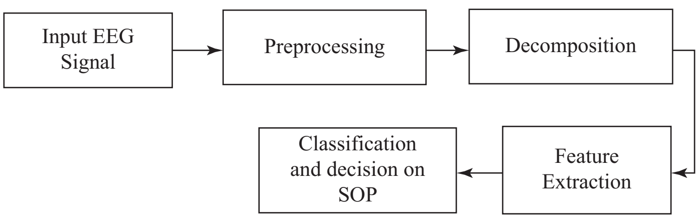

FIGURE 4.1
Block diagram of the proposed system.

4.4 Linear Predictive Coding (LPC)

The LPC coefficients were computed for the non-stationary EEG dataset. Thirteen LPC coefficients are considered for simulation. In linear prediction, the present EEG samples are approximated to the linear combination of past samples, that is:

$$s(n) = \sum_{k=1}^{p} \alpha_k s(n-k) \text{ for some given value of } \rho, S, \alpha_\kappa \tag{4.1}$$

In the above equation:
α_κ, s: Linear prediction coefficients
$s(n)$: Windowed speech sequence

$s(n)$ is calculated by the product of a short time speech frame with either hamming or a similar type of window:

$$S(n) = x(n)^* w(n) \tag{4.2}$$

In the above, the windowing sequence is represented by $\omega(n)$.

4.5 Kurtosis

Kurtosis is used to find heavy-tiredness or light-tiredness of data when compared to a normal distribution. High kurtosis indicates heavy tails and similarly with low kurtosis indicates light tails. For data Y_1, $Y_2 \ldots Y_N$, the kurtosis is given by:

$$\text{Kurtosis} = \frac{\sum_{i=1}^{N} (Y_i - Y)^4 / N}{S^4} \tag{4.3}$$

In the above equation:
S represents standard deviation
$\ddot{y}$ represents mean
N represents the number of data points

To calculate standard deviation, N is used for the denominator in kurtosis.

4.6 Mean

The average of the EEG signal is used as another statistical parameter for multichannel dataset signals. This feature helps to differentiate epileptic seizure event. An arithmetic mean is calculated using the following equation:

4.7 Classification

Classification deals with classifying a new observation; it is based on knowledge gained from training data with known category membership. Various methods, such as neural networks like back-propagation, LVQ, SOM, feed forward, normalized correlation, K-nearest neighbor (KNN), Hamming distance, support vector machine (SVM), weighted Euclidean distance, which are used for classification.

In this chapter, SVM and KNN are used for pattern classification, where the observations are classified based on features.

4.8 K-Nearest Neighbor (KNN)

KNN is one of the popular classification techniques because of its simplicity and robustness. The test observation (feature) is classified by finding a nearest neighbor calculated from training observations (features).

Various distance metric measurement techniques such as Euclidean distance, Chebyshev, city block, correlation, cosine, Spearman, Hamming, and so on, are used to find the distance between the training and testing vectors.

A simple Euclidean distance formula for the distance between training and testing vectors is given by:

$$d(a,b) = \sqrt{\sum_{i=1}^{n} (a_i - b_i)^2} \tag{4.5}$$

The testing vector is assigned with the label of the class which is at the smallest distance.

The features are extracted for training and testing set of observations, representing these observations in some different dimension space. The similarity of two different points can be represented by the distance between them in a space.

The working of the K-nearest neighbor algorithm is as follows [11–13]:

- Define a positive integer value K. Also calculate the features of the new observation.
- For the new observation, calculate the K closest distances.
- Find the closer classification of this training observation.
- This helps us to classify the new observation.

If the result is not satisfactory, change the value of K until the reasonable level of correctness is achieved.

In Figure 4.2, there are two classes, represented by squares and triangles. The new testing feature is introduced in the feature set. The testing feature is classified into the right class using K-nearest neighbor by assigning the label of the higher majority neighbors.

$$A := \frac{1}{n}\sum_{i=1}^{n} a_i \tag{4.6}$$

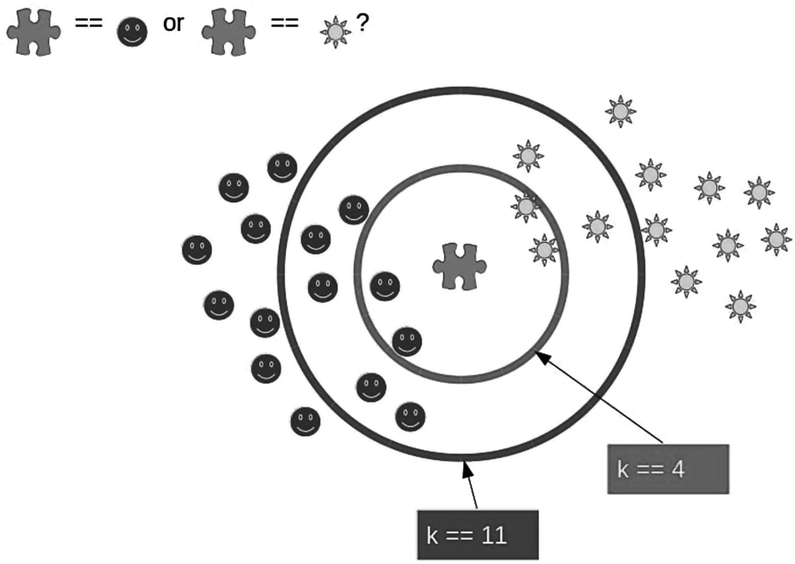

FIGURE 4.2
Classification of query image through a KNN classifier [19].

4.9 Auto-Correlation

The self-correlation of signal is an important feature used to correlate signals based on time difference and average value at origin.

$$R_{XX} = \sum_{i=0}^{\infty} S(n) \tag{4.7}$$

The correlation coefficient R is used to indicate a higher degree of similarity, [−1, 1], where perfect correlation is given by 1 and perfect anti-correlation is given by −1.

4.10 Principal Component Analysis (PCA)

For low artifact signal analysis PCA is mostly used compared to linear discriminant analysis (LDA) and independent component analysis (ICA). Furthermore, PCA is used for mapping data to lower-dimensional space from high-dimensional space.

4.11 Seizure Prediction

Figure 4.3 shows various medical event timelines, including the “no seizure occurrence” time interval referred to as “interictal.”

For preictal, the alarm will be set in forthcoming elapsed time denoted by the Seizure Prediction Horizon (SPH). The Seizure Occurrence Period (SOP) follows the SPH. The seizure is expected to occur at SOP. After the seizure period, during the postictal period, no seizures are observed.

The thumb rule to avoid a harmful situation is that the cumulative time interval of SPH and SOP should be longer than 5 minutes.

The main objective is to identify and detect SOP on the patient side by wearable devices. Further, this data is processed remotely with the help of IoT. The block diagram shows complete wireless Smart Web of Things (SWOT) architecture to monitor epileptic seizure activity at the user end with

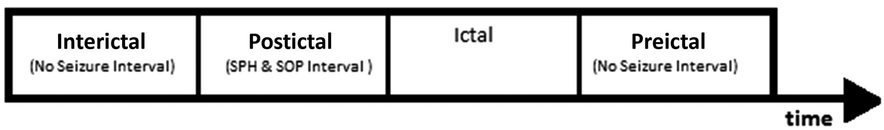

FIGURE 4.3
Seizure event time activity.

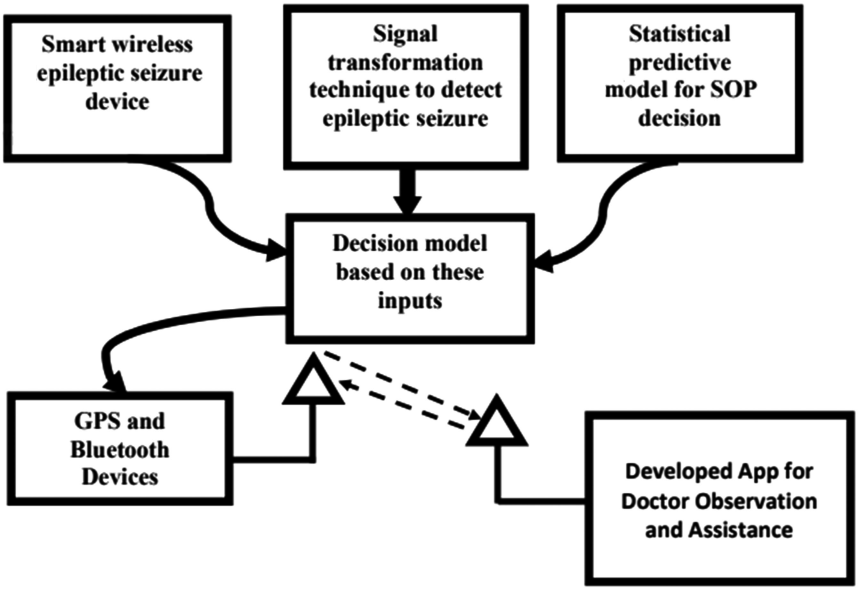

FIGURE 4.4
Web of Things (WOT) to predict epileptic seizure.

a comfortable and portable prototype. The user end observatory device will be connected with the clinical side through existing smartphone devices. Figure 4.4 shows the block diagram of epileptic seizure prediction and detection using signal transform and statistical methods.

Furthermore, such kinds of smart medical devices are today's need, wherein we can monitor health status remotely and also the location of the patient can be traced. The alarm events generated can be further shared to family members as well as to clinical observers for preventive action.

4.12 Results and Discussion

For the simulation analysis, we used the MATLAB tool and the parameters in Table 4.1 are considered. In the proposed work, the EEG signals (Healthy

TABLE 4.1
Simulation Parameter

Descriptions	Parameters
EEG dataset	Set-A and Set-E [15]
Filter	Chebyshev filter
Decomposing technique	Wavelet transform
ANN Classifier	KNN classifier

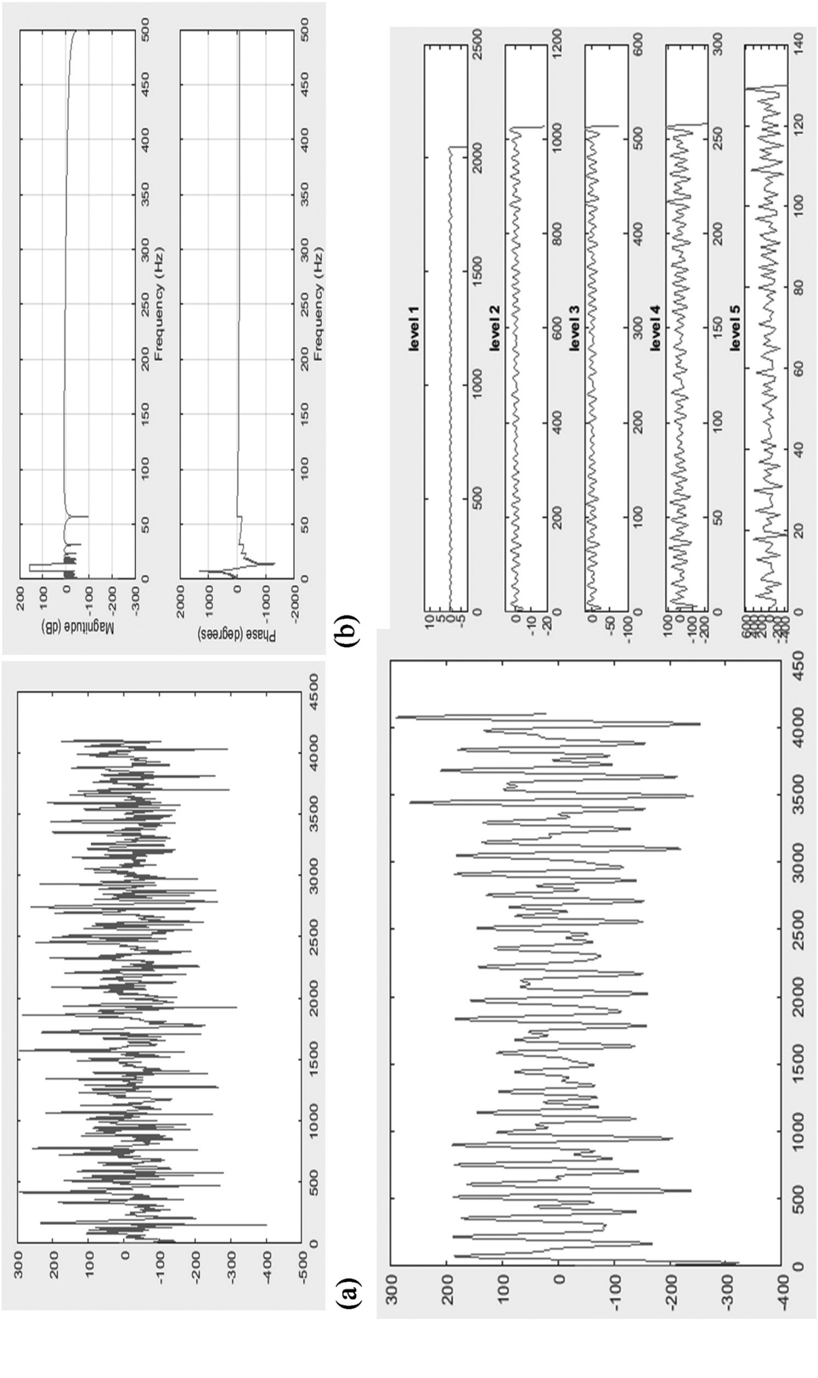

FIGURE 4.5
Qualitative analysis of proposed system on normal EEG signal: (a) input normal EEG signal, (b) magnitude and phase plot of the filter, (c) filtered signal, (d) wavelet decomposition level 5.

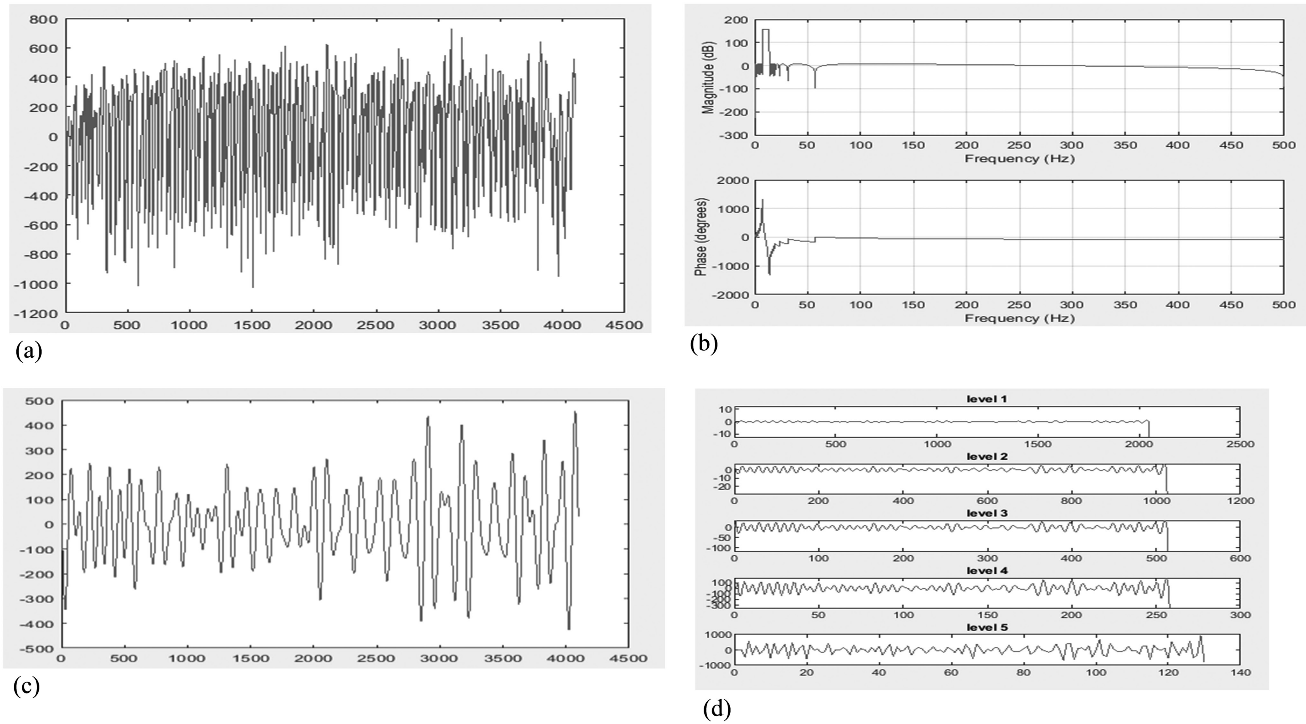

FIGURE 4.6
Qualitative analysis of proposed system on normal EEG signal: (a) input normal EEG signal, (b) magnitude and phase plot of the filter, (c) filtered signal, (d) wavelet decomposition level-5.

TABLE 4.2

Features of Normal and Epileptic EEG Signals

	Features													
Patient	**LPC**					**Kurtosis**	**Mean**	**Auto-Correlation**						**PCA**
1	1	1.258484	...	-0.0455	-0.0818	624.6202	3.57455	-0.1477	1	...	-0.457	-0.025		0.17166
2	1	1.74397	...	-0.3164	-0.2009	1619.425	3.331812	-0.2513	1	...	-0.164	-0.175		0.20423
3	1	1.367817	...	-0.1833	-0.0946	5079.8	4.538614	-1.9356	1	...	-0.497	0.1293		0.12803
4	1	1.112386	...	0.102362	0.009233	6133.054	2.962493	0.378282	1	...	-0.514	-0.091		0.20748
5	1	1.317612	...	-0.15653	-0.10689	7704.017	2.969006	2.610246	1	...	-0.525	0.1911		0.02850
6	1	1.325386	...	-0.01242	0.009085	2564.211	2.695167	-0.20193	1	...	-0.487	-0.087		0.27810

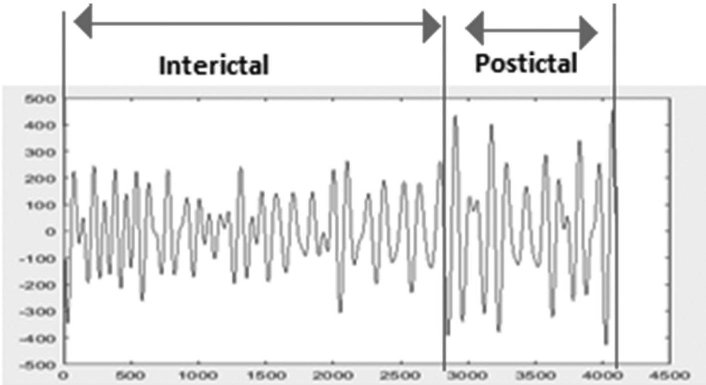

FIGURE 4.7
Seizure activity filtered signal.

person dataset Set-A and epileptic seizure person dataset Set-E) are preprocessed through the Chebyshev filter. The filtered signal is then decomposed by the Discrete Wavelet Transform. Different statistical features were extracted to differentiate the normal and epileptic EEG signal, as shown in Figures 4.5 and 4.6. The outcomes of the proposed system are described below in a qualitative and quantitative manner.

In the qualitative analysis, the graphical results of the EEG signals are presented.

Figure 4.7 shows the seizure timeline activity signal; we need to look at the adaptive seizure prediction algorithm (ASPA) for better real-time performance.

In Table 4.2, the extracted features using five different techniques have been tabulated. Patients 1–4 are the normal patients while Patients 4–6 are the epileptic patients. From Table 4.2, it is observed that the values of the features of the epileptic and normal patients differ. These features are then applied to the Machine Learning Algorithm (KNN). We have programed and designed the Graphical User Interface (GUI) to display the final output whether the person has epilepsy or not. The results are shown in Figures 4.8 to 4.11.

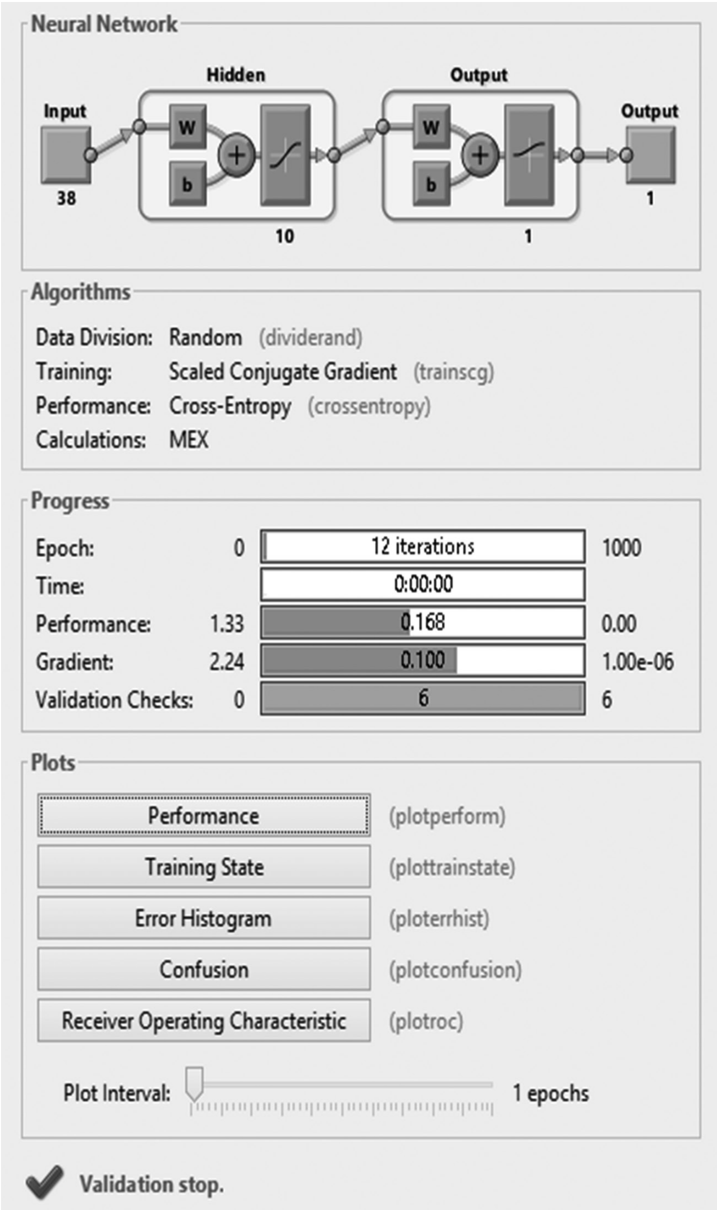

FIGURE 4.8
ANN model.

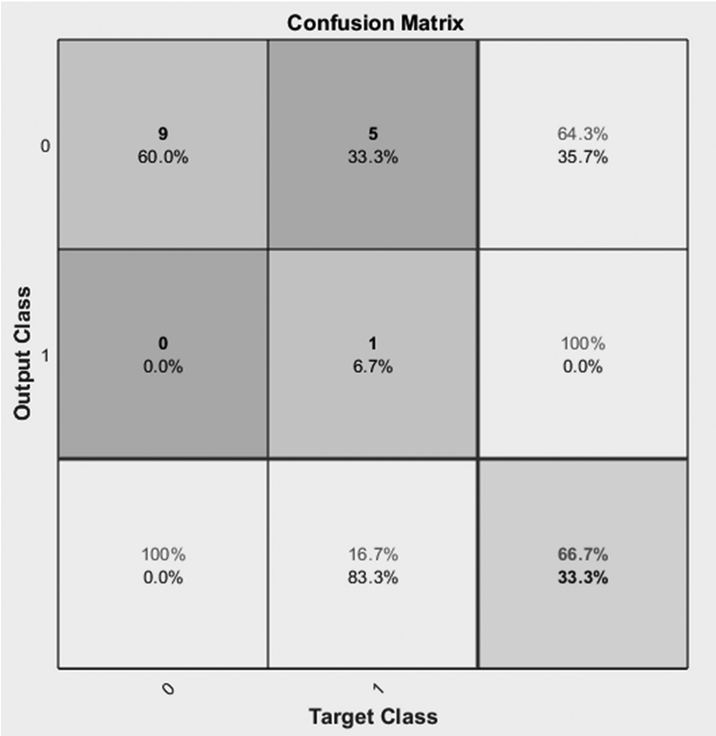

FIGURE 4.9
Confusion matrix.

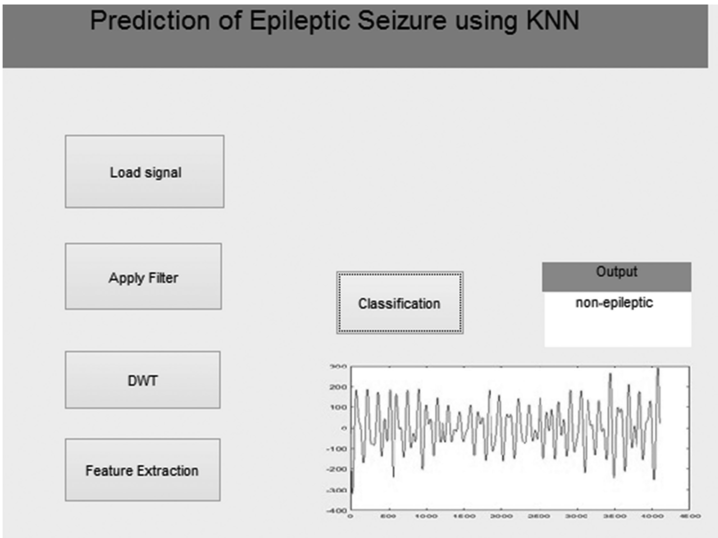

FIGURE 4.10
GUI for non-epileptic patient.

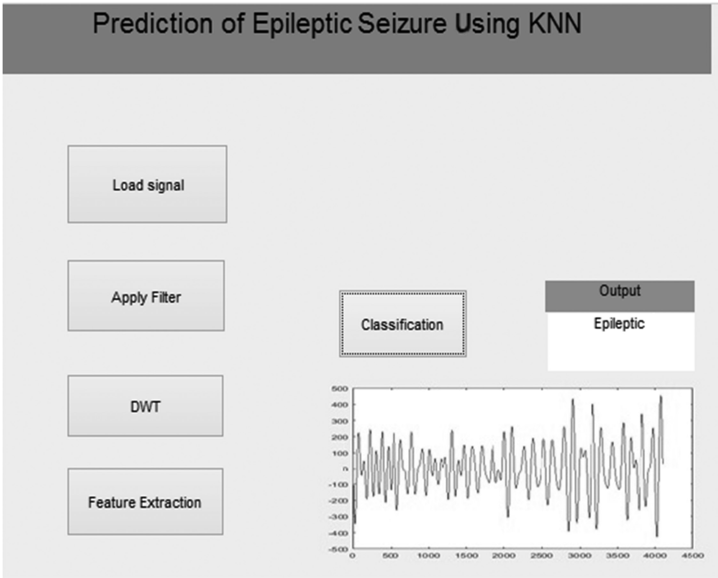

FIGURE 4.11
GUI for an epileptic patient.

4.13 Conclusion

In this chapter, the system for auto-classification of the normal and epileptic EEG signal has been implemented. EEG signals of normal and epileptic patients were collected from online sources. The EEG signals were preprocessed using the Chebyshev filter. From the filtered signal, various features are extracted, viz. LPC, kurtosis, mean, auto-correlation and PCA. Those features are used as an input in a KNN. The final classification of the EEG signals of the existence of seizures or not is done by KNN. A further filtered signal is used to predict the occurrence seizure event. The application can be scaled up using the Web of Things.

4.14 Acknowledgments

We are thankful to Dr. Sanjay Pawar (Fellowship in Stroke Neurology [FISN], Consultant Stroke Physician), for continuous guidance and more insight on epileptic seizure medical terminology.

References

1. World Health Organization. Epilepsy. Available at: https://www.who.int/news-room/fact-sheets/detail/epilepsy.
2. T. Tzallas, M. G. Tsipouras, and D. I. Fotiadis, "Automatic seizure detection based on time-frequency analysis and artificial neural networks," *Computational Intelligence and Neuroscience*, 2, p. 13, 2007.
3. E. Juárez-Guerra, V. Alarcon-Aquino, and P. Gomez-Gil, *Epilepsy Seizure Detection in EEG Signals Using Wavelet Transforms and Neural Networks*. Springer International Publishing, Switzerland, pp. 261–269, 2015.
4. S. S. Zakareya Lasefr, *Epilepsy Seizure Detection Using EEG Signals*. IEEE, p. 6, 2017.
5. J. Rasekhi, M. R. K. Mollaei, M. Bandarabadi, C. A. Teixeira, and A. Dourado, "Pre-processing effects of 22 linear univariate features on the performance of seizure prediction methods," *Journal of Neuroscience Methods*, 217(1–2), pp. 9–16, 2013.
6. C. A. Teixeira, B. Direito, M. Bandarabadi, et al., "Epileptic seizure predictors based on computational intelligence techniques: A comparative study with 278 patients," *Computer Methods and Programs in Biomedicine*, 114(3), pp. 324–336, 2014.
7. C. Brunner, M. Naeem, R. Leeb, B. Graimann, and G. Pfurtscheller, "Spatial filtering and selection of optimized components in four class motor imagery EEG data using independent components analysis," *Pattern Recognition Letters*, 28(8), pp. 957–964, 2007.
8. M. Bandarabadi, C. A. Teixeira, J. Rasekhi, and A. Dourado, "Epileptic seizure prediction using relative spectral power features," *Clinical Neurophysiology*, 126(2), pp. 237–248, 2015.
9. K. Gadhoumi, J. Lina, and J. Gorman, "Discriminating preictal and interictal states in patients with temporal lobe epilepsy using wavelet analysis of intracerebral EEG," *Clinical Neurophysiology*, 123(10), pp. 1906–1916, 2012.
10. S. Zandi, R. Tafreshi, M. Javidan, and G. A. Dumont, "Predicting epileptic seizures in scalp EEG based on a variational Bayesian Gaussian mixture model of zero-crossing intervals," *IEEE Transactions on Bio-Medical Engineering*, 60(5), pp. 1401–1413, 2013.
11. S. J. Roberts, D. Husmeier, I. Rezek, and W. Penny, "Bayesian approaches to Gaussian mixture modeling," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11), pp. 1133–1142, 1998.
12. Bahubali Shiragapur, Nishikant Surwade, et al., "Experimental study on detection of epilepsy," *Helix*, 9(3), pp. 5052–5056, 2019.
13. Vijaylaxmi Jain, Bahubali Shiragapur, et al., "Sleep stages classification using wavelet transform & neural network," in *Proceedings of 2012 IEEE-EMBS International Conference on Biomedical and Health Informatics*, Hong Kong, pp. 71–74, 2012.
14. Song Y. Crowcroft and J. Zhang, "Automatic epileptic seizure detection in EEGs based on optimized sample entropy and extreme learning machine," *Journal of Neuroscience Methods*, 210(2), pp. 132–146, 2012.
15. CHB-MIT Scalp EEG Database. *PhysioNe*t. Available at: https://physionet.org/content/chbmit/1.0.0/.

16. Farzad Samie, Sebastian Paul, et al. "Highly efficient and accurate seizure prediction on constrained IoT devices," IEEE Xplorer, 2018, ISSN-1558-1101.
17. Yinda Zhang, S. Yang, et al., "Integration of 24 feature types to accurately detect and predict seizures using scalp EEG signals," *Sensors*, 18(5), p. 1372, 2018. doi:10.3390/s18051372.
18. V. V. Shete, et al., "Sleep stage classification using wavelet transform & neural network," *Trans-Stellar Journal USA, IJECIERD*, 2(2), pp. 38–45, 2012.
19. K-nearest neighbor classifier. *Python Machine Learning Tutorial*. Available at: https://www.python-course.eu/k_nearest_neighbor_classifier.php.