ANALYSIS AND CLASSIFICATION OF EEG SIGNALS FOR DETECTION OF EPILEPTIC SEIZURES

A THESIS SUBMITTED IN FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY
IN
ELECTRONICS AND COMMUNICATION ENGINEERING
By

MEENAKSHI SOOD
Enrollment No: 116004

Department of Electronics and Communication Engineering
JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY
WAKNAGHAT, SOLAN (H.P.) - INDIA
OCTOBER - 2015

DECLARATION BY THE SCHOLAR

I hereby declare that the work reported in the Ph.D. thesis entitled **"Analysis and Classification of EEG Signals for Detection of Epileptic Seizures"** submitted at **Jaypee University of Information Technology, Waknaghat, Solan, India,** is an authentic record of my work carried out under the supervision of **Dr. Sunil.V. Bhooshan.** I have not submitted this work partially or fully anywhere else for any other degree or diploma.

Meenakshi Sood

Department of Electronics and Communication Engineering
Jaypee University of Information Technology, Waknaghat, India- 173234
October 2015

JAYPEE UNIVERSITY OF INFORMATIO TECHNOLOGY
(Established by H.P. State Legislative vide Act No. 14 of 2002)
P.O. Waknaghat, Teh. Kandaghat, Distt. Solan - 173234 (H.P.) INDIA
Website: www.juit.ac.in
Phone No. (91) 01792-257999
Fax: +91-01792-245362

CERTIFICATE

This is to certify that the work reported in the Ph.D. thesis entitled **"Analysis and Classification of EEG Signals for Detection of Epileptic Seizures"** which is being submitted by **Meenakshi Sood** in fulfillment for the award of degree of Doctor of Philosophy in Electronics and Communication Engineering by the Jaypee University of Information Technology, is the record of candidate's own work carried out by him under my supervision. This work is original and has not been submitted partially or fully anywhere else for any other degree or diploma.

Dr. Sunil .V. Bhooshan
Professor & HOD
Department of Electronics & Communication Engineering
Jaypee University of Information Technology, Waknaghat, Solan.

ACKNOWLEDGEMENT

"Gratitude can transform common days into thanksgivings, turn routine jobs into joy, and change ordinary opportunities into blessings".
Completion of this research would not have been possible without the help of many people, to whom I am very grateful.

I would like to express my sincere gratitude and appreciation to my supervisor, ***Professor Sunil V. Bhooshan****, for his continuous inspiration, support, guidance, encouragement, constructive criticism and individual feedback throughout the course of my PhD study. In him I found that the best are not only efficient, effective and result driven, but at the "core they are persons with the best qualities as human beings". I feel very grateful and blessed to have worked under his supervision. I would not imagine carrying out this research without his constant support and help.*

This feat was possible only because of the unconditional support provided by ***Dr. Vinay Kumar,*** *Thapar University Patiala for serving as excellent advisors, always making himself available and providing invaluable painstaking efforts towards my research work. I express my deepest sense of gratitude and sincere thanks to* ***Dr. Pradeep Naik*** *for his enthusiastic guidance and constructive comments which aided a lot in fulfilling the objectives of this research work.* ***Dr. Shalini Batra****, my childhood friend from Thapar University was kind enough to undertake the editing work and provide me good support system whenever required.*

I gratefully acknowledge Jaypee University of Information Technology, Waknaghat, for offering me the opportunity to perform this research successfully, for providing necessary facilities and support, without which, this work would not have been possible. I am also thankful to faculty especially my friend and colleague ***Dr. Jitendra Virmani*** *and staff of Electronics and Communication department for their support.*

*I also have to thank the members of my PhD **DPMC committee**, for their helpful advice and suggestions in general. I am grateful to the reviewers whose constructive suggestions and invaluable advice improved the quality of publications derived from this work. I would also like to thank all of my friends who supported me in writing, and incented me to strive towards my goal.*

*I express my heartfelt reverence to my **parents** who encouraged and helped me at every stage of my personal and academic life, and longed to see this achievement come true. I wish to express my appreciation to my loving brother **Vikas Batra** for enduring love and faith during my research work. Last and most importantly, I wish to thank my husband, **Dr. Sanjay Sood (IFS)** who stands beside me and encourages me constantly for every step of my work.*

*I shall remain indebted throughout my life to my son **Rittwik Sood** and daughter **Khushbu Sood** for bearing my tantrums, for sharing the time which was their due and giving me happiness, joy and enduring love. I would like to dedicate this thesis to my lovely and sincere children.*

Finally, I humbly bow my head with utmost gratitude before the God Almighty who always showed me the path to follow and without whose endless benevolence and blessings this tedious task could not have been accomplished.

MEENAKSHI SOOD

TABLE OF CONTENTS

Contents	Page No.

Chapter 1

Introduction

Contents		**Page No.**

Contents		Page No.

Contents		**Page No.**

LIST OF ACRONYMS

ACRONYMS	FULL FORM
ACT	Activity
CAD system	Computer Aided Diagnostic system
CA	Classification Accuracy
CM	Confusion Matrix
COMPX	Complexity
CSFV	Combined Signal Feature Vector
ECAD	Ensemble Computer Aided Diagnostic
EEG	Electroencephalogram
FL	Feature length
HCAD	Hierarchical Computer Aided Diagnostic
HE	Hurst Exponent
HJP	Hjorth Parameter
INT-ICT (F)	Inter-ictal
ICT (S)	Ictal
ICA	Individual Classification Accuracy
IMA	Individual Misclassification Accuracy
KNN	K- Nearest Neighbor
MCAD	Module Computer Aided Diagnostic
MLPNN	Multi layer Perceptron Neural Network
MOB	Mobility
NOR (Z)	Normal
NLFV	Non Linear Feature Vector
OCA	Overall Classification Accuracy
PNN	Probabilistic Neural Network
RBFNN	Radial Basis Function Neural Network
ROC	Receiver Operating Characteristics

ACRONYMS	FULL FORM
RSFV	Reduced Signal Feature Vector
Sen	Sensitivity
SFV	Signal Feature Vector
Spec	Specificity
SVM	Support Vector Machine
SBS	Sequential Backward Search
SFS	Sequential Forward Search

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Epilepsy is a persistent, constantly recurring neurological brain disorder characterized by abnormal electrical activity in the brain. Epilepsy is not only a disorder, but rather acts as a syndrome with divergent symptoms involving spasmodic abnormal electrical activities in the brain. Clinical data relevant to such abnormalities is complex, context-dependent, and multi-dimensional, and such data generates an amalgamation of computing research challenges. Electroencephalogram (EEG) is one of the most clinically and a scientifically utilized signal recorded from human brain and is powerful source of providing valuable insight of the brain dynamics. Accurate and careful analyses of these signals play a prominent role in diagnosis of brain diseases and many cognitive processes.

EEG technique has excellent temporal resolution, noninvasiveness, usability, and low set-up costs while capturing brain signals, which makes it popular in this arena of research. The electroencephalogram recordings of epileptic subjects are visually inspected and analyzed by trained neurologists or radiologists for clinical diagnosis of epileptic seizures and possible treatment plans. However, due to the complex, time-consuming, high-dimensional nature of the EEG event recordings, visual inspections of EEG signals often result in errors. Therefore, there is a need to develop automatic systems for classifying the recorded EEG signals.

The driving force behind this thesis is development of automated computer aided classification crucial for restricting subjectivity in the diagnosis of epileptic signals. It should be stressed out that automatic algorithms do not intend to be used instead of an expert, but as a decision support tool as a second opinion. The scope of this work is to detail the variety of Computer Aided Diagnostic (CAD) techniques developed for epilepsy detection using EEG signals. The proposed systems concentrate on detecting epilepsy by classifying only the normal and ictal stages (two-class problem), and other CAD systems are designed to classify three stages, namely, normal, interictal, and ictal (three-class problem).

Nonlinear and non-stationary nature of the EEG signals further increase the complexities related to their manual interpretation. This research work encompasses the inclusion of different informative parameters, with focus on nonlinear features and non stationary nature of EEG signals. The greatest effort is focused on data representation stage, on the potential of extracted feature set, designing, implementing and deciding an appropriate combination machine learning methods in order to enhance the differentiation and identification of epileptic states. The motivation behind this thesis is examination of morphology and topography of waveforms during certain neurophysiologic phenomena. The aim is analysis of hidden dynamic of the EEG time series, extracting more information about pathological vs. normal status of the EEG signals.

The designed automated systems transform the qualitative diagnostic criteria into a more objective quantitative feature classification problem. High sensitivity ensures accurate classification of the signals and high specificity results in no false positive classification. The false positive classifications yield unacceptable effects on the quality of life of the subjects. With this consideration, 100% sensitivity and 100% specificity with 0% misclassification rate was achieved with the attributes found within the search space along with the selected classifiers. The proposed method of epileptic signals classification can be very useful in predicting the action or the intention of action performed on the basis of EEG which leads to more development in human computer interaction.

CHAPTER 1

INTRODUCTION

1.1 OVERVIEW OF EPILEPSY

Epilepsy is a persistent, constantly recurring neurological brain disorder characterized by abnormal electrical activity in the brain. It affects about 1% of world population, out of which 85% is prevalent in the developing countries [1]. Epileptic seizure is the manifestation of neurological disorder that causes excessive and hyper synchronous firing of large number of neurons in the brain. During seizure, occurrence of strange sensations, convulsions and loss of consciousness are noticed. The seizures are sudden, brief and recurrent causing strange sensations, change in emotions, convulsions and loss of consciousness [2]. Simultaneous activity of a group of nerve cells in the cerebral cortex leads to bursts of sudden and excessive electrical energy leading to epileptic seizures. According to the International League Against Epilepsy (ILAE-1981), seizures are classified into two categories (1) Generalized seizures, (2) Partial (or focal) seizures [3]. Generalized seizures start in both hemispheres of the brain simultaneously and are associated with a variety of motor symptoms characterized by generalized spike-and-slow wave discharges. While, partial seizures originate in a localized region and cause relatively mild cognitive and brief sensory symptoms.

Diagnosis of epilepsy is critical as it has overlapping symptomatology with other neurological disorders; moreover, mechanism and cause responsible for epilepsy and seizure progression is not very clear [4]. Despite of intensive research into the causes and medical treatment of epilepsy, we still have little idea of many of the underlying cellular and network properties that can give rise to naturally-occurring seizures. This difficulty stems from both the uniqueness of the disorder in each patient and our still-poor understanding of the human brain. Nonetheless, the detection of the disorder and recognition of the affected brain area is essential for the clinical diagnosis and treatment of epileptic patients.

There are many ways to diagnose epilepsy such as electroencephalography (EEG), Magneto encephalography (MEG), Magnetic Resonance Imaging (MRI), positron emission tomography (PET) etc. Among these methods, EEG is the most useful for diagnosis and identification of epilepsy treatment. EEG has speed, high time resolution, and non-invasive advantages; hence, it is the inexpensive and widespread diagnostic epilepsy method [5].

Usually, the diagnosis of epileptic seizures involves the analysis based on combination of the medical history of the patient and interpretation by an expert neurologist through EEG recordings [6]. Nevertheless, with the emergence of new signal processing techniques, an increased improvement in the analysis of the EEG for prediction of epileptic seizures has been reported. These enhanced algorithms can detect abnormal disorder and malfunctioning of the brain not only during the seizure but also can detect the onset of seizure up to a certain extent. An automatic seizure detection system is used in the diagnosis of epilepsy, which act as a second opinion tool apart from visual inspection of EEG by the physician [7].

1.2 HUMAN BRAIN

In this section a brief description of brain and functionality of its anatomical structures is presented. This part also covers the mechanism of generation of local current flows and electrical activities that can be recorded on the scalp as EEG signals.

1.2.1 Nervous system

The nervous system is a system comprising a network of neurons that collects, communicates, and processes information. Central Nervous System (CNS) and Peripheral Nervous System (PNS) are two parts of CNS [8]. CNS comprises brain and spinal cord, and PNS comprises sensory neurons and interconnected nerves cells. Sensory inputs from PNS processed by the CNS are sent to the organs of the body. The information processing from all mental processes and response initiation is integrated in the CNS and produces ideas, emotions, and other mental processes. Neurons are interconnected to form complex networks that communicate information to and from the brain. These neurons receive impulses from nerves and transfer signals to other nerves individually or collectively [9].

1.2.2 Cerebral cortex

The outermost layer of the cerebrum is known as cerebral cortex and it plays an important role in memory, consciousness, thought and awareness. The increase in the neuronal area is due to convoluted cerebral cortex surface by ridges and valleys. The cortex is divided into two left and right symmetrical hemispheres, which are separated by central sculcus. Each cerebral hemisphere is divided into lobes: the *frontal lobe*, *temporal lobe*, *occipital lobe and* the *parietal lobe*. Each lobe has specific function attached to it; as occipital lobe is involved with vision, frontal lobe deals with decision-making, problem solving, and planning. Similarly, the other two lobes are responsible for reception, comprehension and processing of sensory information from the body; and with memory, emotion, hearing, and language [10].

1.2.3 Physiology of Cerebral Potentials

The electrophysiological properties of the nervous system are the origin of cerebral potentials. These electrical potentials are transmitted by one nerve and cause the production of action potentials in another nerve. The differences of electrical potentials are caused by summed postsynaptic graded potentials and create electrical dipoles between body of neuron and apical dendrites. Electrical charges move within the central nervous system consisting mostly of Na^+, K^+, Ca^{++}, and Cl^- ions, resulting in electrical signals. These ions are pumped in neuron membranes in the direction governed by membrane and a potential of 60-70 mV can be recorded under the membrane of the cell body [10]. EEG records the electric potential from the exposed surface of scalp and measures the current flow in the cerebral cortex. The EEG can be measured directly from the cortical surface and also by using depth probes.

1.3 ELECTROENCEPHALOGRAM (EEG) SIGNALS

EEG results due to an activation of 10000–100000 neurons simultaneously and can be recorded as a potential difference (voltage) between two electrodes placed on the scalp. The measurement of the potential requires approximately 5 cm^2 area of activation on the cortex beneath each electrode. EEG has capability to reflect all the activities of the brain, so it has been proved to be a very powerful tool in the field of clinical neurophysiology [11]. Electroencephalography is an effective non-invasive tool for (*i*) understanding the dynamic complex behavior of the brain, (*ii*) monitoring its different physiological states, and (*iii*) diagnosis of neurological disorders [11].

1.3.1 Acquisition of EEG Signals

As EEG signals are non-invasive electrical brain signal, they are captured with the help of electrodes placed on the scalp (sometimes in form of a cap). Electrodes are cup-shaped and are placed at specific locations of the scalp. The skin never touches the electrode material directly in these electrodes. EEG gel or paste acts as an interface material between the electrode and the skin. The electrodes provide enough volume to contain an electrolyte and capture the electrical signal [12]. The electrode-skin interface impedance depends on the interface layer, area of electrode’s surface, and temperature of the electrolyte. Figure 1.1 shows the electrical equivalent of the combination of skin, electrolyte and electrode. The electrode-tissue interface is resistive and consists of capacitive elements. The ions are accumulated as parallel plates because of the interaction between metallic electrode and electrolyte. The ion-electron exchange occurs between the electrode and the electrolyte that results in voltage given by the Nernst Equation (1.1) [13], simply:

$$\varepsilon = \varepsilon_o - \frac{0.05916}{n_e}\log Q \qquad (1.1)$$

$$Q = \left(\frac{[ions_{inside}]}{[ions_{outside}]} \right)$$

Where ε is half cell potential measured in V, n_e is the transported electron (mol number) and Q gives rate of ions.

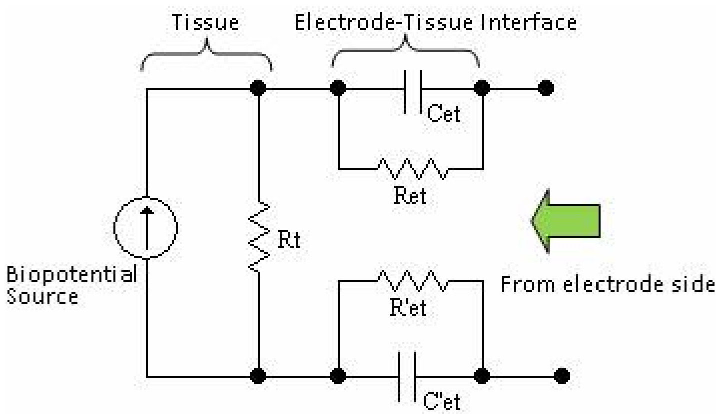


Figure 1.1. Equivalent circuit of biopotential source and electrode-tissue interface from electrode. Biopotential source acts as a current source and tissue resistance is shown as R_t, C_{et} and R_{et} electrode-tissue equivalent elements.

A biopotential source is a current source that causes current flow in the extracellular fluid through the tissue. To record an accurate signal, the contact impedance between the electrode surface and the scalp should be between 1kΩ to 10kΩ. If contact impedance is less than 1kΩ, a possible short between electrodes is indicated, and if impedance is greater than 10kΩ, it can cause distorting artifacts. So the resultant signal developed at the terminals is collected by the electrode and passed on to the electronic circuitry. An EEG machine is a recording device connected by wires to electrodes pasted at key points on the patient's head [14]. The widely used method to describe the location of scalp electrodes is "10-20" system. This is an International Standard of naming and positioning of the electrodes on the cerebral cortex for measuring brain activity [15]. The "10" and "20" refer to the distances between adjacent electrode that are either 10% or 20% of the total front-back or right-left distance of the skull. Each electrode location has a letter to identify the lobe (F:frontal,T:temporal, P:parietal,O:occipital) and a number to identify the hemisphere location; left hemisphere (odd numbers 1,3,5,7) and for right hemisphere (even numbers 2,4,6,8) respectively as depicted in Figure 1.2.

The electronic circuitry of EEG systems comprises amplifiers and filters. They are equipped to store the large data, and to convert analog EEG signal to digital by high sampling rate and large number of quantization levels.

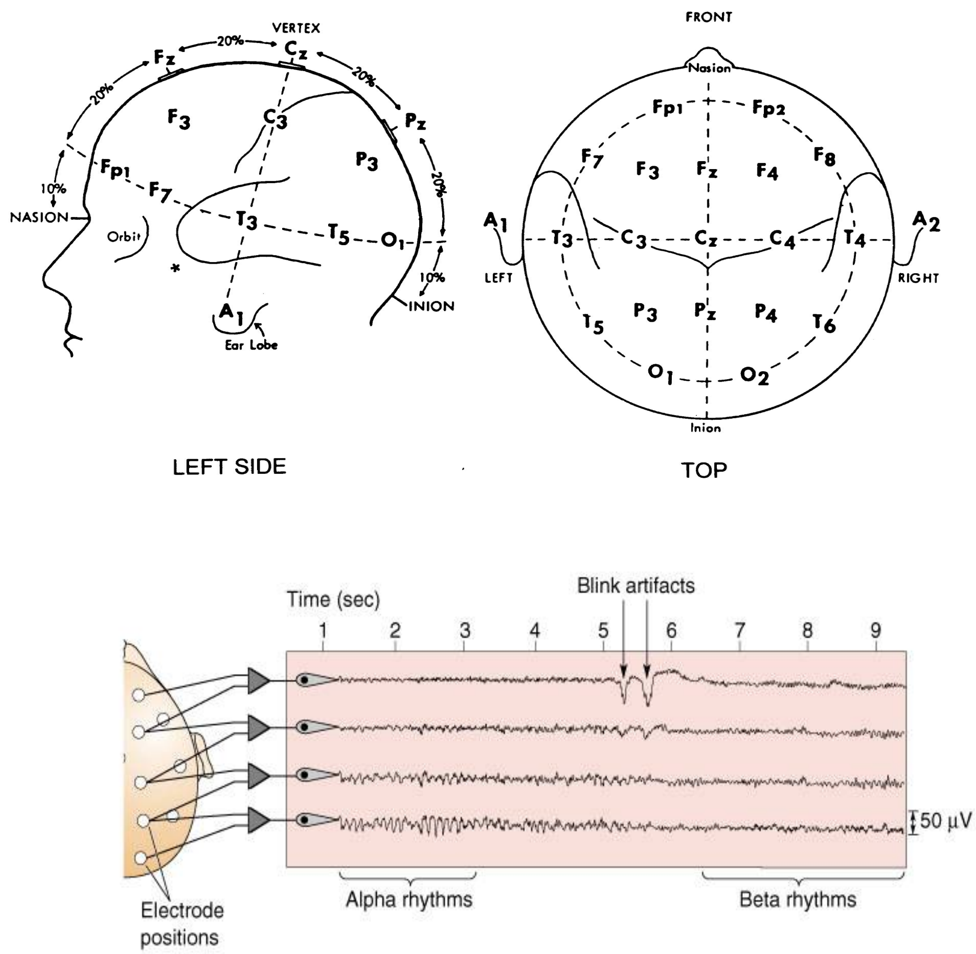


Figure 1.2. Conventional 10-20 EEG electrode positions for the placement of 21 electrodes

As the bandwidth for EEG signal is 50 Hz, at least sampling frequency of 100 samples / sec is required for sampling the EEG signal satisfying Nyquist criterion. After digitization, the voluminous data obtained is recorded in the connected memory units.

1.3.2 Details of the dataset

EEG data set used in this paper consists of EEG recordings taken from Department of Epileptology, University of Bonn, Germany described in [16]. EEGs from five patients for three different conditions were selected, containing 100 single-channel EEG signals of 23.6 seconds duration. Signals were recorded extra-cranially and intra-cranially with 128-channel amplifier system using an average common reference. The signals are digitized using 12-bit resolution and sampled at a rate of 173.61 Hz. Band-pass filter with 0.53–40 Hz (12 dB/octave) cutoff were used. The total number of EEG signals is 300 (100 ictal signals, 100 normal signals and 100 interictal signals). All selected EEG signal segments are cut out from continuous EEG recordings after visual inspection for artifacts due to eye movements. Segments of 4396 samples were first cut out of the recordings, as discontinuities between the end and beginning of a time series are known to cause spurious frequency components. The final segment was chosen so as the slopes at the end and beginning of the time series had the same sign. The amplitude difference of the last and first data points was within the range giving final segments of N = 4096 samples. The data set comprises three different sets (F, Z, and S) with different conditions, the signals of set Z represent normal condition, and set F signals represent interictal condition and signals in set S exhibit ictal activity. Set Z was collected from five healthy subjects with eyes closed. Sets F and S were created from EEG records of the pre-surgical diagnosis of five epileptic patients. Signals in Set F were recorded in seizure-free intervals from the hippocampal formation within the epileptogenic zone. Set S contains the EEG records of epileptic patients during seizure activity. One of the signals with amplitude in micro volts (μV) from each respective category is depicted in Figure 1.3. Visualizing the figure, it can be seen that the amplitude of the signals during seizure is quite high as compared to signals in F state. Two categories of abnormal activity can be observed in an EEG signal: *ictal* (during an epileptic seizure) and *inter-ictal* (between seizures). The amplitude of EEG signals of a normal condition subject is the lowest as compared to rest of the signals. The amplitude of the signals represented in Figure 1.3 is in micro-volts.

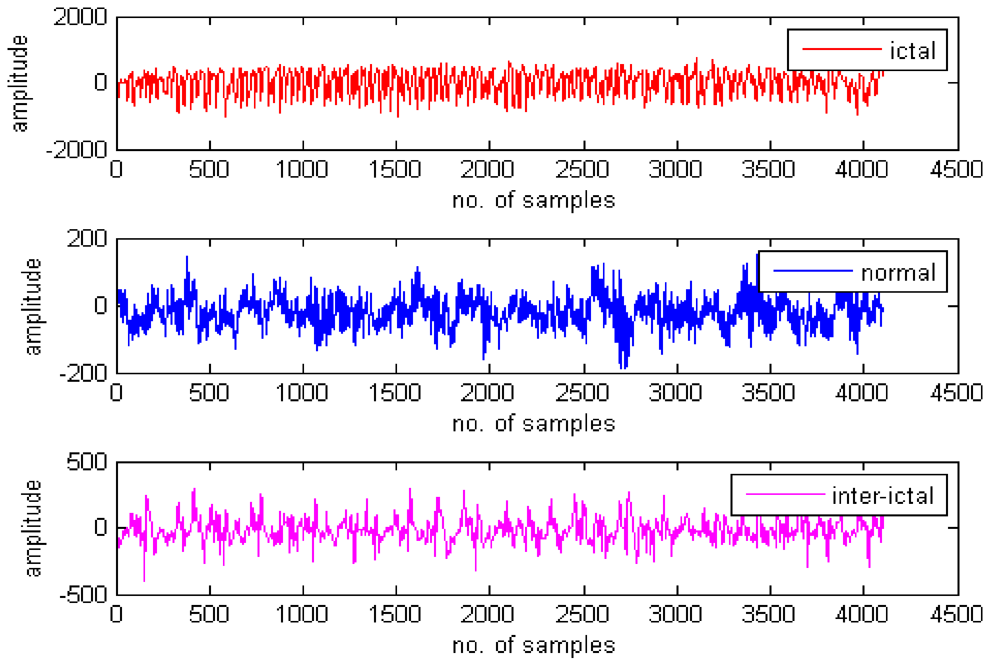


Figure 1.3. EEG signals of a patient in various stages (***a***) ictal state (***b***) normal state (***c***) inter-ictal state with total of 4096 samples. Amplitude of the signals is in µV.

1.3.3 Morphology of EEG Signals

The recorded electrical activity is characterized in terms of specific descriptors and measurements as

1. Frequency or wavelength
2. Voltage
3. Waveform
4. Regulation
5. Occurrence in the (random, serial, continuous) form
6. Reactivity (eye opening, mental calculation, sensory stimulation, movement, affective state).
7. Interhemispheric coherence (homologous areas).

An important element of the recording of EEG signals is responses of the various components of the EEG to certain neurophysiologic changes.

Signal pattern is essential for identification of the brain activity and clearly differentiate one activity with similar characteristics from another. The brain waves when characterized by their frequency bands are divided into five bands depending upon frequency ranges: alpha (α), beta (β), theta (θ), delta (δ), and gamma (γ)[17].

Brief description about the different waves is given below:

- *Delta waves* have the signals that lie within the range of 0.5-4 Hz and are indicative of cerebral damage or brain disease. These waves have large amplitude and appear during deep sleep. These waves are not observed in awake, normal adult.
- *Theta waves* are the signals of the brain having frequency range of 4-7.5 Hz. Theta waves develop in unconscious mind and are also associated with deep meditation. This rhythm occurs during drowsiness and in certain stages of sleep.
- *Alpha waves* occur as a sinusoidal shaped signal with frequency range of 8 to 13 Hz in adults. These signals are predominant in the adults who are awake but not being engaged in intense mental activity. However, sometimes alpha wave may manifest itself as sharp waves.
- *Beta waves* consist of a fast rhythm with low amplitude within the range of 14-26 Hz. They are mainly associated with an activated cortex and observed during mental activities.
- *Gamma waves* (sometimes called the fast beta waves) are frequencies above 30 Hz and up to 45 Hz related to a state of active information processing of the cortex.

Figure 1.4 represents the signal’s frequency distribution in accordance to various bands of frequency content present in the signal. The figure represents the frequency content present in the signal for one patient that clearly indicates the varying frequency bands present in EEG signal. Frequency spectrum of signals is representing signals in Delta band, Theta band, and Alpha band and Beta band. Figure 1.5 shows the quantity of changes in frequency distribution in all samples. It clearly details the amplitude of components of frequency bands present in the signals.

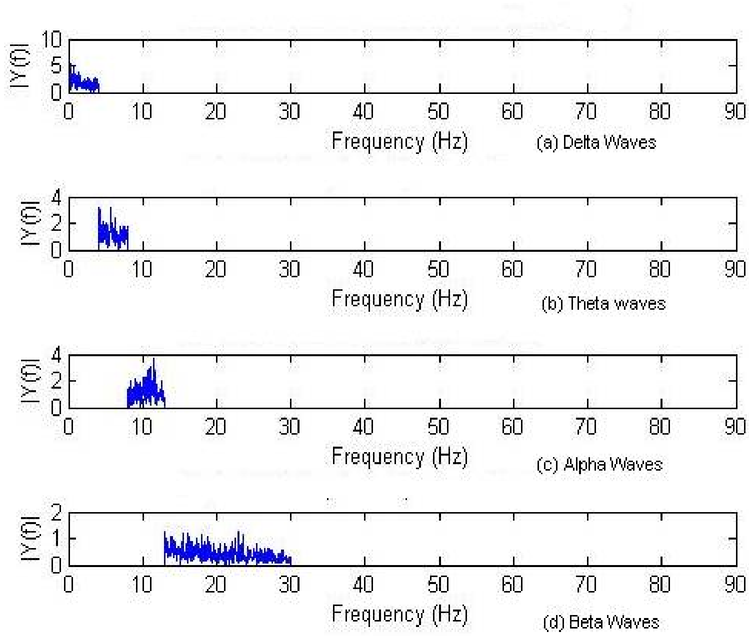


Figure 1.4. Single–sided amplitude spectrum of different frequency band of a signal

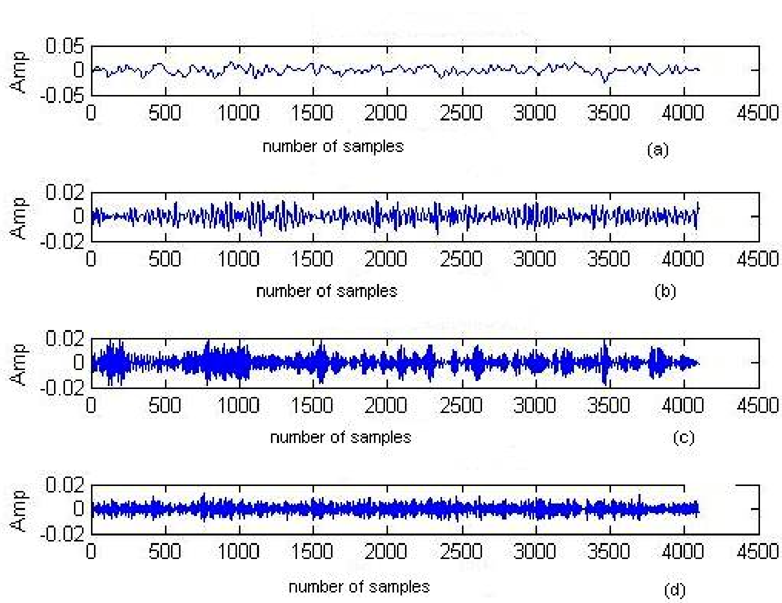
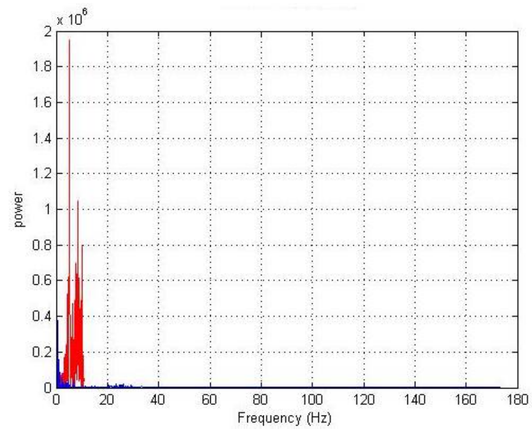
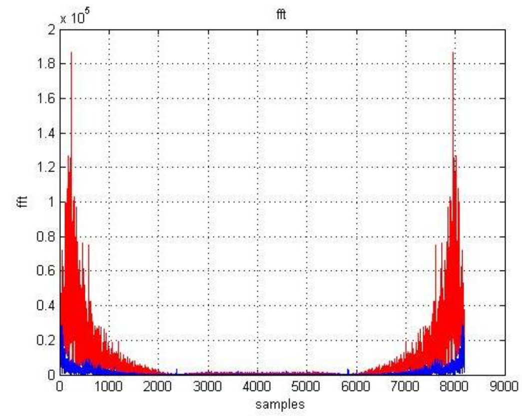


Figure 1.5. Amplitude content of various frequencies present in an EEG signal.



(*a*)



(*b*)

Figure 1.6. (***a***) Power spectrum with reference to frequency, (***b***) FFT of EEG signals of normal and epileptic patient.

Figure 1.6 (*a*) depicts the power spectrum by applying FFT to the autocorrelation sequence and (*b*) represents the power content of a normal and an epileptic patient EEG signal with reference to frequency content in the signals.

1.4 EEG AS A TOOL FOR EPILEPSY

Epilepsy is a disease of the brain caused by spontaneous, intermittent and abnormal electric burst activity in the brain [18]. EEG is one of the main diagnostic tests for epilepsy and an effective clinical tool for monitoring, diagnosing and prognosis of neurological disorders. The onset of a clinical seizure is characterized by sudden changes in the morphology of EEG, but some abnormality in EEG patterns may occur due to different conditions. Epileptic seizures become apparent as characteristic in EEG recordings, their detection can, thus, be used to diagnose, monitor ongoing seizure, and to differentiate epileptic seizures from other paroxysmal, seizure-like symptoms as shown in Figure 1.7. Compared with other measurement methods, EEG is a clean and safe technique for monitoring the brain activity and neurobiological disorders as it provides a visual display of the recorded waveform.

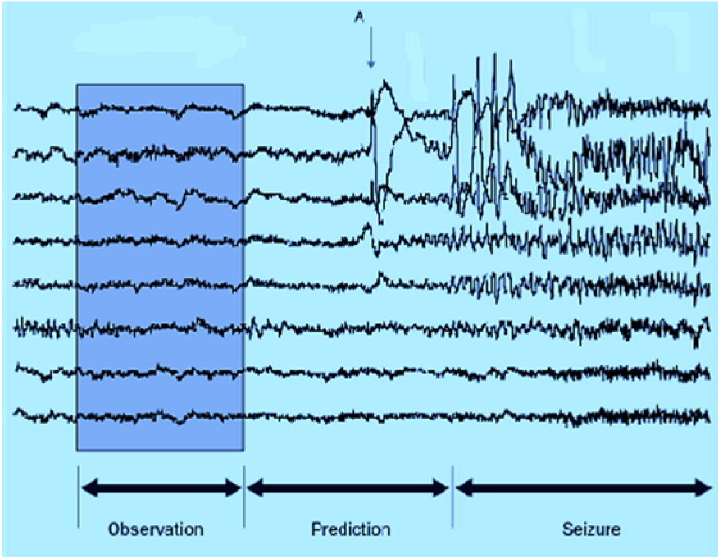


Figure 1.7. EEG signal depicting the beginning of the seizure and during seizure.

Researchers have observed that it may be difficult to identify the components that show momentary changes in the EEG signal in pathological situations such as epileptic seizures [19]. It is not always easy to understand and detect the changes in brain rhythms and waveforms from the scalp EEGs, even with trained eyes. Because of small amplitudes, minute variations, differentiation between foreground and background EEG becomes very subjective and totally depend on the abnormalities. It has also been observed that it is difficult to identify the epilepsy attacks of some patients and the patient has to be monitored asleep and awake. Moreover, the onset of the seizures cannot be predicted in a short interval of time, so a continuous recording of the EEG is required to detect epilepsy. However, the traditional detection or prediction methods including visual and manual scanning of EEG are voluminous, very tedious, time consuming and may be inaccurate and has not yet reached the reliability point to allow clinical translation. The EEG records interpreted by specialists with different types of training may create inconsistent recording of information obtained through the viewing. Hence, it is imperative to analyze the EEG signals using a consistent and appropriate processing method in order to obtain correct diagnoses for the treatment of epilepsy.

Usually, the diagnosis of epileptic seizures involves the analysis based on combination of the medical history of the patient and interpretation by an expert neurologist through EEG recordings [20]. Nevertheless, with the emergence of new signal processing techniques, an increased improvement in the analysis of the EEG for prediction of epileptic seizures has been reported. These enhanced algorithms can detect abnormal disorder and malfunctioning of the brain not only during the seizure but also can detect the onset of seizure. Traditionally, the algorithms for detection of epileptiform activity have been classified as to be either mimetic, linear predictive or template based [21]. However, recent algorithms combine multiple approaches and cannot be classified to these categories anymore.

The automatic detection of epileptic characteristics is arguably the most studied research topic of clinical quantitative EEG. As EEG recordings in any form allows computer aided signal processing techniques to characterize them, and unsatisfactory reliability of the presented methods for clinical use, this topic has been under constant interest. Nevertheless, constantly increasing computation power of computing systems, utilization of more complex algorithms is enhancing the research interest for the detection of epileptic seizures.

1.5 MOTIVATION

Human life is precious and living a good quality life is human right. Epileptic seizures have important public health implications. It is one of the most physically and emotionally destructive neurological disorders affecting population of all ages. Any possibility of alerting a patient and/or his attending staff to an impending epileptic seizure, or anticipating the onset of seizures will have obvious clinical importance.

In the recent years, with the advance signal processing techniques and invasion of this expertise into the field of neurology, considerable effort is invested in detecting and forecasting epileptic patterns. The detection of abnormality should be achieved at an early stage, so that proper and timely action may be taken to avert the impending seizure. An automated analysis and a reliable universal forecaster of seizures can be proving to be very efficient in prognosis of epilepsy. Moreover, by automating the detection of these types of neurological abnormalities, the burden of work on the neurologist can be significantly reduced, response time to the illness can be effectively improved, and suitable medical treatment can be administered within proper time. Also, an automatic seizure detection system if used in the diagnosis of epilepsy, can act as a second opinion tool apart from visual inspection of EEG by the physician.

Therefore the development of accurate computer aided diagnostic system for classification of brain disorders is strongly desired. There is a significant interest in the research community for development of reliable EEG-based automated tools. With the advancement of new signal processing techniques and mathematical algorithms in EEG analysis, supporting methods in medical decision and diagnosis can be developed to avoid tedious analysis of voluminous records and obtain clarity about the brain pathology. This thesis, therefore, investigates and develops a number of promising automatic computer aided diagnostic system for use in these automatic neurological event detection systems. This doctoral thesis, in particular, tries to narrow the gap that exists between present methodology of EEG signal analysis and practical implementation for the benefit of medical fraternity and common man. Main focus lies on developing a system to transform the subjective qualitative diagnostic criteria into a more objective quantitative prognosis criterion and to analyze hidden dynamics of the EEG time series for extracting more information about pathological versus normal status of the EEG signals.

1.6 CLASSIFICATION ALGORITHMS FOR EEG SIGNALS

Classification is one of the main aspects for analysis of EEG signal processing. A proposal for a new classifier for use with our datasets is beyond the scope of this thesis, so standard classifiers have been used to verify the preceding steps and compare them with each other to achieve maximum efficiency and accurate results. There are many different classification methods, but the main focus is primarily on approaches that have low computational requirements, e.g., nearest neighbor, multi layer perceptron neural networks, probabilistic neural networks, support vector machines, radial basis function, Bayes' classifiers.

1.6.1 Nearest neighbor classifier

K-nearest neighbor (*K*-NN) is a simple and robust classifier and performs efficiently with low-dimensional feature vectors [22]. It works by comparing testing data with training data. An identical data in the test and training set is assigned to the same class defined by training data. If the test data is different, a distance measure (e.g., Euclidean distance) is used to determine similarity. Similarity is measured by considering number of closest point. In this algorithm, instead of considering one neighboring point, *K*-nearest neighbour approach is applied that takes *K* points in the training set that are closest to the test point. The test data is assigned the dominant class among its k nearest neighbors within the training set. The default value of *K* is 1, and the default neighborhood object similarity setting is the Euclidean distance, as given by eqn. (1.2)

$$d(x_i, x_l) = \sqrt{\left(x_{i1} - x_{l1}\right)^2 + \left(x_{i1} - x_{l1}\right)^2 + \ldots.\left(x_{ip} - x_{lp}\right)^2} \tag{1.2}$$

K-nearest neighbour classifiers classifies the information based on local information. It draws decision boundaries across the whole space by considering a few local points. The nearness is determined by the measurement of distance between or among matching records. The measurement of distance is also very useful while articulating a confidence of the prediction. It is computationally very efficient at training time, since training simply consists of storing the training set. Testing involves measuring the distance between the test point and every training point, the algorithm works much slower for large data sets.

1.6.2 Naive Bayes classifier

The Naive Bayes classier is simple, computationally efficient, and is often used as a reference classier. It is based on probability theory and asymptotically fastest learning algorithm [23]. It provides an efficient flexible way for dealing with large number of classes. Mathematically, Naive Bayes is a conditional probability model: If x is an instance, represented by a vector $x = (x_1, x_{2,} .....x_n)$ where n represent some features, to be classified in any of class C, then instance probabilities are assigned for each of k possible outcomes or classes

$$p(c_k / x_1, x_2 .....x_n)$$

Using Bayes theorem, Bayesian classifiers says

$$p(c_j / x) = \frac{p(x / c_j) p(c_j)}{p(x)} \tag{1.3}$$

where

$p(c_j / x) =$ probability of instance d being in class c_j

$p(x / c_j) =$ probability of generating instance d given class c_j

$p(x) =$ probability of instance d occurring

$p(c_j) =$ probability of occurrence of class c_j

Bayes classifier assigns a class label $y = C_k$ for some k by:

$$y = \arg\max_k \; p(c_k) \prod_{i-1}^{n} p(x_i / c_k) \tag{1.4}$$

Naïve Bayes has limitation of oversensitivity to redundant and irrelevant data or features. If two or more attributes are highly correlated, the final decision of classification can be misleading as correlated attributes receive too much weight. This may lead to a decline in accuracy of prediction if large number of correlated features is present in the dataset.

1.6.3 Artificial neural network classifier

Artificial Neural Networks (ANN) is a mathematical model inspired by the functional and morphological aspect of biological structure of neurons. It is an adaptive system that changes its behavior based on information provided to the network during the learning phase. The popularity of ANN as a classifier is attributed to their ability (*i*) to be used to generate discriminative likelihood-like scores; (*ii*) to be easily implemented in hardware and embedded platform for its simple structure; (*iii*) to approximate functions and similarity based generalization property; (*iv*) to map complex class distributed features easily [24]. NNs imitate the structure of biological neural network with neurons as processing elements. Each node representing a neuron, receives unknown samples as inputs, processes these inputs by applying threshold functions at every layer and finally generates a single output. The number of layers and neurons, the respective synaptic weights and learning algorithms can be varied in accordance to desired design perspective [25]. Neural network is trained by making it learn, by providing number of inputs through an iterative process of adjustments applied to synaptic weights and thresholds. The neuron model depicted in Figure 1.8 is defined by three basic elements:

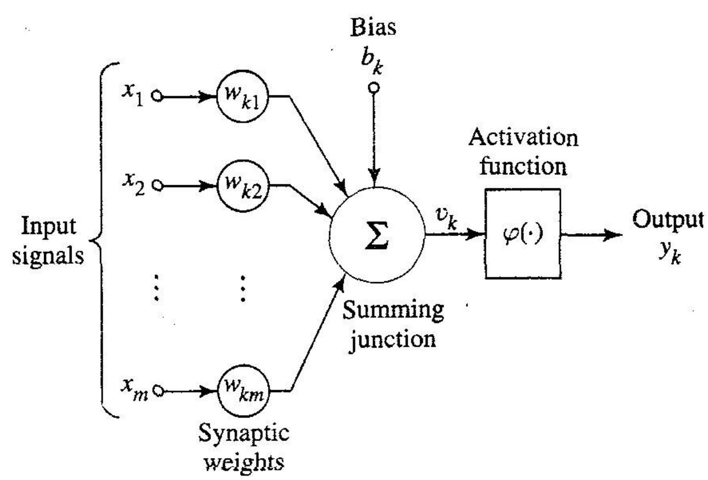


Figure 1.8. Neuron model of Neural network for one node output.

1. Input layer with one or more inputs which accept a signal x_j at the input of synapse j. It is further connected to neuron k after it is multiplied by the synaptic weight w_{kj}. If the associated synapse is excitatory, the weight w_{kj} is positive, and if it is inhibitory the weight will be taken as .negative.

2. An adder for summing the input signals, weighted by the respective weights of the neuron.
3. An activation function or threshold function for limiting the amplitude of the output of a neuron. The model of a neuron also includes an externally applied bias $w_{k0} = b_k$ that has the effect of lowering or increasing the net input of the activation function.

Steps involving NNs process:

Initialization

Determine network topology for to determine nodes in all layers.

Set initial weights for all the connections, w

Set value of η, learning rate parameters

Iterative loop

While exit condition is false, DO:

Input a given vector x

For each k, compute the response $y_k(n)$ produced by input set applied to the first layer of the network in which neuron k is embedded.

$$u_k = \sum_{j=0}^{n} w_{kj} x_j , v_k = u_k + b_k$$

$$y_k = \phi(v_k) \tag{1.5}$$

Calculate the error signal at the output node

$$e_k(n) = y_k(n) - d_k(n) \tag{1.6}$$

where $d_k(n)$ denote desired response for neuron *k* at time *n*, and y_k(n) denotes the value of the actual response (output) of this neuron. The input vector and desired response constitute a pair presented to the network at time *n*. At any time, for any input signal, actual response of neuron k is different from the desired response which is represented as $e_k(n)$.

Minimize and optimize the cost function

The ultimate purpose of this algorithm is to minimize a cost function. The instantaneous value of the mean square- error is the cost function and is given by eqn (1.7)

$$J(n) = \frac{1}{2}\sum_{j=1}^{n} e^{2}{}_{k}(n) \tag{1.7}$$

The network is optimized by minimizing $J(n)$ with respect to the error signal obtained, which in turn depends on synaptic weights of the network. The synaptic weight adjustment is given by

$$\Delta w_{kj}(n) = \eta e_k(n) x_j(n) \tag{1.8}$$

in accordance to the error-correction learning rule (or delta rule).

Update weights to all nodes depending upon the learning process.

If $w_{kj}(n)$ is the value of the synaptic weight at time n, an adjustment $\Delta w_{kj}(n)$ is applied to the synaptic weight $w_{kj}(n)$, yielding the updated value as

$$\Delta w_{kj}(n+1) = w_{kj}(n) + \Delta w_{kj}(n) \tag{1.9}$$

Exit if the condition is true

no change in the weights or error is minimum

Different ANN models are discussed in the following section

1.6.3.1 Multilayer feed forward network

The most commonly used representative of ANNs is the multilayer perceptron (MLPNN) [26]. MLP can solve complex classification tasks, but is sensitive to overtraining, especially with noisy and nonstationary data as EEGs. The input signals (input vector) applied to the input-layer of the network is connected to the neurons in the second layer (i.e. the first hidden layer). The output of the second layer acts as an inputs to the third layer, and so on for the rest of the network till output layer is accessed. The signals achieved from the neurons in the output layer represent the overall response of the network to the activation pattern supplied by the source nodes in the input layer.

1.6.3.2 Back-propagation networks

The back-propagation algorithm is based on delta rule and gradient descent. It includes forward and backward passes; in the forward pass, the information applied to the input moves in the forward direction after computations at each layer. In backward pass, error computed between the resulting output and target value is passed in backward direction for weight updation to achieve minimum error for desired response. The change in weight is given by $\Delta w_{kj}(n) = \eta \delta_j(n) y_i(n)$ where $\delta_j(n)$ is the local gradient at each neuron. The value of local gradient differs depending upon whether the neuron is in hidden layer or outer layer

neuron in output layer $\delta_j(n) = e_j(n)\phi'(v_j(n))$ (1.10)

neuron in hidden layer $\delta_j(n) = \phi'(v_j(n))\sum_k \delta_k(n) w_{kj}(n)$ (1.11)

The significant improvement in performance can be observed by using either Newton's method, conjugate gradient, or the Levenberg–Marquardt (LM) optimization technique [27].

1.6.3.3 **Radial Basis Function NN classifier**

Radial Basis Function (RBFNN) networks feed-forward supervised training algorithm configured in three layer [28]. The input layer comprises of nodes, number is equal to the dimension *n* of the input vector *x*. The hidden layer comprising nonlinear units are connected directly to all of the nodes in the input layer as shown in Figure 1.9. Gaussian function (the width corresponding to the variance, σ_i having peak at zero distance) is taken as the basis functions as activation function and the least squares (LS) criterion is used as objective function. RBF network adjusts iteratively parameters of each node by minimizing the LS criterion.

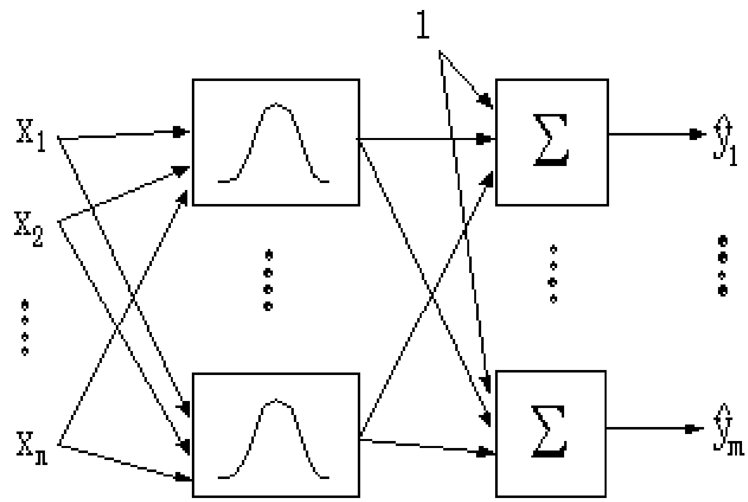


Figure 1.9. Model of RBF neural network

Steps involving RBFNNs process:

The radial distance d_i, between the input vector x and the center of the basis function c_i is computed for each unit i in the hidden layer as $d_i = \| x - c_i \|$

The output h_i of each hidden unit i is computed by applying basis function to this distance $h_i = G(d_i, \sigma_i)$

The hidden unit transforms the input space nonlinearily, and the output space transform hidden unit the in linear fashion. The j^{th} output at the output neuron is computed as

$$y_j = f(u) = w_{0j} + \sum_{i=1}^{L} w_{ij} h_j \quad j = 1,2,\ldots.M \tag{1.12}$$

Summarizing, the mathematical model can be expressed as:

$$y_j = f(u) = w_{0j} + \sum_{i=1}^{L} w_{ij} G(\| x - c_i \|) \tag{1.13}$$

$$y_i = w_0 + \sum_{j=1}^{L} w_j \exp(\frac{\| x_i - x_j \|^2}{d_j^2}) \tag{1.14}$$

where y_i is the output value , input feature vector is x_i, w_0 is a constant bias term, x_j and d_j are the centers and widths for j^{th} hidden neuron in the network, respectively. L is the total number of hidden neurons in the network and w_j are the weight coefficients of connections between hidden neurons and the output neuron. In comparison with other multilayer feed forward neural networks, RBFNNs have simple topological structure, locally tuned neurons, and ability to have a fast learning algorithm.

1.6.4 Probabilistic Neural Network Classifier

The Probabilistic Neural Network (PNN) is an extension of NN that is capable of approximating the Bayes classifier and convergence to Bayes-optimal decision surface. It estimates the similar probability density function for each class based on the training samples [29]

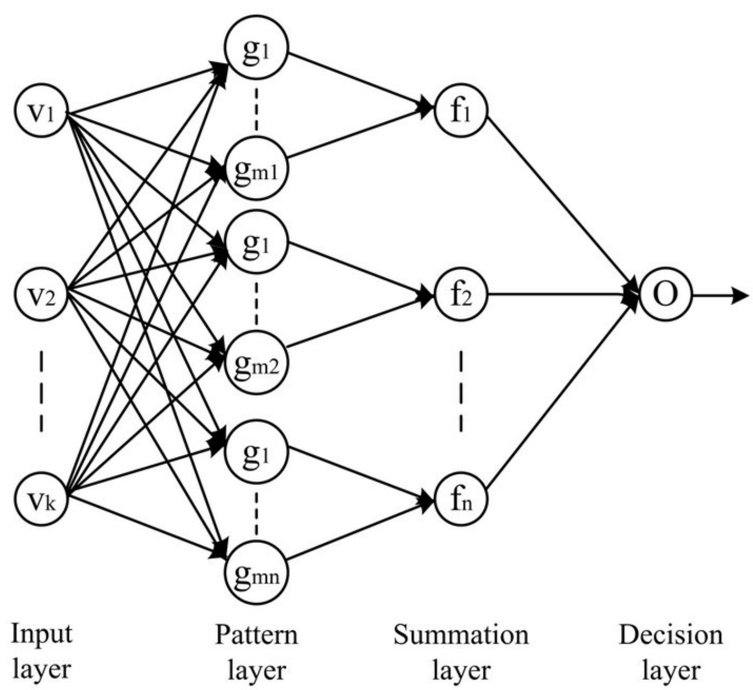


Figure 1.10 PNN Model with four layers for classification of signals.

As shown in Figure 1.10, PNN consists of four layers, each layer performing a particular function; as input layer accepts the input pattern or feature vector, pattern layer consists of the vectors of the training set, summation layer deal with winner takes all approach and output layer represents each of the possible classes. As shown in figure the input layer is fully interconnected with pattern layer, and output layer classifies the data from the information obtained by Summation layer.

The input feature vector receives all inputs, class nodes are connected to the example hidden nodes, so the example vector activations determines the class of the input feature vector. Mathematically, if E is the example vector, and F is the input feature vector, sum of the products of the example vector and the input vector are calculated for each class given by $h_i = E_i F$

The class output activations for each node are then defined by eqn (1.15):

$$c_j = \frac{\sum_{i=1}^{N} e^{\left(\frac{h_i - 1}{\sigma^2}\right)}}{N} \tag{1.15}$$

Where, for i^{th} input feature vector, N is the total number of example vectors for a class, h_i is the hidden-node activation, and σ is a smoothing factor which is chosen through experimentation and determines the performance and generalization of the classifier. The advantage of PNN is fast classification, better generalization as no training is required. It is more adaptable and dynamic as it is easy to add new examples by adding the new hidden node even in the context of noisy data.

1.6.5 Support Vector Machine (SVM)

The fundamental concept of SVM is to transform a set of feature vectors into a higher dimensional space. An optimal hyper-plane is searched in the space that can maximize the margin between classes. In comparison to linear discriminate functions, conversion to higher dimensional space result in simplification of complex classification problems [30].

Let a training feature vector $x_j \in R_d$ with associated labels $y_j \in \{1,-1\}$ belong to linearly separable classes, where hyper plane H_0 is the decision surface to classify a pattern as depicted in Figure 1.11. The problem of classifying a test vector x_k as belonging to one of two classes can be written as:

$$f(x_k) = w.x_k + b \tag{1.16}$$

where $w \in R_N$. The equation f (x) = 0 gives H_0, region of vectors x. If H_a and H_{-a} are two hyper planes parallel to H_0, defined by f (x) = a, and f (x) = - a, then the distance separating these two hyper planes is given by d =2/|| w ||

The distance d (margin), is maximized so as to obtain a classifier boundary that does not over fit to the training data x_j. To be correctly classified, the training vectors should lie outside the margin or on the margin boundary and must satisfy:

$$w.x_k + b \geq 1, \text{ for } y_j = +1, \tag{1.17a}$$

$$w.x_k + b \leq -1 \text{ for } y_j = -1. \tag{1.17b}$$

This can be written more concisely as:

$$y_j(w.x_k + b) \geq 1 \forall j \tag{1.18}$$

As real-world biomedical data are often not linearly separable, with considerable overlap between classes, the decision boundary can be softened by introducing a slack positive variable ξi for each training vector. The conditions that the training vectors must satisfy, defined in eqn (1.17) and (1.18) can now be modified to include ξi such that:

$$(w_i.x + b) \geq 1 - \xi_i, \xi_i \geq 0 \text{ for } y_j = +1,$$

$$(w_i.x + b) \leq 1 - \xi_i, \xi_i \geq 0, \text{ for } y_j = -1. \tag{1.19}$$

The introduction of ξi is problematic in that the constraints in eqn (1.19) will be met for all i if ξi is suitably large, thus giving eqn (1.20)

$$y_i(w_i.x + b) \geq 1 - \xi_i, \xi_i \geq 0 \tag{1.20}$$

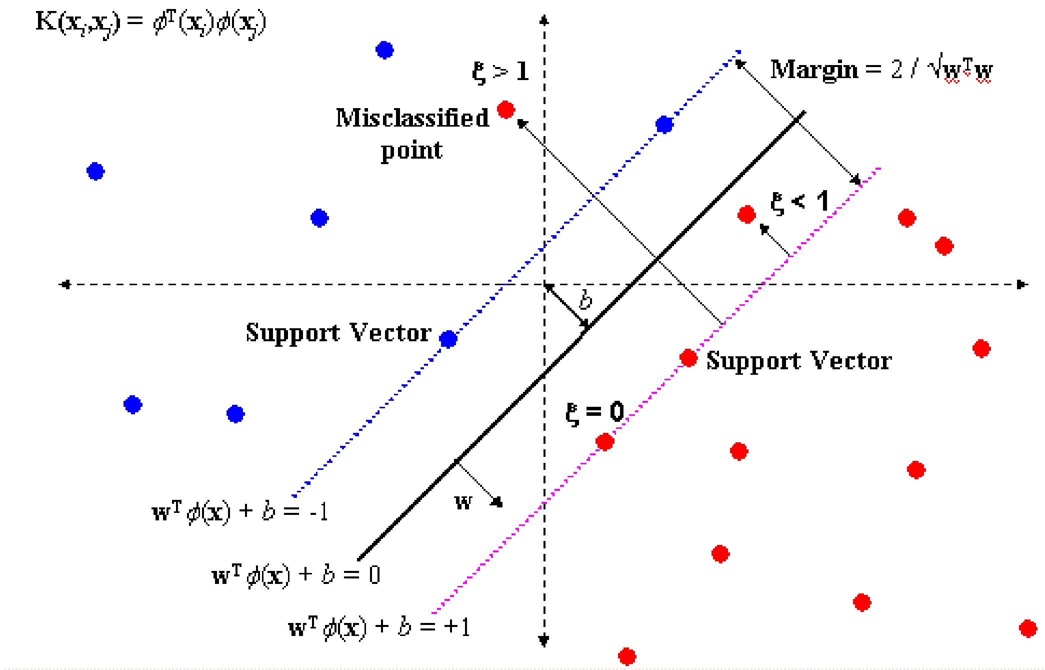


Figure 1.11. SVM classifier with maximized margin with defined hyper plane

To avoid trivial solutions, a regularization constant C is introduced into the objective function which now becomes:

Minimize

$$\frac{1}{2}w^T w + C\sum_{i=1}^{m}\xi_i \qquad (1.21)$$

Subject to

$$y_i(w_i.x + b) \geq 1$$

The regularization parameter C thus controls the degree of penalization introduced by ξi, such that increasing C permits fewer training errors at the expense of reduced generalization. The convex optimization problem outlined by eqn (1.21) is solved with Lagrangian multipliers α_j. Only training patterns lying on the margin surface or within the margin have non-zero α_j; these are the support vectors. The classification process thus consists of assigning one of the two classes to a given input vector x_k of dimension N, such that:

$$f(x) = \sum \alpha_i y_i x_k^T x + b \tag{1.22a}$$

$$w = \sum \alpha_i y_i x \quad b = y_k w^T x_k \quad \text{for any } x_k \text{ such that } \alpha_k = 0 \tag{1.22b}$$

The linear classifier relies on dot product between vectors $K(x_i, x_j) = x_i^T x_j$

If every data point is mapped into high-dimensional space via this transformation

Φ: $x \rightarrow \varphi(x)$, then the dot product becomes: $K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$

where K is a kernel function. A kernel function is any function that corresponds to an inner product in some expanded feature space [31]. The different kernels used in literature are

Linear: $K(x_i, x_j) = x_i^T x_j$, Polynomial of power p: $K(x_i, x_j) = (1 + x_i^T x_j)^p$

Gaussian (radial-basis function network): $K(x_i, x_j) = \exp\left(\frac{\| x_i - x_j \|^2}{\sigma^2}\right)$ and

Sigmoid: $K(x_i, x_j) = \tanh(\beta_o x_i^T x_i + \beta_1)$

In this research work, RBF kernel is used as the kernel function because most of the biomedical researchers consider this function an ideal one. In SVM, classification calculation times are small and it is easy to implement in a real-time system. They are also less prone to over fitting and obtain good generalization performance without feature space dimensionality reduction.

1.6.6 K-Means Algorithm (KMA):

KMA is one of the simplest unsupervised learning algorithms to solve clustering problems. The k-means algorithm is applicable for grouping data points into K clusters according to the distance measure. The input parameter of this algorithm is the number of clusters *K*. If this number is unknown, we can perform the k-means algorithm several times for different numbers of clusters and we can then choose the best one.

The algorithm of KMA is summarized as follows [32]:

Step 1: Choose K initial cluster centers z_1, z_2,z_3…..z_k randomly from the n points {X_1,X_2,X_3….X_k }.

Step 2: Assign point X_i, i= 1,2,3….n to the cluster Cj, $j\ \varepsilon$ {1, 2,…, K}

If $\| X_i - z_j \| < \| X_i - z_p \| \ p = 1,2.....k,\ j \neq p$

Step 3: Compute new cluster centers as follows

$$Z^{new}{}_i = \frac{1}{n} \sum_{x_{j \in C_i}} X_j, i = 1,2.....k \tag{1.23}$$

where n_i is the number of elements belonging to the cluster C_i.

Step 4: If $\| Z^{new}{}_i - z_i \| < \varepsilon, i = 1,2.....k \|$, terminate. (1.24)

Otherwise continue from step 2.

An indisputable advantage of k-means is that it can be used for very large datasets.

1.7 CLASSIFICATION METRICS OF THE CLASSIFIERS

The performance of classifier is assessed by five different statistical measurements, Sensitivity (Sen), Specificity (Spec), Confusion Matrix (CM), Classification Accuracy (CA) and Receiving Operating Characteristic (ROC) curve. The definition of sensitivity, specificity and accuracy [33] and the brief description of the confusion matrix and ROC curve [34] are given in the following section:

1.7.1 Confusion matrix

Table1.1 : Confusion Matrix for two classes classification

	POSITIVE(Seizure)	**NEGATIVE(Normal)**
POSITIVE	*True Positive*	*False Negative*
NEGATIVE	*False Positive*	*True Negative*

Confusion matrix is tabular description of classification of cases in the test dataset. In confusion matrix, the columns denote the actual cases and the rows denote the predicted class. Both the sensitivity and specificity measure classification accuracy against the ground truth diagnosis and are defined by Sensitivity is an assessment of the two upper quadrants while specificity is represented in the lower two quadrants. They are both bounced between 0 and 100%, and higher values indicated better accuracy.

a) Sensitivity: the ratio of number of true positive (TP) decisions divided by the number of actual positive cases.

$$Sensitivity = \frac{TruePositive}{TruePositive + FalseNegative}$$

b) Specificity: the ratio of the number of true negative (TN) decisions divided by the number of actual negative cases.

$$Specificity = \frac{TrueNegative}{TrueNegative + FalsePositive}$$

c) Classifications accuracy: the ratio of the number of correct decisions divided by the total number of applied cases.

$$Accuracy = \frac{TrueNegative + TruePositive}{All}$$

1.7.2 ROC curve

ROC curve is graphical representation of classification of test cases. It represents all the cases, actual and predicted on the plots with sensitivity (true positive rate) represented on the Y-axis and (1-specificity) (false positive rate) on X-axis. Area under the ROC curve is preferred to evaluate the performance of a classifier and its value lies between 0 and 1. When area of the ROC curve is 1, the classifier has a perfect discriminating ability, and if value is 0.5, classifier cannot discriminate between two classes. The AUC value has the advantage that it works just as well for discrimination problems with different class priors.

1.7.3 ***K-fold cross-validation*** is one way to improve the holdout method. The complete data set is divided into ' *k* ' mutually exclusive subsets of an equal size. Each time ' *k*-1' subsets are used as training subset and one subset is used as testing set with sample of every class in each subset. Every data point appears in a test set exactly once and appears in a training set (*k*-1) times. All the datasets are used for training and testing by continuously shifting them in different *k* subsets. The average error across all *k* trials is computed to yield a single measure of the stability of the respective model, i.e., the validity of the model for predicting new observations. The disadvantage of this method is the necessity of rerunning the training algorithm from scratch *k* times, which means it takes *k* times as much computation for the evaluation. The advantage of this method is the ability to independently select the size of each test and number of trials.

1.8 OBJECTIVES

The motivation behind this thesis is the examination of morphology and topography of waveforms obtained in EEG during certain neurophysiologic phenomena. We intend to use power of EEG as a source for detection, diagnosis, treatment and prognosis of several neurological abnormalities. Following the need and being motivated by these factors, following objectives were framed for our research work.

I. To employ quantification and statistical analysis techniques to differentiate and select the significant and contributing features of the EEG signals.
II. To employ quantitative analysis to identify and classify abnormality by developing a prediction CAD model using highly discriminating features.
III. Inception of an ensemble computer-aided diagnostic classifier system (ECAD) for seizure activity.
IV. To design a CAD system while selecting the appropriate module (epoch) to understand neural underpinnings and quasi stationarity nature of EEG signals.

1.9 ORGANIZATION OF THESIS

Chapter 1 lays the foundation as to how and why EEG signals are clinically significant for analysis and classification of epilepsy. The chapter begins with the documenting facts about EEG signals, epilepsy and correlation between EEG signals and epilepsy. This chapter also includes the basic introduction to the classifiers used for designing proposed CAD system and the performance metrics employed throughout the thesis. Finally, motivation, outline and objectives of the present research work are outlined.

Chapter 2 presents a comprehensive literature review of the state of the art of related studies carried out so far. We also discuss the state-of-the-art in dynamics of EEG signals and their classification. The review has been carried out by dividing the problem of automated classification system into different modules.

The goal of **Chapter 3** encompasses the extraction and selection of different features, extracted from clinical acquired EEG signals. In this objective, an efficient statistical analysis is presented to extract informative parameters set for classification of EEG signals that can be further used for deciding and selecting processes to achieve the desired objective. The strength of this study is its rigorous feature selection procedure which when applied to the prediction model allows a high generalization and accurate classification with good interpolation.

Chapter 4 This Chapter discusses the designing of architecture of different machine learning algorithms which are employed for the classification. The purpose of this chapter is to elicit classifiable information from human EEG and to identify the algorithm that provides the highest classification accuracy of epilepsy. In this objective, a novel HCAD system is proposed that focus on modelling the seizure classification by a hierarchical framework, a variation of the classifier system.

In **Chapter 5** we have extended the model introduced in Chapter 4 to infuse the features, in order to assess the advantage of including non linear features along with high order cumulants. The study focuses on the development of a classifier paradigm that consists of ensemble of significant features. Various CAD systems performances are evaluated to investigates which classifier is more efficient in classifying the seizures

Chapter 6 This chapter investigates the possibility designing a CAD system for prognosis of epileptic seizure with less computational complexity and higher accuracy. In this chapter emphasis is laid on quasi stationary nature of signals by dividing complete data set into small segments (epochs) where the stationarity of the signals can be checked.

In **Chapter 7** we summarize the conclusions achieved from the work presented in the previous chapters and we outline possible future scope of the problem and future directions that can be followed for the extension of this work.

CHAPTER 2

REVIEW OF STATE-OF-THE-ART

Clinical neurology is main area of research that deals with diagnosis of epileptic patients through EEG, sleeps analysis and may other abnormalities of brain. As the area of application of EEG is very wide, there are a large number of publications in which authors have dealt with EEG signals, their processing, analysis and applications. The review suggests that the current state of-the-art in the field is progressing with many groups working in this field. It is impossible to mention all the publications that have dealt with processing these biomedical signals, so in the next few pages, only studies closely related to the topic of this thesis are presented. An exhaustive review of the literature relevant to this thesis is attempted and presented. This review primarily focus on artifacts affecting EEG signals, quantitative and qualitative EEG analysis, time and frequency domain analysis, and use of soft computing techniques for signal processing.

2.1. EPILEPTIC SEIZURES

Sudden excessive electrical discharges in the brain cells are the primary cause of epileptic seizures that results in malfunctioning of the brain electrophysiological system [35]. It is difficult to identify variations and changes in the amplitude or frequency of EEG waveform by naked eye observations. So, researchers opt for this field to study the mechanisms behind seizures, design and develop several algorithms and models for the identification of changes during seizure activity. Authors in [5] reported that by studying seizure intensity, it is possible to detect epileptic seizure and can be employed as a clinical tool. Niederhoefer et al. [36] developed algorithms to detect the onset of epileptic seizures using raw data information. EEG has capability to reflect all the activity of the brain, so it has been found to be a very powerful tool in the field of neurology and clinical neurophysiology [37-38].

The extraction of information and its analysis from raw EEG signals becomes difficult as these signals are contaminated by biologically generated and externally generated signals. Since many physiological signals (such as heart beats) are involuntary, some artifacts will always be present in EEG signals [39]. The presence of these kinds of artifacts makes it difficult to differentiate between original brain waves and noise [6]. Recognition and elimination of the artifacts in real – time recordings is a complex task, but essential to the development of practical systems. The artifacts compromise sensitivity of main signal as they are confounded with statistical contrasts [40]. The removal of the artifact becomes essential as the preprocessing step for any type of EEG analysis. Lot of research work has been reported to remove these artifacts occurring from various sources so as to enhance the clinical usefulness of EEG signal [41-43].

2.2. LITERATURE REVIEW OF ARTIFACTS

The presence of artifacts introduce spikes which can be confused with neurological rhythms, making the EEG signals analysis biased and difficult leading to wrong conclusions [44]. The various kind of physiological and extra physiological artifacts which prominently affect true EEG are Instrumental artifacts: generated by the use of an instrument powered from the mains power supply, Analysis artifact: that arise in the course of processing the signal and Biological artifacts: signals arising from different part of body [45]. The other major physiological artifacts are the signals generated from heart, muscles, and eyes, head movement, sweating and breathing [46]. Mains voltage artifact may also appear on the EEG via the use of fluorescent lights or other electrical equipment in the close vicinity of the EEG machine.

An issue of concern in analysis of EEG is the detection and elimination of artifacts leaving underlying background signals due to brain activity intact[47,48]. Accurately evaluating the performance of artifact removal algorithms presents a less straightforward challenge than that of evaluating detection algorithms.

2.2.1. Artifact Recognition and Elimination

The different methods available for reducing and removal of artifacts are application of spatial filters [49], blind source separation [50], and linear regression models, in time as well as frequency domain [51, 52]. In 2003, Durka et al. [53] have used a simple but effective technique for discriminating a “good” EEG and artifacts by optimizing the threshold limits to mark an epoch as an artifact. The optimized parameters are directly related to the signal’s energy distribution, in the frequency or time domain. Authors in [54] have used higher order statistical property, kurtosis, and the 4th cumulant of data to make a clear distinction between non-artifact and artifact signal, and rejecting the later. As the artifacts have overlapping spectra with signal of interest, they have to be removed such that the useful information is not lost. Many researchers have introduced many novel methods for removal of artifacts.

Regression methods were introduced by Quilter et al. [55] and subsequently modified by Verleger et al. [56]. Gratton et al. [57] proposed a time domain regression method in which the scaling factors are computed and averaged separately for each epoch, and further modified the technique by providing separate propagation factors for blinks and saccades. In 1993, Lins et al. [58] and Lagerlund et al. [59] used PCA-based methods to eliminate the artifact components. These components are obtained by decomposing the artifacts contaminating EEG signal. Makeig et al. [60] in 1996 used the algorithm of Bell and Sejnowski [61] proposed in 1995 to report first application of ICA for EEG data analysis. A modification in ICA proposed in 1997 by Hyvärinen and Oja [62], was Fast ICA based on “non-Gaussianity" of underlying components. In 2000, ICA based artifact reduction method was proposed by Everson and Roberts [63] for the artifact removal. [64] Proposed use of principal component analysis (PCA) to remove eye artifacts from EEG.

In 2002, Nicole and Berg [65] used spatial filtering for artifact correction and focused on the pre selection approach for the segments of EEG signals. Another filtering technique which does not involve any reference signal and is totally based on statistical approach is Wiener filtering [66] and based on probabilistically estimation approach is

Bayesian filtering [67]. In 2004, Joyce et al. [68] proposed a method using the second-order Statistics-based Blind Source Identification (SOBI) algorithm for the removal of eye movements. In 2005, Krishnaveni et al. [69] have done an extensive comparison of the entire present ICA algorithm like MS-ICA, SHIBBS(Shifted Block Blind Separation), Kernel-ICA, JADE and RADICAL(Robust, Accurate, Direct ICA) for removal of ocular artifacts from EEG and assessed them in terms of quantitative analysis by using a reliable Mutual Information Estimator. In 2006, Krishnaveni et al. [70] using both the concepts of ICA and wavelets for artifact suppression and elimination, preserved spectral and coherence of neural activity. Correa et al. in 2007 [71] used three adaptive filters in cascade to cancel line interference, ECG and EOG artifacts present in EEG records. A hybrid soft computing technique, Adaptive Neuro-Fuzzy Inference System (ANFIS) is proposed by S.Kezi et al. [72] to estimate the interference and to separate the artifacts from EEG signal. Dewan et al. [73] have performed the separation of ECG artifact by detecting R –peaks using adaptive thresholding method.

In 2008, Devuyst et al. [74] have based their research on a modification of the ICA algorithm using a single-channel EEG and ECG. Their approach gave promising results as compared to earlier proposed techniques. In 2009 algorithm proposed by Jones et al. [75] involves minimizing the error optimally. The choice of algorithm employed determines the efficiency and cost of the filters. In 2010, Gao et al. [76] have used Canonical Correlation Analysis (CCA) technique using correlation threshold to remove the EMG artifacts automatically, without eliminating the signal of interest. Arezki et al.[77] have used LMS with computational complexity using step size to control the rate of adaption. Dong et al. [78] used JADE method for removing ocular artifacts for both eye blinks and saccades and concluded that it is an effective tool for multichannel EEG recordings. In 2011, Babu et al. [79] proposed a method to remove artifacts using wavelet transform and adaptive filter (Fast RLS algorithm). This technique not only improves the quality of EEG signal but also increases the PSNR (Peak Signal to Noise Ratio) value and decreases the elapsed time in comparison to RLS algorithm.

In [80] the recording of a noise-contaminated and a noise-free signal is recorded concurrently. Quantitative studies about suppression of OA distorting the underlying cerebral activity and accurate evaluation on denoising effect of DWICA are covered in [81]. In [82] authors have proposed effective method to achieve the muscle artifact removal from single-channel EEG, by combining canonical correlation analysis with empirical mode decomposition. Authors in [83] conducted the blind source separation on the raw EEG recording by the stationary subspace analysis (SSA) algorithm taking mean and the covariance matrix into account. In [84], automated artifact elimination was achieved by image processing algorithms using linear discriminant analysis (LDA) for classification. Once the acquired EEG signals are preprocessed by removing artifacts, amplifying and normalizing the signals, further processing is done for choosing the classifier.

The two primary considerations for the design of automated detection system are: nature of features selected from the EEG input signal and the analysis techniques applied on extracted features. The review in the upcoming sections covers the research work conducted in these two primary areas.

2.3. FEATURES EXTRACTION AND SELECTION

An efficient way to analyze the signals (with volume and complexity) is feature selection [85, 86], which results in dimensionality reduction by selecting a subset of features from the whole set of inputs. Lot of research work has been carried out to measure the goodness of a feature subset in determining an optimal one. Different feature selection methods can be broadly categorized into the two models: wrapper model [87] and the filter model [88-89]. Both of these models use algorithm to determine the goodness of the selected subsets that are independent of any learning algorithm. To search optimal set of candidate features for evaluation, search strategies such as complete, heuristic, and random search have been studied [90]. High-dimensional raw EEG data often contains many redundant features that affect the computational complexity and processing speed and accuracy of learning algorithms.

Individual rank of features is evaluated and irrelevant features or redundant features those have similar rankings can be removed [91]. Several methods are reported in the literature for extracting quantitative features from EEG signals. Fourier transform is used to find spectral parameters that are used for detection and classification of epileptic seizure [92]. As proven, the signal being analyzed by Fourier transform is assumed to be stationary, but studies have shown that the EEG signal is a non-stationary [93].

In recent years, attempts have been reported on seizure detection and prediction from EEG analysis. The same group further used wavelets and neuro fuzzy system to extract the features [94] Features based on time frequency analysis were presented by many authors. Tzallas et al. [95] and Srinivasan et al. [96] employed time domain and frequency domain features to Elman recurrent neural network for classifying EEG signals. In [97] Guerrero-Mosquera, used stochastic analysis approach for feature extraction and have used hybrid feature selection technique. H. Ocak, et al [98] investigated entropy and approximate entropy for discriminating EEG signals. S Liang et al [99] used time frequency analysis and approximate entropy to detect epilepsy using linear least square method and linear discriminant analysis.

H. Adeli et al. [100] have reported seizure prediction using artificial neural networks with wavelet pre-processing whereas Subasi et al. [94] have used neuro-fuzzy system for seizure detection. Varun Bajaj [101] has classified the EEG signals using intrinsic mode functions generated by empirical mode decomposition using SVM classifiers. In [102] effective and flexible preprocessing is done by appropriate selection of the position electrode channel, the scope of signal source, and frequency bands. They have calculated the maximum Lyapunov exponent and used wavelet packet transform to calculate the average energy. A comprehensive survey paper of the feature extraction methods in the EEG research that formalizes the relevance of the Electroencephalography data analysis in the health applications are covered in [103-104].

Researchers in this field have observed that EEG data has significant non linearity during seizure activity. Iasemidis and Sackellares [105] used principal Lyapunov exponent for predicting seizures after applying nonlinear dynamical techniques methods based. Lehnertz and Elger et al [106] employed nonlinear dynamics to larger datasets, greater numbers of patients for seizure prediction. Lyapunov exponents were further expanded by Güler & Übeyli in [107] and in [108] computed correlation dimension, Lyapunov exponents from depth-EEG signals. EEG signal has the characteristics of randomness, non-stationary, nonlinear, diversity, so only one feature can't describe complete signal. So, features extraction and selection method has attracted author's attention.

An automated system to detect the nature of the seizures and to classify normal, inter-ictal, and ictal states are based on the features extracted using the discussed techniques are reviewed in the next section.

2.4. EEG ANALYSIS AND CLASSIFICATION

Earlier, EEG analysis referred to interpreting the EEG waveforms using descriptive methods. With the advancement in this field, several changes in the EEG signal are analyzed with various methods. The main EEG analysis methods fall under four broad categories: time domain, frequency domain, time–frequency domain, and non linear methods. Various different models have been suggested based on these methods, to assist neurologists in identifying epileptic activities. Pradhan et al. [109] used two layered learning vector quantization networks in order to analyze EEG signals. Petrosian et al. [110] applied EEG signals to recurrent neural networks instead of applying them to statistical properties in the identification of epileptic seizures. Analysis of EEG after artifact extraction and removal has been taken up by A. Saastamoinen et al [111] using RBF.

Classification algorithms that have used features such as standard deviation, median arithmetic mean, zero crossing value, wavelet transform, rényi entropy spectral entropy, are reported in literature. Time-domain approaches for detection of epileptic seizures in EEG signals include approaches such as the linear prediction (LP), fractional linear prediction and principal component analysis (PCA) based schemes. An investigation of recent studies for EEG analysis is carried out in this section and summarized the findings of many automated epilepsy activity classification techniques.

2.4.1. *Automated detection of epileptic seizures using Soft Computing Techniques*

Automated real-time detection of epileptic seizures has always been a challenge, several algorithms have been developed to discriminate the seizure (ictal period) from the normal period of the EEG. Neural networks have the ability to capture the dynamics of complex system; therefore they are exhaustively used to analyze nonlinear systems. H. Adeli et al. [100] have reported seizure prediction using artificial neural networks with wavelet pre-processing. Nigam et al [113] used a nonlinear preprocessing filter for the automated detection of epileptic signals with Artificial Neural Network. Authors in [114] have used nonlinear parameters to characterize the EEG signal and classified EEG signals into normal and epileptic using different entropies using an Adaptive Neuro-Fuzzy Interference System (ANFIS). Further, Guler et al. [108] used largest lyaponuv exponent feature in a feed-forward and Recurrent neural network for classification problem.

Automated detection of seizure using neural network was developed by Tzallas et al. [95] by employing time-frequency methods to analyze EEG signals. In another study [115], nine parameter mixed-band feature spaces were used with Levenberg–Marquardt back propagation neural network. The same group [116] used cosine radial basis function neural network classifier based on principal component analysis to detect epilepsy. Srinivasan et al. [96], employed time domain and frequency domain features to Elman recurrent neural network and probabilistic neural networks [117] for classifying EEG signals. Subasi [118] decomposed the EEG signal into time–frequency representations using Discrete wavelet transform (DWT) and applied to different classifiers, such as artificial neural network, dynamic wavelet network (DWN), dynamic fuzzy neural network (DFNN), for epileptic EEG classification. The same group in [94] used neuro-fuzzy system for seizure detection to enhance accuracy.

An automated epileptic system, which applies interictal EEG data to categorize the epileptic patients using Probabilistic Neural Network (PNN, was developed by Forrest Sheng Bao etal. [119]. In [120] authors used relative wavelet energy based features , and after decomposing original EEG signal into several sub-bands calculated ApEn feature to classify the EEGs using three layer MLPNN. Further same group used wavelet transform and line length feature [121] with K-Nearest Neighbor (KNN) classifier and further bused generic programming [122]. S Liang et al [123] used time frequency analysis and approximate entropy to detect epilepsy using linear least square method and linear discriminant analysis. Polat and Gunes [124] used FFT based Welch method and decision tree classifier to classify EEG signals, further used PCA for dimensionality reduction, and employed AR techniques for feature extraction and C4.5 decision tree classifier for classification. [125].

Majumdar [126] reviews various soft computing approaches of EEG signals which emphasize more on pattern recognition techniques. Majumdar concluded that the neural network and Bayesian approaches are two popular choices. Iscan et al. [127] proposed to use SVM classifier for classification of EEG signals, combined with time- and frequency-feature approach. In other recent studies, highest classification accuracy was achieved by Lima et al. [30] using combination of wavelet transform and SVM, Wang et al. [128] used wavelet packet entropy and KNN classifier, and Orhan et al. [32] used DWT and ANN.

Varun Bajaj [101] has classified the EEG signals using SVM classifiers with intrinsic mode functions generated by empirical mode decomposition. Team of Acharaya [129] used ApEn, SampEn and two phase entropies in a Fuzzy classifier, then employed wavelet coefficients and eigen values to extract features from EEG signals [130]. Wavelet packet decomposition (WPD) was utilised to obtain wavelet coefficients and eigenvalues were determined with the help of principal component analysis algorithm. Authors in [131] have utilized Multi-level local patterns (MLP) and employed Empirical mode decomposition (EMD) to decompose non-stationary EEG signals into intrinsic mode functions (IMFs) and used them further for or classification of seizure and seizure-free electroencephalogram (EEG) signals. Same group in [l32] proposed method that employed a bank of Gabor filters for processing the EEG signals utilizing nearest neighbor. Lately authors in [133] propose usage of wavelet based features and certain statistical features without wavelet decomposition for automatic detection of epileptic seizure.

2.4.2. Automated detection of epileptic seizures using non linear analysis techniques

EEG signal are biological systems that can be represented in an effective way using nonlinear techniques. A number of promising quantitative features derived from EEG are used for processing by using nonlinear time-series analysis techniques for seizure prediction. EEG signals have significant nonlinearity as neurons responsible for generation of signals are non linear. [134]. The various nonlinear parameters of EEG signals for the detection of epilepsy are Higher order statistics, Largest Lyapunov Exponent, Correlation Dimension, Fractal Dimension, Hurst Exponent, Approximate Entropy, Sample Entropy, and Recurrence Quantification Analysis. These attributes measures help understand EEG dynamics and quantify the degree of complexity and underlying chaos in the brain signals.

Lehnertz and Elger [135] studied about epileptic region using correlation dimension techniques, by analyzing the spatial and temporal dynamics. Authors in [16] proposed nonlinear dynamical analysis methods for extracting maximum information from EEG signals. Iasemidis were the first group to apply nonlinear dynamical techniques, based upon the principal Lyapunov exponent (PLE), [136], for predicting epileptic seizures. Authors in [137] proposed usage of correlation dimension nonlinear feature extracted from EEG recordings for the detection of onset of seizures. Freeman et al. [138] proposed EEG models for the domain of neurobiology. An algorithm for calculating LLE was proposed by Wolf et al. [139] which was devised by Rosenstein et al. [140]. Researchers in [141] characterize the EEG signals using nonlinear parameters like CD, LLE, HE, and entropy and the same group used an Adaptive Neuro-Fuzzy Interference System for classification of EEG signals into normal and epileptic using different entropies.

Correlation Dimension is a nonlinear parameter used as a useful indicator of pathologies and widely used measure of fractal dimension and was proposed by Grassberger and

Procassia [142]. Lehnertz and Elger et al. [143-144] have expanded work in nonlinear dynamics and seizure prediction to larger data sets, greater numbers of patients, and a variety of epilepsy types, utilizing parameters based upon the correlation dimension. In [145], chaotic features that include largest Lyapunov exponent (LLE) and correlation dimension (CD) obtained from the wavelet subbands of the EEG signals are shown to be effective in differentiating the signals of various classes as normal, interictal, and ictal waveforms . They found statistically significant differences in the values of these measures. Übeyli classified the EEG signals using Lyapunov exponents and currently, support vector machines were used for classification using Lyapunov spectra [146].

Entropy is a measure of the average information contained in the signal segment and is a suitable feature to characterize EEGs. ApEn was proposed by Pincus [147] and showed that the value of ApEn is more for more complex or irregular data. Diambra et al. [148] have shown that the value of the ApEn drops abruptly due to the synchronous discharge of large groups of neurons during an epileptic activity. Richman and Randall [149] developed the parameter called SampEn that measures the complexity and regularity of the time-series data. N. Kannathal et al [141] investigated entropy; sample entropy and approximate entropy for discriminating EEG signals. The approximate entropy parameters extracted from Fourier transforms of the EEG signals is employed for linear and nonlinear classifiers by Liang et al. [150]. Sharma et al [151] have used entropy for classification of focal EEG signals and observed that epilepsy resulted in reduction of sample entropy and approximate entropy values.

Authors in [152] using fractal dimension and artificial neural networks carried out two-way classification. In [153], linear classifier was employed to distinguish normal and seizure activities using linear statistical measures obtained from the EEG signals. Hurst exponent is another measure of non linearity that is a measure of self similarity, and predictability in a time-series [154].

First and second order statistics are insufficient to evaluate the nonlinear dynamic property of the bio-signals. So high order spectral features are widely used in many applications for EEG analysis [155]. It is researched that third order cumulant can be used for EEG signals that highlights the nonlinear behavior. Chua et al. [156] showed that HOS techniques is useful in cases where the signals are corrupted with Gaussian noise and Acharya et al.[157] used these parameters for the automated classification and detection of epilepsy. A nonlinear method which is simple, adaptive can provide variability in the given time series is Empirical Mode Decomposition (EMD). Martis et al. [158] used EMD techniques to classify normal, interictal and ictal EEG time series. In [159] it is shown that EMD is well suited for analyzing nonstationary and nonlinear signals such as an EEG. Authors [160] reformed their method by proposing a method is based on the empirical mode decomposition and the second-order difference plot (SODP). Same group [l61] have proposed the empirical mode decomposition (EMD) and phase space reconstruction based on the phase space representation (PSR) for classification of epileptic seizure and seizure-free EEG signals. The results of these studies indicate that nonlinear techniques are better applicable for successful EEG analysis and classification of biomedical signals.

From this review article that some methods used small sample data points to represent a large number of data points of EEG recordings. In most of the research work, the reported methods did not select the attributes using a suitable technique and did not observe statistical significance of the attributes, although the parameters significantly affect the classification performance. From the literature, some techniques took more time for computation work, some of them were computationally complex and some of them had a limited success rate.

CHAPTER 3

CHARACTERIZATION OF EEG SIGNALS BY VARIOUS ATTRIBUTES

3.1. INTRODUCTION

The brain signals are generated in all mental states, normal as well as abnormal. EEG has capability to reflect all the activity of the brain, so it has been found to be a very powerful tool in the field of neurology and clinical neurophysiology [38]. Despite the fact that EEG is an important clinical tool for diagnosing, monitoring and managing neurological disorders, distinct difficulties associated with EEG analysis and interpretation hindered its wide-spread acceptance. Hence, it is imperative to analyze the EEG signals using a consistent and appropriate processing method in order to obtain correct diagnoses for the treatment of epilepsy.

While EEG records cerebral activity, it also records electrical activities arising from sites other than the brain. Since EEG signals are very weak (ranging from 1 to 100µV), they can easily be contaminated by other sources. The signals obtained from the scalp are highly contaminated with various unwanted signals, known as *artifacts*, produced by events extraneous to the biological event of interest[43]. The non-physiological artifacts are mainly due to power supply (220 volts), which exceeds the main signal by a factor of 2×10^6 or 126 dB. Interference from the mains power supply is unavoidable in EEG recordings, even if captured within specially equipped shielded rooms. Any source in the body which has an electrical dipole or generates an electrical field is capable of producing physiologic artifacts. These include heart, eyes, muscle, and tongue. Sweating can also alter the impedance at the electrodescalp interface and produce an artifact. Artifacts, compromise investigation by masking effects of interest by masquerading as a neurogenic effect The presence of artifacts introduce spikes which can be confused with neurological rhythms, making the EEG signals analysis biased and difficult leading to wrong conclusions[58].

Recognition and elimination of the artifacts in real – time recordings is a complex task, but essential to the development of practical systems. The spurious 60 Hz power supply signals are typically removed by a band-stop filter, which attenuates frequencies in this specific range to very low levels. EEG amplifiers are equipped with notch filters that suppress signals in a narrow band around the mains frequency. If there is main power supply interference still visible in the signal after activating the notch filter; it may be due to high electrode impedance. The analysis artifacts can be controlled with advanced signal processing techniques, for example, round-off errors due to the quantization of signal samples can be made non-effective by setting the large number of discrete amplitude levels in the quantizer [162]. The artifacts compromise sensitivity of main signal as they are confounded with statistical contrasts. The removal of the artifact becomes essential as the preprocessing step for any type of EEG analysis.

3.2. PREPROCESSING OF THE SIGNAL

Preprocessing of the acquired signals is primarily a two step formulation:

i) Artifact Removal: Artifacts are the unwanted signals components accompanying the main brain signals that may bias the analysis of the EEG and may lead to wrong conclusions. The detection of artifacts can be made on line (during the recording session) or offline (after the recording session has terminated). Different approaches for artifact processing are: avoiding the occurrence of artifacts, detectin of artifacts and rejection of artifacts after identification [53].

ii) Normalization : In order to avoid the bias caused by unbalanced feature values, all the extracted parameters are to be normalized in the range of [0, 1] by using *min–max* normalization procedure [163].

$$\text{Feature}_{\text{norm}} = \frac{\text{Feature}_{\text{value}} - \text{Min}_{\text{value}}}{\text{Max}_{\text{value}} - \text{Min}_{\text{value}}}$$

where $\text{Feature}_{\text{norm}}$ is the normalized value of the feature, $\text{Feature}_{\text{value}}$, $\text{Max}_{\text{value}}$, and $\text{Min}_{\text{value}}$ represents actual value, maximum and minimum value of the parameter under consideration.

3.3. ATTRIBUTES REPRESENTING EEG SIGNALS

The formulative step in EEG signal analysis and processing is to figure out relevant information or "patterns" those should be extracted from the signal. Such information permit the abstraction of hidden information in the signal. Features are represented in terms of quantitative numerical value calculated from EEG signals that represent the brain state. Feature extraction involves finding a set of information that include hidden information embedded in the signals and take into account the dimensionality, noise, time information, non-stationarity, set size and so on of the acquired signals. Feature extraction maps the original feature space to a new feature space with lower dimensions. These extracted features is a novel way of expressing the data, and can be continuous, binary, and categorical and may represent attributes or direct measurements of the signal. There are various algorithms in literature that are used for the identification of features that represent data[86-88]. However, most of the time trial and error method is used to determine which algorithms identify the effective features for the specified problem. Many features have proved to be unique enough to use in all brain related medical applications.

3.3.1. Extracted and Selected Features

As revealed in the literature, numerous features are from time domain, frequency domain, and nonlinear dynamics. In present research work, a set of features were selected which were potentially useful for seizure prediction, and had computational requirements reasonable for real-time implementation. The features are extracted and selected based on expertise, observations, and our understanding of EEG signal characteristics framing Signal Feature Vector (SFV). Extracted features comprises first order, second and higher order cumulants of the raw data. The Cumulate Computation is achieved using first four order moments given by,

$$m_1^x = E[x(n)] \tag{3.1}$$

$$m_2^x(i) = E[x(n)x(n+i)] \tag{3.2}$$

$$m_3^x(i, j) = E[x(n)x(n+i)x(n+j)] \tag{3.3}$$

$$m_4^x(i, j, k) = E[x(n)x(n+i)x(n+j)x(n+k)] \tag{3.4}$$

where $m_1^x, m_2^x, m_3^x, m_4^x$ are the first four order moments, E[.] is the expectation operator, i,j and k are the time lag parameters. The *cumulates* can be computed as non-linear combinations of moments as

$$C_1^x = m_1^x \tag{3.5}$$

$$C_2^x = m_1^x(i) \tag{3.6}$$

$$C_3^x = m_1^x(i, j) \tag{3.7}$$

$$C_4^x = m_4^x(i, j, k) - m_2^x(i)m_2^x(j-k) - m_2^x(k-i) - m_2^x(k)m_2^x(i-j) \tag{3.8}$$

Where C_n^x are the first four order cumulants [18].

The first and second order statistics such as mean, mode, median standard deviation have gained significant importance in the area of biomedical signal processing. For non linear signals, first two order statistics are not sufficient to represent the signal, so we have gone for higher order cumulates in this research work. Brief discription of the features are incorporated in this section.

Amplitude Measure of magnitude of a signal is its amplitude; it is measured in terms of voltage (µvolts). The maxima and minima of the signals are given by the amplitude of the sampled signals, represented by eqn

$$MaxAmp(n) = \max[x_n] \tag{3.8a}$$

$$Min\ Amp(n) = \min[\mathrm{x}_n] \tag{3.8b}$$

Mean attempts to describe a set of data by identifying the central position within that set of data, it is a measure of central tendency given by eqn

$$Mean(n) = \frac{1}{N}\sum_{n=1}^{N} x_i(n) \tag{3.9}$$

Median: It also proveides central tedency of the data, but it is preferred over the mean (or mode) when the data is skewed (i.e., the frequency distribution of the data is skewed). It is given by eqn

$$Median(n) = l + \frac{h}{f}\left(\frac{n}{2} - c\right) \qquad (3.10)$$

Mode: It is the only measure of centre appropriate for nominal data expressed as

$$Mode = L + \left(\frac{f_1 - f_0}{2 f_1 - f_0 - f_2}\right) \mathrm{X} h \qquad (3.11)$$

Standard deviation: It is a statistical feature which indicates the distribution of the data with respective to the mean. The mathematical representation is as follows:

$$STD = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}(x[n] - \frac{1}{N}\sum_{n=1}^{N} X[n])^2} \text{ where } 1 \le n \le 4096 \qquad (3.12)$$

Signal-to-Noise Ratio (SNR): It is defined as the ratio of signal power to the noise power corrupting the signal. An alternative definition of SNR is as the reciprocal of the coefficient of variation, i.e., the ratio of mean to standard deviation of a signal or measurement.

x = var (eeg);

SNR = 20*log10 (1/x) (3.13)

Entropy: It measures the signal complexity and and quantify regularity and order in the signal. It is observed that low entropy value of EEG signals represents less number of dominating process and the EEG signals with high entropy represent large number of dominating processes.

$$E_n(n) = -\sum_{k=n}^{n+N} x(k)\log_2 x(k) \quad \text{where } 1 \le n \le 4096, \qquad (3.14)$$

Coefficient of variation: The coefficient of variation (C_v) is a normalised measure of the variance of a series of data. It is calculated by dividing the standard deviation by the mean of the signal and it is defined by

$$c_v = \frac{\sigma}{\mu} \tag{3.15}$$

Skewness is a measure of the asymmetry. If the probability distribution of a real-valued random variable around its mean is not symmetrical, the data is said to be skewed. The equation for skewness (*SK*) is given as

$$SK = E\left[\frac{(x-\mu)^3}{\sigma^3}\right] \tag{3.16}$$

Where E is the expectancy, μ is the mean and σ is the standard deviation.

Positive skewness indicates a distribution extending toward more positive values. Negative skewness indicates a distribution with an asymmetric tail extending toward more negative values.

Kurtosis: Coefficients of EEG signal do not follow the normal distribution, and have a heavy tail characteristic is justified by the value of kurtosis parameters. Kurtosis (*K*) is the fourth-order central moment of a distribution and is defined by the following equation:

$$K(s) = E(s^4) - 3E(s^2)^2 \tag{3.17}$$

Where *s* is the signal and E is the statistical expectation function of *s*. Kurtosis characterizes the relative peakedness of a distribution compared with the normal distribution. Positive kurtosis indicates a relatively peaked distribution whereas negative kurtosis indicates a relatively flat distribution.

Energy: It signifies the strength of the EEG signal. High Energy implies a seizure activity. Let x(n) sequence be an input signal, then the instantaneous energy of the signal is given by $x(n)^2$. The average energy (*EG*) of the signal is given by eqn :

$$EG(n) = \frac{1}{N}\sum_{n=1}^{N} x_i(n)^2 \tag{3.18}$$

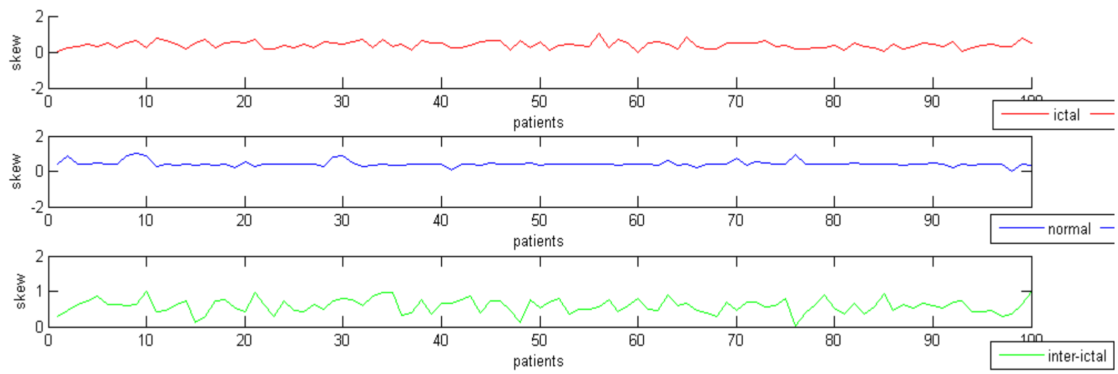
Nonlinear energy(NE) The NE parameter is an instantaneous feature, that provides one value for each value of original data. The nonlinear energy operator is useful for providing an indication as to the spectral content of the signal. For the input signal, in its discrete form, the nonlinear energy (*NE*) operator is represented by

$$NE(x_j) = \sum_{i=2}^{n-1} x_j(i)^2 - x_j(i-1)x_j(i+1) \tag{3.19}$$

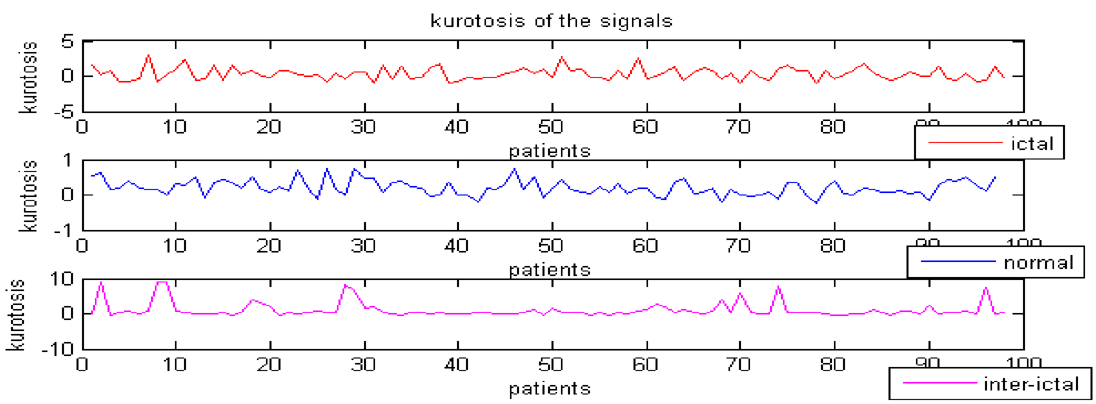
The various features extracted from the signals are informative and apt to analyze the EEG signals. The energy signifies the strength of the signal, entropy quantifies how randomly the seizure signals are distributed as compared to non-seizure signals whereas variance indicates the distribution of the data with respect to mean [164]. Figure 3.1and 3.2 represents some of the features chosen for the study for three different classes of signals. As seen in Figure 3.1(*a*) seizure state is more skewed as compared to the normal state. The kurtosis range levels for all the three classes are shown in Fig 3.1(*b*). The interictal state has the extreme values for kurtosis depicting involvement of large number of dominating process. It has been observed that the signal having high energy lies in the ictal range, whereas the low energy represents the normal signal as depicted in Figure 3.1(*c*).

3.3.2. Features extracted from the dataset

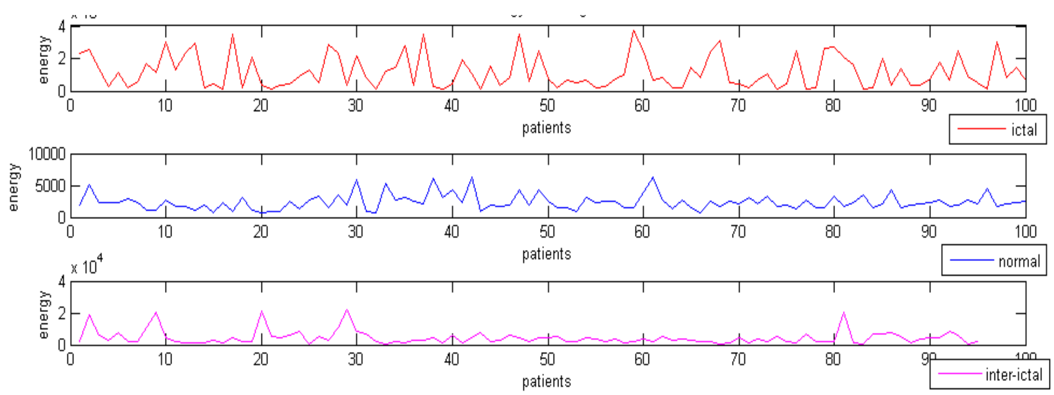
The whole dataset has 300 signals; 100 for each class with 4096 samples in every signal as discussed in chapter 3. Thirteen features are extracted from each signal to frame SFV for comprising precisely F1:Mean, F2:Median, F3:Mode, F4:Coefficient of variation, F5:Minima, F6: Skew, F7: Kurtosis, F8: SNR, F9: Energy, F10:Non linear energy(NE), F11:Maxima, F12:Entropy, F13: Standard deviation.



(a)

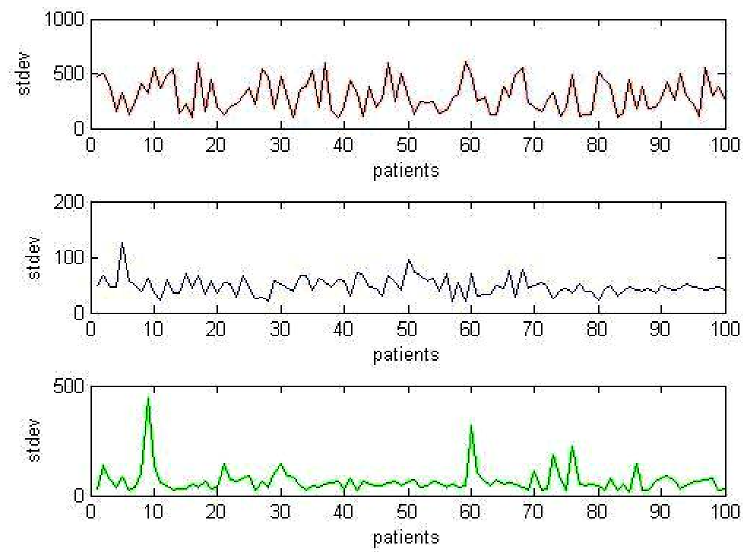


(b)

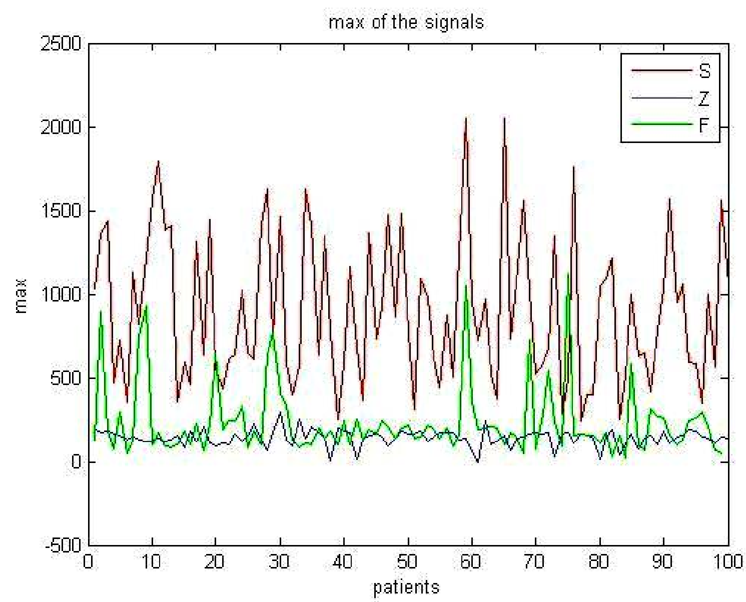


(c)

Figure 3.1. Exemplary Extracted features ***(a)*** Skew of the signals ***(b)*** Kurotosis of the signals ***(c)*** Energy of the signals of subjects in various stages (*i*) ictal state (*ii*) normal state and (*iii*) inter-ictal state.



(a)



(b)

Figure 3.2. Exemplary Extracted features ***(a)*** standard deviation of the signals (***b***) Maximum value of the signals of subjects in various stages.
(S-ictal, Z-Normal, F- interictal)

The variations in the values of the features of EEG signals for normal condintion and seizure condintion are claearly visible in the Figure 3.2. The maximum value of amplitude is large during seizure activity which is self explanatory, represented in Figure 3.2(b). The signals during seizure activity show more dviation as comapred to normal and interictal condintions as shown in Figure 3.2(a). These features are demonstrative of the variations in the signals for three different condintions.

Todate, a common set of features has not been found across individuals that gives adequate performance[122]. An exhaustive search should be conducted to identify an optimal feature set to achieve optimal solution that yields prediction sufficient for clinical application; however, if a suboptimal solution can be obtained from a reasonable feature set, an exhaustive search may not be required. As heterogeneity is a problem for epileptic signals, we believe that it is likely that patient specific features will be found to be more useful for seizure prediction over a large subset of patients.

Table 3.1: Mean value of normalized SFV with variation of variance for normal, interictal and ictal signals

Feature ID	*Mean±var (ictal)*	*Mean±var (normal)*	*Mean±var (inter-ictal)*
F1	0.50007±0.0475	0.46151±0.0348	0.5198±0.0392
F3	0.13562±0.0138	0.48437±0.0255	0.0853±0.0087
F5	0.57462±0.0939	0.54670±0.0332	0.8295±0.0212
F6	0.39269±0.0392	0.58710±0.0408	0.4197±0.0247
F7	0.18853±0.0232	0.30659±0.0300	0.1247±0.0475
F9	0.58902±0.0767	0.31251±0.0483	0.0397±0.0131
F11	0.36365±0.0594	0.47909±0.0271	0.1126±0.0195
F12	0.76060±0.2669	0.94852±0.0784	0.8571±0.0418
F13	0.40082±0.0184	0.27288±0.0242	0.1101±0.0188

Note: *F1 – Mean, F3- Mode, F5- Minima, F6- Skew, F7- Kurtosis, F9 – Energy, F11- Maxima, F12- Entropy F13- Standard deviation.*

Due to inter-variability present in EEG signals, individual features may differ in several orders of magnitude. In order to avoid biased results due to inadequate scaled features, normalization is performed in the range of [0,1]. The comparable mean and variance of all the parameters for the three classes are tabulated in Table 3.1 for comparative outlook.

3.3.3. *Characterization of EEG signals in terms of box plots*

Box-plots are employed in order to graphically present number of statistical parameters of a distribution. Variation in samples of a statistical population are displayed with box plots without making any assumptions of the underlying statistical distribution. Consequently, they are preferred over mean and standard deviation parameters for population distributions that are asymmetric or irregularly shaped and for samples with extreme outliers [165]. Information about the centre and spread is preserved in these plots and moreover, quartiles are insensitive to outliers. Box plots of some of the selected features are depicted in Figure 3.3. The box plots of various features provide a visual aid which allows the variance, degree of dispersion (spread) and skewness in the data to be easily visualised. The box-plot is employed for our work because it facilitates immediate quantification of the Inter Quartile Range (IQR) ranges, median, etc. The box plots of the *mean* feature clearly indicate the overlapping values for three different classes of EEG signals..

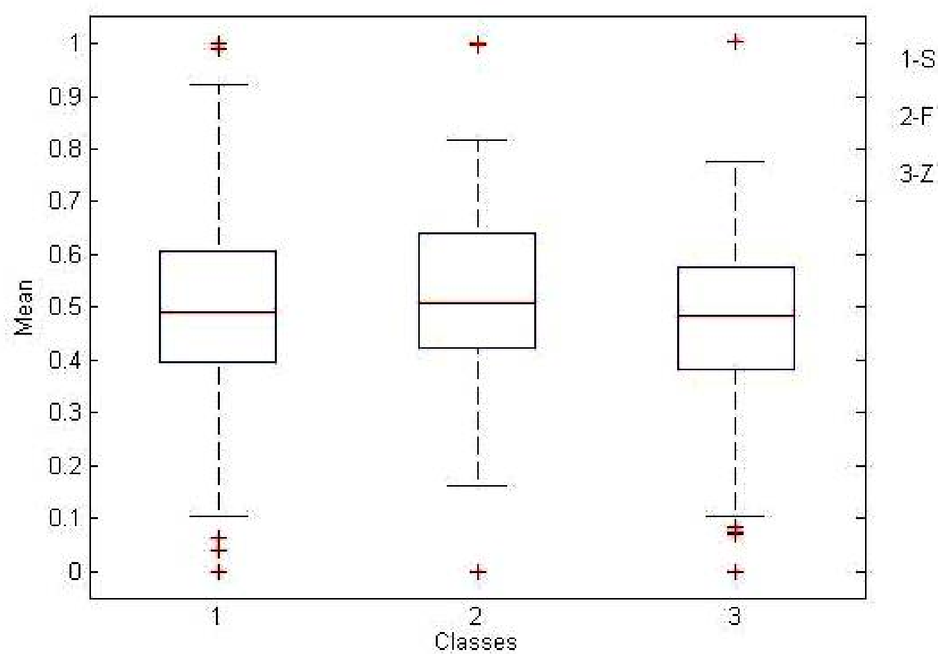


Figure 3.3 (*a*). Box plot of mean function for three different classes, S: ictal, F-inter-ictal, Z-normal

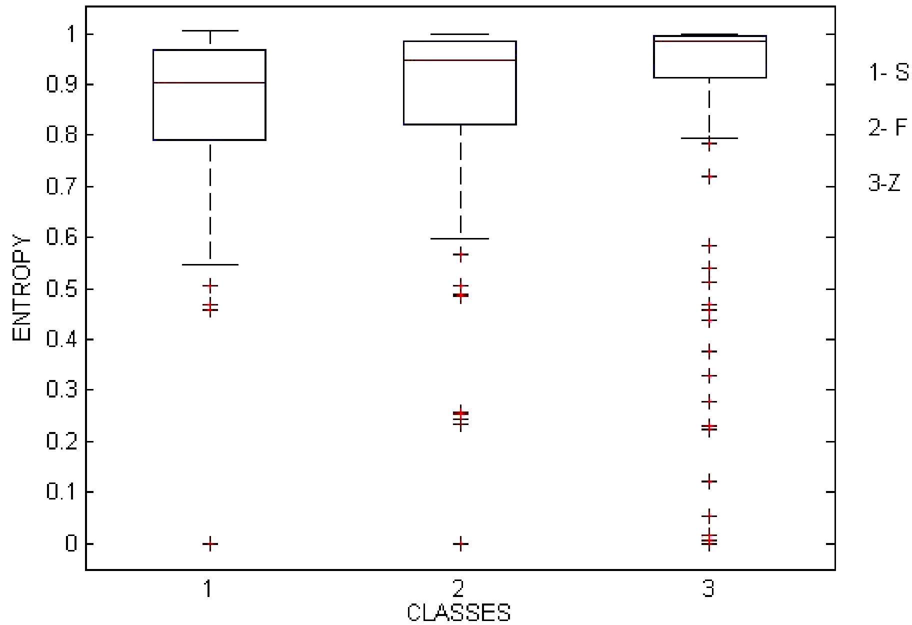


Figure 3.3 (*b*) Box plot of entropy function for three different classes, S: ictal, F-inter-ictal, Z-normal

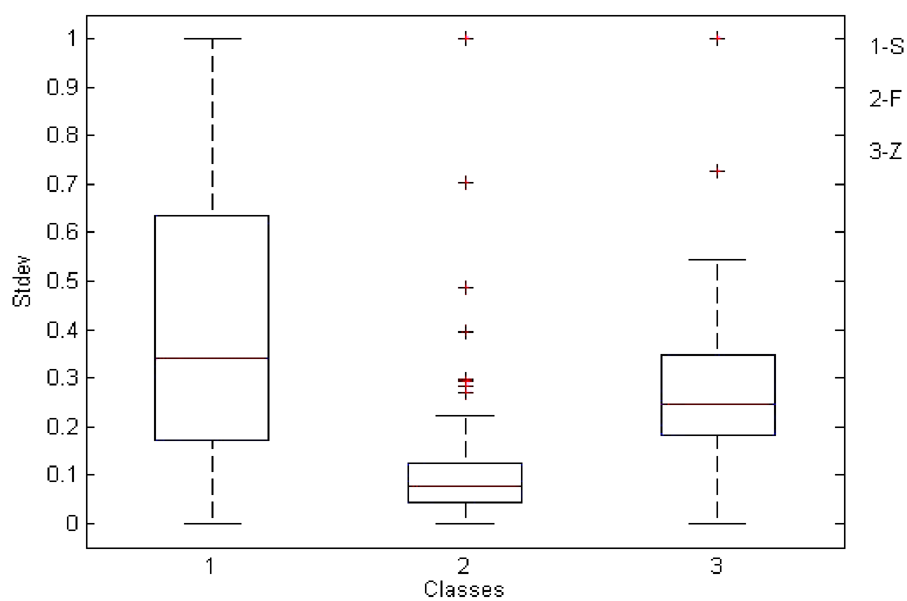


Figure 3.3 (*c*) Box plot of standard deviation function for three different classes, S: ictal, F-inter-ictal, Z-normal

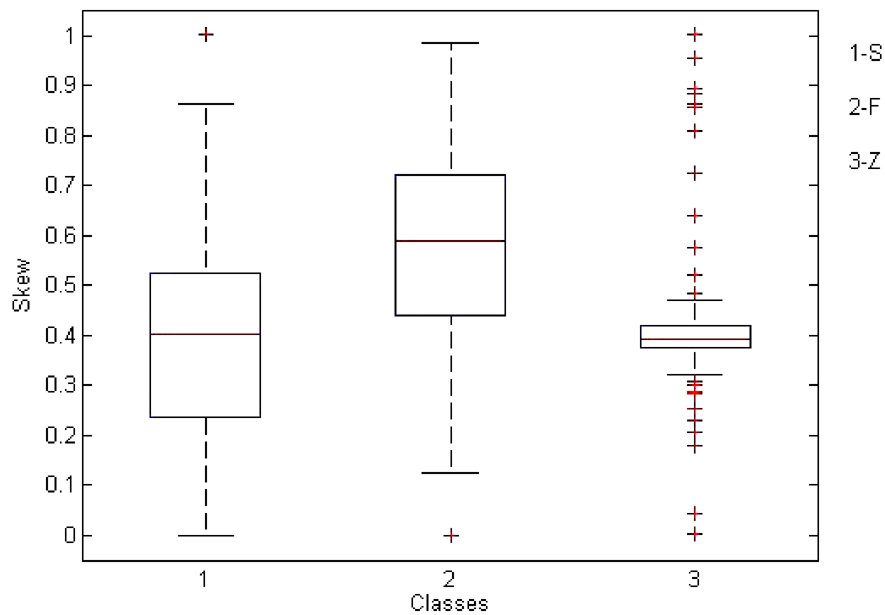


Figure 3.3 (*d*) Box plot of skew function for three different classes, S: ictal, F-inter-ictal, Z-normal

.

It is found that *skew* vlaue associated with the ictal EEG overlaps with the inter-ictal state and normal state of EEG signals. It is also observed that *standard deviation* of ictal state is large as compared to normal and inter-ictal state. Similar observations are recorded for all the feature set of the different classes[166].

3.4. STATISTICAL ANALYSIS OF SFV

For the statistiacl analysis, estimates of variance are computed by ANOVA test. The purpose of analysis of variance is to test differences in means (for groups or variables) for statistical significance. This test is accomplished by analyzing the variance of the data. Total variance due to true random error (i.e., within-group SS) is partitioned into the components and the components that are due to differences between means. These latter variance components are then tested for statistical significance, keeping in mind the hypothesis. If results are significant, the null hypothesis of no differences between means is rejected, and the alternative hypothesis (the means in the population) are different from each other) is accepted. The effectiveness or "fitness" of each individual feature when classes are Gaussian and uncorrelated is measured using Fisher's discriminant ratio (FDR) that determines feature effectiveness[167]. Table 3.2 illustrates the output of the ANOVA analysis depicting F value with parameter p (significance level) to provide statistically significant difference between attributes.

Table 3.2: Summary of ANOVA analysis for linear features in terms of F value for every extracted attribute

Feature	F value	p value
F1	41.986	0.000
F2	22.96	0.011
F3	58.09	0.000
F4	10.55	0.035
F5	49.29	0.000
F6	31.47	0.000
F7	35.28	0.000
F8	45.43	0.063
F9	49.96	0.000
F10	12.81	0.027
F11	99.33	0.000
F12	41.575	0.000
F13	51.145	0.000

Note: *F value = Fisher's Discrimination ratio, p= significant value<0.05 F1 -Mean, F2- Median, F3- Mode, F4- Cov, F5- Minima, F6- Skew, F7- Kurtosis, F8- SNR, F9 – Energy, F10-Nle, F11- Maxima, F12- Entropy F13- Standard deviation.*

On inspecting Table 3.2, some features in particular, standard deviation, mean, minima of the signal have relatively higher value of F. Consequently, they are more significant compared to remaining features. The features having significance value > 0.05 (median, SNR, non linear energy and cov) have less significance for our problem at hand. To verify whether the extracted features are distinct and uncorrelated to each other or not, prediction importance of each feature, in terms of rank and importance parameter is extracted. The inter correlation between these extracted features used in the prediction model was calculated based on variance inflation factor (VIF) indicating multi-collinear analysis.

The VIF value for each feature was calculated using

$$\mathrm{VIF} = \frac{1}{\left(1-\mathrm{R}_j^{\;2}\right)}, \quad \mathrm{Tolerance} = \left(1\text{-}\,\mathrm{R}_j^{\;2}\right) \tag{3.21}$$

where R_j^2 is the multiple correlation coefficient of one feature's effect regressed onto the remaining features. Tolerance value obtained less than 1 for these features indicate that the variable under consideration is almost a perfect linear combination of the independent variables[168]. Figure 3.4 depicts the prediction importance of each attribute arranged in the increasing order of importance for this particular problem at hand.

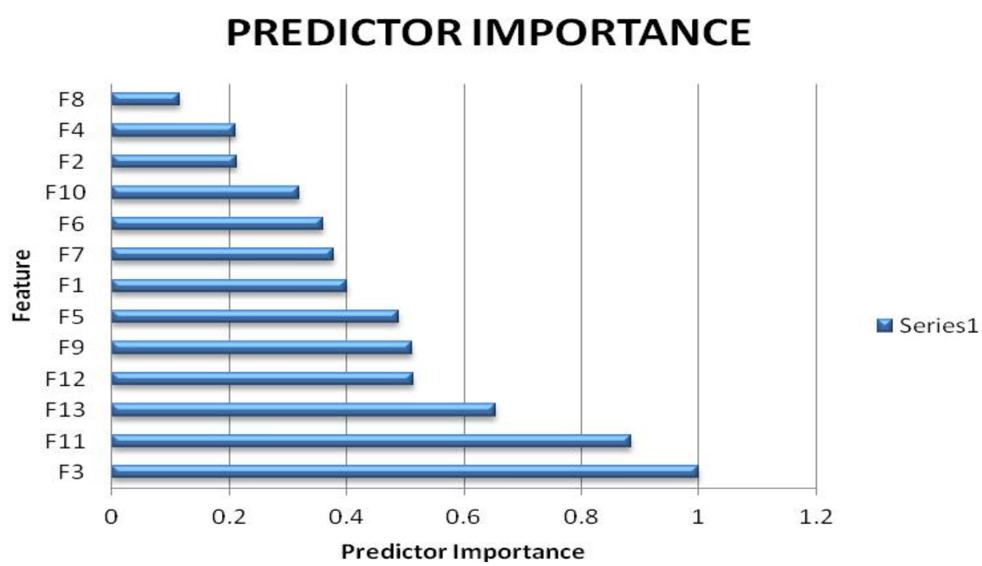


Figure 3.4. Prediction importance of extracted features by calculating VIF for the feature set.

The features are statistical analyzed further by Kruskal Wallis Test, a non-parametric method for testing whether samples are independent, or not related. Table 3 provides chi-square value and significance ($p < 0.05$) of each feature.

3.5. k MEANS CLUSTERING

The goal of the k-means algorithm is to find the optimum "partition" for dividing a number of objects into k clusters. This procedure will move objects around from cluster to cluster with the goal of minimizing the within-cluster variance and maximizing the between-cluster variance. The descriptive statistics are formulated separately for each cluster using every feature and a few of them are displayed in Figure 3.5. In this figure probability density function is plotted for particular feautre for two classes of EEG. The variation between two classes with that particular feature is clearly visible in all the figures.

Table 3.3: Value of Chi- square from Kruskal Wallis Test with significance value < 0.05 for all features

Feature ID	Chi-square	Significance
F1	36.32	0.001
F2	79.306	0.000
F3	252.220	0.000
F4	132.922	0.000
F5	97.964	0.000
F6	44.382	0.001
F7	134.661	0.000
F8	94.818	0.000
F9	54.907	0.000
F10	83.632	0.000
F11	16.340	0.012
F12	25.149	0.001
F13	25.202	0.000

Note: *F1 -Mean, F2- Median, F3- Mode, F4- Cov, F5- Minima, F6- Skew, F7- Kurtosis, F8- SNR, F9 – Energy, F10-Nle, F11- Maxima, F12- Entropy F13- Standard deviation*.

From all the above figures it can be inferred that selected features are capable of discriminating samples that belong to different classes to a certain extent. as there is large overlap between the PDFs the metric which does not facilitate good class separation.

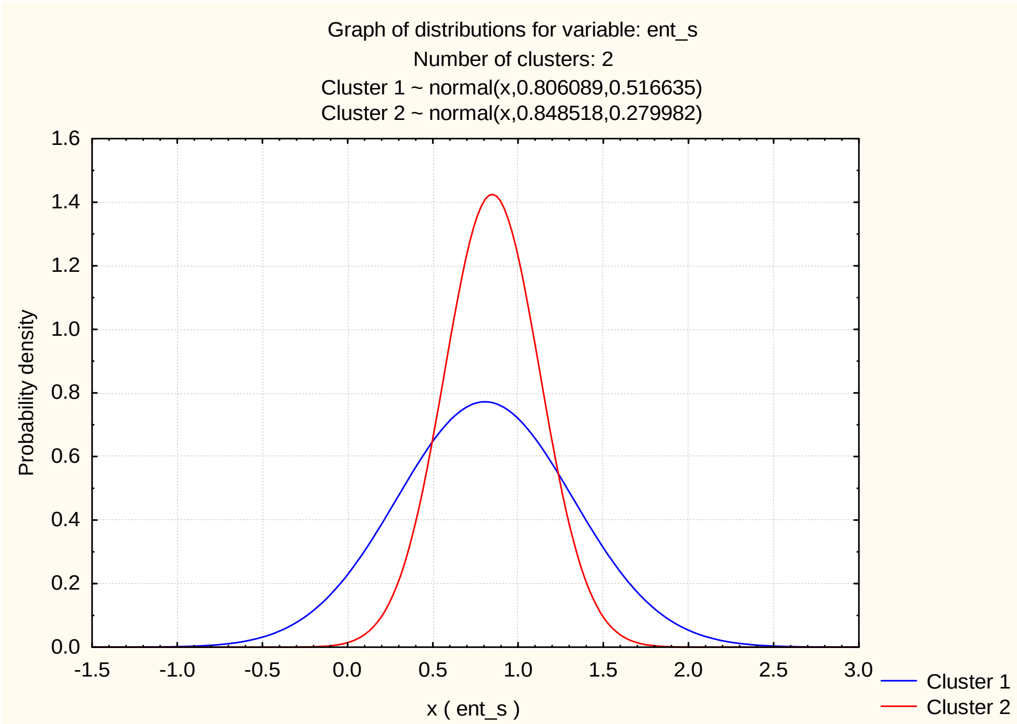


Figure 3.5 (*a*) The distribution of ictal and normal classes around feature F12(Entropy)

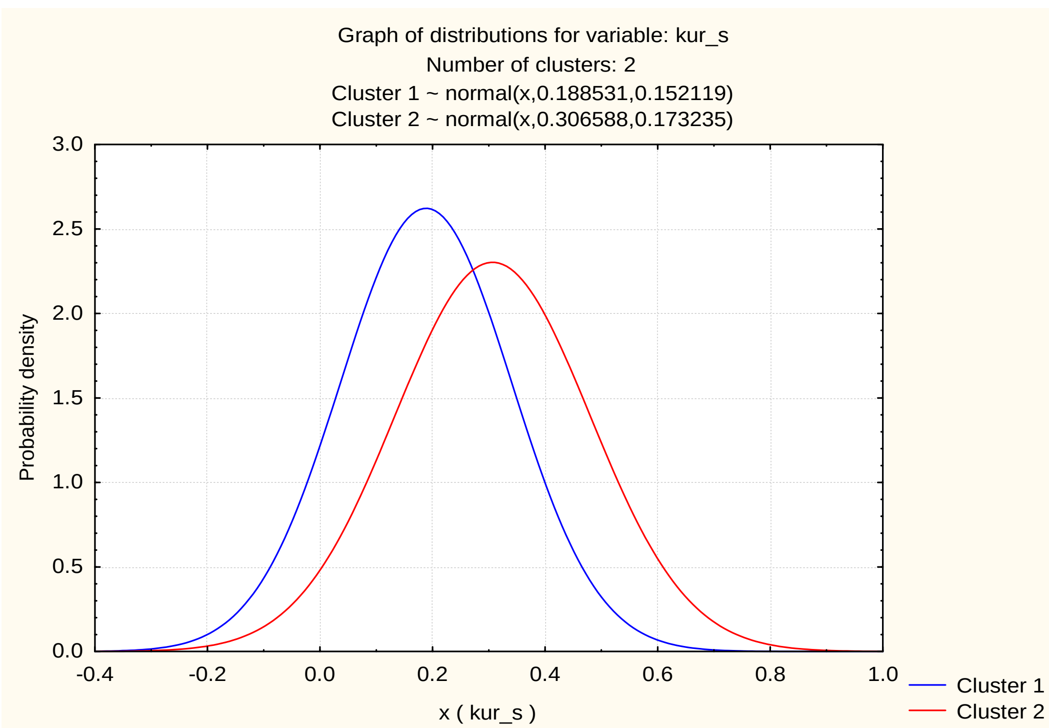


Figure 3.5 (*b*) The distribution of ictal nad normal classes around feature F7(Kurtosis)

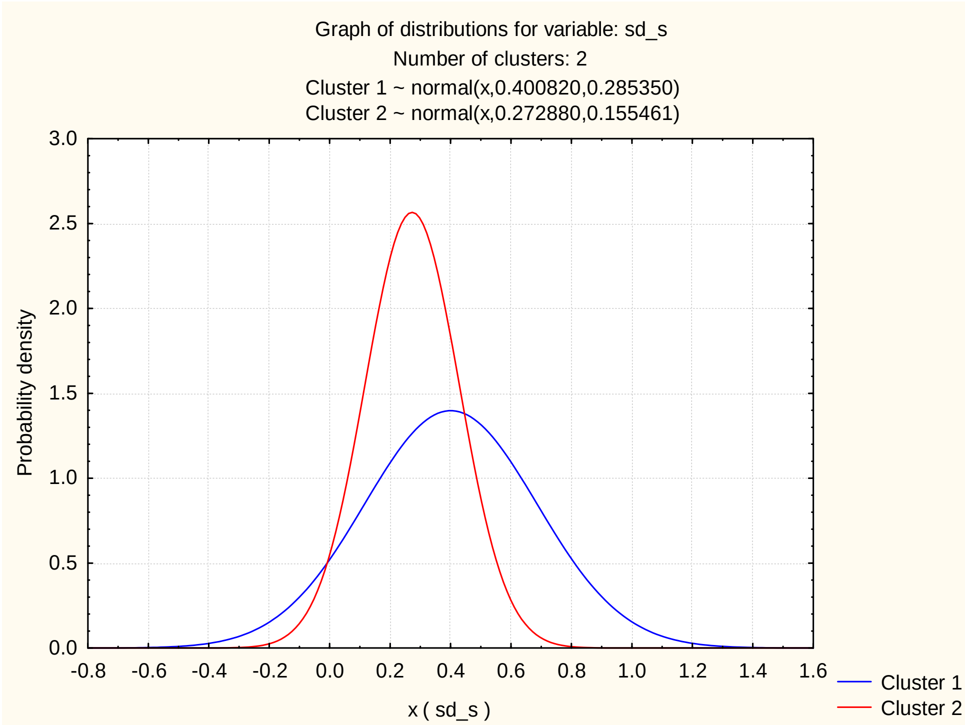


Figure 3.5 (*c*) The distribution of classes around feature F13(Standard deviation)

If a large differences exist between the means of the two curve, there is an indication that the feature might be a suitable candidate for separation between normal and abnormal subjects. The features whose predictor importance parameter and F value was calculated in the previous sections represent the capability of distinguishable clustering, confirming the relevance of features for assessing the capability of distinguishing different classes.

3.6. CONCLUSION

The objective of this study was to identify appropriate specific features for predicting seizures. An exhaustive preliminary evaluation was conducted, based on quantitative EEG analysis, to determine if a generic search process is capable of from a large set of candidate features. The proposed methodology focuses on what features delineates the

EEG signals of a normal patient from an epileptic patient. The list of features chosen for this work may not be considered to be exhaustive, nonetheless it was chosen to be sufficient for proof of computationally efficient classification. These chosen feature set has potential for on-line implementation of the predictor system in low power, implantable environments. In this objective, an efficient statistical analysis is presented to extract useful feature set for classification of EEG signals. The novelty of the proposed methodology lies in the exhaustive statistical analysis of extracted features to come up with prominent feature set for classification purposes. This type of study is instrumental in finding feature extraction algorithm and type of features that can best represent the data.

CHAPTER 4

DESIGN AND DEVELOPMENT OF PREDICTION MODEL TO DETECT SEIZURE ACTIVITY

4.1. INTRODUCTION

Increasing power of computing and enhanced processing capabilities has made analysis of EEG signals efficient and effective. It has become a fundamental tool for diagnosing neural problems, and useful for both physiological research and medical applications. The diagnosis and prediction of epileptic seizures is a daunting challenge, even for the experienced neurologist, due to overlap of the appearances of the normal and abnormal signals. Therefore an automated Computer Aided Diagnostic (CAD) system for classification of the epileptic stages from EEG signals is highly desirable. In light of this fact, present research aims at developing an automated predictive model to diagnose the state of an epileptic patient using EEG signals. A CAD system design is proposed in the current chapter that can detect abnormal disorder and malfunctioning of the brain during the seizure and before the onset of seizure. The potential of two different algorithms of neural network techniques have also been investigated for design of CAD system. This work also includes evaluation of the performance of different classifiers for two-class seizure classification. A novel approach of hierarchical CAD system is proposed for seizure classification.

Two primary considerations for developing any prediction model for detection and classification are the type and nature of features to be extracted from the EEG input signal and the choice of analysis techniques to be applied on these extracted features [169]. The previous research work has dealt with selection of features, different domains of feature selection, artifact handling and concluded with optimum set of features constituting SFV, which can be used further for classification. Extracted features constituting SFV are utilized as the input data for learning algorithm and same set is used in the prediction phase. The chosen set of features is simple but robust for the morphology of EEG data needed for the classification problem.

4.2. PROPOSED CAD SYSTEM DESIGN

The block diagram of the proposed CAD system design for two-class and three-class seizure classification using statistical features is shown in Figure 4.1. Statistical features depicting morphology of EEG signals are utilized to develop an automated soft computing diagnostic system. The quantitative and statistical analysis of the selected features is elaborated in Chapter 3 and the approach is implemented on the same data set. All 300 signals are represented in the form of a vector comprising of thirteen odd features, framing SFV^j_i, where i = 1 to 100 and j = 1 to 3.

4.2.1. Experimental Workflow

Rigorous experimentation has been carried out for designing and evaluating the performance of the proposed CAD system design. The flow of the proposed design is implemented through number of experiments (detailed in Table 4.1).

Table 4.1: Description of experiments carried out for design of CAD system for seizure classification.

Experiment 1	Exhaustive experiments are carried out to develop the architecture of prediction model by varying the number of neurons in the hidden layer and signal feature vector length for deciding the best network topology and ascertain the CAD system with highest overall classification accuracy.
Experiment 2	For evaluating the performance of the proposed CAD system design, rigorous experimentation has been carried out for comparison with other available classifier for the same architecture.
Experiment 3	This experiment deals with two class classification problems with proposed topology and with available binary classifiers to recognize an optimal soft computing paradigm for seizure classification.
Experiment 4	In this work, the focus is on modeling the seizure classification by a hierarchical framework, a variation of the classifier system; HCAD system is proposed.

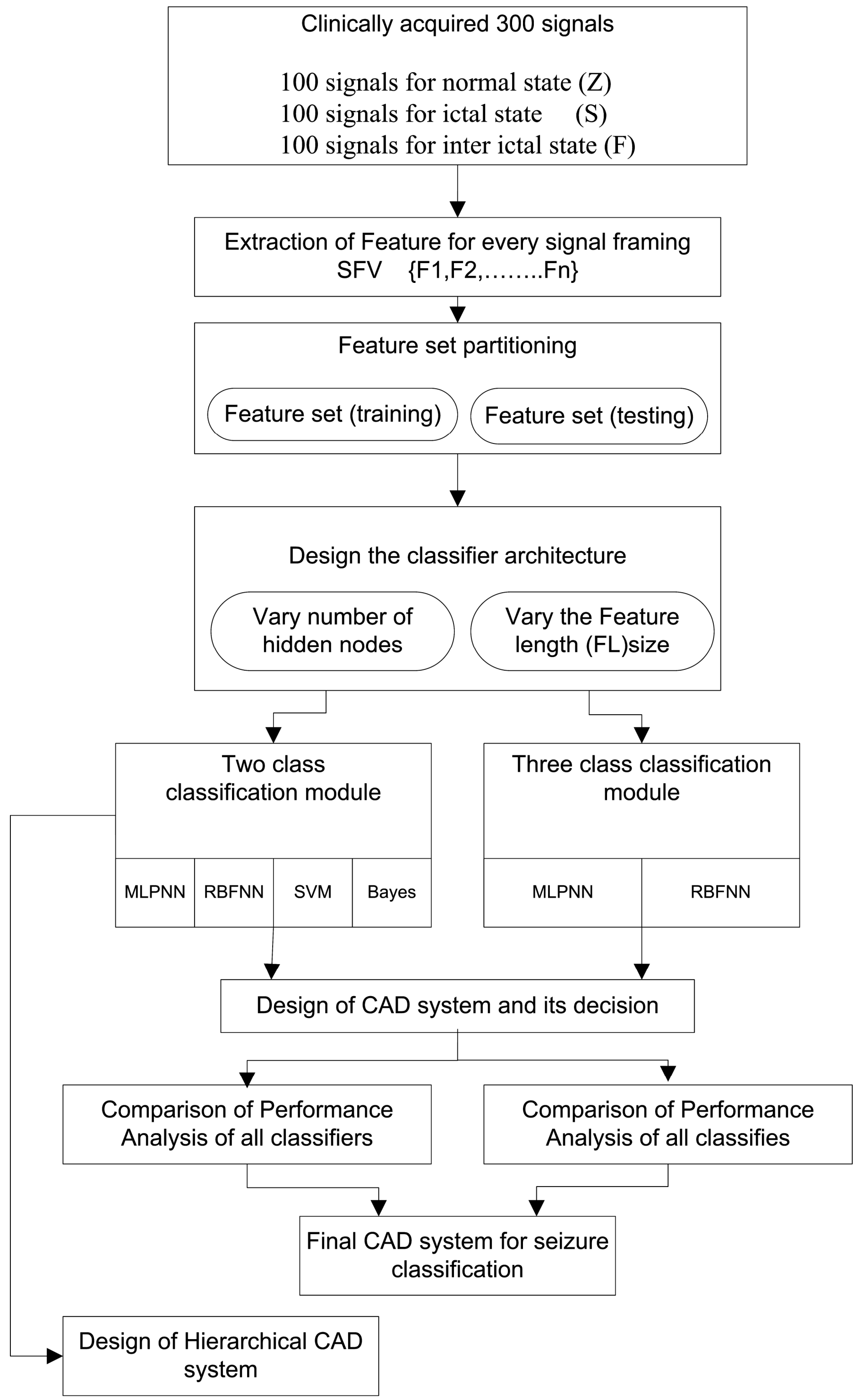


Figure 4.1. Proposed CAD system design using statistical features for two-class and three-class seizure classification

4.2.2. Experiment 1. *To develop the architecture of prediction model to ascertain the CAD system with highest overall classification accuracy.*

4.2.2.1 Design of System Architecture

To analyze the EEG signals with enhanced accuracy and precision, various computational techniques such as neural networks, support vector machines, Bayes classifiers could be useful. However, neural networks have been successfully employed to process EEG signals because of its quality of generalization and great predictive power [26]. There is always a possibility of the network's free parameters adapting to special features of the training data, as there are large number of training samples and relatively larger number of synaptic weights.

In this approach, feed-forward multi-layered perceptron Neural Network (MLPNN) algorithm is employed to obtain a predictive model as this classifier has good generalization performance even without feature space dimensionality reduction and is less prone to over fitting [170]. To solve this problem, multi-layer network of Perceptron is used as it offers explanation in relation to weights and activation functions and obeys convergent theory to gives one solution to execute a problem. The performance of ANN lies in its empirically chosen structure of the network i.e., number of layers, number of neurons in hidden layer, their interconnections and the neurons in the output layer. Multi-layer Perception network can be used with any number of layers, but Kolmogrov Theorem indicates that a three layered perception network is able to separate any kind of space and it can be used for constructing neural networks

The number of neurons in the input layer symbolizes the number of features presented to the network, followed by hidden layer with neurons which transforms the input into nonlinear combinations and passes the signals to the output layer. For the proposed design, the selected features SFV represent thirteen neurons in the input layer. As this is a three classification problem, three neurons are taken in the output layer to classify ictal (S), interictal (F) and normal (Z) categories. Number of neurons in the hidden layer of the neural network has significant effect on the performance of network [24]. More neurons in the hidden layer require more computation; less number of neurons gives high training error and high generalization error due to under fitting. Many techniques have been used to optimize the hidden nodes. These techniques can be divided in to two categories. In the first category, network is generated with the small number of hidden nodes and number of nodes is increased until the maximum accuracy is reached. This method is called the constructive method. While second group of method includes generation of network with large number of hidden nodes initially and it is gradually decreases until the maximum accuracy is reached this is called the destructive method. Present work employs constructive method. Starting with five neurons in the hidden layer, number of neurons was incremented and decremented till maximum classification accuracy of the network was achieved. Each of the architecture with varying hidden neurons is trained, tested and validated; and the performance accuracies of all the tested models are depicted in Figure 4.2.

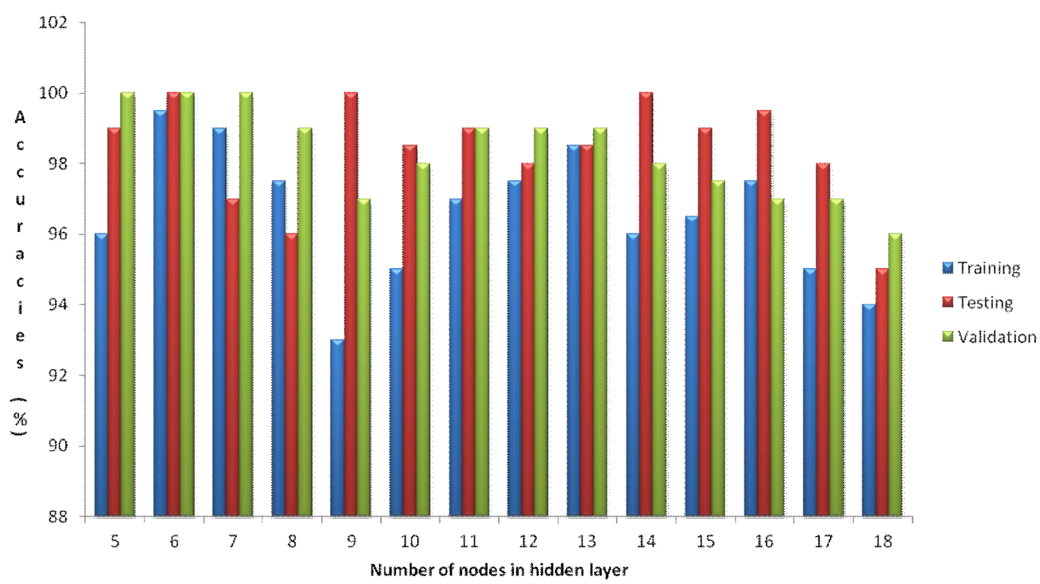


Figure 4.2. Performance accuracies of proposed model in terms of training, testing and validating accuracies; with varying number of neurons in hidden layer.

It was observed that the best classification accuracy was obtained with six neurons in hidden layer. It gives 100% testing and validation accuracy and 99.5% training accuracy. This network performs significantly better and requires a smaller number of iterations to train a neural network. As the number of nodes are increased the training efficiency decreases. In some case even if validating and testing efficiencies are high, as in case with 9 or 13 hidden nodes, the training efficiency is low. Keeping a balance all the three efficiencies, number of hidden nodes is selected as 6. For further experimentation, six hidden nodes will be considered for the designs.

Further, exhaustive experimentation was done for the selection of number of input features or Feature Length (FL). For feature selection techniques two selection methods i.e., Sequential Forward Search (SFS) and Sequential Backward Search (SBS) are employed. In two feature subset selection methods, best combination of predefined selected features is obtained by using some class seperability criterions like FDR, divergence etc discussed in Chapter 3. In SFS approach, a single most discriminatory feature is selected from all the available features based on adopted class seperability criterion that is permanently selected. Further, its combination with all other remaining features is tested, and amongst them the best pair is chosen again in terms of adopted class seperability criterion. This process of selecting features continues until the desired number of features is reached.

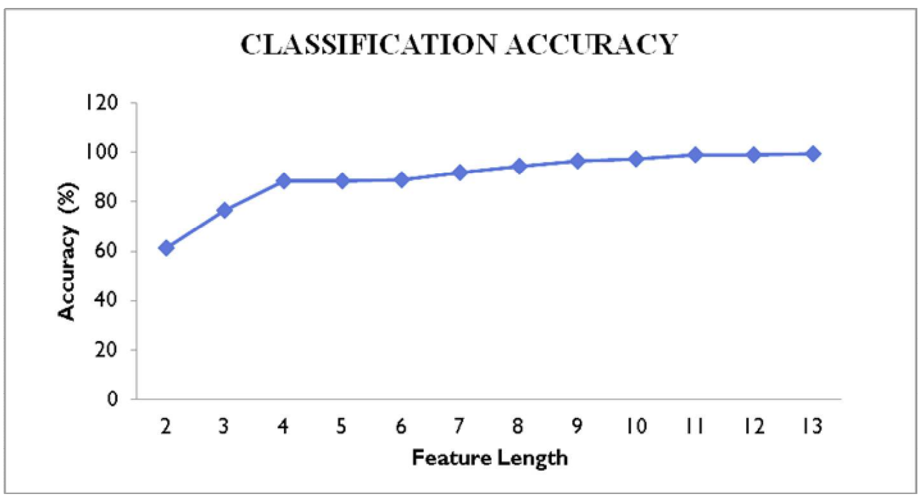


Figure 4.3. Performance analysis with reference to classification efficiency of NN with varying number of feature length.

In this procedure, subsets of features were used to train the network and classify the signals beginning from feature length of two prime features. This process was continued till all the available extracted features were used. After 10-fold cross validation, the prediction model was evaluated for classification with the proposed architecture [171].

For all the designed classifier models, the performance was evaluated in terms of training, testing and validating efficiencies. The classification accuracy with different FL is tabulated in Table 4.2 and summarized in the form of graph in Figure 4.3.

Table 4.2: Classification summary for the CAD system with varying number of features.

Features	FV	Tr. Effi.	Tst. Effi.	Vald. Effi.	CA
Mean-std	200	61.43	71.11	68.88	61.42
Mean-std-eng	300	84.28	82.22	86.66	76.65
Me-std-eng-ent	400	88.15	88.64	88.64	88.29
Me-std-eng-ent-Sk	500	88.32	86.66	88.95	88.94
Me-std-eng-ent-Sk-ku	600	88.63	86.36	90.91	88.68
Me-std-eng-ent-Sk-ku-sn	700	93.36	88.64	86.36	91.63
Me-std-eng-ent-Sk-ku-sn-cov	800	93.84	95.45	95.45	94.31
Me-std-eng-ent-Sk-ku-sn-cov-med	900	95.32	99.55	100	96.72
Me-std-eng-ent-Sk-ku-sn-cov-med-mod	1000	96.21	100	100	97.32
Me-std-eng-ent-Sk-ku-sn-cov med-mod-max	1100	98.58	100	100	98.99
Me-std-eng-ent-Sk-ku-sn-cov med-mod –max-min	1200	98.88	100	100	99.01
All	1300	99.80	100	100	99.3

Note: *Tr. Effi: Training Efficiency, Tst. Effi: Testing Efficiency, Vald. Effi: Validation Efficiency, FV: Feature vector, CA: Classification Accuracy, All efficiencies are in %.*

It was observed that high classification ability for epileptic seizure detection was obtained from NN classifier by feature vector of length 13 in comparison with feature size of lengths 2, 3, 4 and so on as clearly depicted in Table 4.2. Although, testing and validating efficiencies dramatically increases to 100% after the feature length of nine but the classification accuracy does not show much improvement. The variation in the classification accuracy with thirteen features and with 10, 11 and 12 features is very wide. The prime facia of the problem in hand is high classification accuracy and good training efficiency, so the model depicting higher value of these two parameters is to be preferred as compared to the others. Thus, 13 features computed from EEG signal are considered for further analysis with predefined nodes in the hidden layer and required nodes in the output layer of NN. Thus, the final architecture of neural network comprises thirteen input nodes, six hidden nodes and three output nodes as depicted in Figure 4.4 with fully connected architecture.

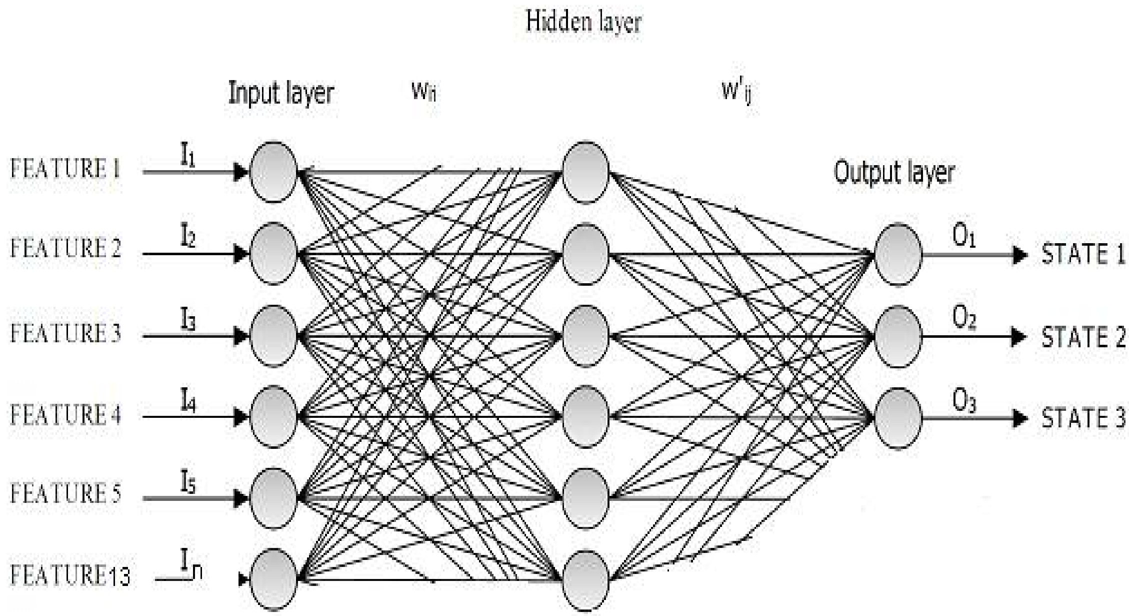


Figure 4.4. Architecture of the proposed neural network for design of CAD system for seizure classification with thirteen nodes in the input layer, six neurons in the hidden layer and three neurons in the output layer.

The overall classification system consist of three layers of artificial NN with tan-hyperbolic and softmax function as the activation function for hidden and output layers respectively with Cross Entropy as error function and BFGS (Broyden-Fletcher-Goldfarb-Shanno) as the technique used for training neural network[170]. To reduce the bias of training and testing data set, bootstrapping technique and 10-fold cross-validation technique are preferred. These techniques provide information about how well the classification model will operate on new data stream. For this work, 70 % data set is used for training, 15% data set for testing and 15% for validation using bootstrapping method with 1000 seed points is used for effective training of the network (primarily to avoid over fitting), to evaluate the average predictive ability of the method and for enhancing prediction accuracy.

4.2.2.2 Results and discussion

The most important aspect of a prediction method is its ability to make correct predictions. Performance metrics for classification and validity of any classifier are: sensitivity, specificity, classification accuracy and ROC curves (discussed in Chapter 1). The confusion matrix and classification summary are useful tools for evaluating the effectiveness of a classification network.

After all the exhaustive experiments undertaken, the prediction model was evaluated for classification with the proposed architecture. The confusion matrix and classification summary for the model are depicted in Table 4.3 and Table 4.4. For medical applications, Individual Classification Accuracy (ICA) should also be high as Overall Classification (OCA) and Individual Misclassification Accuracy (IMA) should be as low as possible. For clinical applications, diagnosis system should not only give high sensitivity and high specificity but also should give almost zero false positive and false negative events [138].

Table 4.3: Classification summary for three class classification using proposed architecture

	Interictal	**Ictal**	**Normal**	**OCA**
Total	99	100	100	299
Correct	98	99	100	297
Incorrect	1	1	0	2
ICA(%)	98.98	99	100	99.33(%)
IMA (%)	1.01	1	0	0.67(%)

Note: *IMA- Individual Misclassification Accuracy, ICA-Individual Classification Accuracy*.

Strategically, we have designed a fully automated neural network model, capable of classifying the seizure activity into ictal, interictal and normal state with classification accuracy as high as 99.3% with misclassification error of 0.67%. The ICA for ictal condition is 99% and IMA is 1%, ICA for normal condition is 100% and IMA is 0% and lastly ICA for interictal is 98.9% with IMA as 1.01%.

Table 4.4: Confusion Matrix for the selected prediction model for three class classification using designed network architecture.

	Inter-ictal	**Ictal**	**Normal**	**OCA**	**Sen(Ictal)**
Inter-ictal	98	0	0		
Ictal	0	99	0	99.33%	99%
Normal	1	1	100		

Note: *OCA – Overall Classification Accuracy, Sen- Sensitivity*

The correct classification accuracy is 99.3% and misclassification accuracy is 0.67%. For the different set of parameters and optimum number of neurons in hidden layer, ANN model revealed a superior model for validating the classification. This network performs significantly better and requires a smaller number of iterations to train a neural network. The promising performances observed are demonstrative of the efficiency and efficacy of systems developed for classification and prediction of normal, ictal and interictal conditions of epileptic patients.

*4.2.3. **Experiment 2** Comparison of Machine Learning Methods for Prediction of Epilepsy*

4.2.3.1 Methodology

In this approach, potential of two different algorithms (back propagation and radial basis function) of neural network technique have been investigated for classification of EEG signals. Classification is based on quantitative parameters obtained from neurophysiologic signals used to train the networks and the performance of the networks is analyzed to confirm the efficacy of the network. To classify the subjects for state of epilepsy using EEG signals and for design of an effective model workflow as shown in Figure 4.5 was maintained.

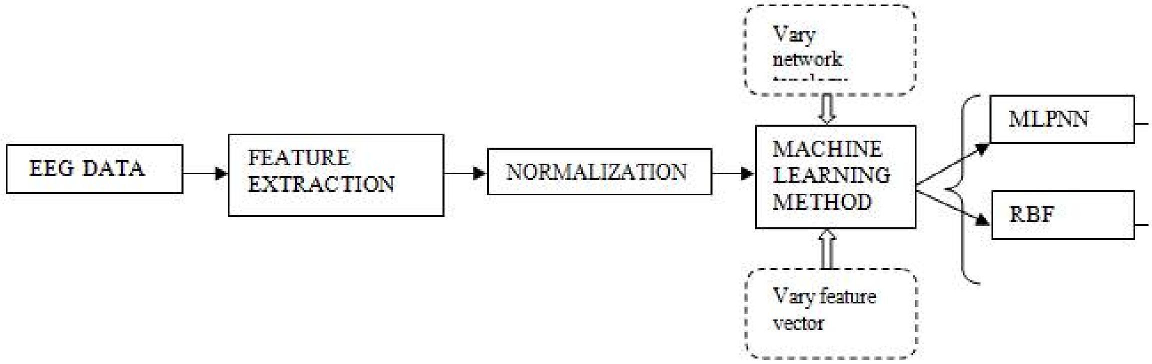


Figure 4.5. Work Flow for comparison of two machine learning methods for three class classification problem

The comparative analysis is based on variation in network topology and in feature vector used for training the networks. The method of selecting feature length is adapted systematically by enumerating all combinations of feature vectors. For defining the topology, the number of neurons in the input layer was taken as thirteen corresponding to FL and number of neurons in the output layer is three to classify three different classes. The number of hidden nodes was varied from 5 to 25 to find out the architecture giving the better performance with high accuracy. To validate predictive model with a good generalization performance, dataset is divided randomly into 70% for training the network, 15% for validation and 15% for testing to assess the predictive performance of the model. For each sequence in the training and testing sets, around 20 networks were trained and best five networks were averaged to get the performance parameters.

4.2.3.2 Results and Discussions

Performance analysis was evaluated for both the networks for different architecture by varying the number of nodes in hidden layer. Varying nodes from five to twenty five in hidden layer, different architectures were trained and their performance and results are as shown in Figure 4.6. Wide variation during training, testing and validating accuracy was observed with two different types of networks with different topology. It is interesting to note that as number of nodes increases, efficiency increases in case of RBF from 50% to around 95% whereas the efficiency does not vary much in case of

MLPNN with increasing number of nodes. Next, the discrimination ability of different features sets are investigated. By varying the FL, efficiency and sensitivity of the methods for classification is determined. Figure 4.7 depicts the comparison of the two methods in terms of sensitivity by considering different length of FL with different number of hidden nodes. It is observed that combined FL consisting of all the features has more efficiency and discrimination ability. Maxima and minima sensitivities obtained are 99.3% and 61.4% for MLPNN and 96.9% and 59.9% respectively for RBF as depicted in Figure 4.6. All the results obtained, used the discrimination ability of all the selected features from all the 300 signals.

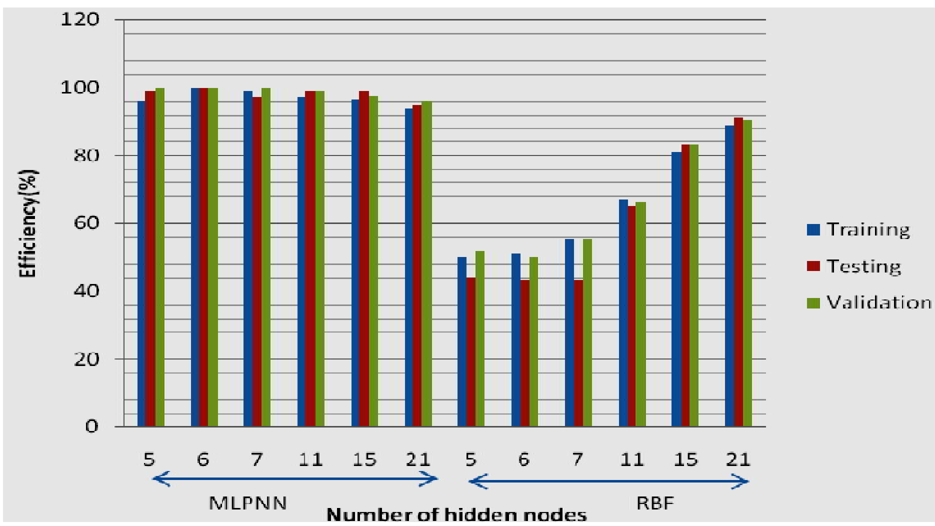


Figure 4.6.Comparative performance analysis for two machine learning methods for three class classification of seizure activity

As observed from the previous results, MLPNN has an edge over RBF for this problem of classification; hence another comparative performance analysis was done. The models were formulated and their efficiencies were calculated by varying number of features but with the same number of hidden nodes. The various comparisons for varying FL are reported in Figure 4.7 in terms of sensitivity of prediction.

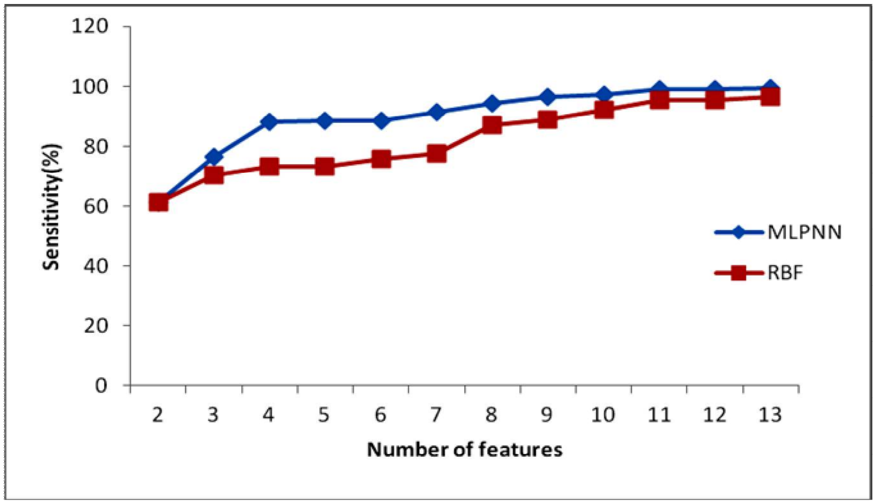


Figure 4.7.Performance Analysis of MLPNN and RBF in terms of sensitivity with same network topology and varying feature index.

Figure 4.8 depicts selected results of subsets of feature vector sets used for classification Figure 4.8 (*a*) depicts the comparison with two features and same number of hidden nodes for both the models; similarly Figures 4.8 (*b-d*) depicts the comparative figures for different FL. It can be visualized from the graph that training accuracy is always better with MLPNN as compared to RBF, but once training is done; there is small variation in testing and validation efficiency accuracy.

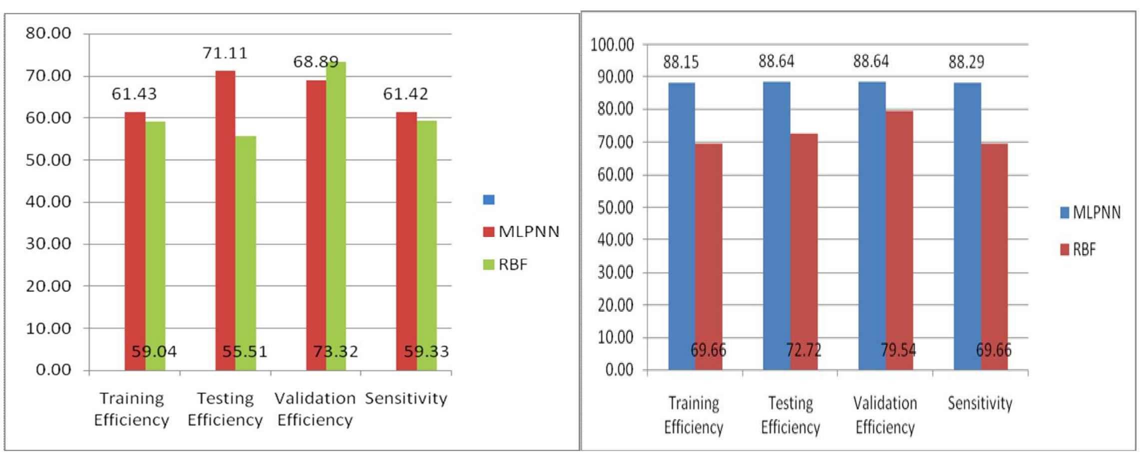


Figure 4.8. Classification Efficiency Analysis of MLPNN and RBF (*a*) Two features with same number of hidden nodes (FL:2) (*b*) Four features (FL:4)

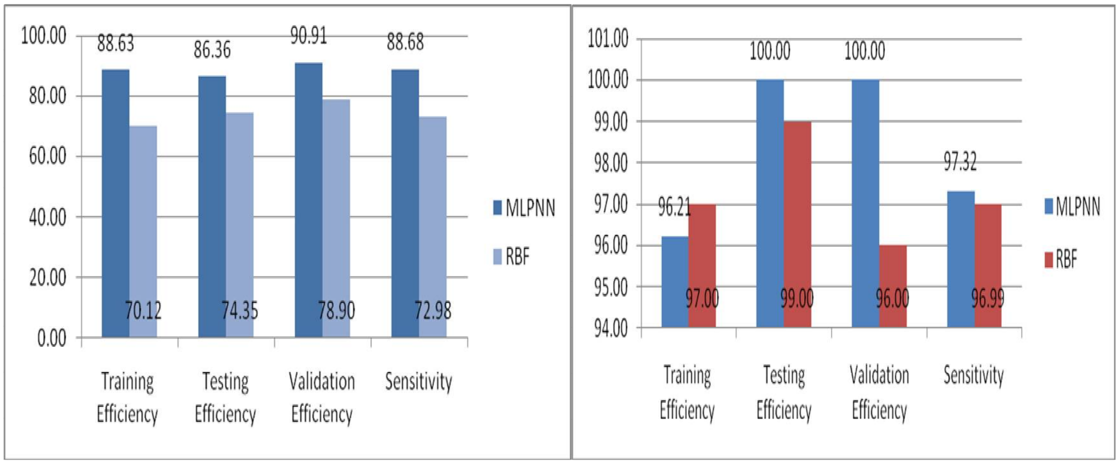


Figure 4.8. Classification Efficiency Analysis of MLPNN and RBF **(*c*)** Six features with same number of hidden nodes (FL:6) (*d*) Ten features (FL:10)

A comparative analysis is done for both the models for seizure classification CAD system. Figure 4.9 represents the training, testing and classification efficiency for the best two models. For these networks the network architecture revealed that number of hidden nodes is different for both the methods.

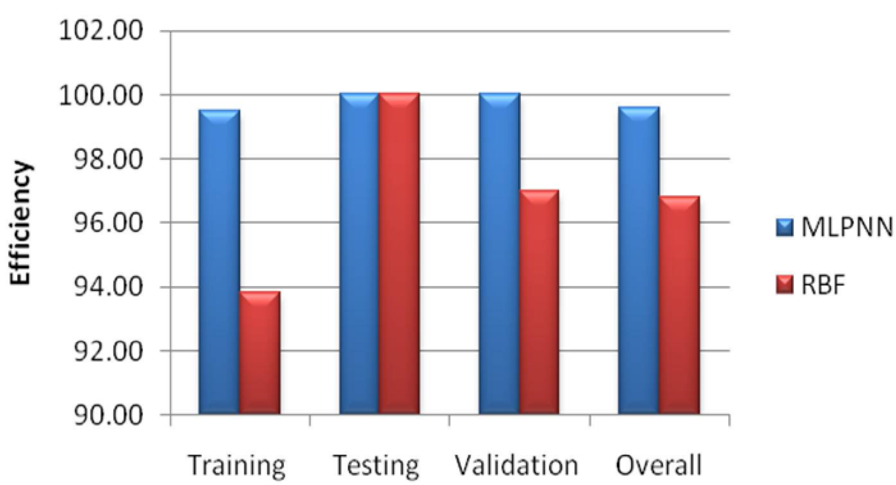


Figure 4.9. Classification Efficiency (in %) for MLPNN and RBF for the final network topology**.**

In the present work, an elaborative comparison has been performed between two machine learning methods for the classification of ictal, inter-ictal and normal state of epileptic patients. Even as prototype, both the ANNs have shown practical performance as demonstrative of the efficiency of the machine learning methods.

Table 4.5: Classification Summary for prediction model with RBFNN providing highest accuracy with topology of 13-30-3

	Normal(Z)	Intict(F)	Ict(S)	Sen Z	Sen F	Sen S	OCA
Normal	97	3	0	97			
INTICT	4	93	2		93.9		96.9
ICT	0	0	100			100	

Note: *Sensitivity, OCA values in %.*

We have demonstrated the feasibility of choosing number of features for classification and concluded that MLPNN revealed a superior model in terms of higher efficiency and number of hidden nodes.

Table 4.6: Confusion Matrix for RBF network with proposed topology

	Normal	Intict	Ict
Total	100	99	100
Correct	97	93	100
Incorrect	3	6	0
ICA (%)	97	93.9	100
IMA(%)	3	6	9

MLPNN could be a very good candidate to achieve the efficiency of 99.3% as compared to 96.9% achieved by RBF with less number of hidden nodes leading to less complexity of the architecture. Results from this study indicate that a classification system based on ANN help in automation of analysis of neurophysiologic signals and the number and type of parameters used as feature set decide the type of network to be used for the better efficiency of the system. RBF gave comparative accuracies in all the experiments. The evaluation parameters for RBF giving maximum accuracy of classification is tabulated in Table 4.5 and confusion matrix for the finalized model is illustrated in Table 4.6. The highest classification accuracy of 96.8% was obtained for 30 hidden nodes with larger number of hidden neurons. Performance of classification techniques is measured in terms of Classification Accuracy and misclassification. The individual classification accuracy is 97%, 93.9% and 100% for normal and interictal and ictal classes with 3%, 6% and 9% misclassification rate. The sensitivity obtained for inter-ictal condition is 93.9% for ictal 100% and for normal 97%. The overall classification obtained is 96.9% for the three class classification problem.

4.2.4. ***Experiment 3*** *This experiment deals with two class classification problems with proposed topology and available binary classifiers for seizure classification*.

4.2.4.1 Methodology

Advanced methods of signal and data analysis as well as increasing powers of computing, provide improved computing tools to record and analyze the EEG signals. The new techniques present a detailed insight to study brain mechanisms; computationally strong signal processing techniques have enhanced the accuracy and precision of analysis of signals. The present methodology studies a bi-class problem to show the generalizability of soft computing paradigm technique. The deployment of the techniques could be used to get deeper information of EEG associated to epilepsy events in an automatic way. In this work thirteen statistical features as extracted and selected from raw signals are chosen in order to investigate the adequacy for the discrimination of two classes of epileptic subject. The results obtained by various classifiers and their classification results are reported.

Support Vector Machine (SVM) as classifier

SVM primarily performs classification tasks by handling multiple continuous and categorical variables by constructing hyper planes in a multidimensional space to differentiate between separates cases of different class labels [31]. It first tries to map the input feature vector into high dimensional feature space, either linearly or by methods depending on kernel type chosen; such that error is minimized over the training dataset. An optimized division is sought so as two classes are separated by the largest margin. It provides users to choose from a number of available modes and kernel functions. In our work we have used primarily Gaussian RBF and polynomial kernels functions because of their localized and finite responses across the entire range of the real x-axis. For our problem the cut-off value used for prediction is 0, i.e. a query vector is regarded as member of positive dataset if its score is greater than 0 and is regarded as member of negative dataset if its score is less than 0 [172]. For polynomial kernel with degree=3.000, gamma=0.077 number of support vectors were found to be 91 (84 bounded). For SVM with radial basis function as kernel with gamma=0.077, number of support vectors was found to be 26. For the performance evaluation and validity of the classifier, three key parameters are considered: Sensitivity, Selectivity and Accuracy which are evaluated by examining the confusion matrix tabulated in Table 4.7.

Table 4.7 Classification performance with SFV using SVM classifier for two-class seizure classification

Feature vector	Classifier	CM			Acc.	Sen.	Spec.	Miss rate
			NOR	**ICT**				
SFV	**SVM**	**NOR**	97	3	98.0%	97.0%	99.0%	0.02
		ICT	1	99				

Note: *CM: Confusion Matrix, Acc. Accuracy for binary classification, Sen: Sensitivity, Spec: Specificity expressed in percentage.*

Evaluation of each confusion matrix is done by computing the efficiency parameters for SVM using every kernel. Considering the problem at hand, the best performance was obtained with SVM with RBF kernel with 98.0% classification accuracy for seizure classification. On analyzing the CM of SVM, the other two prominent results are the sensitivity as 97% and specificity as 99%. Only 2% of cases deviate from the actual classification.

Naive Bayes as classifier

Naive Bayes models are easy to use and interpret and are effective classification tools that incorporate a variety of methods for modelling the conditional distributions of the inputs including normal, lognormal, gamma, and Poisson. It can be clearly interpreted from Table 4.8 that 97.5% of the subjects are rightly classified into three classes with 97% of sensitivity and 98% of specificity. Only 2.5% of subjects are misclassified and give ambiguous results.

Table 4.8: Classification performance with SFV using Naive Bayes classifier for two-class seizure classification

Feature vector	Classifier	CM			Acc.	Sen.	Spec.	Miss rate
			NOR	**ICT**				
SFV	**Naive Bayes**	**NOR**	98	2	97.5%	97.0%	98.0%	2.5%
		ICT	3	97				

Note: *SFV: Signal Feature Vcetor, CM: Confusion Matrix for classification.*

Radial Basis Function neural network (RBFNN) as classifier

In this study, for RBFNN, the Gaussian function and the least squares (LS) criterion are selected as the activation function and the objective function. The inputs to the network are passed to the middle layer kernels followed by output layer. The number of hidden neurons chosen is result of exhaustive training and testing. Seizures are correctly classified with 95.5 % accuracy. The rate of true positive is 96% with true negative rate as 95.0% as depicted in Table 4.9.

Table 4.9: Classification performance with SFV using RBF classifier for two-class seizure classification

Feature vector	Classifier	CM			Acc.	Sen.	Spec.	Miss rate
			NOR	**ICT**				
SFV	**RBF**	**NOR**	95	5	95.5%	96.0%	95.0%	4.5%
		ICT	4	96				

K nearest neighbors (KNN) as classifier

KNN algorithm is a supervised non-parametric classification algorithm that does not require a priori process of training, as it is not necessary to set the value of the parameters. In this algorithm, comparison is carried out with an example set one by one. The pattern is selected in particular class based on the distance function as similarity measure including K nearest neighbors of that class. The optimum values for parameter 'K' is determined empirically by repeated experimentation, in this case k being equal to three and the Euclidean metric is used to calculate the distance between neighboring classes

Table 4.10: Classification performance with SFV using KNN classifier for two-class seizure classification

Feature vector	*Classifier*		*CM*		*Acc.*	*Sen.*	*Sepc.*	*Miss rate*
			NOR	**ICT**				
SFV	**KNN**	**NOR**	96	4	94.5%	96.0%	93.0%	5%
		ICT	7	93				

The classification results obtained with the classifier are tabulated in Table 4.10. The CA obtained is 94.5% for the problem at hand with 96% sensitivity and 93% specificity and error rate came out to be 5%. kNN algorithm has great applicability and has been successfully used in lot of biomedical problems.

Multi layer perceptron neural network (MLPNN) as classifier

The neural network architecture employed in this study is feed-forward network with thirteen neurons in the input layer and two neurons in the output layer. The activation functions used are tan-hyperbolic and softmax for hidden and output layers, with Entropy as error function. BFGS (Broyden-Fletcher-Goldfarb-Shanno) is used for training neural networks. These methods perform significantly better and require a smaller number of iterations to train a neural network given their fast convergence rate and more intelligent search criterion. A clear inference can be drawn from Table 4.11, MLPNN succeeded in classifying the epileptic subjects with the total classification accuracy of 98.5% with misclassification rate of 1.5%. The other two performance metrics are 98.0% and 99.0%.

Table 4.11: Classification performance with SFV using MLPNN classifier for two-class seizure classification

Feature vector	Classifier	CM			Acc.	Sen.	Spec.	Miss rate
			NOR	**ICT**				
SFV	**MLPNN**	**NOR**	98	2	98.5%	98.0%	99.0%	1.5%
		ICT	1	99				

4.2.4.2 Discussions

The classification results obtained for various algorithms as discussed in the previous sections are collectively depicted in Figure 4.10. It shows all variations achieved in performance metrics for different soft computing techniques for this research work.

Analyzing the results comparatively it can be inferred that out of all the tested models MLPNN gave the best results in terms of CA, sensitivity, specificity and misclassification rate followed by SVM and Naive Bayes, and lower classification for the same architecture is given by KNN and RBF. The experimental results show that this classifier promises high classification accuracy, good sensitivity and specificity for two class classification. The proposed model can assist clinicians for diagnosing different epileptic stages in their earlier stage as it has the potential in designing EEG based diagnostic system for detection of electroencephalographic changes

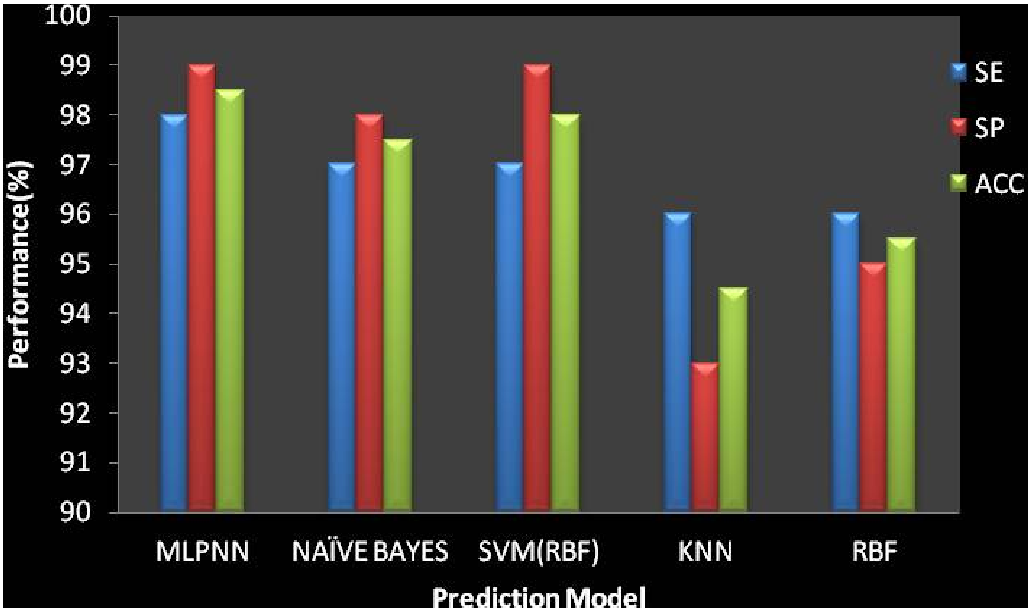


Figure 4.10. Comparative performance analysis for the classification of normal and epileptic subject with various soft computing paradigms.

*4.2.5. **Experiment 4** Hierarchical computer aided diagnostic system for seizure classification*

4.2.5.1 Methodology

In the present work, hierarchical computer aided diagnostic system (HCAD) for classification of normal, ictal and inter-ictal of EEG signals is proposed. In this objective, the focus is on modelling the seizure classification by a hierarchical framework, a variation of the classifier system. The proposed HCAD system for seizure classification comprises of feature extraction and selection module and hierarchical classification module. In continuation with the previous objective, the block diagram of the proposed CAD system is shown in Figure 4.11. The design of classification module is carried out using a hierarchical framework consisting of two stages.

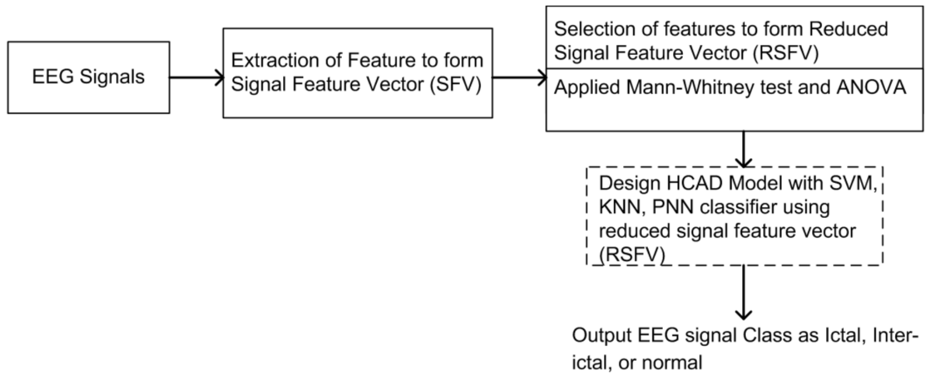
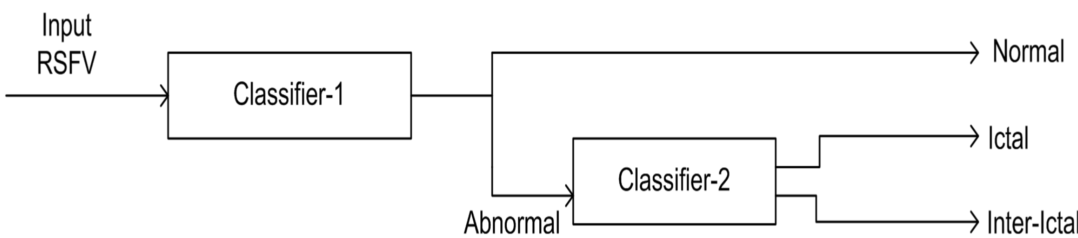


Figure 4.11. (*a*) Proposed HCAD system for classification of seizure using EEG signals

The first stage yields prediction for normal and abnormal class and the second stage yields further classification of abnormal class into ictal or inter-ictal classes. The block diagram of two stage hierarchical classification module is shown in figure 4.11.



(*b*)

Figure 4.11.(*b*) Two stage hierarchical classification module for classification of Ictal, Inter-ictal and normal classes

The proposed model is a two step formulation of classification. The optimally reduced SFV obtained after the feature selection module were passed to the classification module. In first module the classification is to classify signals into normal and abnormal class. If the signal is predicted as belonging to normal class, no further processing is done, and if it is predicted to belong to abnormal class, it is passed to the second classifier for further classification of abnormal class into ictal or inter-ictal class. For the proposed HCAD system, three classifiers SVM, KNN and PNN are extensively used for the classification task as we have inferred from the previous objectives that SVM and NN are best suited for our problem in hand. The mathematical and theoretical details of the classifiers are covered in Chapter 1. The results obtained by these HCAD systems in term of accuracy for binary classification obtained at each stage and the overall classification accuracy (OCA) of HCAD system values are reported in the following section

4.2.5.2 Results and Discussion

HCAD system with kNN Classifier

The unknown data is classified as a class that is most common among its nearest neighbours. The contribution of the nearest samples is more than the farthest samples. The optimum values for parameter '*k*' are determined empirically by repeated experimentation for values of '*k*' ε {1,2,...,9}, and the Euclidean metric is used to calculate the distance between neighboring classes. HCAD system with *k*NN classifier was capable of providing classification accuracy of 93.3% as depicted in Table 4.12. 100 cases were taken as abnormal case for first stage of prediction model and 50 cases for normal condition. Second stage considers these 100 abnormal cases to be classified into ictal and interictal stages giving 94% classification accuracy.

Table 4.12. Confusion matrix and classification accuracy using kNN classifier for HCAD system.

Classifier	CM			Acc. Bin-Class(%)	OCA(%)
KNN 1		**Normal**	**Abnormal**	97.3% (146/150)	**93.3% (140/150)**
	Normal	46	4		
	Abnormal	0	100		
KNN 2		**Ictal**	**Interictal**	94.0% (94/100)	
	Ictal	46	4		
	Interictal	2	48		

Note: *Acc.(Bin-Class): Accuracy for binary classification, OCA: Overall Classification Accuracy*

HCAD system with PNN Classifier

The algorithm defines a probability density function for each class based on the training dataset and the optimized kernel width parameter. Spread parameter (Sp) determines the width of the radial basis kernel function that covers the space of the input features. In the present work, the optimum values for spread parameter 'Sp' is determined empirically by repeated experimentation for values of 'Sp'ε {1,2, . . ,9,10}and the best accuracy results obtained with the proposed algorithm are tabulated in table 4.13.

Table 4.13. Confusion matrix and classification accuracy using PNN classifier for HCAD system.

Classifier		CM		Acc. Bin-Class(%)	OCA(%)
PNN 1		**Normal**	**Abnormal**		
	Normal	48	2	97.3% (146/150)	
	Abnormal	2	98		**95.3% (143/150)**
PNN 2		**Ictal**	**Interictal**		
	Ictal	47	3	95% (95/100)	
	Interictal	2	48		

PNN based HCAD designs yield the binary class classification accuracy values of 97.3% for classification between normal and abnormal classes and 95% for classification of abnormal cases into ictal and inter-ictal cases respectively. The OCA values obtained for PNN based HCAD design is 95.3% that is comparable to OCA obtained with KNN based HCAD design.

HCAD system with Support Vector Machine (SVM) Classifier

A crucial step for obtaining good generalization performance with SVM classifier is the correct choice of the regularization parameter C and kernel parameter γ. The regularization parameter C attempts to maximize the margin while keeping low value for training error. In the present work, extensive grid search is carried out in the parameter space for the values of C and γ using ten-fold cross-validation to obtain optimal values of C and γ for training the SVM model. The SVM classifier has been implemented using LibSVM library [173]. It can be observed from the Table 4.14, that the SVM based HCAD system results in 98% classification accuracy for classifying normal and abnormal cases in the first stage and 97% classification accuracy for classification of abnormal cases into ictal and inter-ictal cases in the second stage. The overall classification accuracy for SVM based HCAD design is 96.6 % which is quiet high as compared to the previous results. Finally, MLPNN was used to design HCAD system, following the same algorithm. It can be observed from Table 4.15, OCA obtained with MLPNN is almost same as OCA obtained with SVM, but accuracy obtained for first stage is more for MLPNN, and for second stage value is more for SVM classifier.

Table 4.14: Confusion matrix and classification accuracy using SVM classifier for HCAD system.

Classifier		CM		Acc. (%)	OCA(%)
		Normal	**Abnormal**		
SVM 1	**Normal**	48	2	98% (147/150)	96.6% (145/150)
	Abnormal	1	99		
		Ictal	**Interictal**		
SVM 2	**Ictal**	48	2	97% (97/100)	
	Interictal	1	49		

.

Table 4.15: Confusion matrix and classification accuracy using MLPNN classifier for HCAD system.

Classifier		CM		Acc. (%)	OCA(%)
		Normal	**Abnormal**		
MLPNN1	**Normal**	49	1	98.6%(148/150)	96.6%(145/150)
	Abnormal	1	99		
		Ictal	**Interictal**		
MLPNN 2	**Ictal**	49	1	96%(96/100)	
	Interictal	3	47		

Note: *CM: Confusion Matrix for classification, Acc.: Accuracy in percentage, OCA: Overall Classification Accuracy*.

From the exhaustive experimentation carried out in the present work, it is observed that the MLPNN based CAD system results in highest classification accuracy of 97% and SVM based CAD system gave 96.6% in comparison with 93.3% and 95.3% as obtained from KNN and PNN based HCAD systems. The promising results obtained from the present work indicate that the proposed MLPNN and SVM based HCAD system can be routinely used for seizure classification in clinical practice. It also provides a direction towards implementation of Hybrid Hierarchical CAD system. The deployment of the techniques could be used to get deeper information of EEG associated to epilepsy events in an automatic way.

4.3. CONCLUSION

This objective aims at developing an automated robust and efficient predictive model to diagnose the state of an epileptic patient using EEG signals. Classification is based on quantitative parameters obtained from neurophysiologic signals which are used to train the networks and the performance of the networks is analyzed to confirm the efficacy of the network. The comparative analysis is based on variation in network topology and in feature vector used for training the networks. Strategically, we have designed a fully automated neural network model, capable of classifying the seizure activity into ictal, interictal and normal state with classification accuracy as high as 99.3% with misclassification error of 0.67%. The stress is also laid on comparison of two machine learning methods (MLM) for prediction of epilepsy with the same dataset. In the field of mathematical modelling, perceptron neural network and radial basis function neural networks have an edge for the classification purposes. In this research work, both the algorithms are tested and evaluated for two class and three class classification and comparative results are analyzed exhaustively.

The accurate classification of the features within two classes is also an important factor to improve the performance of detector. Thus, this objective also aims to develop a computer aided diagnostic system for binary classification of EEG signals. Various methodologies that could be implemented in hardware for monitoring an epileptic patient are checked. The selected classifier must be capable to set a nonlinear decision boundary between the seizure and non-seizure feature vectors Efficacy of techniques is evaluated on the basis of performance measures, sensitivity, specificity and accuracy. It has been observed that artificial neural network and support vector machine with radial basis function kernel are more successful as compared to other soft computing paradigms.

Results from this research work strongly indicate towards classification system based on ANN for automation of analysis of neurophysiologic signals. The number and type of parameters used as feature set are deciding factor to design the topology of network to be used for the better efficiency of the system. The promising results obtained by the proposed SVM based HCAD system in the presence of a diversified dataset used in the present study indicate its usefulness in a clinical environment to assist neurologist for the diagnosis of epileptic seizure during routine clinical practice. CAD system designs with hierarchically placed classifiers improvise the performance by going stepwise from the general classification problem, normal versus abnormal EEG signal to the more particular classification problem which is the identification of exact abnormality. The promising performances observed are demonstrative of the efficiency and efficacy of systems developed for classification and prediction of normal, ictal and interictal conditions of epileptic patients.

CHAPTER 5

DESIGN OF ENSEMBLE CAD SYSTEM FOR CLASSIFICATION OF SEIZURE ACTIVITY

5.1. INTRODUCTION

EEG is a pathological phenomena that has chaotic properties and characterized as a nearly random signal [135] as the neurons are theoretically highly non linear. The non-linear analysis method is effectively applied to electroencephalogram signals to study the dynamics of the complex underlying behavior. Analyzing EEG signals with the aid of nonlinear dynamics takes the advantage of requiring much lower quantity of data. These techniques are superior to traditional linear methods such as Fourier transforms and power spectral analysis [136]. Nonlinear measures like Sample Entropy (SampEn), Hurst Exponent (HE), Approximate Entropy (ApEn), Correlation Dimension (CD), Largest Lyapunov exponent (LLE), Complexity, Mobility quantify the degree of complexity in a time series and help in understanding EEG dynamics and chaotic nature of brain signals [144]. These features can provide additional information about investigated signals and possibly point out characteristics that are not otherwise obvious, but may have clinical relevance. The CAD systems designed and discussed in the previous chapter focused on time domain and frequency domain features. The next objective of our research is to analyze the acquired EEG signals for non linearity by applying signal processing tools and use the non linear features for design of CAD system to classify epileptic seizures. The nonlinear properties of the time series are investigated by calculating Entropy, Hurst exponents and Hjorth parameters during epileptic seizures, and in the interval between the seizures. Ensemble of features including linear and non linear, time and frequency domain and higher order statistics is framed and used for training and testing classifiers. The approach utilized to achieve this objective is depicted in Figure 5.1 and detailed in next section.

5.2. PROPOSED METHODOLOGY

The proposed approach as depicted in Figure.5.1 is potentially applied for the nonlinear analysis and design of CAD system utilizing ensemble of features. The objective of designing CAD systems using feature vectors SFV1 and SFV2 are achieved in objective2. Some features, in particular energy, entropy, skewness, are statistically more significant than the other features as signal to noise ratio, covariance etc. Therefore a reduced set of features comprising contributing features is framed as RSFV for classification purpose. Classification is performed with three classifiers MLPNN, SVM and *k*NN with RSFV as input vector. The EEG signals are chaotic and non linear in nature, so next step leads to extraction of non linear features as Hurts exponent, activity, mobility, entropy to describe the non linearity exhibited by EEG signals. These features after stringent statistical analysis are framed as Non Linear Feature Vector (NLFV) set. To exploit hidden dynamics of EEG signals all the contributing features are ensemble together to obtain a Combined Signal Feature Vector set (CSFV). CSFV is then utilized to design the CAD system using three classifiers which have given better classification results in the previous objectives.

Training set is formed with instances corresponding to all categories with the corresponding labels and remaining instances are framed as testing and validating sets in order to assess the performance of the model. 10-fold cross-validation approach has been used to train the classifier and after all seizure cases are tested once, we compare the predictions and choose the most efficient one for all samples. The model giving the best overall classification is picked and analyzed for performance metrics from confusion matrix. The final step results in evaluating the performance of the classifiers with SFV, RSFV, and CSFV in terms of Accuracy, Specificity and Sensitivity and is tabulated as Confusion Matrix. The same procedure is followed for designing the ECAD system for classification of three classes. The proposed algorithm ensures the soundness and robustness of the design of Ensemble CAD system for seizure classification.

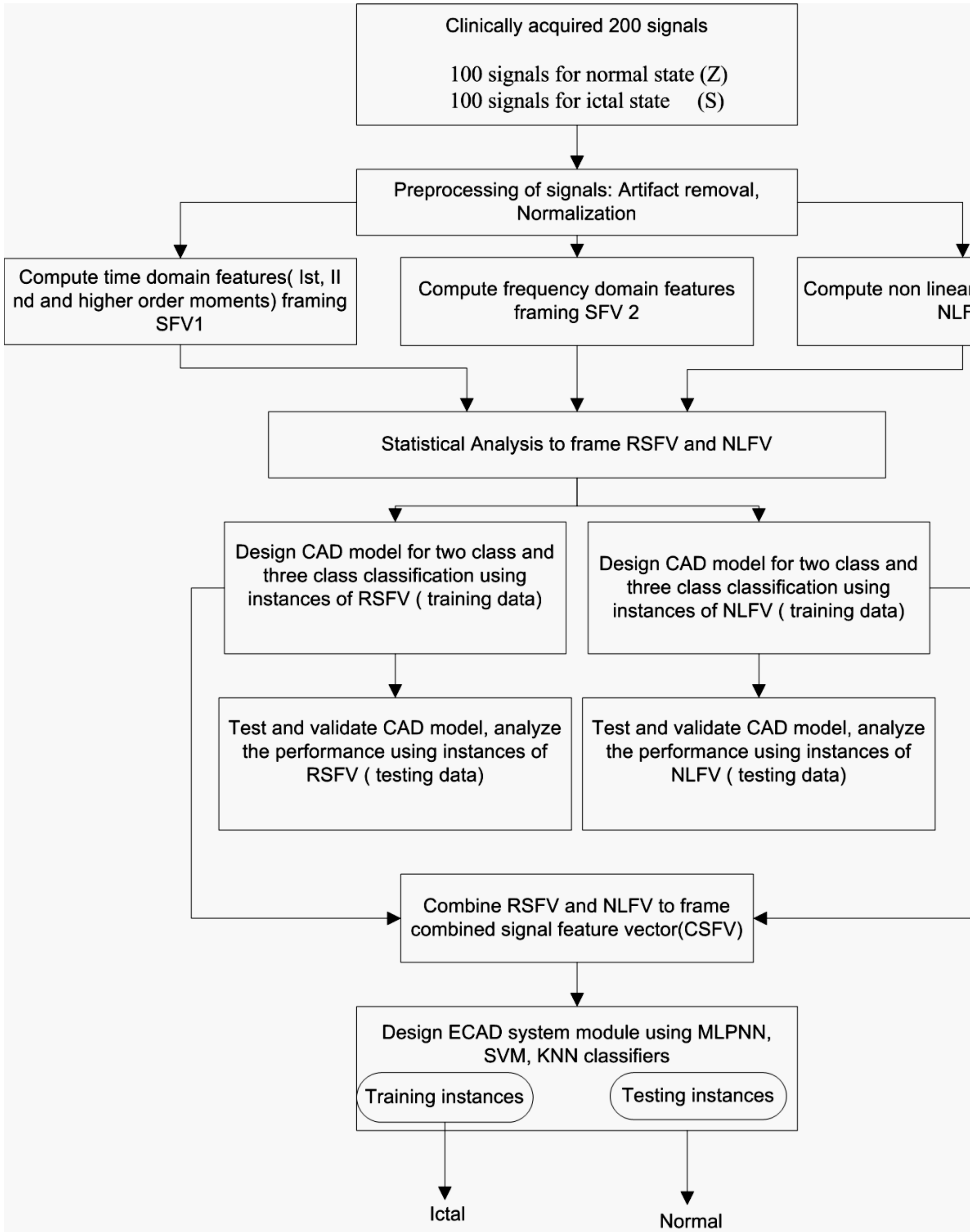


Figure. 5.1 Proposed Methodology for design of ECAD system for seizure classification using EEG signals.

5.2.1. Experimental Workflow

Rigorous experimentation has been carried out for designing and evaluating the performance of the proposed ECAD system for classification. The flow of the proposed design is implemented through two experiments detailed in Table 5.1

Table 5.1: Description of experiments carried out for design of ECAD system for seizure classification

Experiment 1	In this set of experiments, non linear features; Entropy, Hurst exponents and Hjorth parameters for each class of EEG are computed
Experiment 2	This experiment deals with fusion of features to encompass CSFV, to be utilized for three classifiers.

*5.2.2. **Experiment 1** Extraction of Non-Linear Features*

The underlying dynamics of epilepsy and chaotic nature of brain function is investigated by non-linear analysis[174]. Through this work, we intend to present that non-linear analysis can provide a promising tool for detecting relative changes in the human brain signals, which may not be detected by conventional linear analysis. A brief overview of the theoretical background for these features is presented and discussed along with results of the analysis based on these exponents.

5.2.2.1 Entropy

Entropy is a measure of predictability and randomness, higher the values of entropy, the less are system order or more is the system randomness [129]. Entropy provides recognizable variation for physiological signals as EEG signals are considered chaotic and is time varying. ApEn is a statistical parameter, widely used in the analysis of physiological signals, such as epileptic seizure time series data. It is researched that the synchronous discharge of large groups of neurons during an epileptic activity reduces the value of ApEn abruptly. This parameter makes it suitable feature to characterize the EEG signals.

Approximate Entropy Estimator ApEn (m, r, N) of a series is calculated by eqn 5.1 and 5.2 as given below

$$ApEn(m,r,N) = \phi^m(r) - \phi^{m+1}(r) \quad (5.1)$$

$$\phi^m(r) = \frac{1}{N-m+1}\sum_{i-1}^{N-m+1} \ln(C_r^m(i) \quad (5.2)$$

where

$$C_r^m(i) = \frac{N^m(i)}{N-m+1}$$

m is run length, *r* is tolerance window, N represents number of sample points and $C_r(i)$ value measures the frequency of occurrence of patterns similar to a given one of window length *m* [129].

5.2.2.2 Hurst Exponent (HE)

Hurst Exponent is a valuable asset in analysis of EEG signal as it provides a means of classifying time series in terms of predictability."Representation of signals by Hurst exponent, both in numerical or visual form, makes a good basis for the development of stage recognition algorithms as well as for visualization that could ease the work of medical doctors"[175]. Hence, Hurst exponent is proven to be an effective analysis tool for better understanding of the nature of EEG.

It is defined as the relative tendency of a time series to either regress to a longer term mean value or 'cluster' in a direction. It is the measure of similarity between observations as a function of the time lag between them and also a measure of autocorrelation (persistence and long memory). The value of the HE ranges between 0 and 1. A value of HE lying between 0 and 0.5 indicates a time series with negative autocorrelation, meaning thereby, a decrease between values will be followed by another decrease. The value of HE lying between 0.5 and 1 indicates a time series with positive autocorrelation indicating an increase of values will be followed by another increase, and value of H=0.5 indicates a "true random walk", that any particular value of the signal will be arbitrarily followed by a decrease or an increase in the signal [176].

The method which was introduced by Hurst, for the problem stated in the previous paragraph, is known as rescaled range analysis or the R/S statistics or analysis. A brief description of R/S analysis is given in the following section.

Considering a time series of full length N and divide it into a number of shorter time series of length n = N, N/2, N/4. Hurst exponent is defined in terms of the Rescaled Range as given by eqn 5.3 as follows:

$$E\left[\frac{R(n)}{S(n)}\right]=Cn^{H} \qquad \text{as n --}\infty \tag{5.3}$$

Where $\frac{R(n)}{S(n)}$ is the Rescaled Range, E[x] is the expected value, n is the time of the last observation (e.g. it corresponds to X_n in the input time series data.) and H is a constant. The Rescaled Range For a (partial) time series of length n, is calculated for a time series, $X=X_I, X_2, X_{3,}\ldots X_n$ as follows:

1. Calculate the mean of the observations;

$$m=\frac{1}{n}\sum_{i=1}^{n}X_i$$

2. Create a mean-adjusted series;

$$Y_t = X_t - m \qquad \text{for } t = 1,2,3\ldots.n$$

3. Calculate the cumulative deviate series Z;

$$Z_t = \sum_{i=1}^{t} Y_i \qquad \text{for } t = 1,2,3\ldots.n$$

4. Compute the range R;

$$R(n) = \max(Z_1, Z_2 ... Z_n) - \min(Z_1, Z_2 ... Z_n)$$

5. Compute the standard deviation S;

$$S(n)=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i-m)^2}$$

Figure 5.2 Computation of Hurst Exponent of EEG signals

The Hurst values were plotted for the acquired signals in which a unique trend in the EEG activity was observed. The plots of HE for different classes show significant changes before and after the seizure. The results clearly indicate that the brain undergoing epileptic seizure is shows long term positive correlation whereas; the normal brain exhibits randomness, the signals are more or less stochastic.

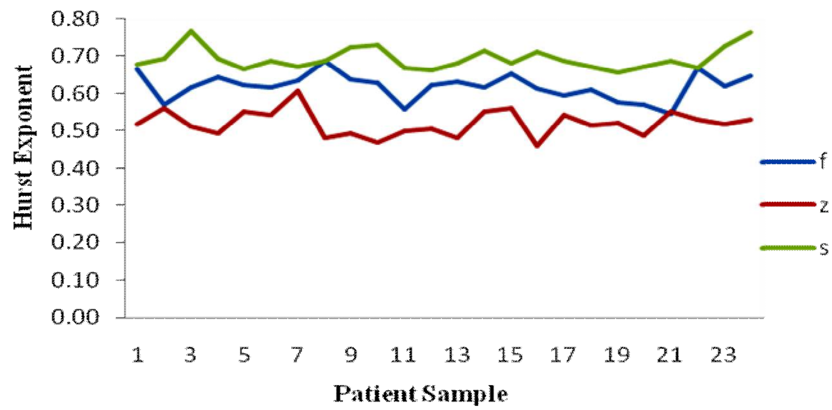


Figure 5.3 Value of Hurst Exponent for different classes for number of patients f-inter-ictal, z - normal, s-ictal.

It was also observed that the HE values of EEG seizure activity decreases before initiation of ictal activity, and the HE value increased during ictal attacks. From Figure 5.3 it is seen that Hurst exponent values obtained for ictal period denoted by S (seizure) is in the range of 0.65-0.78, whereas the HE value for Z (normal) and HE value for (inter-ictal) F period is between 0.57-0.69. The results clearly indicate an increase in the HE values during seizures, pointing to an increase in the degree of self-similarity. The inference drawn from these results are: that an increasing or a decreasing trend of the time series is present during the seizure activity, increase in HE represents the reduction in brain system complexity, and number of the dynamic equations required for description of brain in the seizure state decreases.

5.2.2.3 Hjorth Parameters

The Hjorth parameter is one of the ways of indicating statistical property of a signal and is represented by three parameters shown in Table 5.2 Activity, Mobility, and Complexity. The first parameter, Activity is the variance of the time function that indicates measure of the mean power of the signal. The value of this parameter returns a large value if many high frequency components of the signal exist. Second parameter, Mobility is defined as the square root of the ratio of variance of first derivative of the signal and that of the signal. Mobility can be computed as the ratio of standard deviation of the slope and standard deviation of the amplitude. Complexity parameter indicates change in frequency; it gives the variation in the shape of a signal with reference to pure sine wave. It is expressed as number of standard slopes actually seen in the signal during the average time required for one standard amplitude. The value of Complexity converges to 1 as the shape of signal gets more similar to pure sine wave [177]. Table 5.2 represents high-informative feature in test EEG signals and their significance.

Table 5.2 Hjorth parameters representation

FID	***Feature***	***Formulae***
F14	Activity	$\sigma_j^2 = \frac{1}{n}\sum_{i=1}^{n}(x_j(i)-\mu_j(x_j))^2$
F15	Mobility	$mob(x_j) = \sigma_{\Delta j} / \sigma_j$
F16	Complexity	$compx(x_j) = \frac{\sigma_{\Delta 2j} / \sigma_{\Delta j}}{\sigma_{\Delta j} / \sigma_j}$

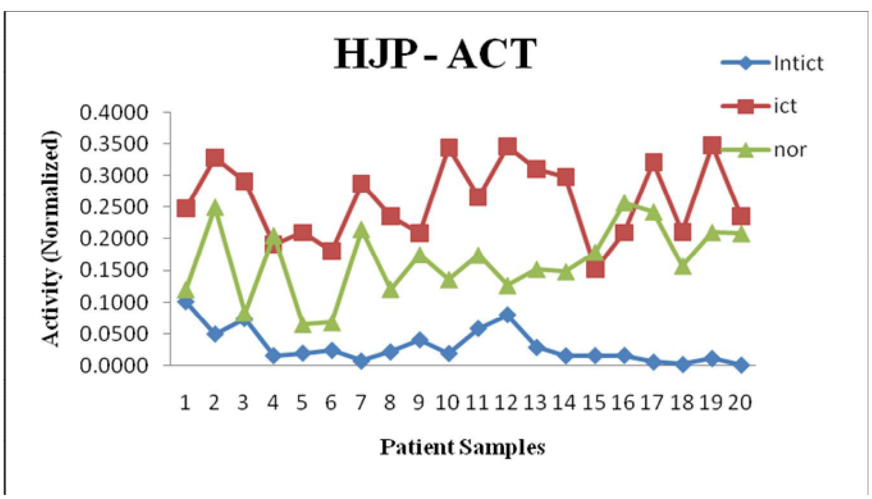


Figure 5.4 (*a*) Comparative graph of Activity for all the three different condition of epileptic subjects

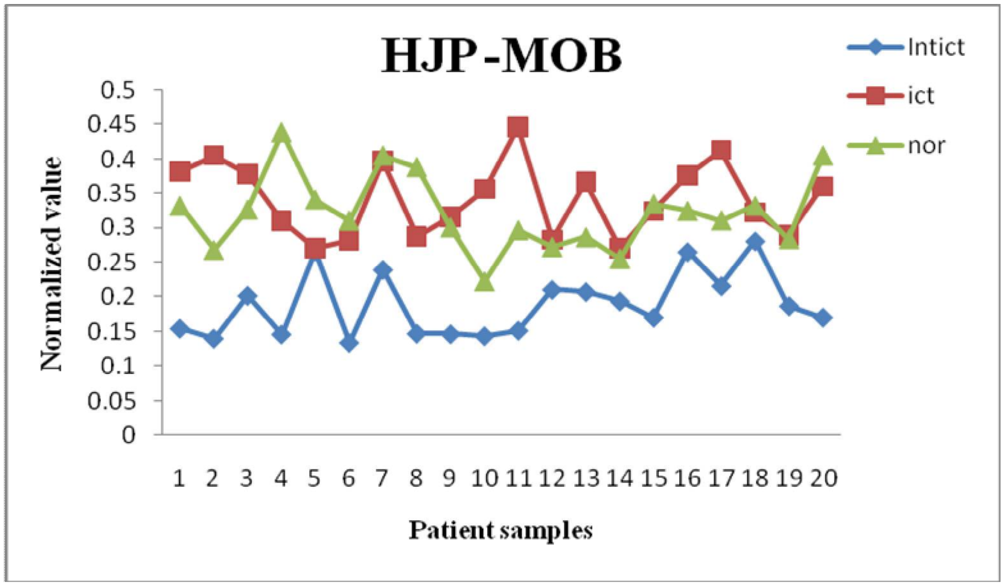


Figure 5.4 (*b*) Comparative graph of Mobility for all the three different condition of epileptic subjects

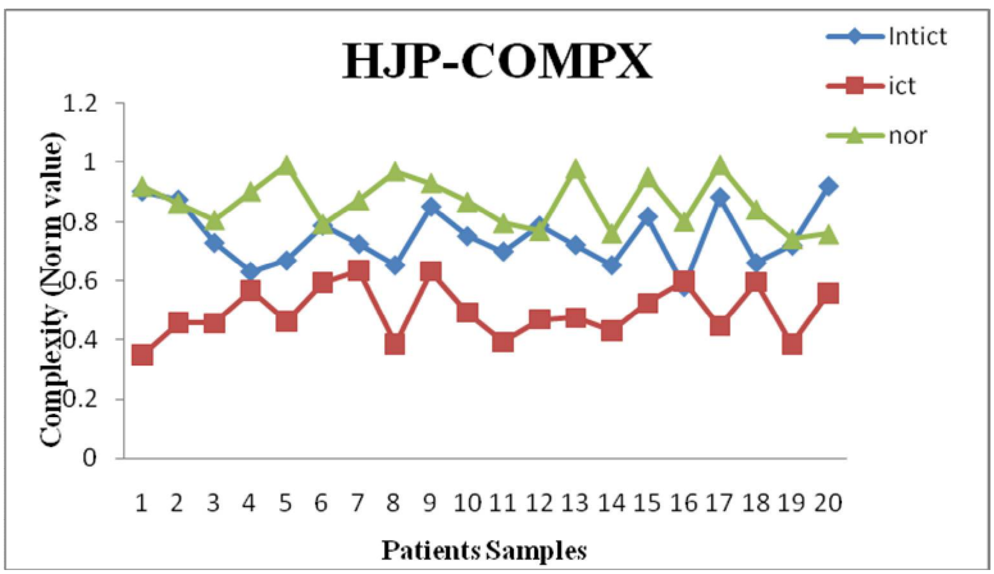


Figure 5.4 (*c*) Comparative graph of Complexity for all the three different condition of epileptic subjects

All these parameters are comparatively depicted for three classes in Figures 5.4. Signals of twenty subjects are taken to represent the complete set. Figure 5.4 (*a*) represents the activity of signals for the three classes of EEG signals. Activity during seizure is more as compared to activity of a normal patient. Figure 5.4 (*b*) depicts mobility, does not give clear cut distinction between normal and ictal conditions.

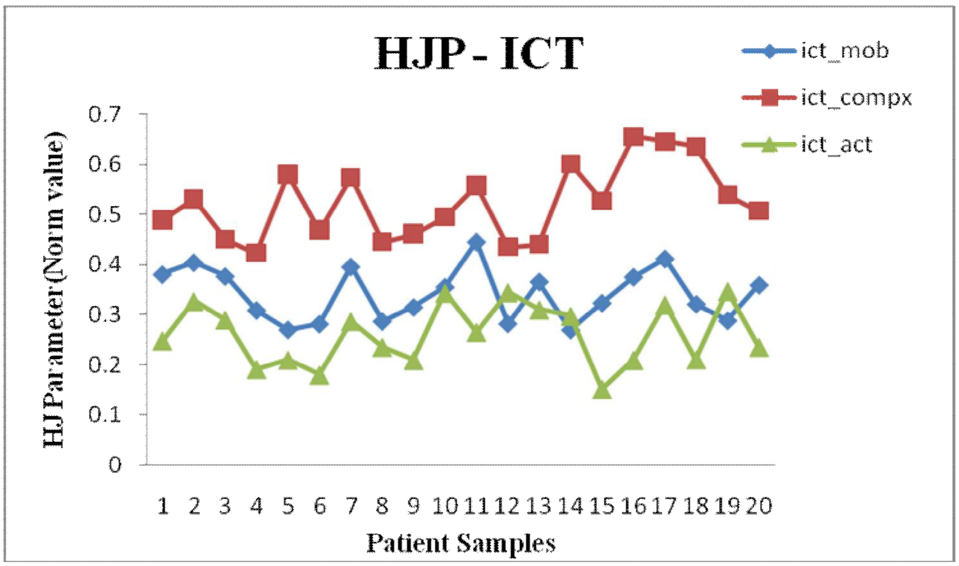


Figure 5.5. **(*a*)** Comparative graph of all the three Hjorth parameters for ictal condition of the subjects

Figure 5.4 (*c*) represents complexity feature of Hjorth parameters for three conditions for different subjects and it is observed that ictal state has lower value of complexity as compared to normal. Further, Figure 5.5 depicts comparative graphs of all conditions with respect to three parameters.

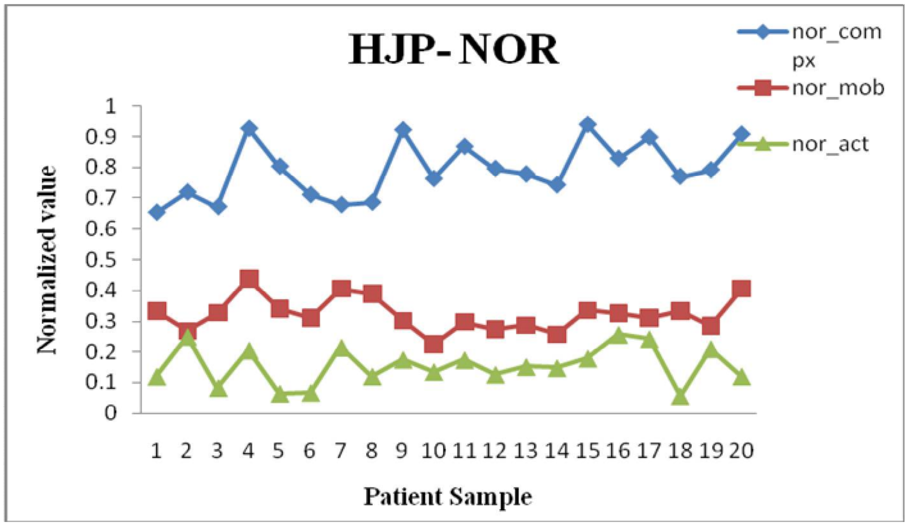


Figure 5.5 (*b*) Comparative graph of all the three Hjorth parameters for normal condition of the subjects

5.2.3. Statistical Analysis of Non linear feature set

Normalization of the parameters is done with max–min approach in order to avoid possible problems caused by inadequately scaled features. This method allows variables to have diverging means and standard deviations but equal ranges (in this case (0, 1)).

Table 5.3: ANOVA analysis of non linear feature set in terms of F ratio and p value.

		SS	df	MS	F	P
Activity	Between Groups	2.082	2	1.041	29.827	0.000
	Within Groups	10.26	294	0.035		
	Total	12.342	296			
Mobility	Between Groups	1.223	2	0.612	109.298	0.000
	Within Groups	1.645	294	0.006		
	Total	2.869	296			
Complexity	Between Groups	4.748	2	2.374	151.893	0.000
	Within Groups	4.595	294	0.016		
	Total	9.343	296			
Hurst	Between Groups	1.223	2	0.612	362.365	0.000
	Within Groups	0.496	294	0.002		
	Total	1.719	296			

Note: *F value = Fisher's Discrimination ratio, p= significant value<0.05, SS=Sum of Squares, MS=Mean Square*

For this work, ANOVA test and Wilcoxon tests were conducted (as we were not sure about the gaussianity of the variables) and the results are tabulated in Table 5.3. From ANOVA analysis, it is well known that if FDR of the extracted feature is high, two classes are distinguishable. The p-value obtained using analysis of variance between groups to decide whether the means are different. This test uses the variation (variance) within the groups and translates into variation (i.e. differences) between the groups, taking into account how many subjects there are in the groups.

It is evident from the table that all these extracted features have high F ratio and low p-value. The p value is less than 0.05, confidence range and F value is quite high, this provides sufficient substance to reject the null hypothesis of equal means. This investigation represents statistically significant features and prominent enough to be used for distinguishing the various classes of EEG signals [178]. Box plots without making any assumptions of the underlying statistical distribution display the variation in samples of distribution. For each subject, obtained values of Mobility and Complexity for each class were used to create box plots. These plots are presented in Figure 5.6. Every figure illustrates values of these components for ictal, interictal and normal stages separately. Again, it can be identified from these figures, how values of Hjorth components differ for three classes. Summarizing Table 5.4 represents the mean values for all non linear features for three conditions.

Table 5.4: Mean values for all non linear features

Feature ID	Ictal	Normal	Inter-ictal
F13	0.76060	0.94852	0.85712
F14	0.23625	0.10008	0.03536
F15	0.32967	0.33987	0.198906
F16	0.49741	0.79078	0.697027
F18	0.69288	0.52559	0.8078

Note: *F13-Entropy, F14- Activity, F15- Mobility, F16-Complexity, F18 Hurst exponent*

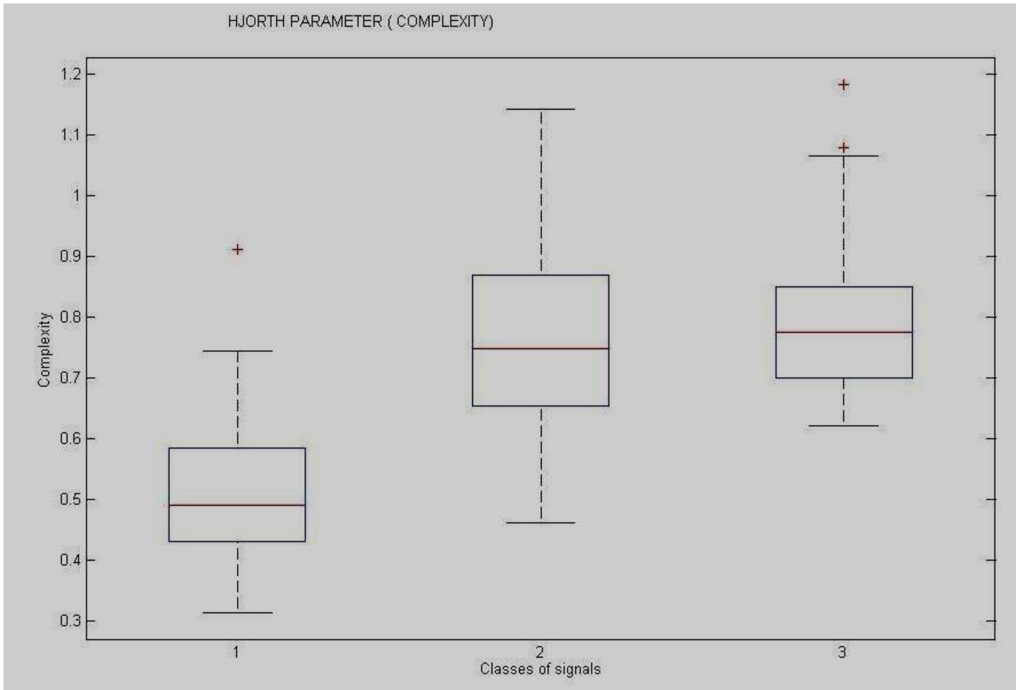


Figure 5.6 (*a*) Box plots for Complexity attribute for three different datasets (Class1: ictal, class 2: inter-ictal, class 3: Normal)

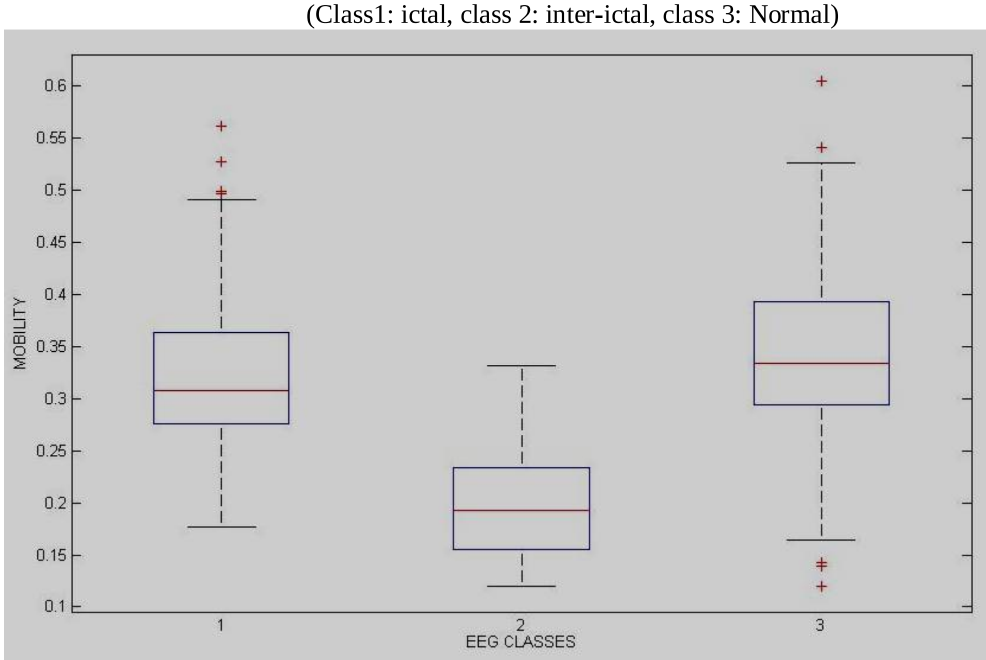


Figure 5.6 (*b*) Box plots for the Mobility attribute for three different datasets (Class1: ictal, class 2: inter-ictal, class 3: Normal).

5.2.4. Classification using non linear parameters

The extracted features are framed as vectors, are used as inputs to the three classifiers that were trained to classify these EEG recordings. The classification analysis was performed on each of the aforementioned parameters and all classifiers were compared based on their performances as illustrated in Table 5.5.

Table 5.5: Classification accuracy of normal and ictal signals using NLFV with MLPNN, KNN and SVM classifiers.

Feature vector	Classifier	CM			Acc.	Sen.	Spc.	Miss rate
			NOR	ICT				
NLFV	SVM	NOR	28	1	96.5%	96.50%	96.50%	3.4%
		ICT	1	28				
			NOR	ICT				
	KNN	NOR	28	1	94.80%	93.00%	96.50%	5.0%
		ICT	2	27				
			NOR	ICT				
	MLPNN	NOR	97	1	98.40%	97.90%	99%	1.5%
		ICT	2	96				

Highest classification accuracy was obtained with MLPNN classifier with high sensitivity and specificity. But, the obtained results are not very encouraging to be used exclusively for detection of epileptic seizures. So we go further to combine non linear features with already assessed attributes in the previous objectives for the same problem. Experiment 2 deals with design of CAD system with ensemble set of linear and non linear features to determine which of these parameters with best classifier yield optimal results.

5.2.5. **Experiment 2.** *Ensemble of feature set*

The building components of this proposed methodology are divided into two sections; fusion of the features in the first and the classification technique in the second. This experiment involves a preliminary evaluation to identify appropriate specific features from a large set of candidate features significant to account for variations exhibited by normal and abnormal signals.

As discussed in Chapter 3, with a large feature set, it is likely that the information captured by some features may be redundant, duplicate and add to computational complexity, so it is beneficial to reduce the number of features by selecting a smaller feature subset. The promising, prominent, statistically analyzed features framed in Chapter 4 used for classification of brain signals into ictal and normal conditions are further improvised by selecting the feature to reduce dimensionality of the data and to improve classification accuracy by framing Reduced Signal Feature Vector (RSFV). Feature reduction is important because it leads to decreases in computational speed, increase in learning accuracy and enhancing learning comprehensibility leads to increased accuracy of the classification.

Fusion of information from EEG signals can be performed at the feature level or at the classifier level. Fusion at the classifier level employs different classifiers for each signal and combines the results thereafter [179]. Feature fusion method is utilized in this approach and features from different modalities are concatenated for further usage. The features from earlier objectives and non linear features from this objective are ensemble together to frame Combined Signal Feature Vector (CSFV). In the attempt to achieve the optimal results of maximum sensitivity on the training data and maximum specificity the proposed combination of consolidated parameters is framed into simple but effective matrix.

As shown in Chapter 4, the simplicity of linear discriminate classifiers provides good results for our problem at hand. Therefore, to investigate the feasibility of using ensemble of features for seizure classification three main classifiers were trained and tested on a subset of the features as outlined in chapter 4. The classification performance and results obtained are discussed in the next section.

5.3. RESULTS AND DISCUSSION

5.3.1. Classification Analysis

The two most important criteria of choosing a best classifier method are accuracy and computational efficiency, so we investigate the recital of our approach in terms of accuracy and computational efficiency. To evaluate the performance of ECAD system; classifiers are designed utilizing RSFV and CSFV. The dataset is divided into subsets for training, testing and validating the classification. As part of the training, 70% of the feature vectors were used to train each classifier in order to prevent the network from being trapped in a local minimum and to rule out the possibility that any differences in the results were obtained just by chance. 30% of data was used then for testing and validating the results with 10-fold cross validation. The performance evaluation for the classification of seizures is performed with three classification algorithms ANN, KNN and SVM, in terms of sensitivity, specificity and accuracy.

To start with, classifiers are trained with RSFV as input and results obtained are tabulated in Table 5.6. It is observed from the results, the highest classification accuracy of 99% is achieved with MLPNN with 99% sensitivity and 99% specificity. Taking SVM as classifier 98.7% of accuracy is obtained with 100% sensitivity and 97.5% specificity.KNN as classifier provides lower value of performance metrics.

Classification with different classifiers with CSFV data set result in higher value of performance measure parameters as reported in Table 5.7. 100% classification accuracy is obtained with MLPNN and SVM with highest value of sensitivity and specificity being 100%, with 0% misclassification rate. KNN classifier provides better results as compared to the previous obtained results.

Table 5.6: Classification accuracy of normal and ictal signals using RSFV with MLPNN, KNN and SVM classifiers

Feature vector	Classifier		CM NOR	CM ICT	Acc.	Sen.	Spec.	Miss rate
			NOR	**ICT**				
RSFV	SVM	**NOR**	40	0	98.7%	100%	97.5%	1.3%
		ICT	1	39				
			NOR	**ICT**				
	KNN	**NOR**	39	1	95.0%	97.5%	92.5%	5.0%
		ICT	3	37				
			NOR	**ICT**				
	MLPNN	**NOR**	39	1	99.0%	99%	99%	1%
		ICT	1	39				

***Note**: RSFV: Reduced Signal Feature Vector, CM: Confusion Matrix for classification.*

Table 5.7: Classification accuracy of normal and ictal signals using CSFV with MLPNN, KNN and SVM classifier

Feature vector	*Classifier*		*CM* NOR	*CM* ICT	*Acc.*	*Sen.*	*Spc.*	*Miss rate*
			NOR	**ICT**				
CSFV	SVM	**NOR**	40	0	100%	100%	100%	0.00
		ICT	0	40				
			NOR	**ICT**				
	KNN	**NOR**	39	1	96.2%	97.5%	95.0%	4.0
		ICT	2	38				
			NOR	**ICT**				
	MLPNN	**NOR**	40	0	100%	100%	100%	0.00
		ICT	0	40				

Note: *CSFV: Combined signal feature vector, CM: Confusion Matrix for classification.*

On observing the achieved results, we can infer that there is net increase of 1.3% in classification accuracy with SVM and an increase of 1% with MLPNN. The misclassification rate has reduced to 0% giving optimal results. We are able to achieve 100% classification accuracy on the EEG epileptic database for the pair of healthy subjects and epileptic patients during seizure activity. Based on these results, it can be concluded that non linear features can differentiate well between normal and epileptic patients, providing high classification accuracy, and zero misclassification rate and thus can be used in the field EEG analysis.

5.3.2. Comparative Performance Analysis

The proposed approach uses ensemble of linear and non linear parameter selection for the design of ECAD system. It is evident from Figures 5.7 and 5.8, that there is much influence on performance of the proposed approach when CSFV is utilized as compared to RSFV. As a result of statistical analysis of all the extracted features and appropriate choice of features, the proposed methodology is able to capture the variations more effectively. More importantly; the proposed approach for classification of seizure and seizure-free EEG signals outperforms the existing methods.

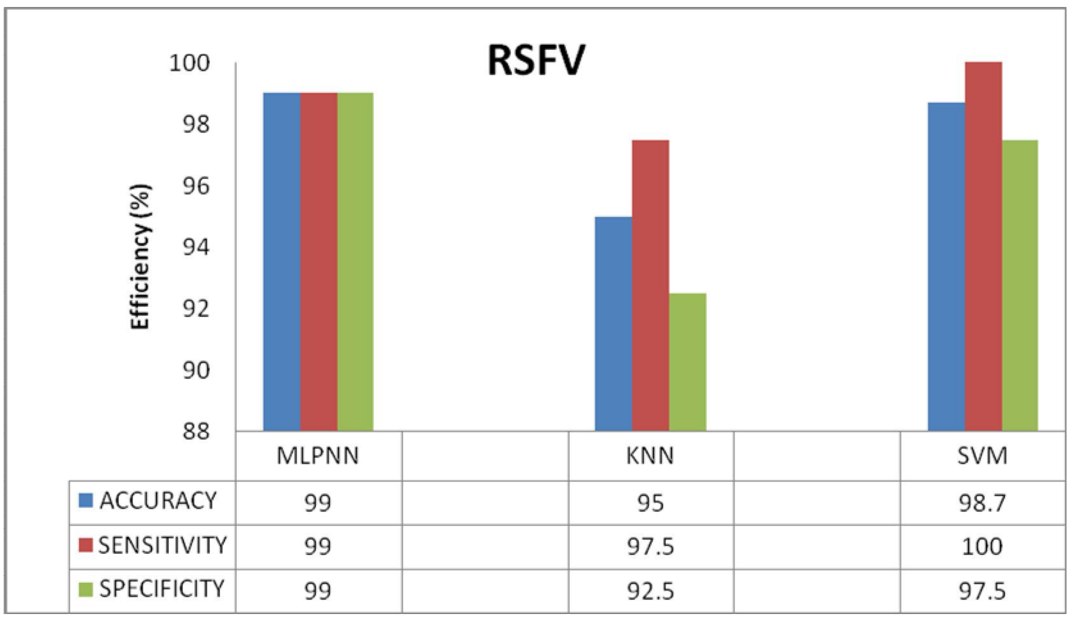


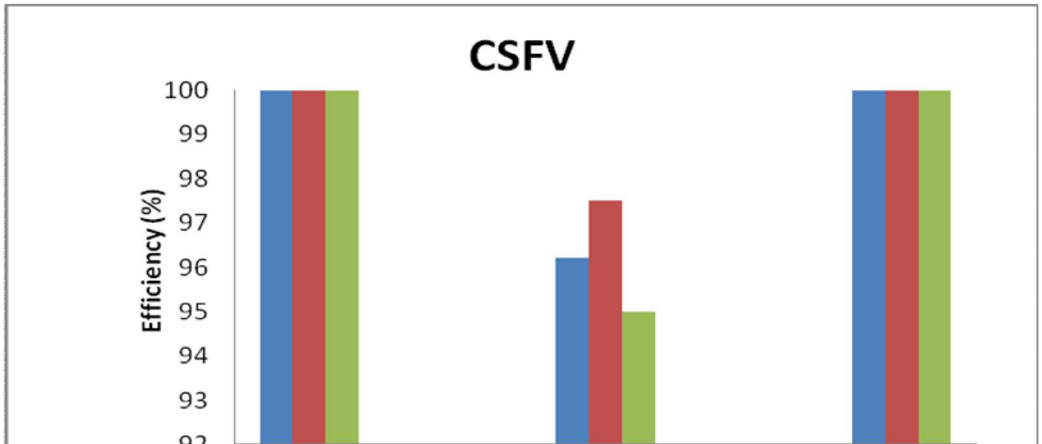
Figure 5.7. Performance analysis of different classifiers in terms of performance metrics with reduced signal feature vector (RSFV).

To ensure the soundness of the proposed method key steps were taken.

- ✓ The choice of features selected and the suitability of these parameters by lying focus on a statistical analysis of these descriptor components.
- ✓ Fusion of features so as to ensure inclusion of all variations of EEG signals.
- ✓ Choosing the best three classifiers in order to assess linear seperability between all

epileptic and non-epileptic cases.

- ✓ Test the results using the classifier based on the three performance metrics used (accuracy, sensitivity, specificity)
- ✓ To evaluate the performance of the classifiers based on ROC parameters



	MLPNN		KNN		SVM
ACCURACY	100		96.2		100
SENSITIVITY	100		97.5		100
SPECIFICITY	100		95		100

Figure 5.8. Performance analysis of different classifiers in terms of performance metrics with combined signal feature vector (CSFV).

Classification results of the algorithm are displayed by a Confusion Matrix for whole dataset for MLPNN in Table 5.8. It is clearly depicted in Table 5.8, that the proposed design is 100% efficient in classifying the EEG signals into normal and epileptic patients. The classification rate is 100% for class 1 and 100% for class 2 and overall classification of the system is 100%. According to the confusion matrix, no misclassification has occurred for any of the classes.

Table 5.8. Confusion Matrix for designed ECAD system for complete dataset.

Actual class Predicted class	NOR	ICTAL	Sensitivity	Specificity	Accuracy
NOR	100	0	100%	100%	100%
ICTAL	0	100			

Figure 5.9 (*a*)-(*b*) depict the comparative analysis of ECAD system designed with MLPNN classifier and SVM classifier based on performance measures with respect to the variations in the feature vector. The comparison is made on the basis of accuracy; sensitivity and specificity. The figures are self explanatory, showing the variation in results achieved due to variation in the feature set chosen for classifying the signals.

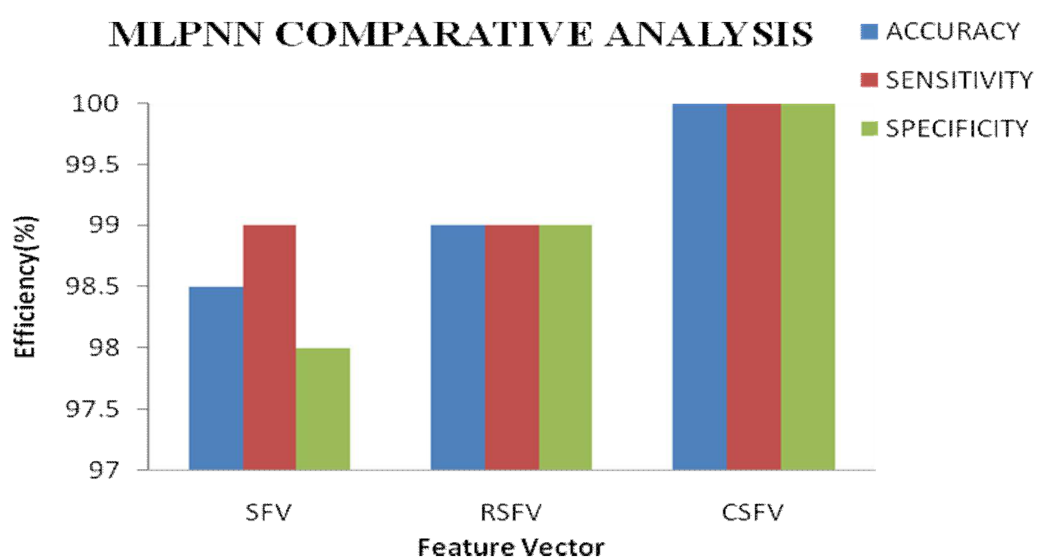


Figure 5.9 (*a*) Comparative depiction of classification efficiency with SFV, RSFV and CSFV for the design of ECAD system

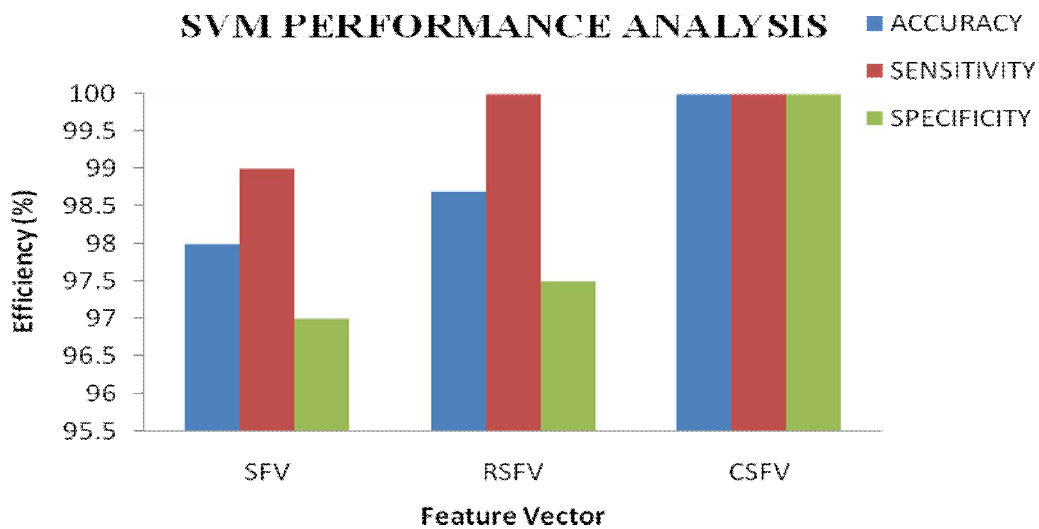


Figure 5.9 (*b*): Comparative depiction of classification efficiency with SFV, RSFV and CSSFV for the design of PS based CAD system

It can be inferred from the results presented in Figure 5.9 that the proposed ECAD system with CSFV as feature vector outperforms the CAD systems with RSFV and SFV with same classifier. A consistent improvement in classification accuracy is observed in both the cases. These results also demonstrate the effect of parameter selection on overall performance of CAD design. The effect on accuracy is found to be more positive, encouraging and convincing to the fact that non linear features are more effective in capturing variations.

One of the advantages of the proposed methodology is the computational efficiency and usability. The experimental results obtained demonstrate that the ECAD design is efficient for extracting and selecting features to represent the EEG signals and MLPNN and SVM classifier have the inherent ability to solve a classification task for these EEG signals.

5.3.3. Receiver Operating Characteristic (ROC) curve

Receiver Operating Characteristic (ROC) curve is a graphical plot of the sensitivity of the classifier (“True Positive Rate”) against (1-specificity) (“False Positive Rate”) [34]. A diagonal line corresponds to no seperability at all; any line far from diagonal line depicts seperability between classes. The area below the curve (between 0 and 1) gives an impression of the discrimination ability of the classifier. A good test for seperability with ROC curve is; sensitivity rises rapidly and 1-specificity hardly increases at all until sensitivity becomes high, in graphical terms the coordinate (0,1) give the maximum seperability.

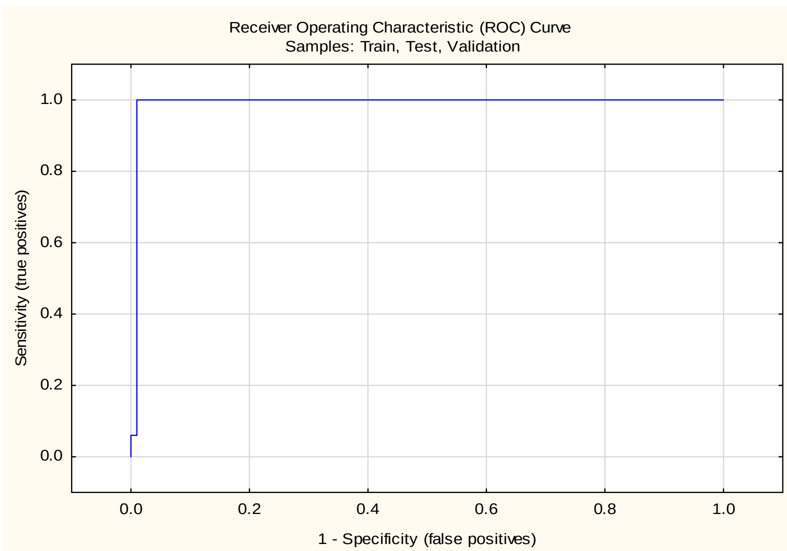


Figure 5.10. Receiving operating characteristic (ROC) curve for the designed CAD system with RSFV.

The Receiving operating characteristic (ROC) curve for the CAD system with RSFV is shown in Figure 5.10 on the testing and validating vector set of EEG data set. The area values under ROC curve is 0.967 for all pairs of EEG data sets. The curve on the y axis is accurate but there is slight change is on x-axis, which is in conformity with the results obtained and depicted in Figure 5.9 (*b*).

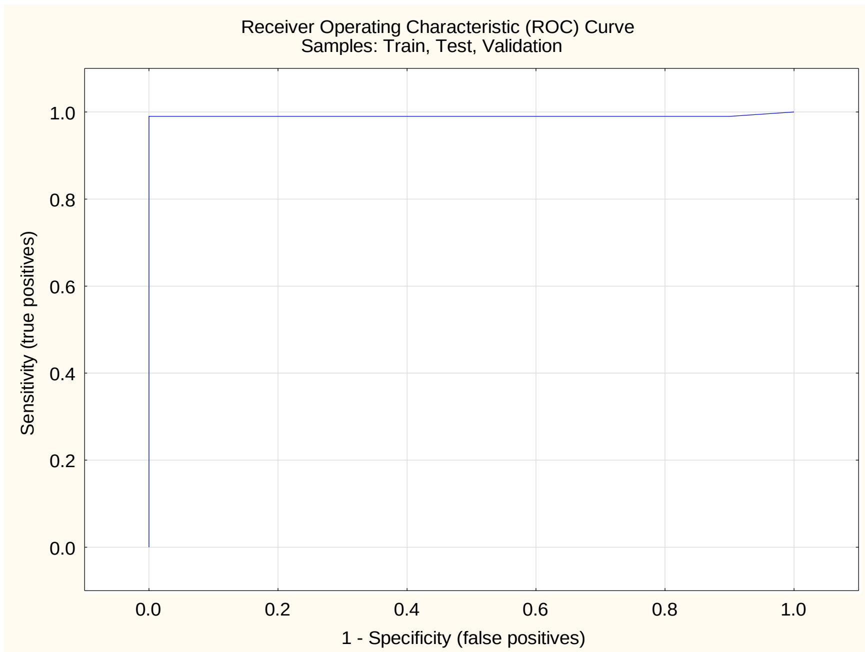


Figure 5.11. Receiving operating characteristic (ROC) curve for the designed ECAD system

The Receiving operating characteristic curve for the designed ECAD system with CSFV is drawn in Figure 5.11 on the testing and validating vector set of EEG data set. The area values under ROC curve is 0.998 for all pairs of EEG data sets. There is 3% increase in Area Under Curve (AUC) in ROC curve as indicated in the figures. Hence it is concluded that the proposed approach has a high discriminating capability to classify EEG signals. The models proposed in this work produces excellent results and demonstrates the superiority of our approach for classifying EEG brain signals into normal and ictal signals.

5.4. CONCLUSION

In this work, we have proposed a novel ensemble CAD system to detect the epilepsy condition using EEG signals. The time domain and feature domain features are ensemble with non linear features and coupled with MLPNN and SVM classifier to yield high classification accuracy results. Experimental results show that the resulting attributes when used for classification results in 100 % classification accuracy with CSFV, 99.5% with RSFV and 98.5% classification with SFV with MLPNN as classifier. The net increase in percentage classification accuracy is 1.5% with MLPNN, 2% with SVM and 1.8% with KNN with CSFV when compared to efficiency obtained with SFV. The increase in the area under ROC curve is 3%. The results obtained are clear indicators of the great potential in classifying epileptic signals and normal signals with 100% classification accuracy. The study presents a unique analysis, both in the identification of the effective features from among the features used for the representation of the problem and also in the identification of effective classification algorithms from among the 5 most popular classification algorithms. The feature extraction and selection are meant to reduce dimensionality without the loss of important information embedded in the signal and to simplify the amount of resources needed to describe a huge set of data accurately. As features selected for the present work were chosen with an eye towards real-time implementation, this method has the potential to be applied in prototype implantable devices for treating epilepsy.

.

CHAPTER 6

DESIGN OF A MODULE BASED CAD SYSTEM (MCAD)

6.1. INTRODUCTION

A stochastic process is known to be stationary if its statistical properties do not change over time. A strictly stationary stochastic process is one where for given $t_1, \ldots, t_\ell$ the joint statistical distribution of $X_{t1}, \ldots, X_{t\ell}$ is the same as the joint statistical distribution of $X_{t1}+\tau, \ldots, X_{t\ell+\tau}$ for all ℓ and τ; that implies that all moments of all degrees (expectations, variances, third order and higher) of the process, anywhere are the same. It also means that the joint distribution of (X_t, X_s) is same as $(X_{t+\tau}, X_{s+\tau})$ and hence cannot depend on s or t but only on s-τ. For the purely random process, the mean and variance are constant functions in time. Weak stationarity of signals imply that mean and variance of a stochastic process do not depend on t (that is they are constant) and auto covariance between X_t and $X_{t+\tau}$ depends only on lag τ (τ is an integer, the quantities also need to be finite) [180]. EEG distribution is mutivariate Gaussian process even if mean and covariance properties change from segment to segment, thus exhibiting non-stationary signal properties. One of the methods that have been used to test stationarity is to divide the signal into segments that fulfill the criterion of shorter stationary time series [181]. To comply, time series is divided into number of epochs or sub samples and dynamics is assumed to be approximately stationary termed as quasi-stationary.

EEG segments chosen from the dataset are recorded from a single channel and are a sequence of time samples, chosen under an inclusion criterion of weak stationarity [16]. One approach for analyzing a time series is to divide the time series into small subsamples (termed as epochs) and inspect whether there are certain underlying dynamics in a particular epoch. In this objective, we intend to work on effect of non stationarity dynamics of EEG signals by dividing complete data set into small segments (epochs) where the stationarity of the signals can be checked. This method encompasses the quantification and analysis of nonstationary neurophysiologic signals. Classification accuracy can be adaptively enhanced, if the underlying reasons for the nonstationarity are known a prior. In this objective a novel framework for nonstationary data analysis and design is presented. The key contribution of this work is introduction of a novel modular approach that aims to identify an algorithm that provides highest level of classification accuracy with selected features and selected epoch size. This type of analysis may also contribute to the better understanding of signals nature and characteristics.

6.2. PROPOSED METHODOLOGY

Two signal processing steps are typically required before the classification of obtained bio signals: signal preprocessing and feature extraction. The signals are prepared for further processing, in the preprocessing step that include filtering out unnecessary frequencies, artifact rejection, signal scaling, and signal transformations [36]. Feature extraction and selection is the most crucial step in a successful bio signal algorithm. Both of these steps are already covered in the previous chapters.

In Chapter 4 and 5, CAD models for the classification of EEG signals have been developed. The experimental results show that the algorithm works well to solve a classification task using representative features for classification. The sample size was taken as the complete signal of 4096 samples. Continuing with the same algorithm, the recorded EEG segments having fixed samples of each class are segmented into different sizes of non overlapping modules (M); that comprises the samples based on binary relation, to ensure reasonable time resolution. For each module, several features are extracted that characterize the dynamics of signals and statistically analyzed as already discussed in previous chapters. The features those were deemed significant as inferred by the previous algorithm are extracted from the segments and reframed to reduce signal feature vector (M_xRSFV) of module M_x, where x is the size of module. These feature vectors are used as inputs for the classifiers which were found to be efficient for this problem; and performance measures; accuracy, sensitivity and specificity are determined [33] as elaborated in the previous chapters.

It is not clear what segment length best captures the signal characteristics and sufficient for stationary check. So, to investigate what epoch length would be most appropriate for classifying epileptic seizure, classification task was performed for different epoch size '*n*' resulting into modules of size '*m*' and comparative system performance was subsequently evaluated. Figure 6.1 depicts the work flow of the proposed method with clear and distinct stages followed in this approach.

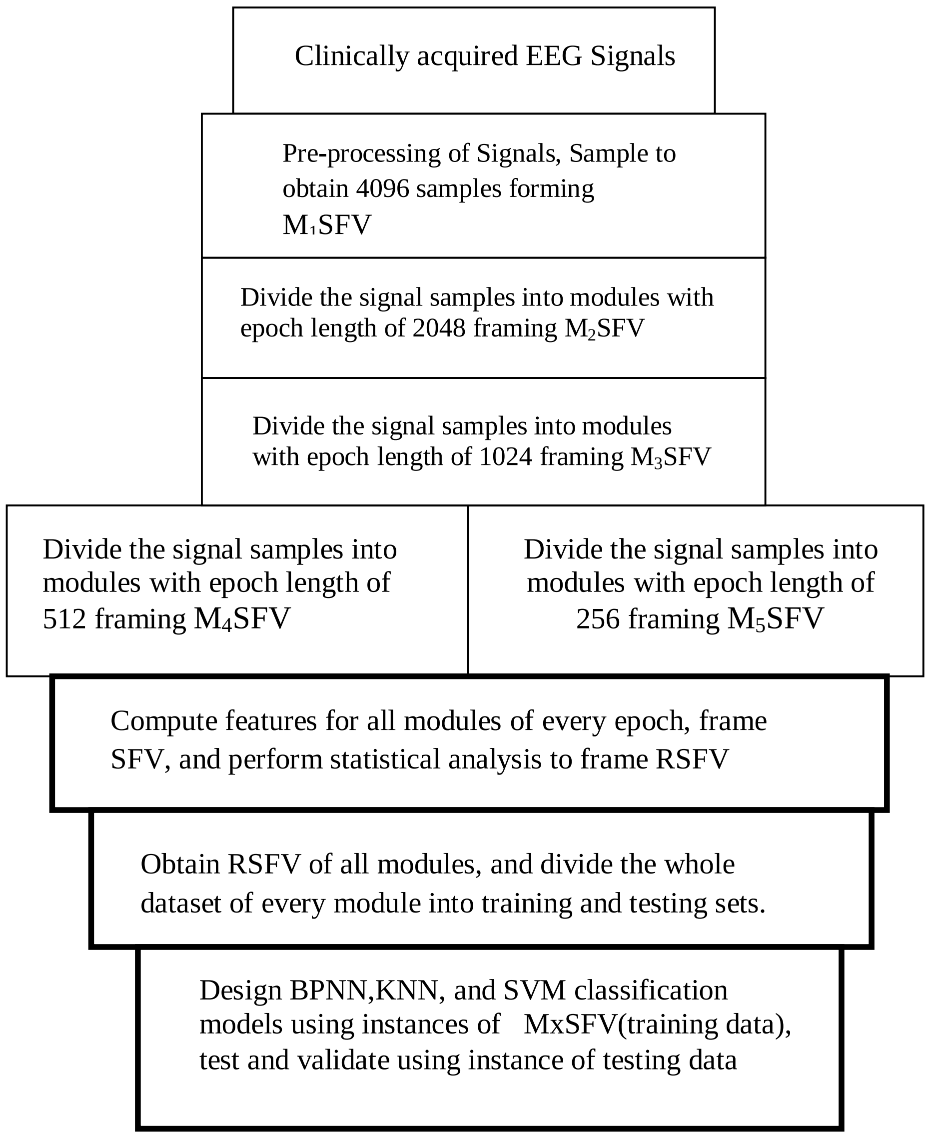


Figure 6.1. Proposed algorithm for design of Module based CAD system

This objective aims at designing Module based Computer Aided Diagnostic (MCAD) system for the classification of EEG signals. The algorithm is epoch based technique in which decision making is performed in two stages. In the first stage, representative features from each epoch are extracted and framed as feature vector, and the second stage utilizes these feature vectors as input to classifiers to classify EEG signals in two or three classes.

6.2.1. Work Flow

To elucidate the proposed algorithm to design ECAD system, workflow is depicted in Figure 6.2. EEG signal is segmented into epochs whose duration is capable of capturing the characteristics of the event to be detected. Data set contains 100 data files of each class with 4096 data samples. Although free to choose the size of epoch, values were chosen according to the binary relation. In proposed method, we have considered size of subsamples n = 256, 512, 1024, 2048 and 4096 for each EEG data of a class corresponding to module size m = 16, 8, 4, 2 and1 matching to time resolution of 1.5, 3, 6, 12 and 23.6 sec. The segment size is decided so as to provide a balance between data segment long enough to provide good frequency resolution and short enough to satisfy the condition of stationarity.

Signal feature vectors comprising thirteen features as discussed in Chapter 3, are extracted from every module and statistically analyzed. The analysis result in RSFV framed out of prominent seven features (*mean, mode, standard deviation, skew, kurtosis, power and entropy*) from each sub-sample. For n sub samples or epochs, m modules are formed with seven features of every module resulting in vector set of size [n x m x 7]. The three different classifiers are engaged to design M_xCAD systems, where M_x denote the module of size x. To achieve this objective and to improve the classification performance of CAD systems, and to understand the underlying dynamics of nonstationarity of the signals, I have restricted myself to the comparison of distinct module size taken from same time series.

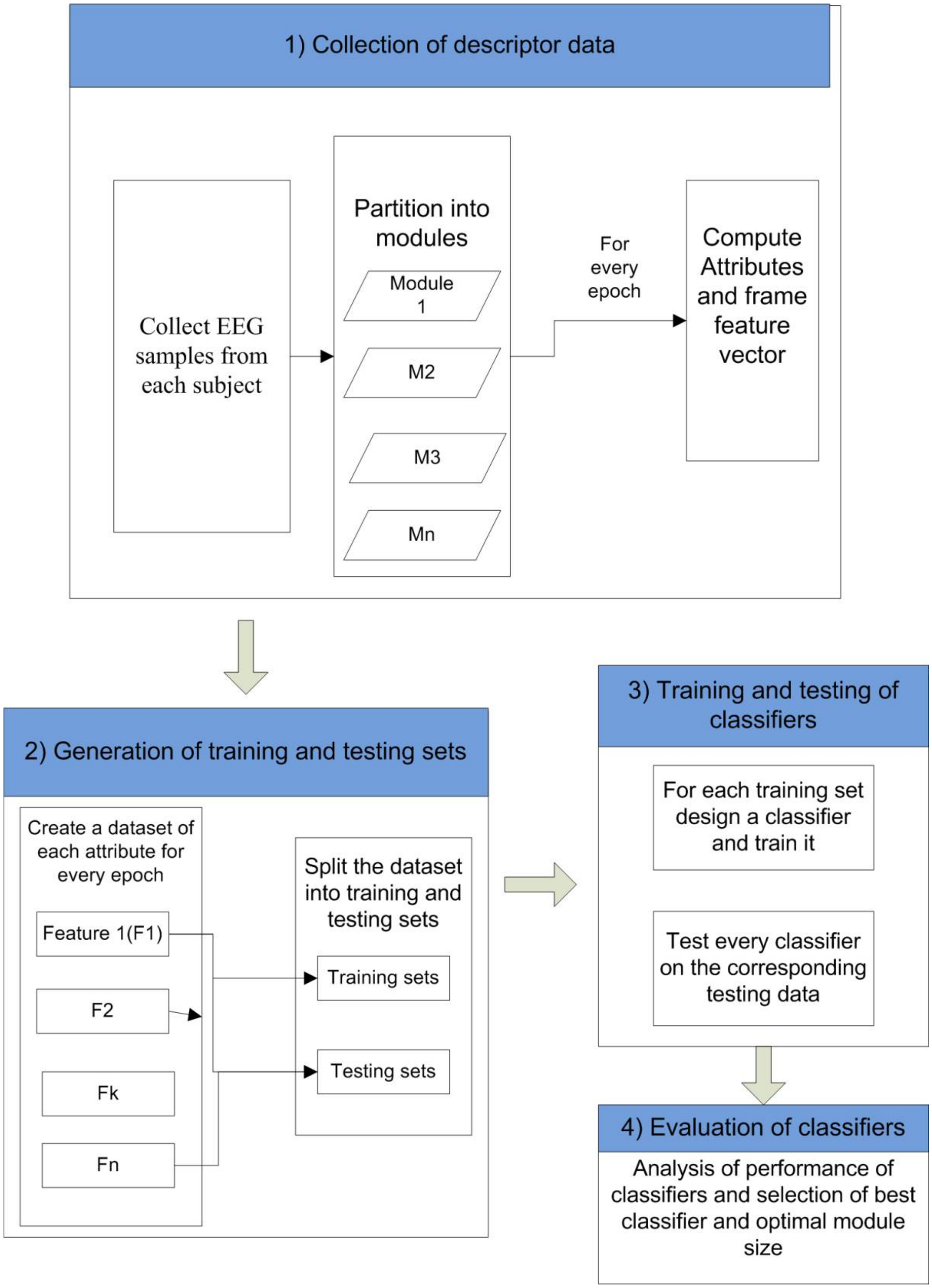


Figure 6.2. Work flow of the proposed classification strategy

To comply with the proposed methodology, we have utilized MLPNN, SVM, and KNN classifiers. The complete feature vector is divided into two sets of training and testing; the classifier is trained with the training vectors, whereas the testing vectors are used to find out classification accuracy and the effectiveness of the trained classifier for the classification of EEG signals [122]. Sensitivity, specificity and classification accuracy of the proposed method are calculated with every module size with bootstrap method. For the classification purpose it not clear what epoch size is most appropriate, therefore, a number of epoch size are investigated and overall system performance is subsequently evaluated. Henceforth, analysis of all of these sets of epochs will qualify whether seizure classification is better obtained with longer or shorter-duration epoch lengths.

6.3. DESIGN OF VARIOUS MCAD SYSTEMS

To investigate whether our method is applicable to a short time series and whether our method depends strongly on the length of data segment, the epoch size respectively takes the value of 256, 512, 1024 and 2048 samples, and the approaches used in the previous sections are employed for classifying EEG epileptic dataset into two or three classes depending upon the chosen module size. The complexity of the classification task increases exponentially with the dimension of the data, so it becomes worthwhile to extract the features and also limit the number of features to only those which are most discriminative for the classification task.

6.3.1. Performance analysis of CAD system with module size of 16 (M_{16}RSFV)

Starting with the largest module size, the complete dataset is divided into epochs of 256 samples. As there are 4096 samples in total for each set, epoch size of 256 samples (*n*) result in 16 modules (*m*) giving 1600 samples. Every module is represented by seven features, consequently resulting in feature vector set of size $[256x16x7]$ for 100 data files of one class. For two class classification problem, the feature vector set is employed for input to the classifiers for training with 75% of dataset and remaining 25% of samples for testing purpose with four chosen classifiers. Overall classification accuracy of the four classifiers is achieved with 5-cross validation and results obtained are tabulated in Table 6.1. It can be observed that classification accuracy obtained with MLPNN classifier is high as 89.25% as compared to OCA obtained with SVM (86.37%), KNN (73.37%), and RBF (79.87%). Sensitivity and specificity obtained with all the classifiers are represented in the table.

Table 6.2 describes the performance with three vector matrices created from data set of healthy subjects, subjects with epileptic activity and inters epileptic activity in order to classify three classes in the study. Out of the 4800 signals, obtained from SFV of module size of 16, 70 % of data samples are used to train the classifiers and remaining 30% of the samples are used to test the validity of the classifier. The OCA and ICA for the three cases with four different classifiers are reported and tabulated in Table 6.2.

Table 6.1: Performance of MCAD system of module size of 16, with four classifiers for two classes

Module (Epoch) Size	Classifier		Confusion Matrix		OCA	Sen	Spec
16(256)			**NOR**	**ICT**			
	SVM	**NOR**	330	70	86.37%	90.25%	82.50%
		ICT	39	361			
			NOR	**ICT**			
	KNN	**NOR**	273	127	73.37%	78.52%	68.25%
		ICT	86	314			
			NOR	**ICT**			
	MLPNN	**NOR**	338	62	86.87%	89.25%	84.5%
		ICT	43	357			
			NOR	**ICT**			
	RBF	**NOR**	310	90	79.87%	82.25%	77.5%
		ICT	71	329			

Note: *CM- Confusion Matrix, OCA – Overall Classification Accuracy,*

Table 6.2: Performance of MCAD system for three classes with module size of 16, epoch size of 256 samples.

Classifier	L(M_{16}SFV)		Confusion Matrix			ICA (%)	OCA (%)
			NOR	INT	ICT		
MLPNN	256*16*7	NOR	311	30	59	77.7	78.02
		INT	32	284	84	71.0	
		ICT	11	48	341	85.2	
			NOR	INT	ICT		
RBF	256*16*7	NOR	298	78	24	74.5	70.58
		INT	62	260	78	65.0	
		ICT	51	60	289	72.2	
			NOR	INT	ICT		
KNN	256*16*7	NOR	289	64	47	72.3	68.40
		INT	39	276	85	69.0	
		ICT	78	66	256	64.0	
			NOR	INT	ICT		
SVM	256*16*7	NOR	293	55	52	73.2	72.08
		INT	57	268	75	67.0	
		ICT	66	30	304	76.0	

Note: *OCA – Overall Classification Accuracy, ICA- Individual classification Accuracy.*

From the results obtained for three class classification problem, out of four classifiers, MLPNN gives the better overall classification accuracy of 78.02% with ICA of 85.25%, 77.7% and 71% for the three different classes. SVM provides 72.08% OCA followed by RBF giving 70.58% and KNN providing 68.40% overall classification accuracy. RBF and KNN do not provide good classification results as compared to SVM and MLPNN, continuing the trend of classification algorithm for two class classification problem.

6.3.2. Performance analysis of MCAD system with module size of 8(M_8CAD)

Following the technique discussed in the previous section, dataset is divided into epochs of 512 samples resulting in 8 modules giving 800 samples of each class. Every module when represented by seven features, result in feature vector of size [512x8x7] for 100 data files of data set. In all 1600 samples are used for training and testing of the various classifiers. The classification results obtained from MCAD system for two class problem are tabulated in Table 6.3. The classification accuracy results obtained with different classifiers do not vary much as compared to the results obtained with epoch size of 256. MLPNN gives 86.87% of OCA, followed by SVM with 84.58%, RBF with 79.75% and KNN classifier provides 73.54% of OCA. Accuracy obtained with this module size is comparable with module size of 16.

Table 6.3: Performance of MCAD system of module size of 8, with four classifiers for two classes

Module (Epoch) Size	Classifier	Confusion Matrix			Acc.	Sen	Spec
8(512)	**SVM**		**NOR**	**ICT**			
		NOR	198	42	84.58%	86.66%	82.5%
		ICT	32	208			
	KNN		**NOR**	**ICT**			
		NOR	178	61	73.54%	74.16%	72.9%
		ICT	65	175			
	MLPNN		**NOR**	**ICT**			
		NOR	203	37	86.87%	89.16%	84.58%
		ICT	26	214			
	RBF		**NOR**	**ICT**			
		NOR	198	42	79.75%	77.08%	82.5%
		ICT	55	185			

Following the previous steps for three class problem, the classifiers are trained with 70% of samples and tested and validated with 30% of dataset, with results depicted in Table 6.4. The highest accuracy is again obtained with MLPNN and the lowest is obtained with KNN. These observations follow the same trend as for module size of 16. The OCA obtained with MLPNN is 79.16% followed by SVM (78.61%), RBF (70.7%) and lastly by KNN (67.76%).

Table 6.4: Classification performance of MCAD system for three classes with module size of 8and epoch size of 512 samples.

Classifier	L (M_8SFV)		Confusion Matrix			ICA (%)	OCA (%)
			NOR	**INT**	**ICT**		
MLPNN	512*8*7	**NOR**	193	34	13	80.4	79.16
		INT	32	165	43	68.75	
		ICT	16	12	212	88.33	
			NOR	**INT**	**ICT**		
RBF	512*8*7	**NOR**	168	22	50	59.7	70.70
		INT	21	149	70	62.08	
		ICT	20	28	192	80.1	
			NOR	**INT**	**ICT**		
KNN	512*8*7	**NOR**	140	42	58	58.33	67.76
		INT	38	158	44	65.83	
		ICT	27	23	190	79.16	
			NOR	**INT**	**ICT**		
SVM	512*8*7	**NOR**	184	17	39	76.66	78.61
		INT	14	188	38	78.33	
		ICT	18	28	194	80.83	

6.3.3. Classification results of CAD system with module size of 4 (M_4CAD)

Following the proposed methodology, complete dataset is further divided into epochs of 1024 samples resulting in 4 modules giving 400 samples of each class. Every module result in feature vector set of size [1024x4x7] for 100 data files of data set. The results obtained are tabulated in Table 6.5 in terms of accuracy, sensitivity and specificity. The obtained results can be compared with the results obtained with different module size and with the same classifiers. Accuracy of 87.50% was obtained with SVM and MLPNN classifier whereas classification accuracy obtained with KNN is 76.6% and 84.16% with RBF following the similar trend.

Table 6.5: Performance of MCAD system of module size of 4, with four classifiers for two classes.

Module (Epoch) Size	Classifier	Confusion Matrix			Acc.	Sen	Spec
4(1024)	**SVM**		**NOR**	**ICT**	90.41%	93.33%	87.5%
		NOR	105	15			
		ICT	8	112			
	KNN		**NOR**	**ICT**	79.58%	82.5%	76.6%
		NOR	92	28			
		ICT	21	99			
	MLPNN		**NOR**	**ICT**	89.5%	91.66%	87.5%
		NOR	105	15			
		ICT	10	110			
	RBF		**NOR**	**ICT**	84.98%	85.8%	84.16%
		NOR	101	19			
		ICT	17	103			

Note: *NOR-Normal, ICT- Ictal, OCA –Overall Classification Accuracy, Sen- Sensitivity, Spec-Specificity*

MCAD systems with module size of 4 were tested for three class problem, whose results are tabulated in Table 6.6. The highest accuracy obtained is with MLPNN and the lowest accuracy is obtained with KNN. These observations follow the same trend as for module size of 16 and 8. The OCA obtained with MLPNN is 86.43% followed by SVM (80.26%), RBF (79.73%) and lastly by KNN (75.25%). The ICA for ictal and normal conditions are more as compared to the interictal class complying with the previous results.

Table 6.6: Classification performance of MCAD system for three classes with module size of 4and epoch size of 1024 samples.

Classifier	L (M_4SFV)	Confusion Matrix	NOR	INT	ICT	ICA (%)	OCA (%)
MLPNN	1024*4*7	NOR	102	12	6	85.00	86.43
		INT	9	100	11	83.53	
		ICT	4	7	109	90.83	
RBF	1024*4*7	NOR	97	12	11	80.08	79.73
		INT	9	92	19	76.6	
		ICT	13	9	98	81.6	
KNN	1024*4*7	NOR	79	19	22	65.83	75.25
		INT	11	94	15	78.33	
		ICT	8	14	98	81.66	
SVM	1024*4*7	NOR	98	9	13	81.66	80.26
		INT	21	91	8	75.80	
		ICT	8	14	98	83.33	

6.3.4. Classification results of CAD system with module size of 2 (M_2CAD)

Lastly, the same steps were followed for the design of MCAD system with epochs of 2048 samples resulting in 2 sets of individual modules. Every module represented by seven features, result in feature vector set of size $[2048x2x7]$. The results obtained with the same set of classifiers are tabulated in Table 6.7. Accuracy of 91.66% was obtained with MLPNN and 90.83% with SVM classifier and 87.5% of classification accuracy is obtained with KNN. The obtained results can be compared with the results obtained with different module size but with same classifiers in the coming section.

Table 6.7: Performance of MCAD system of module size of 2, with four classifiers for two classes

Module (Epoch) Size	Classifier		Confusion Matrix		Acc.	Sen	Spec
2(2048)			NOR	ICT			
	SVM	NOR	52	8	90.83%	95.0%	86.66%
		ICT	3	57			
			NOR	ICT			
	KNN	NOR	50	10	87.50%	91.66%	83.33%
		ICT	5	55			
			NOR	ICT			
	MLPNN	NOR	53	7	91.66%	95.0%	88.33%
		ICT	3	57			
			NOR	ICT			
	RBF	NOR	47	13	85.08%	91.66%	78.33%
		ICT	5	55			

Classification results of MCAD system with two modules for three class classification are tabulated in Table 6.8. The highest accuracy obtained is with MLPNN and the lowest accuracy is obtained with KNN following the same trend of results as followed by the classifier with different module size. The OCA obtained with MLPNN is 98.33% followed by SVM (90.1%), RBF (88.63%) and lastly by KNN (84.45%). The OCA has shown an increasing trend with the increase of epoch size.

6.3.5. Classification results of CAD system with module size of 1 (M_1CAD)

Finally, the epochs of 4096 samples are framed, resulting in single module. The results obtained with these features for two classes and three class classifications with the complete dataset as a whole are tabulated in Table 6.9. For two classes 100% classification accuracy is obtained with 100% sensitivity and specificity and 0% false alarm rate.SVM also provides good results with 98.3% of accuracy followed by RBF with 96.6% and KNN exhibiting 93.3%.

Table 6.8: Classification performance of MCAD system for three classes with module size of 2and epoch size of 2048 samples

Classifier	L(M_2SFV)	Confusion Matrix				ICA (%)	OCA (%)
			NOR	**INT**	**ICT**		
MLPNN	2048*2*7	**NOR**	128	73	14	98.33	98.33
		INT	90	142	11	96.66	
		ICT	40	29	242	100.0	
			NOR	**INT**	**ICT**		
RBF	2048*2*7	**NOR**	54	5	1	90.0	88.63
		INT	0	55	5	91.66	
		ICT	2	7	51	85.01	
			NOR	**INT**	**ICT**		
KNN	2048*2*7	**NOR**	52	5	3	86.66	84.45
		INT	11	47	2	78.33	
		ICT	1	6	53	88.33	
			NOR	**INT**	**ICT**		
SVM	2048*2*7	**NOR**	52	7	1	86.66	90.1
		INT	2	52	6	86.66	
		ICT	0	2	58	96.66	

Presented empirical evidence suggests that by simply considering multiple samples in time, different level of accuracy can be achieved for the classification decision. Temporal structure of EEG yields variation in the classification rate. Accuracy is the mean value of sensitivity and specificity and is included here as a single-figure indication of classification performance [126]. Specificity and sensitivity results from all the MCAD system indicate that the feature set is useful in separating between normal and abnormal EEG. Classification performance for all the CAD systems are compared in the following section and best suited epoch lengths is determined based on performance metrics.

Table6.9: Performance of MCAD system of module size of 1, with four classifiers for two classes.

Module (Epoch) Size	Classifier	CM			Acc.	Sen	Spec	Miss rate
1(4096)	SVM		**NOR**	**ICT**	98.3%	100%	96.66%	1.66
		NOR	29	1				
		ICT	0	30				
	KNN		**NOR**	**ICT**	93.3%	97.0%	90.0%	6.6
		NOR	27	3				
		ICT	1	29				
	MLPNN		**NOR**	**ICT**	100%	100%	100%	0.0
		NOR	30	0				
		ICT	0	30				
	RBF		**NOR**	**ICT**	96.66%	96.66%	96.66%	1.66
		NOR	29	1				
		ICT	1	29				

6.4. RESULTS AND DISCUSSIONS

6.4.1. Two class classification problem

The results obtained for classification of two classes discussed in the previous sections are collectively represented in Figure 6.3. As a comparative study, we have evaluated the performance of the approach with the same attributes that summarizes entire set of dataset. It is observed that SFV of length seven with module size of 4096 yields highest OCA of 100% with MLPNN and 98.3% with SVM classifier. OCA obtained with different epoch size of 256, 512, 1024 and 2048 module size with same SFV with different classifiers are depicted in Figure 6.3. An interesting fact is that by use of optimal module size, sensitivity and specificity values for ictal and normal cases increases. From the results obtained, it is observed that classification accuracies increase with increase in length of 'n' and best performance is achieved with sample size of n = 4096. Increasing the modules overload the classifiers, that affect the classification time and accuracy.

	MLPNN	SVM	RBF	KNN
256	86.87	86.37	79.87	73.37
512	86.87	84.58	79.75	73.54
1024	89.5	90.41	84.98	79.58
2048	91.66	90.83	85.03	87.5
4096	100	98.3	96.66	93.3

Figure 6.3. Comparative performance analysis of various classifiers with different module sizes.

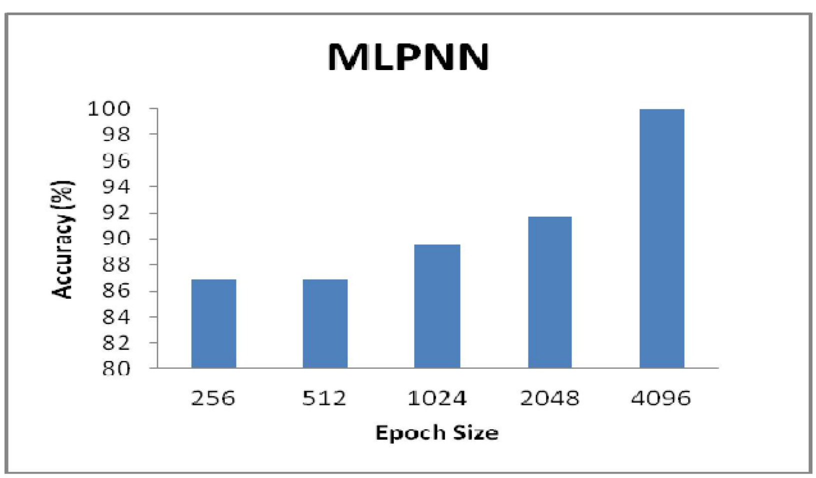


Figure 6.4. Classification Performance of MLPNN for different epoch size for two class classification.

All the classifiers show the similar trend of increase in accuracy with increase in epoch size. Nonetheless, MLPNN classifier outperforms the other algorithms and capable of providing highest OCA for almost all module sizes as depicted in Figure 6.4.

6.4.2. Three class classification problem

The results tabulated in the previous tables for three class classification problem are plotted in Figures 6.5 for comparative study. The comparison is based on the size of modules for various classifiers employed for the classification. The module size of 16 is represented by m5, and sizes of 8, 4, 2, and 1 corresponding to m4, m3, m2 and m1 in Figure 6.5.

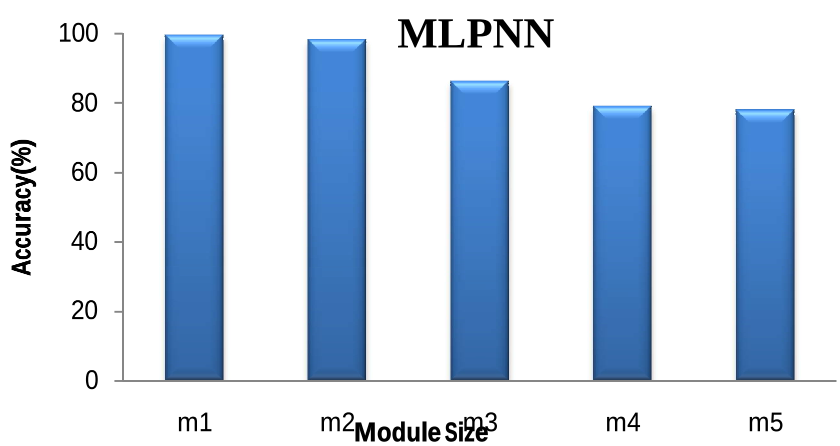


Figure 6.5. (*a*) Classification accuracies achieved with various module sizes with MLPNN classifier

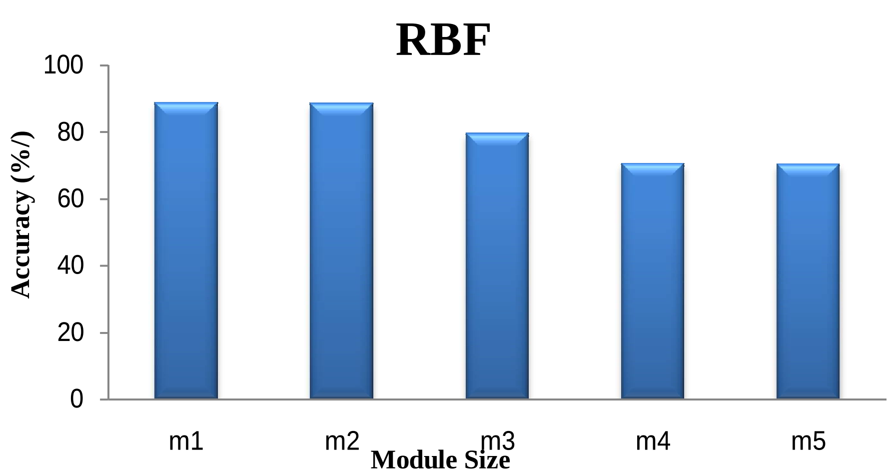


Figure 6.5. (*b*) Classification accuracies achieved with various module sizes with RBF classifier

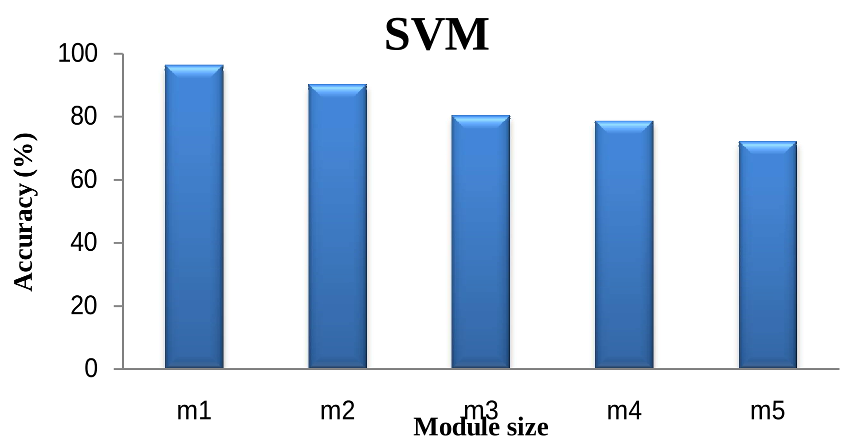


Figure 6.5. (*c*) Classification accuracies achieved with various module sizes with SVM classifier

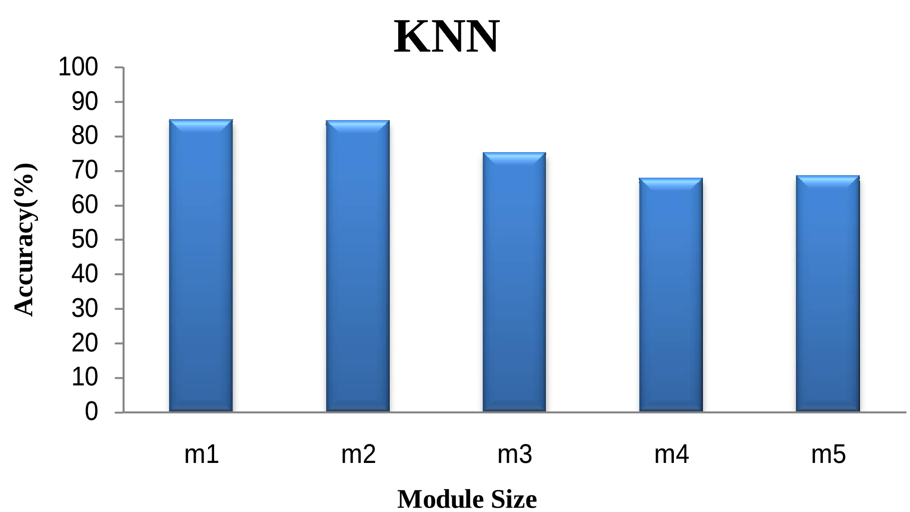


Figure 6.5 (*d*) Classification accuracies achieved with various module sizes with KNN classifier

The proposed technique is evaluated through four classification problems, and the highest classification accuracies obtained for all module sizes is achieved by MLPNN followed by SVM. For highest epoch size, accuracy of these both classifiers is comparable. It is apparent that with increase in number of modules, classification accuracies for all the four classifiers decrease. The minimum performance metrics are achieved for larger size of modules and this trend is uniform for all the four classifiers. Based on these results and observations, important criteria of choosing efficient technique are accuracy and computational efficiency, which are used in this research work for assessment.

6.5. CONCLUSIONS

In this objective, we intended to design a CAD system for prognosis of epileptic seizure with less computational complexity and higher accuracy. The proposed Module based CAD method is employed for distinguishing normal and ictal signals and for classification of sets S, F and Z. The statistical approach is employed for the feature extraction and different classifiers are used for the classification of EEG signals. A practical issue in seizure detection is to determine the length of the signal and number of modules optimal for successful classification. In general, selection of features and module size ensure the best accuracy that can be achieved in practice. The longer the signal, better are chances to be discriminative as more is the information content in it. Moreover, abnormalities of background activity and slow activity in some regions, are much less evident in smaller segments and cannot be easily appreciated by looking at the EEG. A one and half second epoch is long enough to detect any significant changes but short enough to avoid redundancy in the signal and suitable for prognosis of EEG signals. That is why long EEG recordings require the use of reliable and accurate computer programs that are able to extract the hidden information, so that a better diagnosis can be provided. This assertion constitutes the main aim of this objective.

From the results obtained, it is observed that classification accuracies increase with less number of modules and with increase in length of 'n' and best performance is achieved with sample size of n= 4096. The results obtained from MLPNN and SVM are quite encouraging and comparatively higher as compared to other classifiers in all aspects. With 100% OCA, high sensitivity and specificity with selected feature set, we infer that the chosen features and MLPNN classifier are more promising than any other proposed method. We conclude from the demonstrated results that the proposed technique can be used for epileptic seizure detection.

CHAPTER 7

CONCLUSION AND FUTURE WORK

7.1. CONCLUSION

EEG is an important measurement of brain activity and has great potential in helping in the diagnosis and treatment of neurophysiologic diseases and abnormalities. Automated CAD systems cannot, and are not intended to, fully replace neurologists, nonetheless, the primary aim is to support the evaluation of medical doctors and to take the decisions accurately and efficiently. In this research work, an attempt has been taken to develop a robust prediction model for seizure detection which may assist radiologists in making differential diagnosis and provide better prognosis by providing second opinion in case of voluminous and highly complex brain signals.

The work in this thesis has approached the problem of designing of CAD system by considering the relevant features, morphology of brain signals, supervised machine learning methods and incorporating the statsitcal analysis in the procedure of classification of neurophysiological signals. The research work presents a new approach to the use of algorithm in this sense. The study presents a different analysis, both in the identification of the most effective features used for the representation of the problem and also in the identification of the most efficient algorithm in the classification of the problem from among the 6 most popular classification algorithms. This work includes inspection of a comprehensive feature set indicating their usefulness in the representation of EEG signals and to find out, what features delineates the EEG signals of a normal patient from an epileptic patient. The novelty of this objective lies in the exhaustive statistical analysis of extracted features to come up with prominent feature set for classification purposes and to the author's knowledge is the first of its kind. Various statistical tests are applied exhaustively to verify that the extracted features are distinct and uncorrelated to each other. The chosen features are simple but robust for the morphology of EEG data for the classification problem.

The ambit of this thesis is design of different CAD systems for classification of seizure activity with highest accuracy, sensitivity and specificity. Strategically, a fully automated neural network model, capable of classifying the seizure activity into ictal, interictal and normal state with classification accuracy as high as 99.3% has been designed. The promising performances observed are demonstrative of the efficiency and efficacy of systems developed for classification and prediction of normal, ictal and interictal conditions of epileptic patients. A bi-class classification design is also incorporated to show generalizability of soft computing paradigm technique. A novel hierarchical CAD system is modeled for seizure classification. We conclude that EEG based information is sufficient for the proposed method that has higher potential in designing EEG based diagnostic system for detection of electroencephalographic changes.

Further non-linear analysis method is efficasiously applied to EEG signals to study the non linear characteristics in the signals and to enhance the classification performance. The inclusion of non linear features in designing ECAD systems improved classifier performance prominently. It was found that the fusion of signals at both feature and classifier levels improved detection and classification leading to good separation between normal, ictal and interictal EEG signals. In this research work, a classification accuracy of 100% was achieved on EEG epileptic database for healthy subjects and epileptic patients. The experimental results demonstrated that the ensemble of features were good choice for capturing representative characteristics of EEG signals and MLPNN and SVM were promising machine learning methods for the classification of these signals. Improvements in ROC area was observed for ECAD system when compared to the normal CAD systems.

The final part of this thesis demonstrates usage of epochs to assess the non stationarity of the signals. Selection of features and module size ensure the best accuracy that can be achieved in practice. A one and half second epoch is long enough to detect any significant changes but short enough to avoid redundancy in the signal and suitable for prognosis of EEG signals.

Using both epoch and classifier based metrics, it was clear that the use of MLPNN classifiers improved the performance of the automated epileptiform detection system. This type of analysis also contributes to the better understanding of signals nature and characteristics. These types of studies may be instrumental in finding effective solutions to the question “which feature algorithm should be used to acquire the feature that can best represent the data?”

Table 7.1 summarizes all the proposed CAD system in this thesis with the performances. The research findings indicate that this proposed approach can distinguish the categories of EEG signals in two and three-class very efficiently. Taken together, on the basis of our results and our cooperation with medical doctors, it can be concluded that the research presented in this dissertation has found successful methods for the reliable classification of EEG signals and can facilitate more effective evaluation of the complex biomedical signals.

Table 7.1: Proposed CAD Systems with their performance

S.No.	CAD System	Classes	OCA (%)
1.	NN based CAD	S-F-Z S-Z	99.3 98.5
2.	SVM based CAD	S-Z	98.5
3.	SVM based hierarchical CAD	S-F-Z	96.6
4.	SVM based Ensemble CAD NN based Ensemble CAD	S-Z	100
5.	NN based Ensemble CAD	S-F-Z	98.33
6.	NN based MCAD system	S-F-Z	98.03
7.	NN based MCAD system	S-Z	100

Note: OCA- Overall Classification System, S - Ictal class, Z - Normal class, F - interictal class.

7.2. FUTURE WORK

It is believed that the methods presented in this thesis would provide promising outcomes in the EEG signal classification area. Nevertheless, there are many research directions to explore; we intend to examine the possibility of using the proposed technique in the application of other biomedical signals. Moreover, it is, also highly promising idea to apply these methods to other neurophysiologic paradigms.

Extensive future work will be directed in extending the algorithm for multi-channel EEG signal classification and would strive for the robustification of CAD systems. The performance of the classifier is to be augmented by adding more patients and creation of a larger database in cooperation with relevant medical institutions. Further on, the developed and tested methodology should be adjusted and modified to make it able to process EEG signal in the real time, for continuous non-invasive monitoring of brain activity. Exhaustive future work will be oriented towards development of suitable algorithms that can predict the onset of the seizures.

A graphical user interface for the model is to be developed that shall increase the applicability of this algorithm. Extensive future work will examine the possibility of dealing with artifacts effectively; successfully removing artifacts to achieve significant improvement in proposed algorithms for diagnostic systems. Lastly, future research work will be eyeing on integrating modalities that provide good time resolution (EEG and EMG) and good frequency resolution (fMRI and PET), as these will be able to capture dynamic evolution of time varying connectivity patterns.

REFERRED JOURNAL PUBLICATIONS

1. Meenakshi Sood, and Sunil V. Bhooshan, "Parameter-selective Based CAD System For Epileptic Seizure Classification" *International Journal of Applied Engineering Research*, vol 10(12), pp. 28751-28769, 2015. ***(Scopus Indexed)***

2. Meenakshi Sood and Sunil V. Bhooshan, "Design and Development of Prediction Model to Detect Seizure Activity Utilizing Higher Order Statistical Features of EEG signals" *Res. J. Pharm., Biol. Chem. Sci.* vol.5(3), pp.1129-1145, 2014. ***(Scopus Indexed)***

3. Meenakshi Sood,Vinay kumar, S.V.Bhooshan, "Comparison of machine learning methods for prediction of epilepsy by neurophysiological EEG signals" *Int J Pharm Bio Sci*, vol.5 (2 B), pp. 6 – 15, 2014. ***(Scopus Indexed)***

4. Meenakshi Sood, Vinay Kumar, Sunil V. Bhooshan," Review of State of Art in Electrooculogram Artifact Removal from Electroencephalogram Signals," *International Journal of Enhanced Research in Science Technology & Engineering, vol.*2(4), pp.32-41, April-2013.

5. Meenakshi Sood, Sunil V. Bhooshan," A Novel Module based Approach to Automated Epileptic Diagnosis CAD system using EEG ", *Biomedical Signal Processing and Control* ***(Communicated)***

REFERRED CONFERENCE PUBLICATIONS

6. Meenakshi Sood and Sunil V. Bhooshan, "Hierarchical Computer Aided Diagnostic System for Seizure Classification", Proceedings of the 9th INDIACom; INDIACom-2015; IEEE

Conference ID: 35071, 2nd IEEE International Conference on “Computing for Sustainable Global Development”, 11^{th} – 13^{th} March, 2015, (BVICAM), New Delhi (INDIA), pp.6.315-6.320.

7. Meenakshi Sood, and Sunil V. Bhooshan, “Automatic Processing of EEG signals for Seizure Detection using Soft Computing Techniques”, IEEE International Conference on Recent Advances and Innovations in Engineering (ICRAIE-2014), May 09-11, 2014, Jaipur India, pp 1-6. ***(Scopus indexed)***

8. Meenakshi Sood and Sunil V. Bhooshan, “Modular Based Dynamic Analysis of EEG Signals Using Non-Linear Feature”, IEEE 3rd International Conference on Parallel, Distributed and Grid Computing 11-13 Dec 2014, JUIT Waknaghat, Solan. pp. 186-190. DOI: 10.1109/PDGC.2014.7030739 ***(Scopus indexed)***

9. Meenakshi Sood, and Sunil V. Bhooshan, “ Prognosis of Epileptic seizures using EEG signals”, 10th INDIACom; 3rd 2016 International Conference On “Computing for Sustainable Global Development”.

(Communicated)

REFRENCES

[1] Atlas: Epilepsy Care in the World. World Health Organization, 2005.

[2] J.Gotman, “Automatic recognition of epileptic seizures in the EEG”, *Electroen Clin Neuro.,* vol.54, pp.530–540, 1982.

[3] H.Gastaut, "Dictionary of Epilepsy. Part I: Definitions", World Health Organization. 1973.

[4] C. Logar, B. Walzl, and H. Lechner, “Role of long-term EEG monitoring in diagnosis and treatment of epilepsy”, *European Neurology*, vol. 34, pp.29-30, 1994.

[5] J.C. Rajapakse, A. Cichocki and V.D. Sanchez, “Independent component analysis and beyond in brain imaging: EEG, MEG, fMRI, and PET”, In Proc. of IEEE Conf. on Neural Information Processing, vol.1, pp. 404-412, 2002

[6] S. Sanei and J.A. Chambers, *EEG signal processing*. Wiley-Interscience, 2007.

[7] J Gotman, "Automatic seizure detection: Improvements and evaluation", *Electroencephalogr. Clin. Neurophysiol*., vol.76, pp.317–324,1990.

[8] Merlet, I, Garcia-Larrea, et al., “Simplified projection of EEG dipole sources onto human brain anatomy”, *Neurophysiol Clin*., vol.29 (1), pp. 9-52, 1999.

[9] N. R. Carlson, “Structure and Functions of the Nervous System”, *Foundations of physiological psychology*, vol. 5(3), 2002b.

[10] E. Callaway, P.R. Harris, "Coupling between cortical potentials from different areas", *Science* vol.183 (4127), pp. 873–875, 1974.

[11] W. Chaovalitwongse, P. Pardalos, L. D. Iasemidis, D. S. Shiau, and J. C. Sackellares, “Dynamical Qualitative and Quantitative Evaluation of EEG Signals in Epileptic Seizure Recognition”, *MECS I.J. Intelligent Systems and Applications*, vol. 06, pp. 41-46, 2013.

[12] W.J.R. Dunseath, E. F. Kelly, "Multichannel PC-based data-acquisition system for high-resolution EEG", *IEEE Transactions on Biomedical Engineering*, vol.42 (12), pp.1212-1217, 1995.

[13] Ali Bulent Usakli, “Improvement of EEG Signal Acquisition: An Electrical Aspect for State of the Art of Front End”, *Computational Intelligence and Neuroscience,* vol. 2010, Article ID 630649, 7 pages, 2010.

[14] I. Ulbert, E. Halgren, G, Heit, G. Karmos, " Multiple Microelectrode-recording System for Human Intracortical Applications, *Journal of Neuroscience Methods*, and vol.106 (1): pp.69-79, 2001.

[15] Teplan, “Fundamentals of EEG Measurement,” *Measurement Science Review*, vol. 2, pp.1-11, 2002.

[16] Andrzejak, K. Lehnertz, F. Mormann, C.Rieke, P. David, and C. Elger, “Indications of nonlinear deterministic and finite dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state”, *Phys. Rev. E.*, vol.4 (6), 2001.

[17] R. Palaniappan, “Brain computer interface design using band powers extracted during mental tasks”, In Proceedings of the 2nd International IEEE EMBS Conference on Neural Engineering, 2005.

[18] U.R. Acharya, Sree S. Vinitha, G. Swapna, R. Joy Martis, J.S.Suri., “Automated EEG analysis of epilepsy: A review”, *Knowledge based systems*, vol.45, pp.147–165, 2013.

[19] Kannathal, U. R. Acharya, C. M. Lim, and P. K. Sadasivan, “Characterization of EEG—A comparative study”, *Computer Methods and Programs in Biomedicine*, vol. 80(1), pp. 17– 23, 2005.

[20] J. Arnhold, K. Lehnertz, P. Grassberger, C.E. Elger, “A robust method for detecting interdependences: application to intracranially recorded EEG”, *Physica D,* vol.134 (4), pp.419–430, 1999.

[21] Epilepsy: Mechanisms, Models, and Translational Perspectives edited by Jong Rho, Raman Sankar, Carl E. Stafstrom.

[22] Fabien Lotte, Marco Congedo, Anatole L´ecuyer, Fabrice Lamarche, Bruno Arnaldi, “A review of classification algorithms for EEG-based brain–computer interfaces", *Journal of Neural Engineering,* Institute of Physics: Hybrid Open Access, vol. 4, pp.24. 2007.

[23] N. Friedman, D. Geiger, and M. Goldszmidt, “Bayesian network classifiers: Machine Learning”, vol.29, pp.131–163, 1997.

[24] H. Adeli, S.L.Hung, "Machine Learning – Neural Networks, Genetic Algorithms, and Fuzzy Sets", New York: Wiley, 1995.

[25] Claude Roberta, Jean-Franc¸ois Gaudyb,and Aime Limogea, “Electroencephalogram processing using neural networks”, *Clinical Neurophysiology*, vol.113, pp.694–701, 2002.

[26] S. Haykin, "Neural networks: A comprehensive foundation”, New York: Macmillan, 1994.

[27] Dr.R. Sukanesh, and R.Harikumar, “A Patient Specific Neural Networks (MLP) for Optimization of Fuzzy Outputs in Classification of Epilepsy Risk Levels from EEG Signals”, Engineering *Letters*, vol. 13(2), pp.50-56, 2006.

[28] C Lee, P Chung, J Tsai, and C Chang, “Robust radial basis function neural networks”, *IEEE Trans. Syst. Man Cybern.,Part B, Cybern*., vol.29, pp.674–685, 1999.

[29] K. Z. Mao, K.-C. Tan, and W. Ser, “Probabilistic Neural- Network Structure Determination for Pattern Classification”, IEEE *Transactions on neural networks*, vol. 1(4), pp. 1009-1016, 2000.

[30] A Lima, AI. Coelho, and M Eisencraft, “Tackling EEG signal classification with least squares support vector machines: a sensitivity analysis study”, *Comput. BioI. Med*, vol 40(8) pp, 705-714, 2010.

[31] C. J. Burgess, "A tutorial on support vector machines for pattern recognition. Data Min”, *Knowl Discov*, vol., pp.21–47, 1998.

[32] U. Orhan, M. Hekim, and M. Ozer , “EEG signals classification using the K-means clustering and a multilayer perceptron neural network model”, *Expert Systems with Applications.*, vol.38, pp.13475–13481, 2011.

[33] J.Virmani, V. Kumar, N. Kalra and N. Khandelwal, "Characterization of primary and secondary malignant liver lesions from B-mode ultrasound”, *Journal of Digital Imaging*, vol. 26(6), pp.1058-1070, 2013.

[34] Tom Fawcett, “An introduction to ROC analysis” *Pattern Recognition Letters*, vol.27, pp.861–874, 2006.

[35] I. Osorio, M.G. Frei, S.B. Wilkinson, “Real-time automated detection and quantitative analysis of seizures and short-term prediction of clinical onset”, *Epilepsia,* vol.39 (6), pp.615–627, 1998.

[36] R Tetzlaff, C Niederhofer, P. Fischer, “Feature extraction in epilepsy using a cellular neural network based device - first results”, *Circuits and Systems*, Proceedings of the 2003 International Symposium, ISCAS '03, vol.3, pp. III-850 - III-85, 2003.

[37] H. Jokeit and M. Schacher, “Neuropsychological aspects of type of epilepsy and etiological factors in adults”, *Epilepsy & Behavior*, vol.5, pp.14–20, 2004.

[38] E. Hoffmann, “Brain Training against Stress: Theory, Methods and Results from an Outcome Study”, version 4.2., 2005.

[39] T.P. Jung, S. Makeig, M. Westerfield, J. Townsend, E. Courchesne, and T.J. Sejnowski., “Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects”, *Clinical Neurophysiology*, vol.111(10), pp.1745–1758, 2000.

[40] S. Romero, M. A. Mananas, and M. J. Barbanoj, “Ocular Reduction in EEG Signals based on Adaptive Filtering, Regression and Blind Source Separation,” *Ann Biomed. Eng*., vol. 37(1), pp.176–91, 2009.

[41] P. He, G. Wilson, and C. Russell, "Removal of ocular artifacts from electro-encephalogramy adaptive filtering", *Medical and Biological Engineering and Computing*, vol.42(3): pp.407–412,2004.

[42] R. Flink, B. Pedersen, A. B. Guekht, K. Malmgren, R. Michelucci,B. Neville, F. Pinto, U. Stephani, and C. Ozkara, “Guidelines for the use of EEG methodology in the diagnosis of epilepsy international league against epilepsy”, *Acta Neurologica Scandinavica*, vol. 106, 2002

[43] S. O’Regan, S. Faul, and W. Marnane, "Automatic detection of EEG artefacts arising from head movements using EEG and gyroscope signals", *Medical Engineering & Physics*, vol.35 (7), pp.867–874, 2013a.

[44] M. Schröder, T. N. Lal, T. Hinterberger, M. Bodgan, N.J. Hill, and N. Birbaumer, “ Robust EEG channel Selection across Subjects for Brain-computer Interfaces,” *EURASIP Journal on Applied Signal Processing*, vol. 19, pp.3103–3112, 2005.

[45] H. Zeng, A. Song, R. Yan, and H. Qin, “EOG artifact correction from EEG recording using stationary subspace analysis and empirical mode decomposition,” Sensors, vol. 13, no. 11, pp. 14839–14859, 2013.

[46] Md Kafiul Islama, Amir Rastegarniaa, Anh Tuan Nguyena, Zhi Yang, "Artifact characterization and removal for in vivo neural recording", *Journal of Neuroscience Methods*, vol.226, pp.–123, 2014.

[47] Hamed Taheri Gorji, Abbas Koohpayezadeh, Javad Haddadn, " Ocular Artifact Detection and Removing from EEG by wavelet families: A Comparative Study " *Journal of Information Engineering and Applications*, vol. 3(13), 2013.

[48] P. K. Sadasivan., & D. N. Dutt, "SVD based technique for noise reduction in electro-encephalographic signals", *Signal Processing*, vol. 55, pp.179–189, 1996.

[49] T. D. Lagerlund, F. W Sharbourgh, and N.E. Busacker, “Spatial filtering of Multichannel Electroencephalographic Recordings through Principal Component Analysis by Singular Value Decomposition,” *J. Clin. Neurophysiol*, vol. 14, pp.73-82, 1997.

[50] P. Berg and M. Scherg., "A fast method for forward computation of multiple-shell spherical head models", *Electroencephalography and Clinical Neurophysiology*, vol. 90(1), pp.58–64, 1994.

[51] R. J. Croft and R. J. Barry, "Removal of ocular artifact from the EEG: a review", *Clinical Neurophysiology*, vol.30 (1), pp.5–19, 2000.

[52] M.T.Akhtar, W.Mitsuhashi, James J. Christopher, "Employing spatially constrained ICA and wavelet denoising, for automatic removal of artifacts from multichannel EEG data", *Signal Processing*, *vol.92*(2), pp.401-416, 2012.

[53] P. J. Durka, H. Klekowicz, K. J. Blinowska, W. Szelenberger, and Sz. Niemcewicz, "A Simple System for Detection of EEG Artifacts in Polysomnographic Recordings," *IEEE Transactions on Biomedical Engineering*, vol. 50(4), pp.527, 2003.

[54] A. Delorme, T. J. Sejnowski, and S. Makeig, "Enhanced Detection of Artifacts in EEG Data using Higher-order Statistics and Independent Component Analysis," *NeuroImage*, vol.34, pp.1443-1449, 2007.

[55] P. M. Quilter, B. B. Mcgillivray, and D. G. Wadbrook. "The Removal of Eye Movement Artifact from the EEG signals using Correlation Techniques", Random Signal Analysis, *IEEE Conference Publication*, vol.159, pp.93-100.

[56] R. Verleger, T. Gasser, and J. Möcks. "Correction of EOG artifacts in Event-related Potentials of EEG: Aspects of Reliability and Validity", *Psychophysiology*, vol.19, pp.472–480, 1982.

[57] G. Gratton. "Dealing with Artifacts: The EGG Contamination of the Event-related Brain Potential". Behavior Research Methods, Instruments, & Computer, University of Missouri, Columbia, Missouri vol.30 (1), pp. 44-53, 1998.

[58] O. G. Lins, T. W. Picton, P. Berg, and M. Scherg. "Ocular Artifacts in Recording EEGs and Event-related Potentials: II Source Dipoles and Source Components", *Brain Topography*, vol.6 (1), pp.65–78, 1993.

[59] T. D. Lagerlund, F. W. Sharbrough, and N. E. Busacker. "Spatial Filtering of Multichannel Electroencephalographic Recordings through Principal Component Analysis by Singular Value Decomposition", *Journal of Clinical Neurophysiology*, vol.14 (1), pp.73–82, 1997.

[60] S. Makeig, Tzyy-Ping Jung, Anthony J. Bell, and Terrence J. Sejnowski. "Independent Component Analysis of Electroencephalographic data", *Advances in Neural Information Processing Systems*, MIT Press, Cambridge MA, vol.8, pp.145-151, 1996.

[61] A. Bell and T. Sejnowski, "An Information-maximization Approach to Blind Separation and Blind Deconvolution", *Neural Computation*, vol.7, pp.1129–1159, 1995.

[62] A. Hyvärinen and E. Oja, "A Fast Fixed-point Algorithm for Independent Component Analysis", *Neural Computation*, vol.9, pp.483–492, 1997.

[63] R. Everson and S. J. Roberts, "Independent Component Analysis". Artificial Neural Networks in Biomedicine, Springer-Verlag, New York, pp.153-168, 2000.

[64] Tzyy-Ping Jung, Scott Makeig, "Removing electroencephalographic artifacts by blind source separation", *Psychophysiology*, vol.37, pp.163–178, 2001.

[65] Ille, Nicole, Berg, Patrick, Scherg, and Michael, "Artifact Correction of the Ongoing EEG Using Spatial Filters Based on Artifact and Brain Signal Topographies", *Journal of Clinical Neurophysiology*, vol.19 (2), pp.113-124, 2002.

[66] M. Izzetoglu, A. Devaraj, S. Bunce, and B.Onaral, "Motion artifact cancellation in NIR spectroscopy using Wiener filtering", *IEEE Trans. Biomed. Eng*, vol. 52(5), pp.934–938, 2005.

[67] D. Fox, J. Hightower, L. Liao, D. Schulz, and G. Borriello, "Bayesian filtering for location estimation", *IEEE Pervas. Comput*., vol.2(3), pp.24–33, 2003.

[68] C. A. Joyce, Irina F. Gorodnitsky and Marta Kutas, "Automatic Removal of Eye Movement and Blink Artifacts from EEG data using Blind Component Separation", *Psychophysiology*, vol.41 (2), pp.313-325, 2003.

[69] V. Krishnaveni, S.Jayaraman, P. M. Kumar, K. Shivakumar, and K. Ramadoss, "Comparison of Independent Component Analysis Algorithms for Removal of Ocular Artifacts from Electroencephalogram", *Measurement Science Review*, vol.5 (2), 2005.

[70] V. Krishnaveni, S. Jayaraman, S. Aravind, V. Hariharasudhan, and K. Ramadoss, "Automatic Identification and Removal of Ocular Artifacts from EEG using Wavelet Transform", *Measurement Science Review*, vol. 6(2), 2006.

[71] A. Garcés Correa , E. Laciar, H. D. Patiño, and M. E. Valentinuzzi, "Artifact Removal from EEG Signals using Adaptive Filters in Cascade", 16th Argentine *Bioengineering Congress* and the 5th Conference of Clinical Engineering, Conference Series, pp. 90, 2007.

[72] K. S. Vijilal, P. Kanagasabapathy, S. Johnson' and V. Ewards, "Artifacts Removal in EEG Signal using Adaptive Neuro Fuzzy Inference System", *IEEE* - ICSCN 2007, MIT Campus, Anna University, Chennai, India, pp.589-591, 2007.

[73] M. Ali Akber Dewan, M. J.Hossain, and Md. Moshiul Hoque, "Contaminated ECG Artifact Detection and Elimination from EEG using Energy Function

based Transformation", International Conference on Information and Communication Technology ICICT 2007, Bangladesh, pp. 52-56, 2007.

[74] S. Devuyst, T. Dutoit, P. Stenuit, M. Kerkhofs, and E. Stanus, "Removal of ECG Artifacts from EEG using a Modified Independent Component Analysis Approach", *IEEE EMBS Conference*, Vancouver, British Columbia, Canada, pp. 5204-5207, 2008.

[75] D. L. Jones, S. Appadwedula, M. Berry, M. Haun, D. Moussa, and D.Sachs, "Adaptive Filtering: LMS Algorithm", *Tech. Rep.*, 2009.

[76] J. Gao, Pan Lin, P. Wang, and Y. Yang, "Online EMG Artifacts Removal from EEG Based on Blind Source Separation," *IEEE 2nd International Asia Conference on Informatics* in *Control Automation and Robotics*, pp.28-32, 2010.

[77] M. Arezki, A. Benallal, and D. Berkani, "Improvement of the simplified fast transversal filter type algorithm for adaptive filtering", *J. Comput. Sci*, vol. 5(5), pp.347–354, 2010.

[78] J. Dong, T. Wang, A. Zhang, H. Dai, "The Removal of Blink and Saccade Artifact in EEG recordings by Independent Component Analysis", *3rd IEEE International Conference on Biomedical Engineering and Informatics (BMEI 2010*), pp. 1071-1075, 2010.

[79] P. Ashok Babu and Dr.K. V. S. V. R. Prasad, "Removal of Ocular Artifacts from EEG Signals by Fast RLS Algorithm using Wavelet Transform", *International Journal of Computer Applications*, vol. 21(4), pp. 1-5, 2011.

[80] T. Sweeney, Hasan Ayaz, Tomas E. Ward, "A Methodology for Validating Artifact Removal Techniques for Physiological Signals", *IEEE Transactions on Information Technology in Biomedicine*, vol. 16(5), pp.918-927, 2012.

[81] Li Mingai, Yan Cui and Jinfu Yang, "Automatic Removal of Ocular Artifact from EEG with DWT and ICA Method", *Appl. Math. Inf. Sci.*, vol.7(2), pp.809-816, 2013.

[82] Xun Chen, Chen He, and Hu Peng, "Removal of Muscle Artifacts from Single-Channel EEG Based on Ensemble Empirical Mode Decomposition and Multiset Canonical Correlation Analysis", *Journal of Applied Mathematics,* vol.2014, Article ID 261347, 10 pages, 2014.

[83] Hong Zeng and Aiguo Song, "Removal of EOG Artifacts from EEG Recordings Using Stationary Subspace Analysis", *The Scientific World Journal*, vol. 2014, Article ID 259121, 9 pages,2014.

[84] T. Radüntz et al., "EEG artifact elimination by extraction of ICA-component features using image processing algorithms", *Journal of Neuroscience Methods* , vol.243, pp. 84–93, 2015.

[85] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection", *Journal of Machine Learning Research*, vol. 3, pp.1157–1182, 2003.

[86] J.G. Dy and C.E. Brodley, "Feature selection for unsupervised learning", *Machine Learning Research*, vol.5, pp.845–889, 2004.

[87] R. Kohavi and G. H. John, "Wrappers for feature subset selection", *Artificial Intelligence*, vol.97 (1-2), pp.273–324, 1997.

[88] H. Liu and R. Setiono, "A probabilistic approach to feature selection - a filter solution", In *Proceedings of the Thirteenth International Conference on Machine Learning*, pages 319–327, 1996.

[89] H. Liu, H. Motoda, and L. Yu, "Feature selection with selective sampling", In *Proceedings of the Nineteenth International Conference on Machine Learning*, pp. 395–402, 2002b.

[90] L. Yu and H. Liu, "Feature selection for high-dimensional data: a fast correlation-based filter solution", In *Proceedings of the twentieth International Conference on Machine Learning*, pp. 856–863, 2003.

[91] L. Yu and H. Liu, "Redundancy based feature selection for microarray data", In *Proceedings of the Tenth ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 737–742, 2004.

[92] B. Boashash, M. Mesbah, P. Colditz, “Time frequency detection of EEG abnormalities,” in *Time–Frequency Signal Analysis and Processing: A Comprehensive Reference* Elsevier, Oxford, UK, pp. 663–670, article 15.5, 2003.

[93] K. Polat, S. Güneş, “Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast Fourier transform,” *Applied Mathematics and Computation*, vol.187, pp. 1017–1026, 2007.

[94] A. Subasi, "Application of adaptive neuro-fuzzy inference system for epileptic seizure detection using wavelet feature extraction", *Computers in Biology and Medicine,* vol.37 (2), pp.227-244. 2007.

[95] A.T Tzallas, M.G. Tsipouras and D.I. Fotiadis, "Automatic seizure detection based on time-frequency analysis and artificial neural networks", *Comput Intell Neurosci*. vol.7 (3): pp.1-13, 2007.

[96] V. Srinivasan, C. Eswaran, N. Sriraam, "Artificial neural network based epileptic detection using time-domain and frequency-domain features", *J. Med. Syst*, vol.29(6), pp.647–660. 2005.

[97] Guerrero-Mosquera, "Stochastic relevance analysis of epileptic EEG signals for channel selection and classification", Conference proceedings:. Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Conference07/2013, pp.2104-2107, 2013.

[98] H. Ocak , "Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy", *Exp. Syst. Appl*., vol.36(2), pp.2027–2036, 2009.

[99] S. F. Liang, H.C Wang, W.L.Chang, “Combination of EEG Complexity and Spectral Analysis for Epilepsy Diagnosis and Seizure Detection", *EURASIP Journal on* Advances in Signal Processing, pp.853434, 2010.

[100] H. Adeli, Z Zhou N, Dadmehr, "Analysis of EEG records in an epileptic patient using wavelet transform", J. Neurosci. Meth. vol.123 (1), pp.69–87, 2003.

[101] V. Bajaj and R.B. Pachori, "Classification of Seizure and Non-seizure EEG signals using empirical mode decomposition", *IEEE T INF TECHNOL B*. vol.16 (6), pp.1135-1142, 2012.

[102] Heejun Cho, Dongkyoo Shin, Dongil Shin, “A Study on parameter choice system design for EEG classifications", *International Journal of Information Technology & Computer Science*, vol.7 (2), 2013.

[103] Mohammad A Rahman et al., "A comprehensive survey of the feature extraction methods in the EEG research", Proceeding ICA3PP'12 Proceedings of the 12th international conference on Algorithms and Architectures for Parallel Processing- Volume Part II, pp. 274-283, 2012.

[104] A. Faris Abuhashish, Mohd S. Sunar, Hoshang Kolivand, Farhan Mohamed and Dzulkifli B. Mohamad, "Feature Extracted Classifiers Based on EEG Signals: A Survey", *Life Science Journal*, vol. 4, 2014.

[105] L.D. Iasemidis, L.D.Olson, R.S. Savit, and J.C. Sackellares, "Time dependencies in the occurrences of epileptic seizures: A nonlinear approach", *Epilepsy Res*., vol.17, pp.81–94, 1994.

[106] Lehnertz Ket al., “Nonlinear EEG analysis in epilepsy", *J. Clin. Neurophysiol*., vol. 18, pp.209–222, 2001.

[107] Übeyli E., "Analysis of EEG signals using Lyapunov exponents", *Neural Netw. World* vol.16 (3), pp.257–273, 2006.

[108] Nihal Fatma Gu¨ler, Elif Derya Ubeyli, I Gu¨ler, "Recurrent neural networks employing Lyapunov exponents for EEG signals classification", *Expert Systems with Applications* vol.29, pp. 506–514,2005.

[109] N. Pradhan, P. K. Sadasivan and G. R. Arunodaya, "Detection of seizure activity in EEG by an artificial neural network: A preliminary study", *Computers and Biomedical Research,* vol. 29, pp. 303-313, 1996.

[110] A. Petrosian, D. Prokhorov, R. Homan, R. Dashei and D. Wunsch, "Recurrent neural network based prediction of epileptic seizures in intra and extracranial EEG", *Neurocomputing,* vol. 30, pp. 201-218, 2000.

[111] Antti Saastamoinen et al, "Waveform detection with RBF network - Application to automated EEG analysis, J.Neurocomputing, vol.20, pp.1-13, 1998.

[112] D. Yao, "Electroencephalography inverse problem by subspace decomposition of the fourth-order cumulant matrix", *J. Biomed. Eng*., vol.17 (2), pp.174–178, 2000.

[113] V.P, Nigam, and D, Graupe, "A neural-network-based detection of epilepsy," *Neurol. Res,* vol. 26(1), pp. 55-60, 2004.

[114] N. Sadati, H.R. Mohseni, and A. Magshoudi, "Epileptic seizure detection using neural fuzzy networks," in: Proceedings of the *IEEE International Conference on Fuzzy Systems*, pp. 596-600, 2006.

[115] S. Ghosh-Dastidar, H. Adeli, "A new supervised learning algorithm for multiple spiking neural networks with application in epilepsy and seizure detection", *Neural netw*. vol.22 (10), pp.1419–1431, 2009.

[116] S.Ghosh-Dastidar, H. Adeli, and N. Dadmehr, "Principal component analysis enhanced cosine radial basis function neural network for robust epilepsy and seizure detection," *IEEE Transactions on Biomedical Engineering,* vol. 55(2), pp.512-518, 2008.

[117] V. Srinivasan, E. Eswaran, and N. Sriraam, "Approximate entropy based epileptic EEG detection using artificial neural networks," *IEEE Trans, Inform, Technol. Biomed,* vol. 11(3), pp.288-295, 2007.

[118] Subasi, "EEG signal classification using wavelet feature extraction and a mixture of expert model," *Expert Systems with Applications*., vol.32, pp.1084–1093, 2007.

[119] Forrest Sheng Bao, Donald Yu-Chun Lie, and Yuanlin Zhang, "A New Approach to Automated Epileptic Diagnosis Using EEG and Probabilistic

Neural Network", in Pro- ceedings of the 2008 20th IEEE International Conference on Tools with Artificial In-telligence, vol.02, pp. 482-486, 2008.

[120] L. Guo, D. Rivero, J. Seoane, A. Pazos, "Classification of EEG signals using relative wavelet energy and artificial neural networks" In: Proceedings of the first ACM/SIGEVO, Summit on Genetic and Evolutionary Computation (GEC'09),Shanghai: pp.177–184, 2009.

[121] L. Guo, D. Rivero, J. Dorado, J.R. Rabunal, A. Pazos, "Automatic epileptic seizure detection in EEGs based on line length feature and artificial neural networks," J. Neurosci. Methods, vol. 191(1), pp.101–109, 2010.

[122] L. Guo, D. Rivero, J. Dorado, C.R. Munteanu, A. Pazos, "Automatic feature extraction using genetic programming: an application to epileptic EEG classification", *Expert Syst. Appl*., vol.38 (8), pp.10425–10436, 2011.

[123] S.F. Liang, H.C. Wang, W.L. Chang, "Combination of EEG Complexity and Spectral Analysis for Epilepsy Diagnosis and Seizure Detection", *EURASIP Journal on Advances in Signal Processing*, pp.853434, 2010.

[124] K. Polat, S. Gunes, "A novel data reduction method: distance based data reduction and its application to classification of epileptiform EEG signals", Appl. *Math. Comput*. vol.200 (1), pp.10–27, 2008.

[125] K. Polat, S. Gunes, "Artificial immune recognition system with fuzzy resource allocation mechanism classifier, principal component analysis and FFT method based new hybrid automated identification system for classification of EEG signals", *Expert Syst. Appl*. vol.34 (3), pp.2039–2048, 2008.

[126] K Majumdar, "Human scalp EEG processing: various soft computing approaches", *Appl. Soft Comput.*, vol.11 (8), pp.4433–4447, 2011.

[127] Z. Iscan, Z. Dokur, T. Demiralap, "Classification of electroencephalogram signals with combined time and frequency features", *Expert Syst. Appl*. vol.38, pp.10499–10505,2011.

[128] D. Wang, D. Miao, C. Xie, "Best basis-based wavelet packet entropy feature extraction and hierarchical EEG classification for epileptic detection", *Expert Syst. Appl*., vol.38 (11) pp.14314–14320, 2011.

[129] U.R. Acharya, F. Molinari, S. Vinitha Sree, S. Chattopadhyay, Ng. Kwan-Hoong, J.S. Suri, "Automated diagnosis of epileptic EEG using entropies", *Biomed. Signal Process. Control*, vol.7 (4), pp. 401–408, 2012.

[130] U. R. Acharya, S. V. Sree, P. C. A. Ang and J. S. Suri, "Use of principal component analysis for automatic classification of epileptic EEG activities in

wavelet framework”, *Expert Systems with Applications*, vol. 39(10), pp. 9072-9078, 2012.

[131] T.Suneel Kumar, Vivek Kanhangad, Ram Bilas Pachori, " Classification of Seizure and Seizure-free EEG Signals using Multi-Level Local Patterns", Proceedings of the 19th *International Conference on Digital Signal Processing*, pp. 20-23 August 2014.

[132] S. Kumar, V. Kanhangad, and R.B. Pachori, "Classification of seizure and seizure-free EEG signals using local binary patterns", *Biomedical Signal Processing and Control*, vol. 15, pp. 33-40, 2015.

[133] Nabeel Ahammad, T Fathima and Paul Joseph, "Detection of Epileptic Seizure Event and Onset Using EEG", BioMed *Research International* Volume vol.2014, Article ID 450573. 2014.

[134] J.P. Pijan, D.N. Narveen , " Non linear dynamics of epileptic seizures on basis of intracranial EEG recordings", *Brain Topoger*, vol.9(4), pp.249-270, 1997.

[135] K. Lehnertz, C.E. Elger, "Spatio-temporal dynamics of the primary epileptogenic area in temporal lobe epilepsy characterized by neuronal complexity loss", *Electroencephalogr. Clin. Neurophysiol.* vol.95 (2), pp.108–117, 1995.

[136] L. D. Iasemidis, D.-S. Shiau, W. Chaovalitwongse, J. Sackellares, P. M. Pardalos, J. C. Principe, P. R. Carney, A. Prasad, B. Veeramani, and K. Tsakalis, “Adaptive epileptic seizure prediction system,” *IEEE Trans.* Biomed. Eng., vol. 50(5), pp. 616–627, 2003.

[137] J. Lamberts, P.L.C. Van den Broek, J. Bener, J. Van Egmond, R. Dirksen, A.M.L.Cohen, "Correlation dimension of the human electroencephalogram corresponding to cognitive load", *Neuropsychobiology*, vol.41 (3), pp.149–153, 2000.

[138] W.J. Freeman, "A proposed name for a periodic brain activity: stochastic chaos", *Neural Netw.*, vol.13 (1), pp. 11–13, 2000.

[139] A. Wolf, J.B. Swift, H.L. Swinney, J.A. Vastano, “Determining Lyapunov exponents from a time series", Physica D, vol.16, pp. 285–317, 1985.

[140] M. Rosenstein, J.J. Colins, C.J. De Luca, "A practical method for calculating largest Lyapunov exponent from small data sets", *Physica D*, vol.65, pp.117–134, 1993.

[141] N. Kannathal, C.M. Lim, U.R. Acharya, P.K. Sadasivan, "Entropies for detection of epilepsy in EEG", *Comput. Methods Programs Biomed*. vol.80 (3), pp.187–194, 2005.

[142] P. Grassberger, I. Procassia, "Measuring the strangeness of strange attractors", *Physica D* vol.9 (1–2), pp.189–208, 1983.

[143] K. Lehnertz, R. Andrzejak, F. Mormann, J. Wegner, G. Widman, P. David, and C. Elger, “Linear and nonlinear analysis techniques for anticipating epileptic seizures,” *Epilepsia*, vol. 40, pp. 71, 1999.

[144] K. Lehnertz et al., “Nonlinear EEG analysis in epilepsy,” *J. Clin. Neurophysiol.*, vol. 18, pp. 209–222, 2001.

[145] H.Adeli, S. G. Dastidar, and N. Dadmehr, “A wavelet-chaos methodology for analysis of EEGs and EEG sub-bands to detect seizure and epilepsy,” *IEEE Trans. Biomed. Eng.*, vol. 54(2), pp. 205–211, 2007.

[146] I. Gu¨ler and E. D. U¨ beyli, “Multiclass support vector machines for EEG signals classification,” *IEEE Trans. Inf. Technol. Biomed.*, vol. 11(2), pp. 117–126, 2007.

[147] S.M. Pincus, "Approximate entropy as a measure of system complexity", *Proc Nat Acad Sci*, USA, vol. 88, pp.2297–2301, 1991.

[148] L. Diambra, J. Figueiredo, C. Malta, "Epileptic activity recognition in EEG recording", *Phys. A Stat. Mech*. Appl., vol.273 (3–4), pp.495– 505, 1999.

[149] J.S. Richman, M.J. Randall, "Physiological time-series analysis using approximate entropy and sample entropy", Am. *J. Physiol. Heart Circ. Physiol*., vol.278, pp.H2039–H2049, 2000.

[150] S. F. Liang, H. C. Wang, and W. L. Chang, “Combination of EEG complexity and spectral analysis for epilepsy diagnosis and seizure detection,” *EURASIP J. Adv. Signal Process*, vol. 2010, pp. 853434-1–853434-15, 2010.

[151] Sharma, R.B. Pachori, and U.R. Acharya, "Application of entropy measures on intrinsic mode functions for the automated identification of focal electroencephalogram signals", *Entropy*, vol. 17(2), pp. 669-691, 2015.

[152] M. Schneider, P. N.Mustaro, and C. A. M. Lima, “Automatic recognition of epileptic seizure in EEG via support vector machine and dimension fractal,” in *Proc. Int. Joint Conf. Neural Netw.*, pp. 2841–2845, 2009.

[153] T. Fathima, Y. U. Khan, M. Bedeeuzzaman, and O. Farooq, “Discriminant analysis for epileptic seizure detection,” in *Proc. Int. Conf. Devices Commun.*, pp. 1–5, 2011.

[154] W.S. Pritchard, D.W. Duke, "Measuring chaos in the brain: a tutorial review of nonlinear dynamical EEG analysis", *Int. J. Neurosci*. vol.67 (1–4), pp.31–40, 1992.

[155] C.L. Nikias, "Higher-order spectral analysis, engineering in medicine and biology society", in: Proceedings of the 15th Annual International Conference of the IEEE, pp. 319,1993.

[156] K.C. Chua, V. Chandran, U.R. Acharya, C.M. Lim, "Application of higher order spectra to identify epileptic EEG", J. Med. Syst. vol.35 (6), pp.1563–1571, 2011.

[157] U.R. Acharya, S. Vinitha Sree, J.S. Suri, "Automatic detection of epileptic EEG signals using higher order cumulant features", Int. *J. Neural Syst*., vol. 21 (5), pp.1–12, 2011.

[158] R.J. Martis, U.R. Acharya, J.H. Tan, A. Petznick, R. Yanti, K.C. Chua, E.Y.K. Ng, L. Tong, "Application of empirical mode decomposition (EMD) for automated detection of epilepsy using EEG signals", *Int. J. Neural Syst*., vol. 22 (6), 2012.

[159] R. B. Pachori, "Discrimination between ictal and seizure-free EEG signals is using empirical mode decomposition," in *Res. Lett. Signal Process*, vol. 2008, pp. 293056-1–293056-5, 2008.

[160] R.B. Pachori and S. Patidar, "Epileptic seizure classification in EEG signals using second-order difference plot of intrinsic mode functions", *Computer Methods and Programs in Biomedicine*, vol. 113(2), pp. 494-502, 2014.

[161] R. Sharma and R.B. Pachori, Classification of epileptic seizures in EEG signals based on phase space representation of intrinsic mode functions", *Expert Systems with Applications*, vol. 42(3), pp. 1106-1117, 2015.

[162] M.C. Van den Berg-Lenssen, C.H.M. Brunia, J.A. Blom, "Correction of ocular artifacts in EEGs using an autoregressive model to describe the EEG; a pilot study" pp.-4694 , 1989.

[163] A. Jain and D. Zongker, "Feature selection: Evaluation, application, and small sample performance", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 2, pp.153-158, 1997.

[164] Helen T. Ocbagabir, Khald A. I. Aboalayon, Miad Faezipour, " Efficient EEG Analysis for Seizure Monitoring in Epileptic Patients", *Systems, Applications and Technology Conference (LISAT)*, pp.1-8, 2013.

[165] Krzywinski, Martin, and Naomi Altman., "Points of Significance: Visualizing samples with box plots," Nature Methods, 2014.

[166] M. D'Alessandro, "Epileptic seizure prediction using hybrid feature selection over multiple intracranial EEG electrode contacts: a report of four patients. *IEEE Transactions on Biomedical Engineering*, vol. 5, 2003.

[167] S. M. Kay, S. L. Marple, 'Spectrum analysis – A modern perspective', *Proceedings of the IEEE*, vol. 69, pp.1380–1419, 1981.

[168] P.K. Naik, S. Sindura, T. Singh, H. Singh, "Quantitative structure activity relationship (QSAR) of the insecticides: The development of predictive in vivo insecticide prediction model", SAR QSAR in *Env Res*, vol.20, pp.551-556, 2009.

[169] S. Amjed Al-Fahoum, Ausilah A Al-Fraihat, "Methods of EEG Signal Features Extraction Using Linear Analysis in Frequency and Time-Frequency Domains", *ISRN Neuroscience*. Article ID 730218, 7 pages, 2014.

[170] Dourado A, Martins R, J Duarte, B Direito. Towards Personalized Neural Networks for Epileptic Seizure Prediction. Artificial Neural Networks - ICANN 2008 *LNCS Springer-Verlag Berlin Heidelberg, vol.*5164, pp. 479-487, 2008.

[171] S. J. M. Smith, "EEG in the diagnosis, classification, and management of patients with epilepsy", *J.Neurol Neurosurg Psychiatry*, vol.76, pp.ii2-ii7, 2005.

[172] E.D. Ubeyli, "Least squares support vector machine employing model-based methods coefficients for analysis of EEG signals", *Expert Syst. Appl*., vol.37 (1), pp.233–239, 2010.

[173] C.C. Chang and C. J. Lin LIBSVM, A library of support vector machines, 2012.

[174] Li Jun-wei, Jason Gu, XIE Yun, "Analysis and Research on Combination Feature Extraction Method of EEG Signal", *Automation, Control and Intelligent Systems,* vol. 3(2), pp. 26-30, 2015.

[175] B. Qian, K. Rasheed, "Hurst exponent and financial market predictability", in:IASTED Conference Financial Engineering and Applications, pp. 203–209, 2004.

[176] A. Carbone, "Algorithm to estimate the Hurst exponent of high dimensional fractals," *Phys. Rev. E*, vol. 76, pp. 056703, 2007.

[177] Seung-Hyeon Oh, Yu-Ri Lee, and Hyoung-Nam Kim, "A Novel EEG Feature Extraction Method Using Hjorth Parameter" *International Journal of Electronics and Electrical Engineering*, vol. 2(2), pp. 106-110,2014.

[178] S.A. Hosseini, M-R. Akbarzadeh-T, M-B. Naghibi-Sistani, "Dynamical Qualitative and Quantitative Evaluation of EEG Signals in Epileptic Seizure Recognition," *MECS I.J. Intelligent Systems and Applications*, vol. 06, pp. 41-46, 2013.

[179] Deon Garrett, David A. Peterson, Charles W. Anderson, and Michael H. Thaut, "Comparison of Linear, Nonlinear, and Feature Selection Methods for EEG Signal Classification" *IEEE Transactions on Neural systems and Rehabilitation Engineering*, vol.11(2), 2003.

[180] V. Gerla, L. Lhotska, and V. Krajca, “Utilization of time-dependence in EEG signal classification. In *The 3rd European Medical and Biological Engineering Conference EMBEC05*, CLS JEP, vol. 11, 2005.

[181] M.K. Ahirwal, A. Kumar, and G.K. Singh, Improved range selection method for evolutionary algorithm based adaptive filtering of EEG/ERP signals, *Neurocomputing*, vol.144, pp. 282-294, 2014.