BOOK	Enhanced Telemedicine and e-Health: Advanced IoT Enabled Soft Computing Framework
DOI	The final authenticated version is available online at https://doi.org/10.1007/978-3-030-70111-6_14
CITATION	Subasi A. (2021) Disease Prediction Using Artificial Intelligence: A Case Study on Epileptic Seizure Prediction. In: Marques G., Kumar Bhoi A., de la Torre Díez I., Garcia-Zapirain B. (eds) Enhanced Telemedicine and e-Health. Studies in Fuzziness and Soft Computing, vol 410. Springer, Cham.

Metadata of the chapter that will be visualized in SpringerLink

Book Title	Enhanced Telemedicine and e-Health
Series Title	
Chapter Title	Disease Prediction Using Artificial Intelligence: A Case Study on Epileptic Seizure Prediction
Copyright Year	2021
Copyright HolderName	The Author(s), under exclusive license to Springer Nature Switzerland AG

Corresponding Author	Family Name	**Subasi**
	Particle	
	Given Name	**Abdulhamit**
	Prefix	
	Suffix	
	Role	
	Division	Faculty of Medicine, Institute of Biomedicine
	Organization	University of Turku
	Address	Turku, 20520, Finland
	Division	Department of Computer Science, College of Engineering
	Organization	Effat University
	Address	Jeddah, 21478, Saudi Arabia
	Email	abdulhamit.subasi@utu.fi absubasi@effatuniversity.edu.sa

Abstract

Artificial Intelligence uses statistical theory to generate mathematical models from samples. After a model is generated, its depiction and algorithmic solution for understanding require being competent as well. Biomedical data related to different diseases are recorded from a body, which can be at the organ level, cell level or molecular level. Biomedical data is mainly utilized to predict, diagnose or identify particular physiological or pathological conditions. The goal of biomedical data analysis is exact modelling of data by employing feature extraction, feature selection and dimension reduction for the prediction and detection of upcoming pathological problems by utilizing artificial intelligence algorithms. This chapter explains the steps of biomedical data analysis and how artificial intelligence techniques are utilized in disease prediction. An automated epileptic seizure prediction and detection approach based on deep learning is also presented. Since Deep Learning can automatically extract and learn features, the electroencephalography (EEG) time series are fed into the deep learning model. Deep Learning has been utilized in the prediction and detection of epileptic seizures. Since EEG recordings are high dimensional data, a Convolutional Neural Network (CNN) is suitable for this use. The results show that CNN achieved a testing accuracy of 99.09% accuracy for the prediction of epileptic seizures from EEG signals.

Keywords (separated by '-')

Biomedical data analysis - Disease prediction - Artificial intelligence - Deep learning

Disease Prediction Using Artificial Intelligence: A Case Study on Epileptic Seizure Prediction

Abdulhamit Subasi

Abstract Artificial Intelligence uses statistical theory to generate mathematical models from samples. After a model is generated, its depiction and algorithmic solution for understanding require being competent as well. Biomedical data related to different diseases are recorded from a body, which can be at the organ level, cell level or molecular level. Biomedical data is mainly utilized to predict, diagnose or identify particular physiological or pathological conditions. The goal of biomedical data analysis is exact modelling of data by employing feature extraction, feature selection and dimension reduction for the prediction and detection of upcoming pathological problems by utilizing artificial intelligence algorithms. This chapter explains the steps of biomedical data analysis and how artificial intelligence techniques are utilized in disease prediction. An automated epileptic seizure prediction and detection approach based on deep learning is also presented. Since Deep Learning can automatically extract and learn features, the electroencephalography (EEG) time series are fed into the deep learning model. Deep Learning has been utilized in the prediction and detection of epileptic seizures. Since EEG recordings are high dimensional data, a Convolutional Neural Network (CNN) is suitable for this use. The results show that CNN achieved a testing accuracy of 99.09% accuracy for the prediction of epileptic seizures from EEG signals.

Keywords Biomedical data analysis · Disease prediction · Artificial intelligence · Deep learning

---

A. Subasi (✉)
Faculty of Medicine, Institute of Biomedicine, University of Turku, Turku 20520, Finland
e-mail: abdulhamit.subasi@utu.fi; absubasi@effatuniversity.edu.sa

Department of Computer Science, College of Engineering, Effat University, Jeddah 21478, Saudi Arabia

G. Marques et al. (eds.), *Enhanced Telemedicine and e-Health*, Studies in Fuzziness and Soft Computing 410, https://doi.org/10.1007/978-3-030-70111-6_14

1 Introduction

Biomedical data is collected from a biological or medical source that might be at the organ, cell, or molecular levels. On the one hand, biomedical data is primarily collected in hospitals to identify certain clinical or pathological disorders and to evaluate the therapy and diagnose the disorder. On the other hand, biomedical data analysis is used to eliminate noise, construct reliable models and analyze their components, extract characteristics for vital component and predict possible functional or pathological events in the heart, brain, muscle and kidney [1]. Moreover, biomedical data includes information that is useful for understanding the complex mechanisms of pathophysiology and behavior of living systems. The analysis of biomedical data is usually required to improve the relevant knowledge and to define the pathology level for routine clinical diagnosis, recovery or therapy. Several biomedical data processing methods also referred to as preprocessing techniques, are available such as denoising, averaging, sampling, segmentation, spectral estimation and feature extraction [2].

Biomedical health data have been widely employed in recent years for research and clinical decision-making in the healthcare research field, such as diagnosis, treatment, and prediction of diseases. Clinical judgments have historically relied on the finding, expertise, and experience of physicians, with the aid of numerous clinical and diagnostic tests. This approach leads to unintended biases, mistakes and high costs, which has a negative effect on the quality of patient service [3]. Wu et al. [4] suggested combining clinical decision support with computer-based patient records that would eliminate medical errors and unintended changes in procedures and increase patient well-being. Computer-based decision support has become popular and started to combine with information management in healthcare. As a result, this has created a significant amount of biomedical data involving healthcare records, interactions between patient and doctor, medical history, referral details and insurance claim data. With the rapid evolution of methods and techniques for data processing, biomedical healthcare data are being used in an extensive range of healthcare research fields rather than simply for record-keeping, such as healthcare data analysis [5], chronic disease surveillance [6], the creation of disease risk forecasting models [6, 7] and the comparison of disease incidence and drug outcomes [8]. Besides, these data are an essential tool for chronic disease monitoring and observation [6]. Severe health problems such as heart failure and epilepsy are frequently identified with development at a later phase. If predictive models would detect these potential health risks before, patients could be treated better, which could alter the course of the disease or disorder. Because of the availability of massive biomedical healthcare data, prediction of diseases has become a significant research field [9]. Artificial intelligence techniques have had a substantial influence on the lifestyle of humans over the past two decades, by forecasting human activities and forthcoming developments. These techniques are well adapted to transforming collected data into useful knowledge to help decision making in the healthcare system to increase accuracy in diagnosis and speed up testing time [10].

Modern healthcare systems have improved human life expectancy through the advent of medications, medical facilities, and the maintenance of health records for patients. Nevertheless, healthcare support systems face substantial challenges with modern developments in healthcare, such as lack of sufficient medical information, misdiagnosis, inevitable errors, data risk, and delayed communication. Computerized healthcare information management, such as the Electronic Health Record (EHR), Disease Prediction Scheme (DPS), and Clinical Decision Support System (CDSS) draw significant interest. To help the physician in the decision-making process, early-stage monitoring of the healthcare data of patients became effective for the correct diagnosis of the disease. Hence, CDSS gained unexpected prominence due to technical advancement [11]. The advent of Internet of Things (IoT) and medical sensors includes a significant amount of diverse devices, which continuously track and acquire physiological signals from distant users such as heartbeat, brain signals, motor characteristics, body temperature, and insulin [12]. In tracking the health conditions of patients, valuable interpretations can only be extracted from the signals of the medical sensors. Medical sensors have recently been equipped with emergency warnings to inform the patient or his/her family members and emergency service once the biomedical signal reaches the threshold. Finally, a large amount of medical data for patients is obtained by healthcare service providers creating a valuable tool for the support network for clinical decision making.

Cloud computing acts as a foundation for vast storage and data processing because of the space limitation characteristics of IoT and minimal storage facilities. The cloud-related healthcare framework monitors all cloud-based health information while the patient is travelling for specific treatment from one hospital to another. Nevertheless, as the medical data is kept on the cloud server to identify, predict disorder, and maintain the healthcare records, it is suitable for medical data being processed in real-time. Therefore, local processing of health information for a patient and the sending of an instantaneous medical decision or a warning when an emergency happens to allow healthcare services to save human lives [13]. The main objective of this chapter is to present the disease prediction using artificial intelligence.

Deep Learning algorithms are gaining a lot of attention because of their excellent success in different machine learning applications. In the electroencephalography (EEG) study, recurrent neural networks (RNN) were used to learn temporal patterns for the detection of epileptic seizures [14]. Long Short-Term Memory (LSTM) is an evolution of RNNs, which requires gates to deal with the issue of the disappearing gradient. More recently, LSTM has been used to predict epileptic seizures using EEG signals [15]. Noteworthy attention in the EEG signal analysis has also been attracted by Convolutional Neural Networks (CNNs). CNNs are used to both raw data [16] and wavelet space [17] to classify epileptic EEG signals and achieved higher performance. The key benefit of CNNs is the ability to automatically learn new features, producing better results while the amount of data is adequate to train the CNNs [18]. This chapter presents the use of artificial intelligence techniques for the prediction and detection of the epileptic seizure using feature extraction and signal transformations to produce deep learning inputs.

The motivation for applying artificial intelligence to the problem of predicting epileptic seizures is their effectiveness in the classification of biomedical signals. Further exploratory studies are required to utilize EEG recordings for computerized epileptic seizure detection. Moreover, this work is one of the studies used to tackle the issue of seizure prediction using EEG and Deep Learning. Epileptic seizure detection in advance will enable epileptic patients or their caregivers to take measures until an epileptic seizure occurs, thus reducing complications and possible damage associated with the case, where EEG is generally monitored [19]. By using CNN, 99.09% accuracy is achieved for the epileptic seizure prediction and detection.

This chapter is structured as follows: disease prediction is discussed in Sect. 2. Section 3 includes artificial intelligence techniques for disease prediction, which includes a description of the Deep Learning algorithm. Section 4 provides details of the case study of epileptic seizure prediction, and Sect. 5 presents the discussion. Finally, in Sect. 6, conclusions are given.

2 Disease Prediction

Disease prediction and detection is a technique of artificial intelligence that first employs training data to create a model, and then the resulting model utilizes the test data to obtain prediction results. Classification approaches based on artificial intelligence for the chronic disease prediction and detection were applied on disease datasets, and the findings are promising. An original classification strategy that can accelerate and make simpler the process of chronic disease prediction and detection is crucial [20].

Computer-aided decision support systems (CADSS) help the physician monitor the symptom-based risks of disorder incidence in the subjects. A person cannot easily detect signs of diseases such as diabetes and arrhythmia. Therefore, CADSS is used for proper evaluation of the illness, assisted by artificial intelligence techniques, medical sensors, and computing services. The use of artificial intelligence techniques enables improved efficiency with ease of implementation for prediction of complex disease. CADSS is a computerized decision-making system, which offers a patient community with predictive disease and treatment recommendations, rather than patient-centered recommendations. As CADSS does not find clear characteristics of the patients for predicting illness, the recommendations given are unreliable. The emergence of medical big data has driven customized CADSS production to increase the efficiency of existing CADSS [13].

Disease prediction from healthcare data that demonstrate valuable knowledge in large quantities relating to patients with specific diseases is a scientific domain issue [21, 22]. Predicting the risk of the disease includes predicting the likelihood of disease and implementing a preventive plan either to minimize the risk of disease in a specific way or to avoid disease risk altogether. Prediction of illness has several advantages, such as early-stage analysis and treatment of disease, disease drop, and death prevention [23–25].

Artificial intelligence techniques are used extensively in medicine. Different classification algorithms for disease diagnosis have been developed to provide high precision for disease prediction. After analyzing the various characteristics of the disease, several artificial intelligence techniques are built for predicting different types of disease at the early stages. These algorithms are commonly used in breast cancer, kidney disease, thyroid disease, diabetes, other cancers, erythemato-squamous diseases and much more. Different classification algorithms are applied to obtain better prediction accuracy to build a CADSS [26].

3 Artificial Intelligence Techniques for Disease Prediction

Artificial intelligence is a term that uses a collection of data or some knowledge for prediction and classification. The learning process is essentially the implementation of optimization of the model parameter with the training dataset or previous experience. Models may either be predictive, to allow future predictions; descriptive, to derive information from data input or both. Two essential tasks are carried out in artificial intelligence: (1) analysing the volume of data and enhancing the model and (2) testing the model and demonstrating the solution in an effective manner. In some applications, the learning efficacy is as crucial as the precision of the classification. Also, artificial intelligence helps the system to learn about and adapt to changes in various environments. Artificial intelligence supports us in healthcare data, vision, voice, face or other kinds of recognition [27].

In the study of medical data, literature presents numerous scientists who have employed diverse artificial intelligence techniques for predicting chronic diseases to obtain better diagnostic results and prediction performance. Several artificial intelligence algorithms, such as Naïve Bayes, decision trees, artificial neural networks (ANN), k-nearest neighbor, and support vector machines, have been introduced and utilized recently for chronic disease prediction and detection. Intelligent automated systems [28] are essential in other research and scientific applications for prediction and detection of various diseases,. Nonetheless, most conventional methods are not adaptive, that decreases the performance and increases the time for decision-making on disease prediction. Because of these limitations of current traditional classification methods, such as lower performance and longer decision-making time, an improved classification precision is needed to predict diseases [20]. The general framework for disease prediction using artificial intelligence techniques is shown in Fig. 1.

3.1 Artificial Neural Networks

ANN are a learner of greater interest in various fields of machine learning. ANN can generate a high precision prediction for different sorts of problems. Furthermore, ANN can mimic a human brain's role in making predictions, recognizing patterns or

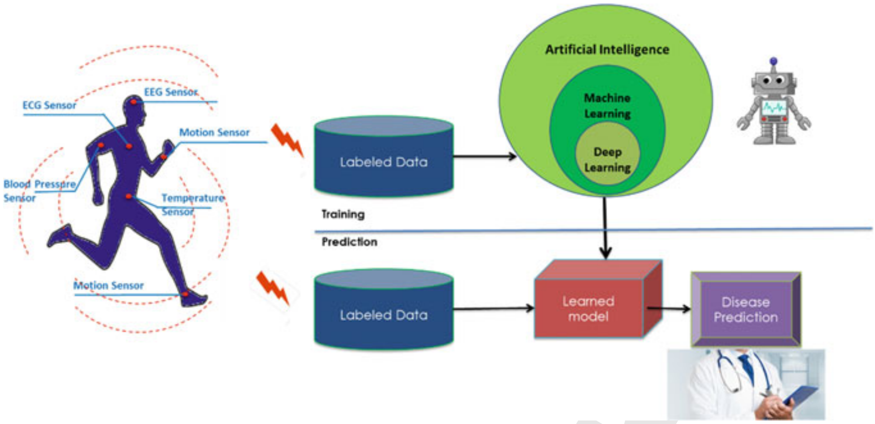

Fig. 1 General framework for disease prediction using artificial intelligence techniques

learning from past experiences. This procedure is expressed in a computer program that applies algorithms for machine learning and pattern recognition to construct an efficient predictive model. When building the ANN, the specialist will determine the number of nodes that are suitable for the problem, the most appropriate weight value for every connection, how the nodes should be connected, and how ANN should be trained. In ANN, input data is obtained through input nodes, the values of which are weighted by the values of the weights and preserved in the links.

In addition, in the hidden layer, a transfer function is utilized to generate the output. This output is transferred to the final node as a prediction. The nodes in ANN can be divided into input and output layers. Input layer contains a group of input values, while the output layer contains target values. Unit is of any input variable, which has its unit within neural networks. Each input layer is connected with the output layer by a minimum of one weight. The main drawback of the ANN is the complexity of the resulting model. The model is built as a “black-box”. As the model’s complexity increases, they may add additional hidden layers. This approach is among the most complicated of predictions to implement. Nevertheless, the advantage of neural networks is that all data forms can be processed without noticeable changes in input data and information of linearity and distribution [29].

ANN can be described as a combination of various models of predictions feeding into one another. There is a hidden layer between the input and output layer. This type of structure is called as multilayer architecture in neural networks. The model is built in propagating forward and backwards. In scientific studies, the back-propagation technique is the most popular. In reality, a finite amount of input variables can be interpreted by neural networks; otherwise, the time needed is high. The success of ANN depends on how the learning process evolves and what initial values are utilized in the training phase [29].

3.2 Deep Learning

For any learning process, if a linear model is not appropriate, new features, which are nonlinear functions of the input are likely to be found, and then a linear model is built in those features space. It requires to differentiate what roles of such a strong basis are. The best approach in ANN is to extract these features in the hidden layer since it has the advantage of extracting informative features and utilizing them to predict the output in a supervised manner. An ANN with one hidden layer has poor capability, whereas an ANN with numerous hidden layers can learn more complicated functions. It is the philosophy behind deep neural networks, in which every hidden layer, starting from simple network input, utilizes the preceding layer values and learns more complex input features.

An additional characteristic of deep neural networks is that, up to the output layer, succeeding hidden layers are connected to more complex structures where these intangible definitions involve the outputs being studied. In deep learning, the idea is to learn features with minimal human interference the characteristic levels of that abstraction, since in some applications it is not known which assembly is present in the input and any kind of dependences should be learned in an automated way throughout the training [30].

One critical problem of training an ANN with several hidden layers is that subsequently multiplying the derivatives in all layers is crucial in order to back-propagate the error to the preceding layer, and the gradient has vanished. This is also the explanation for poor learning in recurrent, extended neural networks. For CNN, this situation cannot occur since the fan-in, and fan-out of hidden units are, of course, unimportant [31]. The entire deep neural network can be trained in a supervised manner if there are enough labeled data and computing resources.

Deep learning methods are primarily popular, as they require less manual intervention. It is intuitive to think of several layers of rising notion lying underneath deep learning. In several applications, the abstraction layers can be understood as an informative discovery of such an abstract illustration and a better understanding of the problem [27, 32].

3.3 Convolutional Neural Networks

CNN are an improved version of deep neural networks. They consist of units named neurons that take a weighted sum of inputs and generate an operating level of outputs. The activity level is often a nonlinear input function, such as a rectified linear unit, in which the operation is equal to the input for all positive inputs, and 0 for all non-positive inputs. What is remarkable about CNN is how this architecture structures the interactions between the neurons. Units are grouped into layers in a feedforward neural network, and the units in a given layer only receive input from units in the layer below. CNN are networks with feedforward features. Unlike regular vanilla feedforward networks, CNN has spatial structure units. In each layer, units are grouped into 2-D grids called feature maps. These maps of features are the product of a convolution made on the layer below. This implies that the same convolutional filter (set of weights) is used at each location in the layer below. Thus, a unit may only obtain input from units in the layer below at a similar position with a specific location on the 2-D grid. In addition, the weights attached to the inputs in a function map are the same for each unit. CNN recognize images as volumes, i.e. three-dimensional objects to be defined only by width and height, rather than flat canvases since digital color images have a red-green-blue (RGB) encoding that perceives objects' color spectrum by mixing those three colors. These images are processed by a convolutional neural network as three different layers of color piled one on top. A convolutional neural network accepts a normal color image as a rectangular box, the height and width of which is determined by the number of pixels in these dimensions with a depth of three layers, one for every letter in RGB. Such layers of depth are called channels. The intensity of R, G and B will be represented by a number for each pixel of an image to form an element in one of the three stacked two-dimensional matrices. These numbers are the original, sensory, raw, features that are served into the convolutional neural network, and the aim of a CNN is to detect relevant signals, which essentially benefit to better distinguish images [33].

4 Case Study of Epileptic Seizure Prediction

Epilepsy is a neurological disorder characterized by continuing inclination to produce recurrent seizures and may disturb people of different age. Epilepsy results from the progressive neurobiological mechanism of 'epileptogenesis' [34] that affects the normal brain network to fire neurons in the cerebral cortex in a self-sustaining, hyper-synchronized manner. Seventy million people worldwide suffer from epilepsy, which is in the list of the most serious brain disorders after stroke, migraine, and Alzheimer's disease [35]. Epileptic seizures are devastating and disrupting patients' daily activities and are related to increased risk of early death. Epilepsy treatment is further complicated by the shortage of neurologists in numerous countries, especially in developing countries.

While in some literature epilepsy and seizures are often used synonymously, not all seizures are epileptic, and seizures can also happen because of the desperate neurological problems without inevitably indicating a long-term inclination to repeated epileptic seizures. The epileptic seizure is triggered by an abrupt irregular, self-sustaining electrical discharge, which takes place in the cerebral networks and typically persists less than a few minutes.

Epileptic seizure attacks are difficult to predict, and even the frequency and length of the attack cannot be predicted. Hence, the prediction and detection of epilepsy attacks in early stages are essential for preventing and counteracting their contrary effects. The brain activity of epilepsy patients can be classified in diverse states: inter-ictal (between seizures), pre-ictal (immediately preceding seizure), ictal (during a seizure), and post-ictal (immediately after a seizure).

Epileptic seizure prediction is a problem of classification based on distinguishing between pre-ictal and inter-ictal states. An epileptic seizure occurs in groups due to the chronic nature of epilepsy and patients affected by seizure clusters may gain benefit by predicting follow-up seizures [36]. Moreover, epileptic seizures, which cause hardly long-lasting injuries or death, may lead to loss of perception or cause only minor mental distress. Seizures have highly variable length of time and incidence rate. Its length can vary from a few seconds to several minutes. There exist epileptic patients who have only a few seizures in their lives, while others have multiple seizures within a single day [37, 38].

EEG is a diagnostic technique especially useful for analyzing the brain anatomy throughout an epileptic seizure attack. Epilepsy diagnosis and treatment has been studied extensively by EEG. Furthermore, EEG signals are non-Gaussian, non-stationary and utilized to identify the form of brain disorders by measuring the electrical activity of the brain. EEG assessment research helps in the classification between normal and abnormal brain activity. To accurately predict epilepsy, longer duration EEG recordings need to be examined.

Expert neurologists analyze epilepsy by examining continuous EEG signals that have been documented over multiple days, weeks, or even months, requiring significant human time and effort. Over the years, different studies have used predictive methods based on Artificial Intelligence to tackle this problem. Deep learning is an advanced machine learning technique, which can more effectively learn patterns from large data sets by processing it via a multilayer hierarchical architecture. Since deep learning can achieve very reliable results, researchers have inspired to eliminate various real-world problems by using deep learning approaches. Moreover, different scientists proposed deep learning-based techniques to the epileptic seizure prediction [36]. The common structure for the epileptic seizure prediction is shown in Fig. 2. This structure contains three key components: (i) signal preprocessing/denoising, (ii) feature extraction/dimension reduction and (iii) classification. In this section, a complete description of each component will be delivered.

Scientists have been working, since the last century, to overcome the challenges associated with epilepsy diagnosis and prediction. Since EEG signals are a crucial source for tracking brain activity before, during, and after an epileptic seizure, epileptic seizure prediction centered first on EEG recording analysis. EEG signals contain several contaminations such as eye-movements, blinks, heart impulses and muscle noise. Various filtering and noise reduction methods are employed to reduce the influence of these sources of noise and artifacts [39]. Substantial features are required for creating an artificial intelligence model for identifying and classifying inter ictal, pre-ictal and ictal phases after removal of noise [36].

EEG is the most popular way of investigating epilepsy and recording changes in the behavior of the electrical brain that may announce an upcoming seizure. LSTM networks are implemented using EEG signals in the prediction of an epileptic seizure. To enhance the lives of patients with tonic seizures and drug-resistant epilepsy,

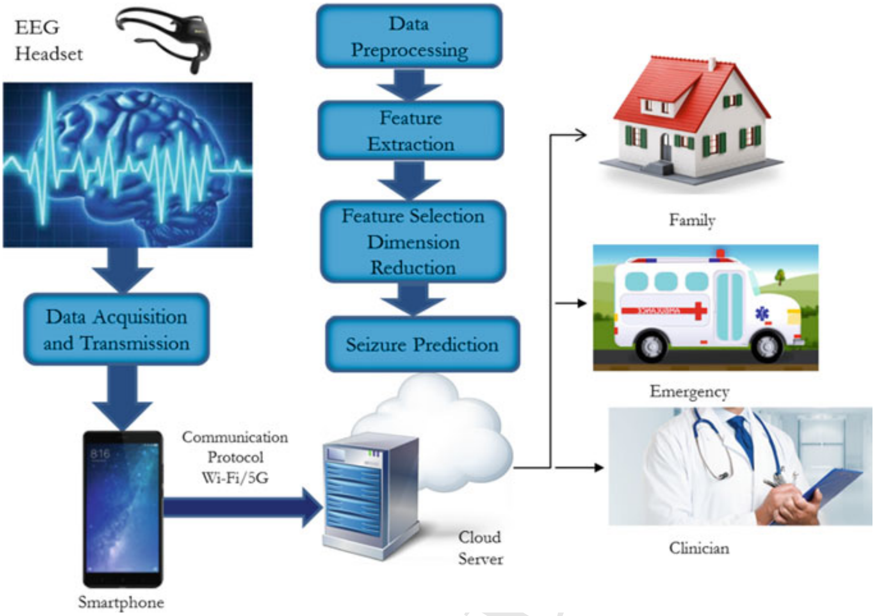

Fig. 2 General framework for epileptic seizure prediction using artificial intelligence techniques

seizure prediction has gained increasing attention as one of the most complicated predictive data analysis techniques. Several studies have achieved good results in delivering sensible alert systems or interactive regulation of neural stimulation over persistent seizures; some of them have accomplished high efficiency. Nevertheless, most of these works include handcraft feature extraction and/or tailor-made feature extraction that is implemented separately for every patient, to yield a low false prediction rate and high sensitivity.

Epilepsy is the second most prevalent brain disorder. Epilepsy makes living a normal life difficult for patients, as the occurrence of seizures cannot be predicted easily. Hence, if a fair amount of time before their occurrence could be predicted for seizures, patients with epilepsy might take precautions against them and enhance their health and life quality [40].

Truong et al. [41] applied CNN to various datasets of intracranial and scalp electroencephalogram and suggested a generalized retroactive and patient-specific technique for seizure prediction. They used the short-term Fourier transform to extract features in both the time domain and the frequency domain on 30-s EEG frames. The algorithm automatically produces optimized features to better recognize the pre-ictal and inter-ictal segments for each patient.

Chu et al. [40] studied a novel seizure prediction framework based on an interpretation of the attractor state. They investigated the transition phase from normal to attractor seizure state and examined the phenomenon realized before entering the attractor seizure state. Moreover, they have described a computed spectral measure for seizure prediction in scalp EEG from the result of an experiment. Six EEG frequency bands of Fourier coefficients are derived from scalp EEG recordings, and the spectral measurement is determined on the basis of the coefficients for each half-overlapped duration of 20 s. The computed spectral measure is implemented using a low-complexity approach to predict seizures. They developed an early warning system before the occurrence of the epileptic seizure from the scalp EEG. Using this algorithm, a low-complexity seizure prediction framework was implemented using scalp EEG.

Tsiouris et al. [15] used a two-layer LSTM network to test the efficiency of seizure prediction employing four different pre-ictal window lengths, ranging from 15 min to 2 h. The proposed LSTM network utilizes a broad range of features, containing time and frequency domain features, the cross-correlation between EEG channels and graph-theoretical features extracted before classification. The results showed that the developed method is capable of predicting all 185 seizures, producing 0.11–0.02 false alarms per hour, based on the duration of the pre-ictal window, with high seizure sensitivity levels and low false prediction levels (FPR).

A complex aspect of the treatment of epilepsy, the exact classification of various epileptic disorders, is of specific concern and has been thoroughly studied. Gao et al. [42] suggested a novel deep learning approach, called as the EEG epileptic signal classification (EESC). First, this approach extracts power spectrum density energy diagrams (PSDEDs) from the epileptic EEG signals, then extracted features from the PSDED are utilized as an input of deep CNNs and transfer learning to classify four types of epileptic states (inter-ictal, pre-ictal to 30 min, pre-ictal to 10 min, and seizure). It outperforms the current methods of epileptic seizure prediction in terms of accuracy and efficacy. The proposed model achieved an average classification accuracy of over 90%.

Alickovic et al. [43] introduced a new model which is based on EEG measurements for automatic seizure onset identification and seizure onset prediction. Two conventional EEG databases, CHB-MIT (scalp EEG) and Freiburg (intracranial EEG), were processed. The proposed model is characterized by four key components: (1) multi-scale principal component analysis for EEG denoising, (2) EEG signal decomposition utilizing time-frequency techniques, (3) statistical values of decomposed sub-bands, and (4) machine learning techniques. In ictal versus inter-ictal EEG, the developed scheme yielded an overall accuracy of 100% for both databases.

A seizure prediction model can detect seizures before they occur and enable clinicians to treat patients with epilepsy on time. Studies on the prediction of seizures have evolved from analysis of signal processing to machine learning. Since numerous prediction techniques are hand-crafted and need high computational complexity, it is difficult to make predictions in real-time.

Wei et al. [44] transformed the EEG time series into two-dimensional multichannel fusion images. To create a deep spatiotemporal learning model for predicting epileptic seizures, a long-term recurrent convolutional network (LRCN) was developed. The CNN block was utilized for deep features extraction from the EEG data in an automated manner. In learning a time series to classify the pre-ictal segments, the LSTM block was integrated into the learning. In the seizure prediction model, new network set-up and a post-processing technique were suggested. The model of deep seizure prediction achieved 93.40% accuracy, 91.88% sensitivity and 86.13% specificity in segment-based evaluations. One hundred sixty-four seizures were predicted for the tests based on the cases. The developed technique provides high sensitivity and a low false prediction (FPR) rate of 0.04 FP/h.

Patients also consider the most troublesome aspect of epilepsy to be the unpredictability of seizures. Hence Liu et al. [45] established an accurate indicator of seizures to alert patients with epilepsy for the forthcoming seizures. They used time and frequency domain features to deliver the same data source with two different views utilizing a multi-view, CNN architecture to predict the occurrence of epileptic seizures. The proposed deep learning model attained an average area of 0.82 and 0.89 area under the curve (AUCs) on two subjects of the CHB-MIT scalp EEG data set.

An effective epileptic seizure prediction framework mitigates the issue and reduces symptoms in a patient's life with epilepsy. Rukhsar et al. [46] developed a model that can predict seizures well ahead of their occurrence. Multivariate statistical process control was used in the long-term scalp EEG signal for seizure predictions. Excessive neuronal activity in the pre-ictal seizure phase has been observed to shift the electrical characteristics of the behavior from chaotic to rhythmic. Employing multivariate statistical process control, commonly known as an anomaly monitoring tool, eight temporal dependent features are utilized to predict the seizures. In total, 90 CHB-MIT EEG seizures of 10 patients are investigated. The findings of the suggested approach showed that 80 out of 90 seizures were precisely predicted before the onset of the seizure, resulting in a sensitivity of 88.89%. The false-positive rate is 0.39 per hour.

4.1 Biomedical Signal Denoising

In raw biomedical signals, noise and artefact detection is a key issue. Filtering these artifacts is required to reduce the impact on the feature extraction. Multiple filtering techniques, such as band-pass filter, wavelet filter, finite impulse response filter, and an adaptive filter, were employed. The analysis is often carried out to normalize the data and to make it compatible with other patient records. Also, in the EEG recording there are several data dropouts or corrupted data due to the limitations of implanted electrodes that lead to lower algorithm performance. Moreover, some outliers can exist in the data due to muscle artefacts and ambient noise. The existence of these outliers affects the extracted features critically [36].

Biomedical signals include various forms of artifacts from internal or external noise interference. These artifacts can be circumvented by using signal denoising methods to filter out artifacts and noise [47]. Preprocessing/denoising is one of the main steps of the analysis of biomedical signals. The key objective of the preprocessing/denoising is to simplify successive processes without losing relevant information by increasing the signal-to-noise ratio (SNR), and to improve the quality of the signal. Researchers utilize these techniques to remove or decrease the undesirable signal components by transforming the signals into another domain. Multiscale Principal Component Analysis (MSPCA) combines the ability of the Principal Component Analysis (PCA) to minimize the relationship between variables with the ability of the wavelet transform to eliminate the relationship between auto-correlated measurements. On the one hand, MSPCA measures the PCA of the wavelet coefficients at each scale with the integration of the effects at appropriate scales. On the other hand, MSPCA is effective, as it requires inputs of events that alter behaviors over time and frequency [38, 48].

4.2 Feature Extraction

All methods for the prediction of seizure include reliable features which are well associated with pre-ictal and inter-ictal levels. These characteristics can be classified as univariate (measures taken separately across every EEG channel) and multivariate (measures taken on two or more EEG channels) according to the number of EEG channels. The categorization of each of these functions is as linear or nonlinear [36].

Feature extraction is one of the main steps in the evaluation of biomedical signals. Thus, useful and informative features of the biomedical signals consisting of several data points can be obtained employing various feature extraction techniques. Such distinguishing and informative parameters define the signal waveform behavior that defines a precise action. Frequencies and amplitudes may reflect the biomedical signaling patterns. Features can be extracted utilizing various feature extraction techniques that is the another signal processing step to simplify the subsequent classification step [49]. Time-frequency (TF) approaches can be utilized to decompose biomedical signals to track changes in time and frequency. [47]. To achieve better efficiency, a smaller number of values that define the informative characteristics of the signals should be taken into consideration. In general, transformation of signals into an informative feature vector is identified as feature extraction. A signal classification framework analyzes the characteristics of a signal, and based on those characteristics, the signal class is determined [50]. Signals can be decomposed into both time and frequency domain using time-frequency methods. In this study Dual Tree Complex Wavelet Transform (DT-CWT) is utilized for features extraction [38].

4.3 Dimension Reduction

Dimension reduction is a technique to reduce the dimension of the original feature vector while preserving the most distinguishing information and eliminating irrelevant information [51]. Numerous features extraction methods generate redundant features. Some types of feature selection/reduction techniques that generate a novel set of features must be implemented to enhance the classifier performance and yield minimum classification error. Several methods used to reduce dimensions and choose features to improve classification accuracy [38, 52].

To interpret the data, the dimensions of biomedical signals must be reduced to obtain more precise results. Small number of parameters are utilized to decrease the dimension of the biomedical signals by various means. In addition, to achieve greater classification accuracy, the features or dimensions must be reduced. For example, the DT-CWT generates wavelet coefficients to designate the signal energy distribution in both time and frequency domains and describes the biomedical signals with a set of wavelet coefficients. A dimension reduction approach must be employed to obtain a smaller number of features from wavelet coefficients, since wavelet-based feature extraction methods include the feature vector, which is too large in size to be used as a classifier input.

Different techniques such as higher-order statistics, entropies and Lyapunov exponents have been used recently for dimension reduction. First, second, third and fourth-order statistics of wavelet decomposition sub-bands can be used for dimension reduction [38]. In this chapter, six statistical features are utilized namely mean absolute values of the coefficients in every sub-band, standard deviation of the coefficients in every sub-band, average power of the coefficients in every sub-band, skewness of the signal coefficients in every sub-band, kurtosis of the signal coefficients in every sub-band, ratio of the absolute mean values of coefficients of adjacent sub-bands.

4.4 Prediction and Classification

Biomedical signals can be decomposed into both time and frequency domain using time-frequency methods such as Dual Tree Complex Wavelet Transform (DT-CWT). Deep Learning methods are the result of advances in artificial intelligence research and deliver a raw data processing capability. Moreover, Deep Learning includes multiple layers of computational modules, which work together to process data and yield the decisive outcome. These layers assist in isolating correct features for the product result and their assessment or analysis.

The essence of Deep Learning algorithms is that they involve modular layers designed to use universal algorithms to learn the data [53]. These layers form deep neural networks (DNN) blocks. Widely known DNN structures are Convolutional neural networks (CNN) and RNN [53].

CNN's architecture is similar to that of neuron communication patterns in the brain. A CNN's convolution process is just like a weight filter that extracts the features from multi-dimensional input data. While using the RNNs, logical sequences are found in the input data. The output of the hidden layer passes through the next layer and also feeds back to itself. The current performance is essentially a cumulative knowledge of the present moment and of the past.

The main difference between the RNN and CNN architecture is that only the current input is considered by CNNs, while the current input and the previous input are considered by RNN, i.e. it includes memory logic. On data from the time series, RNN performs marginally better, whereas CNN is good for tasks like classifying images. Across several scientific fields, deep learning architectures have been used, for example across disease prediction [17], clinical imaging [54], computational biology [55] and genomics and proteomics [56].

Deep Learning algorithms are adequate to detect intricate patterns for the classification of the high-dimensional data, particularly EEG data. CNN is a popular neural network for EEG data training since CNN is very efficient in noise reduction [57]. Deep Learning solves problems in several fields. However, there is a powerful relationship between DNN and nervous system research. ANNs are used as a model for computing brain function [58], whereas CNNs are employed for the processing of visual information and hidden CNN layer activations, neuronal activity is considered to be correlated with visual sensory motor processing in associated brain regions. In neuroscience, deep networks have a valuable computational aspect as they are statistical time series models of brain neuronal activity [59]. The present period of progress is driving the study of artificial intelligence techniques inspired by neuroscience [36, 58].

Artificial intelligence techniques utilize feature vectors that are derived for training from conventional methods of signal processing and provide good accuracy, but these techniques cannot predict a generalized model. The existence of noise and artefacts in data makes detection of features very difficult to manage. Hence, generalizing an automated system with loyal output is a challenging issue, particularly when there are limited samples of training available. At the other hand, in epileptic seizure prediction [60], the ability of deep learning algorithms to automatically learn function opens new ways of investigating. Features learned through Deep Learning techniques are more efficient and robust than features created by hand [61]. Troung et al. [62] developed a prediction method based on CNN by introducing a procedure, which can be extended to all patients with limited preprocessing of EEG data [36]. Tsiouris et al. [15] first used LSTM for an epileptic seizure prediction by comparing the efficiency of different LSTM architectures for 5–30 s segment size of randomly selected inputs. The performance of three LSTM architectures was evaluated using EEG segment feature vectors as inputs to LSTM, in which the feature vector includes specific time-domain features and frequency domain features, besides graph theory local and global features [36].

In biomedical engineering, and especially in analyzing biomedical signals, artificial intelligence algorithms are used to solve problems. Researchers achieved real-time prediction and detection. EEG research has considerably improved with the widespread use of mathematical modelling and artificial intelligence techniques. By classifying patterns within the EEG, artificial intelligence approaches have also provided the opportunity to enhance the recognition, making EEG signals useful for recognizing brain disorders and primary diseases. Various experiments have also been carried out on the features of the EEG signals associated with neurological disorders [38, 63].

An application software aimed to monitor seizures consists of two models [64]: (i) an epileptic seizure prediction technique, which produces an alarm when it detects a seizure coming up, (ii) a model, which monitor a seizure occurrence. In addition, one simple seizure warning will provide for a patient to leave risky circumstances or acts, such as playing sports or climbing stairs. The techniques of seizure prediction are based on the EEG feature extraction, measured over a limited time span from a few seconds to a few minutes [38].

4.5 Experimental Data

EEG data is downloaded from CHB-MIT Scalp EEG Database.[1] All EEG signals, acquired at the Boston Children's Hospital, includes EEG signals of intractable seizures of pediatric subjects. To classify their seizures and determine their candidacy for surgery, subjects were tracked for up to several days following the removal of anti-seizure medication. EEG signal recordings include 23 cases, which were collected from 22 subjects (5 males, ages 3–22; and 17 females, ages 1.5–19). With a resolution of 16 bits, all EEG recordings were sampled at 256 Hz. An International 10-20 EEG electrode positions and nomenclature were employed during the signal acquisition [65, 66].

For the extraction of EEG segments from the EEG dataset of each subject, a rectangular window with a length of 2048 samples (=8 s) was employed. Since the time of onset of each seizure was known, a minimum number of rectangular windows has been used to cover all seizure events. When removing ictal segments, no overlapping was introduced. Since any seizure requires at least 5 min of pre-ictal data, in order to generate 1000 pre-ictal segments, 8-second EEG segments were extracted.

Finally, the experimental dataset containing 3000 8-s EEG segments containing three different classes: inter-ictal, pre-ictal and ictal, which allow testing models of seizure detection and prediction were created. As a result, 1000 inter-ictal, 1000 pre-ictal and 1000 ictal seizures segments were produced. For the time between 30 and 15 min preceding to each seizure onset, 8-s long pre-ictal segments were produced.

4.6 Performance Evaluation Measures

The success of the training set is not a good performance measure on an independent test set. A challenging issue is the predicting performance based on limited data. In most realistic limited-data cases, repeated cross-validation method will be encountered. Another matter that is not as simple as it sounds is to compare the output of various artificial intelligence approaches on a particular problem. It is normal for classification problems to calculate the output of a classifier in terms of the error rate. To predict a classifier's output on new data, it is essential to evaluate its error rate on a dataset that has played no role in the prediction model creation. This distinct

---

[1] https://physionet.org/content/chbmit/1.0.0/.

dataset is termed as test set. It is known that both the training data and the test data are characteristic samples of the essential question. The test data may be different in nature from the training data in certain instances. It is critical that the test information is not used to construct the classifier in any way. One or more learning schemes use the training data to come up with classifiers. The validation data is utilized to tune the classifier's parameters. Then the test data is utilized to assess the error rate of the tuned model. It is important to choose each of the three sets independently. To accomplish better performance in the parameter tuning, the validation set should be apart from the training set, and the test set should be apart from the training and validation set to achieve an accurate estimate of the true error rate [67].

When the volume of data for training and testing is small, it is common for one-third of the data to be kept for testing and the other two thirds to be used for training. To mitigate any bias induced, the unique sample chosen for holdout is utilized to repeat the whole procedure with various random training and testing samples several times. A certain amount of the data is chosen randomly for training in each iteration, likely with stratification, and the remainder is used for testing. A simple version, however, produces the base of an imperative statistical method known as cross-validation [67].

The typical method to estimate a learning technique's error rate given a single, fixed data sample is to employ 10-fold cross-validation. The data are randomly divided into 10 portions where each portion is defined roughly the same amounts as in the complete dataset. In turn, each part is kept out, and the learning system is trained on the remaining nine-tenths; then the holdout set measures its error rate. The learning process is therefore performed on separate training sets a total of 10 times. Finally, the 10 error estimates are averaged to produce an overall error estimate [67].

The right classifications are the true positives (TP) and true negatives (TN). If the answer is wrongly predicted as yes (or positive), a false positive (FP) is when it is actually no (negative). If the answer is wrongly predicted as negative, a false negative (FN) is when it is actually positive [67].

The most common approach to compare algorithms is the classification efficiency without concentrating on a class. Thus, the generally utilized empirical measure is the accuracy, which does not differentiate between the number of correct labels of different classes [68]:

$$Accuracy = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{FN} + \mathrm{TN} + \mathrm{FP}} \times 100\% \tag{1}$$

On the contrary, two measures, which independently estimate the efficiency of a classifier for different classes are Sensitivity (2) and Specificity (3).

$$Sensitivity = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \times 100\% \tag{2}$$

$$Specificity = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}} \times 100\% \tag{3}$$

The positive class measured the indicators of preference [68]. Precision is the relationship between true positives and false positives.

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

A relationship between true positive instances, correctly classified and false-negative instances, misclassified instances called as recall.

$$Recall = \frac{TP}{TP + FN} \tag{5}$$

The correct classification of labels within the various classes is distinguished by all three measures. They focus on positive examples. Recall is a function of correctly classified examples (true positives) and misclassified examples (false negatives). Precision is a result of true positives and instances that are misclassified as positive (false positives). The F-measure is a major indicator that supports algorithms with higher sensitivity and challenges algorithms with higher specificity.

$$F{-}measure = 2 \times \frac{Precision \times Recall}{Precision + Recall} = \frac{2 \times TP}{2 \times TP + FP + FN} \tag{6}$$

This predicted percentage is taken into account by a measure termed the Kappa statistic by deducting it from the successes of the predictor. The Kappa statistic is employed to evaluate the agreement between the categorizations of a dataset predicted and observed while correcting for a relationship, which happens by chance. However, it does not take costs into account, as with the simple success rate [67]. The kappa statistic was defined by Cohen [69] as an agreement index and defined as follows,

$$K = \frac{P_0 - P_e}{1 - P_e} \tag{7}$$

where P_0 is observed agreement and P_e measures the agreement expected by chance [70].

4.7 Experimental Results

The objective of this chapter is to establish a segment-based model for EEG signal classification, which can be implemented using artificial intelligence in automatic interval-based seizure prediction or onset detection systems. Shorter EEG segments have also been extracted from the utilized database that can provide an opportunity to build realistic and computationally efficient systems for seizure prediction. The goal is to build a segment-based model for classifying EEG signals using the CHB-MIT EEG database, which can be used to automatic interval-based seizure prediction.

Thus, the contribution of this chapter is to find a better model by comparing three different commonly utilized artificial intelligence techniques, namely ANN, DNN, CNN.

ANN is a learning algorithm that is stimulated by biological neuron networks to solve different classification problems. The best accuracy is achieved with a hidden layer of 128 neurons in the ANN implementation. ADAM is employed as an optimizer and Rectified Linear Unit (ReLU) as an activation function. In addition, in the model, batch normalization is employed.

A DNN is a particular form of ANN with more than 3 layers that implicitly fuses the extraction and classification of features into a signal learning body and creates a decision-making mechanism directly. In the complex field of machine learning, these types of neural networks have achieved tremendous success in recent years. By using the Tensor- Flow library in the Python programming language, DNN is implemented. There is currently no specific technique for constructing an optimized neural network with the correct number of layers and number of neurons for each layer. Hence, by conducting a large range of trials, DNN is constructed empirically. DNN parameters such as the activation function, the number of hidden layers, the number of learning steps and the number of neurons in each layer are configured manually in each trial. The classification performance of the testing set is assessed for each manual configuration. The best classification accuracy is realized with a DNN consists of 3 hidden layers, with 128, 64, and 32 neurons respectively, after the tedious manual process. The ReLU activation function is utilized in all layers of the DNN classifier. The Softmax function is employed in the output layer, and the cross entropy is used as a cost function.

For the classification of EEG signals, a 7-layer deep convolutional neural network was developed. The deep convolutional neural network model architecture includes the classical CNN layers, but the 1D-CNN structure is prevalent. Feature maps, which represent the EEG segments are exposed to convolution process with weights of different sizes in 1D convolutional layers. 1D convolution is carried out in the first layer of the model with 128 vectors on the input features. This layer's activation outputs are normalized for each batch utilizing the batch normalization layer. New feature maps are created in the 1D max pooling layer by taking the maximum values in the region defined on the feature maps obtained from the previous layers. In the next 1D convolution and 1D max pooling layers, these procedures are repeated. The rectifier linear unit (ReLU) is employed as an activation function. The stride 2 is used for all inputs. The developed deep convolutional neural network has a flattened layer on the 5th layer in order to obtain the feature maps from the 4th layer with an appropriate size as an input to the subsequent layers of the network. Multidimensional input feature vectors are transformed into one-dimensional output data by this layer. A dense-connected neural network layer consisting of 64 units is fed by the features taken from the flattened layer. There are 32 units for the last fully connected layer (layer 7). The number of units is determined as 128 and 64 in layer 5 and layer 6 respectively. Layer 5 and 6 activation functions are still ReLU, while the Softmax function is employed as the last layer activation function. The prediction of the class to which the input data belongs is achieved through the use of the Softmax

Table 1 Classification accuracy of ANN, DNN and CNN for epileptic seizure prediction and detection

Performance	ANN	DNN	CNN
Training accuracy (%)	98.59	99.26	100.00
Validation accuracy (%)	99.40	98.95	99.70
Testing accuracy (%)	98.08	98.38	99.09
Inter-ictal (%)	98.00	98.00	98.18
Pre-ictal (%)	98.16	100.00	100.00
Ictal (%)	98.08	97.12	98.72
Precision	0.982	0.984	0.999
Recall	0.981	0.984	0.991
F-measure	0.981	0.984	0.991
Kappa	0.971	0.976	0.986

layer. The cross-entropy function is utilized for loss assessment during the network training. ADAM is employed for network training. Furthermore, in order to prevent overfitting during the learning process, layers 3, 4 and 5 have a dropout parameter. The probability of a dropout is equal to 0.8. The corresponding training set has been utilized for training the model for each scheme, and the test set was used for final performance assessment. 20% of the original training set was randomly taken as the validation set for tuning the hyperparameters while training the network. After finding the optimal parameters, the model was retrained with the learning rate of 0.01 and a momentum value of 0.90. The network was trained with a batch size of 200 and 25 epochs were used in the training process.

Table 1 presents the performance of three architectures in epileptic seizure prediction and detection, for various phases, inter-ictal, pre-ictal and ictal. Features have been extracted from raw EEG signals using dual tree complex wavelet transform.

With an average accuracy of 98.08%, the least effective approach was ANN. CNN classifier achieved a total classification accuracy of 99.09%, which is the best result obtained. DNN is the second-best among the classifiers and reached a total accuracy of 97.12%. Accordingly, all classifiers achieved good performance in F-measure and Kappa statistics. For pre-ictal state, DNN and CNN achieved 98.38% and 99.09% testing accuracy, respectively.

5 Discussion

In the success of the classification model, feature extraction plays an critical role. The Wavelet-based feature extraction technique, a DT-CWT-based approach, has advantages and drawbacks. The computational burden of examining DT-CWT with an increase in the number of potential levels is one of the difficulties of the DT-CWT-based methodology. Deep Learning has been proposed as a promising approach that can also be used to successfully predict and diagnose epileptic seizures. It should be noted that while in some cases, seizure prediction has been difficult, Deep Learning methods seem to work extremely better and excellent at classifying EEG signals [71]. Compared to the direct use of the raw EEG samples as input to the classifier, feature extraction preceding the classification are mostly preferable. The benefit of feature extraction is that it delivers a simpler and easier way to acquire secret information about the frequency quality of the signal or brain areas.

In principle, if the size of the deep network was significantly increased by adding more layers to compensate for the increased input size of the direct supply of the EEG signals, better EEG signal representation would be learned. The computational cost of training larger CNN models is, however, growing [15].

Instead of training the deep networks with the extracted feature vector, implemented seizure prediction approach helps work in more controllable circumstances by basically simulating the employment of a deep CNN. Deep Learning architectures are assessed utilizing EEG data from the CHB-MIT database. In previous research, the extracted features using DT-CWT have already been successfully used for the prediction of epileptic seizures and according to the DT-CWT created a more informative feature space as a result of this research, provides a considerable gain in classification efficiency. These findings show the benefit of deep networks in EEG-based seizure prediction using CNN and DNN [15].

Within a five-to-30 min′ timeframe (pre-ictal segment), seizure prediction is effectively accomplished, which happens right before seizure onset. Thus, before a seizure happens, there is ample time to make clinical interventions. In convolutional layers, the number of kernel coefficients and the number of coefficient weights are very large. Seizures are detected before they occur, but steps cannot be taken until an epileptic seizure happens. Three distinct groups are defined by one technique: inter-ictal, pre-ictal and ictal.

This chapter explored the use of Deep Learning and EEG recordings to predict epileptic seizures. It is proven that CNN's application to given the existence of EEG, which can be modelled as two- or three-dimensional high-dimensionality [19]. Indication of the efficacy of the implemented method is achieved through the evaluation of various performance indicators, such as 99.09% accuracy.

The high computational time during CNN training in a Python implementation is a drawback of the application of CNNs to EEG recordings. Combining EEG signals with the basic CNN algorithm is the most critical part of obtaining these results. This research is one of the first investigations to discuss the issue of epileptic seizure prediction based on EEG modality and deep learning.

The abilities of the artificial intelligence techniques have so far shown in different practical classification problems. Deep Learning techniques improve the generalization accuracy on a wide variety of problems. Consequently, in this chapter, the performance of popular artificial intelligence techniques is evaluated for epileptic seizure prediction and detection. To evaluate the performance of studied artificial intelligence techniques, a comparative study is realized by employing the publicly available EEG data set containing inter-ictal, pre-ictal and ictal states.

The proposed technique, CNN seems to be suitable for epileptic seizure prediction and detection. According to the presented results in the epileptic seizure prediction and detection, we should emphases the following:

- CNN, and DNN may be successfully applied in the epileptic seizure prediction and detection due to their stable and high performance presented in Table 1.
- CNN is the best-applied method with an average accuracy above 99%.
- DNN has achieved the best result with an average accuracy of 98.38%.
- ANN has achieved the lowest performance, with an average accuracy of 98.08%.
- The drawback of deep learning models is a significant increase in computational cost.

6 Conclusion

Artificial intelligence methods are becoming gradually efficient in epileptic seizure prediction and detection. Several technical and practical solutions still needed to solve their full potential. In this chapter, the use of artificial intelligence methods in the epileptic seizure prediction and detection by employing only EEG signals is presented. Moreover, CNN classification framework is presented to distinguish different states related to epilepsy. This neurological disorder results in severe health-related problems. The experimental validation was implemented with the widely used epileptic EEG dataset. Regarding the comparison of deep learning classification algorithms, CNN revealed better accuracy, while the DNN algorithm showed lower performance.

In summary, different artificial intelligence techniques are compared to assist in the early prediction and detection of the epileptic seizure using CHB-MIT Scalp EEG Dataset. CNN method has achieved the best results in the prediction and detection of the epileptic seizure. The proposed model for the automatic prediction and detection of the epileptic seizure includes 4 steps: (1) denoising with MSPCA (2) feature extraction with DT-CWT (3) Dimension reduction by using statistical values of each sub-bands (4) classification with artificial intelligence techniques.

As the amount and complexity of EEG data increases, Deep Learning algorithms are beginning to show their ability to manage the unpredictable existence of EEG signals and open up new possibilities in demanding biomedical applications such as prediction of epileptic seizures. Deep Learning was presented in seizure prediction and proved to be an ideal instrument for the analysis of EEG pre-ictal and ictal signals. In any pre-ictal period tested, the proposed technique was able to deliver accurate seizure prediction that had not been achieved before.

The proposed approach needs to be thoroughly checked with more EEG data in clinical practice, as the CHBMIT database includes mainly pediatric signals. However, the findings of this chapter deliver clear signs of its effectiveness as an important epilepsy prediction tool and its timely intervention for people. Moreover, the presented approach can be used for any type of disease prediction as well.

The limitation of this study lies in the assessment process, as the small amount of pre-ictal EEG segments resulted in using segmented data to minimize the effect of overfitting and instead force the Deep Learning models to extract more generic pre-ictal information from the entire pre-ictal state period. This technique does not allow the average prediction time to be evaluated. A similar seizure prediction system that would separate timely pre-ictal trends in this direction and provide various warnings to decrease seizure onset could help predict seizures more precisely and minimize false predictions and is intended as future work. It applies additional hyperparameters to the model, such as weights, resulting in an improvement in model training time.

References

1. J. Muthuswamy, Biomedical signal analysis, in *Standard Handbook of Biomedical Engineering and Design*, vol. 14, ed. by M. Kutz (McGraw-Hill Education, New York, 2004), pp. 18
2. L.T. Mainardi, A.M. Bianchi, S. Cerutti, Digital biomedical signal acquisition and processing, in *Medical Devices and Systems* (CRC Press, 2006), pp. 49–72
3. S. Palaniappan, R. Awang, *Intelligent Heart Disease Prediction System Using Data Mining Techniques* (2008), pp. 108–115
4. R. Wu, W. Peters, M.W. Morgan, The next generation of clinical decision support: linking evidence to best practice. J. Healthc. Inf. Manag. JHIM **16**(4), 50 (2002)
5. S.D. Culler, M.L. Parchman, M. Przybylski, Factors related to potentially preventable hospitalizations among the elderly. Med. Care, 804–817 (1998)
6. N. Yiannakoulias, D. Schopflocher, L. Svenson, Using administrative data to understand the geography of case ascertainment. Chronic Dis. Can. **30**(1), 20–28 (2009)
7. T. McCormick, C. Rudin, D. Madigan, A hierarchical model for association rule mining of sequential events: an approach to automated medical symptom prediction (2011)
8. E.S. Fisher, D.J. Malenka, J.E. Wennberg, N.P. Roos, Technology assessment using insurance claims: example of prostatectomy. Int. J. Technol. Assess. Health Care **6**(2), 194–202 (1990)
9. M.E. Hossain, A. Khan, M.A. Moni, S. Uddin, Use of electronic health data for disease prediction: a comprehensive literature review. IEEE/ACM Trans. Comput. Biol. Bioinform. (2019)
10. C. Zhang, L. Zhu, C. Xu, R. Lu, PPDP: an efficient and privacy-preserving disease prediction scheme in cloud-based e-Healthcare system. Future Gener. Comput. Syst. **79**, 16–25 (2018)
11. H. Yin, N.K. Jha, A health decision support system for disease diagnosis based on wearable medical sensors and machine learning ensembles. IEEE Trans. Multi-Scale Comput. Syst. **3**(4), 228–241 (2017)
12. G. Bieber, M. Haescher, M. Vahl, Sensor requirements for activity recognition on smart watches (2013), pp. 1–6
13. D. Malathi, R. Logesh, V. Subramaniyaswamy, V. Vijayakumar, A.K. Sangaiah, Hybrid reasoning-based privacy-aware disease prediction support system. Comput. Electr. Eng. **73**, 114–127 (2019)
14. A. Petrosian, D. Prokhorov, R. Homan, R. Dasheiff, D. Wunsch II, Recurrent neural network based prediction of epileptic seizures in intra-and extracranial EEG. Neurocomputing **30**(1–4), 201–218 (2000)
15. K.M. Tsiouris, V.C. Pezoulas, M. Zervakis, S. Konitsiotis, D.D. Koutsouris, D.I. Fotiadis, A long short-term memory deep learning network for the prediction of epileptic seizures using EEG signals. Comput. Biol. Med. **99**, 24–37 (2018)
16. U.R. Acharya, S.L. Oh, Y. Hagiwara, J.H. Tan, H. Adeli, Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals. Comput. Biol. Med. **100**, 270–278 (2018)

17. H. Khan, L. Marcuse, M. Fields, K. Swann, B. Yener, Focal onset seizure prediction using convolutional networks. IEEE Trans. Biomed. Eng. **65**(9), 2109–2118 (2017)
18. R. San-Segundo, M. Gil-Martín, L.F. D'Haro-Enríquez, J.M. Pardo, Classification of epileptic EEG recordings using signal transforms and convolutional neural networks. Comput. Biol. Med. **109**, 148–158 (2019)
19. R. Rosas-Romero et al., Prediction of epileptic seizures with convolutional neural networks and functional near-infrared spectroscopy signals. Comput. Biol. Med. **111**, 103355 (2019)
20. D. Jain, V. Singh, Feature selection and classification systems for chronic disease prediction: a review. Egypt. Inform. J. **19**(3), 179–189 (2018)
21. S. Huda, J. Yearwood, H.F. Jelinek, M.M. Hassan, G. Fortino, M. Buckland, A hybrid feature selection with ensemble classification for imbalanced healthcare data: A case study for brain tumor diagnosis. IEEE Access **4**, 9145–9154 (2016)
22. M. Chen, J. Yang, X. Zhu, X. Wang, M. Liu, J. Song, Smart home 2.0: Innovative smart home system powered by botanical IoT and emotion detection. Mob. Netw. Appl. **22**(6), 1159–1169 (2017)
23. M. Chen, Y. Zhang, M. Qiu, N. Guizani, Y. Hao, SPHA: smart personal health advisor based on deep analytics. IEEE Commun. Mag. **56**(3), 164–169 (2018)
24. K. He, J. Chen, R. Du, Q. Wu, G. Xue, X. Zhang, Deypos: deduplicatable dynamic proof of storage for multi-user environments. IEEE Trans. Comput. **65**(12), 3631–3645 (2016)
25. M. Usama, B. Ahmad, W. Xiao, M.S. Hossain, G. Muhammad, Self-attention based recurrent convolutional neural network for disease prediction using healthcare data. Comput. Methods Programs Biomed. **190**, 105191 (2020)
26. A.K. Verma, S. Pal, S. Kumar, Comparison of skin disease prediction by feature selection using ensemble data mining techniques. Inform. Med. Unlocked **16**, 100202 (2019)
27. E. Alpaydin, *Introduction to Machine Learning* (MIT press, 2014)
28. N. Barakat, A.P. Bradley, M.N.H. Barakat, Intelligible support vector machines for diagnosis of diabetes mellitus. IEEE Trans. Inf. Technol. Biomed. **14**(4), 1114–1120 (2010)
29. A. Ahlemeyer-Stubbe, S. Coleman, *A Practical Guide to Data Mining for Business and Industry* (Wiley, 2014)
30. Y. Bengio, Learning deep architectures for AI. Found. Trends® Mach. Learn. **2**(1), 1–127 (2009)
31. G.E. Hinton, R.R. Salakhutdinov, Reducing the dimensionality of data with neural networks. Science **313**(5786), 504–507 (2006)
32. A. Subasi, *Practical guide for biomedical signals analysis using machine learning techniques, a MATLAB based approach* (Elsevier, First, 2019)
33. N. Fayyaz Khan, M. Kamil, A. Hussain, M. Sajjad, Detection and classification of vehicle-type by using convolution neural network. Presented at the The 4th International Conference on Next Generation Computing 2018 (2018)
34. S.N. Rakhade, F.E. Jensen, Epileptogenesis in the immature brain: emerging mechanisms. Nat. Rev. Neurol. **5**(7), 380 (2009)
35. M.J. England, C.T. Liverman, A.M. Schultz, L.M. Strawbridge, Epilepsy across the spectrum: promoting health and understanding.: a summary of the Institute of Medicine report. Epilepsy Behav. **25**(2), 266–276 (2012)
36. K. Rasheed et al., Machine learning for predicting epileptic seizures using EEG signals: a review. IEEE Rev. Biomed. Eng (2020), pp. 1–1. https://doi.org/10.1109/rbme.2020.3008792
37. L. Sörnmo, P. Laguna, *Bioelectrical Signal Processing in Cardiac and Neurological Applications*, vol. 8 (Academic Press, San Diego, California, 2005)
38. A. Subasi, Biomedical signal analysis and its usage in healthcare, in *Biomedical Engineering and its Applications in Healthcare* (Springer, 2019), pp. 423–452
39. M.M.N. Mannan, M.A. Kamran, M.Y. Jeong, Identification and removal of physiological artifacts from electroencephalogram signals: A review. IEEE Access **6**, 30630–30652 (2018)
40. H. Chu, C.K. Chung, W. Jeong, K.-H. Cho, Predicting epileptic seizures from scalp EEG based on attractor state analysis. Comput. Methods Programs Biomed. **143**, 75–87 (2017)

41. N.D. Truong et al., Convolutional neural networks for seizure prediction using intracranial and scalp electroencephalogram. Neural Netw. **105**, 104–111 (2018)
42. Y. Gao, B. Gao, Q. Chen, J. Liu, Y. Zhang, Deep convolutional neural network-based epileptic electroencephalogram (EEG) signal classification. Front. Neurol. **11**, 375 (2020). https://doi.org/10.3389/fneur.2020.00375
43. E. Alickovic, J. Kevric, A. Subasi, Performance evaluation of empirical mode decomposition, discrete wavelet transform, and wavelet packed decomposition for automated epileptic seizure detection and prediction. Biomed. Signal Process. Control **39**, 94–102 (2018)
44. X. Wei, L. Zhou, Z. Zhang, Z. Chen, Y. Zhou, Early prediction of epileptic seizures using a long-term recurrent convolutional network. J. Neurosci. Methods **327**, 108395 (2019)
45. C.-L. Liu, B. Xiao, W.-H. Hsaio, V.S. Tseng, Epileptic Seizure prediction with multi-view convolutional neural networks. IEEE Access **7**, 170352–170361 (2019)
46. S. Rukhsar, Y. Khan, O. Farooq, M. Sarfraz, A. Khan, Patient-specific epileptic seizure prediction in long-term scalp eeg signal using multivariate statistical process control. IRBM **40**(6), 320–331 (2019)
47. S. Sanei, *Adaptive Processing of Brain Signals*. The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, (John Wiley & Sons, UK, 2013)
48. B.R. Bakshi, Multiscale PCA with application to multivariate statistical process monitoring. AIChE J. **44**(7), 1596–1610 (1998)
49. B. Graimann, B. Allison, G. Pfurtscheller, Brain–computer interfaces: a gentle introduction, in *Brain-Computer Interfaces* (Springer, 2009), pp. 1–27
50. S. Siuly, Y. Li, Y. Zhang, *EEG Signal Analysis and Classification* (Springer, 2016)
51. A. Phinyomark, F. Quaine, S. Charbonnier, C. Serviere, F. Tarpin-Bernard, Y. Laurillau, EMG feature evaluation for improving myoelectric pattern recognition robustness. Expert Syst. Appl. **40**(12), 4832–4840 (2013)
52. A. Wołczowski, R. Zdunek, Electromyography and mechanomyography signal recognition: Experimental analysis using multi-way array decomposition methods. Biocybern. Biomed. Eng. **37**(1), 103–113 (2017). https://doi.org/10.1016/j.bbe.2016.09.004
53. Y. LeCun, Y. Bengio, G. Hinton, Deep learning. Nature **521**(7553), 436–444 (2015)
54. J.-G. Lee et al., Deep learning in medical imaging: general overview. Korean J. Radiol. **18**(4), 570–584 (2017)
55. C. Angermueller, T. Pärnamaa, L. Parts, O. Stegle, Deep learning for computational biology. Mol. Syst. Biol. **12**(7), 878 (2016)
56. E. Asgari, M.R. Mofrad, Continuous distributed representation of biological sequences for deep proteomics and genomics. PLoS ONE **10**(11), e0141287 (2015)
57. G. Li, C.H. Lee, J.J. Jung, Y.C. Youn, D. Camacho, Deep learning for EEG data analytics: a survey. Concurr. Comput. Pract. Exp. e5199 (2019)
58. D. Hassabis, D. Kumaran, C. Summerfield, M. Botvinick, Neuroscience-inspired artificial intelligence. Neuron **95**(2), 245–258 (2017)
59. V. Jain, H.S. Seung, S.C. Turaga, Machines that learn to segment images: a crucial technology for connectomics. Curr. Opin. Neurobiol. **20**(5), 653–666 (2010)
60. I. Kiral-Kornek et al., Epileptic seizure prediction using big data and deep learning: toward a mobile system. EBioMedicine **27**, 103–111 (2018)
61. A. Antoniades, L. Spyrou, C.C. Took, S. Sanei, *Deep Learning for Epileptic Intracranial EEG Data* (2016), pp. 1–6
62. N.D. Truong, A.D. Nguyen, L. Kuhlmann, M.R. Bonyadi, J. Yang, O. Kavehei, A generalised seizure prediction with convolutional neural networks for intracranial and scalp electroencephalogram data analysis. ArXiv170701976 (2017)
63. R. Begg, D.T. Lai, M. Palaniswami, *Computational Intelligence in Biomedical Engineering* (CRC Press, 2008)
64. M. Winterhalder, T. Maiwald, H. Voss, R. Aschenbrenner-Scheibe, J. Timmer, A. Schulze-Bonhage, The seizure prediction characteristic: a general framework to assess and compare seizure prediction methods. Epilepsy Behav. **4**(3), 318–325 (2003)

65. A.H. Shoeb, *Application of Machine Learning to Epileptic Seizure Onset Detection and Treatment* (2009)
66. A.L. Goldberger et al., PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. Circulation, **101**(23), Art. no. 23 (2000)
67. M. Hall, I. Witten, E. Frank, Data mining: Practical machine learning tools and techniques. Kaufmann Burlingt. (2011)
68. M. Sokolova, N. Japkowicz, S. Szpakowicz, *Beyond Accuracy, F-score and ROC: A Family of Discriminant Measures for Performance Evaluation* (2006), pp. 1015–1021
69. J. Cohen, A coefficient of agreement for nominal scales. Educ. Psychol. Meas. **20**(1), 37–46 (1960)
70. Z. Yang, M. Zhou, Kappa statistic for clustered physician–patients polytomous data. Comput. Stat. Data Anal. **87**, 1–17 (2015)
71. N. LaPierre, C.J.-T. Ju, G. Zhou, W. Wang, MetaPheno: a critical evaluation of deep learning and machine learning in metagenome-based disease prediction. Methods **166**, 74–82 (2019)

Author Queries

Chapter 14

Query Refs.	Details Required	Author’s response
AQ1	Please check the edit made in the sentence ‘Across several scientific fields …’, and correct if necessary.	

MARKED PROOF

Please correct and return this set

Please use the proof correction marks shown below for all alterations and corrections. If you wish to return your proof by fax you should ensure that all amendments are written clearly in dark ink and are made well within the page margins.

Instruction to printer	*Textual mark*	*Marginal mark*
Leave unchanged	• • • under matter to remain	Ⓥ
Insert in text the matter indicated in the margin	⋏	New matter followed by ⋏ or ⋏ⓧ
Delete	/ through single character, rule or underline or ⊢——⊣ through all characters to be deleted	ꝺ or ꝺⓧ
Substitute character or substitute part of one or more word(s)	/ through letter or ⊢——⊣ through characters	new character / or new characters /
Change to italics	— under matter to be changed	⌣
Change to capitals	≡ under matter to be changed	≡
Change to small capitals	= under matter to be changed	=
Change to bold type	∿ under matter to be changed	∿
Change to bold italic	≂ under matter to be changed	≂
Change to lower case	Encircle matter to be changed	≢
Change italic to upright type	(As above)	⊹
Change bold to non-bold type	(As above)	∿⊹∿
Insert 'superior' character	/ through character or ⋏ where required	⋎ or ⋎⋏ under character e.g. ⋎² or ⋎⋏²
Insert 'inferior' character	(As above)	⋏ over character e.g. ⋏₂
Insert full stop	(As above)	⊙
Insert comma	(As above)	,
Insert single quotation marks	(As above)	⋎' or ⋎⋏' and/or ⋎' or ⋎⋏'
Insert double quotation marks	(As above)	⋎" or ⋎⋏" and/or ⋎" or ⋎⋏"
Insert hyphen	(As above)	⊢⊣
Start new paragraph	_⌐	_⌐
No new paragraph	↩	↩
Transpose	⊔⊓	⊔⊓
Close up	linking ⁀‿ characters	⁀‿
Insert or substitute space between characters or words	/ through character or ⋏ where required	Y
Reduce space between characters or words	| between characters or words affected	⊤