EPILEPTIC SEIZURE DETECTION USING ANN ALGORITHM AND K-FOLD CROSS VALIDATION

PROJECT REPORT

MACHINE LEARNING: CSE4020

Slot: E1

Submitted by

Name	Registration Number
JESSICA SAINI	15BCE0164
HARDIKA GOYAL	15BCE0049

Under the guidance of

Prof. Balakrishnan P

Associate Professor, SCOPE,

VIT University, Vellore.

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

Table of Contents

ABSTRACT:

Epilepsy is the most prevalent neurological disorder in humans after stroke. Recurrent seizure is the main characteristic of the epilepsy. Electroencephalogram (EEG) is the recording of brain electrical activity and it contains valuable information related to the different physiological states of the brain. Thus, EEG is considered an indispensable tool for diagnosing epilepsy in clinic applications. Since epileptic seizures occur irregularly and unpredictably, automatic seizure detection in EEG recordings is highly required. Multiwavelets, which contain several scaling and wavelet functions, offer orthogonality, symmetry and short support simultaneously, which is not possible for scalar wavelet. With these properties, recently multiwaveletshave become promising in signal processing applications. Approximate entropy is a measure that quantifies the complexity or irregularity of the signal. This project presents a method for automatic epileptic seizure detection using ANN,which is verified using cross-validation.

INTRODUCTION

Epilepsy is the most prevalent neurological disorder in humans after stroke. About 40 or 50 million people in the world suffer from epilepsy. Epilepsy is characterized by recurrent seizure in which abnormal electrical activity in the brain causes altered perception or behavior. Patients experience varied symptoms during seizures depending on the location and extent of the affected brain tissue. Epileptic seizures may cause negative physical, psychological and social consequences, including loss of consciousness, injuries and sudden death. Until now, the occurrence of epileptic seizure is unpredictable and the mechanisms behind the seizure are little understood. Thus, efforts towards its diagnosis and treatment are of great importance. Electroencephalogram (EEG) is the recording of electrical activity of the brain. EEG recordings contain valuable information for understanding epilepsy. The detection of seizures occurring in the EEGs is an important component in the diagnosis and treatment of epilepsy. However, usually massive amounts of data are included in EEGs and visual inspection for discriminating EEGs is a time consuming and costly process. Thus, developing automatic seizure detection methods is of great significance for reviewing EEGs.

Based on the literature review, there have been numerous works on applying wavelet transform for epileptic EEG signal classificationand recognition. Among those work, only scalar wavelets(wavelets generated by one scaling function) were used to process the EEG signals. Multiwavelets, which consist of more than one scaling and wavelet functions, recently have attracted more attention because of their significant characteristics. Multiwaveletssimultaneously possess orthogonality, short support, symmetry, and a high order of approximation through vanishing moments, that all of them are important for signal processing application.

Artificial neural networks have been used as the most common classifier for classifying EEGs .ANN is a parallelhighly inter-connected structure consisting of a number of simple, non-linear processing elements. ANN can perform computations at a very high speed if implemented on a dedicated hardware. Because of its adaptive nature, it can adapt itself to learn the knowledge of input signals. Thus, an ANN model is also chosen as the classifier system for this research.

In this paper, the proposed epileptic seizure detection method is validated using cross validation. The method consists of three steps. Initially, multiwavelet transform is used to decompose the EEG signal to several subsignals. Then, the approximate entropy feature is extracted from each sub-signal. Finally, the extracted features are used as input to an artificial neural network, which discriminates the EEGs according to the specified classification problems. To the knowledge of the authors, there is no other work in the literature related to using approximate entropy based on multiwavelet transform as the input to an ANN for automatic epileptic seizure detection in EEGs.

A dataset containing 500 EEG segments is employed. The proposed method is evaluated through two classification problems, while different selection of EEGs from the whole dataset is required for each classification problem. The obtained high accuracies indicate the good classification performance of the proposed method. The rest of the paper is organized as follows. Thebasics of multiwavelet transform theory, approximate entropy and artificial neural network are described. The dataset used in this work and the proposed methodology are described in detail. The evaluation procedure and the experimental results are presented, followed by further discussion on the obtained results. Finally, some conclusions and future work are then given.

BACKGROUND OF THE WORK

An artificial neural network is an information-processing system that is based on simulation of the human cognition process. It consists of many computational neural units connected to each other. In ANNs, knowledge about the problem is distributed through the connection weights of links between neurons. The neural network has to be trained to adjust the connection weights and biases in order to produce the desired mapping. ANNs are widely used in the biomedical field for modeling, data analysis, and diagnostic recognition. The ANN's capability of learning from examples, the ability to reproduce arbitrary non-linear functions of input, and the highly parallel and regular structure make them especially suitable for pattern recognition problems.

The training algorithm is an important part of the ANN model. A good topology can be inefficient if trained by an inappropriate algorithm. A suitable training algorithm has a short training process, while achieving better accuracy. One of the most common training algorithms is Bayesian regularization back-propagation, which is also used in current work. This algorithm updates the weight
and bias values according to Levenberg–Marquardt optimization. It minimizes a combination of squared errors and weights and then determines the correct combination so as to produce a network that generalizes well. The process is also called Bayesian regularization.

DATASET DESCRIPTION

The data described by Andrzejak et al. (2001) is used in current work. The whole dataset consists of five sets (denoted as Z, O, N, F and S), each containing 100 single-channel EEG segments of 23.6 s duration, with a sampling rate of 173.6 Hz. These segments were selected and cut out from continuous multi-channel EEG recordings after visual inspection for artifacts (e.g., due to muscle activity or eye movements). Sets Z and O consisted of segments taken from surface EEG recordings that were carried out on five healthy volunteers using a standardized electrode placement scheme. Volunteers were relaxed in an awake state with eyes open (Z) and eyes closed (O), respectively. Sets N, F and S originated from an EEG archive of presurgical diagnosis. Segments in set F were recorded from the epileptogenic zone, and those in set N from the hippocampal formation of the opposite hemisphere of the brain. While sets N and F contained only activity measured during seizure free intervals, set S only contained seizure activity. All EEG signals were recorded with the same 128-channel amplifier system, using an average common reference. The data were digitized at 173.61 samples per second using 12-bit resolution and they have the spectral bandwidth of the acquisition system, which varies from 0.5 Hz to 85 Hz. In this work, two different classification problems are created from the above dataset in order to verify the performance of our method. In the first problem, two sets are examined, normal and seizure, the normal class includes only set Z while the seizure class includes set S. The notation of the problem is simplified as Z–S. In the second problem, all the EEGs from the dataset are used and they are classified into two different classes: sets Z, O, N and F are included in the non-seizure class and set S in the seizure class, which the notation is simplified as ZONF–S. The second classification problem is more close to the clinical applications.

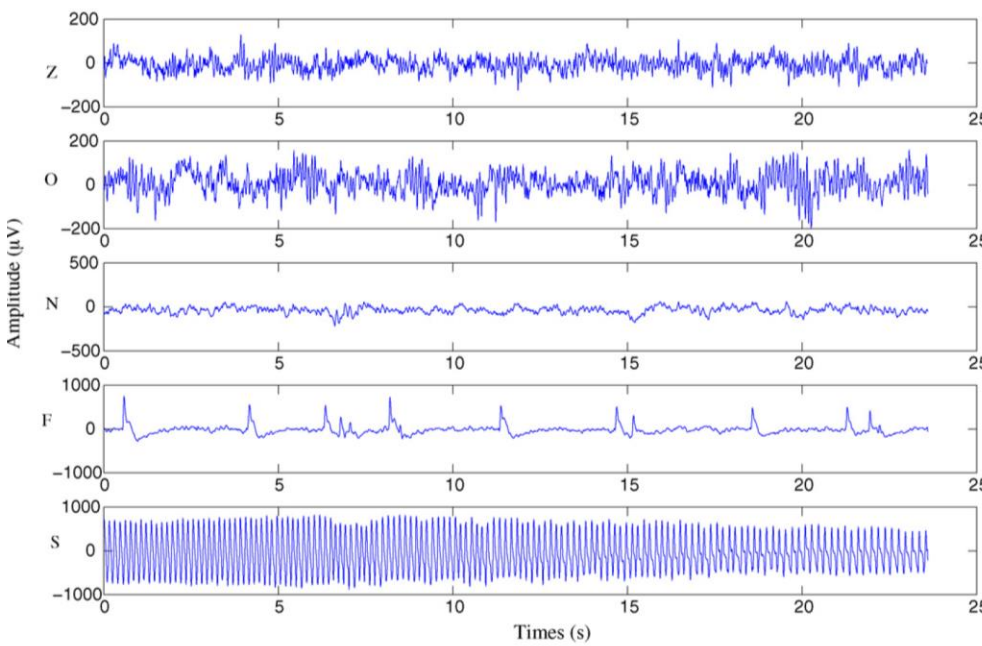


Figure 1: EEG Segments

METHOD

In this work, a novel method based on multiwavelet transform and approximate entropy is proposed for classifying the EEG data into normal/non-seizure and epileptic. The raw EEG data is firstly decomposed into several sub-signals through MWT, then the approximate entropy feature is extracted for each sub-signal to form a feature vector. Finally, the constructed feature vector is put as input to an artificial neural network to classify the EEG into normal/non-seizure and seizure. The block diagram of the proposed approach.

ApEN FEATURE EXTRACTION

Approximate entropy is a statistic parameter that measures the regularity or predictability of a specific time series. It is also known thatApEn possesses good characteristics such as robustness in the characterization of the epileptic patterns. Therefore, in current study ApEn is chosen as the features to discriminate EEGs. ApEn value for each sub-signal of the EEG data decomposed with MWT is calculated to form a feature vector. Before computing ApEn, two important parameters, which are embedding dimension (m) and tolerance window (rr), have to be defined. Based on the suggestions by Pincus (1991), the valuesmandrr are set optimally to 2 and 0.15 times the standard deviation of the data, respectively. After the ApEn values were obtained for 10 sub-signals derived from MWT of the EEG, they construct a feature vector that is used as input to an artificial neural network for classifying EEGs.

ARTIFICIAL NEURAL NETWORK (ANN) CLASSIFICATION

Multi-layer perceptron neural network (MLPNN) is the most widely used ANN structure for classification problems. It has features such as the ability to learn and generalize, smaller training data requirement, fast operation, and easy implementation. In this work, a three-layer MLPNN with Bayesian regularization backpropagation training algorithm is used to classify EEGs based on the previous obtained ApEn feature vector.

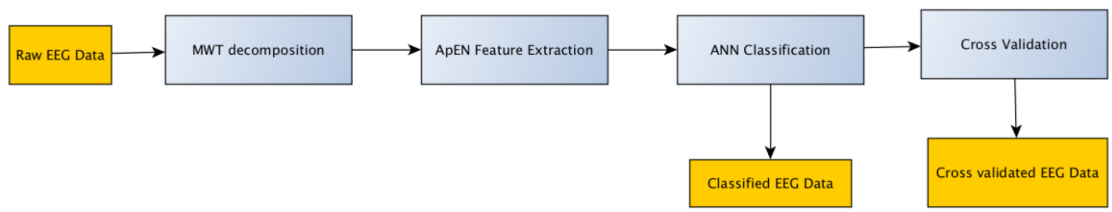
STATISTICAL PARAMETERS

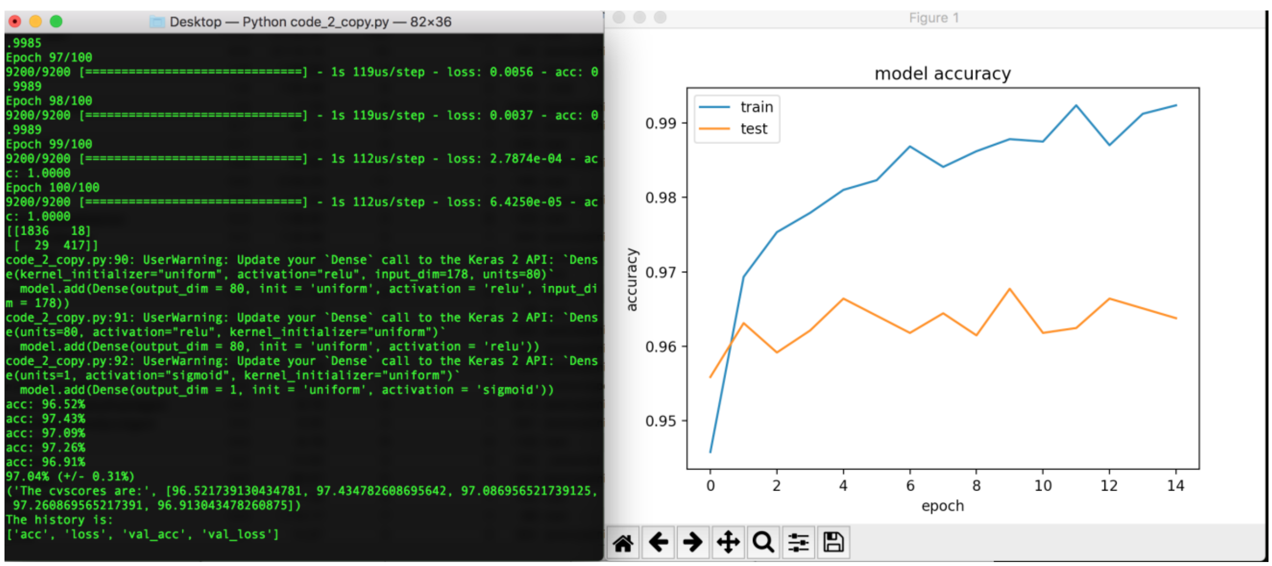
The evaluation of the proposed method on classification problems is determined by computing the statistical parameters of
sensitivity, specificity and classification accuracy. The definitions of these parameters are as:

- Sensitivity: number of correctly detected positive patterns/total number of actual positive patterns. A positive pattern indicates a detected seizure.
- Specificity: number of correctly detected negative patterns/totalnumber of actual negative patterns. A negative pattern indicatesa detected normal/non-seizure.

- Classification accuracy: number of correctly classified patterns/total number of patterns.

RESULTS AND DISCUSSION





Here, we applied 5-fold cross validation to the model by partitioning dataset into five different pairs of training and testing set. We saw that the highest accuracy is obtained in the third iteration i.e. 97.89% while the average accuracy of the model was 97.2%.

We also plotted a graph showing the result obtained from which we can conclude that the model is an over fitted model.

ACKNOWLEDGEMENT

We express our sincere gratitude to Balakrishnan P, Associate Professor, Department of Computer Science & Engineering, Vellore Institue of Technology, Vellore(Tamil Nadu), India, for encouraging and motivating us to work, discover and learn this topic. We are truly thankful to him for guiding us through the entire project and being our mentor.

REFERENCES

[1] Srinivasan, V., Eswaran, C., & Sriraam, N. (2007). Approximate entropy-based epileptic EEG detection using artificial neural networks. *IEEE Transactions on information Technology in Biomedicine*, *11*(3), 288-295.

[2] Guo, L., Rivero, D., & Pazos, A. (2010). Epileptic seizure detection using multiwavelet transform based approximate entropy and artificial neural networks. *Journal of neuroscience methods*, *193*(1), 156-163.

[3] Guo, L., Rivero, D., Dorado, J., Rabunal, J. R., & Pazos, A. (2010). Automatic epileptic seizure detection in EEGs based on line length feature and artificial neural networks. *Journal of neuroscience methods*, *191*(1), 101-109.

[4] Baldassano, S. N., Brinkmann, B. H., Ung, H., Blevins, T., Conrad, E. C., Leyde, K., ... & Litt, B. (2017). Crowdsourcing seizure detection: algorithm development and validation on human implanted device recordings. *Brain*, *140*(6), 1680-1691.

[5] Michael Hills. Seizure detection using FFT, temporal and spectral correlation coefficients, eigenvalues and Random Forest, UPenn and Mayo Clinic's Seizure Detection Challenge.