Epileptic Seizure Detection

Introduction Of Epileptic Seizures

Epilepsy is a serious brain illness that is an endemic neurological disorder all over the world. It is a clinical result that occurs with abnormal neurological electrical discharging of the brain. Epileptic seizures represent the most common positive signs and symptoms of brain disturbance, and epilepsy is one of the most common primary brain disorders . Vascular causes, traumatic causes, infections and brain abscesses, brain tumors, nutritional deficiencies, pyridoxine deficiency, calcium metabolism disorders are lead causes for epilepsy. In diagnosing epilepsy, research is needed for better understanding of mechanisms causing epileptic disorders. The evaluation and treatment of neurophysiologic disorders are diagnosed with the electroencephalogram [EEG]. EEG is crucial for accurate classification of different forms of epilepsy .

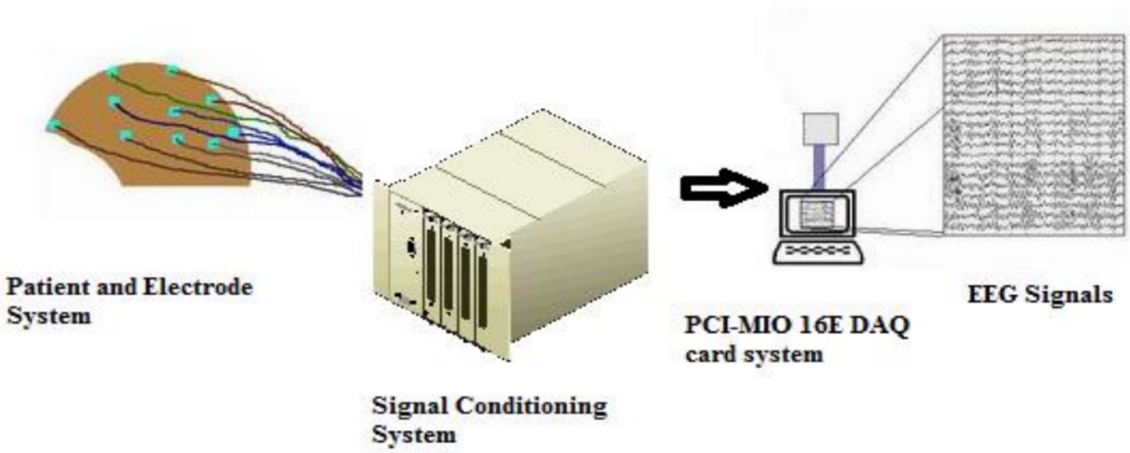
Introduction To This Notebook

The aim of this study is to diagnose epileptic seizures by using different machine learning algorithms. For this purpose, the frequency components of the EEG are extracted by using the discrete wavelet transform (DWT) and parametric methods based on the autoregressive (AR) model. Both these two feature extraction methods are applied to the input of machine learning classification algorithms such as Artificial Neural Networks (ANN), Naive Bayesian, k-Nearest Neighbor (k-NN), Support Vector Machines (SVM), Logistic Regression,Principal Component Analysis.

The results show that k-NN, ANN and SVM were the most efficient method according to test processing of both DWT and AR as feature extraction for recognition of epileptic seizures in EEG.

EEG Data Recording

EEG signals are separated into α, β, δ and θ spectral components and provide a wide range of frequency components. EEG spectrum contains some characteristic waveforms that fall primarily within four frequency bands as follows: δ(0.5-4 Hz), θ(4-8 Hz), α(8-13 Hz), and β (13- 30 Hz) . EEG data set has acquired different age groups in this study. They are known epileptic with uncontrolled seizures and are admitted to the neurology department of the Medical Faculty Hospital of Dicle University1. For this system LabView pro- gramming language has been used and the EEG data used in 400 people who re- ceived 200 of them are epilepsy and with 200 of them are normal. Data set represents signals belonging to several healthy and epileptic patients. The EEG signals are contained by PCI-MIO 16E DAQ card system that provides real time processing and is a data bus of computer, signal processor and personal computer. Fig.shows how to acquire EEG data from a patient [1]. EEG signals are to ensure the accuracy of diagnosing disease that usually takes 8-10 hours in the form of records. EEG sig- nals are used in section and 23.6 seconds, 173 Hz sampling frequency is illustrated with. International 10–20 electrode placement system according to the data collected, 12-bit analog-digital conversion after the samples are recorded subsequently. Data can be passed through the filter 0.53–40 Hz band–pass, the EEG in the presence of clini- cal interest for focusing range is provided. The EEG data used in our study were downloaded from 24-h EEG recorded from both epileptic patients and normal sub- jects. The following bipolar EEG channels were selected for analysis: F7-C3, F8-C4, T5-O1 and T6-O2. In order to assess the performance of the classifier, we selected 500 EEG segments containing spike and wave complex, artifacts and background normal EEG .



Discrete Wavelet Transform

Wavelet transform is a more advantageous spectral analyze method than other spectral analyze methods on non-stationary signals. Because the wavelet transform method changes large low-frequency, high frequency that is narrow for the window size. So, the entire frequency range can be achieved in the optimum time-frequency resolution

Continuous and discrete wavelet transform is analyzed in the scale and variation of parameters due to the continuous wavelet coefficients for each scale is difficult and time consuming. For this reason, discrete wavelet transform is used more often than these non-stationary signals. Wavelet scale is divided into a number of points for the x[n] process as seen in Fig. 2 that is called multi resolution decomposition. It is important that the appropriate wavelet decomposition level is selected, the number of detection and wavelet transform analysis of signals. Because classification accura- cy is dependent on the type of wavelet, dominant frequency components of signals are determined according to the number of decomposition levels. Wavelet coefficients contain important information about EEG signals that provide extraction of feature vectors. Statistical-time frequency of EEG signals sequences are:

The average of the absolute value of coefficients in each sub-band. The maximum absolute value of coefficients in each sub-band. The mean force coefficients of each sub-band. Standard deviation of coefficients in each sub-band. The average absolute value of the ratio of adjacent bands. Distribution of breakdown coefficients in each sub-band.1-3 sequence is signal characteristic; 4-6 sequence is that amount of frequency change. This feature vector of EEG signals that are used as inputs for multi-layer neural network classification.

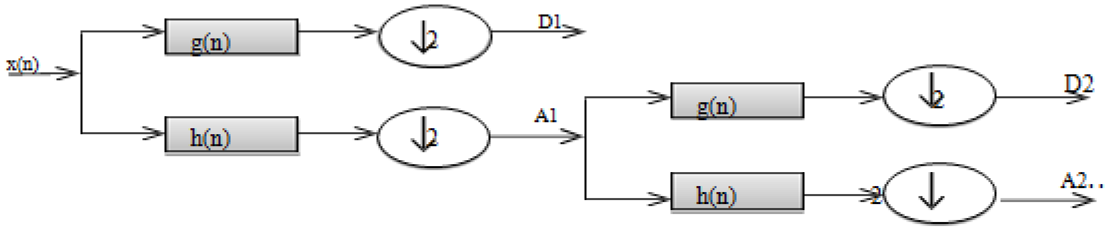


Fig. 2. Realization of discrete wavelet decomposition sub-bands; g[n] is high-pass filter, h[n] is low pass filter [22]

Dataset

- The original dataset from the reference consists of 5 different folders, each with 100 files, with each file representing a single subject/person. Each file is a recording of brain activity for 23.6 seconds.
- The corresponding time-series is sampled into 4097 data points. Each data point is the value of the EEG recording at a different point in time. So we have a total of 500 individuals with 4097 data points for 23.5 seconds.
- We divided and shuffled every 4097 data points into 23 chunks, each chunk contains 178 data points for 1 second, and each data point is the value of the EEG recording at a different point in time.
- So now we have 23 x 500 = 11500 pieces of information(row), each information contains 178 data points for 1 second(column), the last column represents the label y {1,2,3,4,5}.
- The response variable is y in column 179, the Explanatory variables X1, X2, ..., X178

Exploratory Data Analysis

In statistics, exploratory data analysis (EDA) is an approach to analyzing data sets to summarize their main characteristics, often with visual methods. A statistical model can be used or not, but primarily EDA is for seeing what the data can tell us beyond the formal modeling or hypothesis testing task.

Building Machine Learning Models

1. Logistic Regression

Logistic regression, or logit regression, or logit model is a regression model where the dependent variable (DV) is categorical. This article covers the case of a binary dependent variable—that is, where it can take only two values, "0" and "1", which represent outcomes such as pass/fail, win/lose, alive/dead or healthy/sick. Cases where the dependent variable has more than two outcome categories may be analyzed in multinomial logistic regression, or, if the multiple categories are ordered, in ordinal logistic regression.

2.Support Vector Machine (SVM)

Support Vector Machine (SVM) model is a Supervised Learning model used for classification and regression analysis. It is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall.

In addition to performing linear classification, SVMs can efficiently perform a non-linear classification using what is called the kernel trick, implicitly mapping their inputs into high-dimensional feature spaces. Suppose some given data points each belong to one of two classes, and the goal is to decide which class a new data point will be in. In the case of support vector machines, a data point is viewed as a p -dimensional vector (a list of p numbers), and we want to know whether we can separate such points with a $(p-1)$ -dimensional hyperplane.

When data are not labeled, supervised learning is not possible, and an unsupervised learning approach is required, which attempts to find natural clustering of the data to groups, and then map new data to these formed groups. The clustering algorithm which provides an improvement to the support vector machines is called support vector clustering and is often used in industrial applications either when data are not labeled or when only some data are labeled as a preprocessing for a classification pass.

3.*k*-Nearest Neighbors

k -nearest neighbors algorithm (k-NN) is one of the simplest machine learning algorithms and is used for classification and regression. In both cases, the input consists of the k closest training examples in the feature space. The output depends on whether k -NN is used for classification or regression:

- In k -NN classification, the output is a class membership. An object is classified by a majority vote of its neighbors, with the object being assigned to the class most common among its k nearest neighbors (k is a positive integer, typically small). If $k=1$, then the object is simply assigned to the class of that single nearest neighbor.
- In k -NN regression, the output is the property value for the object. This value is the average of the values of its k nearest neighbors.

4. Gaussian Naive Bayes

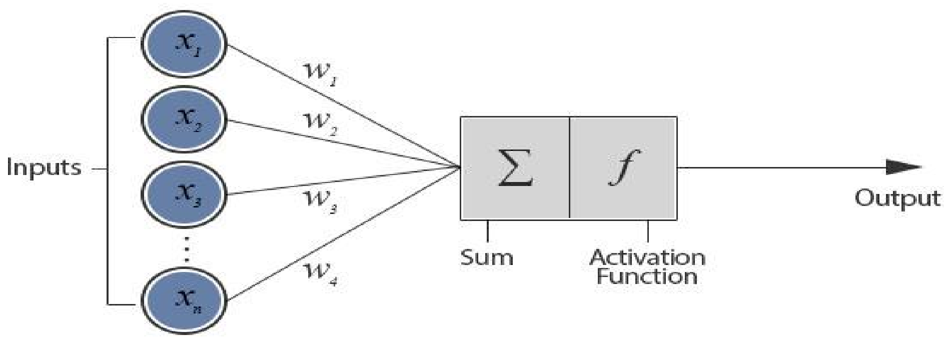
Naive Bayes classifiers are a family of simple probabilistic classifiers based on applying Bayes' theorem with strong (naive) independence assumptions between the features.
Bayes' theorem (alternatively Bayes' law or Bayes' rule) describes the probability of an event, based on prior knowledge of conditions that might be related to the event. For example, if cancer is related to age, then, using Bayes' theorem, a person's age can be used to more accurately assess the probability that they have cancer, compared to the assessment of the probability of cancer made without knowledge of the person's age.
Naive Bayes is a simple technique for constructing classifiers: models that assign class labels to problem instances, represented as vectors of feature values, where the class labels are drawn from some finite set. It is not a single algorithm for training such classifiers, but a family of algorithms based on a common principle: all naive Bayes classifiers assume that the value of a particular feature is independent of the value of any other feature, given the class variable. For example, a fruit may be considered to be an apple if it is red, round, and about 10 cm in diameter. A naive Bayes classifier considers each of these features to contribute independently to the probability that this fruit is an apple, regardless of any possible correlations between the color, roundness, and diameter features.

5. Artificial Neural Networks(ANN)

Computers are great at solving algorithmic and math problems, but often the world can't easily be defined with a mathematical algorithm. Facial recognition and language processing are a couple of examples of problems that can't easily be quantified into an algorithm, however these tasks are trivial to humans. The key to Artificial Neural Networks is that their design enables them to process information in a similar way to our own biological brains, by drawing inspiration from how our own nervous system functions. This makes them useful tools for solving problems like facial recognition, which our biological brains can do easily.

Modeling Artificial Neurons

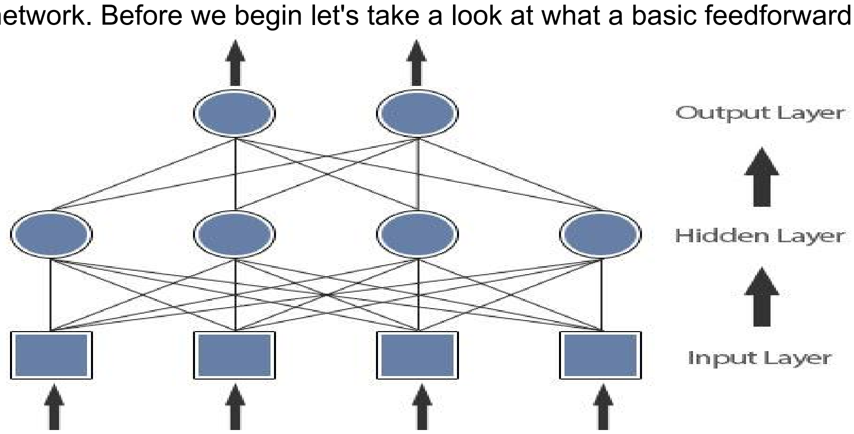
Artificial neuron models are at their core simplified models based on biological neurons. This allows them to capture the essence of how a biological neuron functions. We usually refer to these artificial neurons as 'perceptrons'. So now let's take a look at what a perceptron looks like.



As shown in the diagram above a typical perceptron will have many inputs and these inputs are all individually weighted. The perceptron weights can either amplify or deamplify the original input signal. For example, if the input is 1 and the input's weight is 0.2 the input will be decreased to 0.2. These weighted signals are then added together and passed into the activation function. The activation function is used to convert the input into a more useful output. There are many different types of activation function but one of the simplest would be step function. A step function will typically output a 1 if the input is higher than a certain threshold, otherwise its output will be 0.

Implementing Artificial Neural Networks

So now you're probably wondering what an artificial neural network looks like and how it uses these artificial neurons to process information. In this tutorial we're going to be looking at feedforward networks and how their design links our perceptron together creating a functioning artificial neural network. Before we begin let's take a look at what a basic feedforward network looks like:



Each input from the input layer is fed up to each node in the hidden layer, and from there to each node on the output layer. We should note that there can be any number of nodes per layer and there are usually multiple hidden layers to pass through before ultimately reaching the output layer. Choosing the right number of nodes and layers is important later on when optimizing the neural network to work well on a given problem. As you can probably tell from the diagram, it's called a feedforward network because of how the signals are passed through the layers of the neural network in a single direction. These aren't the only type of neural network though. There are also feedback networks where its architecture allows signals to travel in both directions.

Principal Component Analysis (PCA)

Principal Component Analysis (PCA)is a dimension-reduction tool that can be used to reduce a large set of variables to a small set that still contains most of the information in the large set.
How does PCA work -

1. Calculate the covariance matrix X of data points.
2. Calculate eigenvectors and corresponding eigenvalues.
3. Sort the eigenvectors according to their eigenvalues in decreasing order.
4. Choose first k eigenvectors and that will be the new k dimensions.
5. Transform the original n dimensional data points into k dimensions.

Conclusion

	Model	Score
1	Support Vector Machines	98.08
2	ANN	95.78
4	Naive Bayes	95.78
3	KNN	93.85
5	Principal Component Analysis	91.00
0	Logistic Regression	82.72

Epileptic Seizure Recognition Research Paper

https://iwbbio.ugr.es/2014/papers/IWBBIO_2014_paper_1.pdf

Seizure Recognition and First Aid