# Statistical Feature Extraction and Classification using Machine Learning Techniques in Brain-Computer Interface

**M.Kanimozhi, R.Roselin**

***Abstract: Brain Computer Interface is a paralyzed system. This system is used for direct communication between brain nerves and computer devices. BCI is an imagery movement of the patients who are all unable to communicate with the people. In EEG signals feature extraction plays an important role. Statistical based features are essential feature being used in machine learning applications. Researchers mainly focus on the filters and feature extraction techniques. In this paper data are collected from the BCI Competition III dataset 1a. Statistical features like minimum, maximum, standard deviation, variance, skewnesss, kurtosis, root mean square, average, energy, contrast, correlation and Homogeneity are extracted. Classification is done using machine learning techniques such as Support Vector Machine, Artificial Neural Network and K-Nearest Neighbor. In the proposed system 90.6% accuracy is achieved.***

***Keywords: BCI data, SVM, Neural Network, K-NN, mean min, max, standard deviation, variance, Kurtosis, Skewness, and Root mean Square***

## I. INTRODUCTION

Brain is the most important part in human body. Millions of neuron are interconnected and playing an important role for controlling the human behavior with Internal/external/motor/sensory. Neuron carries the information between brain and human body. Behavior of human can be visualized in team of motor and sensory. Such as eye blinking, hand movement, lip movement, etc.which is related to specific signal frequency help to understand the complex functional behavior of brain. According to Doctors' perception left side brain is actively controlling the right side of the human body and similarly right side brain is controlling the left side human body. EEG (Electroencephalogram) which helps to acquire brain Signal in various states from the scalp surface area. With various frequency levels 0 to 100Hz. Some signal is more than the 100Hz frequency. Medical researchers are analyzing the signal with different perception. In the history of medicine, it is well known that some action or activity is controlling the particular part of brain. Brain functionality and behavior is associated with the place where the electrode is placed in the

brain. Table 1 describes that cranial nerve and functionality [1].

**Revised Manuscript Received on January 5, 2020**

* Correspondence Author

**M.Kanimozhi,** Ph.D research Scholar of Computer Science, Sri Sarada College for Women (Autonomous), Salem - 16, India.

**Dr.R.Roselin**, Associate Professor of Computer Science, Sri Sarada College for Women (Autonomous), Salem - 16, India.

| Name | Function |
| --- | --- |
| Olfactory | Smell |
| OPtic | Vision |
| oculomotor | Eye movement |
| Vestibulocochlear | Hearingand balance |
| Hypogloaasal | Tongue Movement |
| Glossopharyngeal | Taste |

**Table1 1: Cranial Nerve and Functionality**

Signal frequency bands are within four frequency namely Delta, Alpha, theta and Beta and corresponding frequency range is (0-4 HZ), (8-12Hz), (4-7Hz), (13-35Hz). The advantages of using EEG technique are: Most of the other EEG technique are bulky and requires heavy equipment but in this method there is less hardware cost significantly as compared to other technique. These sensors can be used in every field other than medical research where MRI, SPECT, MRS and other techniques are being used. These techniques require large equipment and handling cost and are very costly and bulky. These methods have very high temporal resolution, on the order of milliseconds as compare to other techniques. It is commonly recorded at sampling rates between 250 and 2000 Hz in medical research, but advanced EEG data gathering systems are capable of recording at sampling rates more than this sampling rate if desired we can record and unlike other technique which are non-invasive but on the other hand they provides better resolution.

This paper is organized as follows: Section II gives out the related works Section III explains proposed methodology, Section IV elucidates the results and discussions and Section V concludes this paper with scope for further research.

## II. RELATED WORKS

There are lots of researches in literature which is done in brain computer interface with the Statistical Features and classification Techniques. Prince Kumar Saini and Maitreyee Dutta[2] has proposed a new technique in which signals are decomposed using DWT and statistical features and frequency domain features are extracted which represent the distribution co efficient in both time and frequency domain. Extracted features are input to the feature selection method using ICA and applied various classification algorithm K-NN,SVM and ANN.Ashwini Nakate and P.D. Bahirgonde [3] has de-noised the signal using PCM and signals are extracted using Discrete

Fourier Transform (DFT) and classified the signal using SVM Algorithm.A. Swapnil., S. S. Mhaskeand C. M. Jadhao [4] has detected an epilepsy disorder using the well suited statistical feature and got a promising result. Damien Coyle, Girijesh Prasad and Thomas M. McGinnity [5] have proposed self-organizing fuzzy neural networks (SOFNNs) to extract the mean squared of the predicted signals in time series data. Finally LDA is applied for classification with 94% accuracy. Nitendra Kumar, Khursheed Alam and Abulhasansiddiqi [6] has examined daubechies wavelet. They decomposed the signal and classified using SVM and ANN. Charles Yaacoub, Georges Mhanna and Sandy Rihana[7] has proposed new genetic algorithm for feature selection. Raw signals are prepared for the feature extraction. Features are extracted like autoregressive and discrete wavelet feature coefficients.GA based feature selection applied for the extracted features and classified using ANN with the 90.5% accuracy.SachinBorse [8] has dealt with EEG signal de-nosing. Wavelet transform is used to extract the signal and ICA reduction algorithm is used to separate the multiple signal from the noise and also from each other.S.M. Imran, M.T.F. Talukdar, S. K. Sakib, N. S. Pathan, and S. A. Fattah [9] have employed DWT to extract the essential statistical energy, entropy, maximum, variance for several window sizes.Principal Component Analysis is used for dimension reduction and K-NN is used for classification purpose with 78.26% accuracy level.PariJahankhani, VassilisKodogiannis and Kenneth Revett [10] have applied wavelet transform for feature extraction. The extracted features are sub band of four frequencies such as min, max, mean, std. Then extracted features are classified using NN.AnshulGarg and Rachna Jain [11] has examined various feature selection in BCI data set. Signals are prepared using low-pass filter. The prepared signals are extracted using DWT, EMD and LWT. Extracted features are selected using PCA, ICA, LDA, and various types of DCT waves. Selected feature are applied for classification of signals which resulted in the accuracy of 92% on test data.Irena Koprinska [12] has examined a various feature selection method applied for the BCI data. The signal has noise. So the noise is reduced and the features are selected using information gain ranking, correlation based ranking and Relief.

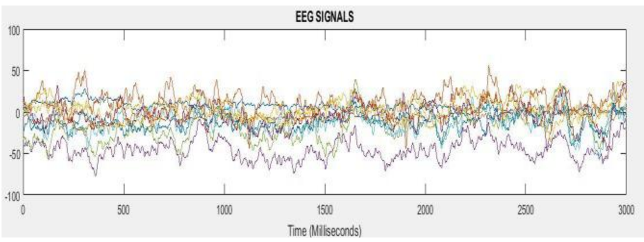



**Figure 1: Graphical representation of EEG Signal**

## III. PROPOSED WORK

The proposed methodology of the BCI signal is illustrated in Figure: 2. Data are taken from the database and Appling feature extraction to derive from the raw EEG signal. The features are Min, Max, Mean, Standard deviation, Variance, skewness, kurtosis, root mean square, energy, contrast, correlation and homogeneity features.

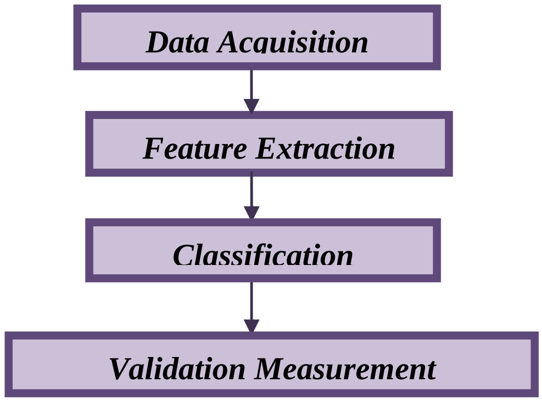



**Figure: 2 Overall Processes**

### A Data Acquisition

Brain Computer Interface (BCI) competition III data set Ia has been taken. Subject asked to do imaginary task such as left small finger or tongue. The data set is basically 3D data. The electrical brain activity is collected during 278 experiments. ECoG platinum electrode which contributed 64 channel pairs from electrodes for every trail. And the frequency level of sample rate is 1000Hz where data are recorded 3 seconds time duration.

### B Feature Extraction

Step 1: The input signal is basically 3 dimension data configuration. So, 3D matrix converts into 2D matrix for the computational cost. Figure1 shows the EEG signal of trail 1 with 10 channels but 278 trails are taken for the purpose of better presentation. In Step 2 the truth full eight statistical features are extracted from the raw signal. Statistical features are Mean, Min, Max, Standard Deviation, Variance, Kurtosis, Skewness and Root mean Square, energy, contrast, correlation, homogeneity are extracted from each channel. In this data set there are 64 channel and 3000 samples. For 64 Channel calculating the 8 statistical features total we extracted 516 features for the each trail. So the extracted feature size is 516× 278.

**B.1 Mean:** The mean of the data set is the arithmetic average of the elements in a data set obtained by adding all the values and dividing it by the number of values. In case of the data if in the form of frequency distribution, then the mean of frequency distribution data can be define as

$$\mu=\sum_{i=0}^{n} a_i\, x_i \qquad (1)$$

**B.2 Variance:**

The variance of data set is the arithmetic average of squared differences between the mean. Again, when we summarize a data set in frequency distribution, the variance of frequency distribution is given by,

$$\text{Variance}=\frac{1}{n}\sum_{i=0}^{n} a_i\ (x_i-\mu) \qquad (2)$$

**B.3 Standard Deviation:**

The standard deviation(STD) of a data set in a frequency distribution can be define by the equation,

$$\text{STD}=\sqrt{\frac{1}{n}\sum_{i=0}^{n} a_i \ (x_i-\mu)} \quad (3)$$

**B.4 Min**
The lowest possible value or the most significant value in each channel is called min value (Min) for each channel.
**B.5 Max:**
The highest possible value or the most significant value in each channel is called maximum value (MAX).
**B.6 Kurtosis:**
It is define as the second measure of peaked distributions of potential values of activity values in each channel.
**B.7 Skewness:**
It is defined as the measure of the similarity or asymmetric distribution. It can be positive as well as negative.
**B.8 Root Mean Square**
Root mean square is function for continues waveform for every channels.

$$\text{Rms}=\sqrt{\frac{\sum_{i=0}^{n} a_i}{n}} \quad (4)$$

**B.9 Energy**

$$\text{Energy}=\sum_{i=1}^{n}\sum_{j=1}^{n} P(I,J)*P(I,J) \quad (5)$$

**B.10 Contrast**

$$\sum_{i,j=1}^{n}\sum_{j=1}^{n} P(i,j)(i-j)^2$$

$$\sum_{i=1}^{n} i^2 \sum_{i,j;|i-j|=n}^{n} P(i,j) \quad (6)$$

**B.11 Correlation**

$$\frac{\sum_{i,j} ijP(i,j)-\mu_x\mu_y}{\sigma_x\sigma_y} \quad (7)$$

$$\frac{\sum_{i,j}(i-\mu_x)(j-\mu_y)-\mu_x\mu_y}{\sigma_x\sigma_y}$$

$$\mu_x=\sum_{i=1}^{n}\sum_{j=1}^{n} P(i,j)$$

$$\mu_y=\sum_{i=1}^{n}\sum_{j=1}^{n} P(i,j)$$

$$\sigma_x^2=\sum_{i=1}^{n}(i-\mu_x)^2\sum_{j=1}^{n} P(i,j)$$

$$\sigma_y^2=\sum_{j=1}^{n}(j-\mu_y)^2\sum_{i=1}^{n} P(i,j)$$

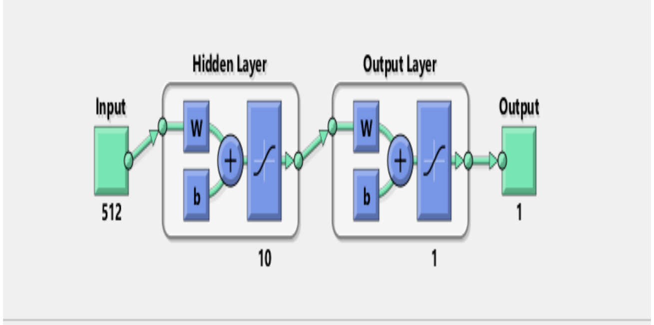
**B.12 Homogeneity**

$$\sum_{i,j}\frac{1}{1+(i-j)^2}P(i,j) \quad (8)$$

**C Classification**

In this work, Machine Learning are artificial neural network three types of support vector machine and K-NN used as classifiers for the EEG signals 60% of the data are trained remaining 40% of the data are tested. Classification done with 10flod cross validation.

**C.1 ANN**
Artificial neural network is an interconnection of the neurons. It simulates the neural structure of the human brain. Basically it is Machine Learning Algorithm. These elements are performing massively parallel computation task for data and pattern recognition. Learning in Neural Network is through the train the data. It is a powerful tool for the decision making, pattern recognition as well as perdition. In this neural network there 10 hidden layers and I output layer.




**C.2 SVM**
SVM is kind of the machine learning algorithm for the supervised learning. It used for the analysis and classified the data. In BCI data there are two type of the class where SVM training a building model for the new set. The set may be one or more set. In SVM classification has linear and non linear kernels for the high dimension of the data. It uses a nonlinear mapping to transform the original training data into a higher dimension. Within this new dimension, it searches for the linear optimal separating hyperplane which means "decision boundary" separating the data of one class from another. With an appropriate nonlinear mapping to a sufficiently high dimension, data from two classes can always be separated by a hyperplane. SVM classifier is not dependent on dimensionality of data. The hyperplane of SVM can be indentified based on the support vectors which help to compute upper bound value on the expected error rate of classifier. Generalization is better when SVM has small number of support vectors. There are three type of the kernel function following
a) Radial Kernel Function (RBF)
b) Linear Kernel Function
c) Polynomial Kernel Function

**C.3 KNN**
K-Nearest Neighbors is simple and easy algorithm. And it is supervised learning. KNN is used for Classification and Regression problem. Initialize K value and find the distance from the current data item.

## IV. RESULT AND DISSUCTION

This experiment is conducted using BCI dataset Ia, implemented in the MATLAB tool for feature extraction and Classification. Extracted feature are classified the two class namely Small finger, tongue. The statistical feature are Mean, Min, Max, Variance, Standard deviation, Kurtosis, Skewness, Root mean Square, energy, contrast, correlationand homogeneity. Twelve statistical features are extracted for each channel that means 64 channels for every trail, totally 516 features are extracted for each trail in the BCI signal. Table 2 shows that sample features of the EEG signal after applying the statistical functions.

**Table 2: Show the sample result of features**

| S.no | Feature1 | Feature2 | | Feature516 |
|---|---|---|---|---|
| 1 | 28.125 | 7.78125 | | 2.83844 3 |
| 2 | 14.71875 | 10.4375 | .......... | 3.249788 |
| 3 | 5.625 | 20.53125 | | 4.030345 |
| 4 | 14.125 | 24.8125 | | 3.241559 |
| 5 | 11.6875 | 8.6875 | | 2.609223 |
| 6 | 15.1875 | 9.21875 | | 3.077414 |
| 7 | 6.90625 | 16.1437 | | 2.89308 |

Figure 3 show the 516 ×278 size of feature signals are plotted in the Matlab tool.

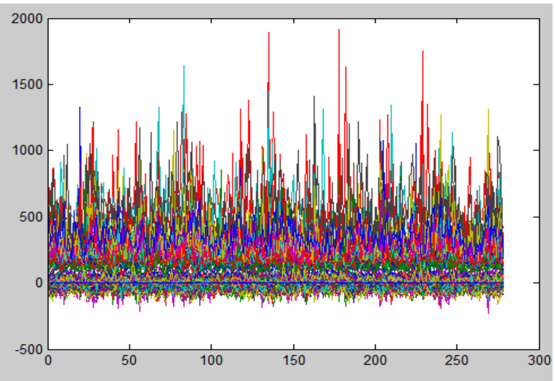



**Figure 3: Graphical representation of the Features**

Table 3 shows the various Machines learning classification with the accuracy measures.

**Table 3: Classification Accuracy**

| Classifiers | Accuracy |
|---|---|
| SVM RFB Kernal | 71.63% |
| SVM Polynomial Kernal | 69.56% |
| SVM Linear Kernal | 85.50% |
| ANN | 90.60% |
| KNN | 79.50% |

ANN gives a better result compared to three SVM classifiers and K-NN. The accurcacy is 90.6% for Artifical Neural Network.

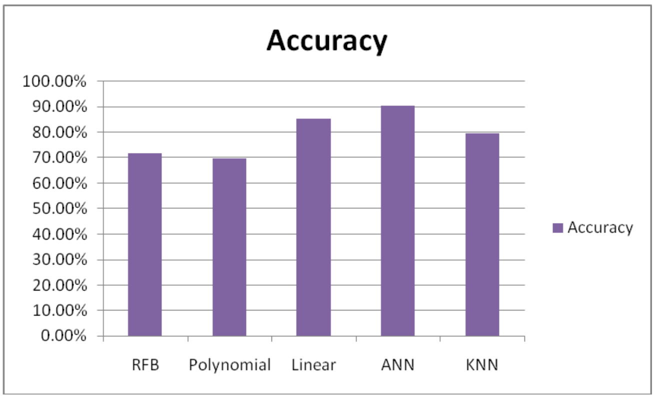



**Figure 4 Graphical representation of Classification Accuracy.**

## V. CONCLUSION

In this paper machine learning algorithms are used for classifying the signal and are implemented for BCI Competition III dataset Ia. In the proposed work 12 fruitful statistical features are extracted for each and every trail from 64 channels in EEG signals. All these features are classified using three types of Support Vector Machine Algorithm (RFB, Linear, and Polynomial), Artificial Neural Network Algorithm and K-NN. ANN gives a better result compared to the SVM classifier. This method successively achieved 90.6% and reduces the time and cost through feature extraction. Further research can be carried out to improve the accuracy in the signal.

## REFERENCE


1. Jsatheeshkumar and P Bhuvaneswari,“ Analysis of Electroencephograpy(EEG) signals and Its Categorization –A Study ”,Elsevier,38,PP : 2525-2536.
2. Prince Kumar Saini, Maitreyee Dutta , “A Hybrid Approach of Feature Extraction and Classification Using EEG Signal ”,International Journal of Innovative Research in Science, Engineering and Technology, vol : 6, No: 8,PP:15595-15603.
3. AshwiniNakate, P.D. Bahirgonde , “Feature Extraction of EEG Signal using Wavelet Transform ”, International Journal of Computer Applications, vol :124 ,No:2 ,PP: 21-24
4. Swapnil A., S. S. Mhaske, Dr. C. M. Jadhao , “ Novel Statistical Feature extraction for detection of epilepsy using EEG signal ”, International Journal of Engineering Science and Computing,vol : 6, No: 4,PP :3387-3391
5. Coyle, Girijesh Prasad and Thomas M. McGinnity , “Extracting Features for a Brain-Computer Interface by Self-Organising Fuzzy Neural Network-based Time Series Prediction ”, IEEE, PP:4371-4374
6. Nitendra Kumar,Khursheed Alam and Abulhasan siddiqi,“Wavelet Transform for Classification of EEG Signal using SVM and ANN”, Biomedical & Pharmacology Journal,Vol:10,No:4,PP: 2061-2069.
7. Charles Yaacoub , Georges Mhanna and Sandy Rihana , “A Genetic-Based Feature Selection Approach in the Identification of Left/Right Hand Motor Imagery for a Brain-Computer Interface”, brain Sciences Article,PP:1-15.
8. Sachin Borse , “EEG De-noising using Wavelet Transform and Fast ICA”, International Journal of Innovative Science, Engineering & Technology, Vol.: 2 ,No: 7, PP:200-205.
9. S.M. Imran, M.T.F. Talukdar, S. K. Sakib, N. S. Pathan, and S. A. Fattah, “Motor Imagery EEG Signal Classification Scheme Based on Wavelet Domain Statistical Features”, IEEE, 2014.
10. PariJahankhani, VassilisKodogiannis and Kenneth Revett,“EEG Signal Classification Using Wavelet Feature Extraction and Neural Networks”, IEEE,2006.

11. AnshulGarg and Rachna Jain ,"Classification of EEG Signals using various Dimensionality Reduction Techniques", International Journal of Computational Intelligence Research,vol:3,No:10,PP: 2309-2331.
12. Irena Koprinska,"Feature Selection for Brain-Computer Interfaces", Springer,2010 ,PP :106-117.

## AUTHOR PROFILE

**M.Kanimozhi,** is currently pursuing Ph. D. at Sri Sarada College for Women (Autonomous), Salem-16. Her area of interest is Data Mining with signal processing.

**R. Roselin,** is currently working as associate professor of computer science in Sri Sarada College for Women (Autonomous), Salem - 16. She has 18 years of teaching experience. Her area of research includes image processing and data mining. Her h-index is 5 and i10 index is 4.