# MULTICLASS SUPPORT VECTOR MACHINE WITH NEW KERNEL FOR EEG CLASSIFICATION

**Mr.A.S.Muthanantha Murugavel,M.E.,M.B.A.,**

*Assistant Professor(SG) ,Department of Information Technology,Dr.Mahalingam college of Engineering and Technology,Pollachi,Tamilnadu,India.*

**D. Akshaya, S. Anitha, M. Manjureka , T. Mohanapriya**

*B.Tech Final year, Information Technology, Dr. Mahalingam College of Engineering and Technology, Pollachi, Tamilnadu, India.*

---------------------------------------------------------------------***---------------------------------------------------------------------

**Abstract -** ***The Electroencephalogram (EEG) is a complex and a periodic time series, which is a sum over a very large number of neuronal membrane potentials. In this Project we have proposed the Multiclass Support Vector Machines with new Kernel (MSVM) for EEG (Electroencephalogram) signals classification problem with hybrid domain features. The Feature Extraction and Classification are performed using the publicly available benchmark datasets. The wavelet transform (WT) is used to extract the time, frequency, wavelet coefficients of discrete time signals. The best among the features extracted are selected using fuzzy logic. Then we classify the selected features using the Multiclass SVM classification technique. Our new classification technique achieves higher classification accuracy and reduces the computational complexity than the existing techniques.***

***Key Words:*** *Electroencephalogram , Fuzzy logic , Wavelet Transform , Multiclass Support Vector Machine*

## 1. INTRODUCTION

The electrical nature of the human nervous system has been recognized for more than a century. It is well known that the variation of the surface potential distribution on the scalp reflects functional activities emerging from the underlying brain. This surface potential variation can be recorded by affixing an array of electrodes to the scalp, and measuring the voltage between pairs of these electrodes, which are then filtered, amplified, and recorded. The resulting data is called the Electroencephalogram (EEG).

In clinical contexts, EEG refers to the recording of the brain's spontaneous electrical activity over a short period of time, usually 20–40 minutes, as recorded from multiple electrodes placed on the scalp. In neurology, the main diagnostic application of EEG is in the case of epilepsy, as epileptic activity can create clear abnormalities on a standard EEG study. A secondary clinical use of EEG is in the diagnosis of coma and encephalopathy. EEG used to be a first-line method for the diagnosis of tumors, stroke and other focal brain disorders, but this use has decreased with the advent of anatomical imaging techniques such as MRI and CT.

Derivatives of the EEG technique include evoked potentials (EP), which involves averaging the EEG activity time-locked to the presentation of a stimulus of some sort (visual, somato sensory, or auditory). Event-related potentials refer to averaged EEG responses that are time-locked to more complex processing of stimuli; this technique is used in cognitive science, cognitive psychology, and psycho physiological research.

### 1.1 NATURE OF EEG

EEG reflects correlated synaptic activity caused by post-synaptic potentials of cortical neurons. The ionic currents involved in the generation of fast action potentials may not contribute greatly to the averaged field potentials representing the EEG. More specifically, the scalp electrical potentials that produce EEG are generally thought to be caused by the extra cellular ionic currents caused by dendritic electrical activity, whereas the fields producing magneto encephalographic signals are associated with intracellular ionic currents.

The electric potentials generated by single neurons are far too small to be picked by EEG or MEG. EEG activity therefore always reflects the summation of the synchronous activity of thousands or millions of neurons that have similar spatial orientation, radial to the scalp. Currents that are tangential to the scalp are not picked up by the EEG. The EEG therefore benefits from the parallel, radial arrangement of apical dendrites in the cortex. Because voltage fields fall off with the fourth power of the radius, activity from deep sources is more difficult to detect than currents near the skull.

Scalp EEG activity shows oscillations at a variety of frequencies. Several of these oscillations have characteristic frequency ranges, spatial distributions and are associated with different states of brain functioning (e.g., waking and the various sleep stages). These oscillations represent synchronized activity over a network of neurons. The neuronal networks underlying some of these oscillations are understood (e.g., the thalamocortical resonance underlying sleep spindles), while many others are not (e.g., the system that generates the posterior basic rhythm)

## 1.2 EPILEPSY

Epilepsy is a neurological condition, which affects the nervous system. Epilepsy is also known as a seizure disorder. It is usually diagnosed after a person has had at least two seizures that were not caused by some known medical condition like alcohol withdrawal or extremely low blood sugar. Sometimes, epilepsy can be diagnosed after one seizure, if a person has a condition that places them at high risk for having another.

Anyone can be affected by seizures. Epilepsy affects approximately 7 per 1000 of the general population around 40 million people globally. The figure for the number of people who may have at least one seizure during their lifetime is even greater - at 5% of the world's population. Epilepsy is more likely to occur in young children or people over the age of 65 year, however it can occur at any time. Epilepsy is usually controlled, but cannot be cured with medication, although surgery may be considered in difficult cases. However, over 30% of people with epilepsy do not have seizure control even with the best available medications. Not all epilepsy syndromes are life long – some forms are confined to particular stages of childhood. Epilepsy should not be understood as a single disorder, but rather as syndromic with vastly divergent symptoms but all involving episodic abnormal electrical activity in the brain. Epilepsy is characterized by recurrent unprovoked seizures.

## 2. DATASET COLLECTION

Here we have used the dataset (digitized EEG signals) for both healthy and epileptic subjects made available online by Dr. Ralph Andrzejak of the Epilepsy Center at the University of Bonn , German (http://www.meb.unibonn.de/epileptologic//physik/eeg dataold.html)

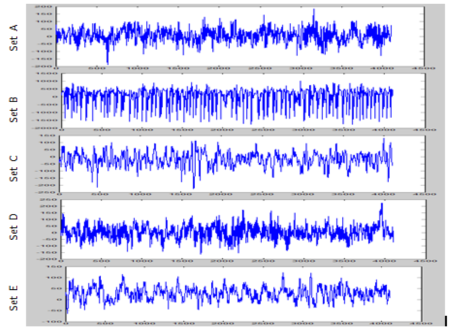


**Fig -1**: Examples of five different sets of EEG signals taken from different subjects

The dataset includes five subsets (denoted as A, B, C, D, and E) each containing 100 single-channel EEG segments of 23.6 s duration. Table 1 lists each class and its description. Sets A and B have been acquired from surface EEG recordings of five healthy volunteers, with eyes open and closed, respectively. Typical EEG segments (one from each category of the dataset) are shown in Figure 4. Signals in subsets C and D have been measured in seizure-free intervals, from five patients in the epileptogenic zone (D) and from the hippocampal formation of the opposite hemisphere of the brain (C). Subset E contains seizure activity, selected from all recording sites exhibiting ictal activity. We used the data described in, which is publicly available.

| Class | Description |
|---|---|
| A | EEG recordings of five healthy volunteers, with eyes open |
| B | EEG recordings of five healthy volunteers, with eyes closed |
| C | Seizure-free intervals, from the hippocampal formation of the opposite hemisphere of the brain |
| D | Seizure-free intervals, from five patients in the epileptogenic zone |
| E | seizure activity, selected from all recording sites exhibiting ictal activity |

Sets A and B have been taken from surface EEG recordings of five healthy volunteers with eyes open and closed, respectively. Signals in two sets have been measured in seizure-free intervals from five patients in the epileptogenic zone (D) and from the hippocampal formation of the opposite hemisphere of the brain (C). Sets A and B have been recorded extra cranially, whereas sets

**Table -1:** Five classes of EEG dataset

C, D, and E have been recorded intracranially. All EEG signals were recorded with the same 128-channel amplifier system, using an average common reference [omitting electrodes containing pathological activity (C, D, and E) or strong eye-movement artifacts (A and B)]. After 12-b analog-to-digital conversion, the data were written continuously onto the disk of a data-acquisition-computer system at a sampling rate of 173.61 Hz. Bandpass filter settings were 0.53–40 Hz (12 dB/oct.).

## 3. FEATURE EXTRACTION

Transforming the input data into the set of features is called feature extraction. The extraction methods are used to reduce the dimensionality of the features. Extracted features represent the characteristics of original signal without redundancy. The features can be extracted from the EEG signal in two different domains such as Time domain features (TDF), Frequency domain features (FDF) and combination of both Time domain and frequency domain features (i.e. Wavelet domain)

### 3.1 Wavelet Transform

Wavelet Transform (WT) is an effective method of time frequency representation of a signal. The attractive feature of WT is that it provides accurate frequency information at low frequencies and accurate time information at high frequencies. This property is important in bio medical applications, because most signals in this field always contain high frequency components with short duration and low frequency components with long time duration.

Abnormalities in the EEG in serious psychiatric disorders are many times too subtle to be detected using conventional techniques, such as Fourier Transform(FT).WT is specific appropriate for analysis of non- stationary signals. It is well suited for locating transient events, which always occur during epileptic seizure.

Wavelet's feature extraction and representation properties can be used to analyze various transient in biological signals. The WT makes use of multi resolution signal analysis technique to decompose EEG signals into a number of frequency bands. The wavelet analysis relies on the introduction of an appropriate basis function and the characterization of signal is reflected through the distribution of amplitude in the basis function. If a proper orthogonal basis function is selected, it has the advantage that a signal can be uniquely decomposed and the decomposition can be inverted.

### Time Domain Features

Time domain analysis process consists of statistical calculations. The time domain features are: Mean, Standard deviation, Maximum and Minimum. These time domain features are calculated for the reconstructed EEG signal amplitude and time duration.

**Mean**

Mean corresponds to the centre of a set of value. The Mean is calculated for each and every sub-band signals.

$$Mean = \frac{1}{N}\sum_{i=1}^{n} x_i$$

**Standard deviation**

Standard deviation is a simple measure of the variability of a data set. The Standard deviation is the root-mean-square (RMS) deviation of its values from the mean.

$$Std = \sqrt{\frac{\sum_{i=1}^{n}(x_i - x)^2}{N-1}}$$

**Maximum and Minimum**

The maximum and minimum values are used to describe the range of observation in the reconstructed signal.

**Trimmed Mean**

A trimmed mean is stated as a mean trimmed by X%, where X is the sum of the percentage of observations removed from both the upper and lower bounds.

**Z-Scores**

A Z-Score is a statistical measurement of a score's relationship to the mean in a group of scores. A Z-score of 0 means the score is the same as the mean. A Z-score can also be positive or negative, indicating whether it is above or below the mean and by how many standard deviations.it is defined as

$$z_i = (x_i - \overline{x})/s$$

**Quantile**

Quantile(X,p) returns quantiles of the values in data vector or matrix X for the cumulative probability or probabilities p in the interval [0,1].

### Frequency Domain Features

The frequency domain features are the power values of each channel from the frequency band. Some of the frequency domain features are Band power, Fractal Dimension and Energy. Band power describes how the power of a signal or time series is distributed with frequency. Fractal dimension is used to approximate dimension of a signal.
Features extracted in frequency domain is one of the best to recognize the mental tasks based on EEG signals.

**Power Spectral Density**

There are several approaches in order to obtain power spectral density, or simply power spectrum. However, the direct approach is the magnitude squared of the Fourier transform of the interested signal, and is equal to

$$PS(f) = |X(f)|^2$$

**Median frequency**

Median frequency is defined as the particular frequency, which di-vides the total area under PS(f) into two parts of equal size.

**Spectral entropy**

Spectral entropy is defined as the same as Shannon entropy with this difference that Pi is the power density of the power spectrum.

$$H_{Sp} = -\sum_{i=f_l}^{f_h} P_i \log P_i,$$

## Wavelet Domain Features

The use of discrete wavelet transform (DWT) both for signal preprocessing and signal segments feature extraction as an alternative to the commonly used discrete Fourier transform(DFT).It includes both time and frequency domain features. The wavelet domain features are:

**Mean Absolute Value:**

Mean Absolute Value (MAV) is similar to average rectified value (ARV). It can be calculated using the moving average of full-wave rectified EEG. In other words, it is calculated by taking the average of the absolute value of EEG signal. It is an easy way for detection of muscle contraction levels. It is defined as

$$MAV = \frac{1}{N}\sum_{i=1}^{N} |x_n|$$

**Variance:**

Variance of EEG (VAR) uses the power of the EEG signal as a feature. Generally, the variance is the mean value of the square of the deviation of that variable. However, mean of EEG signal is close to zero. Hence, variance of EEG can be calculated by

$$VAR = \frac{1}{N-1}\sum_{n=1}^{N} {x_n}^2$$

**Waveform Length:**

Waveform length (WL) is the cumulative length of the waveform over the time segment. WL is related to the waveform amplitude, frequency and time.

## 4. FEATURE SELECTION

**Fuzzy Rule Based Model**

The Extracted parameters Average, Max, Min, Standard Deviation and Variance is considered as input variables to the Fuzzy rule based selection process block. Inference System (FIS) maps an input features to output classes using FL. Fuzzy logic are easy to modify a FIS just by including or excluding rules. The fuzzy rules have written for Extracting Features to get results as Good, Bad and Best data sample values.

**Fuzzy Rules**

Fuzzy rules are linguistic IF-THEN- constructions that have the general form "IF A THEN B" where A and B are propositions containing linguistic variables. A is called the premise and B is the consequence of the rule.

Fuzzy Rule for five input variables and one output variable is defined as few example rules

a) If Average is bad, Maximum is bad, Minimum is bad, Standard Deviation is bad and Variance is bad, the Fuzzy Classification is considered as “BAD”.
b) If Average is good, Maximum is good, Minimum is good, Standard Deviation is good and Variance is good, the Fuzzy Classification is considered as “GOOD”.
c) If Average is best, Maximum is best, Minimum is best, Standard Deviation is best and Variance is best, the Fuzzy Classification is considered as “BEST”. If any one of the Fuzzy Classifier output variable (Bad, Good and Best) is present more number of times in Feature Extracted parameters rule, the Classifier will assign that Fuzzy Classifier output variable to be the final output in the Fuzzy System.

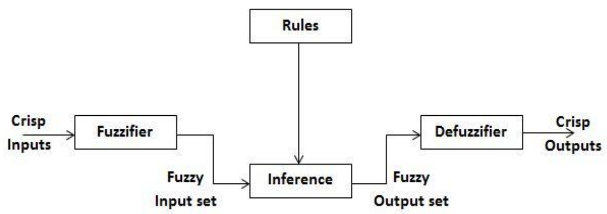


**Fig -2**: A Fuzzy Logic System

## 5. CLASSIFICATION

Multiclass SVM aims to assign labels to instances by using support vector machines, where the labels are drawn from a finite set of several elements.

The dominant approach for doing so is to reduce the single multiclass problem into multiple binary classification problems. Common methods for such reduction include: Building binary classifiers which distinguish between

- one of the labels and the rest (one-versus-all)
- between every pair of classes (one-versus-one).

**One-versus-all**

To get M-class classifiers, construct set of binary classifiers f1, f2, ……,fm, each trained to separate one class from rest. Combine them to get a multi-class classification according to the maximal output before applying the sign function.

**One-versus-one**

It can be used by most of the classifiers. While this involves building classifiers, the time for training classifiers may actually decrease, since the training data set for each classifier is much smaller.

| Features | Accuracy Obtained 50/50 | Accuracy Obtained 100/100 |
|---|---|---|
| Min ,Max ,Mean, Standard Deviation | 83.52 | 99.9 |
| Min, Max, Mean | 81.3 | 99.3 |
| Min, Max, Standard Deviation | 82.38 | 95.94 |
| Mean | 65.63 | 66.083 |
| Standard Deviation | 84.93 | 92.77 |

**KERNEL**

A function that returns the value of the dot product between the images of the two arguments $K(x1, x2)$ = f($x$1), f($x$2). Given a function K, it is possible to verify that it is a kernel. One can use LLMs in a feature space by simply rewriting it in dual representation and replacing dot products with kernels:

$$<x1, x2> \leftarrow K(x1, x2) = <\phi(x1), \phi(x2)>$$

**KERNEL FUNCTIONS**

**Linear Kernel**

The Linear kernel is the simplest kernel function. It is given by the inner product <x, y> plus an optional constant c. Kernel algorithm using a linear kernel are often equivalent to their non-kernel counterparts

$$k(x, y) = x^T y + c$$

**Polynomial Kernel**

The Polynomial kernel is a non-stationary kernel. Polynomial kernels are well suited for problems where all the training data is normalized.

$$k(x, y) = (\alpha x^T y + c)^d$$

Adjustable parameters are the slope alpha, the constant term c and the polynomial degree d.

**RBF Kernel**

A radial basis function network kernel is an artificial neural network kernel type that uses radial basis functions as activation functions. It is a linear combination of radial basis functions. They are used in function approximation, time series prediction, and control.

## 6. RESULT

The result obtained by using HSVM, the proposed classification technique gives higher accuracy when compared with other classification techniques. The comparison is made on the available benchmark datasets as in Table 1.The detail-wavelet coefficients at the first-decomposition level of the five types of EEG signals are different from each other and, therefore, they can serve as useful parameters in discriminating the EEG signals.

In the Table we have taken the datasets A, B, C, D, E and the min, max, mean, standard deviation Features for all types of datasets and the classification accuracy obtained is discussed

**Table -2:** Comparison Results of different combination of features

The Table 3 shows the classification rate of the proposed classification technique over various kernels such as linear, Polynomial, Radial Basis Function and Enhanced Radial Basis Function.

| Kernels | Classification Rate |
|---|---|
| RBF | 83.52 |
| ERBF | 81.82 |
| Polynomial | 80.78 |
| Linear | 79.23 |

**Table -3:** Classification Rate For Various Kernels

## 7. CONCLUSIONS

The Multiclass SVM has shown great performance since it divides the multiclass into several binary class problems. Thus the time complexity reduces to (n-1) for n class problems. The performance of the other neural network was not as high as the Multiclass SVM. This may be attributed to several factors including the training algorithms, the scattered and mixed nature of the features. The results of the present project demonstrated that the Multiclass SVM can be used in the classification of the EEG signals by taking into consideration the misclassification rates. In conclusion, experimental results show that our Multiclass SVM using wavelet based features can well preserve the most discriminant information of EEG signals and improve the performance. In future the work is to achieve still better accuracy by developing a custom (user-defined) kernel. Also the work has to be extended to numerous other classification systems as finger print recognition, ECG classification, cancer diagnostics, etc.,

## REFERENCES

[1] Jie xiang, " the detection of epileptic seizure signals based on fuzzy entropy", elsevier,2015
[2] Varun joshia, ram bilas pachoria, antony vijeshb, "classification of ictal and seizure-free eeg signals using fractional linear prediction",elsevier,2014
[3] Mohammad zavid parvez, manoranjan paul, "epileptic seizure detection by analyzing eeg signals using different transformation techniques", elsevier,2010
[4] Kyong ho lee, "a low-power processor with configurable embedded machine-learning accelerators for high-order and adaptive analysis of medical-sensor signals",ieee,2013
[5] U. Rajendra acharyaa, "automated diagnosis of epileptic eeg using entropies", elsevier,2011
[6] Hasan ocak, "automatic detection of epileptic seizures in eeg using discrete wavelet transform and approximate entropy", elsevier,2009
[7] Varun joshia, ram bilas pachoria, antony vijeshb "classification of ictal and seizure-free eeg signals using fractional linear prediction"
[8] Mohammad zavid parvez n, manoranjan paul, "epileptic seizure detection by analyzing eeg signals using different transformation techniques" elsevier,2008
[9] Shivkumar sabesan, levi b. Good, konstantinos s. Tsakalis, andreas spanias, fellow, david m. Treiman, and leon d. Iasemidis, 'information flow and application to epileptogenic focus localization from intracranial eeg', (2009)
[10] Inan guler and elif derya u beyli 'multiclass support vector machines for eeg-signals classification', (2007)