IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661, p- ISSN: 2278-8727Volume 11, Issue 2 (May. - Jun. 2013), PP 77-84
www.iosrjournals.org

An Approach to Improve Brain Disorder Using Machine Learning Techniques

Swati Sharma1, A Kulothunghun[2]
[1](*Student of SRM University, Computer Science)*
[2](*Assistant Professor of SRM University, Computer Science)*

Abstract: *Machine learning is the important part of the artificial intelligence which plays an important role in medical diagnosis. Attention Deficit Hyperactivity Disorder (ADHD) is one of the most common childhood disorders and continues through adolescence and adulthood. ADHD is a complex brain disorder which is usually difficult to diagnose. The study of different literature reports gives the analytical feature of the disorder as the ADHD has been misdiagnosis with other type of the brain disorders. Here in this report propose a decision system support system in diagnosis of ADHD disorder through brain Electroencephalographic signals. The main goal of this study is to provide different machine learning algorithms that may cure disorder and analysis with different methods of machine learning using MATLAB features. With the input of EEG signals from dataset of patient we apply machine learning techniques like SVM, NAÏVE BAYES, KNN etc and using these algorithms methods, EEG can be decomposed into functionality different components, and this study analysis find the appropriate method is better for brain disorder problem that can reduce the contaminates using MATLAB feature and tools.*

Keywords: *- SVM, KNN, algorithms, EEG signals, supervised learning, neural network, ADHD, feature selection.*

I. Introduction

1.1 WHAT IS ADHD:- ADHD referred as Attention Deficit Hyperactivity Disorder [1] is one of the most common childhood disorder and can be continue through adolescence [2] & adulthood. There are various symptoms that include difficult staying focused & paying attention, difficulty in controlling behavior & hyperactivity. ADHD symptoms change or varies over the time [3] with increasing age, the profile of the symptoms change and difficulties are involves in various functions become more prominent than hyperactivity [4].

There are three types of ADHD that mostly found in children between the ages of 3-6 years.

a) **Predominantly hyperactive impulsive**- children having symptoms of impulsivity may be they may be very impatient, blurt out inappropriate comments, show their emotions without restraint, having difficulty waiting for things they want etc.

b) **Predominantly inattentive**- children having symptoms of inattentive may be easily distracted, miss details, forget things, frequently switch from one activity to another, and become bored with a task after only a few minutes.

c**) Combined hyperactive**-impulsive and inattentive- children who having the symptoms of hyperactivity may be talked nonstop, dash around, touching or playing with anything and everything in sight, be constantly in motion, having difficulty doing easy tasks etc.

It is natural for all children to be inattentive, impulsive, and hyperactive sometimes, but for children who having ADHD, these behaviors are more severe and occur more often. To be diagnosed with the disorder, a child must have symptoms for 6 Months and to degree that is greater than other children of same age. The proper treatment can relief many disorder symptoms with the use of new tools such as brain imaging, supporting diagnostic tools [5] to better understand ADHD & to find effective way to treat it.

1.2 Work of Machine Learning in Artificial Intelligence

The machine learning that usually refers to changes in system and how the system constructs computer programs that automatically improve the experience. The "changes" might be either enhancements to already performing system or synthesize the system. The figure shows how the machine learning works in Artificial Intelligence.

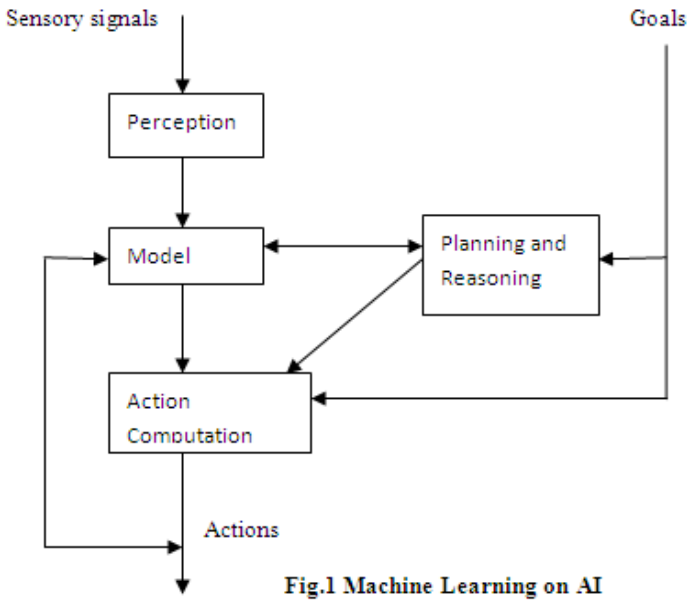

Fig.1 Machine Learning on AI

There are varieties of Machine learning that take to the thing learned is computational structure of some sort. Such varieties are

- Functions
- Logic programs and rule sets
- Finite state machines
- Problem solving system

TYPES OF MACHINE LEARNING

There are two main learning techniques that define the computational feature of the sets. First is **Supervised learning** in the computed form that give the value of *f* for the *m* samples in the training set T and another is semi supervised learning methods that have various *m* samples in finite training sets such that

Training sets T= $\{X_1, X_2, \ldots\ldots X_i \ldots\ldots X_m\}$

The classifier combination approaches can provide solutions to tasks which either cannot be solved by a single classifier, or which can be more effectively solved by a multi-classifier combination scheme. The problem is that we do not know from the beginning which is the best classifier combination method for a particular classification task. In this work, we try to solve this problem by developing a new methodology that combines different combination methods in order to get better performance compared to each individual method. More specifically, in analogy to the combination methods considered, which combine simple classifiers, the proposed meta-classifier approach combines these methods at a higher level aiming at the best classification performance. For the evaluation of approach, we created a medical diagnosis system to classify medical data that have been collected and appropriately inserted into a knowledge base. The basic components used in the system are classifiers such as Neural Networks, Support Vector Machines along with different combination methods. The key feature of the system, from a technical point of view, is that it involves an extra level above the combination of simple classifiers. Specifically, the lowest level consists of simple classifiers, whereas in the middle level there are combination methods that combine the classifiers of the level below.

In this report different and new information is there, which is based in entropy & mutual information measure a different maximal criteria for selecting distinct feature of two groups as well as semi supervised formulation using Co-training with SVM classifier. The use of SVM classifier trained & tested for making different EEG patterns for accurate and great efficiency diagnosis of ADHD group. The combination of Co-training and SVM gives the effective feature selections of different classifiers of patient's data. The decision support algorithms [6] are better diagnosed of ADHD using EEG with different recorded patterns and match the patient's patterns with predefined CO-training and SVM matched patterns and make the model on the basis of symptoms and try to treat it.

There are different classification algorithms that give the periodic burst suppression EEG that applied on the SVM classifier to give the best performance on bursting the tested pattern and followed by the artificial neural network.

The basic need to use SVM and Co-training is that with the help of SVM classifier we find the trained sets which gives us to accurate EEG patterns of patient's brain and with the use of Co-training that forms the training sets and combine the features of linear and non linear classifier to the report which gives the best analytical treatment to ADHD patient and help them to diagnosed.

1.3 EEG feedback on ADHD

EEG is the recording of electrical activity along the scalp. EEG measures voltage fluctuations resulting from ionic current flows within the neurons of the brain. The EEG signals of ADHD brain disorder problems the different algorithms and methods tried to correct or remove EEG contaminates [18]. These techniques attempt to unmixed the EEG signals into some number of underlying components. There are many source separations algorithms often assuming various behaviors or nature [19] of EEG. Using these algorithms, EEG can be decomposed into functionality different components and this study analysis and find which algorithm or method is better for brain disorder problem that can reduce contaminates using MATLAB features. The neurocognitive information processing system that can be defined by components of EEG. This component represents sensory, memory and executive functions.

The standard scalp EEG is recorded at 19 sites. Scalp EEG frequencies are broadly associated with various mental states, as shown in Table 1. With modern computerized systems, experts can map scalp EEG quantitatively by using spectral analysis. Quantitative electroencephalography (QEEG) studies demonstrate deviations from normal patterns in many neuropsychiatric conditions, including ADHD.

Table 1. EEG Rhythms and associated mental states

EEG Rhythm	Frequency (HZ)	Associated Mental States
DELTA	1-4	Sleep, dominant in infants
THETA	3-7	Drowsiness, tuned out
ALPHA	8-12	Alertness, meditation
SMR	12-15	Physically relaxed
BETA	15-21	Focused, sustained attention
HIGH BETA	20-32	Intensity, anxiety
GAMMA	**38-42**	**Important in learning**

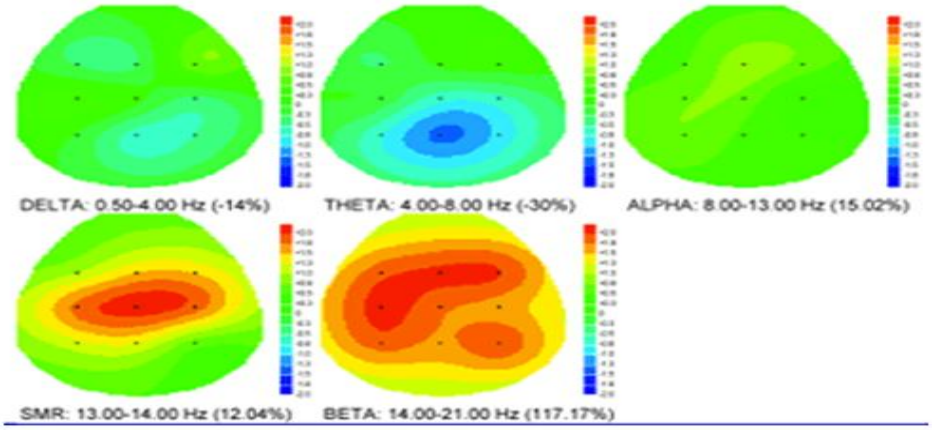

Fig.2 different EEG rhythms in brain of ADHD

These are the different formulation of EEG rhythms that defined in the ADHD related disease. The source is taken by the science direct formulated website.

EEG has been investigated in a large number of studies in the context of sensory & cognitive processing deficits in ADHD.

These patterns have particular frequency ranges and are associated with different states of brain function (e.g., waking and various levels of sleep). These patterns represent synchronized activity over a network of neurons.

Delta waves are the slowest of the known EEG frequencies—no faster than 4 Hz or about ¼ second between each delta wave. Delta waves can be seen in the cerebral cortex of the frontal lobe.

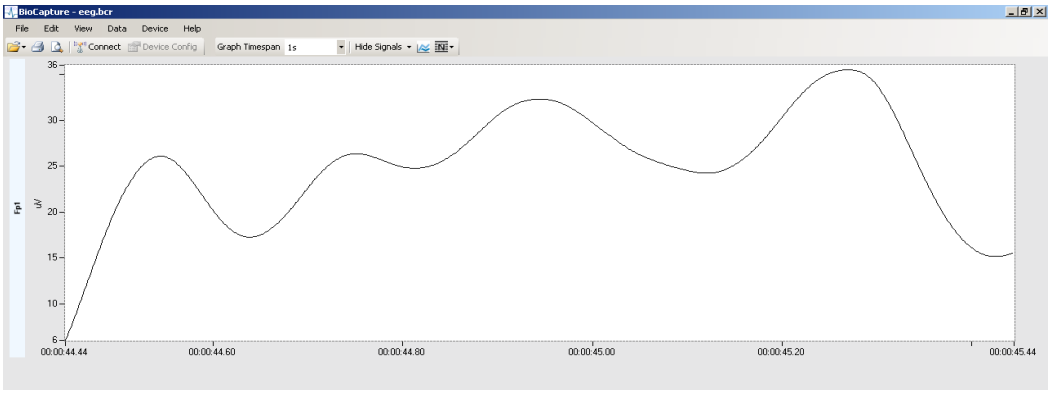

Fig. 3 EEG delta waves

II. Medical Diagnosis using Machine Learning

Machine learning involves different rules or computer models that describe problems on the basis of classifier and patterns. Machine learning provide methods, techniques and tools [11] that help solving diagnostic and prognostic problems in variety of medical domains and specially in brain disorder problem also that define different tools and methods to treat ADHD. Machine learning is divided into three parts supervised learning method, unsupervised learning method and semi supervised learning method [7]

Segmentation techniques are essential for medical diagnosis, as they allow quantifying changes in volume, shape, structure, location etc. Segmentation technique can also be useful in finding markers in medical conditions such as developmental or neurological disorders, helping to diagnose any health problem related to brain disorder or any other health problem.Machine learning techniques have been used for many years of EEG patterns for create the medical modality. Neural networks [9], k-nearest neighbors [10] and many other techniques are used to segment the medical conditions of brain anatomical structures and for also tumors, head injuries and for automatic diagnosis. Various features that may be applied on the segmentation processes that allow preceding different classifier and applicable to brain disorder functioning to sets of medical diagnosis features. The use of EEG patterns that gives an intelligent system which takes real time patient data and create linear and non linear classifier with the help of SVM and Co-training classifiers for detect the variation in patient's conditions.

Learning from patient's data founds various difficulties since the datasets are incomplete, incorrect. And inexactness through this problem machine learning provides various tools for is such as neural networks, decision making etc. But the variant in machine learning depends on training phase and testing phase of the patient medical report and apply various computation model and methods on it to verify the training sets.

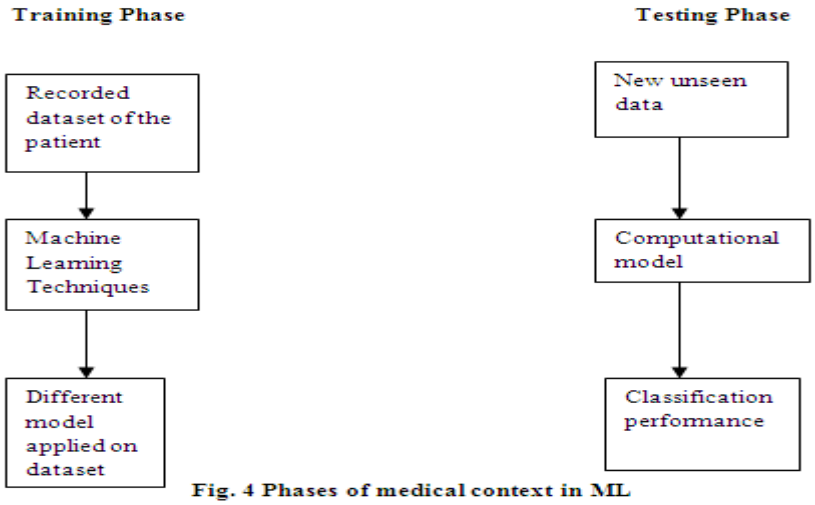

Fig. 4 Phases of medical context in ML

III. Problem Statement

3.1 Problem Definition

The problem definition of the EEG signals that used in ADHD problem is that when being used the factors of Co-training system is divided the problem into two different sets that gives the non linear mapping function to transform the data into higher dimension space where non linear data becomes linear data sets and

another is that the unsupportive function and incomplete data are occurred in during the testing of patient's checkup then use of SVM classifier with linear datasets gives the accurate machine functions of the ADHD system of the report.

The proposed method that may be applied on the data those children who are suffering from ADHD problem gives the best function using SVM and ICA classifier on EEG machine learning datasets. There are some sorts of data that may help to children who are suffering from this problem. Consider the four children who are suffering from ADHD between the ages of 3-6 years. The first goal of this problem is to find the risk estimation and prediction [12] that classifies the patient according to their risk for disease or to make risk predictions. There are many methods for this but machine learning is the best methods to find these predictions and risks. These types of concepts classification is considered as different levels of attention with different groups.

3.2 Problem Solution

The solution of this definition is to define the various algorithms and methods that applied on the pattern of EEG signals and define algorithms on it. The EEG recording can be analyzed using various programs e.g. using free open source toolboxes for MATLAB such as EEGLAB, fieldtrip, NBT etc.

The different machine learning techniques are

Supervised learning method This type of method used the training sets data where each document is labeled by different categories to classify the new text features. The work for training set in text classifier that gives the best performance with new text document [9].

The different algorithms are:-

i) **K-nearest neighbor**: - with the use of KNN we find the training documents to find their k nearest neighbors. The algorithm is based on the assumptions that the characteristics of members of the same class should be same. This method is effective, simple and easy to implement. The disadvantage of these algorithms is that it becomes slow when the size of training set grows and can affect its accuracy.

ii) **Support vector machine (SVM)**: - SVM is one of the fine methods which give the accurate fields. The SVM is based on risk minimization method from computational learning theory [10]. SVM has negative and positive training set which are different from others methods. The performance of SVM remains unchanged if the documents not belong to vector that removed from the training sets.

iii) **Naïve Bayes classifier: -** The naïve bayes classifier is much compatible with the others classifiers often with neural networks. The naïve classifier applies to learning tasks where each instance x is described by a conjunction of attribute values and where the target function

f(x) can take on any value from some finite set V. A set of training examples of
the target function is provided, and a new instance is presented, described by the
tuple of attribute values ***(al, a2.. .a,).***

iv) **Independent Component Analysis**: - ICA is a computational [17] method that separated a set of mixed potentials measured at the scalp into a corresponding set of independent source signals.

A simple application of ICA is the COCKTAIL PARTY PROBLEM where the underlying speech signals are separated from a sample data consisting of people talking simultaneously in a room. It is also good for signals that are not supposed to be generated by a mixing for analysis purposes.

Semi supervised learning method: - This method used for unlabeled data with the few labeled data to classify new unlabeled text document.

Co-training: - The co training [11] method that applied on the data having natural features that divided into two sets. Each sub feature set is enough to train a good classifier. These two subsets are combined and independent to each other. Each classifier classifies the unlabeled data and defines the other classifier. In co training unlabeled data helps to reduce the space size.

These functions deals with the working of supervised method and semi supervised methods using the way of clustering. Generally, supervised and semi supervised methods have been the main focus of research in the area of the artificial intelligence.

The patterns of Electroencephalographic signals EEG used in supervised and semi supervised learning method that with the SVM training sets gives the best analysis method to show the feature selection of the unseen datasets.

IV. Methodology

4.1 Using Matlab Computation

MATLAB stands for MATRIX LABORATORY developed by Cleve Moler in 1984. It is a high level language for technical computing that allow fast prototyping and scripting programming languages such as

integration in C, C++, java code etc it is a object oriented programming function. Matlab compiler is to build executable or a shared library using "Matlab compiler runtime" in order to build applications which run without Matlab.

With the help of Matlab execution we find which is the best classifier that helps to reduce the effects of Brain disorder problem. We have the mathematical parameters of different training sets of classifiers through which we design the wave form and filter it with the help of different classifier. That classifier which has the minimum risk to define the disease in definite time and amplitude gives the best result.

The use of different algorithms that defined the variations of signals and applied the parameters on it that gives the proper result and accuracy.

i) KNN classifiers: - KNN is a supervised learning method for classifying objects based on closest training.KNN is the simplest algorithm in machine learning that an object is classified by a majority vote of its neighbors, with the object being assigned to the class most common amongst its K nearest neighbor.

The program to find K nearest neighbors within a set of points.

a) [Neighbors distances] = kNearestNeighbors (dataMatrix, queryMatrix, k);
b) DataMatrix= (N x D) - N vectors with dimensionality D (within which we search for the nearest neighbors)
c) QueryMatrix = (M x D) - M query vectors with dimensionality D
d) K = (1 x 1) - Number of nearest neighbors desired

The neighbors are taken from a set of objects for which the correct classification (or, in the case of regression, the value of the property) is known. This can be thought of as the training set for the algorithm, though no explicit training step is required. The k-nearest neighbor algorithm is sensitive to the local structure of the data.

When the input data to an algorithm is too large to be processed and it is suspected to be notoriously redundant (e.g. the same measurement in both feet and meters) then the input data will be transformed into a reduced representation set of features (also named features vector). Transforming the input data into the set of features is called Feature extraction. If the features extracted are carefully chosen it is expected that the features set will extract the relevant information from the input data in order to perform the desired task using this reduced representation instead of the full size input. Feature extraction is performed on raw data prior to applying K-NN algorithm on the transformed data in Feature space [21].

4.2 INDEPENDENT COMPONENT ANALYSIS: - ICA is a computational [17] method that separated a set of mixed potentials measured at the scalp into a corresponding set of independent source signals.

A simple application of ICA is the COCKTAIL PARTY PROBLEM where the underlying speech signals are separated from a sample data consisting of people talking simultaneously in a room. It is also good for signals that are not supposed to be generated by a mixing for analysis purposes.

An important note to consider is that if *N* sources are present, at least *N* observations (e.g. microphones) are needed to get the original signals. This constitutes the square case ($J = D$, where D is the input dimension of the data and J is the dimension of the model). Other cases of underdetermined ($J < D$) and over determined ($J > D$) have been investigated.

The simple mathematical definition that linear independent component analysis can be divided into noiseless and noisy cases, where noiseless ICA is a special case of noisy ICA. Nonlinear ICA should be considered as a separate case.

The data is represented by the

Random Vector^{x} = $(x_1,\ldots\ldots,x_m)^T$ and the components as the **Random Vector^{s} = $(S_1,\ldots\ldots S_n)^T$** . The task is to transform the observed **Data^{x}**, using linear static transformation **W as S= W_x** into maximally independent components S measured by some functions **$F(S_1,\ldots\ldots S_n)$** of independence.

With the calculating of different components number F= ($S_1,\ldots\ldots S_{15}$) Random vector=2.37 and Data d= 3.011* function of ICA.

V. Result & Control

The result that comes out using the different techniques on the EEG signals of ADHD brain disorder gives the accuracy and analytical outcome using the MATLAB tools and calculations. It may reduce the contaminates of EEG signals and provide the basic and error free result.

These are the some groups of children that defined the symptoms and machine learning techniques that give the feature extraction. It contains the information regarding the type of extremal point (min/max), which has to be located, and also the information regarding the type of return value (amplitude/latency) the feature provides the signals.

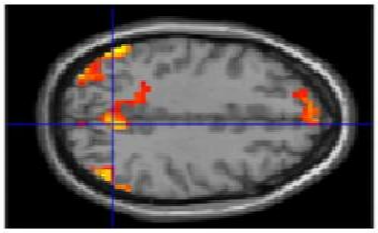

Fig. 5 ADHD disorder in brain

. The red symbol represents the ADHD problem in brain that gives the EEG signals through the EEG machine that record in the machine. With the help of these parameters of the signals we present the parameters in the tools of MATLAB features and applied the methods and techniques of the Machine learning. After that the parameters comes out in the mathematical formation filter them in graph of signal form and gives the result. The table represents the control and the accuracy of the parameters

Table. 2 Different classifiers in MATLAB

Different Classifiers	Parameters (input EEG signals in wave form)	Calculation in MATLAB	Control
KNN	-0.6 to 0.8	Using algorithm	1.56%
ICA	-1.5 to 1.5	Using component BA.5	1.94%

The variations with different groups give the ADHD feature in the problem defined nature and show the analytical feature of the brain disorder.

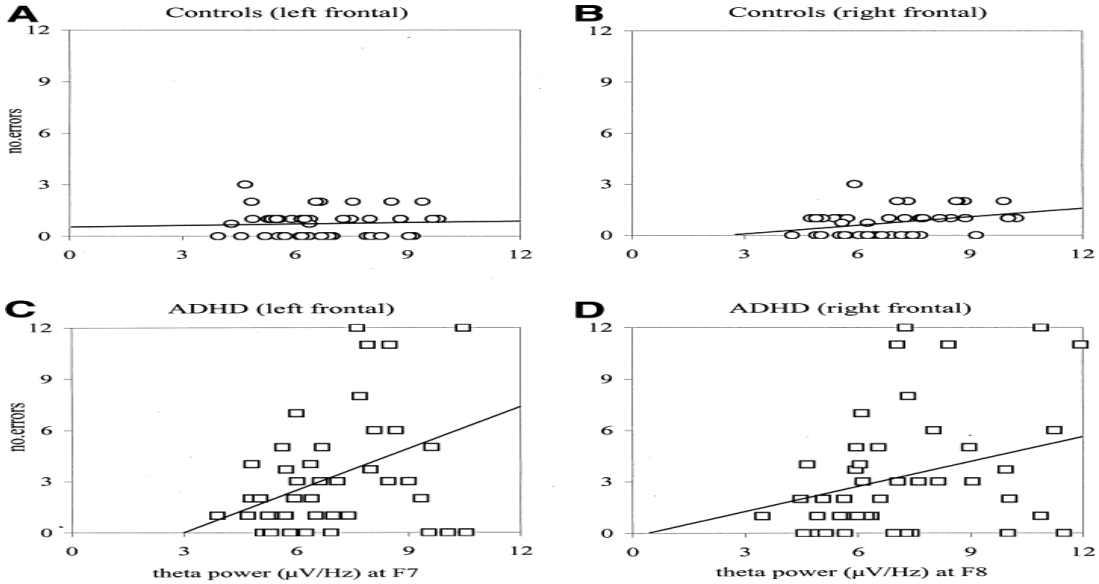

Fig. 6 control of ADHD in theta rhythm

At the end, the set of parameters giving the best performance is selected. The range of parameter values that are going to be tested is properly predefined so as to cover most cases. After selection of parameter values, the system is supposed to have adapted itself to the problem and it is ready to be trained. Due to automatic adaptation, the system does not need an expert's opinion to tune it before putting it to work. So, a doctor can use the system without necessitating technical knowledge and is able to create anytime a new system for a new diagnosis problem.

VI. Conclusion and Future Use

The study of this report gives the result that machine learning techniques applied on the EEG signals may reduce the contaminants of unspecified signals. In terms of future use this can be done in various fields of machine learning implement on different algorithms. ADHD disorder can also be define with Event Related

Potential (ERP) method using different techniques applied on it ca can be diagnosed this major disorder in future. With the EEG that applied on patient's brain, collect the patterns and match that recorded patterns with predefined EEG patterns and give the result. Independent EEG components that have been provide the feature selection that used foe characterized ADHD. In terms of future use this can be done in various fields of machine learning implement on different algorithms. ADHD disorder can also be define with Event Related Potential (ERP) method using different techniques applied on it ca can be diagnosed this major disorder in future.
ADHD can be classified into the using ASD with the parameter of SVM classifier and can be cure in future. There are various methods that can help ADHD patients to recover this disorder. the results of the present study show that independent EEG components can be used for classifying ADHD patients, the use of generic spatial filters limits the significance of the results with regard to the electrophysiological characteristics of ADHD patients With the EEG that applied on patient's brain, collect the patterns and match that recorded patterns with predefined EEG patterns and give the result. Independent EEG components that have been provide the feature selection that used foe characterized ADHD. In terms of future use this can be done in various fields of machine learning implement on different algorithms. ADHD disorder can also be define with Event Related Potential (ERP) method using different techniques applied on it ca can be diagnosed this major disorder in future.
ADHD can be classified into the using ASD with the parameter of SVM classifier and can be cure in future. There are various methods that can help ADHD patients to recover this disorder. The results of the present study show that independent EEG components can be used for classifying ADHD patients, the use of generic spatial filters limits the significance of the results with regard to the electrophysiological characteristics of ADHD patients

References

[1] Attention Deficit Hyperactivity Disorder" book page no- 115-130, National Institute of Mental Health, USA
[2] World Health Organization, "classification of mental & behavioral disorder", 2010
[3] Faraone SV, Blederman J, Mick E, "The age dependent decline of ADHD-a meta analysis of follow up studies" psycho medical 2006 vol 36 page 159-169
[4] Muller et al…., " Classification of ADHD patients on the basic of independent ERP components using machine learning system" joint workshop for cost actions December 2009
[5] Berdakh Abibullow & Jimmy Ao, " Decision support algorithms for diagnosis of ADHD using EEG", Journal of medical system 2012, vol 36, no-4, page 2675-2688, springer.com
[6] "Machine learning techniques for medical analysis", Newton Cheung, October 2001
[7] Vapnik VN, " The nature of statistical learning theory" springer vol 2 2000
[8] Li J, Zhang L, Tau D, Sun H, " A prior neurophysiolyic knowledge for tender based scheme for single trail EEG classification", IEEE trans Neural System Rehab it Engg 2009, page 115-128.
[9] Arzucan Ozgur "Supervised and Unsupervised machine learning techniques foe text document" thesis submitted in Department of computer science, Bogaziki University 2004.
[10] Vladimir N. Vapnik, "The nature of statistical learning theory", springer, New York. 2000
[11] " Semi supervised learning literature survey", Xiaojin Zhu University of Wisconsin computer science December 2007
[12] "Risk estimation and risk prediction using machine learning methods", Jochen Kruppa, Andreas Ziegler, Inke R.Konig June 2012 article is published with open access Springerlink.com
[13] "Describing the Brain in Autism in FiveDimensions—Magnetic Resonance Imaging-Assisted Diagnosis of Autism Spectrum Disorder Using a Multiparameter Classification Approach", Christine Ecker, Andre Marquand, the journal of neuroscience August 2011.
[14] Ward MF, Wender PH, Reimherr FW (1993) The Wender Utah Rating Scale: an aid in the retrospective diagnosis of childhood attention deficit hyperactivity disorder. Am J Psychiatry 150:885– 890.
[15] Han X, Jovicich J, Salat D, "Reliability of MRI-derived measurements of human cerebral cortical thickness: the effects of field strength, scanner upgrade and manufacturer". Neuroimage 32:180 –194. 2006
[16] Gletsos, M., Mougiakakou, S.G., Matsopoulos, G., Nikita, K.S., Nikita, A.S.: A Computer- Aided Diagnostic System to Characterize CT Focal Liver Lesions: Design and Optimization of a Neural Network Classifier. IEEE Transactions on Information Technology in Biomedicine, Vol. 7, No. 3, September 2003.
[17] Hyvärinen, A.; Karhunen, J.; Oja, E. (2001): Independent Component Analysis, New York: Wiley, ISBN 978-0-471-40540-5 (Introductory chapter)
[18] Epstein, Charles M. (1983). "Introduction to EEG and evoked potentials." J. B. Lippincot Co.. ISBN 0-397-50598-1.
[19] Jung, TP; Makeig, S; Humphries, C; Lee, TW; McKeown, MJ; Iragui, V; Sejnowski, TJ (2000). "Removing electroencephalographic artifacts by blind source separation". *Psychophysiology* **37** (2): 163–178.