INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

Classifications & Misclassifications of EEG Signals using Linear and AdaBoost Support Vector Machines

Neelam Rout
Research Scholar, Computer Science Department, S'O'A University
neelamrout732@gmail.com

***Abstract:* Epilepsy is one of the frequent brain disorder due to transient and unexpected electrical interruptions of brain. Electroencephalography (EEG) is one of the most clinically and scientifically exploited signals recorded from humans and very complex signal. EEG signals are non-stationary as it changes over time. So, discrete wavelet transform (DWT) technique is used for feature extraction. Classifications and misclassifications of EEG signals of linearly separable support vector machines are shown using training and testing datasets. Then AdaBoost support vector machine is used to get strong classifier.**

***Keyword:* Electroencephalography (EEG) signals, Feature extraction and selection, Linearly separable support vector machines (SVM), AdaBoost support vector machines.**

I. INTRODUCTION

Human brain is very complex system. Electroencephalography (EEG) signals which is the recording of spontaneous electrical activity of the brain gives information about the function of the brain and it is a direct measurement of cortical activity with millisecond temporal resolution. EEG signals are recorded in a short time by placing electrodes at various positions on the scalp. Its temporal resolution of the EEG signal is high and it is also a non-invasive procedure. The uses of EEG signals can be clinical and research uses [1].

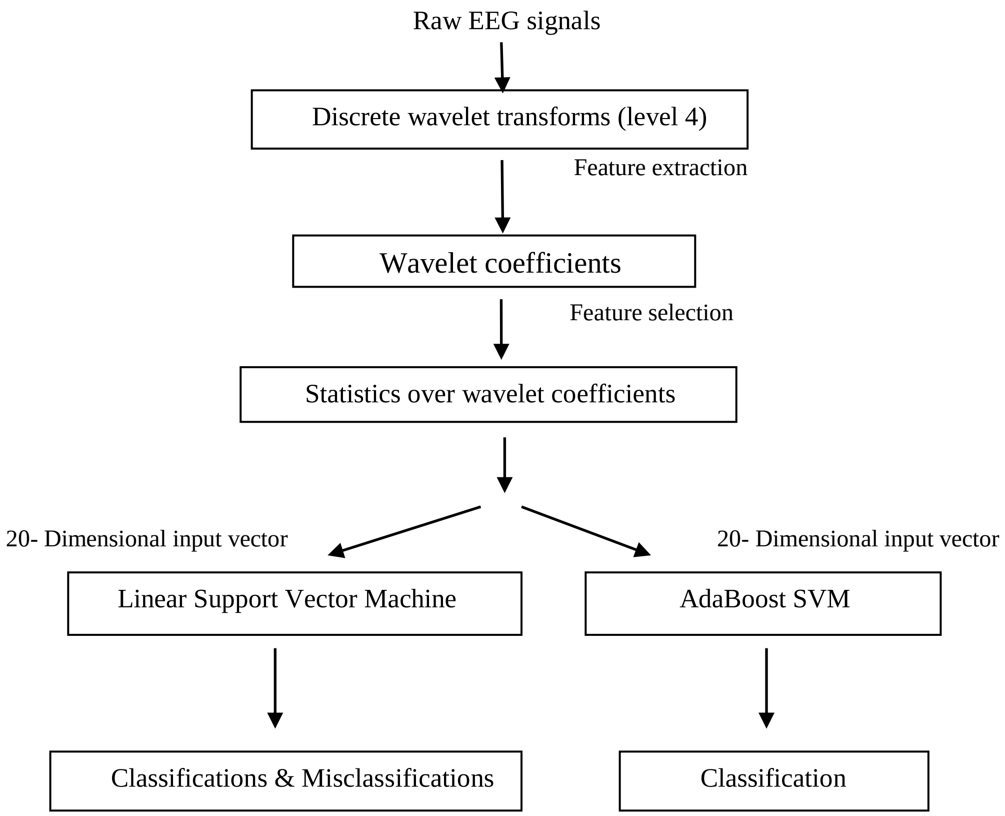
II. MATERIALS AND METHODS

A. *Description of EEG signals*

The raw EEG signal is taken [2] which consists of total 5 sets (classes) of data (SET A, SET B, SET C, SET D, and SET E) but three data sets are selected from 5 data sets in this work i.e. SET A contains recordings from healthy volunteers with open eyes, SET D contains recording of epilepsy patients in the epileptogenic zone during the seizure free interval, and SET E contains the recordings of epilepsy patients during epileptic seizures by using Standard Electrode placement scheme also called as International 10-20 system [1]. Most of EEG waves range from 0.5-500 Hz, these bands are clinically relevant: (i) delta, (ii) theta, (iii) alpha, (iv) beta and (v) gamma.

B. *Feature Extraction and Selection*

Feature extraction is the collection of relevant information from the signal. A discrete wavelet transform (DWT) is any wavelet transform for which the wavelets are discretely sampled. For classification of EEG signals there are four different successive stages followed i.e. normalization, segment detection, feature extraction, and classification. In this each of the EEG signals is decomposed into four detail wavelet coefficients (D1, D2, D3, and D4) and one approximation wavelet coefficients (A4). There are 20 dimension feature vectors for single EEG segment of Class A, Class D, and Class E are used. Different statistics methods are used to reduce the number of features are, maximum, minimum of wavelet, mean and standard Deviation of wavelet coefficients in each sub band [3].



[Fig 1: Generalized Structure of SVM Classifier]

III. CLASSIFIER

A. *Linearly Separable support Vector Machines*

Support vector machines (SVMs, also support vector networks) are supervised learning models. A Support vector machine constructs a hyper plane or set of hyper planes in a high or infinite dimensional space, that optimally separates the data into two categories. It is based on the concept of decision planes that define decision boundaries. The Linearly separable figures are shown in the result section [6] [9].

B. *AdaBoost Support Vector Machines*

Ensemble method is used for unstable classifier to achieve improve performance. AdaBoost (Adaptive Boosting), is a machine learning meta-algorithm was proposed by Y. Freund R. Schapire who won the prestigious "Godel prize" in 2003. It is used in conjunction with other learning algorithms ('weak learns') is combined into a weighted sum that represents the final output of the boosted classifier for a given set of training sample. It maintains weight for each sample. For each iteration , it is adaptively adjust weights. AdaBoost uses higher weights, which seems difficult for component classifier. It combines all the component classifiers to make decision. It is sensitive to noisy data. It is less susceptible to the over-fitting.

It is very fast, simple and easy to program, versatile, and no parameters to tune (except t). Its disadvantages are weak classifiers too complex leads to over-fitting, weak classifiers too weak can lead to low margins, and can also lead to over-fitting, and from empirical evidence, AdaBoost is particularly vulnerable to uniform noise [13][14].

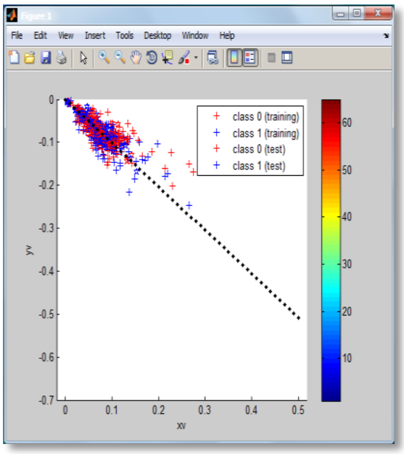
C. *Connection and Comparison between Linear & AdaBoost SVM*

The Boosting and Support Vector Machines methods are "essentially the same" except the measuring techniques for the margin. Boosting relies only the most salient dimensions and SVMs uses kernel tricks (l_2 norm) to compute scalar products in feature space and boosting uses l_1-norm [17].

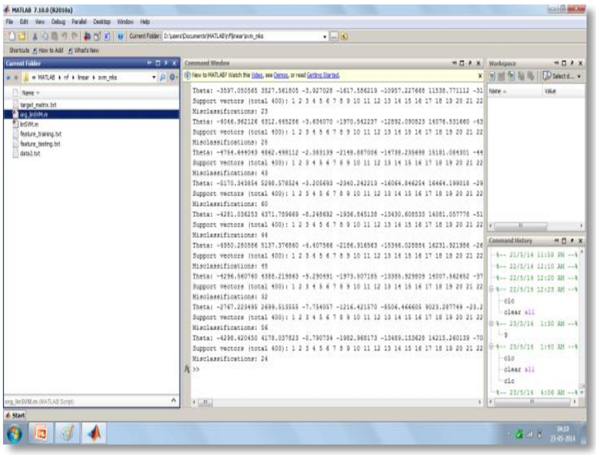
Using any desired linear or non-linear hyper-plane, SVMs globally and explicitly maximize the margin while minimizing the number of wrongly classified examples while in a greedy fashion with lowest error, AdaBoost combines a set of weak learners in order to form a strong classifier.

IV. DISCUSSION OF PRACTICAL IMPLEMENTATION & RESULTS OF SVM

A. *Linearly separable SVM*



[Fig 2: Linearly separable SVM 1]

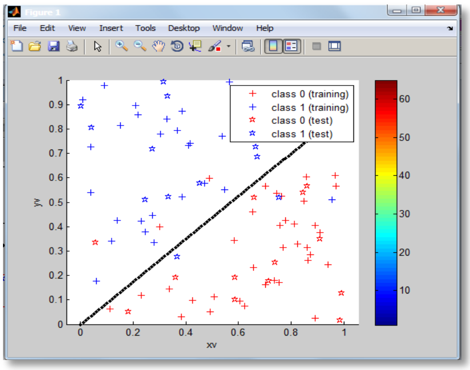


[Fig 3: Result of linearly separable SVM 1]

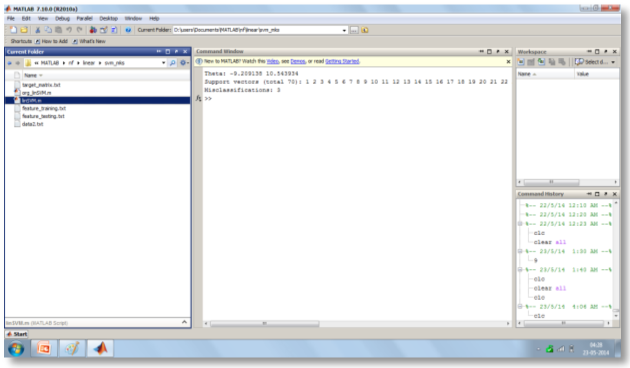
TABLE 1
LINEARLY SEPARABLE SVM 1

1	Theta Values (In between)	-33578.583537 to 33578.583537
	Support Vectors	400
	Misclassifications	21
	Classifications	379
2	Theta Values (In between)	-39322.949057 to 29046.264406
	Support Vectors	400
	Misclassifications	53
	Classifications	347
3	Theta Values (In between)	-40324.554936 to 27440.340592
	Support Vectors	400
	Misclassifications	43
	Classifications	357

4	Theta Values (In between)	-38202.099223 to 35163.947194
	Support Vectors	400
	Misclassifications	60
	Classifications	340
5	Theta Values (In between)	-37586.419054 to 34423.035527
	Support Vectors	400
	Misclassifications	35
	Classifications	365
6	Theta Values (In between)	-24524.845600 to 28915.782878
	Support Vectors	400
	Misclassifications	54
	Classifications	346



[Fig 4: Linearly separable SVM 2]

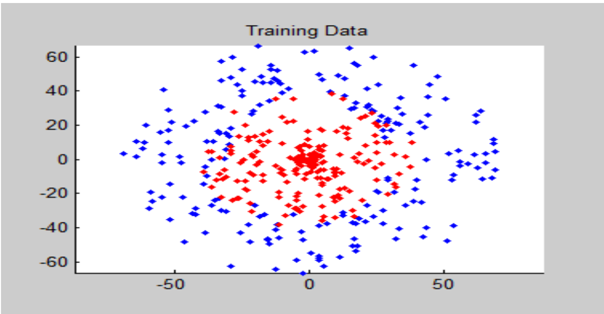


[Fig 5: Result of linearly separable SVM 2]

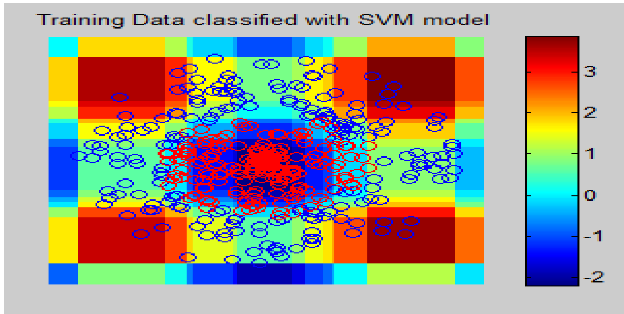
TABLE 2
LINEARLY SEPARABLE SVM 2

Theta Values	**-9.209138 & 10.543934**
Support Vectors	**70**
Misclassifications	**3**
Classifications	**67**

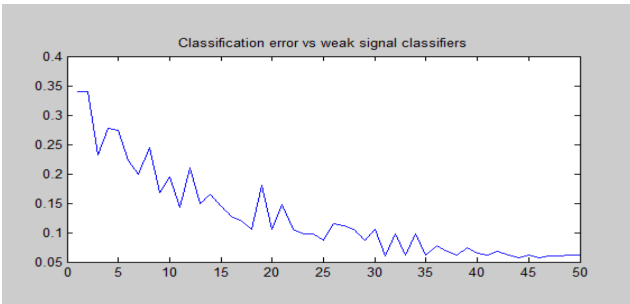
B. *AdaBoost SVM*



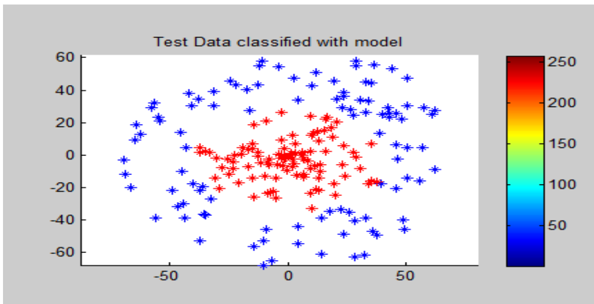
[Fig 6: Training data for AdaBoost SVM model]



[Fig 7: Training data classified with SVM model]



[Fig 8: Classification error versus weak signal classifier]



[Fig 9: Test data classified with model]

V. DISCUSSION OF RESULTS

In linear SVM 1, different theta values are taken but support vectors are fixed i.e. 400 and the results of classifications and misclassifications are shown. In linear SVM 2, by using training and test data sets, the results of classifications and misclassifications are shown where theta values and support vectors are fixed.

In the experiment of AdaBoost SVM, AdaBoost tries to generate a strong classifier. It is done by using a linear combination of a set of weak classifiers which tries to find the best threshold to separate the data into two classes. In each classification step, the boosting part changes the weights of miss-classified data. So that, “weak classifiers” behaves as a “strong classifiers”.

VI. CONCLUSIONS & FUTURE WORKS

Electroencephalography (EEG) signals give important information about neurobiological disorders and for feature extraction, DWT method is used. Different wavelet coefficients like maximum, minimum, mean and standard deviation are used. Then the result of classifications and misclassifications of Support Vector Machines (Linear and AdaBoost) for EEG signals are shown. AdaBooost SVM is used to convert a weak signal to a strong signal. EEG signal processing is a vast area of research. Different types of feature extraction techniques can be tried to reduce the computational complexity. Other new methodologies can be implemented in this area and further different types of classifiers like hybridize pattern classifiers, kernel SVM can be tested.

ACKNOWLEDGMENT

I would like to express my special appreciation and thanks to my advisors, Parents and God.

REFERENCES

[1]. Mohammad Zavid Parvez, Manoranjan Paul, Epileptic seizure detection by analyzing EEG signals using different transformation techniques, *Volume 145, 5 December 2014, Pages 190–200.*

[2]. http://www.meb.unibonn.de/epileptologie/science/physik/eegdata.html.

[3]. S.Soltani, On the use of the wavelet decomposition for time series prediction, *Neurocomputing, vol. 48, pp. 267 - 277, 2002.*

[4]. Qi Xu*, Hui Zhou, YongjiWang, Jian Huang, Fuzzy support vector machine for classification of EEG signals using wavelet-based features, *Medical Engineering & Physics 31 (2009) 858–865.*

[5]. Deng Wang*, Duoqian Miao, Chen Xie, Best basis-based wavelet packet entropy feature extraction and hierarchical EEG classification for epileptic detection, *Expert Systems with Applications 38 (2011) 14314–14320.*

[6]. Abdulhamit Subasi*, M. Ismail Gursoy, EEG signal classification using PCA, ICA, LDA and support vector machines, *Expert Systems with Applications 37 (2010) 8659–8666.*

[7]. Elif Derya U¨ beyli*, Decision support systems for time-varying biomedical signals: EEG signals classification, *Expert Systems with Applications 36 (2009) 2275–2284.*

[8]. Suryannarayana Chandaka, Amitava Chatterjee*, Sugata Munshi, Cross-correlation aided support vector machine classifier for classification of EEG signals, *Expert Systems with Applications 36 (2009) 1329–1336.*

[9]. Elif Derya U¨beyli*, Support vector machines for detection of electrocardiographic changes in partial epileptic patients, *Engineering Applications of Artificial Intelligence 21 (2008) 1196–1203.*

[10]. Babak Mohammadzadeh Asl*, Seyed Kamaledin Setarehdan, Maryam Mohebbi, Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal, *Artificial Intelligence in Medicine (2008) 44, 51—64.*

[11]. Sebastián Maldonado , Richard Weber , Fazel Famili, Feature selection for high-dimensional class-imbalanced data sets using Support Vector Machines, *Information Sciences 286 (2014) 228–246.*

[12]. Elif Derya Übeyli*, Analysis of EEG signals by combining eigenvector methods and multiclass support vector machines, *Computers in Biology and Medicine 38 (2008) 14 – 22.*

[13]. Xue-Fang Chen, Hong-Jie Xing*, Xi-Zhao Wang, A modified AdaBoost method for one-class SVM and its application to novelty detection, *Systems, Man, and Cybernetics (SMC), 2011 IEEE International Conference on , Page(s): 3506 – 3511, ISBN: 978-1-4577-0652-3.*

[14]. Chengming Liu, Manchun Li*, Yongxue Liu*, Jieli Chen, Chenglei Shen, Application of Adaboost Based Ensemble SVM on IKONOS Image Classification, *Geoinformatics, 2010 18th International Conference on , Page(s):1 – 5, ISBN: 978-1-4244-7301-4.*

[15]. Enrique Romero*, Llu12s M3arquez, Xavier Carreras, Margin maximization with feed-forward neural networks: a comparative study with SVM and AdaBoost, *Neurocomputing 57 (2004) 313 – 344.*

[16]. Xuchun Li*, Lei Wang, Eric Sung, AdaBoost with SVM-based component classifiers, *Engineering Applications of Artificial Intelligence 21 (2008) 785–795*.

[17]. Hanif Mohaddes Deylami, Yashwant Prasad Singh, Adaboost and SVM based cybercrime detection and prevention model, *Artificial Intelligence Research, December 2012, Vol. 1, No. 2.*

[18]. E. D.U beyli, Combined neural network model employing wavelet coefficients for EEG signals classification, *Digital Signal Processing, vol. 19, pp. 297-308, 2009.*

