

Determination of Increased Mental Workload Condition From EEG by the Use of Classification Techniques

Sınıflandırma Teknikleri Kullanılarak EEG'den Artan Mental İşyükü Durumunun Belirlenmesi

Adil Deniz DURU¹ 

¹ Marmara University, Faculty of Sport Sciences, Department of Sports Coaching Education, İstanbul, Turkey

Abstract

Extraction of the information hidden in the brain electrical signal enhances the classification of the current mental status. In this study, 16 channel electroencephalogram (EEG) data were collected from 15 volunteers under three conditions. Participants were asked to rest with eyes open and eyes closed states each with a duration of three minutes. Finally, a task has been imposed to increase the mental workload (MW). EEG data were epoched with a duration of one second and power spectrum was computed for each time window. The power spectral features of all channels in traditional bands were calculated for all subjects and the results were concatenated to form the input data to be used in classification. Decision tree, K-nearest neighbor (KNN) and Support Vector Machine (SVM) techniques were implemented in order to classify the one-second epochs. The accuracy value obtained from KNN was found to be 0.94 while it was 0.88 for decision tree and SVM. KNN was found to outperform the two methods when all channel and power spectral features were used. It can be concluded that, even with the use of input features formed by concatenating all subject's data, high classification accuracies can be obtained in the determination of the increased MW state.

Keywords: Mental Workload, Classification, Machine Learning, EEG

Öz

Beyin sinyallerinin sınıflandırılması kişilerin mental durumu hakkında bilgi sahibi olmamızı kolaylaştırır. Bu çalışma kapsamında onbeş katılımcıdan 16 kanallı EEG verisi üç farklı durum için toplanmıştır. Katılımcılardan üçer dakika gözleri kapalı ve açık şekilde dinlenmeleri istenmiştir. Sonrasında mental işyükünü arttıracak bir ödevde uygulanmıştır. EEG verileri birer saniyelik pencereye ayrılmış ve her pencere için güç spektrumları hesaplanmıştır. Geleneksel frekans bantlarına ait güç spektrumları her katılımcı için hesaplanmış ve sonuçlar sınıflandırmada kullanılmak üzere arka arkaya getirilmiştir. Karar ağaçları, K-en yakım komşuluk (KNN) ve Destek Vektör Makinaları (DVM) teknikleri birer saniyelik pencere verilerinin sınıflandırılması için uygulanmıştır. KNN için doğruluk değeri 0.94 olurken karar ağaçları ve DVM için 0.88 değerine ulaşılmıştır. KNN yöntemi diğer iki yöntemle göre tüm kanalların güç spektrumu kullanıldığı durumda daha yüksek doğruluk oranı vermiştir. Sonuç olarak, tüm katılımcıların verilerinin bir havuza konulup sınıflandırma işlemi yapıldığında dahi yüksek mental işyükü durumunun ayrıştırılabildiği görülmüştür.

Anahtar Kelimeler: Mental İşyükü, Sınıflandırma, Makine Öğrenme, EEG

1. INTRODUCTION

Brain electrical activity includes relevant information about the current status of a subject. Electroencephalogram (EEG) with its high temporal resolution property enables researchers to monitor the underlying mechanisms of endogenous or exogenous stimulus. Besides this, EEG is used for diagnosis of mental diseases. EEG can provide quantitative information about memory loss, sleep disorder and stroke. Recently, the classification between healthy brain and illness were studied by many researchers. Mumtaz et al. performed a classification for the diagnosis of the major depressive disorder using synchronization likelihood metric computed from EEG as input feature and exhibited more than 90% accuracy values with the support vector machine (SVM), logistic regression (LR) and Naive Bayesian (NB) techniques [1]. SVM was used in the recognition of

Alzheimer's disease (AD) with the multi-channel features obtained through the use of wavelet transform on the eyes closed (EC) resting EEG data. In that study, linear and radial basis kernels were used and the linear kernel approach was shown to have 92% accuracy in the discrimination of the AD subjects from normal control ones [2]. In another study, classification rate was found to be higher than 80% in the discrimination between the AD and control group subjects, when the cortical source connectivity parameters were used after the solution inverse problem of EEG (measured from 19 channels) [3]. Both scalp level and source level features extracted from EEG yield researchers and clinicians to distinguish the healthy subjects from the subjects that are suffering from mental diseases. For instance, alpha band activity was used as a feature in order to discriminate the Huntington's disease using a supervised neural network [4]. These features can provide valuable information for the classification studies of mental diseases by eliminating irrelevant measurements. Besides this, scalp EEG measurements are subjected to various studies that explore the emotional state and the planned motion [5]. Obermaier et al. used Hidden Markov Models to distinguish whether the subject is performing a left-hand or right-hand imagination [6]. Similar to this study, left-hand/right-hand imagination and word generation tasks were tried to be identified by the use of principal component analysis as a preprocessing tool in order to reduce the input feature set and then a neural network based classifier was performed [7]. Moreover, the information content of the EEG was used to identify the object, that has been presented as visual stimuli [8]. In several studies explored above, the relation between the human behaviour and the EEG measurements has been investigated. One of the challenging problems of the design industry and behavioral science is the real time recognition of the human mental state. Nowadays there is an increased attention to control robotic systems by the thought of the human via the noninvasive measurement of the electrical activity originated from the brain tissue [9]. Classification techniques derived from machine learning can be used to estimate an instantaneous state of a subject based on EEG. Amin et al. achieved high classification accuracy in the identification of the human mental state in two conditions. In their study, Discrete Wavelet Decomposition (DWT) is applied to the EEG data in order to extract the relative band power of the time series measured during the eyes open (EO) and under increased mental workload (MW) conditions. Then, these relative band power values were used as input features for the classification [10].

From several neuroscience studies, it is well known that when the subject closes his eyes without falling asleep, parieto occipital alpha power spectral peak is observed.

Moreover, a clear difference in the power spectral alpha peak is expected to be present when the subject's mental workload increases. Generally, power spectral density is computed as an average of the power spectrum values computed using epoched data and the window length of each epoch is selected as 1 or 2 seconds where EEG data is assumed to be stationary. On the other hand, the autoregressive approximation can be used to compute the power spectrum as a parametric approach. But in both computation techniques, the whole signal or a large content of the signal is used to represent the power spectrum. Thus a relatively long duration of a measurement is needed to identify the current mental status of the subjects. In real-time applications, especially for brain-computer interfaces, the duration is a crucial parameter since their speed is on the order of 1 bit in a second. Thus, a decision should be given in a second in order to achieve this aforementioned speed.

The aim of this work is to classify the mental status of the subject among three cases, EO, EC and MW. EO, EC and increased MW conditions are used and epochs having a window length of a second are used as input to classifiers. KNN, SVM and Decision Tree classifiers are used as classifiers. Classifier performances have been computed and compared.

II. METHODS

2.1 Experimental Paradigm and Data Preprocessing

Subjects were asked to rest in EO state for a duration of 3 minutes and then they were asked to rest in EC state for another 3 minutes. Finally, they were asked to perform a mental arithmetic task where an increase in their MW was aimed. In MW task, subjects counted starting from a given number (3 digit) backward by subtracting another one-digit number which was instructed just before start. After the process, the last number was asked to the subject to check the performance of the given task. During the experiments, 16 channel (F3, Fp1, Pz, O1, C3, Fz, T7, Cz, Fp2, F4, C4, T8, P3, P4, O2, Oz) EEG measurements were performed with a sampling rate of 1kHz using V-amp biopotential amplifier (Brain Products, GmbH). Active electrode cap was used and data was bandpass filtered with a low frequency cutoff of 0.5 Hz and a high frequency cutoff of 40 Hz. Fifteen subjects were voluntarily participated in the experiments. All of them approved and signed a written informed consent. The study had been approved by Ethical Review Board of the Medical Faculty, Marmara University (approval number: 2013-0191). The measurements have been performed in Marmara University, Faculty of Sport Sciences.

Three classes were named as EO (Class 0), EC (Class 1) and increased MW (Class 2) in the concept of this study. The power spectra of the EEG epochs have been computed using Eq. 1. For each channel (c), delta band (f1=1, f2=4), theta (f1=5, f2=7), alpha (f1=8, f2=13) and beta (f1=14, f2=30) power values were computed for each epoch (i), having a length (N) of 1 sec with a sampling rate (fs) of 1kHz. Thus, for each condition, Y_{band} parameter had a dimension of 16x180.

$$Y_{band}(c, i) = \sum_{w=f1}^{f2} \frac{1}{N^2} \left| \sum_{t=(N-1)*i+fs+1}^{(N-1)*i+fs} y(t, c) e^{-\frac{i2\pi w t}{N}} \right|^2 \quad (1)$$

The spectral features deduced from each epoch of EEG measurement for each band and channels were concatenated to form the input attributes of the classification.

$$X_i = [Y_{\delta}(1, i) Y_{\delta}(2, i) \dots Y_{\delta}(16, i) Y_{\theta}(1, i) Y_{\theta}(2, i) \dots Y_{\theta}(16, i) Y_{\alpha}(1, i) Y_{\alpha}(2, i) \dots Y_{\alpha}(16, i) Y_{\beta}(1, i) Y_{\beta}(2, i) \dots Y_{\beta}(16, i)] \quad (2)$$

Thus the input vector for an epoch had a dimension of 1x64. Since the measurements lasted for 3 minutes for a task, input matrix had the size of 180x64 for a subject. Each realization vector X_i was normalized in each band separately and all of the vectors computed for each subject were pooled to be used as input matrix for the classification process. The input matrix had the size of 8100x64. Prior to classification process, the artefacts occurred in the EEG due to eye movements have been marked and the epochs having these markers are removed from the input matrix. Finally, 7778 epochs remained in the input matrix; where randomly chosen 60% of it was used for training and 40% was used for testing.

2.2 Supervised Classification

In the concept of this paper, for the classification process, K-Nearest Neighbours (KNN), SVM and Decision Tree techniques were used. K-Nearest Neighbours (KNN) is a supervised algorithm such that for a test element (X_i), the K – closest points between all elements in the training dataset are computed. Then, among the labeled K-points, the maximum probability of a label is searched and assigned for the test element. The number K that is used in the classifier can be selected experimentally by searching minimum error rate of the classifier. The second method that is used throughout the paper is the SVM. It can be used to classify multiple classes for both linear and nonlinear data. Training dataset is transformed into a higher dimension based on a nonlinear mapping. In SVM, it is aimed to find a decision boundary that is called as hyperplane (for n-dimensional input space)

which minimizes the classification error [11]. For three classes, one against one approach is adopted to construct 3 classifiers where each classifier is trained with the data of two classes. Thus, the class pairs (i.e., EC vs EO, EC vs MW, and EO vs MW) are used to train the classifiers. Finally, the output of each classification is used to count the labels of the input data set in order to find the most observed class label.

As a predictive modeling approach, classification trees are used to model an output value using input variables. A tree structure similar to a flowchart is formed to represent the rules for reaching a class label that is stored in the leaf nodes of the tree. In this structure, internal nodes correspond to a test on an attribute while each edge forms a sub-tree. When an input data enters the tree, the attribute values are used to find out its class label. Among the decision tree algorithms, Iterative Dichotomiser (ID3) is based on information gain metric that is used to select attributes [12]. Information gain is based on entropy metric and is defined as the difference between the information content of the training data tuples and the information content of those when separated with the given attribute (Eq. 5). For all attributes, information gain is computed and the attribute which has the maximum gain is termed as the splitting attribute. The same procedure is then repeatedly applied in order to select attributes using the grouped data.

$$I_x = - \sum_{i=1}^c p_i \log_2 p_i \quad (3)$$

In Eq.3, the amount of expected information required to classify a finite sequence of items in X is denoted with I_x , c is the number of classes and p_i stands for the probability that an arbitrary tuple in X belongs to class i.

$$I_{A,x} = \sum_{i=1}^v \frac{|X_i|}{|X|} I_{X_i} \quad (4)$$

$$Gain_A = I_x - I_{A,x} \quad (5)$$

In Eq.4, v is the number of unique values of attribute A, X_i is the weight of i.th part in X. $I_{A,x}$ represents the amount of information needed to reach the exact classification. Finally, information gain is computed by the subtracting $I_{A,x}$ from I_x . If the attributes are continuous values, which is the case for EEG measurements, a threshold is selected which minimizes the $I_{A,x}$ such that the distinct values of the attribute are sorted and for each pair of adjacent values, a splitting point is computed by taking the average of their values. Then, the attribute is divided into two classes where the values that are lower than the splitting point is assigned to first class and the higher values are assigned to second

class. For each of possible splitting points, $I_{A,X}$ is computed. However, the information gain computed by unique valued identifiers have a biased effect on the classification. This issue had been overcome by the use of gain ratio in the concept of the C4.5 algorithm [16]. C4.5 uses maximum value of the gain ratio for the splitting process. When the data is thought to include outliers, Classification and regression trees (CART) algorithm can perform better than the C4.5. CART algorithm uses Gini index that measures the impurity of X by using Eq. 6.

$$Gini_X = 1 - \sum_{i=1}^c p_i^2 \quad (6)$$

For each attribute, a binary separation is assumed by the Gini index. Thus, if an attribute has n distinct values, then all of the subsets of these values are considered. If an attribute divides X into two groups, the Gini index is going to be computed by Eq.7. The attribute having the maximum value of Gini differences is selected as the splitting parameter (Eq. 8). By the computation of this approach, leaves of the decision tree is constructed.

$$Gini_{A,X} = \sum_{i=1}^c \frac{|I_i|}{|I|} Gini_{I_i} \quad (7)$$

$$\Delta Gini_A = Gini_X - Gini_{A,X} \quad (8)$$

The above mentioned methods are implemented in Python and the performance of the classifiers are investigated using accuracy, precision, recall (sensitivity) and f1-score metrics by use of the confusion matrix (Table 1). Precision is defined as the ratio of number of true positives (T_p) to the sum of T_p and false positives (F_p). On the other hand, recall parameter is computed by dividing the number of T_p to the summation of T_p and false negatives (F_N). The harmonic mean of precision and recall parameters are named as F_1 score. In addition to these parameters, classification accuracy is termed as the ratio of the number of correct predictions to that of all predictions.

Table 1. Confusion matrix regarding to the case of three classes.

	True: EO	True: EC	True: MW
Predicted:EO	Correct	False EO->EC	False EO->MW
Predicted: EC	False EC->EO	Correct	False
Predicted: MW	False MW->EO	False	Correct

EO, EC and MW stand for eyes open, eyes closed and mental workload conditions, respectively.

III.RESULTS

After the implementation of the KNN, performance metrics were computed using the randomly chosen training set. The mean accuracy was found as 0.94 as shown in Fig. 1. Values regarding to Precision, recall, F1-score and support parameters are summarized in Table 2 for each condition and for each classifier used in this study.

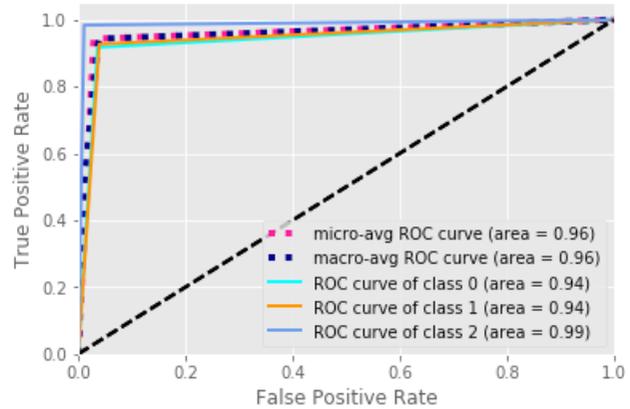


Figure 1. Receiver operating characteristics curves (ROC) computed after the implementation of KNN. The block diagonal black line indicates random classifiers.

Table 2. Classification performance metrics for KNN/SVC/ Decision Tree (DT) methods.

Case	Precision (%)	Recall (%)	F1-score (%)	Support
	KNN/SVC/DT	KNN/SVC/DT	KNN/SVC/DT	KNN/SVC/DT
EO	92/86/87	93/89/89	92/87/88	1039/1016/1031
EC	93/82/87	92/89/86	92/85/87	1019/933/1030
MW	98/94/93	98/85/93	98/90/93	1054/1163/1051
Avg/ Total	94/88/89	94/88/89	94/88/89	3112/3112/3112

The accuracy value of SVC with the radial basis function kernel was found to be 0.875 while it was 0.88 for the decision tree. The ROC curves for SVC and decision tree are visualized in Figure 2 and Figure 3, respectively. Decision tree classification was performed for several maximum depth level values (2-13) and the trees were tested using the remaining part of the data which has not been used to form the tree. Initial layers of the decision tree are shown in Figure. 4. Beta band power values of P3, T7, T8, F3, alpha band power value of right occipital and delta band power value of

C4 were found to discriminate the classes in the first three layers. From decision tree analysis, rules for discrimination between classes can be formed.

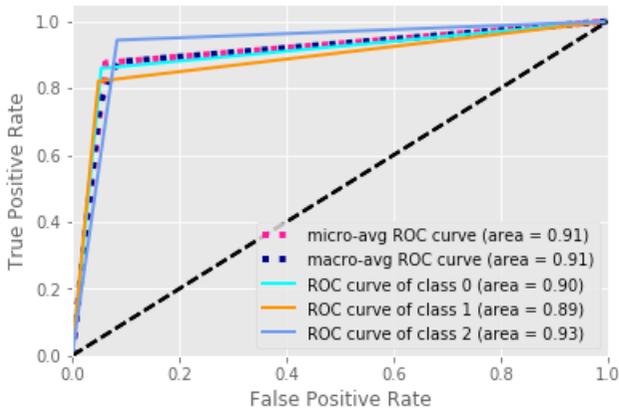


Figure 2. ROC curves computed after the implementation of SVC for each class.

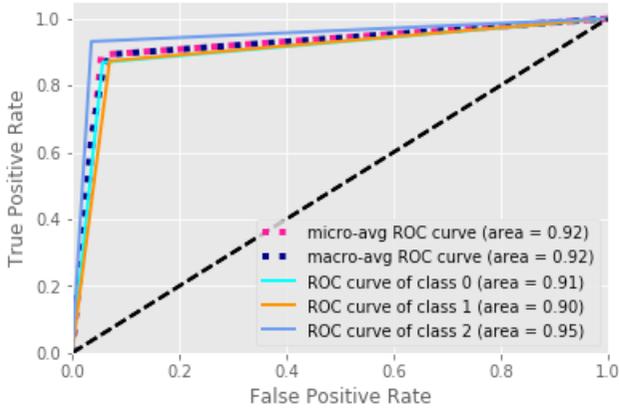


Figure 3. ROC curves computed after the implementation of Decision tree for each class.

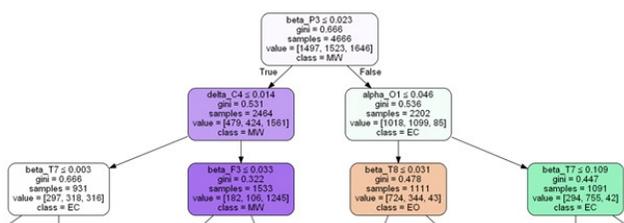


Figure 4. First three levels of the decision tree. Each node includes parameter name, gini value, number of samples, and number of items in each class.

IV.DISCUSSION

In this study, EO, EC, and increased MW states were tried to be identified using the scalp EEG measurements. Power spectrum values of the EEG signal were used as features as input parameters. Epochs having one-second time duration were selected to compute the power spectrum. In a recent study, one second epoch duration was found to have the best classification accuracy when the power spectrum of the EEG was used as input features [13]. However, in that study, classification has been performed using SVM for each subject’s data separately. Nevertheless, Wang et al. reported the classification accuracy as 88% on average among subjects when all the frequency bands were used as input features to discriminate the positive and negative emotion classes.

In our study, the classification accuracy was obtained higher than 96% when each subject’s data were used separately. Since the aim is to generalize the classification among the subjects by incorporating inter-subject variability, all of the data were pooled and the training/testing processes have been performed on the pooled data. The general idea behind the brain-computer interface techniques is to form a structure which uses a reduced amount of training data as well as less number of measurement sensors [14]. Thus, pooling the whole data of the subjects enables us to classify an epoch measured from a subject whose data has not been used in the training phase.

Selection of the features which are going to be used in classification can be performed either using preprocessing tools or using the physiological a priori information about the tasks. For instance, it is relatively easy to discriminate the subject’s cognitive level from the resting baseline. However, when several conditions (tasks) are present, the accuracy can decrease [10]. Amin et al. implemented time-frequency decomposition but the time information has been used in the classification. In the absence of rapidly changing stimuli, usage of the time-frequency components does not seem to be relevant for the classification.

In KNN, since the input attributes are assumed to be equally weighted, accuracy values can be observed low when the noisy attributes are present. But in the present study, eye movement-related artefacts were removed and the annotated tasks were accomplished by subjects successfully which decreased the level of noise. Average scores of precision, recall and f1-score values of all classes were found to be greater for KNN than SVC and decision tree. Parvinnia et al. proposed an adaptive weighted distance nearest neighbor algorithm which they used to classify schizophrenia and healthy subjects. In that study, they have reported higher accuracy values of the proposed method than the accuracy of

KNN. Moreover, they also reported that, nearest neighbor based methods outperform SVM and NaiveBayes [15]. Our results are in agreement with that finding.

In our further studies, decision trees are going to be used as a preprocessing tool for the selection of most relevant features. Then those reduced features are going to be used as input for the classification. Moreover, less number of channels and frequency components are going to be searched for the discrimination of several tasks.

Acknowledgement

This work was supported by Research Fund of Marmara University. Project No: SAG-A-100.713.0296.

References

- [1] Mumtaz W, Ali S.S.A., Yasin M.A.M. & Malik A.S., (2018) A machine learning framework involving EEG-based functional connectivity to diagnose major depressive disorder (MDD). *Med Biol Eng Comput.* Feb;56(2):233-246. doi: 10.1007/s11517.017.1685-z.
- [2] Kanda P.A.F.M., Trambaiolli L.R., Lorena A.C., Fraga F.J., Basile L.F.I., Nitri R., & Anghinah R. (2014), Clinician's Road Map to Wavelet EEG as an Alzheimer's disease Biomarker, *Clinical EEG and Neuroscience*, doi: 10.1177/155.005.9413486272
- [3] Babiloni, C., Triggiani, A. I., Lizio, R., Cordone, S., Tattoli, G., Bevilacqua, V., Soricelli, A., Ferri, R., Nobili, F., Gesualdo, L., Millán-Calenti, J. C., Buján, A., Tortelli, R., Cardinali, V., Barulli, M. R., Giannini, A., Spagnolo, P., Armenise, S., Buena, G., Scianatico, G., Logroscino, G., Frisoni, G. B., ... Del Percio, C. (2016). Classification of Single Normal and Alzheimer's Disease Individuals from Cortical Sources of Resting State EEG Rhythms. *Frontiers in neuroscience*, 10(47). doi:10.3389/fnins.2016.00047
- [4] Bellotti R., De Carlo F., Massafra R, de Tommaso M., Sciruicchio V., (2004) Topographic classification of EEG patterns in Huntington's disease. *Neurol Clin Neurophysiol.* Nov 30;2004:37.
- [5] Cruz-Garza, J. G., Hernandez, Z. R., Nepal, S., Bradley, K. K., & Contreras-Vidal, J. L. (2014). Neural decoding of expressive human movement from scalp electroencephalography (EEG). *Frontiers in human neuroscience*, 8, 188. doi:10.3389/fnhum.2014.00188
- [6] OberMaier B., Guger C., Neuper C. & Pfurtscheller G., (2001) Hidden Markov models for online classification of single trial EEG data, *Pattern Recognition Letters*, 22, 1299-1309
- [7] Lin C.J. & Hsieh M.H., (2009) Classification of mental task from EEG data using neural networks based on particle swarm optimization, *Neurocomputing*, 72, 1121-1130
- [8] Nemrodov D., Niemeier M., Patel A., Nestor A., (2018) The Neural Dynamics of Facial Identity Processing: insights from EEG-Based Pattern Analysis and Image Reconstruction. *eNeuro* 29, ENEURO.0358-17.2018; DOI: 10.1523/ENEURO.0358-17.2018
- [9] Iturrate I., Chavarriaga R., Montesano L., Minguez J. & Millan J.D.R, (2015) Teaching brain-machine interfaces as an alternative paradigm to neuroprosthetics control, *Scientific Reports*, 5
- [10] Amin H. U., Mumtaz W., Subhani A. R., Saad M. N. M. & Malik A. S., (2017) Classification of EEG Signals Based on Pattern Recognition Approach, *Front. Comput. Neurosci.* 11(103) doi: 10.3389/fncom.2017.00103
- [11] V. N. Vapnik, (1999) An Overview of Statistical Learning Theory, *IEEE Transactions on Neural Networks* 10(5)
- [12] J. R. Quinlan. (1986) Induction of decision trees. *Machine Learning*, 1:81-106
- [13] Wang X.W., Nie D. & Lu B.L., (2014) Emotional state classification from EEG data using machine learning approach, *Neurocomputing* 129 94-106
- [14] Gutierrez D. & Escalona-Vargas D.I., (2010) EEG data classification through signal spatial redistribution and optimized linear discriminants, *Computer Methods and Programs in Biomedicine* 97, 39-47
- [15] Parvinnia E., Sabeti M., Jahromi M.Z. & Boostani R., (2014) Classification of EEG Signals using adaptive weighted distance nearest neighbor algorithm, *Journal of King Saud University – Computer and Information Sciences*, 26(1), Pages 1-6
- [16] Quinlan J.R., (1993) Programs for Machine Learning, Morgan Kaufman, San Francisco, USA