Investigation Of Feature Selection Algorithms On A Cognitive Task Classification: A Comparison Study

S. G. Eraldemir, M.T. Arslan, and E. Yildirim

Abstract—**In this study, the effects of feature selection on classification of the electrical signals generated in the brain during numerical and verbal operations are investigated. 18 healthy university/college students were chosen for the experimental study. EEG signals were recorded during silent reading and mental arithmetic operations without using any pen and paper. A total of 60 slides, 30 of which contained reading passages and the rest contained arithmetic operations, were presented in the experiment. EEG signals recorded from 26 channels during the slide show. The recorded EEG signals were analyzed by Hilbert Huang Transform (HHT), and then features were extracted. 312 features were classified by Bayesian Network algorithm without applying feature selection with 92.60% average accuracy. Consistency measures and Correlation based Feature Selection methods were, then, used for feature selection and the numbers of selected features are 8 and 39 on average, respectively. Classification accuracies by using these feature selection algorithms were obtained as 93.98% and 95.58%, respectively. The results showed that feature selection algorithms contribute positively to the classification performance.**

Index Terms— **Hilbert Huang Transform, Consistency Measures, Correlation based Feature Selection, EEG Classification.**

I. INTRODUCTION

Brain-computer interfaces (BCI) are systems that allow to communicate the brain with external devices via a computer. These systems are based on the principle of real-time control of various systems to ease the lives of people who have suffered cognitive, sensory and motor functions such as paralysis or muscular dysfunction [1, 2, 3, 4]. EEG based BCI systems are used intensively for reasons such as high time resolution, low cost, and portability. Recorded EEG signals need to be analyzed, classified and transmitted to the system very quickly [5, 6]. In BCI systems, in general, the following steps are followed: Recording EEG signals, pre-processing, feature extracting and classifying. Finally, In the BCI system, a command is sent to the relevant system considering the classification results.

S. G. ERALDEMIR, is with Dep. of Computer Programing, Iskenderun Technical University, Hatay, Turkey, (e-mail: sgoksel.eraldemir@iste.edu.tr)

M. T. ARSLAN, is with Department of Computer Technology, Mustafa Kemal University, Hatay, Turkey, (e-mail: mtarslan@mku.edu.tr)

E. YILDIRIM, is with Department of Electrical-Electronics Engineering, Adana Science and Technology University, Adana, Turkey, (e-mail: esenyildirim@gmail.com)

Manuscript received August 15, 2017; accepted Nov 16, 2017.

DOI: 10.17694/bajece.419549

In the preprocessing process, the EEG signals are cleaned out noises resulting from the network or subject by various methods. Feature extraction process is an important field which many operations are performed on the signal processing. The most important point in signal processing is whether the signal is stationary or not. Fourier transform is a suitable signal processing method if the signal is a stationary signal whose frequency value does not change by time. Wavelet Transform and Hilbert Huang Transform (HHT) are shown to be more efficient methods to analyze non-stationary signals [7,8]. The Hilbert Huang Transformation, which is preferred in this study, has been used in many studies in the analysis of EEG signals in the literature [9,10,11,12,].

R. Wang et al. showed that HHT gives better results than Fast Fourier Transform and Continuous Wavelet Transform (CWT) in sleepy EEG signals [10]. While sleeping and waking, K. Rai et al. analyzed EEG signals using HHT based features and then they have classified EEG signals by Fuzzy Logic method [11]. Swarnalatha. R. and Prasad D.V in 2015 analyzed the recorded EEG data using HHT method for early diagnosis of bruxism disease, called “interdental rubbing disorder” [12]. In another study, J. Kortelainen et al. have developed a novel approach to early diagnosis of hypoxic ischemic encephalopathy (HIE), permanent illness in infants who do not have adequate blood in the brain, using HHD [9].

In BCI systems, features obtained by various signal processing methods can be used directly for classification process. Furthermore, it is possible to eliminate some of those features that may adversely affect the classification result and thus improve classifier performance. Feature selection methods reduced features without changing the signal's characteristic. Feature selection eliminates irrelevant and unnecessary features and allows selection of features that will provide the actual contribution in the classification process so that both training and test times are shortened during classification. Thus, problems that can arise from the delays in real-time BCI systems are reduced. Feature selection techniques have been used many studies in classification of EEG signals [13, 14, 15, 16, 17]. Feature selection approaches are known as filter, wrapper and embedded systems. Filter-based methods provide fast results because they perform selection based on statistical data in the analysis of large data groups. Wrapper methods choose features that performs well for the chosen classifier. Embedded systems combine the operating logic of the two systems mentioned above.

In this study, we used correlation-based feature selection (CFS) and consistency measures, which are filter based feature selection methods which achieve faster results specifically for high dimensional data.

Ji et al. [14] analyzed EEG signals recorded during sleep from 16 volunteers. In the study, they acquired features by applying the Fourier Transform and the short-time Fourier Transform to the EEG data for the detection of different states of sleeping. Afterwards, they carry out CFS and the genetic algorithm-based feature selection methods to reduce EEG features. As a result, the best results were found by CFS. Hu et al. [15] collected EEG signals to investigate the effects of attention on distance education methods. They reported that 63.90% was achieved for EEG data without feature selection while 80.84% was achieved by applying CFS for feature selection [15]. Onan A. and Korukoglu S. preferred filter based feature selection and wrapper based feature selection methods to reduce the number of features in text classification [17]. The highest result was yielded by filter based feature selection method with a classification accuracy of %89.72 [17].

In this experimental study, EEG data collected from 18 subjects were used, and then the signals were analyzed by HHT. Then HHT based features are classified by a probability based approach, Bayesian Network algorithm. To reduce the cost of operation and achieve better classification rates, consistency measures and correlation-based feature selection methods are applied and the classification performances are compared to the results obtained without feature selection.

II. MATERIALS AND METHOD

EEG data were obtained from 18 volunteer all-male university/college students. EEG data were collected from 32 electrodes with a sampling frequency of 1000Hz. Feature extraction was applied on EEG data by means of HHT, and then consistency-measures and CFS were applied for feature selection.

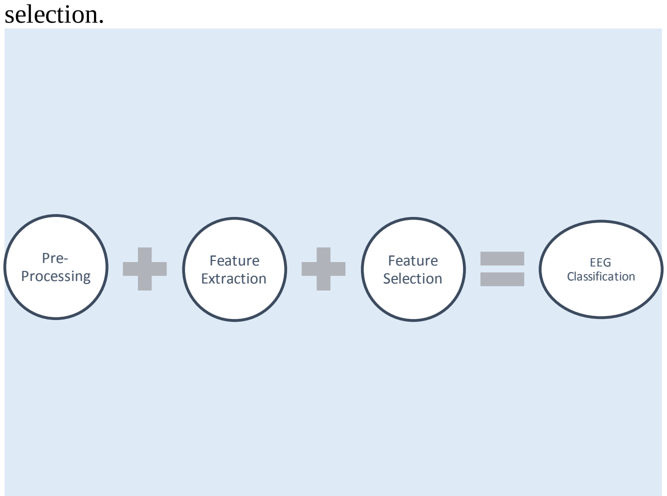

Fig. 1 The approach followed in the study

A. *Collection of EEG data*

The EEG markers were collected from volunteers using electrodes placed on the scalp as shown in Figure 2 in accordance with the international 10-20 system.

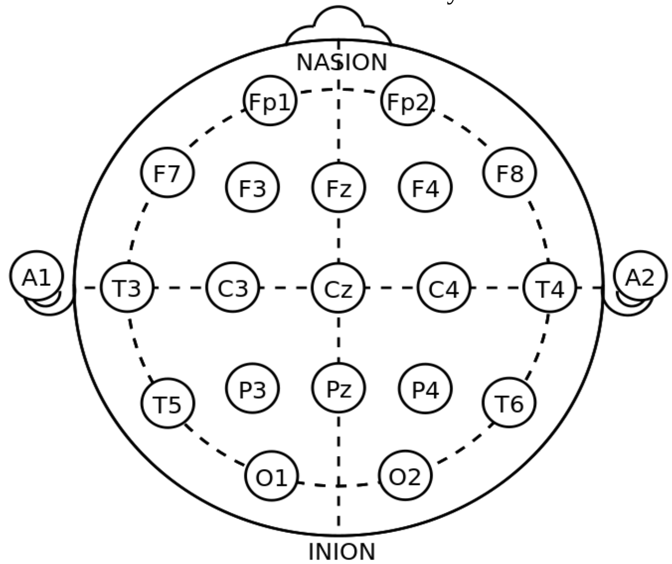

Fig. 2. The position of the electrodes relative to the 10-20 system for EEG recording (Top View)

The important considerations before and during the recording of EEG data are expressed as follows:

- Subjects are warned about their hair being clean and short, and not to use any hair styling products.
- Subjects were told not to take any medication prior to EEG recording.
- It has been stated that subjects should focus only on numerical slides or verbal slides on the screen.
- Subjects were told not to move their body parts such as hands, arms, head, legs and eyes during recording as shown in Figure 3. In addition, they were initially seated in a comfortable position.

A total of 60 slides, 30 numeric and 30 verbal texts, were shown during the recording. The duration of each slide is set to 13.25 seconds.

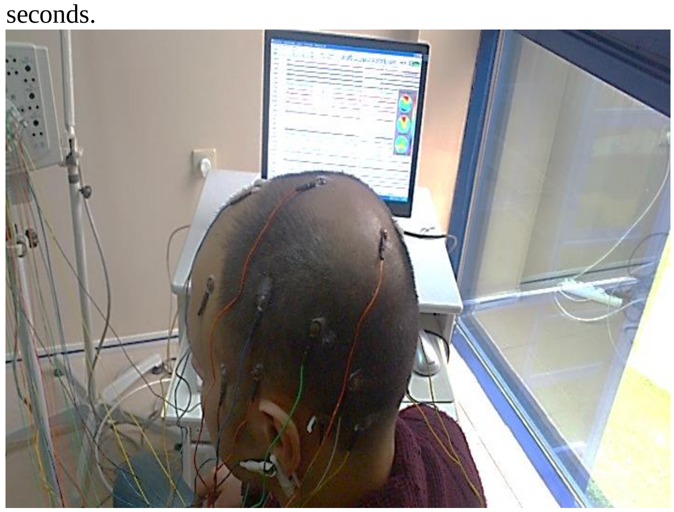

Fig. 3 EEG recording environment

Figure 4 is an example of a slide containing numerical operations.

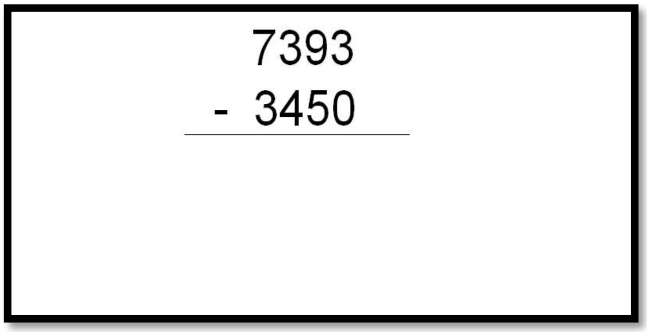

Fig. 4. A Numerical Slide Example

Figure 5 shows an example of slides containing verbal text.

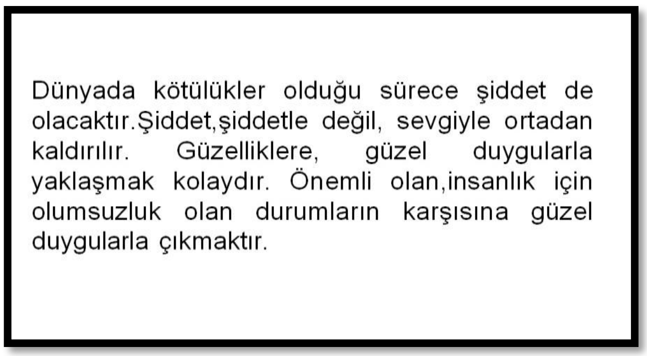

Fig 5. A Text Slide Example

B. Pre-Process

During recording, 50 Hz noise originating from the network is automatically cleaned by the device. After the recordings, signals are bandpass-filtered between 0.1 and 120 Hz and are divided into 60 pieces, each of which is a 13.25 second long EEG segment. Afterwards, the last two slides and the first slide in each section deactivated to prevent any problems due to synchronization. For this reason, 27 verbal and numerical slides were used for experimental study on EEG recordings.

C. Feature Extraction

Features are extracted from pre-processed EEG signals using HHT. HHT is an adaptive method that allows the analysis of non-linear, non-stationary signals as well as signals whose frequency and amplitude change by time [8]. By means of this method, complex, non-linear and non-stationary signals such as especially EEG signals, can be better analyzed than other signal processing methods. HHT is an empirical and adaptive method which combines Empirical Mode Decomposition (EMD) and Hilbert Transform. Hilbert Transform can be applied directly to single-component signals since these signals have only one frequency content at a given time interval. However, this is not possible for multi-component signals such as EEG. In order to solve this problem, Huang first divides the signal into the Intrinsic Mode Functions (IMFs) to include only one frequency value at a certain time [8]. Thereafter, the signal is expressed as the sum of these functions, each of which has a single frequency content instantaneously. In Empirical Mode Decomposition (EMD), the signal is divided into IMFs. In the last stage of the HHT, The Hilbert transform is applied to the obtained IMFs to generate the energy-frequency-time distribution known as the Hilbert spectrum.

EEG signals were analyzed with a sliding window method where each window is one second long with 50% overlap. For each segment, 312 features were extracted by calculating the average amplitude and maximum amplitude values for delta (0.5-4Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-30 Hz) and gamma (31-120 Hz) bands apart from the average amplitude and the maximum amplitude in the whole spectrum.

D. Feature Selection

Features were selected by applying Consistency Measures and Correlation-based Feature Selection (CFS) methods which are filter methods to the feature matrix extracted using HHT. In the analysis of large data, these methods have been used because of the reduction of training and test time.

Consistency Measures

This method was used in many studies in the literature for feature selection [18, 19, 20]. Consistency Measures (CM) is a filter based method that considers consistency levels of class values to examine the usefulness of the subset of attributes [21].

The CM method begins with a subset of all the attributes. Then, a random subset of attributes is generated from a subset space of attributes. If the randomly generated subset of attributes contains attributes less than or equal to the current subset of attributes, then the consistency ratings of the existing and newly created subset of attributes are compared. If the newly created subset of attributes has a better consistency rate, this subset is selected. The process is repeated during a parameter set by the user.

TABLE I.
NUMBER OF SELECTED FEATURES

No of Subject	Total Number of Features	Correlation based Feature Selection	Consistency Measures
Subject1	312	42	8
Subject 2	312	46	9
Subject 3	312	38	9
Subject 4	312	35	9
Subject 5	312	31	5
Subject 6	312	31	8
Subject 7	312	37	6
Subject 8	312	36	11
Subject 9	312	48	8
Subject 10	312	43	7
Subject 11	312	55	8
Subject 12	312	51	7
Subject 13	312	41	6
Subject 14	312	39	8
Subject 15	312	44	6
Subject 16	312	28	6
Subject 17	312	32	8
Subject18	312	32	6
Average	**312**	**39**	**8**

Correlation based Feature Selection

This method creates a new subset of attributes by selecting the attributes that have the highest correlation with the class within the set of attributes extracted from the signal. A subset of the

attributes which has the highest correlation with class but the lowest correlation each other is generated [22, 23, 24].
The number of selected attributes is shown in Table I for each Subject. When Table I is analyzed, it is seen that a very small number of attributes are selected in different numbers for each subject.

III. Results and Discussion

In this study, we reviewed the effect of feature selection algorithms such as CFS and CM on the experimental result, considering that HHT based features show better classification performance. The features were extracted by using HHT from EEG signals and were binary classified as arithmetical operation / silent reading.
The performance of Bayesian Network was worked out using the most commonly used parameters such as Accuracy, Precision and area under the ROC Curve (AUC) [25].

Examining Table II, it is seen obviously that arithmetical and reading operations were classified with an average accuracy of 92.60% and an average precision of 92.95%, without applying feature selection. Although the highest accuracy was 97.40%, the lowest accuracy is found as 85.50%. It is observed that there are a total of 13 subjects whose accuracy is 90% or better.

Table II
THE CLASSIFICATION RESULTS WITH ALL FEATURES

No of Subject	Accuracy	Precision	AUC
Subject 1	0.9100	0.9100	0.9630
Subject2	**0.8550**	**0.8630**	**0.9370**
Subject3	0.8590	0.8610	0.9380
Subject 4	0.8990	0.9040	0.9710
Subject 5	0.9670	0.9670	0.9910
Subject 6	0.9590	0.9590	0.9940
Subject7	**0.9740**	**0.9740**	**0.9920**
Subject 8	0.9010	0.9070	0.9380
Subject 9	0.9440	0.9470	0.9780
Subject 10	0.9490	0.9500	0.9880
Subject 11	0.9390	0.9400	0.9810
Subject 12	0.8870	0.8940	0.9650
Subject 13	0.9700	0.9700	0.9970
Subject 14	0.8770	0.8990	0.9880
Subject 15	0.9700	0.9710	0.9980
Subject 16	0.9630	0.9630	0.9920
Subject 17	0.9340	0.9380	0.9830
Subject 18	0.9110	0.9140	0.9700
Average	**0.9260**	**0.9295**	**0.9758**

The results while applying consistency measure are given in Table III. The average accuracy and precision were found as 93.98% and 94.08%, respectively with an average number of 8 features. Comparing the results, the features selected by consistency is more successful than all features in classification. The lowest classification accuracy was 86.60% and the lowest precision was 86.90% as well as the highest classification accuracy and precision were 97.90%. It is also seen that a total of 16 subjects are classified, with accuracy of 90% and above.

Table IV shows the classification results achieved when applied correlation based feature selection.

Table III.
THE CLASSIFICATION RESULTS USING CONSISTENCY MEASURES

No of Subject	Accuracy	Precision	AUC
Subject1	0.9300	0.9300	0.9800
Subject2	**0.8660**	**0.8690**	**0.9500**
Subject3	0.8770	0.8770	0.9540
Subject4	0.9410	0.9420	0.9890
Subject5	0.9720	0.9720	0.9970
Subject6	0.9300	0.9300	0.9860
Subject7	**0.9790**	**0.9790**	**0.9980**
Subject8	0.9100	0.9140	0.9730
Subject9	0.9400	0.9410	0.9830
Subject10	0.9630	0.9630	0.9960
Subject11	0.9460	0.9460	0.9860
Subject12	0.9010	0.9040	0.9740
Subject13	0.9700	0.9700	0.9960
Subject14	0.9450	0.9470	0.9910
Subject15	0.9720	0.9730	0.9970
Subject16	0.9580	0.9580	0.9930
Subject17	0.9390	0.9420	0.9910
Subject18	0.9770	0.9770	0.9950
Average	**0.9398**	**0.9408**	**0.9849**

Table IV.
THE CLASSIFICATION RESULTS USING CORRELATON BASED FEATURE SELECTION

No of Subject	Accuracy	Precision	AUC
Subject1	0.9390	0.9400	0.9850
Subject2	0.8900	0.8930	0.9740
Subject3	**0.8670**	**0.8680**	**0.9510**
Subject4	0.9680	0.9680	0.9970
Subject5	**0.9860**	**0.9860**	**0.9990**
Subject6	0.9570	0.9570	0.9950
Subject7	0.9810	0.9820	0.9990
Subject8	0.9120	0.9190	0.9890
Subject9	0.9590	0.9610	0.9940
Subject10	0.9830	0.9830	0.9980
Subject11	0.9760	0.9760	0.9910
Subject12	0.9360	0.9380	0.9890
Subject13	0.9840	0.9850	0.9990
Subject14	0.9630	0.9630	0.9950
Subject15	0.9810	0.9820	0.9990
Subject16	0.9820	0.9820	0.9990
Subject17	0.9620	0.9640	0.9970
Subject18	0.9790	0.9790	0.9960
Average	**0.9558**	**0.9570**	**0.9914**

The average accuracy and precision were found as 95.58% and 95.70% as shown in Table IV. The results were found to be higher than the result gained using all the features. In addition, it is demonstrated that the poorest accuracy and precision values were 86.70% and 86.80%, respectively for Subject3 even though Subject5 had the highest precision and the accuracy value were 98.60%. In general, it is clearly seen that the findings in Table IV are more successful than those of Table II and Table III.
Examining the all results, for BCI systems, it is indicated that higher classification performance can be achieved by reducing the number of features with correlation-based feature selection. As a result of the all results, it is understood that the correlation-based feature selection method is more compatible with the

Bayesian Network algorithm for the features from database we used in this study.
The results show the importance of selecting features for the development of faster and higher performance BCI systems and the method to be selected for this process.

IV. CONCLUSION

In this paper, we presented an approach to examine the effect of feature selection algorithms on cognitive tasks based on EEG signals.
The features were extracted by using Hilbert Huang Transform from EEG signals and were binary classified as arithmetical operation / silent reading. For feature selection, we used Correlation based Feature Selection and Consistency Measures algorithm which are filter methods. Bayesian Network was employed for classification. Effect of feature selection algorithms are evaluated for this cognitive task analysis. The classification results indicated that Feature Selection algorithms have a positive effect on EEG signal classification performance. CFS and CM feature selection algorithms are a powerful and useful tools to select EEG features.

REFERENCES

[1]. T. Aflalo. "Decoding motor imagery from the posterior parietal cortex of a tetraplegic human." Science. vol. 348. no. 6237. pp. 906–910. May 2015.
[2]. E. Demirci, "Playing games with brain waves". TUBITAK Science Technical Journal. 44 (520). 18-24p, 2011
[3]. L. R. Hochberg. "Reach and grasp by people with tetraplegia using a neurally controlled robotic arm." Nature. vol. 485. no. 7398. pp. 372–375. 2012.
[4]. F. Cabestaing. T. M. Vaughan. D.J. Mcfarland..J. R. Wolpaw, "Classification of evoked potentials by Pearson ' s correlation in a Brain-Computer Interface". Matrix. 67. 156-166pp, 2017.
[5]. A. Dogan. M.H. Calp .E.S. ARI. H. Ozkose. "An examination on Brain Computers Interfaces in Human Computer Interaction:Characteristics and Working Principle"
[6]. E. Sevinç. "Brain Computer Interfaces", http://www.rehabilitasyon.com/action/makale/1/Beyin_Bilgisayar_Arayuzleri-2299 (2006)
[7]. O. Rioul, and M. Vetterli. "Wavelets and signal processing." IEEE signal processing magazine 8.4 (1991): 14-38.
[8]. N. E. Huang et al.. "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis." Proc. R. Soc. A Math. Phys. Eng. Sci.. vol. 454. no. 1971. pp. 903–995. Mar. 1998.
[9]. J. Kortelainen, E. Vayrynen, U. Huuskonen, J. Laurila, J. Koskenkari, J.T. Backman, S. Alahuhta, "Using Hilbert-Huang Transform to assess EEG slow wave activity during anesthesia in post-cardiac arrest patients.", 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Orlando. FL. 2016. pp. 1850-1853.
[10]. R. Wang. Y. Wang and C. Luo. "EEG-Based Real-Time Drowsiness Detection Using Hilbert-Huang Transform." 7th International Conference on Intelligent Human-Machine Systems and Cybernetics. Hangzhou. 2015. pp. 195-198.
[11]. K. Rai. V. Bajaj and A. Kumar. "Hilbert-Huang transform based classification of sleep and wake EEG signals using fuzzy c-means algorithm." International Conference on Communications and Signal Processing (ICCSP). Melmaruvathur. 2015. pp. 0460-0464.
[12]. R. Swarnalatha. and D. V. Prasad, "Detection of Sleep Bruxism Based on EEG Hilbert Huang Transform".5th International Conference on Biomedical Engineering and Technology (ICBET 2015) .IPCBEE vol.81 (2015) (2015) IACSIT Press. Singapore
[13]. R. Jenke. A. Peer and M. Buss. "Feature Extraction and Selection for Emotion Recognition from EEG." in IEEE Transactions on Affective Computing. vol. 5. no. 3. pp. 327-339. July-Sept. 1 2014.
[14]. Y. Ji, X. Bu, J. Sun and Z. Liu, "An improved simulated annealing genetic algorithm of EEG feature selection in sleep stage." 2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA), Jeju, 2016, pp. 1-4.
[15]. B. Hu, X. Li, S. Sun and M. Ratcliffe, "Attention Recognition in EEG-Based Affective Learning Research Using CFS+KNN Algorithm." in IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. PP, no. 99
[16]. C. Cîmpanu. L. Ferariu. T. Dumitriu and F. Ungureanu. "Multi-Objective Optimization of Feature Selection procedure for EEG signals classification.". 2017 E-Health and Bioengineering Conference (EHB). Sinaia . 2017. pp. 434-437.
[17]. A. Onan, S. Korukoglu. "Evalution of feature selection methods in text classification", Academic Knowledge 2016, Turkey (2016).
[18]. S. Zhang1, Z. Zhao. "Feature Selection Filtering Methods for Emotion Recognition in Chinese Speech Signal". 2008
[19]. L. Jalali. M. Nasiri and B. Minaei. "A hybrid feature selection method based on fuzzy feature selection and consistency measures." 2009 IEEE International Conference on Intelligent Computing and Intelligent Systems. Shanghai. 2009. pp. 718-722.
[20]. B. Chakraborty and G. Chakraborty. "Fuzzy Consistency Measure with Particle Swarm Optimization for Feature Selection." 2013 IEEE International Conference on Systems. Man. and Cybernetics. Manchester. 2013. pp. 4311-4315
[21]. Liu, H., Setiono, R., "A probabilistic approach to feature selection: a filter solution", Proceedings of the Thirteenth International Conference on Machine Learning, 319-327 (1996).
[22]. S. Zhang. and Z. Zhijin, "Feature selection filtering Methods for emotion recognitionin Chinese speech signal.". Signal Processing. 2008.
[23]. T.J. Lee, S.M. Park, K.B. Sim, "Electroencephalography Signal Grouping and Feature Classification Using Harmony Search for BCI". Journal of Applied Mathematics 2013. pp. 1-9.
[24]. S. G. Eraldemir, E. Yildirim, S. Yildirim, Yakup Kutlu, "Feature selection for cognitive EEG signals based channel selection and classification". Innovations and Applications in Intelligent Systems. 2014 (ASYU 2014). 1-6.
[25]. A. Subaşı, M. I. Gursoy, M. "EEG signal classification using PCA, ICA, LDA and support vector machines. Expert Systems with Applications", 37(12), 8659–8666. https://doi.org/10.1016/j.eswa.2010.06.065, 2010

BIOGRAPHIES

Server Göksel ERALDEMİR was bron in iskenderun, Turkey. He received the Bachelor degree in Computer Education and Instructional Technology from the University of OMU, Turkey, in 2002. Then he worked computer teacher in Reyhanlı high school from 2002 to 2009 and After in late 2009 he started to work as a lecturer at the university. Then he received Masters Degree in electrical and electronics engineering from the University Mustafa Kemal, in 2014. His current research interests are biomedical signal processing and data mining.

BIOGRAPHIES

MUSTAFA TURAN ARSLAN was born in Kilis, Turkey. He received the B.S. and M.S. degrees in computer engineering from the University of Erciyes, Turkey, in 2011 and 2016, respectively. Currently, he has been a Ph.D. student in electrical-electronics engineering department in Adana Science and Technology University since February 2018. His current research interests are biomedical signal processing, data mining and artificial intelligence.

BIOGRAPHIES

ESEN YILDIRIM, received the B.S. degree in Electrical and Electronics Engineering from Cukurova University, Adana, Turkey, in 1997, and the M.S. and Ph.D. degrees in Electrical Engineering from University of Southern California (USC), Los Angeles, in 2000 and 2006, respectively. She is currently an Associate Professor of Electrical and Electronic Engineering, at Adana Science and Technology University, Adana, Turkey. Her general research interests include biomedical signal processing, epileptic seizure detection, functional connectivity, learning methods, and emotion recognition from physiological signals.