ISSN 1870-4069

Feature Subset Selection in Electroencephalographic Signals Using Typical Testors

Alexis Gallegos, Dolores Torres, Aurora Torres, Eunice Ponce de León

Benémerita Universidad Autónoma de Aguascalientes, Aguascalientes, Mexico

alexisedm@gmail.com
{mdtorres, atorres, eponce}@correo.uaa.mx

Abstract. Motor imagery (MI) is a mental representation of movement without performing or tensing any muscles. MI requires a conscious activation of the same brain regions involved in actual movement. Brain signals have been explored for multiple applications in biomedical engineering, such as the development of brain-computer interfaces (BCI). BCI systems are designed to translate users' intentions into control signals, commands, or codes. Nevertheless, the major challenge in BCI system development is classifying MI signals recorded by an electroencephalogram (EEG). This paper focuses on applying the testor theory and the logical combinatorial pattern recognition approach for feature selection to reduce the feature representation space for classification tasks. The EMOTIV EPOC+ EEG device recorded the MI-EEG signals with 14 electrodes.

Keywords. Typical testors, feature subset selection, electroencephalographic signals, motor imagery

1 Introduction

This paper focuses on the application of testor theory and the logical combinatorial pattern recognition approach for feature selection. The problem of selecting the subset of features that best describes a phenomenon from a larger set allows to reduce the size of solution space, so that results close to the optimum or the optimum itself are obtained with less resources (time and memory) [1].

Therefore, the aim of the paper was to reduce the MI-EEG (Motor Imagery Electroencephalographic) signals feature representation space for classification tasks. These signals were recorded by the Emotiv EPOC+ device which describes the signals by means of 14 electrodes distributed over the scalp.

The EEG signals represent the electrical brain activity created by billions of neurons [2]. This activity represents the communication between the body and the brain. The analysis of EEG signals is highly relevant in health research for diagnosis, treatment, and monitoring of different diseases [2], [3].

On the other hand, motor imagery, or MI, is a mental performance of movement without any physical activation. The movements analyzed were opening and closing of the hand. The practice of MI is used in the context of sports as well as in rehabilitation treatments because it requires the activation of the same brain areas. It has also shown positive results in learning of physical skills and strength gain [4].

The present document is composed of three more sections. The second section describes the basic concepts used: typical testors, electroencephalographic signals and motor imagery. Section 3 provides the framework that allowed for the typical testors analysis. The framework includes the recording of MI-EEG signals taken with the support of six test subjects, the preprocessing of the data and the selection of the minimal subset of feature that will allow for a correct classification.

Finally, section 4 provides the typical testors found, which represent the minimum subset of features for describing objects. In addition, the informational weight (IW), which represents a measure of significance for each feature involved, is described. Thus, the higher its value, the more determinant it is to differentiate classes of objects.

2 Important Concepts

2.1 Typical Testors

Technological advances have led to the generation of large amounts of data at an unprecedented speed. These data describe objects or phenomena with high number of features, resulting in a challenge for machine learning and data mining research [5]. Feature Selection is the area of pattern recognition *"responsible for identifying those features that provide relevant information for classification purposes"* [6].

There are different tools or methodologies applied in addressing feature selection. One of them is the logical combinatorial pattern recognition approach, where testor theory is applied for such task [6]–[8].

Testor theory was introduced in the 60s to locate faults in electronic circuits [7].

Dimitriev, Zhuravlev and Krendeleiev's testor theory approach [9] for classification and feature selection establishes that classes are disjoint sets, the criterion of comparison between features is Boolean, and the criterion of similarity between objects accepts that two objects are different if at least one of their features is also different [10].

Testor theory defines a testor as a feature subset that does not confuse object belonging to different classes. I.e., no object belonging to class T_0 can be confused with any object of class T_1 according to the values in its features [11].

According to Shulcloper et al. [12], the definition of testor can be extended to more than two classes.

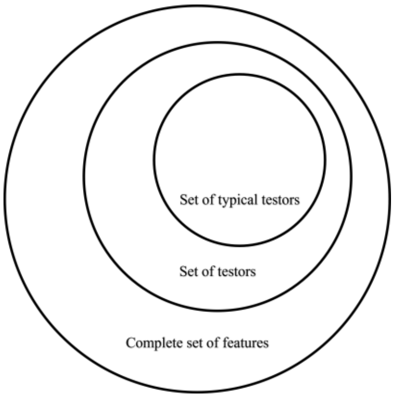

Fig. 1. Sets of features, testors and typical testors [8].

Within the set of testors there is the set of typical testors (see Fig. 1), also known as irreducible testors. A testor can be a typical testor if by eliminating one of its features, the remaining subset is no longer a testor. Therefore, a typical testor is the minimum feature subset needed to distinguish objects of different classes [10], [12].

The importance of the typical testor calculation lies in the reduction of the feature representation space, feature selection, as support for classification and pattern recognition systems [10].

As described so far, the main aim of the testor theory is feature selection. However, testor theory can be used to determine the relevance of each feature by computing the informational weight from the set of typical testors [13]. The informational weight is calculated by means of the relative frequency. Let be τ the number of typical testors found and $\tau(i)$ the number of typical testors in which the feature x_i appears, the informational weight is given by [8]:

$$P(x_i) = \tau(i) / \tau. \quad (1)$$

The obtained score represents a measure of significance for each feature. This means that the higher the score of the feature, the greater its relevance in class distinction [8], [14], [15].

2.2 Electroencephalographic Signals

Brain is a complex part of the human body which plays an important role for controlling behavior of human body according to different stimuli. The study of the functional and cognitive behavior of the human brain has been an important area of medical research to find better diagnoses and treatments for brain related issues [3]. These studies can be performed by processing electrical brain activity, specially through computational modeling. The electrical brain activity is created by billions of interconnected neurons across different areas of the brain [16]. These neurons *"act as information carriers between the body and brain"* [3]. Voltage potential resulting from current flow in and around neurons can be recorded by electrodes placed on the scalp and reported as electroencephalographic (EEG) signals [16], [17].

As a definition*, "electroencephalography (EEG) is the non-invasive measurement of the brain's electric fields"* [17]. According to Keenan et al. [18], the EEG recording allows for data analysis from the frequency and amplitude domains, i.e., time and voltage. EEG signals are complex because they correspond to a mixture of information, physiological artifacts (eye movements, muscle movements, heartbeats, sweat) or technical artefacts (power supply line, electrode disconnection) [19].

Electroencephalography has applications in several domains such as health, education and, entertainment [16]. For example, EEG is decisive for in the diagnosis, treatment and monitoring of epileptic syndrome patients, and the study of sleep patterns, depth of anesthesia, and attention deficit hyperactivity disorder. There are also applications in cognitive and affective monitoring such as level of fatigue, mental workload, mood, or emotions, even stress control [16], [20], [21]. Outside the medical field, there are also applications such as BCI systems that allow the translation of EEG signal patterns into messages or commands for applications or interactive devices [16] with the aim of making human computer interaction more natural, especially for people with neuro-muscular disabilities [22].

2.3 Motor Imagery

Motor imagery (MI) is defined by Mokiento et. al. [23] as the mental performance of movements without by any kind of peripheral muscular activity. I.e., MI is a mental representation of movement without any body movement [4], e.g., opening or closing the left or right hand without executing it [24]. The practice of MI requires the conscious activation of the same brain regions involved in the preparation and execution of movement. In addition, MI allows motor development and the learning of motor skills, including the gain of strength in specific muscle groups [4], [25].

MI has been applied in the sports context and positive effects have been reported in speed, performance accuracy, muscle strength, movements dynamics and motor skill performance. From the medical point of view, there have been positive results in rehabilitation in patients with neurological conditions, e.g., stroke, spinal cord injury, or Parkinson disease [4].

Computer science has studied motor imagery for the development of brain computer interface (BCI). A BCI system allows the communication between the brain and external devices without the involvement of peripheral nerves or muscles [26]. The most important BCI systems are the MI-based BCI systems, which involves real-time applications such as contactless writing, prosthetic arms, virtual reality systems, gamming apps, wheelchairs, etc.[27].

The major challenge in the design of MI-based BCI systems is the classification of EEG signal due to events such as eye blink, eye movement, muscular movements, teeth grinding, and heart rhythm interfere with the EEG signal recording resulting in a noisy signal [28]. Consequently, multiple frameworks have been developed in the literature with the aim of making EEG signal processing more efficient [28], [29].

3 Framework

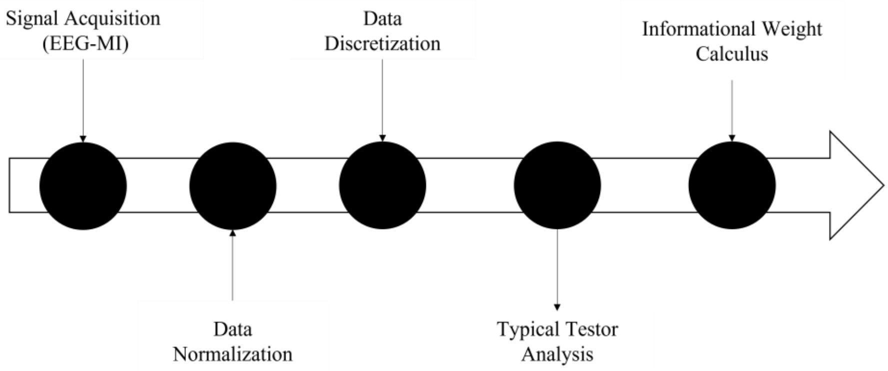

Fig. 2. Framework for typical testors analysis.

This section of the document describes the framework used to perform the typical testors analysis applied to EEG signals with motor imagery. This framework aims to select a feature subset thar allows the correct discrimination of signals corresponding to the motor imagery of "open" and "close" the right hand.

As shown in Figure 2, the process began with the EEG signal recording by means of the Emotiv EPOC+ EEG device with 14 channels (AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4), distributed according to the 10-20 electrode placement system, and a sampling rate of 128 samples per second.

Fig. 3. Emotiv EPOC+ device, 14 signals and 10-20 system [30].

Six test subjects were used for signal sampling. For each one, six five-second samples were taken: three correspond to the intention to "open" and three more to "close" the hand. Additionally, the open-source application Cykit was used to record EEG signals in csv files, which were used as parameters for the following phases of the methodology.

The second phase of the framework consisted of normalizing the data set since, although the fourteen features describe a brain electrical signal, there are small variations that may cause some features to be dominated by others. To avoid this issue, every feature was standardized using z-score normalization. In this way, the set of features has the same scale.

The EEG signals are recorded with continuous values; therefore, the third phase of the framework was to apply a discretization. This, in addition, with the objective of making the typical testors analysis easier.

Finally, the phases four and five properly involve a feature subset selection (FSS) process using the logical combinatorial approach by means of typical testor analysis (see section 2.1). This process used in data mining provides tools for the efficient reduction of the number of features describing objects, with the purpose of removing irrelevant features resulting in more stable representations.

The results obtained from the application of the framework described above are shown in section 4 below.

4 Results and Conclusions

As a result of phase 1, a database of 23,846 records described by the signals from the 14 electrodes of the Emotiv device was obtained. Specifically, the database included 12,190 records corresponding to the intention to open the hand and 11,656 to the intention to close the hand.

As mentioned in the previous section, the data set was preprocessed through a normalization and discretization process. Once completed, a random sampling of 500 records per class was performed, ending the preprocessing a smaller data set that will be the parameter for the feature subset selection application, i.e., the computing of typical testors.

Table 1. Testors and Typical Testors

Typical Testors													
AF3	**F7**	**F3**	**FC5**	**T7**	**P7**	**O1**	**O2**	**P8**	**T8**	**FC6**	**F4**	**F8**	**AF4**
1	1	1	1	1	1	1	1	1	1	1	0	1	1

The typical testors were calculated by means of a Python library of the exhaustive method using the Python library TestoresTipicos.py developed by Daniel Barajas, PhD student at Autonomous University of Aguascalientes, Mexico. The testors and typical testors found are shown in Table 1. As can be seen, the typical testors are represented as binary values where zero represents the absence of the feature, i.e., it does not provide relevant information and one means that the feature provides essential information.

For this study, two testors were obtained, one of which is a typical testor. This typical testor is the minimum feature subset needed to distinguish objects of different classes ("open" and "close"). This typical testor is used to calculate the informational weight of each feature as described in section 2.1.

Table 2. Informational Weight

Feature	IW	Feature	IW
AF3	100%	**O2**	100%
F7	100%	**P8**	100%
F3	100%	**T8**	100%
FC5	100%	**FC6**	100%
T7	100%	**F4**	0%
P7	100%	**F8**	100%
O1	100%	**AF4**	100%

Table 2 shows the informational weight calculated from the typical testor resulting from Table 1. Each percentage represents a measure of significance for each feature involved. In this sence, 13 of the 14 features are essential (100% of informational weight) to defferentiate the two classses, while the feature F4 with 0% of informational weight presents no relevant information for this process.

In this manner, the objective of reducing the dimensionality of the problem is achieved by describing the clases with a smaller number of features, making the representation of objects easier and as a support for classification systems.

As future work, it is proposed to test the typical testor found by evaluating the performance of different machine learning algorithms sucha as artificial neuronal networks or support vector machines.

References

[1] M. D. Torres Soto, A. Torres Soto, and E. Ponce de León Sentí, “Algoritmo Genético y Testores Típicos en el Problema de Selección de Subconjuntos de Características.” 2006, [Online]. Available: http://www.iiisci.org/journal/CV$/risci/pdfs/C415DR.pdf.

[2] W. H. Wolberg, W. N. Street, D. M. Heisey, and O. L. Mangasarian, "Computer-derived nuclear features distinguish malignant from benign breast cytology," *Hum. Pathol.*, vol. 26, no. 7, pp. 792–796, 1995, doi: http://dx.doi.org/10.1016/0046-8177(95)90229-5.

[3] J. S. Kumar and P. Bhuvaneswari, "Analysis of Electroencephalography (EEG) Signals and Its Categorization–A Study," *Procedia Eng.*, vol. 38, 2012, doi: 10.1016/j.proeng.2012.06.298.

[4] R. Dickstein and J. E. Deutsch, "Motor Imagery in Physical Therapist Practice," *Phys. Ther.*, vol. 87, no. 7, pp. 942–953, Jul. 2007, doi: 10.2522/ptj.20060331.

[5] J. Cai, J. Luo, S. Wang, and S. Yang, "Feature selection in machine learning: A new perspective," *Neurocomputing*, vol. 300, Jul. 2018, doi: 10.1016/j.neucom.2017.11.077.

[6] V. I. González Guevara, S. Godoy Calderon, E. Alba Cabrera, and J. Ibarra Fiallo, "A Mixed Learning Strategy for Finding Typical Testors in Large Datasets," in *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, 1st ed., Montevideo, Uruguay: Springer, 2015, pp. 716–724.

[7] E. Alba-Cabrera, R. Santana, A. Ochoa, and M. Lazo-Cortés, "Finding typical testors by using an evolutionary strategy," *Proc. Fith Ibero Am. Symp. Pattern Recognit.*, 2000.

[8] A. E. Gallegos Acosta, F. J. Álvarez Rodríguez, M. D. Torres Soto, and A. Torres Soto, “Identificación de factores de riesgo en patologías médicas mediante métodos de selección de subconjuntos de características [recurso electrónico],” Universidad Autónoma de Aguascalientes, Aguascalientes, México, 2018.

[9] A. N. Dmitriev, Y. I. Zhuravlev, and F. P. Krendeleiev, "On the mathematical principles of patterns and phenomena classi-" cation," *Diskretn. Anal*, vol. 7, no. 3, p. 15, 1966.

[10] Y. Santiesteban Alganza and A. Pons Porrata, “Lex: Un nuevo algoritmo para el calculo de los testores tipicos,” *Ciencias Mat.*, vol. 21, no. 1, Jan. 2003, [Online]. Available: https://dibpxy.uaa.mx/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=edsgii&AN=edsgcl.147016161&lang=es&site=eds-live&scope=site.

[11] J. Ochoa Somuano, “Técnicas de Selección de Atributos para la Categorización Automática de Escenas Visuales,” Centro Nacional de Investigación y Desarrollo Tecnológico, Cuernavaca, Morelos, 2005.

[12] J. Ruíz Shucloper, E. Alba Cabrera, and M. Lazo Cortés, “Introducción a la Teoría de Testores.” Departamento de Ingeniería Electrica, CINVESTAV-IPN, p. 197, 1995.

[13] G. Díaz Sánchez *et al.*, "Typical Testors Generation Based on an Evolutionary Algorithm," *Intell. Data Eng. Autom. Learn. - IDEAL 2011*, vol. 6936, Sep. 2011, doi: 10.1007/978-3-642-23878-9.

[14] J. A. Santos, A. Carrasco, and J. F. Martínez, “Feature Selection using Typical Testors applied to Estimation of Stellar Parameters,” *Computación y Sistemas*, vol. 8. Instituto Politécnico Nacional, México, pp. 15–23, 2004.

[15] M. D. Torres, E. Ponce de León, C. A. Ochoa, A. Torres, and E. Díaz, “Mecanismos de Aceleración en Selección de Características Basada en el Peso Informacional de las Variables para Aprendizaje no Supervisado,” *Revista Iberoamericana de Sistemas, Cibernética e Informática*, vol. 6. Revista Iberoamericana de Sistemas, Cibernética e Informática, pp. 29–34, 2009.

[16] G. Coro, G. Masetti, P. Bonhoeffer, and M. Betcher, "Distinguishing Violinists and Pianists Based on Their Brain Signals," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2019, vol. 11727 LNCS, pp. 123–137, doi: 10.1007/978-3-030-30487-4_11.

[17] A. Biasiucci, B. Franceschiello, and M. M. Murray, "Electroencephalography," *Curr. Biol.*, vol. 29, no. 3, Feb. 2019, doi: 10.1016/j.cub.2018.11.052.

[18] S. A. Keenan, O. Carrillo, and H. Casseres, "Electroencephalography," in *Encyclopedia of Sleep*, Elsevier, 2013.

[19] A. Santillan Guzman, H. Z. Ramírez Uriarte, J. J. Oliveros Oliveros, M. M. Morin Castillo, and H. Ramírez Díaz, “Interfaz Gráfica Intuitiva para el Procesamiento de Señales EEG,” *Memorias DEl Congr. Nac. Ing. Biomédica*, vol. 5, no. 1, pp. 146–149, 2018, doi: dx.doi.org/10.24254/CNIB.18.20.

[20] Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T. H. Falk, and J. Faubert, "Deep learning-based electroencephalography analysis: a systematic review," *J. Neural Eng.*, vol. 16, no. 5, p. 51001, Aug. 2019, doi: 10.1088/1741-2552/ab260c.

[21] F. Ramos Argüelles, G. Morales, S. Egozcue, R. M. Pabón, and M. T. Alonso, “Técnicas de electroencefalografía: principios y aplicaciones clínicas,” *An. Sist. Sanit. Navar.*, vol.

32, pp. 69–82, Jun. 2009, [Online]. Available: http://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S1137-66272009000600006&nrm=iso.

[22] V. P. Oikonomou, K. Georgiadis, G. Liaros, S. Nikolopoulos, and I. Kompatsiaris, "A Comparison Study on EEG Signal Processing Techniques Using Motor Imagery EEG Data," in *2017 IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS)*, Jun. 2017, doi: 10.1109/CBMS.2017.113.

[23] O. A. Mokienko, L. A. Chernikova, A. A. Frolov, and P. D. Bobrov, "Motor Imagery and Its Practical Application," *Neurosci. Behav. Physiol.*, vol. 44, no. 5, Jun. 2014, doi: 10.1007/s11055-014-9937-y.

[24] Z. Tayeb *et al.*, "Validating Deep Neural Networks for Online Decoding of Motor Imagery Movements from EEG Signals," *Sensors*, vol. 19, no. 1, Jan. 2019, doi: 10.3390/s19010210.

[25] T. Mulder, "Motor imagery and action observation: cognitive tools for rehabilitation," *J. Neural Transm.*, vol. 114, no. 10, Oct. 2007, doi: 10.1007/s00702-007-0763-z.

[26] S. Kumar and A. Sharma, "A new parameter tuning approach for enhanced motor imagery EEG signal classification," *Med. Biol. Eng. Comput.*, vol. 56, no. 10, Oct. 2018, doi: 10.1007/s11517-018-1821-4.

[27] S. R. Sreeja and D. Samanta, "Classification of multiclass motor imagery EEG signal using sparsity approach," *Neurocomputing*, vol. 368, Nov. 2019, doi: 10.1016/j.neucom.2019.08.037.

[28] S. Kumar, A. Sharma, K. Mamun, and T. Tsunoda, "A Deep Learning Approach for Motor Imagery EEG Signal Classification," in *2016 3rd Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE)*, Dec. 2016, doi: 10.1109/APWC-on-CSE.2016.017.

[29] S. R. Sreeja, J. Rabha, K. Y. Nagarjuna, D. Samanta, P. Mitra, and M. Sarma, "Motor Imagery EEG Signal Processing and Classification Using Machine Learning Approach," in *2017 International Conference on New Trends in Computing Sciences (ICTCS)*, Oct. 2017, doi: 10.1109/ICTCS.2017.15.

[30] J. Gwizdka, R. Hosseini, M. Cole, and S. Wang, "Temporal dynamics of eye-tracking and EEG during reading and relevance decisions," *J. Assoc. Inf. Sci. Technol.*, vol. 68, Aug. 2017, doi: 10.1002/asi.23904.