Analyzing brain signals using decision trees: an approach based on neuroscience

Josimara de A. Silveira[1], **Diana F. Adamatti**[1], **Fernanda A. H. Carvalho**[1]

[1]Programa de Pós-Graduação em Modelagem Computacional
Universidade Federal do Rio Grande (FURG) - Brasil

josisilver@hotmail.com, dianaada@gmail.com, fahc@vetorial.net

Abstract. *This paper presents a case study of treatment of brain signals using decision trees to classify of these signals, and they are analysed based on neuroscience. We have collected brain signals for 3 subjects during an imagination task and we classify these signals using decision trees, a supervised machine learning method. To analyse the processing data in neuroscience, we defined the matching between the electrodes position and the corresponding functions. The results are promising, because we can better understand the brain e its functionalities.*

1. Introduction

The man-machine interfaces, such as keyboard and mouse on computers, cameras and joystics on video games, are commonly used by the most of people. However, there is a part of population that they may need other ways to interact with electronic equipment due to health problems or accidents, that they lead to severe paralysis, impeding the use of common interfaces. The development of Brain Computer Interfaces (BCI) aims to provide this communication without the use of muscle movements [Wolpaw et al. 2002].

BCI systems are a communication method based on neural activity generated by the brain, independing of the achievement of the muscle movements. In the case of people with severe paralysis, and that they have not compromised the brain, the neural activity to perform a movement is generated, but it does not reach the corresponding muscles, or those muscles are weak to obey such commands [Vallabhaneni et al. 2005]. There are several interfaces designed for people with disabilities [Wickelgren, 2003 *apud* [Vallabhaneni et al. 2005]]. Most of these systems, however, use some type of muscle control, as the movement of the head, neck, eyes or other facial muscles.

Neural activity used in a BCI system can be recorded using invasive or non-invasive techniques. While requiring only brain activity, BCI systems use signals are generated voluntarily by the user. Interfaces based on involuntary brain activities as generated during a seizure, utilize many of the same principles and components of BCI systems, but are not included in this area.

The Neuroscience studies the most varied methods nervous system and defines coordination as the ability to produce muscle contraction resulting in a movement biologically appropriate to the context. The performance on unknown motion increases as the subject trains the realization of such action [Pfurtscheller and da Silva 1999].

Decision trees (DT) are classifiers that adopt a conditional tree structure, where each node specifies a test to be performed on a single attribute [Duda et al. 2002, Mitchell 1997, Alpaydin 2010]. DT are tools for classification and prediction. They provide rules that explain the behavior of the system being modeled, and provide a classification model. One of the most important features of decision trees is that rules inferred through them are easy to understand.

Apply decision trees in the context of BCI systems is an approach that is presented. There are some studies that discuss this technical type using the feature vectors extracted from brain signals of motor imagination [Koprinska 2010, Aparna et al. 2010]. However, the methodology presented in this work does not involve the creation of feature vectors.

This paper presents the analysis of brain signals collected with the Emotiv Epoc[1], using decision trees for the classification, the DT generated are analysed based on neuroscience, with the help of an expert. We have checked electrodes activated during the task accomplishment of imagination of hand movement from decision trees and we observed the functions of the brain regions where the electrodes are placed.

The work is divided into five sections: in Section II is presented some concepts of neuroscience and BCI systems and DT. The Section III presents the methodology. In the Section IV presents the results obtained and Section V concludes the paper.

2. Theoretical Background

2.1. Neuroscience

Neuroscience is the set of disciplines that study the most varied methods nervous system, its structure, development, function, evolution, compared with the behavior and the mind, and also their changes. It defines coordination as the ability to produce muscle contraction in concert, resulting in a movement biologically appropriate to the context. The realization of a movement emerges as the subject uses complex dynamic parameters with diagrams or motor representations constructed by the association of information generated in different brain areas [Farwell and Donchin 1988].

From the anatomical point of view, the human brain is divided into two hemispheres, right and left. Each hemisphere has the following components: one thalamus, one hippocampus, an amygdala, basal ganglia, one cerebral cortex and four lobes (occipital, parietal, temporal and frontal), according Figure 1. The occipital lobe is primarily involved with the sense of sight; the parietal includes information coming from the sense of touch and somatosensory perception; the temporal lobes process audio data; and the frontal lobes is essential for planning cognitive actions and movement.

The last portion of the frontal lobes is called the primary motor cortex and it is involved in the execution of movements. The remainder is called the prefrontal cortex, and it occupies about 30 % of the human brain and it is responsible for rational behavior, because that is where we process sensory information that allow us to perceive stimuli, dealing with information and emotions, plan actions and understand and issue appropriate behaviors [Ratey 2002, Gazzaniga and Heatherton 2005, Lent 2008].

Experiments revealed that the specialities of the hemispheres may differ, but rarely this specialization means exclusivity functional. The left hemisphere is associated with global functions and the right one to specific functions [Lent 2001].

[1] https://emotiv.com/epoc.php

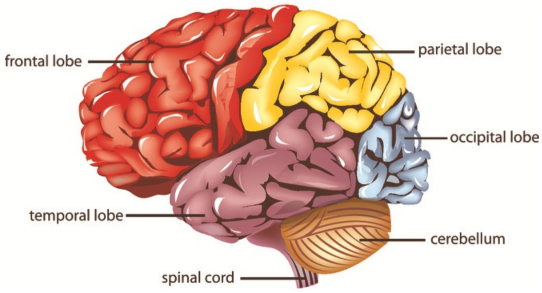

Figure 1. Parts of the human brain.

The procedural memory is habits, skills and rules, something that often memorized without feeling, and it is used without being aware. During the learning of a new task, many specific circuits of the frontal cortex are activated, but on repeat playback on a reasonable number of times, the standards acquire stability and behavior becomes automatic, requiring no more conscious attention. It is a procedure that, “ filed” in memory, acquires great strength and durability[Ratey 2002].

In this way, the regularity of the sequence of movements and the performance of an individual to perform a movement that it was unusual for him/her gradually improves. Motor skills employed in order to achieve goals result from implicit memories and it is not required attention, happen automatically, without the awareness of remembering something and effortlessly deliberative [Lent 2001].

About regard to motor activities, the regular execution leads to building a framework that can be used whenever the activity requires certain procedural domain. However, it is pertinent to remember that the engram (memory set formed from our learning and experiences throughout life) is unique, each person has one [Squire and Kandel 2003, Corrêa 2010].

The imagination motor is defined as a dynamic process in which the subject accesses the motor plan of action and actively monitors its deployment. Studies have demonstrated the existence of marked parallelism between imagination and execution of movement, and the temporal characteristics of the mental simulation of a movement, extremely similar to its execution. So, that the time spent in the imagination of a walk to a certain place is similar to spending in carrying out such action. Imagination motor also modulates the autonomic nervous system. During the simulation of motion is observed, for example, increased heart rate proportional to the exercise load imagined, even if there is no evidence of metabolic or electromyographic muscle activity [Lent 2008].

2.2. BCI Systems

BCI systems are tools that can help users to communicate and realize daily activities, although they have limited success and they are still mainly in research environments. The primary users of BCI system are individuals with mild to severe physical disabilities. These systems have also been developed for users with cognitive, such as autism, and even for people who do not have any disabilities, primarily as games industry.

The goal of a BCI system is to allow the user to interact with a device. This interaction is enabled through a variety of functional intermediates, control signals and feedback loops, as detailed in Figure 2. Functional intermediaries components play specific roles in converting intentions into action. By definition, this means that the user and the device are also integral parts of a BCI system. The interaction is also possible through feedback loops that serve to inform each component of the system, the status of one or more components.

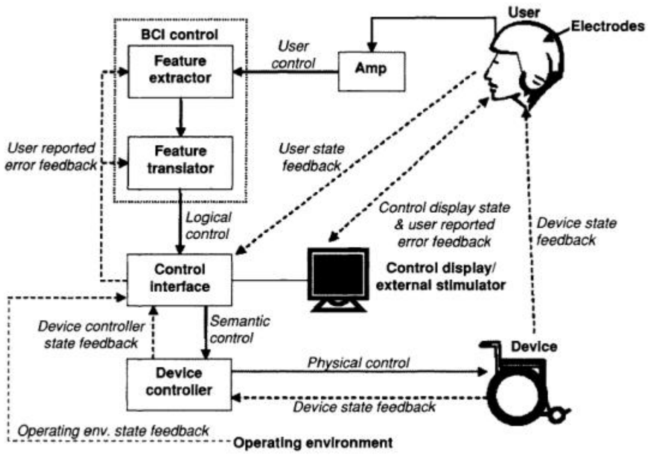

Figure 2. Intermediaty componentes and feedback loops in a BCI System [Vallabhaneni et al. 2005].

The development of a BCI system involves four basic steps. The first is the collection of brain signals, then there is the preprocessing of these signals. After, the relevant characteristics are extracted and the last step is the classification of signals.

The signal acquisition can be performed in brain by an invasive way, which electrodes are inserted into the user's scalp; or noninvasive, with electrodes placed on the scalp of the user. Invasive methods are not commonly used in humans because of the risks involved, but these methods can provide higher resolution signals collected and they need less training time user. The non-invasive techniques, such as electroencephalography (EEG), electromyography (EMG), computed tomography (CT), among others, can be used to collect signals.

However, the EEG is the most widely used method in the acquisition signal to BCI systems. EEG has a high temporal resolution and it is able to measure the activity every millisecond. Modern devices of EEG also have a spatial resolution signals with reasonable (up to 256 electrodes simultaneously).

The most of EEG-based BCI systems use an array of electrodes on the scalp suggested by the international 10-20 system, which contains 64 electrodes. For a better spatial resolution, it is also common to use a variant of the 10-20 system, which fills the spaces between the electrodes with additional electrodes [Malmivuo and Plonsey, 1995 *apud* [Vallabhaneni et al. 2005]].

The effectiveness of a BCI system depends on the user's ability to control his/her brain activity and his/her persistence or goodwill. Unlike motor tasks, the control of brain activity is more difficult to achieve because the user can not identify or evaluate the activity. The user can only understand his/her brain activity through feedback received from the BCI system components [Curran and Stokes, 2003; Kostov and Polak, 2000;

Laubach et al., 2000 *apud* [Vallabhaneni et al. 2005]].

Therefore, the purpose of training is to get users to voluntarily produce EEG detectable signals and it can be altered to achieve a specific result. While the user is not aware of how and when the signals are generated, the process of signal generation can only be activated by voluntary actions of the user. However, these signals can be produced voluntarily through conscious mental activity, such as adding numbers, or as an automatic response to a situation that requires little conscious effort, as riding a bike.

The users are able to generate signals detectable voluntarily to broadcast their intentions. Methods of signal acquisition, however, capture noise generated by activities unrelated to the user's intent, and that noise can come from inside or outside the brain. Appropriate characteristics must be extracted to maximizing the signal to noise ratio.

The purpose of preprocessing and extraction of features (or attributes) is to describe an item through its attributes, which should be very similar for the same category, but very different categories of items. This characterization is done by choosing the most relevant characteristics within the numerous options available. This selection process is necessary because not related characteristics can make the translation algorithms have poor generalization, increase the complexity of calculations and require more training samples to achieve a given level of accuracy.

Translation techniques are algorithms developed to convert the input characteristics (independent variable) in commands to control device (dependent variable) [Wolpaw et al. 2002]. Some translation techniques which are widely used in other areas of signal processing are adapted to systems BCI.

There are types of algorithms of translation characteristics. Some use simple features, such as amplitude or frequency. Others use individual characteristics, while advanced algorithms use a combination of spatial and temporal characteristics produced by one or more physiological processes. Algorithms currently in use include, but are not limited to: linear classifiers, Fisher discriminant and CSSD (Common Spatial Subspace Decomposition), classifiers based on Mahalanobis distance, neural networks (NN - Neural Networks), support vector machines (SVM - Support Vector Machines) and hidden Markov models (HMM - Hidden Markov Models) [Vallabhaneni et al. 2005].

2.3. Decision Trees

Machine learning is an area of Artificial Intelligence that deals with computational techniques on learning and building systems with the ability to acquire knowledge automatically [Quinlan 1988].

There are three cases in which machine learning can be divided: reinforcement, unsupervised and supervised learning[Russel and Norvig 2004].

The agent reinforcement learning to learn from the rewards they receive, and not being informed by an instructor which attitude to take. In reinforcement learning is included in the sub-problem of learning how the environment works.

In unsupervised learning, the instances input not have a rating. The agent does not learn what to do, it does not have the information of which is the desired state or what the correct action to be taken [Russel and Norvig 2004]. In this method the data are rearranged so that they form classes or groups of patterns.

In supervised learning, there is a learning algorithm, or inducer, a set of training examples for which the label of the associated class is known. This method involves a function learning its inputs and outputs. The purpose of the inductor is, as of functions or induced rules, generated with data already labeled, classify new instances that do not yet have that label. Induction of decision trees is a supervised learning method, since the examples should be sorted before being presented to the inductor. But if such examples are scarce, inconsistent or poorly sorted the result may not be satisfactory.

Algorithms that induce decision trees use divide and "conquer" approach, where each node in the tree represents a test of some attribute of the instance being analyzed and each branch from that node corresponds to a possible value for that attribute. This process of testing and branching continues until the leaves, where the label of the instance is determined.

The choice of the best attribute is given in order to use one that has better ability to classify isolated examples. There are different types of selection criteria, but the most algorithms decision tree induction works with univariate division functions, that is, each internal node of the tree is divided according to a single attribute. In this case, the algorithm tries to find the best feature to accomplish this division.

The validation of decision trees is made in the form of division of the set of samples so that this part is used for training another part of the trees and to test the ability of the tree obtained by classifying unknown samples by it. Among the many existing validation techniques. In this work, we use Cross Validation. The cross validation method consists in dividing the number of samples in *n* equal parts using *n-1* parties to train the tree and one for testing part. This process is repeated *n* times, leaving each round, a different piece for testing. After the *n* repetitions will be made the estimated error.

2.4. Related Works

Koprinska (2010) presents an evaluation of five methods for features extraction and ten classification algorithms. The data used are the imagination of movements of the right and left hands, tongue or feet. The best classifiers with correct answers above than 80% were SVM, AdaBoost, Bagg, RBF and RF. Three of these classifiers (AdaBoost, RBF RF) using the J48 algorithm to generate decision trees [Koprinska 2010].

Aparna *et.al.* (2010) presents a system that allows the classification of mental tasks based on statistical data obtained in the frequency domain using discrete cosine transform and extracting useful features from the frequency and the application of decision tree algorithms for classification. The data used are from the imagination of movements of fingers or tongue. The maximum value, minimum value and the average calculated for each of the channels, for all times, were recorded as features. The data were used as input to the decision tree classifier. This study concluded that the decision trees are able to rank reasonably well in the proposed method and CART (81.3 %) performed better than QUEST (78.4 % accuracy) in preprocessing method proposed [Aparna et al. 2010].

3. Methodology

The methodology adopted in this work has five steps, as shown in Figure 3.

Figure 3. Steps to methodology.

3.1. Data Acquisition

The brain signals analysed were collected from 3 subjects, 13 years-old [2]. Data collection was performed during 4 consecutive days, using the equipment *Emotiv EPOC*. The subjects were oriented to open and close the hand holding the ball for a minute, first with the left hand and then with the right hand. They were oriented to imagine the same task. Between each task (do and imagine) there was a pause of approximately 30 seconds. We used data from the imagination task.

The Emotiv EPOC is a device that allows monitoring of 14 channels (beyond the landmarks CMS/DRL (*Common Mode Sense/Driven Right Leg*), terminals P3/P4 and channels of signal quality), the acquisition of neural signals and processing through a wireless interface.

3.2. Preprocessing and Feature Extraction

The preprocessing of the signals was done in EEGLAB, which were removed from the channels that were not of interest, being saved only the 14 signs of the electrodes.

EEGLAB is a suite of tools used in Matlab for data processing of EEG, MEG and other electrophysiological data and event-related continuous incorporating independent component analysis (ICA), analysis of time/frequency, noise rejection, statistics related to events and several useful ways of viewing data [Delorme and Makeig 2004].

For each signal feature extraction was filtered in frequency bands*delta*(0 a 4 Hz), *teta* (4.01 a 8 Hz),*alfa* (8.01-13 Hz) and *beta* (13.01 a 30 Hz), using EEGLAB.

Files of imagination of left and right hands movements were merged, which they were selected 8,000 instances of each, generating files with 16,000 instances. Moreover, a label is inserted to identify whether that instance was imagination of movement of the right hand (ImaD) or left hand (imaE).

3.3. Classification

We have used WEKA software, configured to J48 algorithm. DT were generated with at least 1000 instances per sheet for data sets without filtering of the subject 1. We have chosen this data set because the success rates of unfiltered data were close to success rates of trees set of filtered data with greater accuracy filter (0-4 Hz), which got hit around 90%.

The J48 algorithm uses the concept of information gain criterion branching. The gain information is based on entropy index to measure the uniformity of each node [Dalagassa et al. 2008].

WEKA is a collection of machine learning algorithms used in data mining and implemented in Java. It has tools for preprocessing, classification, regression, clustering, association rules and data visualization [Hall et al. 2009].

[2]In literature, the number of subjects is general small, because the amount of generated instances of each subject is big, and the treatment and analysis is complex. We have collected just three days, but we intend, as future works, collect more days.

3.4. Validation

The results obtained from the previous step were validated based on neuroscience and evaluating hit rates in the classification. Based on the literature, we have formulated the Table I, where the brain regions, the electrodes position (according to international standard “ 10-20”) and the brain features are presented.

4. Results

Here, we present the data of one of the subject, by space limitation. This subject has higher rates (Table I) above 90 % for decision trees with at least two instances per leaf, and the third day had best values. The Table I presents the higher rates of the subject in the four days of data collection with values for trees with at least 2, 200, 500 and 1000 instances per leaf. The values present in Table I are promising, because a classification tax higher than 70% to DT is considered very good. The graphical DT will be analysed with 1000 instances per leaf, being leaner and easy viewing.

Table 1. Hit rates of the subject for the first four days gathering

Instances\Day	day 1	day 2	day 3	day 4
2	96.0563%	94.7438%	98.7375%	91.1125%
200	86.8688%	83.5813%	93.9875%	81.4313%
500	82.0813%	80.625%	89.2875%	77.5313%
1000	**78.8125%**	**75.5438%**	**85.825%**	**73.2313%**

The first tree (Figure 4), referring to the first day, presented six possible paths for classification tasks (imagination of the right and left hands). The path 1 indicates that the subject performed a preview of the activity (FC6). In the second path, there was a high concentration in planning movements (AF3), and viewing images. The same happens with the path 3, in which the subject, in addition to the activities above, performs a preview image during a dialogue (O2). It is thought that the subject has mentally verbalized his/her action. The path 4 can be considered the most complete, because it involves the electrode P8, where spatial representations are formed, tactile perceptions and movement of the hands, which is enabled if the physical realization of the movement. On the path 5, the subject uses viewing and linguistic processing, possibly he/she verbalized an action (T7 and FC5). On the path 6, in addition to certain actions on the path 5, there was a high concentration and planning of movements (AF3).

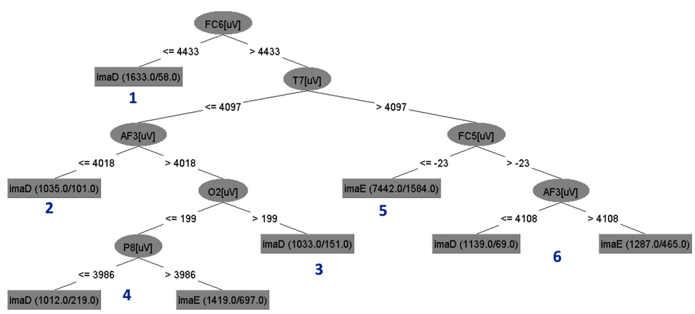

Figure 4. DT on first day.

The second tree (Figure 5) presents the second day collection. It presented five possible paths for classification tasks. The path 1 presents a verbalization of linguistic processing (F3), followed by spatial representations, tactile perceptions and hand movement (P8). The path 2 involves, beyond the path 1 (F3 and P8), again verbalization of linguistic processing (FC5). The path 3 involves, besides the verbalization of linguistic processing, a high concentration in the planning of movements (AF4). The path 4 involves the path 3 and a new verbalization of linguistic processing (F7). And the path 5 may be considered as complete as it involves, in addition to verbalization of linguistic processing (F3, F7), and high concentration of motion planning (AF4) and the representation of movement (P7).

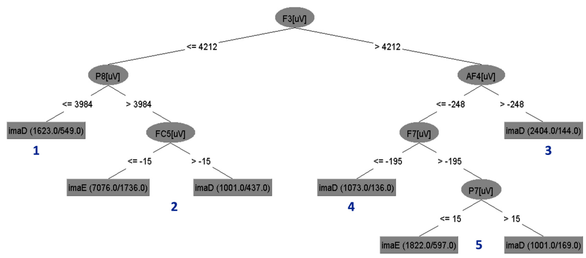

Figure 5. DT on second day.

The third tree (Figure6) presents the third day collection. It has only two paths. For classification of the tasks on the path 1, the subject presented a high concentration and planning of movements (AF4). The path 2 involves, in addition to high concentration, viewing images. It is thought that the subject has looked at his/her hand or mentally visualized the hand moving (O2).

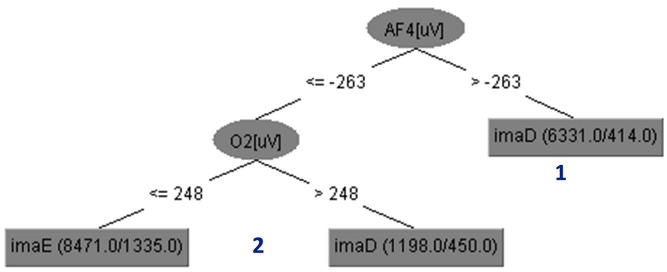

Figure 6. DT on third day.

The fourth tree (Figure 7) presents the fourth day collection. It presented 7 paths. This tree is incremental, i.e., paths are complementary (new nodes are inserted), but the previous sequence remains the same. In paths 1, 2 and 3 the subject performed only one linguistic processing to perform the task (T7 and T8). In path 4, beyond the linguistic processing, was performed visual-spatial processing (F8). The path 5 involves the previous activities and high concentration and planning of movements (AF3). The path 6 can be considered the most complete, because besides the previous paths, involves performing imagination of movement (P7). Path 7 backs to the subject hold a high concentration (AF4).

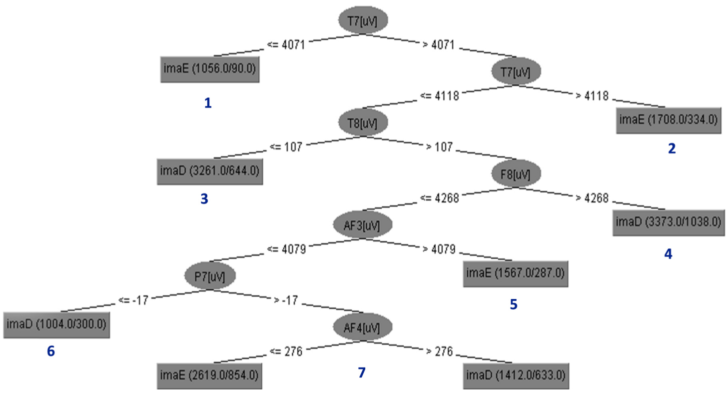

Figure 7. DT on forth day.

5. Conclusions

In this work, we presented four decision trees, one for each collected day. Important to note that the results are totally different in each day, because the signals were collected only for 4 days, without previous training of the subject and without an adaptation period to the tasks (these tasks are not daily life tasks).

According to neuroscience [Gazzaniga and Heatherton 2005], if the subject was "trained" for a long period of time for an action (for example, more than 30 days), it becomes an automatic action, meaning the use of the same brain regions (generating decision trees very similar).

The feature extraction step, much emphasized in some studies, no significant improvement in the classification of signals. Frequency bands used, only one had higher rates than the rates without filter. This band, *delta* (0-4 Hz), is related to deep sleep, with little relevance in the BCI systems area.

The contribution of this work is the suggestion of an automated way to signal processing, minimizing the feature extraction and focusing on classification.

As future work we intend to analyse the performance of data movement, which has been collected, comparing it with imagination motion, seeking similarities in brain areas used.

We intend to collect the brain signals for a longer period of time (30 days, for example), where there may be a training or learning by individuals resulting in process automation. In this case, there is a possibility of the result (generated decision trees) to be similar in the last days of each subject, where they have established a constant for the imagination of a movement. And it can also be found similarities between the decision trees of different subjects, since long time spent performing the same actions in the same place.

References

Alpaydin, E. (2010). *Introduction to Machine Learning*. The MIT Press, 2nd edition. 537 p.

Aparna, C., Murthy, J. V. R., Babu, B. R., Chandra, M. V. P., and Rao, S. (2010). Reducing dataset size in frequency domain for brain computer interface motor imagery classification. *International Journal on Computer Science & Engineering*, 02(09):2924 – 2927.

Corrêa, A. C. O. (2010). *Memória, aprendizagem e esquecimento: a memória através das ciências cognitivas*. Editora Atheneu, São Paulo.

Dalagassa, M., S.H, S., and Carvalho, R. (2008). Avaliação de modelos para a classificação de beneficiários com indicativos para o diabetes mellitus tipo 2. In *Anais do XI Congresso Brasileiro de Informática em Saúde*. XI Congresso Brasileiro de Informática em Saúde.

Delorme, A. and Makeig, S. (2004). Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134:9–21.

Duda, R., Hart, P., and Stork, D. (2002). *Pattern Classification*, chapter Chapter 8. Willey Interscience, 2nd edition.

Farwell, L. A. and Donchin, E. (1988). Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and Clinical Neurophysiology*, 70:510–523.

Gazzaniga, M. S. and Heatherton, T. (2005). *Ciência psicológica: mente, cérebro e comportamento*. Artmed, Porto Alegre.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I. H. (2009). The weka data mining software: An update. *SIGKDD Explorations*, 11(1).

Koprinska, I. (2010). Feature selection for brain-computer interfaces. In *Proceedings of the 13th Pacific-Asia international conference on Knowledge discovery and data mining: new frontiers in applied data mining*, PAKDD'09, pages 106–117, Berlin, Heidelberg. Springer-Verlag.

Lent, R. (2001). *Cem bilhões de neurônios: conceitos fundamentais de neurociência*. Atheneu, São Paulo.

Lent, R. (2008). *Neurociência da mente e do comportamento*. Guanabara Koogan, Rio de Janeiro.

Mitchell, T. (1997). *Machine Learning*. WCB McGraw-Hill.

Pfurtscheller, G. and da Silva, F. L. (1999). Event-related eeg/meg synchronization and desynchronization: basic principles. *Clinical Neurophysiology*, 110(11):1842 – 1857.

Quinlan, J. R. (1988). An empirical comparison of genetic and decision-tree classifiers. In *Proceedings of the Fifth International Conference on Machine Learning*, pages 135–141, Ann Arbor, MI: Morgan Kaufmann.

Ratey, J. J. (2002). *O cérebro: um guia para o usuário*. Objetiva, Rio de Janeiro.

Russel, S. and Norvig, P. (2004). *Inteligência Artificial*. Elsevier/Campus, Rio de Janeiro.

Squire, L. R. and Kandel, E. R. (2003). *Memória: da mente às moléculas*. Artmed, Porto Alegre.

Vallabhaneni, A., Wang, T., and He, B. (2005). Brain-computer interface. In He, B., editor, *Neural Engineering*, pages 85–121. Kluwer Academic, New York.

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., and Vaughan, T. M. (2002). Brain-computer interfaces for communication and control. *Clinical neurophysiology*, 113(6):767 – 791.