Support Vector Machines in the Diagnosis of Alzheimer's Disease

Lucas R. Trambaiolli #1, Ana C. Lorena #2, Francisco J. Fraga *3, Renato Anghinah %4

Centro de Matemática Computação e Cognição, Universidade Federal do ABC
Rua Santa Adélia, 166 CEP: 09.210-170 Santo André SP Brasil
[1] lucasrtb@yahoo.com.br
[2] aclorena@gmail.com

* Centro de Engenharia e Ciências Sociais, Universidade Federal do ABC
Rua Santa Adélia, 166 CEP: 09.210-170 Santo André SP Brasil
[3] franciscojfraga@gmail.com

% Centro de Referência em Distúrbios Cognitivos, Hospital das Clnicas,
Faculdade de Medicina da Universidade de São Paulo
Rua Arruda Alvim, 206 - CEP: 05.410-020 - São Paulo - SP - Brasil
[4] anghinah@terra.com.br

***Abstract*—The diagnosis of Alzheimer's disease involves neurological examinations, among which one can mention the electroencephalogram (EEG). The waveforms collected by the EEG scalp electrodes are then analyzed by a specialist, to identify patterns that indicate the presence of dementia. Some studies show the use of Machine Learning (ML) techniques in extracting these patterns, providing therefore an aid to the specialist in this type of analysis. In this study, Support Vector Machines, a ML method based on Statistical Learning Theory, were used to induce a model that, from EEG waves get an automatic diagnosis of Alzheimer's disease, becoming a tool that supports the diagnosis made by neurologists. The performance obtained from analysis of EEG epochs was 79.9% accuracy and 83.2% sensitivity, whereas the analysis considering the diagnosis of each patient was obtained 87.0% accuracy and 91.7% sensitivity.**

***Index Terms*—Alzheimer's disease, Support Vector Machines, electroencephalogram, coherence.**

I. Introduction

Alzheimer's disease (AD) is considered the main cause of dementia in Western countries [1], manifested by memory loss and impairment of at least another area of cognition (calculation, praxis, gnosis, executive functions, language, etc.). As the definitive diagnosis of AD can only be established with a histo-pathological analysis of the brain (autopsy or biopsy) [2], the challenge to search for a biological marker to determine an early diagnosis of the disease remains open. Currently, the correct diagnosis by neuropsychological evaluations ranges from 85 to 93% accuracy in university hospitals. But not always these cognitive screening batteries are easy to apply, requiring experienced people and often lengthy sessions for its realization [3].

One of the tests used as an alternative for the diagnosis of dementia, when it is still uncertain, is the electroencephalogram [4], which is the record of different brain rhythms manifested by electric potentials from synaptic activity between neurons [5]. The long-term electroencephalogram has been of great importance for the diagnosis of dementia and encephalopathy, mainly in cases where this diagnosis cannot be made based only on the initial clinical evaluations [4].

These tests generate a large amount of data, which must then be examined by a specialist. Many studies can be found using ML techniques in the analysis of EEG data to differentiate AD patients from normal patients. The data sets in these cases consist of EEG records of several patients, previously labeled according to their diagnosis [6], [7], [8], [3], [9], [10], [11], [12], [13].

The objective of this study is to test different combinations of characteristics extracted from the EEG signals as input for a ML technique [14]. We used a ML technique known as Support Vector Machines (SVM) [15] in the analysis of EEG data derived from normal individuals and patients with AD. This technique is known by its good generalization ability and robustness to high dimensional data, like those used in this paper. The goal was to develop models to aid in the extraction and classification of patterns found in these records and thus assist in an initial diagnosis of AD.

II. Materials and Methods

A. Data Acquisition

The data set used in this study was extracted from a clinical database, composed of EEG examinations of individuals with AD and without the disease (control). Two samples of volunteers were selected: the first (S1) corresponds to a population of 19 healthy subjects, 14 females and 5 males, with an average age of 71.63 years. The second sample (S2) corresponds to a population of 16 individuals diagnosed with probable AD, 14 females and 2 males, with an average age of 73.44 years.

The diagnosis of AD was made according to NINCDS-ADRDA criteria [16], classified as mild to moderate, according to the DSM III-R [17]. Subjects had no history of diabetes mellitus, kidney diseases, thyroid diseases, alcoholism, liver disease, lung disease or lack of vitamin B12 to prevent the occurrence of other causes for diagnosing cognitive impairment.

The EEG records were obtained by an equipment brand EMSA, with 32 channels, 12 bits A/D converter and sampling rate of 200 Hz. Placement of scalp electrodes (referential montage) followed the international 10-20 system illustrated in Figure 1. The biauricular referential electrodes (A1 and A2) were attached, as recommended by the Brazilian Society of Clinical Neurophysiology and the American EEG Society. During the examination, the records were obtained with subjects awake and resting with eyes closed.

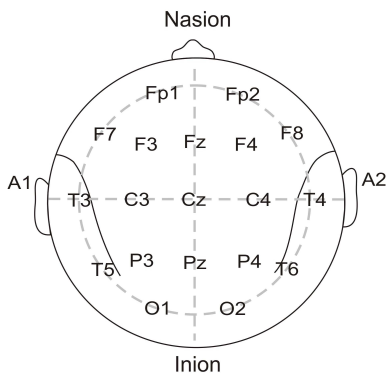

Fig. 1. The international 10-20 system (the "10" and "20" refer to the 10% or 20% interelectrode distance).

B. Preprocessing

The main objective of the preprocessing of EEG signals is making more obvious events that can differentiate the two classes (Alzheimer's versus healthy), thus enabling a better posterior classification of data using ML techniques. First, the tests were evaluated by an experienced physician in order to eliminate portions of artifacts (electrical interference, which may stem from the movement of the wires, bad contact, eye movement, etc.). The segments of EEG signals free of artifacts were selected from all 20 electrodes, each divided at intervals of 8 seconds, known in the literature as epochs [18].

The digitized signals containing selected EEG epochs were subjected to a low pass digital filter with a cutoff frequency at 50Hz. It was an infinite impulse response elliptic filter with a zero in the frequency of 60Hz, thus eliminating completely the interference of the power grid. The frequency analysis was performed using a 512-point Fast Fourier Transform (FFT), applied after sub-segmenting the signal (epochs) with Hamming windows of 2.5 s and 90% overlap between successive windows [18].

The frequency bands were divided into sub-bands $\delta1$ (0,1-2,0Hz), $\delta2$ (2,5-4,0Hz), $\theta1$ (4,5-6,0Hz), $\theta2$ (6,5-7,5Hz), $\alpha1$ (8,0-10,0Hz), $\alpha2$ (10,5-12,0Hz), $\beta1$ (12,5-12Hz), $\beta2$ (15,5-21,0Hz), $\beta3$ (>21,0Hz) [19]. A parameter of analysis widely used in EEG studies and also used in this work is the EEG coherence. The coherence between a pair of EEG channels is obtained dividing the estimated cross spectrum of two channels by the auto-spectra of each channel [20] according to equation 1.

$$Coh_{ij}^2 = \frac{E||C_{ij}(\omega)|^2|}{E|C_{ii}(\omega)|E|C_{jj}(\omega)|} \tag{1}$$

Where $C_{ij}(\omega)$ is the cross spectral density and $C_{ii}(\omega)$ and $C_{jj}(\omega)$ are the power spectral densities of signals i and j (EEG channels). The average spectral windows of each epoch were calculated using the periodogram method of Welch [21].

The combination of electrodes used in the calculation of coherence was between:

- Inter-hemispheric counterpart electrodes [19];
- Intra-hemispheric frontal electrodes [22];
- Intra-hemispheric rear electrodes [22];
- Equidistant electrodes [22];
- Rear electrodes in bipolar montage [23].

The electrodes pertaining to each combination are listed in Table I, where the coherence operation is represented by a dot character $(.)$ and the bipolar montage is represented by a dash character $(-)$ as usual.

Another innovative attribute under consideration in this work was the peak of the spectrum coming from bipolar montages. A bipolar signal can be obtained from a referential montage simply subtracting two signals for elimination of common biauricular reference [24]. The electrodes used in the calculation of the bipolar spectra peaks were F3-F4, F7-F8, C3-C4, T3-T4, P3-P4, T5-T6 O1-O2.

C. Support Vector Machines

SVMs constitute a ML technique based on concepts from the Statistical Learning Theory [25]. They separate data by a hyperplane, considering bounds in the generalization ability of a linear classifier [26]. Accordingly, given a training data set *T* containing *n* pairs $(\mathbf{x}_i, y_i)$, where $\mathbf{x}_i \in \Re^m$ is a data point with m dimensions and $y_i \in \{-1, +1\}$ is the class of $\mathbf{x}_i$, SVMs seek the linear classifier $g(\mathbf{x}) = sgn(\mathbf{w} \cdot \mathbf{x} + b)$ separating data from classes $+1$ and -1 with minimum error while also maximizing the margin of separation between the classes (Figure 2) [15]. For instance, in the application considered in this paper each data point represents the EEG of a patient (in fact, the characteristics extracted from his/her EEG according to the pre-processing techniques previously described), while the classes represent the presence or absence of the Alzheimer disease.

The maximization of the margin is accomplished through the minimization of the norm $\|\mathbf{w}\|$. The following optimization problem is solved in this process:

$$\text{Minimize: } \|\mathbf{w}\|^2 + C\sum_{i=1}^{n} \xi_i$$
$$\text{Under the restrictions:}$$
$$y_i\left(\mathbf{w} \cdot \mathbf{\Phi}(\mathbf{x_i}) + b\right) \geq 1 - \xi_i \text{ and } \xi_i \geq 0, \text{ for } i = 1, \ldots, n$$

Where C is a constant that imposes a different weight for training error over the generalization of the classifier and ξ_i are slack variables. The restrictions are originally imposed to ensure that no training data should be within the margins of

TABLE I
COHERENCE OPERATIONS PERFORMED IN EACH COMBINATION OF ELECTRODES

Combination	**Electrodes and coherence operations**
Inter-hemispheric counterpart electrodes	Fp1.Fp2, F7.F8, F3.F4, C3.C4, P3.P4, T5.T6 and O1.O2
Intra-hemispheric frontal electrodes	Fp1.F7, Fp2.F8, Fp1.F3, Fp2.F4, Fp1.C3, Fp2.C4, F7.C3, F8.C4, F3.C3 and F4.C4
Intra-hemispheric rear electrodes	O1.P3, O2.P4, O1.T5, O2.T6, O1.C3, O2.C4, P3.C3, P4.C4, T5.C3 and T6.C4
Equidistant electrodes	O1.Fp1, O2.Fp2, O1.F7, O2.F8, O1.F3, O2.F4, P3.Fp1, P4.Fp2, P3.F7, P4.F8, P3.F3, P4.F4, T5.Fp1, T6.Fp2, T5.F7, T6.F8, T5.F3e T6.F4
Rear electrodes in bipolar montage	T3-C3.T4-C4, C3-P3.C4-P4, T5-P3.T6-P4, T3-T5.T4-T6, P3-O1.P4-02 and T5-O1.T6-02

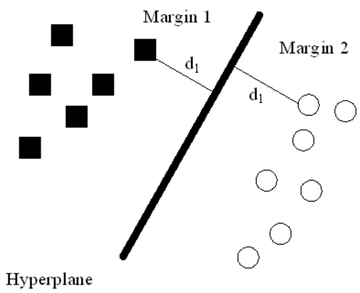

Fig. 2. Example of a linear frontier induced by a SVM.

separation between the classes. The slack variables relax these restrictions on the margins, in order to avoid an overfitting to training data and also for dealing with noisy data. The number of training errors and data between the margins is controlled by the minimization of the term $\sum_{i=1}^{n} \xi_i$ [15].

Nevertheless, the classifier obtained is still limited, since there are many data sets where data can not be satisfactorily divided by a hyperplane, making a non-linear frontier more adequate to the problem [15]. The Cover theorem states that, if it is possible to increase the dimensions of data through a non-linear mapping function $\mathbf{\Phi}$, with a high probability they will become linearly separable in the new space, which is usually called characteristics space [27].

The mathematical toll employed for the computation of $\mathbf{\Phi}$ is named *Kernel*. The Kernel K is a function which takes two variables $\mathbf{x}_i$ and $\mathbf{x}_j$, representing two data points, and calculates the dot product between them in the characteristics space. Since all computations involving data points in SVMs are in the form of dot products, the non-linearization of SVMs can be easily accomplished through the use of a proper Kernel function. In this paper the non-linear RBF (Radial-Basis Function) Kernel function was used:

$$K(\mathbf{x}_i, \mathbf{x}_j) = e^{-(\gamma < \mathbf{x}_i - \mathbf{x}_j, \mathbf{x}_i - \mathbf{x}_j >^2)} \quad (2)$$

The classifiers Ripper [28], Decision Tree [29] and SVM with linear Kernel were also tested in this study, but SVMs with the RBF kernel achieved a best overall performance. This may be attributed to the large number of input features in the data sets investigated.

III. RESULTS

All SVMs in this paper were induced using the Weka tool [30] with default parameters values, in order to allow a fair comparison between the different combinations of data considered. The parameters employed were $C = 1.0$ and $\gamma = 0.01$ in the RBF Kernel.

The total data set with 3305 epochs extracted from the EEG exams was divided into two parts: one for training the classification models (approximately 68.06% of the data) and one for testing the classification models induced (approximately 31.94% of the data). The division was made so that different patients were present in the training and testing sets - and that a patient used for testing was not used for training, and vice-versa. The experiments were repeated three times, varying the data sets for training and testing, but maintaining the same proportions and configurations previously discussed. This procedure was adopted in order to better assess the generalization ability of the classifiers, for different partitions of data.

Table II presents the mean and standard deviation (considering the different partitions of data) of the following performance measures obtained by SVMs in the classification of epochs of the EEG exams: accuracy rates, AUC (Area Under the ROC Curve), Sensitivity (correct classifications for epochs from patients with AD) and Specificity (correct classifications for epochs from normal patients). Different combinations of characteristics extracted from the EEG exams were considered as input for the SVMs: frontal coherences, rear coherences, long distances coherences, homologous coherences, rear-bipolar coherences, bipolar peaks and all coherences combined to the bipolar peaks. Therefore, each line of the table represents a different configuration of the data sets (different electrodes, coherences, or combinations of them). The best results obtained for each performance measure in the table are highlighted in boldface, while the worst results are highlighted in italics.

The best performing classifiers in Table II were chosen for a further analysis of the performance in the diagnosis of the test individuals. This analysis allows to verify the accuracy rate obtained by patient, rather than the performance by epochs previously considered. For such analysis, for each patient the ratio between the number of correctly classified epochs divided

TABLE II
PERFORMANCE BY EPOCHS OF EEG EXAMS

Data	**Accuracy(%)**	**AUC**	**Sensitivity(%)**	**Specificity(%)**
Frontal coherences	65.2 ± 3.5	0.66 ± 0.01	69.1 ± 9.4	63.2 ± 12.0
Rear coherences	67.4 ± 3.4	0.68 ± 0.03	64.6 ± 3.5	70.4 ± 12.0
Long distances coherences	54.8 ± 2.4	0.55 ± 0.03	47.8 ± 8.5	$\mathit{62.3 \pm 12.5}$
Homologous coherences	58.4 ± 9.2	0.58 ± 0.09	46.9 ± 21.5	69.1 ± 6.1
Rear-bipolar coherences	$\mathit{53.4 \pm 8.0}$	$\mathit{0.52 \pm 0.05}$	$\mathit{21.9 \pm 16.8}$	$\mathbf{81.7 \pm 12.7}$
Bipolar peaks	76.4 ± 3.6	0.76 ± 0.04	71.0 ± 0.1	80.5 ± 3.6
All coherences + bipolar peaks	$\mathbf{79.9 \pm 3.9}$	$\mathbf{0.80 \pm 0.04}$	$\mathbf{83.2 \pm 3.6}$	76.4 ± 8.5

by his/her total number of epochs was calculated. Considering the threshold of $50\%+1$ for diagnosis, the accuracy, sensitivity and specificity performance measures could then be calculated. Table III presents these results.

IV. DISCUSSION

First it is important to describe the meaning of the performance measures to be evaluated. Accuracy is the rate correct classifications for the whole data set. Sensitivity is the rate of diagnoses with a positive result, given that the individual has the disease. Specificity is the rate of diagnoses with a negative result, given that the individual does not have the disease. And the area under the ROC curve (AUC) can be interpreted as the probability that, given two individuals (one with the disease and other without the disease), the prediction made will be largest for the individual with the disease.

From Table II, it is possible to verify that, in the data sets containing coherences, the frontal and rear configurations had a better accuracy rate when compared to the other coherences, as also observed by Locatelli *et al.* (1998) [22]. On the other hand, the rear-bipolar coherences, despite showing a low accuracy, had a good specificity rate when compared to the other coherences, as also verified in Trambaiolli *et al.* (2009) [23]. The set of bipolar peaks presents better results than those obtained for the coherences. This may be an indication that reference electrodes interferes in the EEG signal.

By considering all the previous attributes together as input for the SVM classifier, an improvement of accuracy, AUC and sensitivity is obtained, standing out their low standard deviation values.

Nevertheless, all these results refer to classifications by epochs of exams. For medical interpretation, the epoch-based classification is not exactly helpful, since it does not give a definitive diagnosis for a given patient, but instead it diagnoses excerpts of his/her EEG exam. Therefore, an analysis of the classification per patient is necessary, considering the total set of epochs contained in his/her exam. Since all entries involving the spectrum's peaks alone and the combination of these peaks with all coherences achieved the best performance in the classification of the epochs, an analysis per patient was carried out for these particular data sets.

Both sets of attributes had an accuracy improvement in the per patient evaluation scenario. For both data sets the accuracy, sensitivity and specificity rates were all above the level of 80%, and the sensitivity for the bipolar peaks data exceeds 90%. The phenomenon of slowing of EEG signals in patients with AD, which occurs mainly in the alpha and beta bands as noted by several authors [31], [32], [33], become more evident when the difference between more localized electrical potentials is considered, i. e., between electrodes next to each other, as performed by bipolar recording. Despite the relatively high standard deviation values in sensitivity and specificity of tests with the peaks, a sensitivity of 91.7% is extremely important, since it can help reducing the risk of AD patients stop treatment by the lack of a correct diagnosis.

As can be seen in Table II, tests for the rear-bipolar coherence favor normal class in detriment of class AD, which reflects on the great difference between the values of specificity and sensitivity obtained for this data set. In the case of all coherences + bipolar peaks, there is a balance between sensitivity and specificity, suggesting that no class is strongly favored in detriment of the other.

It is also noteworthy that the standard deviation values obtained for the analysis per patient are quite high. Nevertheless, this can be attributed to the fact that we have relatively few patients for testing (13 ± 2), which causes an elevation in the standard deviation values for sensitivity and specificity. For instance, when you consider that each class corresponds to about half of the patients, the misclassification of a patient can result in differences of performance of approximately 15%.

V. CONCLUSION

Although more tests are needed, involving, for example, a larger number of subjects, this study demonstrated that: the combination of frontal coherences, rear coherences and long distances coherences, as proposed by Locatelli *et al.* (1998) [22]; the homologous inter-hemispheric coherences proposed by Anghinah (2003) [19]; the rear-bipolar coherences proposed by Trambaiolli *et al.* (2009) [23]; and the peaks of spectrum, investigated in this paper, are good input attributes for SVM classifiers in the classification of EEG epochs.

Nevertheless, the main result presented in this paper is that the classification per patient was best when using as inputs for SVMs only the frequencies of the peaks of spectrum obtained by bipolar recording. Therefore, we can conclude that the use of bipolar peaks are a good tool for providing a set of characteristics from EEG signals for SVMs in the classification of patients with AD. Given the simplicity of this pre-processing, allied to the high sensitivity obtained in the

TABLE III
PERFORMANCE BY PATIENTS

Data	Accuracy(%)	Sensitivity(%)	Specificity(%)
Bipolar peaks	$\mathbf{87.0 \pm 3.8}$	$\mathbf{91.7 \pm 14.4}$	$\mathbf{84.9 \pm 14.4}$
All coherences + bipolar peaks	81.2 ± 5.9	82.5 ± 4.8	80.2 ± 7.7

classification experiments, this can be considered a promising result.

Future work shall consider tuning the parameter values of the SVM classifiers, since this procedure can lead to a further increase in the classification performances achieved.

ACKNOWLEDGMENTS

To “Universidade Federal do ABC” (UFABC) and “Conselho Nacional de Desenvolvimento Científico e Tecnológico” (CNPq) for the financial support provided.

REFERENCES

[1] T. D. Bird, “Alzheimers disease and other primary dementias,” in *Harrisons principles of internal medicine*, E. Braunwald, A. S. Fauci, D. L. Kasper, S. L. Hauser, D. L. Longo, and J. L. Jameson, Eds. McGraw-Hill, 2001, pp. 2391–2399.

[2] R. D. Terry, “Neuropathological changes in alzheimer disease,” *Prog. .Brain Res.*, vol. 101, pp. 383–390, 1994.

[3] D. Parikh, N. Stepenosky, A. Topalis, D. Green, J. Kounios, C. Clark, and R. Polikar, “Ensemble based data fusion for early diagnosis of alzheimers disease,” *Proceed. of the 2005 IEEE Eng. in Med. and Biol. 27th Ann. Conf.*, pp. 2479–2482, 2005.

[4] F. L. C. Luccas, N. I. O. Braga, L. C. Fonseca, and M. L. Frochtengarter, “Recomendaes para o registro e interpretao do mapeamento topogrfico do eletroencefalograma (eeg) e potenciais evocados sensoriais (pes) parte i: aspectos gerais [in portuguese],” *Brazilian J. Epilepsy Clin. Neurophysiol.*, vol. 2, pp. 175–182, 1996.

[5] C. M. Epstein, “Digital eeg: choosing the right equipment for your lab,” *30th Ann. Course Clin. EEG and Electrophysiol. of Am. EEG Soc.*, pp. 1–11, 1995.

[6] V. Schetinin, “Polynomial neural networks learn to classify eeg signals,” *NIMIA-SC2001 - 2001 NATO Adv. Study Inst. on Neural Networks for Instrum., Meas., and Related Ind. App.: Study Cases*, pp. 9–20, 2001.

[7] S. Yagneswaran, M. Baker, and A. Petrosian, “Power frequency and wavelet characteristics in differentiating between normal and alzheimer eeg,” *IEEE Eng. in Med. and Biol. 24th Ann. Conf.*, pp. 46–47, 2002.

[8] S. Y. Cho, B. Y. Kim, E. H. Park, J. W. Kim, W. W. Whang, S. K. Han, and H. Y. Kim, “Automatic recognition of alzheimers disease with single channel eeg recording,” *Proceed. of the 25th Ann. Intern. Conf. of the IEEE EMBS*, pp. 2655–2658, 2003.

[9] T. H. Kima, B. Y. Kim, E. H. Park, J. W. Kim, E. W. Hwang, S. K. Han, and S. Cho, “Computerized recognition of alzheimer disease-eeg using genetic algorithms and neural network,” *Future Gen. Comp. Syst.*, vol. 21, pp. 1124–1130, 2005.

[10] N. Stepenosky, A. Topalis, H. Syed, D. Green, J. Kounios, C. Clark, and R. Polikar, “Boosting based classification of event related potentials for early diagnosis of alzheimers disease,” *Proceed. of the 2005 IEEE Eng. in Med. and Biol. 27th Ann. Conf.*, pp. 2494–2497, 2005.

[11] A. Cichocki, S. L. Shishkin, T. Musha, Z. Leonowicz, T. Asada, and T. Kurachi, “Eeg filtering based on blind source separation (bss) for early detection of alzheimers disease,” *Clin. Neurophys.*, vol. 116, pp. 729–737, 2005.

[12] J. M. Abe, H. F. S. Lopes, and R. Anghinah, “Paraconsistent artificial neural networks and alzheimer disease,” *Dem. and Neuropsy.*, vol. 1, pp. 241–247, 2007.

[13] C. Lehmann, T. Koenig, V. Jelic, L. Prichep, R. E. John, L. O. Wahlund, Y. Dodge, and T. Dierks, “Application and comparison of classification of alzheimers disease in electrical brain activity (eeg),” *J. of Neuro. Methods.*, vol. 161, pp. 342–350, 2007.

[14] T. Mitchell, *Machine Learning*. McGraw Hill, 1997.

[15] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines an other kernel-based learning methods*. Cambridge University Press, 2000.

[16] G. McKhann, D. Drachman, M. Folstein, R. Katzman, D. Price, and E. M. Stadlan, “Clinical diagnosis of alzheimer’s disease: report of the nincds-adrda work group under the auspices of department of health and human services task force on alzheimer’s disease,” *Neurology.*, vol. 34(7), pp. 939–944, 1984.

[17] R. L. Spitzer, J. B. W. Williams, M. Gibbon, and M. B. First, *User’s guide for the structured clinical interview for DSM-III-R: SCID*. American Psychiatric Press, 1990.

[18] S. Sanei and J. A. Chambers, *EEG signal processing*. Wiley-Interscience, 2007.

[19] R. Anghinah, “Estudo da densidade espectral e da coerncia doeletroencefalograma em indivduos adultos normais e com doena de alzheimer provvel [in portuguese],” Ph.D. dissertation, Univ. de So Paulo, 2003.

[20] J. Gotman, P. Gloor, and W. F. Ray, “A quantitative comparison of traditional reading of the eeg and interpretation of computer-extracted features in patients with supratentorial brains lesions.” *Electroencephalogr. Clin. Neurophysiol.*, vol. 38, pp. 623–639, 1975.

[21] P. D. Welch, “The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms,” *IEEE Trans. Audio Electroacoust.*, vol. 15(2), pp. 70–73, 1967.

[22] T. Locatelli, M. Cursi, D. Liberati, M. Franceschi, and G. Comi, “Eeg coherence in alzheimers disease,” *Electroencephal. and Clin. Neurophysiol.*, vol. 106, pp. 229–237, 1998.

[23] L. R. Trambaiolli, A. C. Lorena, F. J. Fraga, and R. Anghinah, “Uso de aprendizado de mquina no auxlio ao diagnstico de alzheimer [in portuguese],” *REIC - Revista Eletrnica de Iniciao Cientfica*, 2009.

[24] P. L. Nunez and R. Srinivasan, *Electric fields of the brain: the neurophysics of EEG*. Oxford University Press, USA, 2006.

[25] V. N. Vapnik, *The nature of Statistical learning theory*. Springer-Verlag, 1995.

[26] C. Campbell, “An introduction to kernel methods,” *Studies in Fuzziness and Soft Computing*, vol. 66, pp. 155–192, 2001.

[27] S. Haykin, *Neural Networks A Compreensive Foundation*. Prentice-Hall, 1999.

[28] W. W. Cohen, “Fast effective rule induction,” *Proceedings of the Twelfth International Conference on Machine Learning*, p. 115123, 1995.

[29] J. R. Quinlan, “Induction of decision trees,” *Machine Learning*, vol. 1(1), pp. 81–106, 1986.

[30] I. A. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques*. Elsevier, 2005.

[31] T. Dierks, R. Ihl, L. Frolich, and K. Maurer, “Dementia of the alzheimer type: effects on the spontaneous eeg described by dipole sources,” *Psych. Res.*, vol. 50, pp. 151–162, 1993.

[32] G. G. Yener, A. F. Leuchter, D. Jenden, S. L. Read, J. L. Cummings, and B. L. Miller, “Quantitative eeg in frontotemporal dementia,” *Clinical EEG (electroencephalography)*, vol. 27, no. 2, p. 61, 1996.

[33] D. Cibils, “Dementia and qeeg (alzheimers disease),” *Clin Neurophysiol (Supplements Advances in Clinical Neurophysiology)*, vol. 54, pp. 289–294, 2002.