

Towards the Early Diagnosis of Alzheimer's Disease Through the Application of a Multicriteria Classification Model

Amaury Brasil, Plácido Rogério Pinheiro
and André Luís Vasconcelos Coelho

*University of Fortaleza (UNIFOR), Graduate Program in Applied Informatics,
Fortaleza (CE)
Brazil*

1. Introduction

The Alzheimer's disease (AD) is a progressive and degenerative disease of the brain which causes a serious impairment over its two main activities: thinking and memory. According to Celsis (Celsis, 2000), AD is the most common form of dementia among the elderly population, comprising up to 75% of all dementia cases. AD causes a gradual loss of intellectual abilities with deterioration in cognition, function, and behavior, affecting many aspects of an individual life.

This way, with the decline of the normal functioning over the nervous and other bodily systems, and with the natural behavioral and personality changes, the identification of what constitutes abnormal impairment becomes a hard task. Davidoff (Davidoff, 1986) argues that the problem over the AD diagnosis is not only related to the current level of understanding of the disease, but also to the comprehension of the normal process involving the patients age. For the author, there are yet no consistent established set of values for what would be a normal level of impairment in the elderly. To overcome these difficulties, some researchers (Braak & Braak, 1997; Elias et al., 2000; Kawas et al., 2003) have demonstrated that the AD first symptoms appears relatively early in life, and it evolves during lifetime. This fact raises the chances of identifying the pathology decades before a clinical diagnosis of dementia can be made.

Trying to detect potential patients with AD as early as possible, many studies (Brasil, Pinheiro & Coelho, 2009; Castro et al., 2007a;b; Mortimer et al., 2005; Nestor et al., 2004) have investigated potential tests and exams that, through a functional and cognitive analysis, may help the early AD detection. In this context, to evaluate the effectiveness of our MCDA classification approach in the early AD detection, we have developed a special-purpose AD-related database by following the recommendations of the Scientific Department of Cognitive Neurology and Aging of the Brazilian Academy of Neurology (Nitrini et al., 2005) and by making use of a neuropsychological battery of exams made available by the well-known Consortium to Establish a Registry for Alzheimer's Disease (CERAD) (Fillenbaum et al., 2008). Various experiments have been performed over this database in

a manner as to either fine-tune the components of the MCDA model or to compare its performance level with that exhibited by other state-of-the-art classification algorithms.

In the present study, two Multicriteria Decision Analysis (MCDA) classification approaches, which are developed upon the method recently proposed by Goletsis et al. (Goletsis et al., 2004) (referred to hereafter as *gMCDA classifier*) and the well-known MCDA PROAFTN classification procedure (Belacel, 2000), are employed towards the effective early diagnosis of Alzheimer's disease. The classifiers make use of the concept of prototypes, that is, special alternatives representing the classes of a problem, and has associated with itself some control parameters related to the expert's preference modeling process. As some of the experiments reported here reveal, the appropriate selection of prototypes as well as the calibration of control parameters are key issues to leverage the classifiers' performance. This way, our approach combines two complementary techniques, one based on ELECTRE IV methodology (Roy, 1996) and the other on a customized genetic algorithm (Eiben & Smith, 2003), in order to select the best prototypes and effectively calibrate the control parameters, respectively.

2. Alzheimer disease classification

The very early detection of Alzheimer's disease (AD) has been deeply investigated in numerous studies in the past years. These studies have demonstrated that the pathology usually arises decades before the clinical diagnosis is effectively made, and so a reliable identification of AD in its earliest stages is one of the major challenges clinicians and researchers face nowadays. The following researches evolved to tackle the classification of AD and other dementia.

A classification problem refers to the assignment of a group of alternatives to a set of predefined classes, also known as categories. During the last decades these problems have been tackled using a high variety of statistical and machine learning techniques. Recently, the area of Multicriteria Decision Aid (MCDA) (Figueira et al., 2005; Roy, 1996) has also brought new methodologies and techniques to solve these problems.

The main difference between the MCDA classification methods and others coming from related disciplines, as artificial neural networks (ANN), Bayesian models, rule-based models, decision trees, etc. (Witten & Frank, 2005), lies in the way that the MCDA methods incorporate the decision maker's preferences into the categorization process. In the ANN field, for instance, the work of French et al. (French et al., 1997) performs a comparison between an ANN model and a linear discriminant analysis (LDA) algorithm to classify and to stage the degree of dementia. The results demonstrated that the ANN algorithm clearly outperformed the LDA one in terms of classification accuracy, highlighting the utility of using ANN for group classification of patients with AD and staging dementia severity using neuropsychological data.

Figueiredo et al. (Figueiredo et al., 1995) present an algorithm that classifies individuals into four different groups (i.e., clinically diagnosed groups of elderly normal, demented, AD, and vascular dementia subjects). The classification is performed after the analysis of computer tomography image data from brain and using an optimal interpolative neural network. Another classification work related to dementia disorders among the elderly (Zaffalon et al., 2003) uses a naïve credal classifier to address two different classification problems: discrimination between demented and control patients, and the assignment from among the

different types of dementia. The dataset was developed from a set of measures collected among of a series of computerized tests (tasks), which assess some cognitive faculties of the patient.

Sandip et al. (Sandip et al., 2007) realize the AD classification based on a molecular test that evaluates characteristic changes in the concentrations of signaling proteins in the blood, generating a detectable disease-specific molecular phenotype. By this way, through a molecular biomarker in blood plasma, the model classifies the patients into AD or non-AD and identifies those presymptomatic individuals with mild cognitive impairment which will eventually convert to Alzheimer's disease.

In the MCDA field, a decision making model has been recently proposed by Castro and Pinheiro (Castro et al., 2007a;b; 2008) to assist the specialist in the early diagnosis of the Alzheimer's disease. Differently from our approach, this model uses the Macbeth software (Bana et al., 2003) to construct the judgement matrices and the value scales for each fundamental point of view (FPV) already defined. Each patient's information is judged by the decision maker for each FPV; then the Macbeth software generates the value scales that will be used in the final judgment of the patient's diagnosis. Instead of providing the classification itself, this sort of model gives the possibilities of a patient acquiring or not a certain type of dementia in the future.

In the present chapter, some experiments were performed enclosing the Multicriteria Decision Analysis (MCDA) classification techniques, that will be described in next section, towards the effective early diagnosis of the Alzheimer's disease.

3. Multicriteria decision analysis

Zopounidis and Doumpos (Zopounidis & Doumpos, 2002) define that the decision making problems, according to their nature, the policy of the decision maker, and the overall objective of the decision, may require the choice, ranking, or the assignment of the considered alternatives into predefined classes.

The practical approach that concerns the classification problems motivated researches in developing different methods and mathematical models to solve these problems trying to achieve the highest classification rate. A substantial overview on MCDA methods can be found in (Belacel, 2000; Massaglia & Ostanello, 1991; Mousseau et al., 1999) where the authors address the definitions and the problems that are involved in the decision making process.

These methods have been successfully applied to real world problems. The major difficulty during their employment, however, is that, in order to produce models that comply with the decision maker's expectations, a set of control parameters, such as threshold variables, weights, coefficients, etc., needs to be properly set in advance, which turns out to be a hard task to be dealt with. Some authors, like Belacel (Belacel, 2000) and Jacquet-Lagrèze & Siskos (Jacquet-Lagrèze & Siskos, 2001), have already provided some alternatives to counter this sort of drawback, although their solutions seem to be rather specific to the contexts that were investigated and yet no general recipes are available to be deployed in all methods and circumstances.

As pointed out by Zopounidis and Doumpos (Zopounidis & Doumpos, 2002), the great majority of works conducted on the MCDA classification theme has focused on the development of novel MCDA classification methods, not giving much emphasis on

characterizing and comparing their distinctive problems. Likewise, the authors also advocate that future research on this field should consider a more deep investigation into some important practical issues, such as the analysis of the interdependencies of the control parameters of the algorithms, the statistical validation of the generated models, the analysis of performance over large data sets, and the establishment of links between MCDA classifier models and those coming from related disciplines, such as Pattern Recognition, Machine Learning, and Data Mining (Witten & Frank, 2005).

In this context, we have developed an approach, to assist the doctors during the early stages of AD, upon a specific MCDA classification method, which is also composed of two complementary techniques: one responsible for eliciting the values of the classifier's control parameters and the other in charge of selecting the best prototypes from the dataset in accordance with the decision maker's preferences. The chosen classifiers and the associated techniques are detailed in the sequel.

3.1 MCDA classification models

The first MCDA classification method that we have chosen to investigate was that proposed by Goletsis et al. (Goletsis et al., 2004). This method makes use of prototypes to serve as references against the new alternatives compared (matched) with it. One distinctive aspect of this scheme with respect to other MCDA-based classifiers is that it presents less control parameters to be adjusted (only some thresholds and criteria weights). In what follows, we provide further details of gMCDA.

As described by Goletsis et al. (Goletsis et al., 2004), during the comparison between an alternative and a prototype the first thing to be computed is the Similarity Index ($SI_j(a, b_p^h)$). This index is calculated for each criterion, and its objective is to model the criteria into a five zone similarity index. In order to compute this index, two thresholds must be specified.

The first threshold that needs to be specified is the similarity threshold, q_j , which represents the maximum allowed criterion difference $|g_j(a) - g_j(b_p^h)|$ between the alternatives and the prototypes. Using this, the alternatives can be judged similar under a specific criterion.

The second threshold used by the ($SI_j(a, b_p^h)$) computation is the dissimilarity threshold, p_j , representing the minimum allowed criterion difference between an alternative a and prototype b_p^h . This threshold needs to be defined in order to consider the criteria totally dissimilar.

The Similarity Index(SI) ($SI_j(a, b_p^h)$) is computed as described below:

$$SI_j(a, b_p^h) = \begin{cases} 1, & \text{if } |g_j(a) - g_j(b_p^h)| \leq q_j \\ \left(\frac{|g_j(a) - g_j(b_p^h)| - p_j}{q_j - p_j} \right), & \text{if } q_j < |g_j(a) - g_j(b_p^h)| < p_j \\ 0, & \text{if } |g_j(a) - g_j(b_p^h)| \geq p_j \end{cases} \quad (1)$$

After the computation of the similarity index, the next step is to compute the concordance index (CI). This index indicates the overall similarity concordance of an alternative a with a prototype b_p^h . This index is computed as follows:

$$CI(a, b_p^h) = \sum_j w_j SI_j(a, b_p^h), \quad (2)$$

where w_j is the weight of a specific criterion and $\sum_j w_j = 1$.

Each alternative will have its CI computed for all prototypes of all classes. After that, the next step is the computation of the membership degree (MD) of an alternative a to a category C^h . The best CI of a to all prototypes of C^h is given to the Membership Degree (MD). The MD is computed as follows:

$$MD(a, C^h) = \max\{CI(a, b_1^h), \dots, CI(a, b_{l_h}^h)\} . \quad (3)$$

Finally, the last step is the assignment of the alternative a to the category $C(a)$ with the maximum MD calculated in relation to all the groups of prototypes. The formula is presented below.

$$C(a) = \arg \max_h MD(a, C^h) . \quad (4)$$

The gMCDA method was first applied to the ischemic beat classification problem (Goletsis et al., 2004). According to the authors (Goletsis et al., 2004) the main difficulty encountered when applying their method is the specification of weights and thresholds p_j and q_j . Aiming to achieve better performances and trying to obtain an automated beat classification, the authors have incorporated a genetic algorithm for the adjustment of the parameters of the multicriteria method. Further information concerning the implementation of this algorithm can be found in (Goletsis et al., 2004).

3.2 PROAFTN classification model

The other MCDA classification method implemented also makes use of prototypes to serve as references against which the new alternatives are compared (matched) with. Differently from the gMCDA, the PROAFTN method (Belacel, 2000) substitutes the similarity relation by the outranking relation, through the calculation of an indifference index based on an alternative and a reference profile (prototype).

The assignment procedure consists of calculating the degree of membership of each alternative to be assigned to each class based on the fuzzy indifference relation between this alternative and each prototype. Belacel (Belacel, 2000) defines the principle in the following way: when the alternative a is judged indifferent to a prototype b_p^h according to the majority of criteria (majority principle) and there is no criterion which uses its veto against the affirmation " a is indifferent to b_p^h " (minority respect principle), the action a is considered overall as indifferent to a prototype b_p^h . In order to calculate the fuzzy indifference relations it is necessary to build the partial indifference indices using the concordance and non-discordance concepts to aggregate them. As in the gMCDA algorithm the alternative will be assigned into the class with the maximal membership degree value (Belacel, 2000).

In general, the prototype scores are given by intervals, so for each criterion g_j , we associate to each prototype b_p^h the interval $[S_j^1(b_p^h), S_j^2(b_p^h)]$, with $S_j^2(b_p^h) \geq S_j^1(b_p^h)$. The comprehensive indifference index is determined by aggregating the partial indifference indices. These indices indicate if the action a is indifferent or not to a prototype b_p^h according to a criterion g_j . The partial indifference relation is given as follows:

$$aI_j b_p^h \iff g_j(a) \in [S_j^1(b_p^h), S_j^2(b_p^h)], \quad (5)$$

If the value of the alternative a according to the criterion g_j is equal to $S_j^1(b_p^h)$ or to $S_j^2(b_p^h)$, the alternative a will be indifferent to prototype b_p^h according to Equation 5. However, considering

the imperfection and imprecision of the data, the alternative a on the criterion g_j can be assessed as: $g_j(a) = S_j^1(b_p^h) - \epsilon$ or $g_j(a) = S_j^2(b_p^h) - \epsilon$, where ϵ is a real number, which take very small values (Belacel, 2000). In this case, the application of the Equation 5 leads to transform the indifference situation into a non-indifference situation between the alternative a and prototype b_p^h according to criterion g_j , despite the fact that variation is not significant. In order to remedy this inconvenience, Belacel (Belacel, 2000) introduced two discrimination thresholds $d_j^-(b_p^h) \geq 0$ and $d_j^+(b_p^h) \geq 0$, which correspond, respectively to two functions of $S_j^1(b_p^h)$ and $S_j^2(b_p^h)$.

Formally, three comparative situations between the action a and prototype b_p^h according to criterion g_j are obtained using the two discrimination thresholds:

- If $S_j^1(b_p^h) \leq g_j(a) \leq S_j^2(b_p^h)$, then a is clearly indifferent to b_p^h ;
- If $[g_j(a) \leq S_j^1(b_p^h) - d_j^-(b_p^h)]$ or $[g_j(a) \geq S_j^2(b_p^h) + d_j^+(b_p^h)]$, then a is not indifferent to b_p^h ;
- If $[S_j^1(b_p^h) - d_j^-(b_p^h) < g_j(a) < S_j^1(b_p^h)]$ or $[S_j^2(b_p^h) < g_j(a) < S_j^2(b_p^h) + d_j^+(b_p^h)]$, then there is a weak indifference between a and b_p^h .

For each alternative a from the set of alternatives A to be classified and each reference alternative b_p^h of the class C^h , the partial concordance index on the criterion j is computed as follows.

$$C_j(a, b_p^h) = \min\{C_j^-(a, b_p^h), C_j^+(a, b_p^h)\}, \tag{6}$$

where,

$$C_j^-(a, b_p^h) = \frac{d_j^-(b_p^h) - \min\{S_j^1(b_p^h) - g_j(a), d_j^-(b_p^h)\}}{d_j^-(b_p^h) - \min\{S_j^1(b_p^h) - g_j(a), 0\}} \tag{7}$$

$$C_j^+(a, b_p^h) = \frac{d_j^+(b_p^h) - \min\{g_j(a) - S_j^2(b_p^h), d_j^+(b_p^h)\}}{d_j^+(b_p^h) - \min\{g_j(a) - S_j^2(b_p^h), 0\}} \tag{8}$$

The second step of PROAFTN computes the partial discordance indices. The aim of determining the discordance index $D_j(a, b_p^h)$ of the criterion g_j is to apprehend the fact that such a criterion is more or less discordant with the assertion " a is indifferent to b_p^h " (Belacel, 2000). The discordance index is maximum ($D_j(a, b_p^h) = 1$) when the criterion g_j uses its veto against this assertion aIb_p^h . It is minimum ($D_j(a, b_p^h) = 0$) when the criterion is not in discordance with this indifference (i.e. $C_j(a, b_p^h) \notin 0$). If the criterion g_j is in discordance (i.e. $C_j(a, b_p^h) = 0$ with indifference and it does not use its veto against this indifference, we have: $0 < D_j(a, b_p^h) < 1$, which represents the intermediary zones between the non-discordance and discordance situations.

The veto thresholds $v_j^-(b_p^h)$ and $v_j^+(b_p^h)$ such as $v_j^+(b_p^h) \geq d_j^+(b_p^h)$ and $v_j^-(b_p^h) \geq d_j^-(b_p^h)$, $j = 1, \dots, n$, are used to define the values from which the action a is considered as very different to prototype b_p^h for criterion g_j (Belacel, 2000).

The discordance index ($D_j(a, b_p^h)$) is represented between the values $S_j^1(b_p^h) - v_j^-(b_p^h)$ and $S_j^1(b_p^h) - d_j^-(b_p^h)$ on one hand and $S_j^2(b_p^h) + v_j^+(b_p^h)$ and $S_j^2(b_p^h) + d_j^+(b_p^h)$ on the other hand,

by the linear interpolation function.

$$D_j(a, b_p^h) = \max\{D_j^-(a, b_p^h), D_j^+(a, b_p^h)\}, \quad (9)$$

where,

$$D_j^-(a, b_p^h) = \frac{g_j(a) - \max\{g_j(a), S_j^1(b_p^h) - d_j^-(b_p^h)\}}{d_j^-(b_p^h) - \max\{S_j^1(b_p^h) - g_j(a), v_j^-(b_p^h)\}} \quad (10)$$

$$D_j^+(a, b_p^h) = \frac{g_j(a) - \min\{g_j(a), S_j^2(b_p^h) + d_j^+(b_p^h)\}}{-d_j^+(b_p^h) + \max\{-S_j^2(b_p^h) + g_j(a), v_j^+(b_p^h)\}} \quad (11)$$

After that, the next step calculates the fuzzy indifference relation as:

$$I(a, b_i^h) = \left(\sum_{j=1}^n w_j^h C_j(a, b_p^h)\right) \times \left(\prod_{j=1}^n (1 - D_j(a, b_p^h))^{w_j^h}\right), \quad (12)$$

where w_p^h is a positive coefficient that indicates the importance of an attribute g_j to a class C^h .

The fourth step evaluate the fuzzy membership degree $d(a, C^h)$. The membership degree is computed for each class from the set of categories C by selecting the maximal values of indifference indices from the reference alternatives of each class:

$$d(a, C^h) = \max\{I(a, b_1^h), I(a, b_2^h), \dots, I(a, b_{L_h}^h)\}, h = 1, \dots, K. \quad (13)$$

The final step assigns the alternative to the class with the maximal membership degree:

$$a \in C^h \Leftrightarrow d(a, C^h) = \max\{d(a, C^1) / l \in \{1, \dots, K\}\} \quad (14)$$

As it can be seen, the implemented methods differ in the way they create relations between the alternatives in order to provide the classification. The PROAFTN method is based on the outranking relation while the gMCDA classifier provides a similarity relation with the application of similarity and dissimilarity thresholds. According to (I. Yevseyeva, 2007), these thresholds represent the maximal difference on the criterion value that is still small enough for two alternatives to be considered similar. In addition, it can be noted that the gMCDA classifier presents less control parameters to be adjusted (only some thresholds and criteria weights). On the other hand, the PROAFTN method encloses a veto threshold, which discriminates situations of discordance with the indifference relation.

The works of Brasil et al. (Brasil et al., 2010; Brasil, Pinheiro, Coelho & Costa, 2009a;b) demonstrated the relevance of choosing the best prototypes for both classifiers. These studies have evidenced that the methods presented above are indeed very sensitive to the choice of prototypes and that their proper choice may be instrumental for leveraging the performance levels.

3.3 ELECTRE IV

One of the complementary techniques applied jointly with the gMCDA classifier tackles the problem of prototype selection. This technique is also based on the MCDA principles, but conversely is based on the concept of sorting of alternatives and criteria.

According to Zopounidis and Doumpos (Zopounidis & Doumpos, 2002), the indirect techniques are widely used for developing sorting models that employ the outranking concept. To apply this technique, the decision analyst specifies the parameters based on an interactive inquiry process with the decision maker. This process ensures that the decision maker preferences will be correctly captured in the model.

Differently from other similar algorithms (Keeney & Raiffa, 1993), the ELECTRE IV method (Roy & Hugonard, 1982) does not require the specification of a weight value for each criterion. Conversely, the decision analyst chooses the criterion that it wants to work with and then ELECTRE IV combines them to give birth to the outranking relations. This approach avoids the problem of trying to quantify how important a criterion is. Each criterion can be either defined as a benefit or cost criterion. When the decision analyst considers a cost criterion, the lower the criterion value, the higher its merit; the converse is true for a benefit criterion.

To employ this method to rank the alternatives of a class, the decision analyst should define only the preference and indifference thresholds for each criterion. Specifically in our MCDA approach, the ELECTRE IV method will assume the role of the indirect technique responsible for the prototype selection activity.

Basically, the ELECTRE IV method can be divided into five stages: 1) criteria selection; 2) calculus of the relative thresholds; 3) construction of weak and strong outranking relations; 4) construction of the downward and upward ranks; and 5) elicitation of the final rank.

The first step to employ the ELECTRE IV algorithm is to select the criteria that will be used during the ranking process. The second stage is the determination of the relative thresholds. This phase basically sets the relation of two alternatives under some criterion. It can be defined that two alternatives are indifferent, strictly preferred, or weakly preferred over a criterion k . After that, it is necessary to construct the weak and strong outranking relations for every pair of alternatives (Ukkusuri et al., 2007). At this point, an alternative i will either strongly or weakly outrank an alternative j based on several restrictions that compares the relative ranks and the thresholds defined (Ukkusuri et al., 2007). The next step determines the strengths, weaknesses and the qualification of each alternative, and, based on these numbers, defines the downward and upward ranks. Finally, the final rank is set using the mean of the upward and downward ranks.

3.4 Genetic algorithm

Evolutionary computation is the field of research that draws ideas from evolutionary biology in order to develop search and optimization techniques (known as evolutionary algorithms) for solving complex problems (Back et al., 1997). Most of these techniques are rooted on the neo-Darwinian theory of evolution, which states that a population of individuals, is capable of reproducing and subjected to (genetic) variation followed by selection, result along time in new populations of individuals increasingly more fit to their environment.

Genetic algorithms (GAs) comprise the class of evolutionary algorithms that uses a specific vocabulary borrowed from natural genetics (Eiben & Smith, 2003). The data structures representing the individuals (genotypes) of the population are often called chromosomes; these are one-chromosome (haploid) individuals encoding potential solutions to a problem. In standard GAs, the individuals are represented as strings of bits. Each unit of a chromosome is

termed a gene, located in a certain place in the chromosome called locus. The different values a gene can assume are the alleles. The problem to be solved is captured in an objective (fitness) function that allows evaluating the adequacy of any potential solution.

As each chromosome corresponds to the encoded value of a candidate solution, it has to be decoded into an appropriate form for evaluation and is then assigned a fitness value according to the objective. For each chromosome is assigned a probability of reproduction, so that its likelihood of being selected is proportional to its fitness relative to the other chromosomes in the population. If the fitness of each chromosome is a strictly positive number to be maximized, selection is traditionally performed via an algorithm called Roulette Wheel selection (Eiben & Smith, 2003). The assigned probabilities of reproduction result in the generation of a population of chromosomes probabilistically selected from the current population. The selected chromosomes will generate offspring via the use of probabilistic genetic operators, namely, crossover (recombination of gene blocks) and mutation (perturbation through genetic variation) each one associated with a specific rate. Each new generation contains a higher proportion of the characteristics of the previous generation good members, providing a good possibility to converge to an optimal solution of the problem.

According to (Eiben & Smith, 2003), GAs have successfully been applied to a wide variety of problems, including those which are hard to be solved by other methods. In the MCDA field, their application primarily concerns the task of control parameter optimization (Brasil, 2009; Brasil, Pinheiro & Coelho, 2009; Goletsis et al., 2004; Gouvenir & Erel, 1998), the same investigated in this work.

4. Diagnosis of Alzheimer's disease

Despite the fact that the final AD diagnosis is performed by a microscopic brain tissue exam, through a biopsy or necropsy, Chaves (M.I.F. Chaves, 2000) presents evidence that it is not necessary to wait for the patient's death to know, with certainty, what is causing the symptoms and the perceived behavior.

In this context, the doctors can make a "probable" or "possible" AD diagnostic. Pinholt et al. (Pinholt et al., 1987) attests the difficult associated with the process of AD detection, wherein the authors highlight that, despite the high incidence of dementia in the elderly population, doctors fail to detect them in 21 to 72% of the cases. There is a vast number of clinical instruments that assist the clinician to obtain the diagnosis. Castro's work (A. K. Castro, 2008) summarizes some instruments as follows:

- Complete patient's and family's medical history;
- Neurological exam and neuropsychological tests;
- Review of the use of drugs;
- Physical examination;
- Psychiatric evaluation;
- Laboratorial tests like blood and urine exams;
- Image exams like: computed tomography, magnetic resonance, single photon emission computed tomography (SPECT) and positron emission tomography (PET).

Although these exams are not specific for the AD detection, they can increase its diagnosis accuracy by showing some issues related to its cognitive activity. As a neurodegenerative

disease (once the brain cells are lost, they cannot be replaced), the current research challenge is to make a premature diagnose, before the cognitive functioning is impaired (K.R. Daffner and L.F.M. Scinto, 2000). In this context, Rentz e Weintraub (D.M. Rentz and S. Weintraub, 2000) stresses that the neuropsychological deficits are still the best way to detect the AD early symptoms.

It is a common sense that the AD early diagnosis can bring benefits to the patients and their families. With the constant development of drug therapies and therapeutic advances, an early treatment can delay the progress of functional decline. The families and all the stakeholders can feel the benefits as they prepare for the patient management, raising their overall quality of life (A. K. Castro, 2008).

Besides the quality of life issue, another major factor that stimulates the early diagnosis is the financial one. According to some studies of the Alzheimer's Association, the Alzheimer disease generates high costs, being one of the most expensive diseases, losing only to the cancer and cardiovascular diseases (R. Koppel, 2002). The main reason that raises this cost is the need of a multidisciplinary treatment. Normally, the AD affects one or more cognitive areas, such as: attention, perception, memory, reasoning, sense, imagination, thinking and language. For such cases, it may be necessary treatments with psychologists, neurologists, neuropsychologists, geriatricians, psychiatrists, physiotherapists, occupational therapists, etc. Advanced cases can require the constant presence of the family and/or care providers, that are responsible for assisting the patients as they lose their abilities to interpret and express what is happening with their bodies and minds (A. K. Castro, 2008; L.P. Gwyther, 1985).

As mentioned before, this case study seeks to assist the decision maker (clinician) in the early AD diagnosis. To achieve this objective, we have manually designed a specific dataset of cases taking as reference the neuropsychological battery of CERAD standardized assessments and the Brazilian consensus of cognitive and functional evaluation. These are discussed in the following two subsections.

4.1 CERAD

The Consortium to Establish a Registry for Alzheimer's Disease (CERAD) was founded in 1986 after the Health Research Extension Act of 1985 with a specific focus on issues of diagnosis and diagnostic standardization (Fillenbaum et al., 2008). At that time, besides the fact that there had been an increasing interest over the illness, there was no uniform guideline over some issues, like diagnostic criteria, testing methods, and classifications of the disease severity, that could be followed. CERAD is a distinctive collaborative initiative to attend to this need.

CERAD has developed some standardized assessment instruments from different manifestations of Alzheimer's disease: clinical neuropsychology; neuropathology; behavior rating scale for dementia; family history interviews; and assessment of service needs. In this way, the CERAD battery improved the ability of specialists and researchers to describe and correlate clinical, neuropsychological, and neuropathologic aspects of AD.

4.2 The novel dataset

In order to provide a way to detect the presence of AD as soon as possible, we have followed the recommendations of the Scientific Department of Cognitive Neurology and

Aging of the Brazilian Academy of Neurology (Nitrini et al., 2005) while crafting our dataset of cases. This consensus specifies the recommendations over the clinical diagnosis of AD through a functional and cognitive perspective, and therefore the database was designed by following the strategy of correlating clinical and neuropsychological assessments of CERAD with recommendations provided by the Brazilian consensus.

In particular, the language evaluation exams allow for both a quantitative and qualitative diagnosis, showing the profile of the linguistic disorder (Nitrini et al., 2005). For the Brazilian consensus, the Boston Naming Test is one of the recommended tests that can be applied to break down the language aspects of a patient. This way, the first criterion (attribute) considered in the dataset relates to the amount of right answers given by each patient.

According to (Nitrini et al., 2005), the dementia diagnosis should be established in a clinical exam, documented as the Mini-Mental State Examination. To comply with the consensus, we turned this assessment into the second criterion associated with each case. This criterion reflects the sum of answers correctly assigned by each patient.

The third AD cognitive criterion designates a set of cognitive skills related to social relationships and that guarantee a proper, responsible, and effective conduct of the patient (Nitrini et al., 2005). Among the tests available in CERAD battery, we have used the Verbal Fluency exam. This test requests the patient to verbalize the highest number of animals as possible during a certain period of time. The criterion is defined by the number of items mentioned in a minute, excluding the repeated ones.

One of the main characteristics of AD is the impairment of memory. The Brazilian consensus stresses the importance of the memory evaluation and suggests the memorization of lists of words as an exam that can be applied to detect any sort of brain impairment during the early stages of the disease. This exam asks the patient to remember a ten-word list after a short period of time to evaluate the status of the short-term memory. The CERAD assessment applies three lists of ten words, so the database criterion we have devised specifies the overall number of words that were remembered by the patient.

The last criterion introduced relates to the concept of constructional ability. This CERAD assessment provides a non-verbal measure of the patient's mental health through the manipulation of geometric figures. The criterion denotes the number of elements correctly-assigned by the patient.

Besides the fact that the neuropsychological assessments available in the CERAD battery of exams were applied to more than 5,000 patients, only 119 cases could be effectively used in our experiments. This number was achieved after cross-correlating the neuropsychological and clinical assessments in order to certify whether the patient had effectively developed AD or not. By these means, the resulting dataset encompasses 5 criteria and 119 alternatives (cases).

5. An integrated approach to diagnose the Alzheimer's disease

In this section, we provide details of the experiments we have conducted so far over the database introduced previously. First, we concentrate on the prototype selection and control parameter calibration tasks conducted, respectively, by the ELECTRE IV and GA engines. The achieved results and a flavor of comparison with some state-of-the-art classifiers are presented. The gMCDA and PROAFTN classification methods were applied in order to provide the results.

Criteria	Description	+p	+q	-p	-q
C1	Boston Naming Test	0.9	0.39	-0.9	-0.39
C2	Mini-Mental State Examination	0.9	0.39	-0.9	-0.39
C3	Verbal Fluency	1.1	0.45	-1.1	-0.45
C4	Word List	1.1	0.45	-1.1	-0.45
C5	Constructional Praxis	0.9	0.35	-0.9	-0.35

Table 1. Criteria preference and indifference thresholds

The intent of this experiment is to provide a comparison of the gMCDA and PROAFTN performance levels over the same circumstances (prototypes and optimized parameters). In such regard, we have decided to compare the gMCDA and the PROAFTN classifiers assisted with the ELECTRE IV and GA engines with their performances when acting alone.

5.1 ELECTRE IV engine

The ELECTRE IV method (Roy & Hugonard, 1982) has been applied to assist in the prototype selection task through an indirect perspective. In such case, the decision analyst is responsible for providing the system with his/her preferences, which are effectively captured through the preference and indifference parameters (thresholds) associated with ELECTRE IV, so that the method can sort the alternatives.

Since the alternatives are ranked, the number of prototypes chosen while conducting the experiments was 7% of the original dataset. From that number, the prototypes were separated into their classes respecting their original distribution in the dataset. It is interesting to note that the application of the ELECTRE IV method can vary depending on the type of dataset that is under consideration. In cases where the problem presents more than two classes, the ELECTRE IV should be applied for each class separately, sorting the best alternatives of each class. This occurs because the classification problems often present conflicting criteria.

When applied to our AD dataset, as it only presents two categories, the ELECTRE IV engine needs to be applied only once to sort the patients from the most probable of not having Alzheimer to those most probable of manifesting the disease. In our experiments, we have ranked the patients from the non-AD to the AD category. For this purpose, we have established the same preference and indifference thresholds for all criteria, as they are all benefit criteria and have the same numerical ranges. For this dataset, all criteria were considered as relevant, so we have avoided discarding any attribute. Table 1 shows the preference and indifference values that were elicited for each criterion from the decision maker (clinician).

5.2 The genetic algorithm engine

According to our approach, after the best prototypes are selected by the ELECTRE IV engine, a customized GA is then employed in order to automatically estimate the gMCDA classifier's control parameters (thresholds). The GA components (Eiben & Smith, 2003) have been configured as follows: a population of 50 individuals (which initially is randomly generated) is evolved at each generation; the Roulette Wheel operator is used to select individuals to reproduce; individuals are recombined through a single-point crossover and the offspring is mutated according to a uniform distribution over the parameters' ranges; the crossover and

	1	2	3	4	5	6	7	8	9	10	Mean
gMCDA	91.66	84	83.33	95.83	91.66	95.83	87.5	95.83	91.66	85.71	90.28
Random gMCDA	58.82	56.3	64.71	58.82	55.46	58.82	67.86	67.86	67.86	70.83	62.73
PROAFTN	79.17	76	83.33	79.17	83.33	83.33	75	79.17	83.33	67.86	78.97
Random PROAFTN	70.83	74	73.75	75.42	72.92	76.67	76.25	77.92	77.92	68.57	74.42

Table 2. Performance of the classifiers when applied to the 10 test sets (results are show in percentage).

mutation rates are 80% and 15%, respectively; and the stop criterion adopted is to go through 500 generations of evolution.

To experiment with the GA, we have randomly generated 10 pairs of stratified training/test datasets from the original database, allocating 80% of the samples for training and the remaining for test. After the training phase, the best chromosome (configuration of thresholds) discovery is applied to the test data.

Classification Algorithm	Classification Rate (%)
J48	75.63%
NBTree	84.033%
OneR	82.352%
NaiveBayes	75.63%
gMCDA Classification Model	90.28%
PROAFTN Classification Model	78.97%

Table 3. Performance measures for the AD diagnosis.

5.3 Classification results

Table 2 shows the performance levels achieved by the classifiers when they had the parameters and prototypes optimized by the developed methodology. It is easily noticeable that the classifiers show a high sensitivity to the choice of prototypes and cut-off threshold values. As it was demonstrated in following works (Brasil et al., 2010; Brasil, Pinheiro, Coelho & Costa, 2009b), where the impact of the prototype selection is evidenced, it can be seen that for the AD dataset, the choice of the prototypes and control parameters seems indeed to be a key issue to be properly dealt with in order to leverage the classifiers' performance.

By contrasting the results without the application of the model with those produced by our approach in Table 2, it is possible to observe that, for some sets of prototypes, the proposed model could improve the classifiers' performance by more than 20%, taking the mean results over the 10 sets of random prototypes. Moreover, in some runs, the gMCDA classification rate could increase for as high as 33%.

Differently from the last experiment, the results produced by the gMCDA and PROAFTN classifiers can be compared here. Considering only the results produced by the model, the gMCDA classifier surpasses the PROAFTN in 90% of the executions. If we take a single execution, for example the last one (tenth), the difference comes close to 18%.

Finally, to provide a flavor of comparison with other classification algorithms, we have resorted to some well-known classification models available in the WEKA workbench (Witten & Frank, 2005). Table 3 brings the average accuracy levels achieved with each contestant model over the 10 derived datasets. The performance level achieved by the gMCDA classifier

was superior to those achieved by the other models. It should be emphasized that for each of the four additional classifiers we performed some preliminary experiments in order to manually calibrate its associated control parameters. However, we can not guarantee that the sets of parameters effectively obtained were in fact the optimal ones at all. From the results discussed above, one can conclude that the ELECTRE IV and GA engines have demonstrated good potential in solving the prototype and parameter selection problems.

6. Conclusion

The Alzheimer's Disease is a global health problem that is attracting the attention of the public authorities. During the last decades, the worldwide elderly population presented a continuous growth, increasing the incidence of AD. The high costs to treat this disease and the significant deterioration in the families' quality of life have brought a new challenge to researchers and the medical community: identifying the presence of this illness when its first signs appear.

In such regard, we designed a new database that takes as reference the functional and cognitive recommendations of the Scientific Department of Cognitive Neurology and Aging of the Brazilian Academy of Neurology and the CERAD's neuropsychological battery of exams. Unlike other studies over the AD, in this case study, our purpose was to assess the performance achieved by an extended version of two MCDA classification algorithms (Belacel, 2000; Goletsis et al., 2004) while coping with the AD early diagnosis.

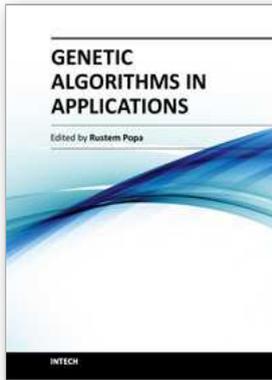
In this context, the employment of the ELECTRE IV algorithm revealed that the prototype selection task really exerts an important role over the MCDA classification process. Along with the ELECTRE IV, a GA engine was deployed to assist in the automatic calibration of the control parameter values (weights and thresholds) associated with both classifiers. In a general way, the devised MCDA approach achieved satisfactory levels of accuracy during the patient classification process over the conducted experiments, leveraging the performance of the classifiers and even comparing favorably against some well-known methods (Witten & Frank, 2005).

7. References

- A. K. Castro (2008). Um Modelo Híbrido Aplicado ao Diagnóstico da Doença de Alzheimer, *Master thesis*, Master Program in Applied Informatics, University of Fortaleza.
- Back, T., Fogel, D. & Michalewicz, Z. (1997). *Handbook of Evolutionary Algorithms*, Oxford Press.
- Bana, C. A., Corte, J. M. & Vasnick, J. C. (2003). Macbeth. LSE-OR Working Paper.
- Belacel, N. (2000). Multicriteria assignment method PROAFTN: Methodology and medical applications, *European Journal of Operational Research* 125: 175–183.
- Braak, H. & Braak, E. (1997). Frequency of stages of Alzheimer-related lesions in different age categories, *Neurobiology of Aging* 18: 351–357.
- Brasil, A. T. (2009). A novel approach based on multiple criteria decision aiding methods to cope with classification problems, *Master thesis*, Mestrado em Informática Aplicada, Universidade de Fortaleza (UNIFOR).
- Brasil, A. T., Pinheiro, P. R. & Coelho, A. L. V. (2009). Towards the early diagnosis of alzheimer's disease via a multicriteria classification model, *5th Evolutionary Multi-Criterion Optimization (EMO 2009)*, Lecture Notes in Computer Science, pp. 393 – 406.

- Brasil, A. T., Pinheiro, P. R. & Coelho, A. L. V. (2010). *Innovations and Advances in Computer Sciences and Engineering*, Springer Verlag, chapter The Impact of the Prototype Selection on a Multicriteria Decision Aid Classification Algorithm.
- Brasil, A. T., Pinheiro, P. R., Coelho, A. L. V. & Costa, N. C. (2009a). Comparison of two prototype-based multicriteria classification methods, *IEEE Symposium on Computational Intelligence in Multicriteria Decision-Making (MCDM 2009)*, IEEE Symposium Series on Computational Intelligence 2009, Nashville, USA.
- Brasil, A. T., Pinheiro, P. R., Coelho, A. L. V. & Costa, N. C. (2009b). Selecting prototypes for two multicriteria classification methods: A comparative study, *World Congress on Nature and Biologically Inspired Computing (NABIC 2009)*, Coimbatore, India.
- Castro, A. K., Pinheiro, P. R. & Pinheiro, M. C. (2007a). Applying a decision making model in the early diagnosis of Alzheimer's disease, *Rough Sets and Knowledge Technology, Second International Conference, RSKT 2007*, Vol. 4481 of *Lecture Notes in Computer Science*, pp. 149–156.
- Castro, A. K., Pinheiro, P. R. & Pinheiro, M. C. (2007b). A multicriteria model applied in the early diagnosis of Alzheimer's disease: A Bayesian approach, *Procs. of VI International Conference on Operational Research for Development*, pp. 9–19.
- Castro, A. K., Pinheiro, P. R. & Pinheiro, M. C. (2008). A multicriteria model applied in the early diagnosis of Alzheimer's disease, *Rough Sets and Knowledge Technology, Third International Conference, RSKT 2008*, Vol. 5009 of *Lecture Notes in Computer Science*, pp. 612–619.
- Celsis, P. (2000). Age-related cognitive decline, mild cognitive impairment or preclinical alzheimer's disease?, *Annals of Medicine* 32: 6–14.
- Davidoff, A. D. (1986). Issues in the clinical diagnosis of alzheimer's disease, *American Journal of Alzheimer's Disease and Other Dementias* 1(1): 9–15.
- D.M. Rentz and S. Weintraub (2000). *Early Diagnosis of Alzheimer's Disease*, Humana Press, chapter Neuropsychological Detection of Early Probable Alzheimer's Disease, pp. 169 – 189.
- Eiben, A. E. & Smith, J. E. (2003). *Introduction to Evolutionary Computing*, Springer-Verlag.
- Elias, M. F., Beiser, A., Wolf, P., Au, R., White, R. F. & D'Agostino, R. B. (2000). The preclinical phase of Alzheimer disease: A 22-year prospective study of the Framingham cohort, *Arch Neurol* 57: 808–813.
- Figueira, J., Mousseau, V. & Roy, B. (2005). ELECTRE methods, *Multiple Criteria Decision Analysis: State of the Arts Surveys* pp. 133–162.
- Figueiredo, R. J. P. et al. (1995). Neural-network-based classification of cognitively normal, demented, alzheimer disease and vascular dementia from single photon emission with computed tomography image data from brain, *Medical Sciences* 92: 5530–5534.
- Fillenbaum, G. G. et al. (2008). Consortium to establish a registry for Alzheimer's disease (CERAD): The first twenty years, *Alzheimer's & Dementia* 4: 96–109.
- French, B. M., Dawson, M. R. W. & Dobbs, A. R. (1997). Classification and staging of dementia of the alzheimer type: A comparison between neural networks and linear discriminant analysis, *Arch Neurol* 54(8): 1001–1009.
- Goletsis, Y., Papaloukas, C., Fotiadis, D., Likas, A., & Michalis, L. (2004). Automated ischemic beat classification using genetic algorithms and multicriteria decision analysis, *IEEE Transactions on Biomedical Engineering* 51(10): 1717–1725.
- Gouvenir, H. A. & Erel, E. (1998). Multicriteria inventory classification using a genetic algorithm, *European Journal of Operations Research* 105(1): 29–37.
- I. Yevseyeva (2007). Solving Classification Problems with Multicriteria Decision Aiding Approaches, *Phd thesis*, Faculty of Information Technology, University of Jyväskylä.

- Jacquet-Lagréze, E. & Siskos, J. (2001). Preference disaggregation: Twenty years of MCDA experience, *European Journal of Operational Research* 130: 233–245.
- Kawas, C. H. et al. (2003). Visual memory predicts Alzheimer's disease more than a decade before diagnosis, *Neurology* 60(7): 1089–1093.
- Keeney, R. L. & Raiffa, H. (1993). *Decisions with Multiple Objectives: Preferences and Value Trade-Offs*, Cambridge University Press.
- K.R. Daffner and L.F.M. Scinto (2000). *Early Diagnosis of Alzheimer's Disease*, Humana Press, chapter Early Diagnosis of Alzheimer's Disease: An Introduction, pp. 1 – 27.
- L.P. Gwyther (1985). Care of Alzheimer's patients: A manual for nursing home staff, *Technical report*, American Health Care Association.
- Massaglia, M. & Ostanello, A. (1991). N-TOMIC: A decision support for multicriteria segmentation problems, *Lecture Notes in Economics and Mathematics Systems*, Vol. 356, pp. 167–174.
- M.I.F. Chaves (2000). Diagnóstico Diferencial das Doenças Demenciais, in O. Forlenza & P. Caramelli (eds), *Neuropsiquiatria Geriátrica*, Atheneu, pp. 81 – 104.
- Mortimer, J. A. et al. (2005). Very early detection of Alzheimer neuropathology and the role of brain reserve in modifying its clinical expression, *Journal of Geriatric Psychiatry and Neurology* 18(4): 218–223.
- Mousseau, V., Slowinski, R. & Zielniewicz, P. (1999). ELECTRE TRI 2.0a: Methodological guide and user's documentation, *Technical report*, Université de Paris-Dauphine.
- Nestor, P. J., Scheltens, P. & Hodges, J. R. (2004). Advances in the early detection of Alzheimer's disease, *Neurodegeneration* 5: S34–S41.
- Nitrini, R., Caramelli, P., Bottino, C. M., Damasceno, B. P., Brucki, S. M. & Anghinah, R. (2005). Diagnóstico de doença de Alzheimer no Brasil: Avaliação cognitiva e funcional, *Arquivos de Neuro-Psiquiatria* 63: 713–719.
- Pinholt, E. M., Kroenke, K., Hanley, J. F., Kussman, M. J., Twyman, P. L. & Carpenter, J. L. (1987). Functional assessment of the elderly: A comparison of standard instruments with clinical judgment, *Arch Intern Med* 147: 484–488.
- R. Koppel (2002). Alzheimer's disease: The costs to U.S. businesses in 2002, *Report*, Alzheimers Association.
- Roy, B. (1996). *Multicriteria Methodology for Decision Aiding*, Kluwer Academic Publishers.
- Roy, B. & Hugonard, B. (1982). Ranking of suburban line extension projects on the Paris metro system by a multicriteria method, *Transportation Research* 16: 301–312.
- Sandip, R. et al. (2007). Classification and prediction of clinical Alzheimer's diagnosis based on plasma signaling proteins, *Nature Medicine* 13: 1359–1362.
- Ukkusuri, S. V., Karoonsontawong, A. & Kockelman, K. M. (2007). *Congestion Pricing Technologies: A Comparative Evaluation in New Transportation Research Progress*, Nova Science Publishers.
- Witten, I. H. & Frank, E. (2005). *Data Mining: Practical Machine Learning Tools and Techniques*, 2 edn, Morgan Kaufmann, San Francisco.
- Zaffalon, M., Wesnes, K. & Petrini, O. (2003). Reliable diagnoses of dementia by the naive credal classifier inferred from incomplete cognitive data, *Artificial Intelligence in Medicine* 29: 61–79.
- Zopounidis, C. & Doumpos, M. (2002). Multicriteria classification and sorting methods: A literature review, *European Journal of Operational Research* 138(2): 229–246.



Genetic Algorithms in Applications

Edited by Dr. Rustem Popa

ISBN 978-953-51-0400-1

Hard cover, 328 pages

Publisher InTech

Published online 21, March, 2012

Published in print edition March, 2012

Genetic Algorithms (GAs) are one of several techniques in the family of Evolutionary Algorithms - algorithms that search for solutions to optimization problems by "evolving" better and better solutions. Genetic Algorithms have been applied in science, engineering, business and social sciences. This book consists of 16 chapters organized into five sections. The first section deals with some applications in automatic control, the second section contains several applications in scheduling of resources, and the third section introduces some applications in electrical and electronics engineering. The next section illustrates some examples of character recognition and multi-criteria classification, and the last one deals with trading systems. These evolutionary techniques may be useful to engineers and scientists in various fields of specialization, who need some optimization techniques in their work and who may be using Genetic Algorithms in their applications for the first time. These applications may be useful to many other people who are getting familiar with the subject of Genetic Algorithms.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Amaury Brasil, Plácido Rogério Pinheiro and André Luís Vasconcelos Coelho (2012). Towards the Early Diagnosis of Alzheimer's Disease through the Application of a Multicriteria Classification Model, Genetic Algorithms in Applications, Dr. Rustem Popa (Ed.), ISBN: 978-953-51-0400-1, InTech, Available from: <http://www.intechopen.com/books/genetic-algorithms-in-applications/towards-the-early-diagnosis-of-alzheimer-s-disease-through-the-application-of-a-multicriteria-classi>

INTECH
open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the [Creative Commons Attribution 3.0 License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.