



## Article

# An Intelligent Multicriteria Model for Diagnosing Dementia in People Infected with Human Immunodeficiency Virus

Luana I. C. C. Pinheiro <sup>1,\*</sup>, Maria Lúcia D. Pereira <sup>1</sup>, Evandro C. de Andrade <sup>2</sup>, Luciano C. Nunes <sup>2</sup>, Wilson C. de Abreu <sup>3</sup> , Pedro Gabriel Calíope D. Pinheiro <sup>2</sup>, Raimir Holanda Filho <sup>2</sup> and Plácido Rogerio Pinheiro <sup>1,2,\*</sup> 

<sup>1</sup> Graduate Program in Clinical Care in Nursing and Health, Ceará State University, Fortaleza 60714-903, Brazil; maria.duarte@uece.br

<sup>2</sup> Graduate Program in Applied Informatics, University of Fortaleza (UNIFOR), Fortaleza 60811-905, Brazil; evandro.andrade@aluno.uece.br (E.C.d.A.); lcominn@uni7.edu.br (L.C.N.); pedrogcdpinheiro@edu.unifor.br (P.G.C.D.P.); raimir@unifor.br (R.H.F.)

<sup>3</sup> Porto School of Nursing, 4200-072 Porto, Portugal; wjabreu@esenf.pt

\* Correspondence: luana.ibiapina@aluno.uece.br (L.I.C.C.P.); placido@unifor.br (P.R.P.)

**Abstract:** Hybrid models to detect dementia based on Machine Learning can provide accurate diagnoses in individuals with neurological disorders and cognitive complications caused by Human Immunodeficiency Virus (HIV) infection. This study proposes a hybrid approach, using Machine Learning algorithms associated with the multicriteria method of Verbal Decision Analysis (VDA). Dementia, which affects many HIV-infected individuals, refers to neurodevelopmental and mental disorders. Some manuals standardize the information used in the correct detection of neurological disorders with cognitive complications. Among the most common manuals used are the DSM-5 (Diagnostic and Statistical Manual of Mental Disorders, 5th edition) of the American Psychiatric Association and the International Classification of Diseases, 10th edition (ICD-10)—both published by World Health Organization (WHO). The model is designed to explore the predictive of specific data. Furthermore, a well-defined database data set improves and optimizes the diagnostic models sought in the research.

**Keywords:** cognitive dementia; HIV; machine learning; verbal decision analysis; multicriteria model; medical diagnostic optimization



**Citation:** Pinheiro, L.I.C.C.; Pereira, M.L.D.; Andrade, E.C.d.; Nunes, L.C.; Abreu, W.C.d.; Pinheiro, P.G.C.D.; Holanda Filho, R.; Pinheiro, P.R. An Intelligent Multicriteria Model for Diagnosing Dementia in People Infected with Human Immunodeficiency Virus. *Appl. Sci.* **2021**, *11*, 10457. <https://doi.org/10.3390/app112110457>

Academic Editors: Giorgio De Nunzio and Kiril Tenekedjiev

Received: 3 May 2021

Accepted: 29 October 2021

Published: 7 November 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The continuous search for clinical diagnoses with greater accuracy has driven the use of hybrid models of computational technologies. These hybrid models use concepts of Artificial Intelligence (AI), Machine Learning (ML), and qualitative methods based on Verbal Decision Analysis (VDA), among others. The Machine Learning (ML) algorithms added agility to the diagnostic processes and have brought innovations to the field of health [1].

These hybrid models can be used to accelerate the diagnosis of neurological disorders such as cognitive and motor dementia in individuals infected with the Human Immunodeficiency Virus (HIV) and thus establish an earlier, more agile, and effective treatment [2].

HIV Associated Dementia (HAD) is the foremost source of morbidity and mortality in infected individuals and the most common in the younger population with HIV. Research has shown that age is not a prevalent factor for predisposition to dementia. Many individuals who are 50 years old or more became infected when they were much younger and have lived with the AIDS virus for an extended period asymptotically without seeking professional help or the use of antiretrovirals. When this age group has signs and symptoms of the disease, such as a cognitive deficit, which causes a negative impact on their quality of life and life expectancy, it causes great concern to health professionals and their relatives.

The difficulty of an early and accurate diagnosis of dementia, especially among HIV-infected people, is due to the rarity of symptoms at the beginning of the infection, requiring a prolonged observation period to reach a diagnosis. Finding various likely events and characteristics to establish a diagnostic pattern for symptoms associated with dementia in HIV-infected individuals is a significant challenge; however, this would allow for early treatment [3].

The social relevance of this research is due to the advantages of using Information and Communication Technologies to predict the high risk of dementia in people with HIV. These technologies can point out the necessary components to assess problems through the identification of respective classified and ordered characteristics related to HAD. Furthermore, such technologies can be used to plan health strategies and referrals to specialists instead of neuropsychological tests, which are considered time consuming and inadequate for some patients.

Although research into HIV has evolved over the last three decades, AIDS still has no cure. In addition, this disease is stigmatized in society, and as an aggravating factor, HIV-infected individuals are discriminated against in many situations, especially in work environments. Furthermore, society considers that competitiveness is a differential in terms of an individual's profile and productive capacity, and consequently, AIDS has become an exclusion factor in the process of wealth generation [4].

People living with HIV have an increased risk of cognitive impairment due to the neurotropism of the virus for the Central Nervous System (CNS), which is considered the second most affected site after the lymphatic system [5].

However, an early diagnosis points to promising horizons for these individuals. Early clinical interventions help control HIV in the body and allow these people to live longer and have a better quality of life [6].

The main characteristic of Machine Learning (ML) is its ability to learn through experience (historical data) and consequently improve its performance [7]. Machine Learning algorithms require consistent and accurate data to define whether a diagnosis is positive or negative.

Decision support methods aim to assist in the decision-making process to simplify the analysis of a problem and its alternatives and thereby justify the choice of a particular action. A decision process using verbal methods analyzes the decisionmaker's preferences more reliably. The basis for decision making is the exposure of the pros and cons and their analysis [8].

Research into the use of machine learning and multicriteria methods to make diagnoses [9] has been carried out for diagnosis prediction of Alzheimer. A battery of tests of patients with this diagnosis was performed, and the data were used to structure a decision tree based on the characteristics that played the main role in the diagnosis, and a scale of preference was established by analyzing the resulting tree [10]. Another work applied the deductive method and the scanning research technique to study a Wisconsin Breast Cancer data set case study. Here, the authors aimed to evaluate and compare the performance and effectiveness of machine learning models using TOPSIS [11]. On the other hand, the authors in [12] presented an automated model for cognitive workload assessment to accurately categorize the intensity of workload induced by multitasking situations. In [13,14], the authors employed machine learning techniques and a multicriteria decision on the Pima Indian diabetes data set for early diagnosis predicting of diabetes mellitus.

Technologies including machine learning and Multiple Criteria Decision Analysis cannot replace a physician's expertise, but they can support them in taking care of straightforward and time-consuming diagnostic tasks and assist in more demanding procedures.

The Verbal Decision Analysis (VDA) is a methodological approach of Multiple Criteria Decision Analysis (MCDA) that verbally supports solving decision making problems more realistically from the decisionmaker's point of view. When applying VDA methods, one is not required to assign numeric degrees of preferences to a criteria value concerning another. Thus, the procedure used to determine the importance of a decisionmaker in

a given context is psychologically valid. In addition, operations that acquire the verbal preferences of a decisionmaker are more stable and consistent [15,16].

The present study proposes a hybrid model based on Machine Learning (ML) and associated with a method of Verbal Decision Analysis (VDA) in order to improve the predictive power of the data characteristic from a set of structured information. The diagnostic models of neurological disorders, such as cognitive and motor dementia, were applied in this work to people infected with HIV (Human Immunodeficiency Virus).

This work is divided into five sections. Section 1 provides the contextualization and motivation for the study. Section 2 describes a theoretical framework that addresses the disease and concepts of the technologies used in the study. Section 3 presents the results of the hybrid model. Section 4 discusses the Results. Section 5 gives the Conclusion and an indication of future work on this subject.

## 2. Materials and Methods

This section discusses specific topics of AIDS and the consequent cognitive and motor dementia in people infected with HIV. In addition, it presents a model based on Information Technology (IT) methodologies to uncover the characteristics that are more determinant for early diagnosis of dementia in people with HIV/AIDS.

### 2.1. AIDS/HIV and Cognitive Dementia Considerations

In the early 1980s, the world was surprised by the appearance of a pathological agent, until then little known, which caused a disease called Acquired Immunodeficiency Syndrome (AIDS). This is, to date, one of the most significant public health challenges, as it mainly contaminates young people and, to a lesser extent, the elderly. The spread of the disease is partly due to medications such as sexual stimulants and hormone replacements, injectable drugs, and the low acceptance of the usage of condoms [17].

The World Health Organization (WHO) estimates that approximately 40 million people are infected with HIV. From the first cases of HIV/AIDS in the early 1980s to 2019, 76 million people have been infected with HIV, and 35 million have died from AIDS-related illnesses. In the early stages of the disease, 5% of AIDS patients have cognitive and motor dementia. This percentage reaches 20% in the later stages of the disease [18–20].

Furthermore, a recent analysis by the Joint United Nations Program on HIV/AIDS (UNAIDS) shows the potential damage caused by the COVID-19 pandemic in low- and middle-income countries worldwide due to insufficient healthcare for HIV-infected individuals. To prevent the dissemination of COVID-19, measures that restrict the movement of people, such as blockades and the closure of borders, have been introduced, and these are impacting both the production and distribution of AIDS drugs and generating increased costs and supply problems. Generic antiretroviral drugs are used to treat people infected with HIV. The final cost of antiretroviral medicines exported from India, for example, has increased at least 25% compared to regular prices [20].

In the mid-1990s, pharmacological research developed antiretroviral drugs, and these initiated a new form of treatment known as Highly Active Antiretroviral Therapy (HAART). This therapy provided an increase in the quality of life of HIV-infected individuals. However, even with HAART, the rate of individuals with AIDS and dementia increased, as cognitive impairments can be influenced by unemployment, social exclusion, and feelings of depression, among others [21]. Notably, AIDS is a syndrome that seriously compromises the immune system of HIV-infected individuals, and consequently, they are prone to other diseases and even death. HIV acts mainly on the lymphatic system, responsible for the body's defenses. However, the damage caused to an individual's cognitive and motor skills comes from the fact that HIV also attacks the Central Nervous System (CNS) [22].

The presence of the virus can be detected in brain tissue and the cerebrospinal fluid early on, extending throughout the evolution of the disease, even if it is asymptomatic to neurological disorders [23]. Cognitive and motor alterations in individuals with AIDS show that HIV crosses the blood–brain barrier through a Trojan-horse-like mechanism using

infected macrophages [24]. HIV infects glial cells that secrete neurotoxins by accessing the brain, causing the neuron to die and, consequently, causing a clinical neurological deficit. In addition, the virus is free in the acellular cerebrospinal fluid due to the difficulty of antiretroviral drugs to cross the blood brain barrier, making it difficult to fight the virus in this region [23,25]. Thus, the lack of access to the CNS and the cerebrospinal part by antiretroviral drugs give the virus a certain freedom of action in these tissues, enabling the development of neurological disorders such as cognitive dementia [22]. Thus, once the virus accesses the CNS, it can affect an individual's cognitive, sensory, and motor functions, leading to mental impairment with deficits in attention, learning, memory, information processing capacity, and problem solving [26].

Necropsy tests in HIV patients have indicated viruses in cortical and subcortical structures such as the frontal lobes, subcortical white matter, and basal ganglia [27]. The mechanisms that lead to cognitive impairment are not yet fully understood. However, neurotoxins released by microglia and periventricular macrophages cause the release of cytokines, which leads to modifications in the synaptic architecture of the cortex. The most common mechanism that leads to cell failure is apoptosis or programmed cell death [28]. Individuals with an advanced degree of this disease have deficits in several cognitive domains.

HIV-infected and asymptomatic individuals may manifest mood disorders associated with psychomotor slowness and decreased verbal memory. Another form of manifestation is a decrease in verbal and nonverbal cognitive functioning but without any manifestation of mood disorders [29].

According to the American Academy of Neurology, neurological disorders associated with HIV infection comprise the following categories: "minor motor disorder" and "HIV-associated cognitive dementia" [30]. Based on the two criteria mentioned and the research and observations made by publications after the introduction of HAART, a new review of the nosology of HIV-associated neurological disorders was proposed. This established the following categories: asymptomatic neurological disorder, mild HIV-associated neurological disorder, and HIV-associated dementia [31].

The criteria for defining a diagnosis are the deficits in cognitive, behavioral, and motor functions. The essential characteristic of cognitive and motor dementia is the loss of intellectual and motor functions, interfering seriously with social and occupational engagements [30]. At less advanced stages, a detailed neuropsychological assessment to determine the degree and nature of the cognitive impairment is necessary to identify any psychological morbidities such as depression and anxiety [32].

## 2.2. Information Technologies Used in the Proposed Model

### 2.2.1. Fundamentals of the "Random Forest" Algorithm

The Random Forest algorithm consists of decision trees capable of predicting the expected values of regressions and classifications.

Breiman introduced the Random Forest method, and the technique has since become an essential data analysis tool. The algorithm is generally recognized for its accuracy and ability to handle small sample sizes, capacity resource spaces, and complex data structures [32,33].

Theoretically, the analyses are less conclusive, and despite its wide use in practical contexts, little is known about the mathematical properties of Random Forest. So far, most studies have focused on isolated parts, whose main theoretical result offers an upper bound for the generalization error of forests in terms of correlation and strength of individuals trees [34].

The bagging procedures and the Classification and Regression Trees (CART) division criteria critically analyze the Random Forest [35]. Like a bootstrap aggregation contraction, bagging is a general grouping scheme that proceeds by generating subsamples from the original data set, building a predictor for each sample and deciding on the mean. Bagging is one of the most effective computationally intensive procedures to improve unstable estimates, especially for large and high-dimensional data sets. However, it is impossible

to find a suitable one-step model due to the complexity and scale of the problem. On the other hand, the CART-split algorithm builds individual trees to choose the best cuts perpendicular to the axes. The best value is selected at each node of the tree based on Gini impurity and the squared error of the regression.

In Breiman's Random Forest [33], each leaf, a terminal node of the individual trees, contains a fixed, preset number of observations, (this parameter is usually between 1 and 5).

CART was chosen for the induction of decision trees in this study. This choice was motivated by the fact that the algorithm does not restrict the nature of the explanatory variables because it generates results that are easy to interpret and easy to find interactions between the variables.

### 2.2.2. Fundamentals of the "Logistic Regression" Algorithm

Logistic Regression, also known in the literature as "logit regression", "maximum entropy classification (MaxEnt)", or "log-linear classifier", consists of a Machine Learning algorithm for classification problems and predictive analysis based on the probability concepts [36]. Logistic Regression is used to assign observations to a discrete set of classes. Logistic Regression is one of the most popular Machine Learning algorithms for binary classification, given a set of independent variables, and is used to predict a binary result (1 or 0, Yes or No, True or False) [32]. It has been applied successfully in various areas, such as Medicine [37], Finance [38], and Economics [39]. In this model, the probabilities describing the possible outcomes of a single trial are modeled using a logistic function, namely the logical relationship between a dichotomous response variable and a series of numerical (continuous, discrete) or categorical explanatory variables. [37–40].

Logistic Regression is a type of linear regression in which the probability log is used as a dependent variable. It predicts the probability of an event occurring by adjusting data to a logit function [41]. The algorithm makes several assumptions, such as independence, answers (logits) at all subpopulation levels, where the explanatory variable is normally distributed. The variance is constant between reactions and all the values of the explanatory variable. Intuitively, a transformation for the response variable is applied to produce a continuous probability distribution over the output classes bounded between 0 and 1; this transformation is called a "logistic" or "sigmoid" function [42].

A logistic function or logistic curve is an S, sigmoid curve, generated by the Equation (1) [43]:

$$f(x) = \frac{L}{1 + e^{-k(x-x_0)}} \quad (1)$$

where:

$x_0$  = value of the midpoint of the sigmoid curve,

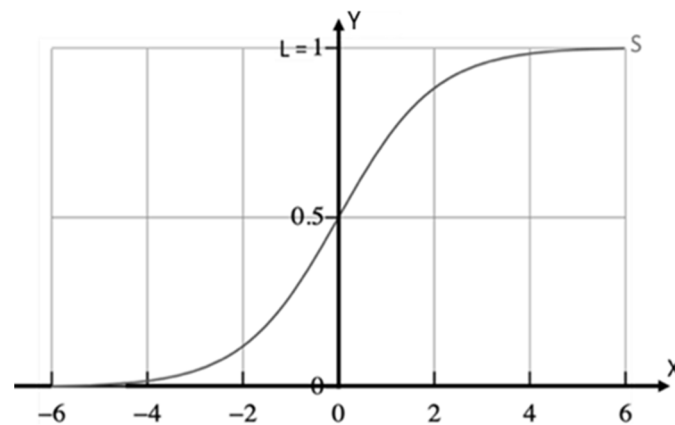
$L$  = maximum value of the curve,

$k$  = the logistic growth rate or slope of the curve.

For values of  $x$  in the real-number domain from  $-\infty$  to  $+\infty$ , the S curve, standard logistic sigmoid function, i.e.,  $L = 1$ ,  $k = 1$ ,  $x_0 = 0$ , as shown in Figure 1, is obtained in the graph with  $f$  approaching  $L$  when  $x$  approaches  $+\infty$  and with  $f$  approaching 0 (zero) when  $x$  approaches  $-\infty$ .

Logistic regression is more of a classifier than a regression technique, despite its name. This statement is especially true in medicine or the psychosocial sciences, focusing on predicting and explaining [44].

Moreover, for Logistic Regression, the response variable is quantitative; that is, the log of the chance of being classified in the  $i$ -th group of a binary or multiclass answer [45]. The probability of a positive event or diagnosis can be estimated directly in Logistic Regression. In the case of a dependent variable  $Y$  assuming only two possible responses or binary states of interest (True or False), there is also a set of  $p$  independent variables,  $X_1, X_2, \dots, X_p$ .



**Figure 1.** Sigmoid curve as defined in Equation (1), for  $L = 1, k = 1, x_0 = 0$ .

The logistic regression model can be written as follows:

$$P(Y = 1|X_1, \dots, X_p) = \frac{\exp(\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p)}{1 + \exp(\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p)} \quad (2)$$

The coefficients  $\beta_0, \beta_1, \dots, \beta_p$  are the regression coefficients that are estimated by the maximum likelihood of the data set. The probability of  $Y = 1$  for a new instance is then estimated by replacing  $\beta$  with its estimated counterparts and  $X$  with its realizations for the new instance considered in Equation (2). The new instance is then assigned to class  $Y = 1$  if  $P(Y = 1) > c$ , where  $c$  is a fixed limit, and class  $Y = 0$  otherwise. The commonly used limit  $c = 0.5$  produces a classifier called Bayes [46,47].

For discrimination of two groups, the classification rule is as follows:

- if  $P(Y = 1) > 0.5$ , then  $Y = 1$  is classified.
- if  $P(Y = 1) < 0.5$ , then  $Y = 0$  is classified.

In this study,  $P$  can be interpreted as the probability of an individual with HIV having a positive diagnosis for high risk of dementia, given the characteristics qualified during the assessment by the health professional, represented by,  $X_1, X_2, \dots, X_p$ .

The objective is to find a set of weights. The negative likelihood log is minimized over the defined training set using optimization techniques such as descending gradient or descending stochastic gradient [42]. Minimizing the negative likelihood log also means maximizing the  $p_i$  parameter estimate of selecting the correct class. The loss function that measures the difference between the fundamental truth label and the predicted class label is called the cross-entropy. If the forecast is too close to the true label, the amount of loss is low. Alternatively, the resulting log loss is more significant if the prediction is far from the true label.

The simplicity and interoperability of logistic regression can occasionally outperform other nonlinear models. Still, if the response variable is taken from a small sample, logistic regression models become unsatisfactory and perform poorly for binary responses [48].

The Logistic Regression can be categorized as follows:

- i. Binomial or binary:
  - Here, it addresses situations where the observed result for a dependent variable can have only two possible types, "0" or "1" ("sick" or "healthy").
  - They are usually referred to as "Logistic Regression".
  - It predicts the probability that an observation falls into one of two categories of a dichotomous dependent variable, based on one or more independent variables, which can be continuous or categorical [46,49].
- ii. Multinomial:

- This is an extension of a simple linear regression, used when we want to predict the value of a variable based on the value of two or more other variables.
  - Alternatively, when we want to predict the variable called a dependent variable (or sometimes an outcome, objective, or criterion variable).
  - The variables we are using to predict the value of the dependent variable are called independent variables (or sometimes predictive, explanatory, or regressive variables) [41,44].
- iii. Ordinal:
- This works with dependent and ordered variables.
  - It is usually called “ordinal regression”.
  - It is used to predict an ordinal dependent variable, given the independent variables [48].

### 2.2.3. Fundamentals of the “Naïve Bayes” Algorithm

Bayesian networks can be applied to classification problems using Bayesian classifiers. The simplest structure of the Bayesian Network is a classification algorithm called Naive Bayes, also known as the Bayesian classifier with k-dependence and the structure of Simple Bayesian Networks with K-Dependency [47].

Therefore, Naive Bayes is a simple probabilistic algorithm based on Bayes’ theorem. This algorithm uses training data to form a probabilistic model based on evidence of specific characteristics in an extensive database. The algorithm assumes independence between the features of the model, implying the absence of relationships between a given quality and the other attributes of the database [47].

A Naive Bayes classifier is structured in the calculation of the posterior probability distribution  $P(Y | X)$ , where  $Y = (y_1, y_2, \dots, y_p)$  is the random variable to be classified presenting categories “k”, and  $X = (x_1, x_2, \dots, x_p)$  is a set of discrete explanatory variables presenting “p”. In addition, to calculate the conditional probability  $P(Y | X)$ , this method assumes a probabilistic independence between the explanatory variables, which facilitates the computational approach following Equation (3):

$$P(Y = y_k | x_1, \dots, x_p) \propto P(Y = y_k) \prod_{i=1}^p P(x_i | Y = y_k) \quad (3)$$

Therefore, a Naive Bayes classifier is based on calculating the probability that a particular observation belongs to a specific category and classifies that observation into the most suitable variety, as shown in Figure 2.

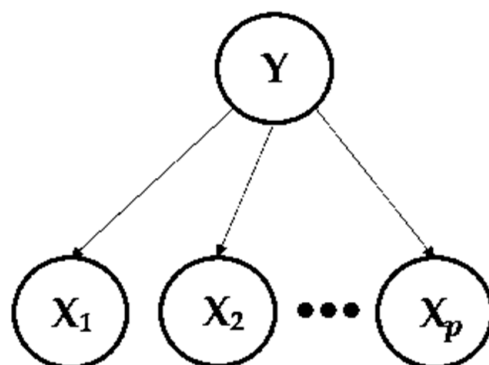


Figure 2. Bayesian Network.

### 2.2.4. Fundamentals of the Verbal Decision Analysis

The Verbal Decision Analysis (VDA) assumes that most decision-making problems can be expressed in natural language and consist of the decision-making process based on the representation related to the problem [50]. Notably, this methodology addresses

unstructured issues, characterized by the absence of logical and well-defined procedures, to be applied in its solution.

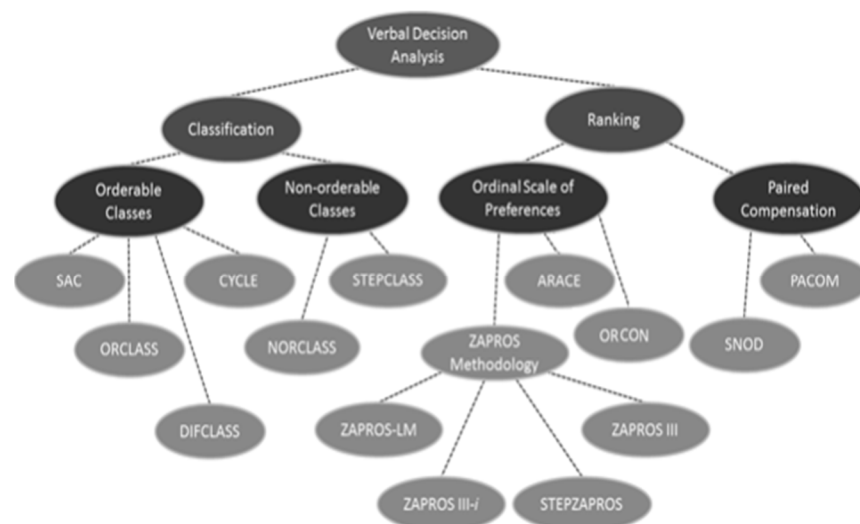
These problems are qualitative, and therefore, it is complex to organize, formalize, and numerically measure them. Furthermore, it is unusual to have all the data required to solve them. Thus, the analysis process is subjective, which requires collecting information from the decision-maker. In this research, the decision-makers were highly qualified health professionals with scientific experience and in clinical practice with HIV/AIDS individuals.

According to Larichev [50], the methods that make up the structure of Verbal Decision Analysis are ZAPROS-III, ZAPROS-LM, PACOM, and ORCLASS and their characteristics and applications. These methods require a large quantity of data to be analyzed by humans as follows:

1. Comparison of two assessments on a verbal scale by two criteria.
2. Assignment of multicriteria alternatives to decision classes.
3. Comparative verbal evaluation of alternatives according to different criteria.

These methods constitute models for DSS (Decision Support Systems) that help a decisionmaker classify or rank the alternatives of multiple attributes.

Figure 3 shows a simplified visualization of the various methods that make up the Verbal Decision Analysis methodology according to the problems.



**Figure 3.** Verbal Decision Analysis methodology. Source: adapted from [8].

This work uses the ZAPROS method, which adapts to the characteristics of the problem and allows the structuring of a decision rule used to compare the alternatives that are not changed, even if the set of alternatives is modified. In addition, this applies to problems with many alternatives [51,52].

In the Verbal Decision Analysis (VDA) structure, the ZAPROS-III-i method can be considered an evolution of the ZAPROS-III and ZAPROS-LM. Like the ZAPROS-LM and PACOM methods, the ZAPROS-III-i method aims to order a group of alternatives, from the most preferable to the least preferable [52].

Although ZAPROS-III-i applies a similar procedure to obtain the preferences of its successor, it implements modifications that make it more effective and more accurate for inconsistencies. Furthermore, the number of alternatives is substantially less than in ZAPROS-III and ZAPROS-LM.

The ZAPROS-III-i method is applied in three steps: Problem Formulation, Obtaining Preferences, and Comparing Alternatives [52].

#### **A. Problem Formulation:**

Given:



1.  $K = 1, 2, \dots, N$ , representing a set of  $N$  criteria;
2.  $n_q$  represents the number of possible values on the scale of  $q$ -th criterion, ( $q \in K$ ); for ill-structured problems, as in this case, usually  $n_q \leq 4$ ;
3.  $X_q = \{x_{iq}\}$  represents a set of values to the  $q$ -th criterion, and this set is the scale of this criterion;  $|X_q| = n_q$  ( $q \in K$ ), where the values of the scale are ranked from best to worst, and this order does not depend on the values of other scales;
4.  $Y = X_1 \times X_2 \times \dots \times X_n$  represents a set of vectors  $y_i$  (every possible alternative: hypothetical alternatives + real alternatives) in such a way that:  $y_i = (y_{i1}; y_{i2}; \dots; y_{iQ})$ , and  $y_i \in Y, y_{iq} \in X_q$  and  $Q = |Y|$ , such that  $|Y| = \prod_{q=1}^Q n_q$ .

**B. Elicitation of Preferences:**

The scale of preferences for quality variations (Joint Scale of Quality Variations JSQV) is constructed in this stage. The elicitation of preferences follows the order of steps shown in Figure 4 [53]. This structure is the same as proposed in [54]; however, substages 2 and 3 (numbered on the left side of the figure) were put together in just one substage.

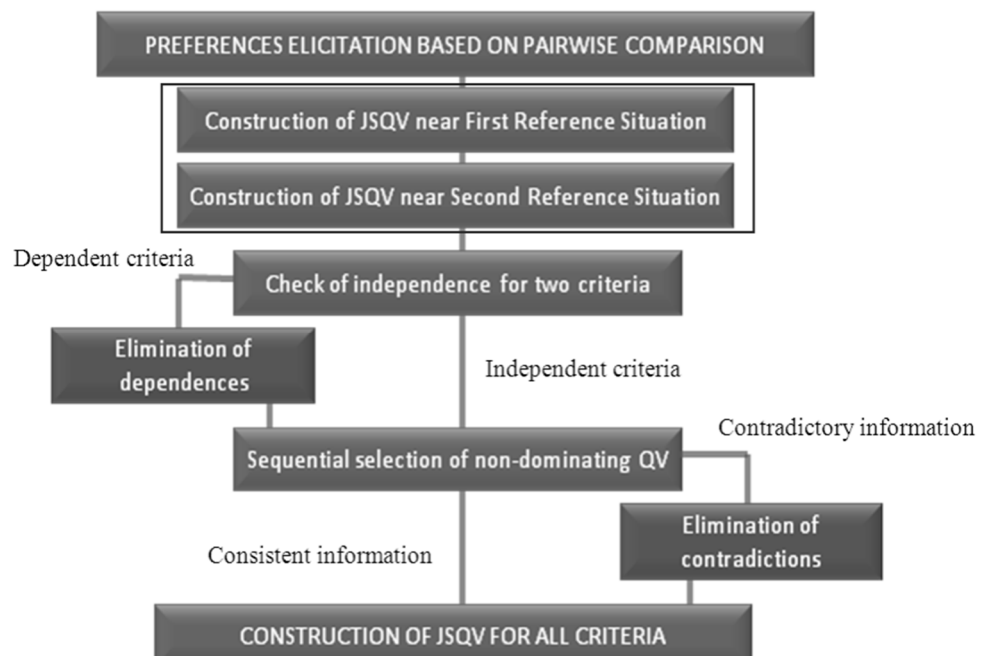


Figure 4. Elicitation of preferences process.

Instead of basing the decisionmaker’s preferences on the first reference situation and then establishing another scale of preferences using the second reference situation, we propose that the two substages be transformed into one. The questions made considering the first reference situation are the same as those considering the second. Thus, both situations are presented and must be regarded as answering the question not to cause dependence of criteria. The alteration reflects on an optimization of the process: instead of making  $2n$  questions, only  $n$  are made. The questions, about Quality Variations (QV) belonging to just one criterion, are made as follows: supposing a criterion  $A$  with  $X_A = A_1, A_2, A_3$ , the decisionmaker is asked about their preferences between the QV,  $a_1-a_2, a_1-a_3$  and  $a_2-a_3$ . Thus, there is a maximum of three questions to a criterion with three values ( $n_q = 3$ ).

The question for two criteria of preferences is formulated differently due to difficulties in understanding and a delay in the decisionmaker’s answers when the QV was exposed to different criteria. The question is made by dividing the QV into two items. For example, having the set of criteria  $k = A, B, C$ , where  $n_q = 3$  and  $X_q = q_1, q_2, q_3$  considering the pair of criteria  $A, B$  and the QV  $a_1$  and  $b_1$ , the decisionmaker should analyze which imaginary alternative would be preferable:  $A_1, B_2, C_1$  or  $A_2, B_1, C_1$ . However, this answer must be the same as the alternatives:  $A_1, B_2, C_3$  and  $A_2, B_1, C_3$ . If the decisionmaker answers that

the first option is better, then  $b_1$  is preferable to  $a_1$ , because it is preferable to have  $B_2$  on the alternative instead of  $A_2$ .

**C. Comparison of Alternatives:**

In order to reduce the number of cases of incomparability, we applied the same structure as proposed in [54] but modified the comparison of pairs of the alternative substage according to the one proposed in [55]. Figure 5 shows the structure of the comparison of the alternative process.

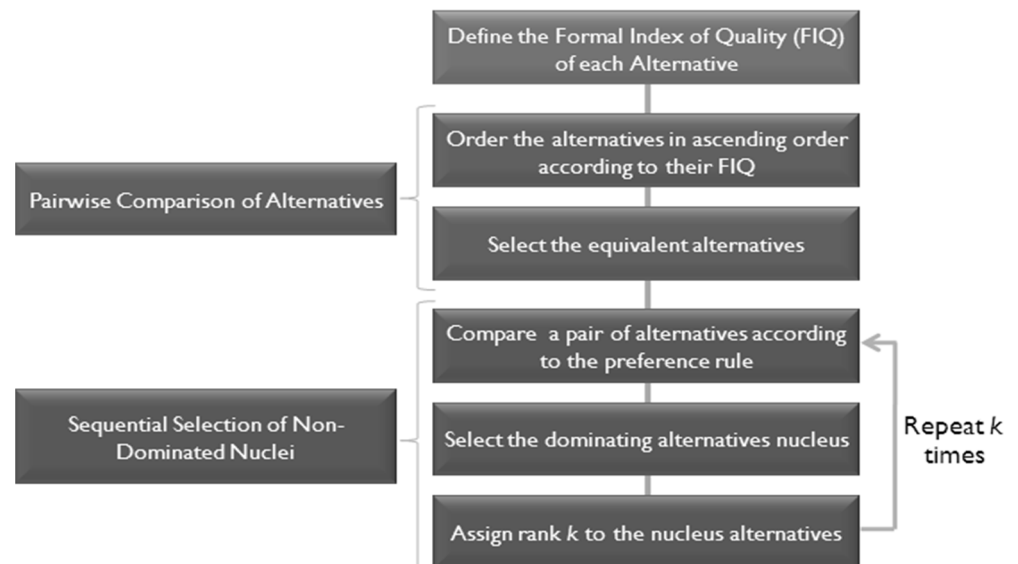


Figure 5. Alternative comparison process.

According to [31] and [15], the Decision Support Systems (DSS) are computer-based systems that aid the decisionmaker in informing and exploring the implications of their preferences, and this leads to a decision based on the problem comprehension. Figure 6 presents a flowchart with the steps to apply the ZAPROS-III-i/VDA.

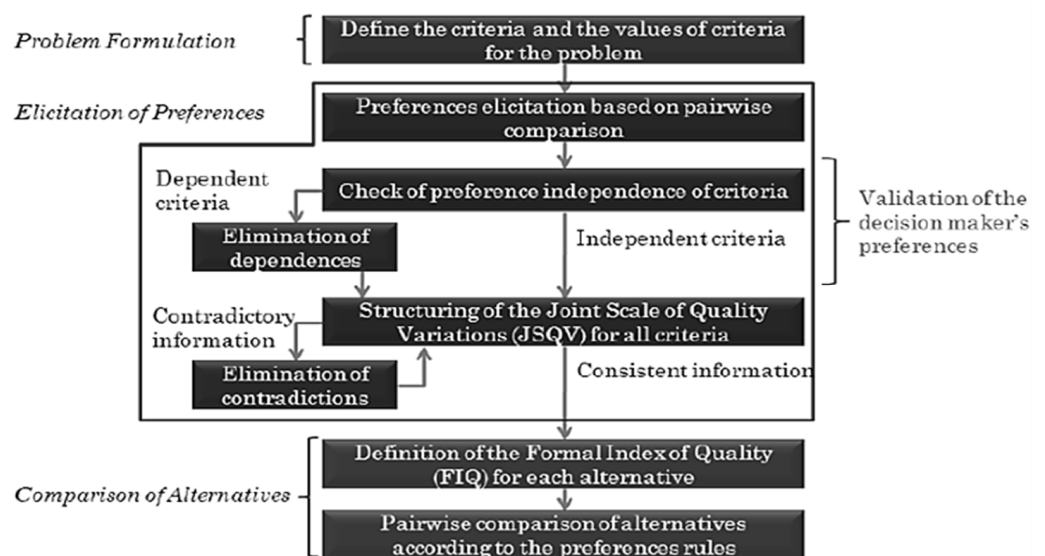


Figure 6. Procedure for applying the ZAPROS-III-i method. Source: adapted from [8].

A disadvantage of the method arises when there is exponential growth in the number of alternatives of the problem as well as the amount of information required to obtain the

preferences. To control this complexity, the number of criteria and values of the treated criteria are limited.

The method proposes that the two reference situations generated by the AI Algorithm and the one corresponding to the health professionals' decisions are analyzed simultaneously to optimize the process and avoid criteria dependence [53].

According to Tamanini and Pinheiro [52,53], these modifications increased the comparability of the method so that several alternatives defined as incomparable, when applying the ZAPROS method in its pure form, can now be compared directly or indirectly. However, these changes in the method process do not alter its computational complexity [56].

### 2.3. Proposed Hybrid Model Used in this Study

There is great interest in applications for Artificial Intelligence (AI) in healthcare in order to be able to acquire an early diagnosis and early treatment. Furthermore, AI can work with large amounts of data and complex systems and generate relevant results. Health events are more stochastic and less deterministic; therefore, they need a centralized predictive control system to learn from their own experience, improving the same algorithms they use to provide recommendations [57].

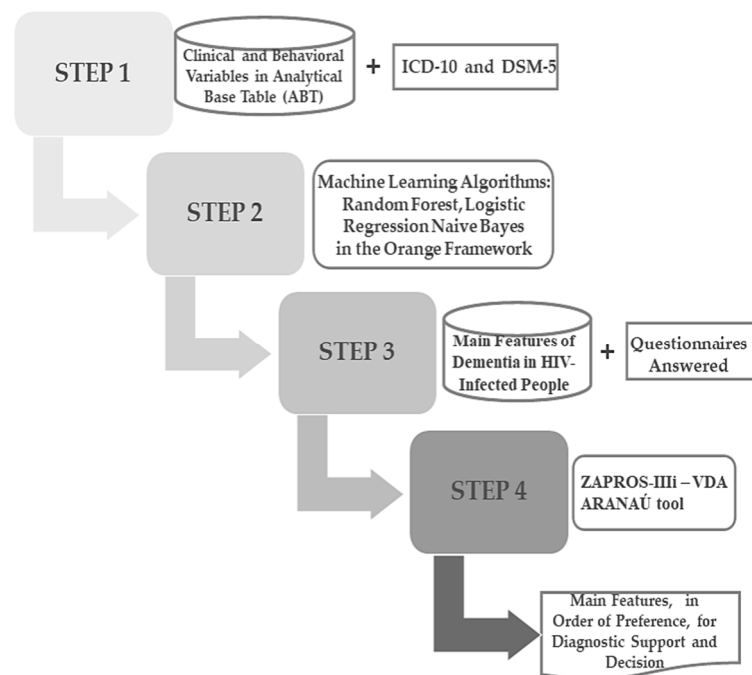
The main characteristic of Machine Learning technology, a subset of AI, is its learning from experiences recorded in historical data, and its performance improves according to this learning [58].

In this study, a large amount of structured data that was composed of many rows and columns containing unique numbers or values was converted to extract only the information able to improve the proposed model, refined from several test bases, one for each learning cycle. The origin of each test base of each cycle counted on the exploration of a large and varied amount of data structured in rows and columns. The learning cycle had to be repeated until the results were satisfactory, consequently making the algorithm more accurate.

The Dementia in People with HIV database was constructed using data collected from HIV-positive individuals aged fifty years and older with mental and clinical conditions suitable to participate in the database [59]. The data collection and monitoring took place at the outpatient clinic of the Reference Hospital for Infectious and Parasitic Diseases located in the city of Fortaleza, CE-Brazil. Data from 294 people were collected; unfortunately, due to the COVID-19 pandemic and the need for social isolation, no further increase in the sample was possible. For the training and testing basis, we applied k-fold cross-validation (with  $k = 10$ ), which performed the fitting procedure a total of ten times, with each fitting being performed on a training. The instruments used to acquire the data were: Evaluation of Clinical and Immunological Aspects; International HIV-Associated Dementia Scale; Brief Symptom Inventory; and Social Support and Instrumental Activities of Daily Living Scale. The individuals at high risk of dementia were identified after data collection. A total of 70% of the data was used for training and 30% for the tests. The collection was started after approval by the Research Ethics Committee of the Federal University of Ceará /Brazil, respecting the ethical precepts of nonmaleficence, beneficence, autonomy, and justice for its participants, under No. 3038967.

The main objective of this study is to propose a hybrid model based on Machine Learning, associated with a multicriteria Verbal Decision Analysis method for the early diagnosis of dementia in HIV-infected individuals. The VDA method chosen was the ZAPROS-III-i, as presented in Section 2.2.4, due to its presence in several multicriteria decision support models in the health area. In turn, three fundamental Machine Learning algorithms were applied that, according to [60], have been widely successfully implemented in various areas, mainly in the health field, for classification and regression purposes. The Random Forest algorithm was presented in Section 2.2.1, the Logistic Regression was presented in Section 2.2.2, and the third Machine Learning algorithm, Naïve Bayes, was presented in Section 2.2.3. This latter algorithm was included to serve as a third option to the Random Forest and Logistic Regression.

The hybrid model proposed in the present study is shown in Figure 7, where the operational steps of the process can be seen.



**Figure 7.** Structure of the proposed hybrid model operating process.

As can be seen in Figure 7, the process has four steps. The first step is made up of visual research and data filters; they are compiled from the DSM-5 (Diagnostic and Statistical Manual of Mental Disorders, 5th edition), the American Psychiatric Association, and the ICD-10 (International Classification of Diseases, 10th edition), and relevant information on cognitive and motor dementia from the HIV-infected individuals. This information is added to these individuals' clinical and behavioral data and to the research database, forming the Analytical Base Table (ABT).

In the second step, the ABT embedded in Step 1 is submitted for analyses by each of the Machine Learning algorithms used in this study: Random Forest, Logistic Regression, and Naive Bayes. It is essential to highlight that, in this study, the Orange Framework was used (version 3.29.3) for applying Machine Learning algorithms with default parameters.

Moreover, the Orange Framework is a set of machine learning and data mining tools for analyzing databases using Python script and visual programming. Orange is aimed at experienced users and programmers, as well as researchers using Data Mining. This framework is based on the C++ language, allowing search solution developers to work in the Python language [61].

Practitioners assessed patients before using Machine Learning and applied an instrument called HIV-Associated Dementia, which reveals whether the patient is at high or low risk of dementia. This instrument has four items to assess memory and psychomotor disorders in people with HIV, whose full assessment is 12 points and with a cutoff value used in this research <10 points for high risk of dementia. This criterion is considered the gold standard for targeting the dementia field. Thus, patients who scored less than 10 after the application of this instrument were classified as 1.0 (high risk), and those who obtained a score greater than 10 were rated 2.0 (low risk).

Regarding the database (information from the widget files on Orange): 294 instances, the lines being registered patients, 58 features (columns), one target (Dementia Field) and four meta-attributes (reason\_for\_admission, CV\_start, TCD4\_start, current\_TCD4). In the database, there are 176 cases with high risk for dementia and 118 low risks for lower

dementia. Using the Orange tool, we selected the Dementia field as a target in the widget file as shown here in Figure 8:

Columns (Double click to edit)				
	Name	Type	Role	Values
25	<b>Memory</b>	<b>N numeric</b>	<b>feature</b>	
26	<b>Constructive S...</b>	<b>N numeric</b>	<b>feature</b>	
27	<b>Sum_risk</b>	<b>N numeric</b>	<b>skip</b>	
28	<b>Dementia</b>	<b>C categorical</b>	<b>target</b>	<b>1.0, 2.0</b>
29	<b>SupSocEmocional</b>	<b>N numeric</b>	<b>feature</b>	

Figure 8. Dementia field.

The target variable is the one that defines the event record with high-risk dementia or not. To determine the risk for dementia, we analyzed the main symptoms or most recurrent characteristics presented in the decision tree generated by Random Forest. That is, if a particular case has the same characteristics identified with dementia equal to 1, it may indicate that the patient is at high risk of dementia.

Figure 9 shows the execution process of each Machine Learning algorithm in Step 2 of the hybrid model.

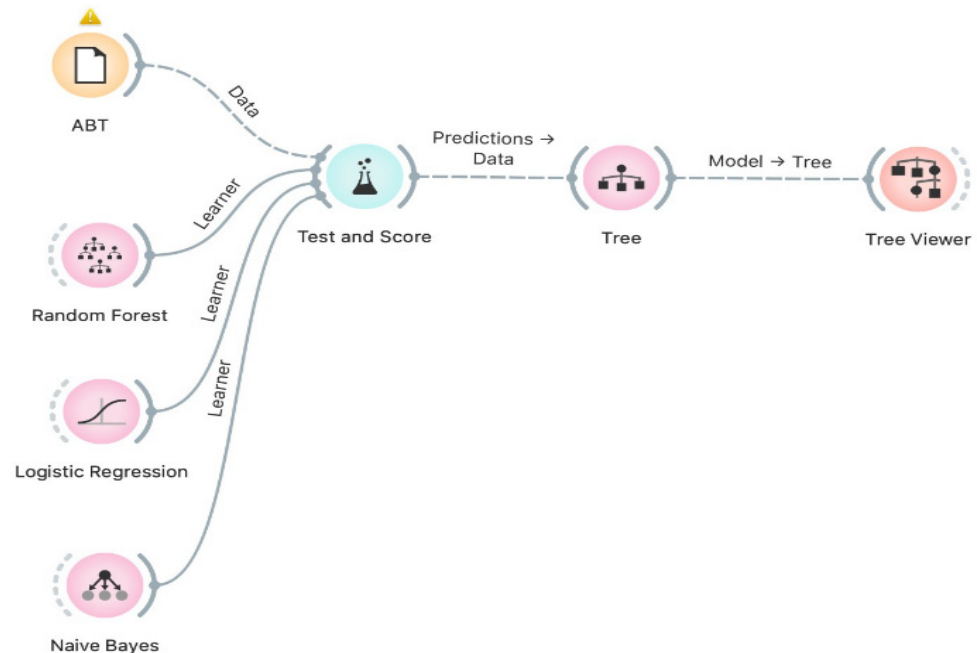


Figure 9. Application of Machine Learning algorithms.

Step 2 of the hybrid model classifies the main characteristics of cognitive and motor dementia, the object of the study. Step 3 is based on the features identified in the previous step. An online questionnaire was designed to contain the main characteristics, and their respective qualifiers (mild/moderate; severe; complete), and criteria (A, B, C) based on the International Classification of Functioning, Disability and Health developed by the World Health Organization (2001), which aims to standardize the description of health to its clinical situation [62].

The creation of this form contains the main characteristics classified by the best performing algorithm, Random Forest, for application to decisionmakers. Due to the mandatory social isolation established during the COVID-19 pandemic, the questionnaire was structured in Google Forms and sent along with the Informed Consent Form (FICF) by email to specialized healthcare professionals with clinical experience scientific research in infectious diseases. These professionals were selected through the Lattes Platform, explaining the purpose of the work, the benefits, and safety when participating in the research. The questionnaire aims to obtain information for the formulation of criteria that facilitate the diagnosis of dementia in HIV-infected individuals. The professionals used their free will to assess according to each one's experience. Some points of facilities and difficulties were addressed with the aim of ordering which characteristics are considered most important in the process for assessing high risk for HAD. Features have multiple choices based on the degree of difficulty defined by the ICF.

The questionnaire addressed the characteristics of patients at high risk for HAD so that the professional, when answering the questions and carrying out their assessment, would qualify the questions from three alternatives. In addition to collecting data that meet the concept, it is essential to obtain them so that the necessary treatment can be applied later to test the hypotheses. Therefore, it is essential to pay attention to the design of the collection instrument, with the type of information it provides and the type of analysis that can be conducted after obtaining the answers on the form. With the collection of useful or relevant data to test the hypotheses, the analysis model was confronted with the collected data, considering the requirements to put it into practice.

Finally, in the fourth step, the answers to the questionnaires were submitted to the ZAPROS-III-i method, which was implemented in the ARANAÚ tool, and then the Verbal Decision Analysis (AVD) was performed to define the order of the main HAD characteristics. These characteristics were used in the proposed model.

Thus, the ZAPROS-III-i places these characteristics in the decisionmaker's preference and constructs a protocol structured in Machine Learning. Notably, the ZAPROS-III-i method is implemented in the multiplatform Aranaú tool. It is developed in Java, in a ".Jar" format with portable code characteristics. Furthermore, this can be compiled and executed on Operating Systems without the need to change the code [63].

Step 4 produces a list of critical features of preference by the decisionmaker to support the diagnosis.

### 3. Results

This section gives the results obtained from the hybrid model using Machine Learning with the Random Forest algorithm, which was identified as the best performance with an accuracy of 87.6%, compared to Logistic Regression 83.9% and Naive Bayes 83.5%.

Random Forest showed the best performance of the model; the precision indicates the amount of correct choices among all the class classifications for a high risk of dementia. Recall indicates one of the high-risk situations, and F1 reveals the harmonic mean between Precision and Recall, as seen in Figure 10:

Evaluation Results									
Model	Train time [s]	Test time [s]	AUC	CA	F1	Precision	Recall	LogLoss	Specificity
Random Forest	1.830	0.263	0.876	0.827	0.820	0.837	0.827	0.492	0.766
Naive Bayes	0.191	0.053	0.835	0.731	0.734	0.745	0.731	1.017	0.742
Logistic Regression	3.581	0.115	0.839	0.789	0.789	0.790	0.789	0.869	0.775

Figure 10. Test and score measures.

Figure 11 demonstrates the Confusion Matrix generated in the Orange tool, revealing the number of instances correctly/incorrectly classified by Random Forest, which also stood out in relation to the other algorithms trained and tested in the ABT data. Random Forest classified with an accuracy of 94.9% of the data as high risk of HAD (1.0).

		Predicted		$\Sigma$
		1.0	2.0	
Actual	1.0	94.9 %	5.1 %	176
	2.0	35.6 %	64.4 %	118
$\Sigma$		209	85	294

Figure 11. Classification with Random Forest.

The result obtained by Random Forest showed that the main characteristics for a high risk of dementia in people with HIV were depression, interpersonal sensitivity, age, phobic anxiety, viral load, and constructive capacity.

Criteria were based on the International Classification of Functioning, Disability and Health (ICF) for the six characteristics identified in the Orange Canvas. The characteristics of depression, interpersonal sensitivity and phobic anxiety were placed in criterion A, which corresponds to emotional functions, interactions, and interpersonal relationships that present a barrier or difficulty. Constructive capacity was placed in criterion B regarding psych motivating functions, learning, and application of knowledge that present an obstacle or difficulty. Finally, age and viral load were placed in criterion C, which shows that the immune response presents a barrier or difficulty.

According to the ICF manual [62], the assessment qualifiers specify the extent to which the functionality of a given feature is compromised. For this research, qualifiers of the Mild/Moderate type present from 0% to 49% of neurological impairment caused by HIV; Severe from 50% to 95%; and Complete identifies that the interference of the infection in cognition and behavior can compromise from 96% to 100% of the person living with the AIDS virus.

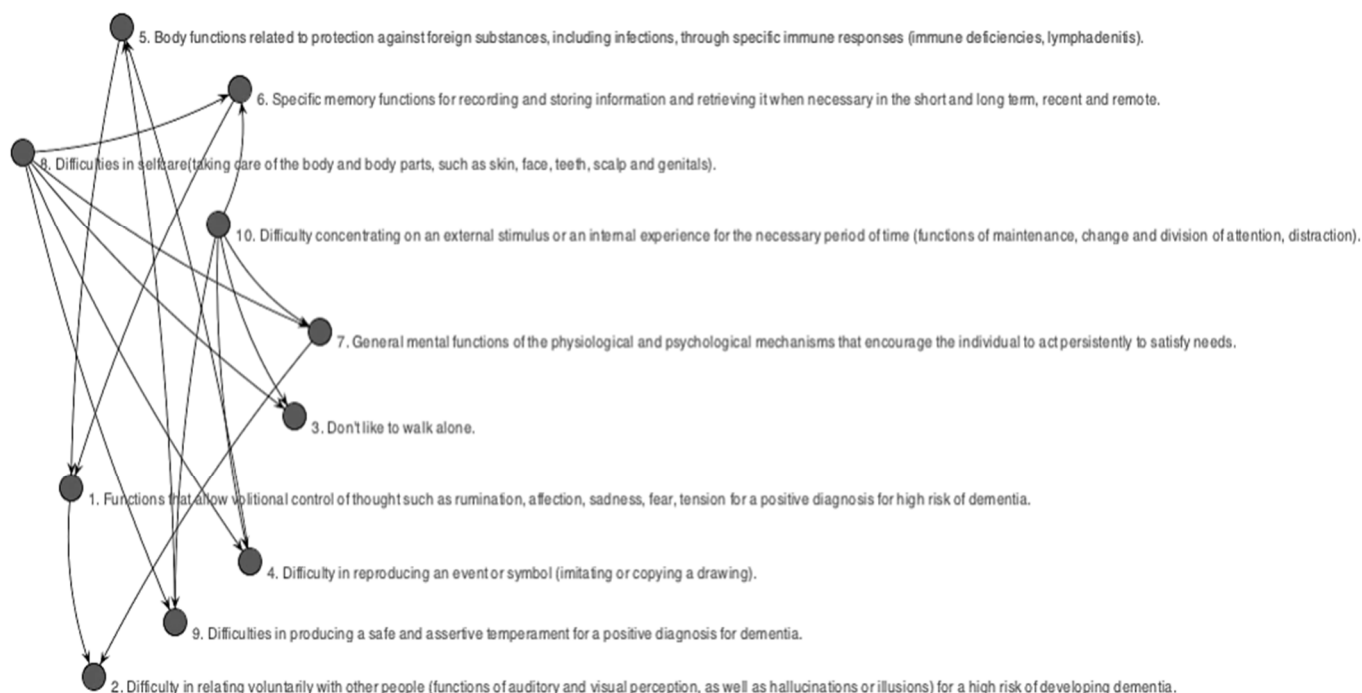
The questionnaire was analyzed by 40 health professionals; most (37.5%) were aged between 31 and 40 years old, 56.3% were women, 22.9% were physicians, 45.8% had specialist qualifications, 29.2% had an MSc, and 2.5% had PhD degrees related to infectious diseases. In addition, the items were based on the main characteristics identified by the Random Forest algorithm in cases of older people with HIV. We verified the relationship in the ICF for each of the characteristics according to functionality and disability. Finally, we have added classification options to indicate their extent or magnitude.

The results of the questionnaire can be seen in Table 1:

After the Elicitation of Preferences, each professional's preferences in the research generated the results shown in Table 1, which shows the search of the ideal result that would significantly affect the detection of dementia in people with HIV, which were compared with the results obtained through decision trees. Then, the data with the values of criteria and alternatives acquired in the research were loaded in Figure 12. The ARANAÚ tool was used to define the order of preference of the main characteristics, which were established by the ZAPROS-III-i method and their qualifiers most likely to determine the definitive diagnosis. The ZAPROS-III-i compared the Quality Variations (QV), the minimum distances of two alternatives, and the Formal Index of Quality (FIQ), which shows the minimum number of qualifiers pairs the decisionmaker also compared.

**Table 1.** Main characteristics of high risk of dementia according to decisionmakers.

Question	Criteria	Alternatives		
		Mild/Moderate	Severe	Complete
1. Functions that allow volitional control of thought such as rumination, affection, sadness, fear, tension for a positive diagnosis for high risk of dementia.	A	5	17	18
	B	16	20	4
	C	15	16	9
2. Difficulty in relating voluntarily with other people (functions of auditory and visual perception and hallucinations or illusions) for a high risk of developing dementia.	A	5	16	19
	B	18	19	3
	C	21	14	5
3. I do not like to walk alone.	A	3	18	19
	B	7	17	16
	C	19	15	6
4. Difficulty in reproducing an event or symbol (imitating or copying a drawing).	A	7	16	17
	B	3	15	22
	C	19	16	5
5. Body functions related to protection against foreign substances, including infections, through specific immune responses (immune deficiencies, lymphadenitis).	A	6	16	18
	B	11	19	10
	C	4	9	27
6. Specific memory functions for recording and storing information and retrieving it, when necessary, in the short and long term, recent and remote.	A	6	14	20
	B	14	18	8
	C	12	14	14
7. General mental functions of the physiological and psychological mechanisms that encourage the individual to act persistently to satisfy needs.	A	3	16	21
	B	7	17	16
	C	17	17	6
8. Difficulties in self-care (taking care of the body and body parts, such as skin, face, teeth, scalp and genitals).	A	6	17	17
	B	3	16	21
	C	13	19	8
9. Difficulties in producing a safe and assertive temperament for a positive diagnosis for dementia.	A	1	18	21
	B	5	17	18
	C	24	9	7
10. Difficulty concentrating on an external stimulus or an internal experience for the necessary period (functions of maintenance, change and division of attention, distraction).	A	6	14	20
	B	5	13	22
	C	12	20	8



**Figure 12.** Order of preference of the main characteristics for dementia in people with HIV.



These data enabled the combination of vectors formed by “criteria × alternatives,” that is, “rows × columns” of Table 1. The graphical representation in Figure 12 shows that the decision-makers’ most preferable characteristics have more outgoing arcs than incoming arcs.

Thus, the combination of the classifications resulted in the final order of preferences of the characteristics in descending order, namely:

10—Difficulty in voluntarily relating to other people (auditory and visual perception functions, as well as hallucinations or illusions) due to a high risk of developing dementia;

9—Difficulties in producing a confident and assertive temperament for a positive diagnosis of dementia;

8—Difficulty in reproducing an event or symbol (to imitate or copy a drawing);

7—Functions that allow volitional thought control such as rumination, affection, sadness, fear, and tension for a positive diagnosis of high risk of dementia;

6—I do not like to walk alone;

5—General mental functions of physiological and psychological mechanisms that encourage the individual to act persistently to satisfy needs;

4—Difficulty in focusing on an external stimulus or an internal experience for the necessary period of time (maintenance functions, change and division of attention, and distraction);

3—Difficulties in self-care (taking care of the body and parts of the body, such as skin, face, teeth, scalp, and genitals);

2—Specific memory functions to record and store information and retrieve it when needed in the short and long term, recent and remote. General mental functions of physiological and psychological mechanisms encourage the individual to act persistently to meet needs;

1—Body functions related to protection against foreign substances, including infections, through specific immune responses (immunodeficiencies, lymphadenitis).

Figure 12 above shows the graphic representation of the final order of preference of the main alternatives established by the ZAPROS-III-i method to positively diagnose the high risk of dementia in people with HIV.

#### 4. Discussion

This section discusses the characteristics of the most accurate algorithm, Random Forest (87.6%), which was more determinant for the high risk of dementia and organized in descending order respected the classification result obtained by the VDA and by the ZAPROS III-i methods.

When aided by AI computational algorithms that can show possible neurological changes caused by the AIDS virus, the health professional can optimize the clinical reasoning that leads to an early and accurate diagnosis of any possible Neurocognitive Disorders. Thus, health professionals can offer a safe, agile, and effective therapeutic follow-up to search for improved quality of life and a good prognosis for HIV. Patients with untreated HIV have an average life expectancy of less than six months without dementia [64].

According to the results of this study, the immune response was considered the preferable characteristic for a high risk of dementia in individuals with HIV. Despite the advent of antiretroviral therapy capable of delaying immunodeficiency, people undergoing treatment are susceptible to neurological changes because the AIDS virus is neurotropic. In addition, other factors, such as viral resistance, drug toxicity, and the need for a high level of treatment adherence, are considered the main difficulties in reducing the viral load, affecting the patient’s cognitive and behavioral aspects [64].

According to the Centers for Disease Control and Prevention (CDC), the most susceptible determinant to HIV-Associated Neurocognitive Diseases (HAND) is the degree of immunosuppression of the infected individual, whose values are the nadir of TCD4 Lymphocytes (LCTD4) + < 350 cells/mm<sup>3</sup> or LTCD4 + current < 350 cells/mm<sup>3</sup> [65,66].

Neuroinflammation can trigger different cognitive and motor declines. In addition, these are classified according to neuropsychological assessments and the impact of the disease on the performance of daily activities, clinical diagnosis, and analysis of cerebrospinal fluid and plasma. These methods reveal a more detailed knowledge of CNS changes during HIV infection and differentiate HAND from other CNS diseases [67–69].

In addition to monitoring the medication adherence, health professionals need to know and observe the subtle signs of psychomotor changes reported and manifested by patients. When identified quickly, they can be reversed by HAART. Studies show that follow-ups based on a cognitive-behavioral approach generate confidence for patients to adhere to interventions guided by the professional, such as the continuous use of antiretrovirals, searching for information about their clinical condition, and improving self-perception of the infected person and their social relationships [70].

Studies regarding social and emotional factors show that infected individuals, when they have feelings of loneliness, negative thoughts about living with HIV, and low social support from family and friends, imply that the progression of the disease makes them more resistant to the guidelines and makes it more difficult to carry out the treatment correctly. These social and emotional factors also hinder them in performing daily activities, even in the early stages of the disease. In addition, they enhance negative thoughts due to depression related to infection and anxiety and the thoughts of death [71–73].

Monitoring stress-induced behavioral changes is essential due to the progression of chronic neuroinflammation, neuroendocrine changes, and anomalous neurotransmission signaling. Following, HIV-positive individuals with psychiatric disorders such as depression, interpersonal sensitivity, and phobic anxiety tend to present increased levels of inflammatory cytokines (TNF-alpha, IL-1beta, and IL-6), alterations in the glutamatergic and monoaminergic pathways in the cortical and subcortical regions of the brain. When such symptoms go underdiagnosed, they can cause irreversible neurocognitive damage such as dementia [74].

The high risk of dementia is clinically significant when it affects the functioning of daily activities. Infected individuals with mild cognitive impairments can have occupational problems, even in the early stages of the disease. These individuals have difficulties in the learning process, and this can be identified due to low performance at work, attention, and memory, especially in performing instrumental activities of daily living, causing difficulties in self-care, such as caring for the body and its parts, such as skin, face, teeth, scalp, and genitals [75–77].

It is essential to emphasize that the analysis of cognitive and behavioral conditions related to HIV is still difficult and requires great attention from professionals and researchers. These conditions represent factors that influence a high risk of dementia, and therefore mechanisms for their delay and control by the health professional must be performed to control and prevent the progression of dementia caused by neuroinflammation.

## 5. Conclusions

Identifying mental disorders is a complex process, given the subjectivities experienced by each individual, and when dealing with the immunosuppressed, it becomes something broader. The presence of HIV, in addition to the neurological impairment, can impact the psychological and social factors of the infected person, as it causes a mixture of feelings of fear, panic, suicidal tendencies, anxiety, and social isolation, which can interfere with adherence to antiretroviral therapy and the individual's quality of life.

The hybrid approach with the Machine Learning algorithm, Random Forest, combined with Verbal Decision Analysis, accurately obtained the characteristic signs of dementia. The application of the hybrid approach proved to be efficient and revealed ten characteristics that influence the prediction of a high risk of dementia in people with HIV.

When analyzed by health professionals, the characteristic “immune factor” came first, with functions that allow volitional control of thought, such as rumination, affection, sadness, fear, and tension, for a positive high-risk diagnosis with insanity in the last

position. The importance of combining the algorithm and health professional's experience to obtain the diagnostic framework is essential, given the need for clinical experience and subjective responses.

A future work will apply the hybrid model to individuals with HIV/AIDS over eighteen years old to expand the universe of classifications and the risk of developing dementia in the younger population. In addition, we intend to incorporate other AI algorithms to classify patients and develop clinical protocols, using methodologies such as Bayesian networks and cognitive maps. Finally, another future work could apply only the Verbal Decision Analysis (VDA) of a decision to the bases presented.

As a limitation of the study, the participation of health professionals was affected as the online submission of the questionnaire coincided with the high incidence of people infected and hospitalized for complications caused by SARS-COV2 during the COVID-19 pandemic.

**Author Contributions:** Conceptualization, L.I.C.C.P. and M.L.D.P.; methodology, P.R.P. and L.C.N.; software, P.G.C.D.P. and E.C.d.A.; validation: L.I.C.C.P.; formal analysis, R.H.F.; investigation, L.I.C.C.P.; resources, R.H.F.; data curation, L.I.C.C.P. and W.C.d.A.; writing—original draft preparation, L.I.C.C.P.; writing—L.I.C.C.P. and P.R.P., Preview and editing, P.G.C.D.P.; visualization, P.G.C.D.P.; supervision, W.C.d.A.; project administration, M.L.D.P.; funding acquisition, R.H.F. All authors have read and agreed to the published version of the manuscript.

**Funding:** The authors Plácido Rogerio Pinheiro and Raimir Holanda Filho are grateful to National Council of Technological and Scientific Development (CNPq) via Grants #304272/2020-5 and 306389/2020-7.

**Institutional Review Board Statement:** Approval was obtained from the Research Ethics Committee of the Federal University of Ceará, Brazil, respecting the ethical precepts of nonmaleficence, beneficence, autonomy, and justice for its participants, under No. 3038967. The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the or Ethics Committee of São José Hospital for Infectious Diseases—HSJ/Health Department, CAAE: 98103217.0.3001.5044, 26 November 2018.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Li, Y.; Jiang, J.; Shen, T.; Wu, P.; Zuo, C. Radiomics features as predictors to distinguish fast and slow progression of Mild Cognitive Impairment to Alzheimer's disease. In Proceedings of the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Honolulu, HI, USA, 18–21 July 2018; pp. 127–130. [[CrossRef](#)]
2. Qi, Y.; Xu, M.; Wang, W.; Wang, Y.Y.; Liu, J.J.; Ren, H.X.; Li, H.J. Early prediction of putamen imaging features in HIV-associated neurocognitive impairment syndrome. *BMC Neurol.* **2021**, *21*, 1–10. [[CrossRef](#)] [[PubMed](#)]
3. Sacktor, N. Changing clinical phenotypes of HIV-associated neurocognitive disorders. *J. Neurovirology* **2018**, *24*, 141–145. [[CrossRef](#)]
4. Serrano-Villar, A.R.S.; Gutierrez, F.; Mirall, E.S.C. HIV as a chronic disease: Evaluation and management of non-AIDS defining, Management of HIV-Associated Disease in HIV. *Open Forum Infect. Dis.* **2016**, *3*, 1–15.
5. Saylor, D.; Dickens, A.M.; Sacktor, N.; Haughey, N.; Slusher, B.; Pletnikov, M.; McArthur, J.C. HIV-associated neurocognitive disorder—pathogenesis and prospects for treatment. *Nat. Rev. Neurol.* **2016**, *12*, 234–248. [[CrossRef](#)] [[PubMed](#)]
6. Rom, S.; Gajghate, S.; Winfield, M.; Reichenbach, N.L.; Persidsky, Y. Combination of HIV-1 and Diabetes Enhances Blood Brain Barrier Injury via Effects on Brain Endothelium and Pericytes. *Int. J. Mol. Sci.* **2020**, *21*, 4663. [[CrossRef](#)]
7. Bini, S.A. Artificial intelligence, machine learning, deep learning, and cognitive computing: What do these terms mean, and how will they impact health care? *J. Arthroplast.* **2018**, *33*, 2358–2361. [[CrossRef](#)] [[PubMed](#)]
8. Tamanini, I. Hybrid Approaches of Verbal Decision Analysis Methods. Ph.D. Thesis, University of Fortaleza, Fortaleza, Brazil, 2015.
9. Hjalmarsson, V. Machine Learning and Multi-Criteria Decision Analysis in Healthcare: A Comparison of Machine Learning. Ph.D. Thesis, Mid Sweden University, Sundsvall, Sweden, 27 June 2018.
10. de Castro, A.K.A.; Pinheiro, P.R.; Dantas Pinheiro, M.C.; Tamanini, I. Towards the applied hybrid model in decision making: A neuropsychological diagnosis of alzheimer's disease study case. *Int. J. Comput. Intell. Syst.* **2011**, *4*, 89–99. [[CrossRef](#)]

11. Vazquezl, M.Y.L.; Peñafiel, L.A.B.; Muñoz, S.X.S.; Martinez, M.A.Q. A framework for selecting machine learning models using TOPSIS. In *Advances in Artificial Intelligence, Software and Systems Engineering, Proceedings of the AHFE 2020 Virtual Conferences on Software and Systems Engineering, and Artificial Intelligence and Social Computing, Florida, FL, USA, 16–20 July 2020*; Springer Nature: Basingstoke, UK, 2020; Volume 1213, p. 119.
12. Mohdiwale, S.; Sahu, M.; Sinha, G.R.; Bajaj, V. Automated cognitive workload assessment using logical teaching learning-based optimization and promethee multi-criteria decision making approach. *IEEE Sens. J.* **2020**, *20*, 13629–13637. [[CrossRef](#)]
13. Juneja, A.; Juneja, S.; Kaur, S.; Kumar, V. Predicting diabetes mellitus with machine learning techniques using multi-criteria decision making. *Int. J. Inf. Retr. Res.* **2021**, *11*, 38–52. [[CrossRef](#)]
14. Menezes, A.C.; Pinheiro, P.R.; Pinheiro, M.C.D.; Cavalcante, T.P. Towards the Applied Hybrid Model In Decision Making: Support the Early Diagnosis of Type 2 Diabetes. In *International Conference on Information Computing and Applications*; Springer: Berlin/Heidelberg, Germany, 2012; Volume 7473, pp. 648–655.
15. Pinheiro, P.; Tamanini, I.; Pinheiro, M.C.D.; Albuquerque, V.H.C. Evaluation of the Alzheimer’s disease clinical stages under the optics of hybrid approaches in Verbal decision analysis. *Telemat. Inform.* **2018**, *35*, 776–789. [[CrossRef](#)]
16. Silva, T.; Pinheiro, P.; Poggi, M. A more human-like portfolio optimization approach. *Eur. J. Oper. Res.* **2017**, *256*, 252–260. [[CrossRef](#)]
17. Gurgel, S.N. Vulnerability to HIV/AIDS in the Elderly: A Comparative Study. Master’s Thesis, Federal University of Paraíba, João Pessoa, Brazil, 2014.
18. UNAIDS. Joint united nations program on HIV/AIDS. Communities at the centre: Defending rights, breaking barriers, reaching people with HIV services. *Glob. AIDS* **2019**. Available online: [https://www.unaids.org/sites/default/files/media\\_asset/2019-global-AIDS-update\\_en.pdf](https://www.unaids.org/sites/default/files/media_asset/2019-global-AIDS-update_en.pdf) (accessed on 28 October 2021).
19. Livingston, G.; Huntley, J.; Sommerlad, A.; Ames, D.; Ballard, C.; Banerjee, S.; Brayne, C.; Burns, A.; Cohen-Mansfield, J.; Cooper, C.; et al. Dementia prevention, intervention, and care: 2020 report of the lancet commission. *Lancet* **2020**, *396*, 413–446. [[CrossRef](#)]
20. Deus, D.M.V.d.; Possas, C.C. Updates on cortical dementia in HIV-infected individuals. In *Humanae: Controversial Issues in The Contemporary World, Recife, PE*; 2018; Volume 12, pp. 1–11. Available online: <https://revistas.esuda.edu.br/index.php/humanae/article/view/621/193> (accessed on 28 October 2021).
21. Antinori, A.; Arendt, G.; Becker, J.T.; Brew, B.J.; Byrd, D.A.; Cherner, M.; Clifford, D.B.; Cinque, P.; Epstein, L.G.; Goodkin, K.; et al. Updated research nosology for HIV-associated neurocognitive disorders. *Neurology* **2007**, *69*, 1789–1799. [[CrossRef](#)] [[PubMed](#)]
22. Dubé, B.; Benton, T.; Cruess, D.E.; Evans, D.L. Neuropsychiatric manifestations of HIV infection and AIDS. *J. Psychiatry Neurosci.* **2005**, *30*, 237–246. [[PubMed](#)]
23. Baker, L.M.; Paul, R.H.; Heaps-Woodruff, J.M.; Chang, J.Y.; Ortega, M.; Margolin, Z.; Usher, C.; Basco, B.; Cooley, S.; Ances, B.M. The effect of central nervous system penetration effectiveness of highly active antiretroviral therapy on neuropsychological performance and neuroimaging in HIV infected individuals. *J. Neuroimmune Pharmacol.* **2015**, *10*, 487–492. [[CrossRef](#)] [[PubMed](#)]
24. Brew, B.J.; Rosenblum, M.; Price, R.W. AIDS dementia complex and primary HIV brain infection. *J. Neuroimmunol.* **1988**, *20*, 133–140. [[CrossRef](#)]
25. Navia, B.A.; Jordan, B.D.; Price, R.W. AIDS dementia complex I: Clinical features. *Ann. Neurol.* **1986**, *19*, 517–524. [[CrossRef](#)]
26. Sanford, R.; Fellows, L.; Ances, B.M.; Collins, D.L. Association of brain structure changes and cognitive function with combination antiretroviral therapy in HIV-positive individuals. *JAMA Neurol.* **2018**, *75*, 72–79. [[CrossRef](#)]
27. de Almeida, S.M. Cerebrospinal fluid analysis in the HIV infection and compartmentalization of HIV in the central nervous system. *Arq. Neuro-Psiquiatr.* **2015**, *73*, 624–629. [[CrossRef](#)]
28. Underwood, J.; Robertson, K.R.; Winston, A. Could antiretroviral neurotoxicity play a role in the pathogenesis of cognitive impairment in treated HIV disease? *AIDS* **2015**, *29*, 253–261. [[CrossRef](#)] [[PubMed](#)]
29. Simioni, S.; Cavassini, M.; Annoni, J.-M.; Abraham, A.R.; Bourquin, I.; Schiffer, V.; Calmy, A.; Chave, J.-P.; Giacobini, E.; Hirschel, B.; et al. Cognitive dysfunction in HIV patients despite long-standing suppression of viremia. *AIDS* **2010**, *24*, 1243–1250. [[CrossRef](#)] [[PubMed](#)]
30. Working Group of the America Academy of Neurology AIDS Task Force. Nomenclature and research case definitions for neurologic manifestations of human immunodeficiency virus type 1(HIV-1) infection. *Neurology* **1991**, *41*, 778–785. [[CrossRef](#)] [[PubMed](#)]
31. Sartori, G.P.; Domínguez, C.I.; Rodríguez, V.G.F.; Dansilio, S.; Presentado, J.C.M. Trastornos neurocognitivos en pacientes VIH positivos Datos preliminares de una cohorte prospectiva uruguaya. *Rev. Médica Del Urug.* **2019**, *35*, 171–180. [[CrossRef](#)]
32. Breiman, L.; Friedman, J.; Olshen, R.; Stone, C. Classification and regression trees. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2011**, *1*, 14–23.
33. Scornet, E.; Biau, G.; Vert, J. Consistency of random forests. *Ann. Stat. Inst. Math. Stat.* **2015**, *43*, 1716–1741. [[CrossRef](#)]
34. Breiman, L. Bagging predictors. *Mach. Learn.* **1996**, *24*, 123–140. [[CrossRef](#)]
35. Beğenilmiş, E.; Uskudarli, S. Classification of organized behavior of sets of tweets using supervised learning methods. In *Proceedings of the 8th International Conference on Web Intelligence, Mining and Semantics (WIMS '18)*, Association for Computing Machinery, Novi Sad, Serbia, 25–27 June 2018; Volume 36, pp. 1–9. [[CrossRef](#)]
36. Pant, A. Introduction to Logistic Regression. Average. Towards Data Science. 2019. Available online: <https://towardsdatascience.com/introduction-to-logistic-regression-66248243c148> (accessed on 21 February 2021).

37. Andrade, E.; Portela, S.; Pinheiro, P.R.; Nunes, L.C.; Simão Filho, M.; Costa, W.S.; Pinheiro, M.C.D. A Protocol for the Diagnosis of Autism Spectrum Disorder Structured in Machine Learning and Verbal Decision Analysis. *Comput. Math. Methods Med.* **2021**, *2021*, 1–15. [[CrossRef](#)] [[PubMed](#)]
38. Han, D.; Ma, L.; Yu, C. Financial prediction: Application of logistic regression with factor analysis. In Proceedings of the 2008 4th International Conference on Wireless Communications, Networks and Mobile Computing, Dalian, China, 12–14 October 2008; pp. 1–4. [[CrossRef](#)]
39. Strzelecka, A.; Kurdyś-Kujawska, A.; Zawadzka, D. Applying logistic regression models to assess household financial decisions regarding debt. *Procedia Comput. Sci.* **2020**, *176*, 3418–3427. [[CrossRef](#)]
40. AVT—Vidhya Analytics Team. Simple Guide to Logistic Regression in R and Python. 2015. Available online: <https://www.analyticsvidhya.com/blog/2015/11/beginners-guide-on-logistic-regression-in-r/> (accessed on 15 February 2021).
41. Andrew, N.G. CS229 Lecture Notes. Stanford University, 2012. Available online: <http://cs229.stanford.edu/syllabus.html>. (accessed on 15 February 2021).
42. Bacaer, N. Verhulst and the logistic equation. In *A Short History of Mathematical Population Dynamics*, 2011th ed.; Springer: Berlin/Heidelberg, Germany, 2010; p. 172. [[CrossRef](#)]
43. Shmueli, G. To explain or predict? *Stat. Sci.* **2010**, *25*, 289–310. [[CrossRef](#)]
44. Hastie, T.; Tibshirani, R.; Friedman, J. The elements of statistical learning: Data mining, inference, and prediction. *Math. Intell.* **2005**, *27*, 83–85. [[CrossRef](#)]
45. Hosmer, D.W.; Lemeshow, S. Introduction to the logistic regression model. *Appl. Logist. Regression.* **2000**, *2*, 1–30. [[CrossRef](#)]
46. Motrenko, A.; Strijov, V.; Weber, G.W. Sample size determination for logistic regression. *J. Comput. Appl. Math.* **2014**, *255*, 743–752. [[CrossRef](#)]
47. Liu, L.; Özsü, M.T. Bayes classifier. In *Encyclopedia of Database Systems*; Springer: Boston, MA, USA, 2009. [[CrossRef](#)]
48. Lund’s Research Binomial. Binomial Logistic Regression Using SPSS Statistics. 2018. Available online: <https://statistics.laerd.com/spss-tutorials/binomial-logistic-regression-using-spss-statistics.php> (accessed on 15 February 2021).
49. Krause, P.J. Learning Probabilistic Networks, Philips research laboratories. *Knowl. Eng. Rev.* **1999**, *13*, 321–351. [[CrossRef](#)]
50. Larichev, O.I.; Moshkovich, H.M. *Verbal Decision Analysis for Unstructured Problems*; Kluwer Academic Press: Boston, MA, USA, 1997.
51. Tamanini, I.; De Castro, A.K.; Pinheiro, P.R.; Pinheiro, M.C.D. Verbal decision analysis applied on the optimization of alzheimer’s disease diagnosis: A case study based on neuroimaging. *Adv. Exp. Med. Biol.* **2011**, *696*, 555–564. [[CrossRef](#)] [[PubMed](#)]
52. Tamanini, I.; Pinheiro, P.R. Reducing incomparability in multicriteria decision analysis: An extension of the zapros methods. *Pesqui. Oper.* **2011**, *31*, 251–270. [[CrossRef](#)]
53. Tamanini, I.; Pinheiro, P.R. Challenging the incomparability problem: An approach methodology based on zapros. In *International Conference on Modelling, Computation and Optimization in Information Systems and Management Sciences*; Springer: Berlin/Heidelberg, Germany, 2001; pp. 338–347.
54. Larichev, O. Ranking multicriteria alternatives: The method ZAPROS III. *Eur. J. Oper. Res.* **2001**, *131*, 550–558. [[CrossRef](#)]
55. Moshkovich, H.; Mechitov, A.; Olson, D. Ordinal judgments in multiattribute decision analysis. *Eur. J. Oper. Res.* **2002**, *137*, 625–641. [[CrossRef](#)]
56. Tamanini, I.; Pinheiro, P.R.; dos Santos, C.N. A hybrid approach to verbal decision analysis and machine learning. In *International Conference on Rough Sets and Current Trends in Computing*; Springer: Berlin/Heidelberg, Germany, 2012.
57. Castro, A.K.A.; Pinheiro, P.R.; Pinheiro, M.C.D. A hybrid model for aiding in decision making for the neuropsychological diagnosis of Alzheimer’s disease, RSCTC ’08. In Proceedings of the 6th International Conference on Rough Sets and Current Trends in Computing, Akron, OH, USA, 23–25 October 2008; pp. 495–504. [[CrossRef](#)]
58. Thabtah, F.; Peebles, D. A new machine learning model based on induction of rules for autism detection. *Health Inform. J.* **2019**, *26*, 264–286. [[CrossRef](#)] [[PubMed](#)]
59. Pinheiro, L.I.C.C.; Pereira, M.L.D.; Fernandez, M.P.; Filho, F.M.V.; de Abreu, W.J.C.P.; Pinheiro, P.G.C.D. Application of data mining algorithms for dementia in people with HIV/AIDS. *Comput. Math. Methods Med.* **2021**, *2021*, 1–8. [[CrossRef](#)] [[PubMed](#)]
60. Couronné, R.; Probst, P.; Boulesteix, A.-L. Random forest versus logistic regression: A large-scale benchmark experiment. *BMC Bioinform.* **2018**, *19*, 1–14. [[CrossRef](#)] [[PubMed](#)]
61. Demšar, J.; Curk, T.; Erjavec, A.; Gorup, Č.; Hočevar, T.; Milutinovič, M.; Zupan, B. Orange: Data mining toolbox in python. *J. Mach. Learn. Res.* **2013**, *14*, 2349–2353.
62. CIF, 2001. International Classification of Functioning, Disability and Health. Available online: [https://psychiatr.ru/download/1313?view=name=CF\\_18.pdf](https://psychiatr.ru/download/1313?view=name=CF_18.pdf) (accessed on 11 April 2021).
63. Tamanini, I.; Pinheiro, P.R.; Carvalho, A.L. *Aranaú Software: A New Tool of the Verbal Decision Analysis*; Technical Report; University of Fortaleza: Fortaleza, Brazil, 2007.
64. Angelim, R.C.D.M.; Brandão, B.M.G.D.M.; Marques, S.C.; de Oliveira, D.C.; Abrão, F.M.D.S. Representações e práticas de cuidado de profissionais de saúde às pessoas com HIV. *Rev. Esc. Enferm. USP* **2019**, *53*, e03478. [[CrossRef](#)]
65. Scutari, R.; Alteri, C.; Perno, C.F.; Svicher, V.; Aquaro, S. The role of HIV infection in neurologic injury. *Brain Sci.* **2017**, *7*, 38. [[CrossRef](#)]

66. Ayele, B.A.; Amogne, W.; Gemechu, L. HIV-associated neurocognitive disorder and HIV-associated myelopathy in a patient with a preserved CD4, but high viral load—a rarely reported phenomenon: A case report and literature review. *BMC Infect. Dis.* **2020**, *20*, 1–6. [[CrossRef](#)]
67. Ceccarelli, G.; Borrazzo, C.; Lazzaro, A.; Innocenti, G.P.; Celani, L.; Cavallari, E.M.; Pinacchio, C.; Santinelli, L.; Mastroianni, C.M.; d'Éttore, G. Problems in diagnosing asymptomatic neurosyphilis in HIV-positive patients: A retrospective study. *Brain Sci.* **2019**, *9*, 278. [[CrossRef](#)]
68. Alves, V.S.; Leite-Aguiar, R.; da Silva, J.P.; Coutinho-Silva, R.; Savio, L.E.B. Purinergic signaling in infectious diseases of the central nervous system. *Brain Behav. Immun.* **2020**, *89*, 480–490. [[CrossRef](#)] [[PubMed](#)]
69. Nath, A. Neuroinfectious diseases. *JAMA Neurol.* **2015**, *72*, 143–144. [[CrossRef](#)] [[PubMed](#)]
70. Alford, K.; Vera, J.H. Cognitive impairment in people living with HIV in the ART era: A review. *Br. Med. Bull.* **2018**, *127*, 55–68. [[CrossRef](#)] [[PubMed](#)]
71. Masquillier, C.; Wouters, E.; Mortelmans, D.; Wyk, B.V. On the road to HIV/AIDS competence in the home: Building a health-friendly environment for people living with HIV/AIDS. *Int. J. Environ. Res. Public Health* **2015**, *12*, 3264–3292. [[CrossRef](#)]
72. Iyidogan, P.; Anderson, K.S. Current perspectives on HIV-1 antiretroviral drug resistance. *Viruses* **2014**, *6*, 4095–4139. [[CrossRef](#)] [[PubMed](#)]
73. Watkins, C.C.; Pieper, A.A.; Treisman, G.J. Safety considerations in drug treatment of depression in HIV-positive patients: An updated review. *Drug Saf.* **2011**, *34*, 623–639. [[CrossRef](#)] [[PubMed](#)]
74. De Francesco, D. Cognitive function and depression in HIV-positive individuals and matched controls. In Proceedings of the Presented at the National AIDS Treatment Advocacy Project Conference, Glasgow, Scotland, 23–26 October 2016; pp. 23–26.
75. Wärtl, I.; Kaufer, C.; Gerhauser, I.; Chhatbar, C.; Ghita, L.; Kalinke, U.; Loscher, W. Microglia have a protective role in the development of viral encephalitis-induced seizures and hippocampal damage. *Brain Behav. Immun.* **2018**, *74*, 186–204. [[CrossRef](#)]
76. Livingston, G.; Huntley, J.; Sommerlad, A.; Ames, D.; Ballard, C.; Banerjee, S.; Mukadam, N. Prevention, intervention and care for dementia: 2020 Lancet Commission report. *Lancet* **2020**, *396*, 443–446. [[CrossRef](#)]
77. Hipolito, R.L.; De Oliveira, D.C.; Da Costa, T.L.; Marques, S.C.; Pereira, E.R.; Gomes, A.M.T. Quality of life of people living with HIV/AIDS: Temporal, socio-demographic and perceived health relationship. *Rev. Latino-Am. Enferm.* **2017**, *25*, 1–10. [[CrossRef](#)] [[PubMed](#)]