Alzheimer is a neurodegenerative disease that consists of the continuous and gradual loss of neurons generating as a result memory loss and difficulty to speak coherently, among other side effects; in general, a patient with Alzheimer’s disease presents strong changes in behavior and personality [3], [4],[5],[6] [7]. The Alzheimer disease is the most common dementia; the principal Alzheimer risk factor is the age, the incidence of the Alzheimer disease (AD) duplicates each 5 years, in patients older than 65; also, 1275 new cases per each 100.000 people older than 65 are diagnosed every year [4] [5]. The use of new techniques and the application of rigorous methodologies becomes increasingly important. With this in mind, the proposed approach seeks to achieve the identification of patterns that allows to have a better diagnosis. According to different studies, the use of biomarkers and genetic markers is necessary to achieve a better classification of the patients. In other words, more suitable diagnosis and patient analysis can be performed using clinical features, biological, genetic and other risk factors [4],[5],[6]. The Alzheimer disease is highly influenced by genetic factors; in the past three decades, the main associated genes were mainly APP, PSEN1, PSEN2, APOE. Nowadays, the number of genes has increased and there are some databases including information about them on-line. Such databases are dedicated exclusively to the collection of genetic information related with AD. The late development onset of Alzheimer’s (LOAD) is associated to both genetic and environmental factors. A risk factor is the allele e4 of the apolipoprotein E (APOE) which is encoding an important protein in the cholesterol metabolism [33]. Some evidence shows that the APOE 4 modulates the metabolism and aggregation of the amyloid beta peptide. The APOE is the cholesterol apolipoprotein transporter, the most important in the brain [33]. There are three alleles of the APOE gene: e2, e3, and e4, which generate 6 different genotypes: e2/e2, e3/e3, e4/e4, e2/e3, e2/e4 and e3/e4. e3 is the most frequent allele (77%) and e2 is the least frequent (8%), e4 with frequency of 15% has a strong association with the development of neurodegenerative diseases [33]. APOE4 is considered a strong risk factor for Alzheimer type (LOAD); some of the implications may be associated with the production of A beta, A beta clearance, formation of fibrils and tangles, cholesterol homoeostasis, repair and synaptic plasticity and neuronal toxicity [33]. According to the latest report, the top ten major genes associated with AD are the APOE gene with its alleles e/2/3/4, BIN1, CLU, ABCA7, CR1, PICALM, MS4A6A, CD33, MS4A4E, CD2AP [39],[30]. However, this number may increase as new studies are developed.

The Genetic Institute of the National University of Colombia is performing association studies of Alzheimer disease (AD) and Parkinson disease (PD). There are samples from patients with neurodegenerative disease and psychiatric problems which have been taken for several years. The polymorphisms studied in colombian patients were selected according to some criteria such as identification and selection of HapMap database, genes that had previously been studied by the research group and for which there is some evidence; the APOE gene and those close to this gene (in a close region) represent a higher risk for the development of the disease. Some of the studied genes by the research group are CLU, PICALM, CR1, TOMM40, PVRL2, APOE, SORL1, CR1, GWA-14q32.13. However, in this research only APOE genotypes were included. The diagnosis of a patient is based on clinical experience, after analyzing the medical records and the neuropsychological tests such as the Mini mental state examination (MMSE), Yesavage test, Lawton and Brody scale.
The use of techniques based on computational intelligence to facilitate the diagnosis is becoming increasingly important [16]. In this study, Bayesian networks were selected as classification method. This method was applied to a case-control sample of colombian patients including individuals with Alzheimer disease as well as healthy individuals. This data set also included clinical variables such as age, gender, educational level, categorized according to the international standard classification of education CINE 1976, marital status, family history of dementia, APOE genotype and the diagnosis. It is important to mention that for the selection and identification of variables, it was necessary to use statistical techniques for analysis of independence between variables and diagnosis. Furthermore, the use of algorithms of association, such as the a priori algorithm, to identify the main rules of association were used to identify the most representative variables used to design a Bayesian classifier.

In addition, some rules related to the diagnosis of AD were obtained as follows: family history of dementia is related to AD diagnosis; diagnosis of patients older than 65 age and with a family history of dementia are related to AD. Female individuals with a family history of dementia are related to the diagnosis of AD. Another very important association found was that the diagnosis of an individual who has a family history of dementia and genotype APOE 34 is related to AD. Also, for those individuals who are married and have 34 APOE genotype their diagnosis is related to AD.


