Early Diagnosis of Alzheimer's Disease using Machine Learning Techniques
A Review Paper

Aunsia Khan and Muhammad Usman
Dept. of Computing, Shaheed Zulfikar Ali Bhutto Institute of Science and Technology (SZABIST), Islamabad, Pakistan

Keywords: Alzheimer's Disease, Machine Learning, Computer Aided Diagnosis, Pathologically Proven Data, Early Diagnosis, Class Imbalance.

Abstract: Alzheimer's, an irreparable brain disease, impairs thinking and memory while the aggregate mind size shrinks which at last prompts demise. Early diagnosis of AD is essential for the progress of more prevailing treatments. Machine learning (ML), a branch of artificial intelligence, employs a variety of probabilistic and optimization techniques that permits PCs to gain from vast and complex datasets. As a result, researchers focus on using machine learning frequently for diagnosis of early stages of AD. This paper presents a review, analysis and critical evaluation of the recent work done for the early detection of AD using ML techniques. Several methods achieved promising prediction accuracies, however they were evaluated on different pathologically unproven data sets from different imaging modalities making it difficult to make a fair comparison among them. Moreover, many other factors such as pre-processing, the number of important attributes for feature selection, class imbalance distinctively affect the assessment of the prediction accuracy. To overcome these limitations, a model is proposed which comprise of initial pre-processing step followed by imperative attributes selection and classification is achieved using association rule mining. Furthermore, this proposed model based approach gives the right direction for research in early diagnosis of AD and has the potential to distinguish AD from healthy controls.

1 INTRODUCTION

Alzheimer's disease (AD), a type of dementia, is characterized by progressive problems with thinking and behavior that starts in the middle or old age. The pathologic characteristics are the presence of neuritic plaques in the brain and degeneration of explicit brain cells. The symptoms usually develop slowly and get serious enough to interfere in daily life. Although the paramount risk factor is oldness but AD is not just an old age disease. In its early stages, the memory loss is mild while in the later stages, the patient's conversation and their ability to respond degrades dramatically. The current treatments cannot stop Alzheimer's disease (AD) from developing but early diagnosis can aid in precluding the severity of the disease and help the patients to improve the quality life. It has been reported that the number of individuals effected with AD will double in next 20 years (Zhang, 2011), while in 2050, 1 out of 85 individuals will be effected (Ron Brookmeyer, 2007). Thus the accurate diagnosis especially for the early stages of AD is very important.

Machine learning is used to interpret and analyze data. Furthermore it can classify patterns and model data. It permits decisions to be made that couldn't be made generally utilizing routine systems while sparing time (Mitchell T, 1997) and endeavors (Duda RO, 2001). Machine learning methodologies have been extensively used for computer aided diagnosis in medical image formation mining (Supekar, 2008) and retrieval (Bookheimer, 2000) with wide variety of other applications (Cruz, 2006) especially in detection and classifications of brain disease using CRT images (Cruz, 2006) and x-rays (Petricoin, 2004) It has just been generally late that AD specialists have endeavored to apply machine learning towards AD prediction. As a consequence, the literature in the field of Alzheimer's disease prediction and machine learning is relatively small. However, today's imaging technologies and high throughput diagnostics have lead us overwhelmed with large number (even hundreds) of cellular, clinical and molecular parameters. In current circumstances, the standard measurements and human instinct don't frequently work. That is the reason we must depend on

Khan, A. and Usman, M..
Early Diagnosis of Alzheimer's Disease using Machine Learning Techniques - A Review Paper.
In *Proceedings of the 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2015) - Volume 1: KDIR*, pages 380-387
ISBN: 978-989-758-158-8

intensively computational and non-traditional approaches such as machine learning. The custom of using machine learning as a part of disease prediction and visualization is a fragment of an expanding shift towards prescient (Weston, 2004) and customized prescription (Cruz, 2006). This drift is important, not only for the patients in increasing their quality of life and life style, but for physicians in making treatment decisions and also for health economists.

In evaluating and analyzing the existing studies, a number of common trends and gaps has been identified. The most evident trends include a rapid growth in the AD detection and prognosis using machine learning methods. Among the major gaps was an imbalance of events with attributes (few instances and too many attributes), the use of pathologically unproven data set (which cause uncertainty in results), class imbalance (too few instances in one class and too many instances in other class), overtraining and lack of external testing or validation. Nevertheless, the better designed and validates studies made it clear that machine learning methods, in comparison to standard statistical methods, could improve the accuracy of AD prediction. Besides, machine learning play an important role in AD prediction and prognosis.

To overcome these limitation, a model is proposed for effective diagnosis of onset of AD. While considering the pathologically proven data set, the proposed model involve a pre-processing step for eliminating the class imbalance issue. Important attributes selection using machine learning method help avoiding the problem of too few instances and too many attributes, known as curse of dimensionality (Cruz, 2006). The model divides the dataset into training and testing data. Training data on a limited testing data leads to a phenomenon of over-training (Chaves, 2010). Thus, training data should be selected to span a representative fragment of the actual data. The model presents classification using association rule mining with minimum support and minimum confidence. The paper is organized in a manner that Section II describes the different machine learning techniques. Section III and IV describes the literature review and critical evaluation. Proposed model is explained in Section V. Finally conclusions are drawn in Section VI.

1.1 Machine Learning Methods

Before starting the detailed analysis of machine leaning methods, it is significant to have a better understanding of what actually machine learning is and what machine learning techniques are commonly used AD prognosis. Machine learning comes under the umbrella of artificial intelligence and has variety of tools to make statistical, probabilistic decisions based on previous learning. It uses past learning (training) to classify new event and predict new patterns. Machine learning is very powerful as compared to standard statistical tools. In machine learning, a good understanding of a problem and limitations of the algorithms are needed to be understood well to get effective results. Therefore, it has a good chance for success if an experimentation is properly conducted and training is carefully and correctly employed and results are vigorously validated. Furthermore, all the algorithms and methods in machine learning are somewhat made different. For instance, few methods are designed on the basis of certain assumptions or for certain type of data which make it inapplicable for other type of data. That is why it is crucial to apply more than one machine learning method on given training data.

Machine learning generally have three types of learning algorithms: 1. Supervised learning 2. Unsupervised learning (Duda RO, 2001) 3. Reinforcement learning (Mitchell T, 1997). In supervised learning, a training data is given whereas the program tries to learn it and learns how to draw the input to the required output. The unsupervised learning algorithms employs self-learning based on unclassified and unlabeled data. Interestingly, the algorithms used in AD prognosis and diagnosis are almost all supervised learning algorithms including Artificial Neural Networks, Decision Trees, genetic algorithms and linear discriminant analysis.

Other techniques which are generally in use are SVM, AR mining, and Ensemble methods. In comparison to the above, SVM or support vector machine is somewhat newer technique (Duda RO, 2001) and is world known machine learning technique now but it is almost unidentified in AD prognosis field. The other methods such as KNN (K-Nearest Neighbors) and DTs (decision trees), are not widely used in AD predictions. Although, many high quality papers were studied for this review. However, almost all of them lacked a valid proven dataset for AD, lacked external or internal validation, were using too many attributes (causing over training) and no well-defined standard was made with which results were compared. These issues are further discussed in Section IV.

2 LITERATURE REVIEW

A detailed study on classification and diagnosis of AD has been proposed by many researchers. This section includes a brief review of the related work.

2.1 Single Modality Approach

The computer aided diagnosis of AD at the early stage of dementia is more challenging that lead R. Chaves et al., (2010) to introduce a classification method for effective and early diagnosis of Alzheimer's disease. Using association rule mining, they found out the associations between attributes of the pre-processed data sets. The proposed method was based on the tri-dimensional activated brain regions of interests (ROIs). These ROIs were obtained through a series of steps such as voxels of each image were considered as features (VAF) and the activation estimation using a certain threshold. For this purpose, a SPECT dataset of 97 instances was used out of which 43 were normal controls and remaining 54 were AD patients. The authors made comparisons with other techniques like VAF, PCA-SVM and GMM-SVM, and results revealed a classification accuracy of 95.87% (100% sensitivity, 92.86 specificity) with a claim of reducing the computational cost. This results show negligible difference in the accuracies with better efficiency in terms of computational time. The author claim it to be an "Effective" approach rather than efficient diagnosis of AD.

Distinguishing the early stage of the disease in AD patients using clinical conventions remained a diagnostic challenge. R. Chaves et al. (2011), later on, continued with his work by finding the associations among attributes while characterizing the perfusion patterns in SPECT images of normal subjects. For this purpose, complete image database was evaluated to reproduce the knowledge of medical experts. The pathologically unproven dataset from ADNI of 97 participants was used, of which 41 were labeled as healthy controls and 56 were labeled as AD patients by expert physicians. Comparisons were made with other techniques like PCA-SVM, GMM-SVM, output revealed the classification accuracy of 94.87% with 91.07% sensitivity and 100% specificity. The class imbalance was minimized as possible while the results were based on pathologically unproven data with no discussion about missing values.

The pathological unproven data sets of AD, made it applicable to different imaging technologies, as well, to diagnose other neuro-degenerative diseases. To address this, R. Chaves et al. (2012) introduced a mining technique using association rule mining defined over discriminant regions using pre-processed SPECT and PET imaging modalities. 97 participants contributed for the datasets, 42 were labeled as healthy controls and 55 were labeled as AD patients by expert physicians. The proposed method was compared with other techniques like PCA-SVM, VAF-SVM and results of this paper out proved them with accuracy of 92.78% with 87.5% sensitivity & 100% specificity for SPECT and 91.33% accuracy with 82.67% sensitivity & 100% specificity for PET. With no discussion about the missing values, the class imbalance have been reduced.

The study by Veeramuthu et al. (2014) developed a CAD tool for decision making about the presences of abnormalities in human brain. The author suggested preprocessing of PET dataset for instance, spatial normalization and intensity normalization. Fisher Discriminants ratio (FDR) was used for feature extraction to get ROIs. The instances were classified to normal if the extracted number of verified rules were above the final threshold otherwise image was classified as AD. The authors claimed 91.33% accuracy with 82.67% sensitivity & 100% specificity in comparison with other methods as VAF, PCA+SVM, and NFM+ SVM. It is observed that the authors did not mention the number of instances used in dataset. The methods adopted for dealing the missing data and class imbalance are also ignored. The dataset taken for the proposed study is not pathologically proven. Support and confidence, effective parameters of AR mining, are not discussed as well as no method for validation has been mentioned by the authors.

R. Chaves et al. (2012), impressed from the findings of PET data, tried to improve the prediction accuracy of the AD especially in early stage which has been of most concern to the researchers. The aim was the improvement in diagnosis of AD using Apriori AR progression and to develop new treatments and monitor their effectiveness while reducing the computational time and cost of clinical trials. The authors have introduced a method for analyzing of Alzheimer's disease by incorporating more detailed PET for instance, FDG-PET and PiB-PET. The data set comprised of 103 participants where 19 were control (CTRL), 19 were AD patients and 65 were with Mild cognitive impairment (MCI). The authors came with good results for PiB PET having classification accuracy of 97.37% and in combination with FDG it achieved the classification accuracy of 94.74% while FDG PET alone received 92.11% accuracy. The proposed method worked with a very small sized pathologically unproven data set with a class imbalance problem which produces uncertainty

in the acquired accuracies.

Similarly, Liu, Zhang et al. (2012) also contributed for early diagnosis of AD by implementing ensemble sparse method for the classification. The study revealed that noise and small sample size is very challenging to achieve good classification accuracy. As cited by the author, the high feature dimensionality will probably reduce the classification capability with standard classifier models, such as linear discriminant analysis, SVM and decision trees. The proposed study used Sparse representation-based classifier (SRC) to generate local patch based classifiers which are fused later on to give more robust and accurate classification. The authors found out that individual sub classifier can be easily trained thus dimensionality to subject ratio can be substantially improved. The study revealed that if the patches are from AD regions then classification accuracy will be high otherwise it will be low. Furthermore, pathologically unproven data set and class imbalance will demonstrate uncertainty in results.

Chaves et al., (2013) elaborated the early diagnosis later on by discretizing the continuous attributes of feature selection. Mean of control images were used to obtain a mask using histogram segmentation. AR mining used those RIOs as input and Control subject images were used to fully characterize the normal pattern of the image. The data of 97 participants was collected for SPECT, of which 41 were normal controls while 56 were AD. Moreover for PET, data of 150 participants was collected which comprised of 75 AD and 75 healthy controls. The results revealed 96.61% accuracy for SPECT while 92% accuracy for PET while comparison was made with VAF-SVM and PCA-SVM. To the best of our knowledge, it is the highest accuracy achieved for SPECT so far. However, the missing values have not been considered and mentioned in this study while the data used for this study was unproven which may degrade the results.

Klöppel et al, (2008) used the structural MRI to distinguish Alzheimer's disease from healthy controls at early stage. The authors applied SVM to MRI for the reliable detection of disease while distinguishing it from normal aging. This research was based on the pathologically proven data sets, collected from different centers as an input for effective classification. Finally, proposed method was implemented using normalized datasets from 67 AD and 91 healthy controls from different scans. Using the whole brain images, 96% of AD patients, who were pathologically verified, were correctly classified using leave one out cross validation. The proposed research showed generalization by combining data from different centers however, the data set is too small for fair comparisons with other methods.

2.2 Multimodal Approach

Although the use of different single biomarkers yield promising results but they are designed to characterize group differences and are not for individual classification. D. Zhang et al. (2011) came up with a method of combing all the three biomarkers for Alzheimer's disease diagnosis i.e. MRI, PET, CSF etc. to discriminate between healthy and AD participants. The authors made use of baseline data set with total 202 instances, out of which 51 were AD, 99 were MCI and 52 were healthy controls. Different tests were conducted for MRI, PET and CSF and the combination of these using 10 fold cross validation. The classification accuracy of 93.2% with 93% sensitivity and 93.3% specificity was achieved with combination of these modalities while individual test yielded highest accuracy of 86.5%. Authors claimed that multimodal classification method (using all MRI, PET, and CSF) achieves consistent improvement and is more robust over those using individual modality, for any number of brain regions selected.. These results directed that CSF and PET have the highest complementary information, while MRI and PET have the highest similar information for classification. Furthermore, it is noted that the availability of data of individual subject on all the modalities is too small for reasonable classification. The knowledge of missing values and how they are handled are not mentioned in this study. Class imbalance is another prominent limitation in this paper.

In support to the above, Westman, Muehlboeck et al. (2012) studied the combination of baseline MRI and CSF data to enhance the classification of AD while making comparison to individual modality. The data from 369 participants was collected to study regional subcortical volumes and cortical thickness measures. The data set comprised of 96 AD and 273 healthy controls, labeled by expert physicians. As cited by the author, FDG-PET can be expensive and it would have been interesting to see how the method of Zhang et al. performed with just the combination of MRI and CSF, but this data was not presented. Orthogonal partial least squares to latent structures (OPLS) multivariate analysis was used for 60 variables (57 from MRI and 3 from CSF). The proposed method resulted in classification accuracies of 91.8% for combined MRI and CSF which is slightly lower than those of (Cuingnet R1, 2011). The study also revealed that SVM and LDA have previously been utilized by others while OPLS showed more

Table 1: Summary and Critical Evaluation of techniques and limitations of different machine learning based AD studies.

	Modality	Technique	Data Set Details	Pathologically proven Data set	Accuracy	Limitation	Validation performed (No. of Folds)
(Chaves, Ramírez et al. 2013)	SPECT PET	Apriori- AR mining	SPECT: AD = 56 CTRL = 41 PET: AD = 75 CTRL = 75	No	SPECT: **96.91%** PET: 92%	Pathologically unproven data with no justification about missing values	Leave one out Cross Validation
(Klöppel, Stonnington et al. 2008)	MRI	Linear SVM	3-groups AD= 67 CTRL= 91	**Yes**	**96%**	Sample size is too small with no justification of missing values.	Leave one out Cross Validation
(Chaves, Ramírez et al. 2010)	SPECT	Apriori- AR mining	AD = 54 CTRL = 43	No	**95.87%**	Did not mention the how they limited the effect of missing values	Leave one out Cross Validation
(Chaves, Górriz et al. 2011)	SPECT	Apriori- AR mining	AD = 56 CTRL = 41	No	**94.87%**	The data may contain Missing values which will cause uncertainty	Leave one out Cross Validation
(Chaves, Ramirez et al. 2012)	FDG- PET + PiB-PET	Apriori- AR mining	AD = 19 CTRL = 84	No	**94.74%**	Unproven data with missing values	Leave one out Cross Validation
(Zhang, Wang et al. 2011)	MRI+ FDG-PET + CSF	SVM	AD = 51 CTRL = 151	No	**93.2%**	Class Imbalance and missing values	10-fold Cross Validation
(Chaves, Ramírez et al. 2012)	SPECT PET	Apriori- AR mining	SPECT: AD = 55 CTRL = 42 PET: AD = 75 CTRL = 75	No	**92.78%**	Unproven data with missing values	Leave one out Cross Validation
(Westman et al., 2012)	CSF MRI	Apriori- AR mining+ SVM	AD = 96 CTRL = 273	No	**91.8%**	Class Imbalance and missing values	7-fold Cross Validation
(Chaves, Ramírez et al. 2012)	SPECT PET	Apriori- AR mining for feature selection PCA, SVM	SPECT: AD = 56 CTRL = 41 PET: AD = 75 CTRL = 75	No	**91.75%**	Unproven data with missing values	Leave one out Cross Validation
A. Veeramuthu et al. (2014)	PET	AR mining	Not Given	No	**91.33%**	No dataset details, missing values or any preprocessing steps highlighted	No
Robi Polikar et al. (2010)	EEG + MRI + PET	Ensemble based decision fusion	AD = 37 CTRL = 36	No	**85.55%**	Unproven data with missing values	5-fold Cross Validation

similarities with SVM except for the ability to separate structured noise from the correlated variation modeling. Previous studies like (O. Kohannim, 2010) has shown that the combination of MRI and CSF significantly improves classification accuracy. However, CSF measures are highly invasive and could cause distress for patients which may provide a basis for combination of MRI and PET rather than MRI and CSF. Furthermore, the data set is not pathologically proven and author did not mention anything regarding missing data which may decrease the overall accuracy of the proposed method.

Polikar, Tilley et al. (2010) supported the use of CSF biomarker for being most promising in early diagnosis of AD. In contrast to that they revealed the costly and highly invasive nature of CSF biomarker along with its potentially painful lumber puncture. The proposed study investigated the fusion of non-invasive biomarkers such as PET, MRI as well as EEG to check their feasibility for the early diagnosis of Alzheimer's

disease. Using ensemble method, each classifier was trained on each datasets from different sources. Classifiers were then combined using an appropriate combination rule (Parikh, D. and R. Polikar, 2007). The Sum and simple majority voting (SMV) rules were used to obtain the data fusion diagnostic accuracies. Followed by the 5-fold cross validation, the outcome indicated the classification accuracy of 85.55% which is 10% -20% improvement as compared to fusion of any of two mentioned modalities. The Ensemble method is promising (Westman, 2012) however, the resultant accuracy is below as compared to the accuracies achieved in previous studies (Chaves, 2013). Although the authors reduced the class imbalance effect but they did not mentioned how they dealt with missing data.

3 CRITICAL EVALUATION

A detailed study on the early diagnosis and classification of AD has been proposed by many researchers. This segment contains a brief critical review and analysis of the related work.

3.1 Limitations

These studies outlined in the previous section are just few examples of how well machine learning experiments should be conducted and obviously there are other good and equally impressive studies with good results. These studies exemplified how the outcomes should be validated and described especially in the prognosis and prediction of AD. However, being able to identify the potential issues wither in the input data, experimental design, validation or the implementation is very critical especially for those who evaluate different studies as well as for those aiming to use machine learning.

Through the analysis of the above studies in this review, the most common problems among them were the input size, attributes and validation. It is easier to get higher accuracies with smaller datasets, such methods could not be used to represent larger population of data. It has been noted that small sample size is prone to overtraining and large data size ensures several effects on robustness, accuracy and reproducibility. It is impressive that 96.6% accuracy is attained, but unproven data used as input and the given size of the data put some doubt on the robustness of the model. Most of the research is done using pathologically unproven data which consequently may introduce uncertainty in the results. While such data can be obtained from specialist centers so no reason of not using such data have been identified in the related studies. The attributes to instance ratio also effects the results. In the above studies, lack of attention paid towards the number and general information of attributes.

Data quality and important attribute selection is also very important for effective results generation in machine learning. Unfortunately, the authors rarely described the methods used to ensure the data integrity and quality. Feature selection is also too important as data quality. However, the features chosen for some clinical data, for instance histological assessments, may not be applicable over time. Therefore, a classifier must be able to update feature sets with respect to time. Similarly, the details about the training and testing data should be clearly mentioned. Most of the algorithms focus more on the classification of major class whereas misclassify or ignore the minority class. Such class imbalance results in choosing the dominant class with poor class prediction, damaging the quality of classification.

4 PROPOSED MODEL

The proposed method consists of four steps as presented in Fig. 1: 1. Pre-processing, 2. Attribute selection, 3. Classification and 4. Class Threshold

For effective classification of the AD data, the first step is preprocessing. The pathologically proven data set is processed to avoid class imbalance and then it is converted to readable data type. Machine learning algorithms works very well when the number of instances of one class are almost equal to the number of instances of other class. Class imbalance damage the classification result severely so to avoid class imbalance, data is over sampled using machine learning technique for instance, synthetic minority oversampling technique (SMOTE). The input data type is converted from numeric into nominal/numeric to nominal values so that the algorithms which uses said data type as input can be implemented.

Attribute selection involves searching through all possible combinations of attributes in the data to find which subset of attributes works best for prediction and classification. It is helpful in the dimensionality reduction and omitting improper attributes. For classification tasks, it can lead to increased accuracy or to reduced computational costs. The third step is based on classification using AR mining with minimum support and minimum confidence. Classification is done using 10-fold cross validation that is, data is divided into 10 parts. One part is used as test and remaining 9 are used as training data and

the process is repeated 10 times to validate the results. The training set is used for classification in order to identify the specific parameters. The association rules results in unique associations among the attributes which are exploited in next step. In the last step, a certain threshold is used over the resultant rules to classify the instances into one of the two classes such as Control and AD.

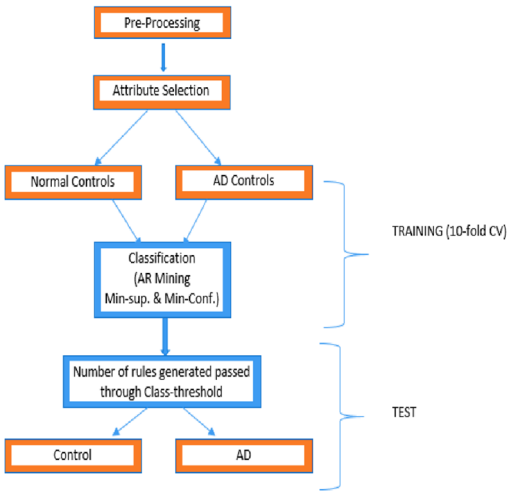

Figure 1: A proposed model for early detection of AD.

5 CONCLUSIONS

This study is based on the comparison and evaluation of recent work done in the prognosis and prediction of Alzheimer's disease using machine learning methods. Explicitly, the recent trends with respect to machine learning has been revealed including the types of data being used and the performance of machine learning methods in predicting early stages of Alzheimer's. It is obvious that machine learning tends to improve the prediction accuracy especially when compared to standard statistical tools. However, based on the review, the clinical diagnosis were not 100% accurate as pathological verification was not provided which consequently introduce uncertainty in the predicted results. The proposed method deals with pathologically proven data and overcomes the class imbalance and overtraining issues. Proposed model is based on single modality to overcome the increased cost of computing and combining different modalities. We believe that pathologically proven data may increase the accuracy and validity, while a balanced class will help the classifiers to give accurate results. This model is can help to improve the prediction performance by physicians and cover the limitations pointed out in the previous research.

REFERENCES

D, Zhang, 2011. Multimodal classification of Alzheimer's disease and mild cognitive impairment. *Neuroimage*, 55(3), 856-867.

Brookmeyer, Ron , Elizabeth Johnson, Kathryn Ziegler-Graham, and H. Michael Arrighi, 2007. Forecasting the Global Burden of Alzheimer's Disease. *Alzheimer's and Dementia*, 3.3, 186-191.

Brookmeyer, Duda RO, Hart PE, Stork DG. , 2001. Pattern classification. *Alzheimer'(2nd edition). New York: Wiley.s and Dementia*, 2nd edition.

T, Mitchell , 1997. Machine Learning. *New York: McGraw Hill*, 2nd edition, 135.

SY, Bookheimer, Strojwas MH, Cohen MS, Saunders AM, Pericak-Vance MA, Mazziotta JC, 2000. , et al. Patterns of brain activation in people at risk of Alzheimer's disease. *N Engl J Med*, 6, 343:450.

Supekar, K, Menon V, Rubin D, Musen M, Greicius MD, 2008. Network analysis of intrinsic functional brain connectivity in Alzheimer's disease. . *PLoS Comput Biol* , 4(6), 1–11.

Cruz, J.A. and D.S. Wishart, 2006. Applications of Machine Learning in Cancer Prediction and Prognosis.*Cancer Informatics*, 2, 59-77.

EF, Petricoin, Liotta LA. , 2004. SELDI-TOF-based serum proteomic pattern diagnostics for early detection of cancer. *Curr Opin Biotechnol*, 15, 24-30.

Bocchi, L, Coppini G, Nori J, Valli G, 2004. Detection of single and clustered microcalcifications in mammograms using fractals models and neural networks. *Med Eng Phys*, 26, 303-12.

Weston, AD, Hood L., 2004. Systems biology, proteomics, and the future of health care: toward predictive, preventative, and personalized medicine. . *J Proteome Res*, 3, 179-96.

Bellman, R, 1961. Adaptive Control Processes: A Guided Tour. *Princeton University Press*, 1, 45.

Rodvold, DM, McLeod DG, Brandt JM, 2001. Introduction to artificial neural networks for physicians: taking the lid off the black box. *Prostate*, 46, 39-44.

Chaves, R., J. Ramírez, et al., 2010. Effective Diagnosis of Alzheimer's Disease by Means of Association Rules. *Hybrid Artificial Intelligence Systems, Springer*, 1, 452-459.

Chaves, R., J. Górriz, et al. (2011). Efficient mining of association rules for the early diagnosis of Alzheimer's disease. *Physics in medicine and biology* 56(18): 6047.

Chaves, R., J. Ramírez, et al. (2012). Association rule-based feature selection method for Alzheimer's disease diagnosis. *Expert Systems with Applications* 39(14): 11766-11774.

Chaves, R., J. Ramírez, et al. (2012). Functional brain image classification using association rules defined over discriminant regions. *Pattern Recognition Letters* 33(12): 1666-1672.

Veeramuthu, A., S. Meenakshi, et al. (2014). A New Approach for Alzheimer's Disease Diagnosis by using Association Rule over PET Images. *International Journal of Computer Applications* 91(9), 9-14.

Chaves, R., J. Ramirez, et al. (2012). FDG and PIB biomarker PET analysis for the Alzheimer's disease detection using Association Rules. *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, IEEE.

Liu, M., D. Zhang, et al. (2012). Ensemble sparse classification of Alzheimer's disease. *Neuroimage* 60(2), 1106-1116.

Hinrichs, C., V. Singh, et al. (2009). Spatially augmented LPboosting for AD classification with evaluations on the ADNI dataset. *Neuroimage* 48(1), 138-149.

Cuingnet R1, G. E., Tessieras J, Auzias G, Lehéricy S, Habert MO, Chupin M, Benali H, Colliot O; (2011). Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database. *Neuroimage* 56, 766–781.

Chaves, R., J. Ramírez, et al. (2013). Integrating discretization and association rule-based classification for Alzheimer's disease diagnosis. *Expert Systems with Applications* 40(5), 1571-1578.

Klöppel, S., C. M. Stonnington, et al. (2008). Automatic classification of MR scans in Alzheimer's disease. *Brain* 131(3), 681-689.

Westman, E., J.-S. Muehlboeck, et al. (2012). Combining MRI and CSF measures for classification of Alzheimer's disease and prediction of mild cognitive impairment conversion. *Neuroimage* 62(1),229-238.

Westman, Eric, et al. (2011). Multivariate analysis of MRI data for Alzheimer's disease, mild cognitive impairment and healthy controls. *Neuroimage* 54.2, 1178-1187.

O. Kohannim, X. Hua, D.P. Hibar, S. Lee, Y.Y. Chou, A.W. Toga, C.R. Jack Jr., M.W. Weiner, P.M. Thompson (2010) Alzheimer's Disease Neuroimaging Initiative Boosting power for clinical trials using classifiers based on multiple biomarkers, *Neurobiology of Aging*, 31 (8), 1429–1442.

Polikar, R., C. Tilley, et al. (2010). Multimodal EEG, MRI and PET data fusion for Alzheimer's disease diagnosis. *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE,* IEEE.

Parikh, D. and R. Polikar (2007). An ensemble based incremental learning approach to data fusion. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions*, 37(2), 437-450.