Detecting Alzheimer's Disease on Small Dataset: A Knowledge Transfer Perspective AI

Mrs padma joshi[1] K.pranavi [1]l.saipraneetha [1]m.srinidhi[2]
[1]Assistant Professor, Department of CSE,[2]Final year B.Tech. Students, Department of CSE,
Sreyas Institute of Engineering and Technology, Hyderabad, Telangana, India

ABSTRACT

In that paper, we proposed as checking the whole patient Alzheimer using SVM and DT classification . Alzheimer's a progressive disease that demolishes brain's memory and its regular functioning as well. There isn't any single test till date to diagnose this disease and brain scans alone can't determine whether the person is possessed by it. Currently, the physician believes that a person is affected by Alzheimer's by based the reports of the family members about the behavioural tendencies and observations of the past medical history. AI combined with Machine Learning algorithms might now be able to change this situation. Big Data processing, as the information comes from multiple, heterogeneous, autonomous sources with complex and evolving relationships, and keeps growing. So in that, we will take results of how much percentage patients get disease as a positive information and negative information. The proposed shows a Bi processing model, from the data mining perspective. Using classifiers, we are processing Alzheimer percentage and values are showing as a confusion matrix. We proposed a new classification scheme which can effectively improve the classification performance in the situation that training dataset is available. In that dataset, we have nearly 500 patient details.

1.EXISTING SYSTEM

In this existing method is we used Naïve Bayes Algorithm for predict that Alzheimer details. Naïve Bayes algorithm makes only clustering process and it will segregate the person who are all suffering from Alzheimer**.** Naïve Bayes Algorithm is slow process for classify all the given details. In that main disadvantage is time efficiency. Patient's Alzheimer details start with large-volume, heterogeneous, autonomous sources with distributed and decentralized control, and seek to explore complex and evolving relationships among data.

1.1.LIMITATIONS

- It will take more time to calculate our dataset which we having our Alzheimer dataset.
- Discover patterns using traditional data mining tools may not used for predictions.
- They concluded that drug treatment for patients in the young age group can be delayed whereas; patients in the old age group should be prescribed drug treatment immediately.
- Prediction and classification of various type of Alzheimer using classification algorithm was carried out in Alzheimer dataset.

2 .PROPOSED SYSTEM

The raw data set is given as input to the system. The unstructured voluminous input data can be obtained from various Electronic Health Record (EHR) / Patient Health Record (PHR), Clinical systems and external sources (government sources, laboratories, pharmacies, insurance companies etc.). Algorithms used is , SVM and DTree algorithm to predict the results accurately. To explore Big Data, proposed system analyzed several challenges at the data, model, and system levels. To support Big Data mining, high-performance computing platforms are required, which impose systematic designs to unleash the full power of the Big Data. At the data level, the autonomous information sources and the variety of the data collection environments, often result in data with complicated conditions, such as missing/uncertain values. In other situations, privacy concerns, noise, and errors can be introduced into the data, to produce altered data copies.

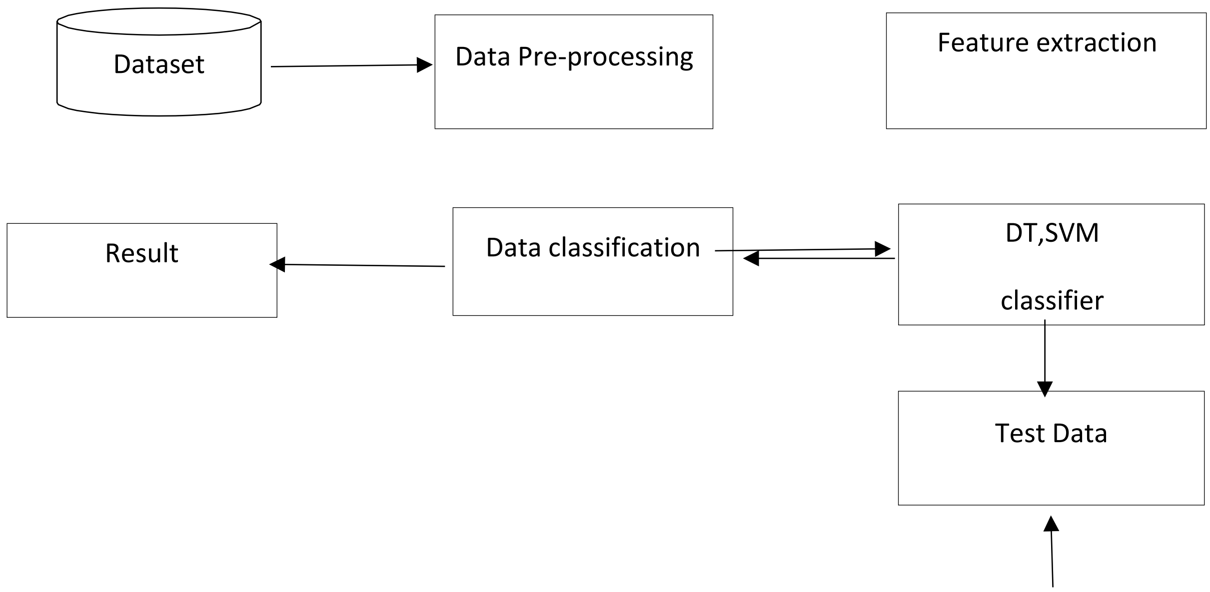
2.1.ADVANTAGES

- Fastest path to business value from raw big data .
- Machine learning models to detect the Alzheimer early.
- Discover patterns using machine learning algorithm that identify the best mode of treatment for Alzheimer across different age.
- Novel, high value business insights driving growth and profitability
- Leverage existing skills and investments
- Minimal time, cost and effort spent
- These technical challenges are common across a large variety of application domains, and therefore not cost-effective to address in the context of one domain alone.

3.OBJECTIVE

To developed and implement a Machine learning approaches that characterizes the features of the Alzheimer , and proposes a Alzheimer processing model, from the machine learning perspective using Support Vector Machine and Decision Tree Classifiers for Classification .

4.SYSTEM ARCHITECTURE



5.SUPPORT VECTOR MACHINE

Support Vector Machine (SVM) is a supervised machine learning algorithm which can be used for both classification or regression challenges . It is generally utilized in characterization issues. In the SVM calculation, we plot every datum thing as a point in n-dimensional space (where n is number of highlights you have) with the estimation of each element being the estimation of a specific arrange. At that point, we perform order by finding the hyper-plane that separates the two classes quite well. Bolster Vectors are essentially the co-ordinates of individual perception. The SVM classifier is a wilderness which best isolates the two classes (hyper-plane/line)

5.1.Working with Support vector machine in Python:

```
#Import Library
# Import other necessary libraries like pandas, numpy…
from sklearn import tree
# Assumed you have, X (predictor) and Y (target) for training data set and
x_test (predictor) of test_dataset
# Create tree object
model = SVC (criterion='gini') # for classification
# model = SVR() for regression
# Train the model using the training sets and check score
model.fit(X, y)
model.score(X, y)
#Predict Output
predicted = model.predict(x_test)
```

6.DECISION TREE

Decision tree is the most powerful and popular tool for classification and prediction. A Decision tree is a flowchart like tree structure, where each internal node denotes a test on an attribute, each branch represents an outcome of the test, and each leaf node (terminal node) holds a class label. Tree based learning algorithms are considered to be one of the best and mostly used supervised learning methods. Tree based methods empower predictive models with high accuracy, stability and ease of interpretation. Unlike linear models, they map non-linear relationships quite well. They are adaptable at solving any kind of problem at hand (classification or regression). Decision Tree algorithms are referred to as **CART (Classification and Regression Trees)**.

6.1.Common terms used with Decision trees:

1. **Root Node:** It represents entire population or sample and this further gets divided into two or more homogeneous sets.

2. **Splitting:** It is a process of dividing a node into two or more sub-nodes.
3. **Decision Node:** When a sub-node splits into further sub-nodes, then it is called decision node.
4. **Leaf/ Terminal Node:** Nodes do not split is called Leaf or Terminal node.
5. **Pruning:** When we remove sub-nodes of a decision node, this process is called pruning. You can say opposite process of splitting.
6. **Branch / Sub-Tree:** A sub section of entire tree is called branch or sub-tree.
7. **Parent and Child Node:** A node, which is divided into sub-nodes is called parent node of sub-nodes whereas sub-nodes are the child of parent node.

6.2.Working with Decision Trees in Python:

```
#Import Library
# Import other necessary libraries like pandas, numpy…
from sklearn import tree
# Assumed you have, X (predictor) and Y (target) for training data set and
x_test (predictor) of test_dataset
# Create tree object
model = tree.DecisionTreeClassifier(criterion='gini') # for classification,
here you can change the algorithm as gini or entropy (information gain) by
default it is gini
# model = tree.DecisionTreeRegressor() for regression
# Train the model using the training sets and check score
model.fit(X, y)
model.score(X, y)
#Predict Output
predicted = model.predict(x_test)
```

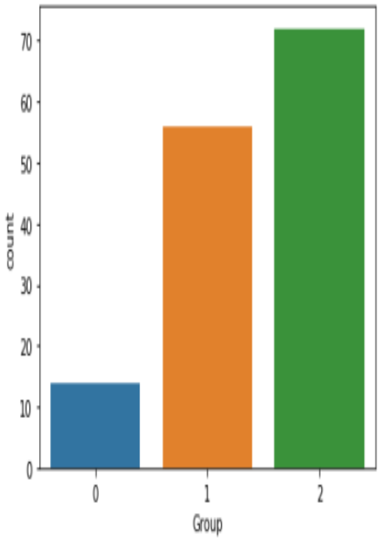
7.DATASET DESCRIPTION

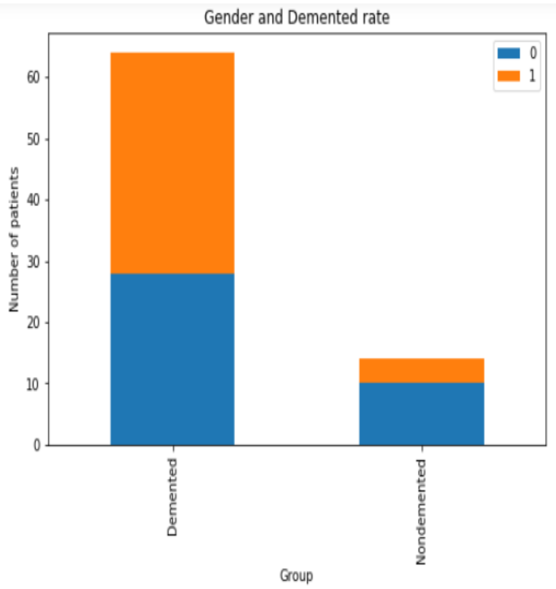
The team has found MRI related data that was generated by the Open Access Series of Imaging Studies (OASIS) project that is available both, on their **website** and **kaggle** that can be utilized for the purpose of training various machine learning models to identify patients with mild to moderate dementia.

- We will be using the longitudinal MRI data.
- The dataset consists of a longitudinal MRI data of 150 subjects aged 60 to 96.
- Each subject was scanned at least once.
- Everyone is right-handed.
- 72 of the subjects were grouped as 'Nondemented' throughout the study.
- 64 of the subjects were grouped as 'Demented' at the time of their initial visits and remained so throughout the study.
- 14 subjects were grouped as 'No ndemented' at the time of their initial visit and were subsequently characterized as 'Demented' at a later visit. These fall under the 'Converted' category.

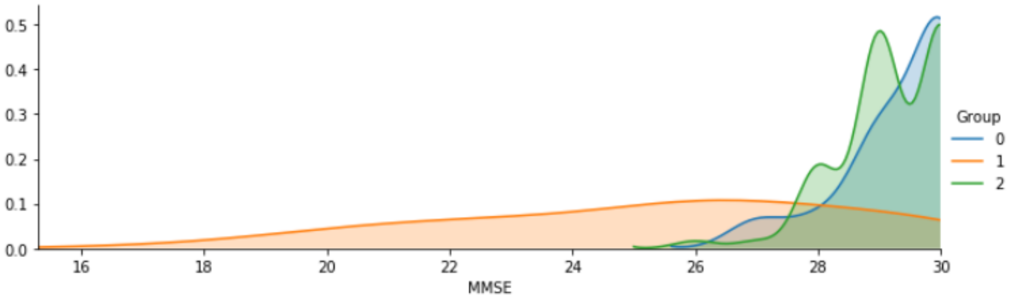
COL	FULL-FORMS
EDUC	Years of education
SES	Socioeconomic Status
MMSE	Mini Mental State Examination
CDR	Clinical Dementia Rating
eTIV	Estimated Total Intracranial Volume
nWBV	Normalize Whole Brain Volume
ASF	Atlas Scaling Factor

ISSN: 2278-4632
Vol-10 Issue-5 No. 6 May 2020

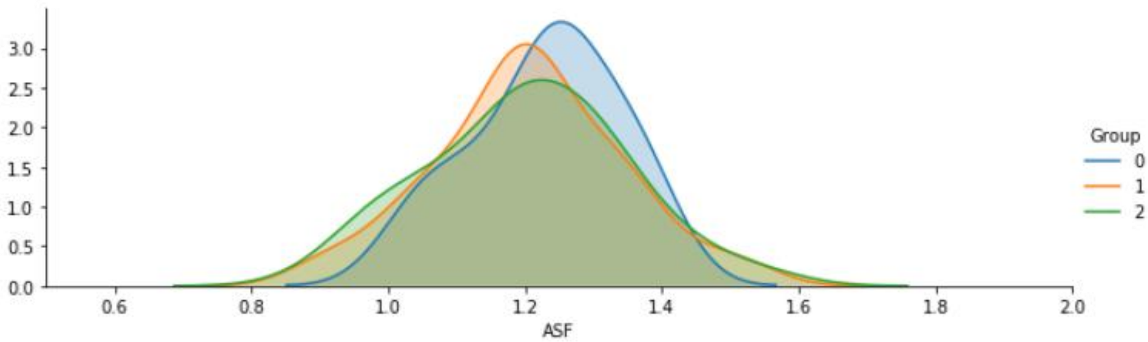




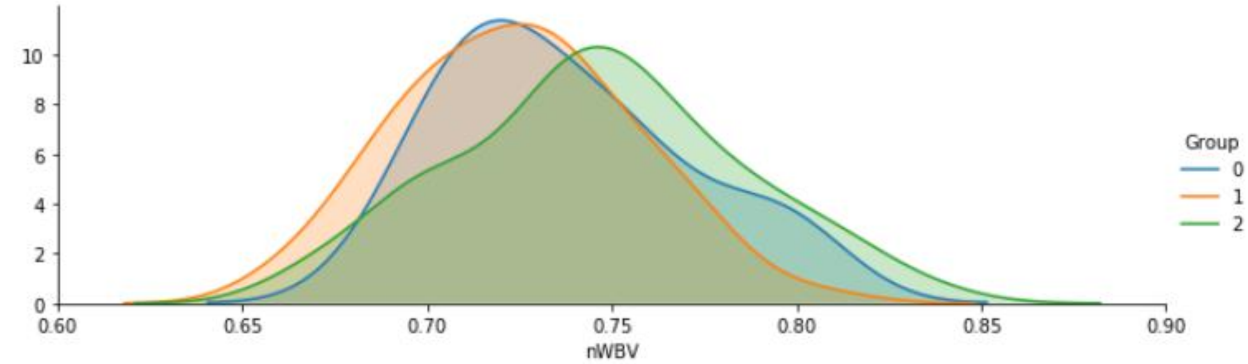
(15.3, 30.0)



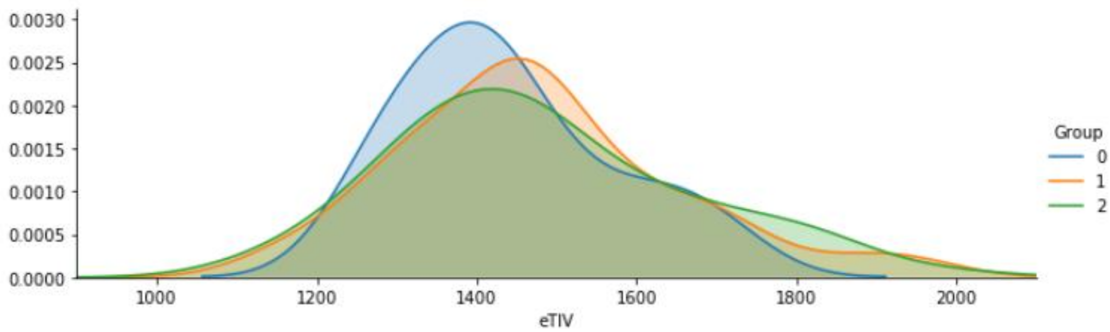
(0.5, 2)



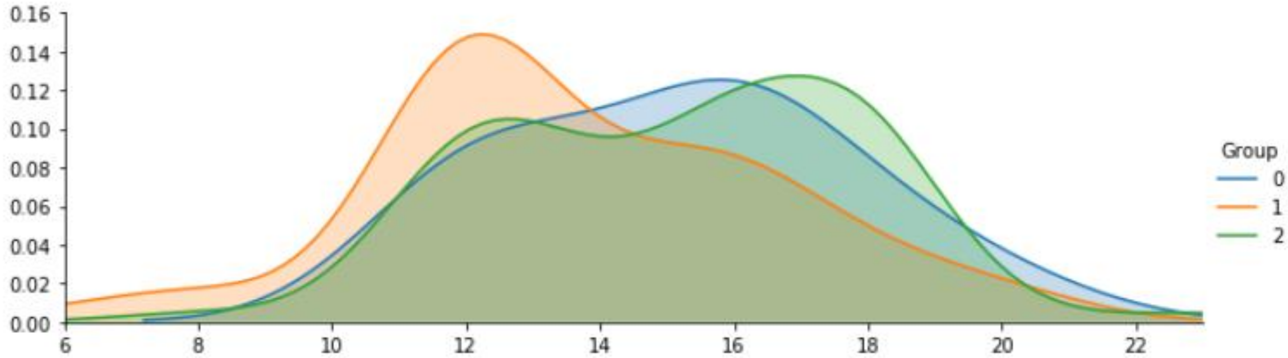
(0.6, 0.9)

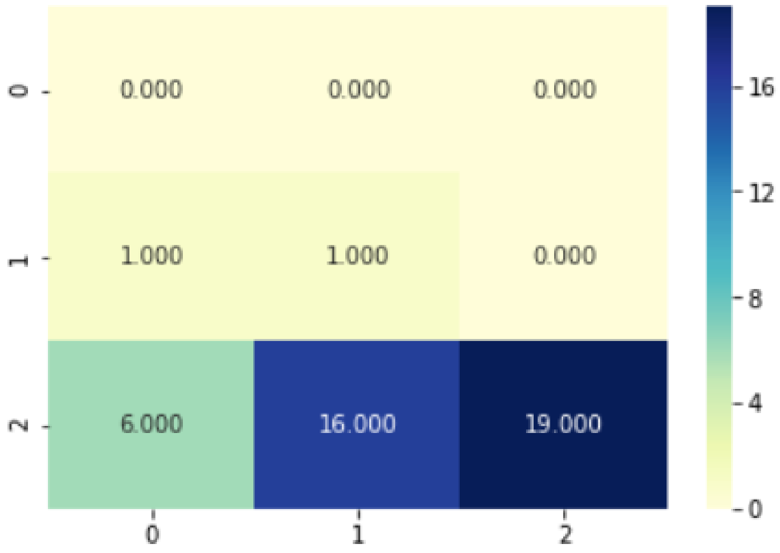


(900, 2100)

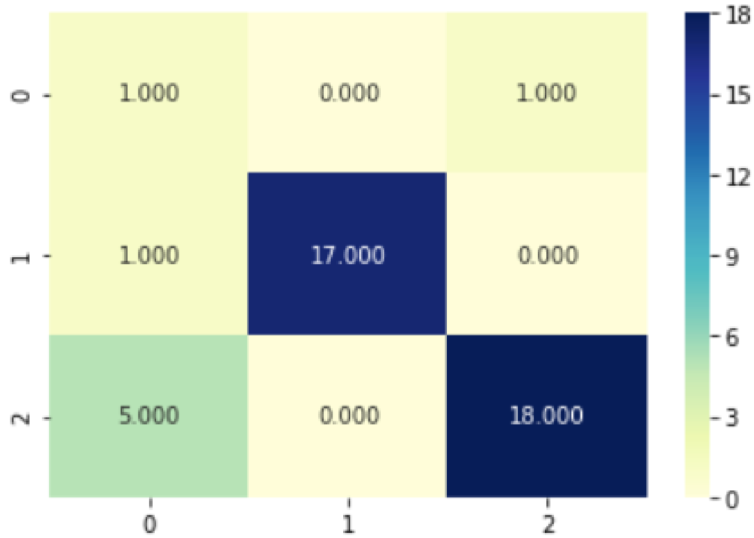


(0, 0.16)

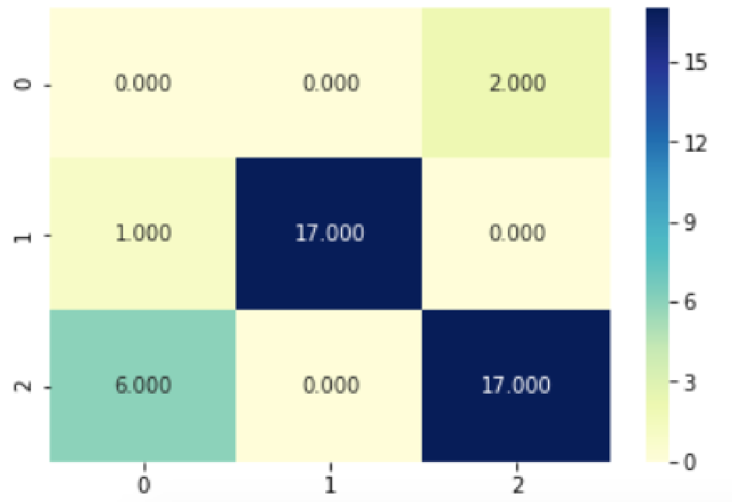




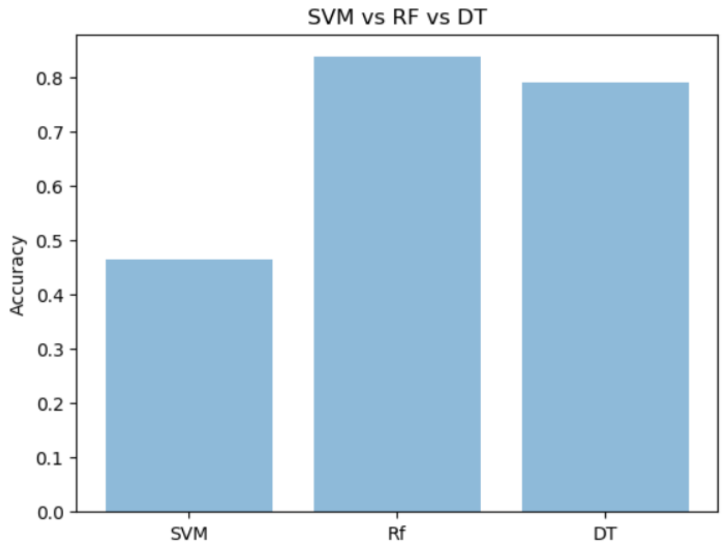
Confusion matrix for SVM algorithm



Confusion matrix for Random forest classifier



Confusion matrix of Decision tree



8.CONCLUSION:

In the present study, we have proposed a supervised classification model based on a SVM and DTree algorithm that can efficiently discriminate between patients with MCI and healthy subjects using clinical CSF biomarkers. The ability of the model to predict patients with early stages of the AD was based on appropriate training attributes selected using feature selection method. The efficiency of the model built using SVM and DTree were evaluated using various statistical performance evaluators and compared. Therefore, the present study is a step forward in predicting the early stages of Alzheimer disease using the ML-based classification model based on early stage.

9.REFERENCES

[1] M. Prince, A. Wimo, M. Guerchet, A. Gemma-Claire, Y.-T. Wu, and M. Prina, "World Alzheimer Report 2015: The Global Impact of Dementia - An analysis of prevalence, incidence, cost and trends," Alzheimer's Dis. Int., p. 84, 2015.

[2] M. Graña et al., "Computer Aided Diagnosis system for Alzheimer Disease using brain Diffusion Tensor Imaging features selected by Pearson's correlation," Neurosci. Lett., vol. 502, no. 3, pp. 225–229, 2011.

[3] M. Dyrba et al., "Combining DTI and MRI for the automated detection of Alzheimer's disease using a large European multicenter dataset," in International Workshop on Multimodal Brain Image Analysis, 2012, pp. 18–28.

[4] S. Haller et al., "Individual classification of mild cognitive impairment subtypes by support vector machine analysis of white matter DTI," Am. J. Neuroradiol., vol. 34, no. 2, pp. 283–291, 2013.

[5] W. Lee, B. Park, and K. Han, "Classification of diffusion tensor images for the early detection of Alzheimer's disease," Comput. Biol. Med., vol. 43, no. 10, pp. 1313–1320, 2013.

[6] T. M. Nir et al., "Diffusion weighted imaging-based maximum density path analysis and classification of Alzheimer's disease," Neurobiol. Aging, vol. 36, no. S1, pp. S132–S140, 2015.

[7] R. Cuingnet et al., "Automatic classification of patients with Alzheimer's disease from structural MRI: A comparison of ten methods using the ADNI database.," Neuroimage, vol. 56, no. 2, pp. 766–81, 2010.

[8] E. Westman, J.-S. Muehlboeck, and A. Simmons, "Combining MRI and CSF measures for classification of Alzheimer's disease and prediction of mild cognitive impairment conversion," Neuroimage, vol. 62, no. 1, pp. 229–238, 2012.

[9] C. Aguilar et al., "Different multivariate techniques for automated classification of MRI data in Alzheimer's disease and mild cognitive impairment," Psychiatry Res. - Neuroimaging, vol. 212, no. 2, pp. 89–98, 2013.

[10] A. Ortiz, J. M. Górriz, J. Ramírez, F. J. Martínez-Murcia, and A. D. N. Initiative, "LVQ-SVM based CAD tool applied to structural MRI for the diagnosis of the Alzheimer's disease," Pattern Recognit. Lett., vol. 34, no. 14, pp. 1725–1733, 2013.

[11] L. Khedher, J. Ramírez, J. M. Górriz, A. Brahim, and F. Segovia, "Early diagnosis of Alzheimer's disease based on partial least squares, principal component analysis and support vector machine using segmented MRI images," Neurocomputing, vol. 151, no. P1, pp. 139–150, 2015.

