

# Alzheimer's Dementia: Diagnosis and Prognosis using Neuro-Imaging Analysis

Sountharajan S<sup>1\*</sup>, Karthiga M<sup>2</sup>, Suganya E<sup>3</sup>, Rajan<sup>4</sup>

<sup>1</sup>Department of CSE, Amrita School of Computing, Amrita Vishwa Vidyapeetham, Chennai, India

<sup>2</sup>Department of CSE, Bannari Amman Institute of Technology, Sathyamangalam, Tamilnadu, India

<sup>3</sup>Department of Information Technology, SSN College of Engineering, SH 49A, Kalavakkam, India

<sup>4</sup>Department of Information Technology, K. S. Rangasamy College of Technology, Tiruchengode, Namakkal, Tamil Nadu

Email: s\_sountharajan@ch.amrita.edu<sup>1\*</sup>

## Abstract

Alzheimer's disease (AD) is the most common type of progressive neurological disorder that leads to the death of brain cells over the time. It causes memory loss and decline in the cognitive skills among the elderly subjects. Early diagnosis of the progressive diseases plays a vital role in the healthcare community. Machine learning (ML) algorithms and various multivariate data exploratory tools are employed in the field of AD research. The main purpose of this work is to analyse the importance of features selection which in turn enhances the classification accuracy of the models. The hyper parameter tuning for Support Vector Machine (SVM) classification and Boruta algorithm for Random Forest (RF) classification are applied for the selection of optimal set of features. In this work, a five-stage ML pipeline with each stage further categorized into different sub-levels is proposed. Initially, the data collected from the Open Access Series of Imaging Studies (OASIS-2) dataset of Magnetic Resonance Imaging (MRI) brain images is explored and pre-processed using the imputation technique. Feature scaling of the pre-processed data is done using the Min-max scaling technique. Then, the classification techniques such as logistic regression, Decision Tree (DT) classification, Random Forest (RF) classification, Support Vector Machine (SVM) classification and AdaBoost Classification are applied to classify the data and finally the performance of the classifiers are compared in terms of accuracy, Area under the curve of the Receiver Operating Characteristic (AUC) curve and recall measures. From the performance analysis, it is concluded that the Random Forest (RF) classifier yields maximum accuracy, recall and AUC values. The hyperparameter tuning and Boruta algorithm added significance to the SVM and RF classification, thereby resulting in a F-score of 91% and 92% respectively.

**Keywords:** Random Forest (RF), Support Vector Machine (SVM), Hyperparameter tuning, Boruta Algorithm, Feature selection

## INTRODUCTION

Alzheimer's disease is a neurodegenerative disorder that hinders performance in the regular activities of an individual. The victim survives without any symptoms for years till they experience loss of memory and speech issues [1]. The neurons in the learning, thinking and cognitive parts of the brain worn out with time resulting in the death of the brain cells. This results in the development of dementia, a form of the Alzheimer's disease [2]. In general, Alzheimer's disease is categorized into three stages including preclinical biomarker, mild cognitive disorder and Alzheimer dementia.

The preclinical biomarker is a developmental stage where there are no symptoms like memory loss still there is a minor change in the fluid and blood in the brain. The mild cognitive disorder results in light cognitive impairment without hindrance in the day-to-day activities. It cannot be assured that mild cognitive disorder is due to Alzheimer it may also happen due to medications right after a trauma [3]- [5].

**Address for correspondence:** Sountharajan S, Department of CSE, Amrita School of Computing, Amrita Vishwa Vidyapeetham, Chennai, India,  
Email: s\_sountharajan@ch.amrita.edu

Received date: 10 August 2022

Accepted: 19 September, 2022

Published: 10 October, 2022

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

**For reprints contact:** pnrjournal@gmail.com

**How to cite this article:** S S, M K, E S, Rajan, Alzheimer's Dementia: Diagnosis and Prognosis using Neuro-Imaging Analysis, J Pharm Negative Results 2022;13(4):46-63.

Access this article online

Quick Response Code:



Website:

www.pnrjournal.com

DOI:

10.47750/pnr.2022.13.04.006

Dementia is the final stage in which Alzheimer disease reports positive with a drastic change in the symptoms in the form of cognitive impairment over a period of time. The Alzheimer Dementia (AD) is further classified as mild, moderate and severe stage which can be represented as AD1, AD2 and AD3 respectively. The mild stage the individual can lead an independent life without any assistance. The problem in their language, memory loss, agitation and lack of performance in their routine shows the moderate Alzheimer. In the AD2 stage, they require little assistance from the care takers in moving on with their routine life. The blood clots in the brain and organ failure makes the victim bed bound which requires a care taker around the clock in AD3 stage [6].

Overall, the Alzheimer disease suppresses the entire body due to depression, thyroid problems, sleep apnea and leads to excessive alcohol intake. The dark side of the Alzheimer's disease is that for every 65 seconds a person is diagnosed with Alzheimer in United States [7]. Still, the medicine for AD is in research hence it is better to diagnose the disease in the early stages. The Magnetic Resonance Imaging (MRI) of the brain can be used for the diagnosis of suspected patients. The local and generalized shrinkage of tissues in the brain is a factor of evaluation [8]. The researchers have found that the features obtained from the MRI brain images greatly helps in devising the further treatment for Alzheimer. The electronic patient records maintained at present facilitates the computer aided diagnosis that ensures accurate predictions. The health care

departments are enhanced by the application of machine learning techniques to improve the quality and to reduce the manual resources. The machine learning algorithms derive the hidden insights from the existing dataset which helps in the classification of new records [9]. Instead of explicit programming, the machine learning derives the patterns using features. The electronic health records contain several features in a non-linear fashion which leads to curse of dimensionality [10]. The main objective of this research is to classify the patient records with dementia and non-dementia using different machine learning techniques from Open Access Series of Imaging Studies (OASIS-2) dataset of Magnetic Resonance Imaging (MRI) brain images [11]. Several researches consider raw MRI data whereas this work deals with three to four biomarkers derived from the MRI images. The classification techniques such as logistic regression, Decision Tree (DT) classification, Random Forest (RF) classification, Support Vector Machine (SVM) classification and AdaBoost Classification are compared in terms of accuracy, Receiver Operating Characteristic (ROC) curve and recall measures. Feature selection phase of machine learning enhances the classification accuracy in case of large dataset. The minimal optimal issue can be solved by choosing the significant features from the dataset. The random search feature selection on SVM and Boruta algorithm for RF classification is applied to show difference in the performance.

The remaining sections of the paper are formulated as follows: Section II presents the related study of machine

learning based Alzheimer's disease prediction. Section III presents the feature selection strategies opted in this paper, namely, hyperparameter tuning and Boruta algorithm. In addition it describes various classification models evaluated using OASIS dataset. Section IV has a detailed discussion of results obtained at each stage. Section V concludes the paper with feature work.

## Literature Survey

OASIS and MIDAS are the popular databases that have multimodal images of scanned body parts, especially known for brain images. This biomedical dataset has gained its importance in various developed nations to attain benchmarking research results in biomedical science. The latest version of OASIS is the third version which contains highly performing multimodal images [12]. The OASIS dataset is used for the evaluation of deformation-based features including modular grey matter, geodesic anisotropy, Jacobian determinant and deformation vector norm. The optimal features are selected using Welch's t-test, Pearson correlation and Bhattacharyya distance methods. The linear SVM classification is applied to classify the Alzheimer's disease as dementia or non-dementia. The research shows that Welch's t-test has selected the optimal features that results in higher prediction accuracy [13]. The missing values are handled by the eXtreme Gradient Boosting (XGBoost) algorithm for multiclass classification. On comparison with RF and SVM classification, XGBoost handles the missing values in an efficient manner [14]. Machine learning techniques including K-means, Naïve Bayes, SVM, K-Nearest Neighbor, and Artificial Neural Networks (ANN) is applied for classifying Alzheimer's disease. These techniques are applied on OASIS, Alzheimer's Disease Neuroimaging Initiative (ADNI) and the European DTI Study on Dementia (EDSD) dataset [15].

An optimized kernel parameter is chosen to develop principal component analysis based on downsized kernel, which normalizes the feature for attaining better classification results. The classification of normalized features is done using a multiclass support vector machine. Hence, the principal component analysis reduced the dimensionality of synthetic data and it is evaluated on the OASIS MRI database [16]. The classification of Alzheimer's disease is automated using three feature extraction techniques including Freesurfer, NiftyReg and Multi-Atlas Labeled Propagation with EM refinement (MALPEM). Region of Interest (ROI) is extracted from each image and the extracted features are classified using K-Nearest Neighbor classification technique. The issue of binary classification is resolved by the multimodal feature and it attained an accuracy of 90% [17]. On comparison with the texture based features extracted using Gray Level Co-occurrence Matrix (GLCM), the statistical features provide better results. The classification is carried out using SVM, Random forest and ensemble based classifiers, among which Random forest provided better results on OASIS dataset [18].

Several machine learning algorithms are applied on the OASIS dataset for the early detection of Alzheimer's disease [19]. The five staged multi modal classifier has the feature selection strategies embedded in it. The results showed that Random forest outperformed the other classification algorithms [20]. Three class SVM classifier using Bag of Features is applied to classify Alzheimer's disease in both normal and tumor brain images. Upon the selection of interest point, bag of features method is applied for feature extraction. The histogram calculations are performed on the extracted features before classification [22]. The SVM classifier is incorporated with recursive feature elimination method, namely, universum SVM that facilitates classification with global data. The significant regions of brain image to attain better classification accuracy are obtained in global data rather than local data [21]. The morphometry analysis is done based on volume and voxels features. The Universum SVM outperformed the conventional SVM with recursive feature elimination [23]. A similar kind of brain MRI based prediction for Parkinson's disease using variational autoencoders is proposed in [25] wherein the authors suggested an accuracy increase of 6% compared to the existing. A comparative based machine learning models are used to determine the accuracy outperform level in the proposed work for Alzheimer's

disease in [26].

This research mainly focused to build an efficient feature reduction and classification algorithm to diagnosis the Alzheimer's dementia. We executed the novel feature selection techniques to diagnose and extract the optimized features. Various Machine Learning algorithms are applied to the optimized features and their accuracy are compared. From the results, it is determined that classification with optimized features outperforms well in terms of efficiency than the classification with all the features. So, this model is best suited in diagnosing the Alzheimer's disease with greater accuracy. [24]

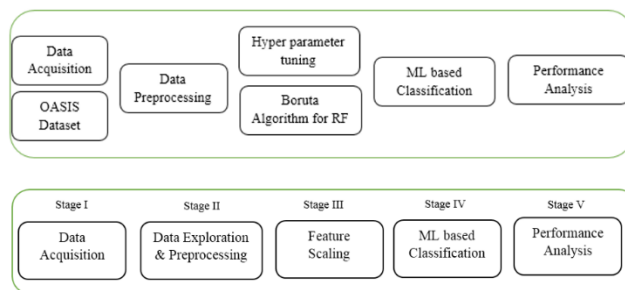
## Methodology

In this section, the detailed description of the work is described in 4 level machine learning based approach. The methodology layout is depicted in Fig.1.

## Data acquisition

The dataset for training the machine learning based classification model is collected from the Open Access Series of Imaging Series (OASIS) website [11]. The longitudinal MRI brain images in Comma Separated Values (CSV) format for around 150 patients are utilized.

## Dataset description



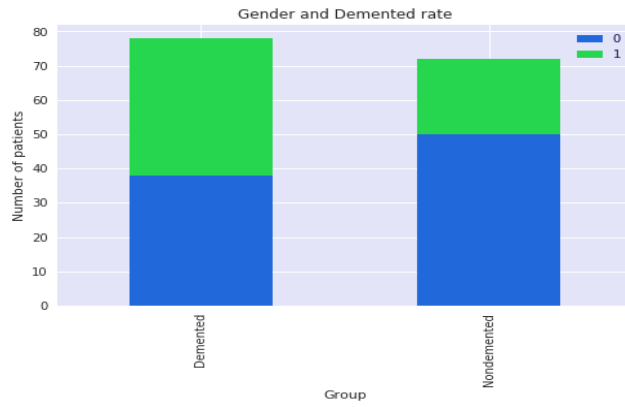
**Figure 1:** Flow of the methodology

The OASIS [11] project includes the cross-sectional and longitudinal MRI data of brain images. In this research, the longitudinal MRI data in demented and non-demented older adults is utilized. The dataset consists of medical record of around 150 suspected patients who lies in the age gap of 60 years to 96years. The dataset includes both male and female patients who has right-handed behaviour. The T1 weighted MRI scans are given in as data in OASIS project in which each individual is scanned at least once. Among the 150 subjects, 72 of them are non-demented, 64 of them are demented and the remaining 14 subjects are converted from non-demented to demented during the

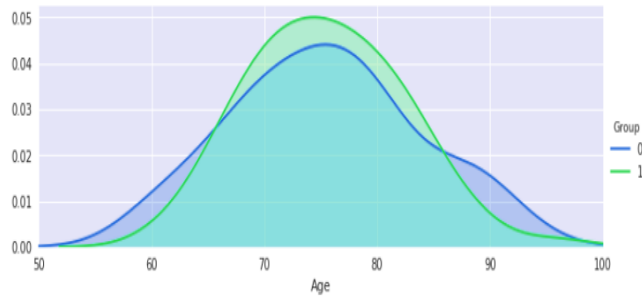
study.

## Data exploration

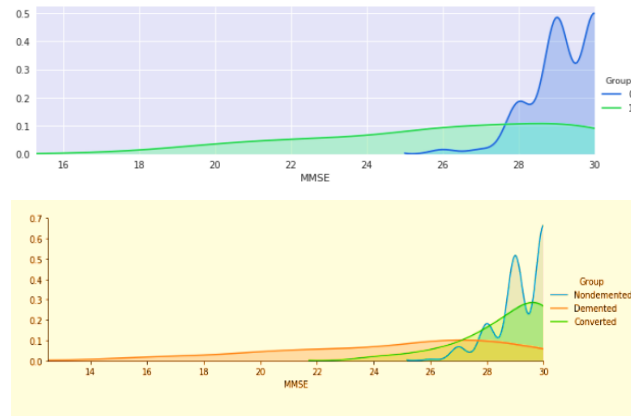
The data analysis finds the relationship between the data using which correlations can be assumed before data extraction. The best suited analysis method can be chosen further by understanding the nature of the data. Initially, the maximum, minimum and average values for all the features in the dataset are determined. The graphical representation of the inferences is shown in Fig 2(a) (b) (c) and (d) respectively.



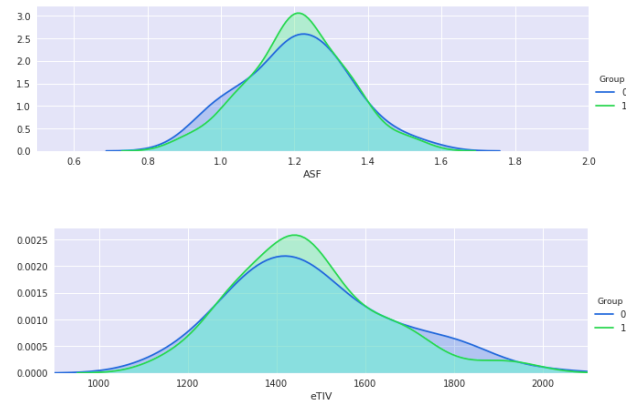
**Figure 2(a):** Gender comparison for demented and non-demented group

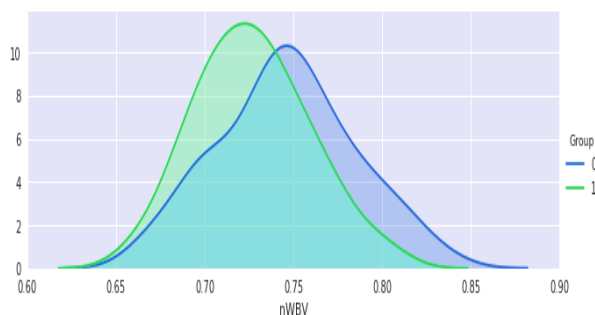


**Figure 2(b):** Age group concentration of demented patients



**Figure 2(c):** Peak values of MMSE scores for non-demented group





**Figure 2(d):** Analysis of Clinical Parameters including Atlas scaling Factor, Normalized Whole Brain Volume and Estimated Total Intracranial Volume

The following inferences are made through data exploration

- On the analysis of the gender and group features having demented and non-demented values, it is predicted that men are more probable to the dementia when compared to women.
- The analysis of the MMSE values shows that the non-demented group has higher MMSE values ranging from 15.3 to 30, when compared to the demented group.
- The normalized brain volume is lower in the demented group due to the shrinkage of tissues with the progress of AD, when compared to the non-demented group.
- The dementia affects the patients in an age group of 70 to 80 years and a few patients of 90 years.
- The survival rate of the patients suffering from dementia is guessed to be low. Hence, there are only few of patients above 90 years old.

**Data Pre-processing and Feature Scaling**

The dataset is preprocessed to remove the null values and to eliminate the inappropriate values for further processing. The Type I and Type II errors arise due to the missing values which in turn reduces the statistical power. In this OASIS longitudinal MRI dataset, 8 missing values are found in the SES feature. The imputation is used to fill the missing values of SES attributes which has discrete values. Each feature in the dataset is scaled to a specific range using the min-max scaling technique, which is one of the normalization techniques. The min-max normalization technique is preferred as it transforms the data in a linear fashion that helps in sustaining the relationship within the original data.

A predefined range can be fixed in min-max scaling, so that the values will not exceed the specified boundary.

The general equation of min-max normalization is given as

$$D' = \left( \frac{D - \text{Minimum value of } D}{\text{Maximum value of } D - \text{Minimum value of } D} \right) \times (B_1 - B_2) + B_2$$

(1)

Where,

D' - Normalized data

D - Original data

B<sub>1</sub> and B<sub>2</sub> - Boundaries.

**5-fold cross validation**

The dataset is segregated into training and testing dataset in such a way that no two records overlap using 5-fold cross-validation technique. The performance of a machine learning model can be evaluated by the cross-validation technique. The dataset is shuffled in a random manner and it is split into 5 groups. By considering one set as test set and the other 4 sets as training set, a model is trained and evaluated. This evaluation is carried out on all the groups to summarize the evaluation scores of the model.

**Hyperparameter Optimization**

The process of selecting optimal parameters to enhance the results of learning algorithms is termed as hyperparameter optimization. The parameters from the hyperparameter search space can be identified using two common strategies namely, grid search or random search. The configuration of hyperparameter tuning in a specific dataset D is given as shown in the equation.

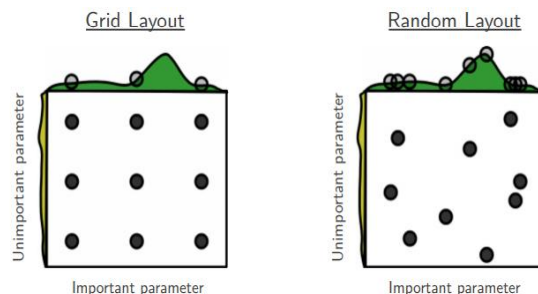
$$\varphi^{(D)*} = \arg \arg R^{(D)}(\varphi) \quad \varphi \in O \tag{2}$$

where R is the risk factors considered for choosing the hyperparameters from the dataset.

Grid search is a conventional method for optimizing the hyperparameters which continues to search in a selected search space repeatedly. As the search is exhaustive it requires large time consumption and computational complexity. The combination of important and unimportant parameters is processed to attain the optimized parameters both in case of grid and random search. The grid layout chooses three unique places to test the nine trials in the dataset, whereas in random search the unique values are explored from all the nine trials. The grid search fails in the case of large dataset. The high dimensional dataset results in inefficiency for the grid search even with good granularity. The search dimensions are highly exponential to the waster trials in the grid search. The random, search is utilized so that

the subspace undergoes efficient search without the wastage of trials. Fig 3 shows the random and grid layout with sample

trials in the search space.



**Figure 3:** Sample trials of grid and random layout in the search space

### Boruta Feature Selection for Random Forest Classification

The random forest classification is enriched with a wrapper package of R, namely, Boruta algorithm. A number of unbiased classifiers that are weak in nature are utilized for classification based on the voting principle. Various bagging samples are applied to train the individual tree before the training phase. The different attributes undergo permutation

to identify the accuracy error using which the feature importance can be measured. The mean and standard deviation along with the Z-score is computed for all the features. The original values of the attributes in the dataset are shuffled to predict a shadow attribute for supporting the Z-score. The correlation and random fluctuations are reduced by grouping the results for ensemble weak classifiers.

---

#### Boruta Algorithm for Feature Selection in Random Forest Classification

---

Step 1: The shadow features are added to the information system

Step 2: The correlations are eliminated by shuffling the attributes in the system

Step 3: The random forest classifier is executed and the calculated Z scores are recorded

Step 4: The score of the attribute better than the maximum Z score of the shadow attributes is set as hit attribute.

Step 5: The attributes whose score is below the value of maximum Z score of shadow attributes is eliminated as insignificant attributes.

Step 6: The remaining attributes are considered as features of importance.

Step 7: The shadow attributes are eliminated.

Step 8: The process is iterated till the feature importance of all the attributes are computed.

---

### ML Based Classification Models

#### Logistic Regression

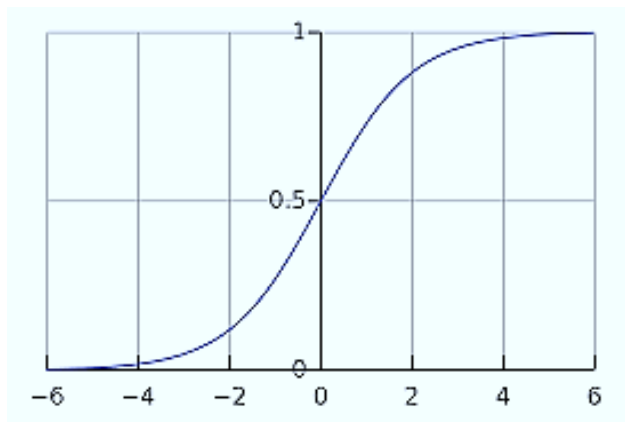
Logistic regression is one of the significant statistical model that facilitates binary classification of dataset based on probabilities. Even the imbalanced dataset works well with logistic regression. A logistic regression model with 'm' number of records and 'f' number of parameters with binary outcomes can be represented as  $L \in LR^{m \times f}$ . The binary outcome is represented as Y in which it has two values  $y_i = 1$  or  $y_i = 0$ . The value 1 represents the positive class which represents the demented patients and 0 represents negative class which indicates the non-demented patients. The

objective of a model 'L' is to classify the patients either under demented or non-demented category. The probability of outcome is based on the regression coefficients and the variables as shown below:

$$\log \log \left( \frac{P(i)}{1-P(i)} \right) = \gamma_0 + \gamma_1 x_1 + \dots + \gamma_n x_n \quad (3)$$

Where,  $P(i)$  is the event probability and  $\gamma$  represents the regression coefficients.

Instead of fitting a line to the data, the logistic regression fits an 'S' shaped curve termed as logistic function.



**Figure 4:** Logistic Response Function

The logistic response function in Fig. 4 is to identify the demented patients is described as,

$$f[x_i, m] = P(y_i) = \frac{\exp^{x_i m}}{1 + \exp^{x_i m}} = \frac{1}{1 + \exp^{-x_i m}} \tag{4}$$

The positive response in the logistic transformation is  $\delta = f(p_i) = \log \log \left( \frac{p_i}{1-p_i} \right) = x_i m$

$$\tag{5}$$

The matrix representation of the logistic function is defined as

$$\delta = X \gamma, \delta = X m \tag{6}$$

The joint probability density function of the data is use to estimate the maximum likelihood method and the likelihood function is given as

$$ML(m) = \prod_{i=1}^k (p_i)^{y_i} (1 - p_i)^{1-y_i} = \prod_{i=1}^k \left( \frac{\exp^{x_i m}}{1 + \exp^{x_i m}} \right)^{y_i} \left( \frac{1}{1 + \exp^{x_i m}} \right)^{1-y_i} \tag{7}$$

By taking log on both sides,

$$\log ML(m) = \sum_{i=1}^k \left( y_i \log \left( \frac{\exp^{x_i m}}{1 + \exp^{x_i m}} \right) + (1 - y_i) \log \left( \frac{1}{1 + \exp^{x_i m}} \right) \right) \tag{8}$$

The logistic regression with inverse regularization strength influences the training time by optimizing the feature weights. Overfitting is a scenario where there are large number of features and less amount of data. The problem of over fitting can be resolved using regularization in order to get highly accurate prediction results. For each feature, the magnitude will increase by applying a penalty via regularization. The error difference of the actual and predicted values is minimized.

### Support Vector Machine

The data points that belong to different classes are grouped in a hyperplane and the margin between the classes are maximized to indicate the segregation. It can have various types of kernels, among which linear kernel is the best option in case of high dimensional data. The spatial pattern which includes the voxels and its weights are used to train the algorithm. A 5-fold cross validation is uses to split the training and testing data that alters the groups in each out of 5 trails. The support vectors are obtained from the hyperplane that classifies the datasets. A straight line segregates the datasets into two groups that lie opposite to each other in a two-dimensional plane. The hyperplane should be noise resistant and it should generalize the data in an accurate manner. Hence, the process of selection an optimal hyperplane is a critical task. The optimal hyperplane and the maximum margin between two groups are shown in Fig 5.

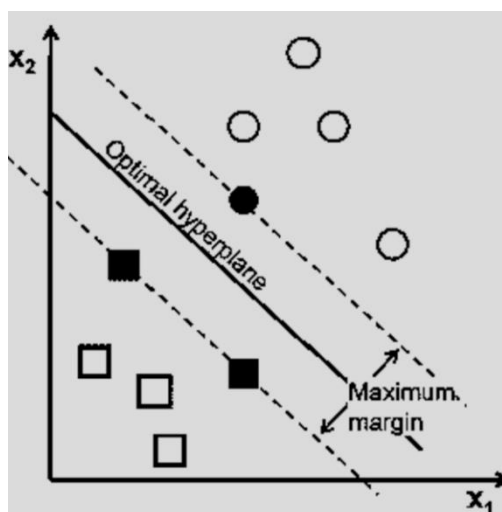


Figure 5: SVM Hyperplane

The hyperplane adapts to the straight-line equation  $y=mx+c$ , where  $m$  is the slope and  $c$  is the constant. The remodeling of line equation gives  $mx-y+c=0$ , in which the coefficients are  $(m,-1)$  and the variables are  $(x,y)$ . Let the coefficients be termed as  $\delta$  and the variables be  $X$  then the hyperplane equation is described as

$$\delta \cdot X + c = 0 \tag{9}$$

The hyperplane having large geometric margin  $GM$  will be considered significant and it plays a vital role in the estimation of  $\delta$  and  $c$  of the hyperplane.

$$GM \text{ where } \pi_i \geq GM, i = 1 \dots n$$

(10)

Let us say  $GM=FM/(||l||)$ , where  $FM$  is the functional margin and  $l$  is the length of the vector, then the equation can be written as

$$GM \text{ where } fm_i \geq FM, i = 1 \dots n$$

(11)

The boosting of  $GM$  by rescaling the values of  $\delta$  and  $c$  do not have any impact on the result. Let the value of  $FM=1$  then the equation is described as

$$\frac{1}{||l||} \text{ where } fm_i \geq 1, i = 1 \dots n$$

(12)

The maximization and the minimization issues are proportionate to each other hence can be defined as

$$\frac{1}{||l||} \text{ where } fm_i \geq 1, i = 1 \dots n$$

(13)

The minimization equation can be further optimized as

$$\frac{1}{2} ||l||^2 \text{ where } y_1 = (\delta x + c) - 1 \geq 0, i=1 \dots n$$

(14)

In the SVM classification, the hypothesis function is represented as

$$H(x) = \begin{cases} 1 & \text{if } \delta \cdot x \geq 0 \\ -1 & \text{if } \delta \cdot x < 0 \end{cases}$$

(15)

### Decision Tree Classification

Decision tree is a directed rooted tree that splits a single decision into two or more paths. The subspaces are further divided until a discrete condition is met. The partition in the instance space is done based on the values of attributes. The target values are represented in the leaf nodes of the tree that has a corresponding probability value. The attributes are given in the nodes whereas the values in its branches. The pruning and stopping criteria are utilized to reduce the complexity of the decision tree. The number of leaves, nodes, attributes and depth of the tree are the important criteria for controlling the complexity.

The dataset is constructed as a decision tree to identify a suitable solution with reduced error. The depth of the tree is reduced by reducing the number of nodes. The decision tree solves the classification problem using greedy strategy which splits the dataset into several subsets thereby increasing the tree depth. The major phases in decision tree classification are induction and pruning. The tree construction based on the splitting condition is done in the induction phase. The over fitting is resolved by pruning phase which in turn enhances the accuracy of classification. The decision tree with Gini score of OASIS dataset is depicted in Fig 6.

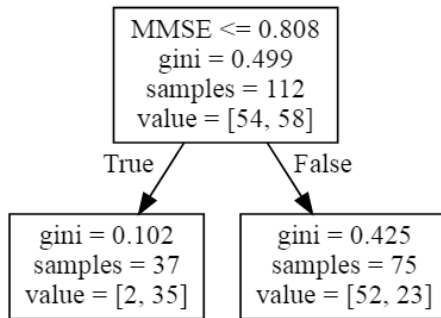


Figure 6: Decision tree from OASIS dataset

The subset of the tree in the internal nodes is denoted as  $T_i$  in which the right and left nodes are denoted as  $T_i^r$  and  $T_i^l$  respectively. The following conditions are satisfied by a decision tree.

- The decision tree is the union both right and left subtrees

$$T_i = T_i^r \cup T_i^l$$

- The intersection of right and left subtree is null

$$T_i^r \cap T_i^l = \emptyset$$

The Gini index is used to partition the decision tree into left and right subspaces. The impurity of the dataset is estimated using Gini index that ranges from 0 to 1. The value 0 represents purity whereas 1 represents impurity. Gini index for a dataset with n number of attributes is defined as

$$G(D) = \sum_{i=1}^n p(X = x_i) \{1 - \sum_{j=1}^m p(Y = y_j / X = x_i)\}^2$$

(16)

in which, the number of classes is represented by ‘n’ and the attribute values are denoted by ‘m’. X is the attributes in the dataset and Y is the output. The impurity of the dataset is reduced by the objective function as shown

below.

$$GI(X = x_i = \{G_X(D)\})$$

(17)

### Random Forest Classification

RF is defined as a general principle of randomized ensembles of decision trees. The general structure of RF is depicted in Fig 7. Generally, RF is built by the recursive partitioning of binary tree into similar nodes. The similarity of the child node is enhanced by the inheritance of data from the parent node. The original data is sampled by bootstrap sampling to generate a large number of trees to grow RF. Each tree in the RF generates a response with respect to the set of predictor values provided as input to those trees. The missing values in the dataset can be easily managed by the RF. Bootstrap samples are derived from the training dataset. Each sample is constructed as tree in which ‘n’ random features are selected to split a node into two children. The tree is subdivided until the node size becomes minimum value. An Out-Of-Bag (OOB) error rate is computed using the data that is not present in the bootstrap sample.

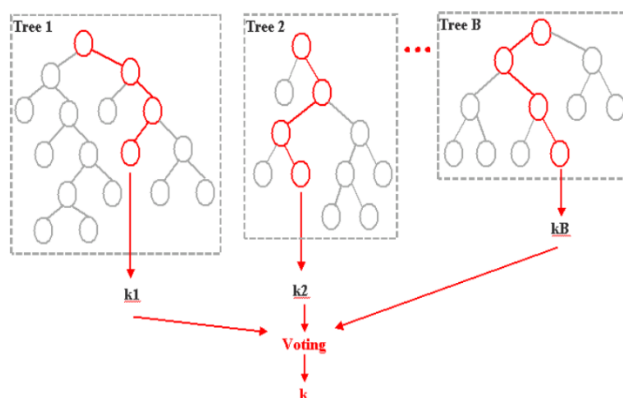


Figure 7: General Architecture of RF algorithm

The probability of a node is the ratio of number of samples that move towards it to the total number of samples in the tree. The node with minimum impurity and higher probability value is considered as the important feature.

$$n(i) = w_i I_i - w_l(i) I_l(i) - w_r(i) I_r(i)$$

(18)

Where,  $I_i$  is the impurity of node  $I(i)$  and  $I_r(i)$  is the impurity value of left and right subtree.

$w_i$  is the weight of samples.

$$I(f) = \frac{\sum_{i: \text{node } i \text{ splits on the } i^{\text{th}} \text{ feature of } n} f_i}{\sum_{k \in \text{all nodes } n} f_k} \tag{19}$$

The feature importance is normalized as a value that ranges from 0 to 1 which is described as

$$Z(f_i) = \frac{f_i}{\sum_{i \in \text{all features}} f_i} \tag{20}$$

In random forest, the feature importance is the ratio of the sum of feature importance of all the trees to the total number of trees in the forest.

$$RF(f_i) = \frac{\sum_{i \in \text{all features}} f_i}{T} \tag{21}$$

**AdaBoost Classification**

The classification performance is enhanced by combining various weak learners together in AdaBoost algorithm as shown in Fig 8. The weight of the training samples is

adjusted adaptively to improve the classification. Let

$Z = \{(x_1, y_1), (x_2, y_2), \dots (x_n, y_n)\}$  be the training samples each class has a label represented as  $y_i \in \{-1, 1\}$ . The weak decision tree which has a minimum error rate of below 50% is considered as weak learners in AdaBoost. In the sampling step, a set of samples in 'n' iterations are chosen to train various classifiers and the error rate are estimated. The weight of the samples is computed in each iterations, where the weights are same at the initial stage.

The weight of samples at the initial stage is shown as

$$w^i = [w_1^i, \dots, w_N^i], w_j^i \in [0, 1] \tag{22}$$

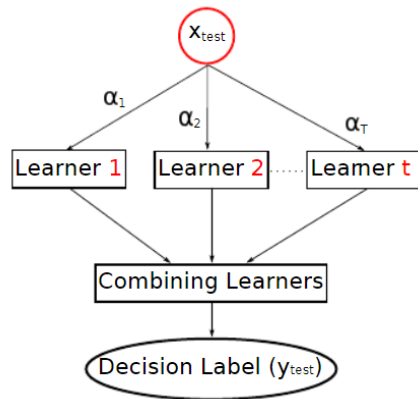
The sum of weights is equal to 1 as represented in the following equation

$$\sum_{j=1}^N w_j^i = 1 \tag{23}$$

$w_j^i$  is the weight of all the samples in the iteration.

The weight of the samples are updated

$$w_i^{n+1} = \frac{w_i^n}{Z_n} \begin{cases} e^{at} & \text{if } x_i \neq h_n(y_i) \\ e^{-at} & \text{else} \end{cases} \tag{24}$$



**Figure 8:** Combination of weak learners in AdaBoost

The correct classification of samples provides positive sign where  $x_i = h_n(y_i)$  and the incorrect classification gives negative sign where  $x_i \neq h_n(y_i)$ .

Hence, the updated weights equation can be reformulated as

$$w_i^{n+1} = \frac{w_i^n}{Z_n} (-\alpha_n x_i h_n(y_i)) \tag{25}$$

Weight of the samples in the upcoming iteration is represented as  $w_i^{n+1}$  and the term  $\alpha_n = \frac{1}{2} \ln \ln \frac{1-\epsilon_t}{\epsilon_t}$  in

which the training error is denoted by  $\epsilon_t$ .

The normalized form of updated weights are described as

$$Z_n = \sum_{i=1}^N w_i^n e^{-\alpha_n x_i h_n(y_i)} \tag{26}$$

The test sample is passed on to the combined weak classifiers which provide a perfect decision as the result. The final decision estimation is defined as

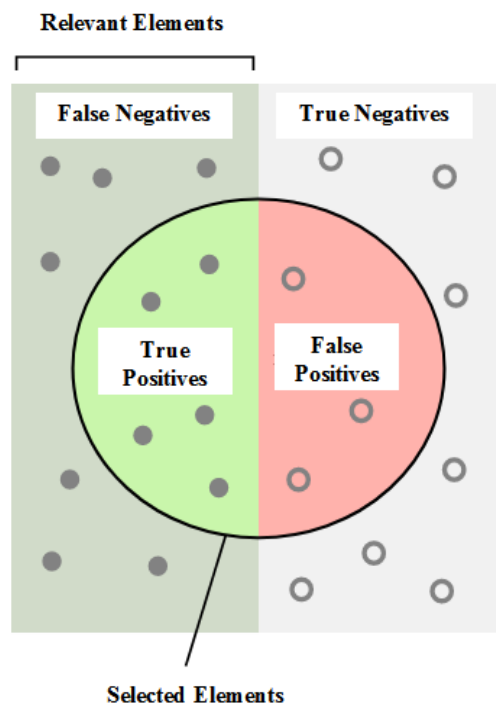
$$D_{final} = \text{sign}\left(\sum_n \alpha_n h_n(x_{test})\right)$$

(27)

### Performance Evaluation

The performance of the classifiers is evaluated using the OASIS dataset in terms of Accuracy, Recall, AUC and feature importance. During the medical diagnosis of the non-life-threatening diseases like neurodegenerative

diseases, it is highly important to have a high True Positive Rate (TPR) so that the early identification of the patients with AD could be possible. Also, the False Positive Rate (FPR) should be as low as possible, to avoid the misdiagnosis of the healthy adult as a demented patient.



**Figure 9:** True and False Positives of the selected items

In this case, AUC is the ideal choice. Fig 9 depicts the true and false positives of the selected items. Accuracy is used to determine the relevancy of the selected items and recall is used to evaluate the selection rate of the relevant items.

- True Positives (TP): Correct prediction of the demented patients as demented.
- True Negatives (TN): Correct prediction of the non-demented patients as non-demented.
- False Positives (FP): Incorrect prediction of the demented patient as non-demented. This is also known as “Type I error”.
- False Negatives (FN): Incorrect prediction of the non-

demented patient as demented. This is referred as “Type II error”.

In Logistic Regression, the Input values (x) are combined linearly using the weights or coefficient values to predict an output value (y) as binary values (either 0 or 1) despite of the numeric value as obtained from the linear regression procedure. Table 2 presents the Logistic Regression results with Imputation and dropping the missing items on the OASIS Dataset. From the table, it is clearly understood that the Logistic Regression with Imputation yields higher accuracy and recall than the Logistic Regression by dropping missing values.

**Table 2:** Results for Logistic Regression on OASIS dataset

Logistic Regression Results on OASIS Dataset				
	Accur acy	Best Regularizati on Parameter C	Test Accuracy with C	Test Recall with C
Logistic Regression with Imputation	75.41	10	78.94	75
Logistic Regression by dropping missing Values	70.09	10	75	70

**Table 3:**  
Analysis of SVM

Performance Analysis of Decision Tree Classifier	
Validation Accuracy	77.85
Test Accuracy	81.57
Test Recall	0.65
Best parameter for max Depth	1

Performance  
Classifier

Table 3 shows the performance analysis of the SVM classifier. The validation accuracy of the SVM classifier is 77.18%, test accuracy is 81.57% and test recall is 82.22%. Table 4 depicts the performance analysis of the Decision Tree classifier. The validation accuracy of the Decision Tree classifier is 77.85%, test accuracy is 81.57% and test recall is 0.65%. Table 5 shows the Feature Importance in

the Decision Tree computed using Gini Index. Gini index measures the degree or probability of a particular item being wrongly classified when it is randomly chosen. The Gini index degree varies between 0 and 1, where 0 denotes that all the items belong to a certain class, and 1 denotes that the items are randomly distributed across various classes. The Mini-Mental State Examination (MMSE) value is 1.0.

**Table 4:** Performance Analysis of Decision Tree Classifier

Performance Analysis of SVM Classifier	
Validation Accuracy	77.18
Best parameter for C	100
Best parameter for Gamma	0.1
Best parameter for Kernel	Radial Basis Function (rbf)
Test Accuracy	81.57
Test Recall	82.22

**Table 5:** Feature Importance in Decision Tree computed using Gini Index

Feature Importance in Decision Tree computed using Gini Index	
M/F	0.0
Age	0.0
EDUC	0.0
SES	0.0
MMSE	1.0
eTIV	0.0
nWBV	0.0
ASF	0.0

**Table 6: Performance Analysis of Random Forest**

Performance Analysis of Random Forest	
Validation Accuracy	79.63
Test Accuracy	84.21
Test Recall	80
Best parameter for M,d, m	14, 5, 7

**Table 7: Feature Importance in Random Forest**

Feature Importance in Random Forest	
M/F	0.035
Age	0.095
Educ	0.062
SES	0.06
MMSE	0.4
eTIV	0.075
nWBV	0.146
ASF	0.124

**Table 8: Performance Analysis of AdaBoost**

Performance Analysis of AdaBoost	
Validation Accuracy	77.85
Test Accuracy	84.21
Test Recall	65.1
Best parameter for M	2
Best Parameter for LR	0.0001

Table 6 presents the performance analysis of the Random Forest. The validation accuracy is 79.63%, test accuracy is 84.21% and test recall is 80%. The feature importance of the Random Forest algorithm is shown in Table 7. The

MMSE value is 0.4. Table 8 shows the performance analysis of the Adaboost algorithm. The validation accuracy is 77.85%, test accuracy is 84.21% and test recall is 65.1%. The feature importance of the Adaboost algorithm is depicted in Table 9. The MMSE value is 0.142.

**Table 9: Feature Importance in AdaBoost**

Feature Importance in AdaBoost	
M/F	0.071
Age	0.142
Educ	0.21
SES	0.074
MMSE	0.142
eTIV	0.214
nWBV	0.146
ASF	0.01

**Table 10: Comparative Analysis of Accuracy, Recall and AUC values**

ML Model	Accuracy	Recall	AUC
Logistic Regression (w/ imputation)	78.95	70.00	79.16
Logistic Regression (w/ dropna)	75.00	70.00	70.00
SVM	81.58	70.00	82.22
Decision Tree	81.58	65.00	82.50
Random Forest	84.21	80.00	84.44
AdaBoost	84.21	65.00	84.50

Table 10 illustrates the comparative analysis of the accuracy, recall and AUC values. Random Forest and AdaBoost algorithms yield maximum accuracy of about 84.21% than the Logistic Regressions, SVM and Decision Tree classifiers. The recall value of the Random Forest Algorithm is 80%. The AUC values of the Random Forest and AdaBoost algorithm

are higher when compared to the Logistic Regressions, SVM and Decision Tree classifiers. The comparative analysis shows that the Random Forest Algorithm is the optimal choice used for the classification during the Prognosis of the Alzheimer's disease.

**Table 11: Boruta Algorithm for Feature Selection in Random Forest**

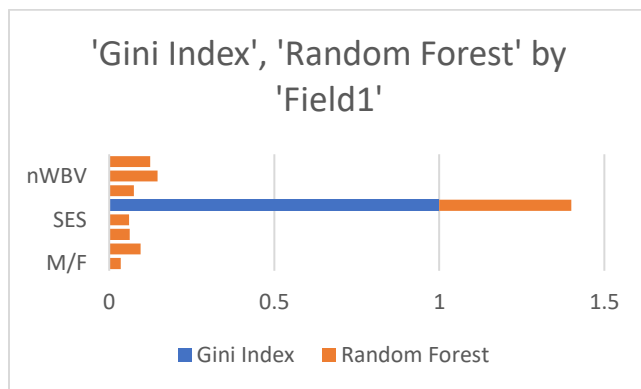
Feature Ranking using Boruta Algorithm		
Rank	Features	Score
1	MMSE	0.27
2	nWBC	0.19
3	ASF	0.14
4	eTIV	0.12
5	Age	0.09
6	EDUC	0.09
7	SES	0.06
8	Gender	0.05

**Table 12: Feature Importance Ranking using Random Forest and Adaboost Algorithm**

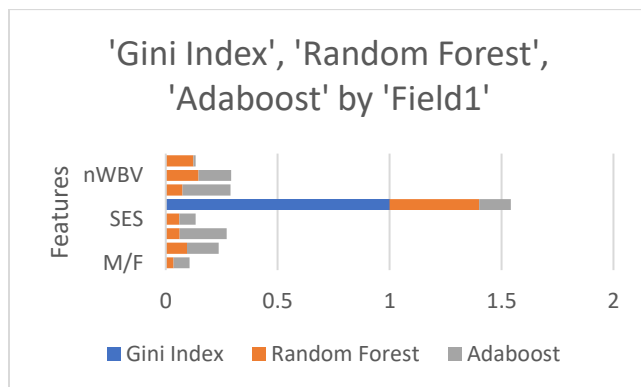
	Feature Importance		
	Gini Index	Random Forest	Adaboost
M/F	0.0	0.035	0.071
Age	0.0	0.095	0.142
EDUC	0.0	0.062	0.21
SES	0.0	0.06	0.074
MMSE	1.0	0.4	0.142
eTIV	0.0	0.075	0.214
nWBV	0.0	0.146	0.146
ASF	0.0	0.124	0.01

From the above discussed performance evaluation, the Random Forest classification model attained an accuracy of 84.21% and the Support Vector Machine classification model reached a level of 82.58%. Considering these two models for further study, the hyperparameter tuning process is applied to enhance the classification accuracy. The random search CV is applied to select the features of SVM classification and the Boruta algorithm of R package is utilized for RF

classification. The features ranked in accordance with Boruta algorithm is depicted in Table 11 and with other algorithms in Table 12 which shows MMSE has high significance. The comparison results of feature ranking by Gini Index and Random Forest method are illustrated in Fig 11 which shows MMSE score has the highest ranking. Fig 12 depicts the feature selection ranking using Gini Index, Random Forest and Adaboost algorithms.



**Figure 11:** Comparison results of feature ranking



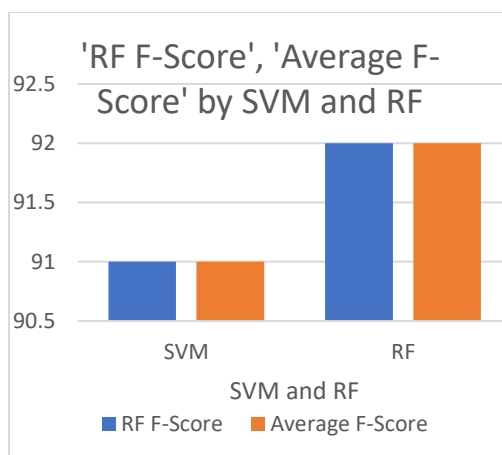
**Figure 12:** Feature selection method by Gini Index, Random Forest and Adaboost methods

The Boruta algorithm is incorporated in Random Forest classification for computing the feature importance. All the attributes are given a chance to find a compatible feature for selection. The depth of the tree is set to 5 for attaining the minimal set of optimal features from the dataset. The maximum iteration is set to 100 and the alpha value is considered 0.05. The performance evaluation of SVM with Random search-based feature selection and RF with Boruta

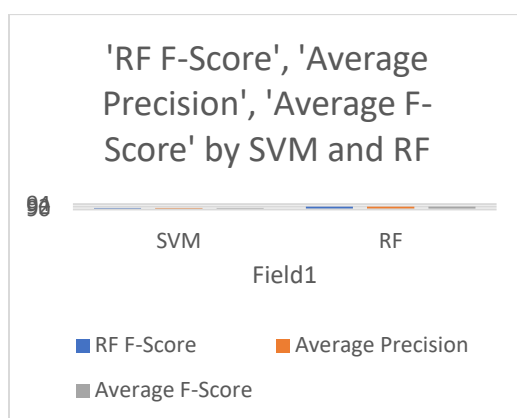
algorithm based feature selection is shown with their graphical representation of the various comparison results depicted in Fig 13 to Fig 14. The performance analysis results of SVM, Decision Tree, Adaboost and Random Forest classification algorithms are depicted in the below Table 13. From Table 13, it is observed that Random Forest algorithm holds the maximum accuracy score. The confusion matrix of RF and SVM is represented in Fig 15 and Fig 16 respectively.

**Table 13:** Performance Analysis score

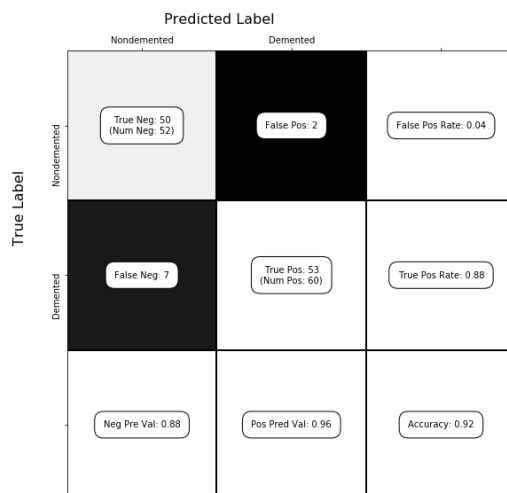
Performance Analysis	SVM	Decision Tree	Random Forest	Adaboost
Validation Accuracy	77.18	77.85	79.63	77.85
Best parameter for C	100			
Best parameter for Gamma	0.1			
Best parameter for Kernel	Radial Basis Function (rbf)	1		
Test Accuracy	81.57	81.57	84.21	84.21
Test Recall	82.22	0.65	80	65.1
Best parameter for max Depth		1		
Best parameter for M,d, m			14, 5, 7	
Best parameter for M				2
Best Parameter for LR				0.0001



**Figure 13:** Comparison results of F-score and Average F-score for SVM and RF methods



**Figure 14:** Comparison results of Average Precision and Average F-score for SVM and RF methods



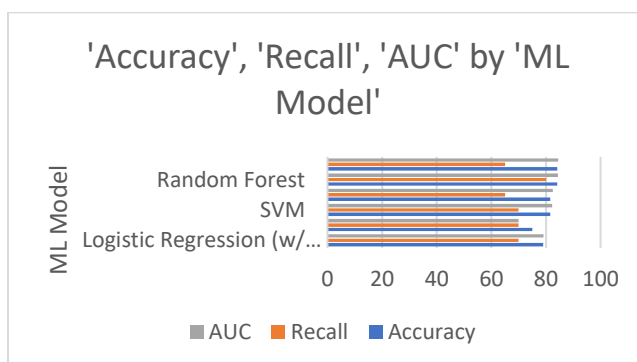
**Figure 15:** Confusion Matrix for Random Forest Classification with Boruta Feature Selection

		Predicted Label		
		Nondemented	Demented	
True Label	Nondemented	True Neg: 50 (Num Neg: 52)	False Pos: 2	False Pos Rate: 0.04
	Demented	False Neg: 8	True Pos: 52 (Num Pos: 60)	True Pos Rate: 0.87
		Neg Pre Val: 0.86	Pos Pred Val: 0.96	Accuracy: 0.91

**Figure 16:** Confusion Matrix for Support Vector Machine Classification with Random Search Feature Selection

The precision and recall of Dementia group identified using SVM classification is 96% and 86% respectively. For non-dementia group the precision resulted in 87% whereas the recall is 96%. Both the groups attained an F-score of 91%. Similarly, the RF classification is evaluated separately for demented and non-demented group which provided an F-score of 92%. There is a slight variation in recall and precision when compared to SVM classification. The recall and precision value of demented group is 88% and 96%

whereas for non-demented group 96% and 88% respectively. Sensitivity is considered as the indicator for the evaluation of model. On the comparison of overall performance of the model, the RF classification is slightly better with 87% and 98% for dementia and non-dementia groups respectively. Whereas, the SVM resulted in 87% for dementia and 96% for non-dementia groups. The comparison results of Accuracy, Recall and Area Under Curve values by the various algorithms is shown in Fig 17.



**Figure 17:** Accuracy, Recall and AUC score by various ML models

## CONCLUSION

In this work, a framework based on the machine learning approach for the diagnosis of AD is devised to accurately classify the longitudinal brain MRI data of the elderly subjects into two categories: demented and non-demented. The proposed framework is completely automated that ensures better classification for the diagnosis of AD. The objective of this work is to find the optimal algorithm for classifying the MRI data with better accuracy and recall. Despite of the traditional classification method, the proposed pipeline approach provides a combination of multiple features facilitating the detection of the disease. The hyperparameter tuning and Boruta algorithm for feature selection improved the performance of SVM and RF classification. The efficiency of the machine learning algorithm is evaluated using the accuracy, recall and AUC values on the OASIS-2 dataset. From the comparative analysis, it is concluded that the Random Forest Algorithm

is the optimal classifier for the Prognosis of Alzheimer's disease.

## REFERENCES

- Heneka, Michael T., et al. "Neuroinflammation in Alzheimer's disease." *The Lancet Neurology* 14.4 (2015): 388-405.
- Ewers, Michael, et al. "Neuroimaging markers for the prediction and early diagnosis of Alzheimer's disease dementia." *Trends in neurosciences* 34.8 (2011): 430-442.
- Alzheimer's Association, William Thies, and Laura Bleiler. "2013 Alzheimer's disease facts and figures." *Alzheimer's & dementia* 9.2 (2013): 208-245.
- Alzheimer's Association. "2016 Alzheimer's disease facts and figures." *Alzheimer's & Dementia* 12.4 (2016): 459-509.
- Alzheimer's Association. "2015 Alzheimer's disease facts and figures." *Alzheimer's & Dementia* 11.3 (2015): 332-384.
- Sperling, Reisa A., et al. "Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease." *Alzheimer's & dementia* 7.3 (2011): 280-292.

- Reisberg, Barry, et al. "Behavioral symptoms in Alzheimer's disease: phenomenology and treatment." *The Journal of clinical psychiatry* (1987).
- Mueller, S. G., N. Schuff, and M. W. Weiner. "Evaluation of treatment effects in Alzheimer's and other neurodegenerative diseases by MRI and MRS." *NMR in Biomedicine: An International Journal Devoted to the Development and Application of Magnetic Resonance In vivo* 19.6 (2006): 655-668.
- Gianfrancesco, Milena A., et al. "Potential biases in machine learning algorithms using electronic health record data." *JAMA internal medicine* 178.11 (2018): 1544-1547.
- Rajkomar, Alvin, Jeffrey Dean, and Isaac Kohane. "Machine learning in medicine." *New England Journal of Medicine* 380.14 (2019): 1347-1358.
- <https://www.oasis-brains.org/#data>
- Marcus, Daniel S., et al. "Open Access Series of Imaging Studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults." *Journal of cognitive neuroscience* 19.9 (2007): 1498-1507.
- Azam, Muhammad Adeel, et al. "Analysis of the MIDAS and OASIS Biomedical Databases for the Application of Multimodal Image Processing." *International Conference on Intelligent Technologies and Applications*. Springer, Singapore, 2019.
- Savio, Alexandre, and Manuel Graña. "Supervised classification using deformation-based features for Alzheimer's disease detection on the OASIS cross-sectional database." *KES*. 2012.
- Aghili, M., et al. "Prediction Modeling of Alzheimer's Disease and Its Prodromal Stages from Multimodal Data with Missing Values." *International Journal of Medical and Health Sciences* 13.2 (2019): 36-40.
- Soliman, Sarah A., et al. "Exploiting of Machine Learning Paradigms in Alzheimers Disease." *International Journal of Psychiatry and Psychotherapy* 4 (2019).
- Neffati, Syrine, et al. "An improved machine learning technique based on downsized KPCA for Alzheimer's disease classification." *International Journal of Imaging Systems and Technology* 29.2 (2019): 121-131.
- Toshkhujaev, Saidjalol, et al. "Automatic classification of Alzheimer's disease using different neuroimaging tools".(2019): 58-61.
- Karthiga, M., Sountharajan, S., Nandhini, S. S., & Sathis Kumar, B. (2020, May). Machine Learning Based Diagnosis of Alzheimer's Disease. In *International Conference on Image Processing and Capsule Networks* (pp. 607-619). Springer, Cham.
- Afzal, Sitara, et al. "A Segmentation-Less Efficient Alzheimer Detection Approach Using Hybrid Image Features." *Handbook of Multimedia Information Security: Techniques and Applications*. Springer, Cham, 2019. 421-429.
- Sountharajan, S., E. Suganya, M. Karthiga, and C. Rajan. "Automatic glioblastoma multiforme detection using hybrid-SVM with improved particle swarm optimisation." *International Journal of Biomedical Engineering and Technology* 26, no. 3-4 (2018): 353-364.
- Khan, Afreen, and Swaleha Zubair. "An Improved Multi-Modal based Machine Learning Approach for the Prognosis of Alzheimer's Disease." *Journal of King Saud University-Computer and Information Sciences* (2020).
- Marghalani, Bashayer Fouad, and Muhammad Arif. "Automatic Classification of Brain Tumor and Alzheimer's Disease in MRI." *Procedia Computer Science* 163 (2019): 78-84.
- Richhariya, B., et al. "Diagnosis of Alzheimer's disease using universum support vector machine based recursive feature elimination (USVM-RFE)." *Biomedical Signal Processing and Control* 59 (2020): 101903.
- Madan, Yamini, Iswarya Kannoth Veetil, Gopalakrishnan EA, and Soman KP. "Synthetic Data Augmentation of MRI using Generative Variational Autoencoder for Parkinson's Disease Detection." In *Evolution in Computational Intelligence*, pp. 171-178. Springer, Singapore, 2022.
- Sudharsan, Deepthi, S. Isha Indhu, Kavya S. Kumar, Lakshaya Karthikeyan, L. Srividhya, V. Sowmya, E. A. Gopalakrishnan, and K. P. Soman. "Analysis of Machine Learning and Deep Learning Algorithms for Detection of Brain Disorders Using MRI Data." In *Artificial Intelligence on Medical Data*, pp. 39-46. Springer, Singapore, 2023.