

Smart Diagnosis: An Ensemble Machine Learning Web Application for Early Detection of Alzheimer's Disease

Yetunde D. Otun*

Department of Computer Science, Redeemer's University, Ede, Osun, 232101, Nigeria

Email: akinwunmi7127@run.edu.ng

ORCID iD: <https://orcid.org/0009-0003-6810-9096>

*Corresponding author

Abosede O. Oguntunde

Department of Computer Science, Redeemer's University, Ede, Osun, 232101, Nigeria

Email: oguntunden@run.edu.ng

ORCID iD: <https://orcid.org/0000-0001-6508-9737>

Samson A. Arekete

Department of Computer Science, Redeemer's University, Ede, Osun, 232101, Nigeria

Email: areketes@run.edu.ng

ORCID iD: <https://orcid.org/0000-0003-1266-0535>

Oluwale B. Olajide

Faculty of Computing, University of Ibadan, Ibadan, Oyo, Nigeria

Email: oolajide4174@stu.ui.edu.ng

ORCID iD: <https://orcid.org/0000-0003-0772-8339>

Benjamin S. Aribisala

Department of Computer Science, Lagos State University, Ojo, Lagos, 102101, Nigeria

Email: benjamin.aribisala@lasu.edu.ng

ORCID iD: <https://orcid.org/0000-0002-6290-1707>

Received: October 5, 2025; Revised: February 13, 2026; Accepted: March 14, 2026; Published: June 8, 2026

Abstract: Alzheimer disease is a chronic neurodegenerative disorder and the primary cause of dementia among the population, which has a huge burden to the patients, their caregivers and the health care system. Timely intervention is necessary to reduce disease progression, facilitate timely intervention and improve the quality of life. But the traditional forms of diagnostic are frequently costly and non-available especially in resource-deficient environments. The research paper proposes an interpretable and cost-efficient machine-learning model that can be used to identify the presence of Alzheimer disease at its early stages based on clinical and demographic metrics based on the Open Access Series of Imaging Studies cross-sectional dataset, which contains 436 participants. The data consists of seven numeric and two categorical variables, whereas the Clinical Dementia Rating was changed into two categories namely demented and non-demented. An extensive preprocessing pipeline was used, which entailed missing value imputation, categorical encoding and elimination of irrelevant variables, as well as class balancing with the Synthetic Minority Oversampling Technique. A number of machine learning models were tested, which comprise Logistic Regression, Support Vector Machine, Random Forest, Gradient Boosting, and Extreme Gradient Boosting. The results show that the highest accuracy of 92% was attained using the model implemented by the ensemble and the tree, with the most accuracy being returned by the Random Forest and the ensemble model. Random Forest, too, had a sensitivity of 95%, whereas Gradient Boosting and Extreme Gradient Boosting had the highest area under the receiver operating characteristic curve of 98%. The models were implemented as a lightweight web application on the Flask framework, which can make real-time predictions and color coded. The system illustrates the possibility of combining interpretable machine learning with web technologies to make it possible to conduct easy and effective early screening of Alzheimer disease under resource-limited healthcare conditions.

Index Terms: Alzheimer's Disease, Clinical Dementia Rating, Machine Learning, SMOTE Analysis, Web Application

1. Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative disorder and the most common cause of dementia worldwide. It currently affects more than 50 million people globally, a figure projected to triple by 2050 as populations age, posing immense challenges for patients, caregivers, and healthcare systems alike [1, 2] in Fig. 1. The condition is characterized by memory decline, cognitive impairment, and behavioral disturbances, which worsen over time and ultimately lead to complete dependence on care. Beyond its devastating personal and familial impact, AD constitutes a major public health and economic burden, accounting for billions of dollars annually in healthcare costs and lost productivity [3].

Early detection of AD is widely recognized as crucial for improving patient outcomes. Timely diagnosis enables patients to access symptomatic therapies, participate in clinical trials, and plan proactively for future care, thereby enhancing quality of life [4]. Furthermore, it allows healthcare systems to initiate interventions that may delay institutionalization and reduce long-term costs. However, despite these advantages, AD is frequently underdiagnosed or diagnosed too late, often because traditional diagnostic methods remain heavily reliant on specialist expertise, advanced neuroimaging, cerebrospinal fluid biomarkers, and detailed neuropsychological testing. While accurate, these approaches are often expensive, time-consuming, and inaccessible in many low- and middle-income settings [5]. Consequently, there is an urgent need for affordable, scalable, and user-friendly tools to support early detection in diverse healthcare environments.

In recent years, machine learning (ML) has emerged as a transformative approach in medical diagnostics. By analyzing complex patterns in large datasets, ML algorithms can generate predictions that rival or even surpass traditional diagnostic techniques in speed and accuracy [6-8]. For Alzheimer's disease, predictive modeling based on demographic, clinical, and cognitive data has shown promise in distinguishing between demented and non-demented individuals, as well as in predicting disease progression [9]. Among the suite of available algorithms, logistic regression has appeal in clinical contexts. Unlike "black box" models such as deep neural networks, logistic regression provides interpretable results, enabling clinicians to understand the relationship between input variables and diagnostic outcomes [10, 11]. This interpretability fosters transparency, trust, and accountability, which are essential when introducing artificial intelligence-driven tools into medical practice.

Equally important is the mode of delivery of such predictive systems. Web-based applications provide a powerful means of translating machine learning models into practical, user-accessible tools. By leveraging lightweight architectures such as Flask for backend processing and HTML/CSS/JavaScript for frontend interaction, predictive models can be deployed in real time and accessed across devices with minimal computational resources. Such applications enable clinicians and even community healthcare workers to input readily available patient data and obtain immediate diagnostic predictions. In addition to enhancing accessibility, these platforms support integration into clinical workflows, reduce reliance on expensive imaging, and extend decision support into low-resource settings [12].

The issue of identifying early onset of Alzheimer disease is tackled in the present project, as the web application based on machine learning will be trained to use both clinical and demographic variables, using the OASIS cross-sectional dataset. The system categorizes individuals into demented and non-demented with first models being founded on the Logistic Regression (LR) because of its transparency and interpretability. Synthetic Minority Oversampling Technique (SMOTE) and other preprocessing techniques are used to deal with class imbalance to enhance sensitivity to minority cases without significantly reducing the overall accuracy.

This paper acknowledges the drawbacks of single models and goes on to explore the concept of an ensemble, which is the use of Logistic Regression, Support Vector Machine (SVM), as well as Gradient Boosting (GB) classifiers. The ensemble takes advantage of the complementary nature of both linear and non-linear models to create probabilistic outputs, which are pooled to come up with strong predictions. This design models the complex and non-linear decision boundaries and preserves clinically relevant characteristics which enhance predictive and interpretative quality. The resultant ensemble is incorporated into an interface based on Flask, which enables real-time, color-coded predictions, which can be used and accessed in a clinical and community healthcare setting. The following research questions and hypotheses were developed: (1) Can machine learning models effectively detect early Alzheimer's disease using clinical data? (2) Which model provides the best balance between accuracy and interpretability? (3) Does ensemble learning improve predictive performance compared to individual models?

Finally, the present research will benefit the sphere of AI-based healthcare interventions by providing a sufficiently methodologically solid, interpretable, and clinically viable instrument of early dementia diagnosis, especially in the settings, where the resources of specialists are scarce. The ensemble modeling and lightweight web implementation are the reasons that depict how machine learning can be used to facilitate accessible and real-time clinical decision-making.

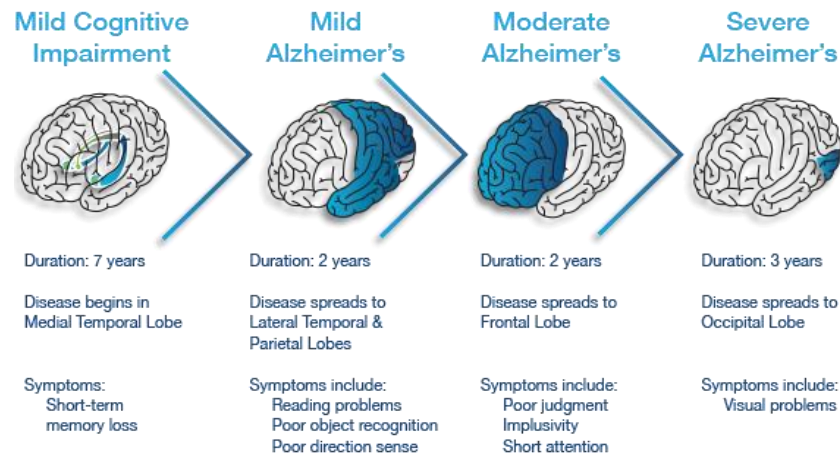


Fig. 1. Alzheimer's disease progression chart (<https://www.mccare.com/education/alzprogression.html>).

2. Literature Review

Previous literature has shown that machine learning algorithms can reliably predict Alzheimer's disease (AD) based on numerical data (e.g., cognitive test scores, demographic data, and clinical assessments). However, in a number of low-resource countries where neuroimaging is expensive and mostly unavailable, clinical data tend to present a more realistic option, although these machine learning techniques are not yet clinically deployable and accessible. Kloppel et al. [13] proposed an automated method of AD diagnosis based on structural MRI and support vector machines (SVMs). In their work, the authors addressed the issue of limited accuracy and subjectivity of traditional diagnostic methods by suggesting a fully automated machine learning model to distinguish between AD and normal aging, as well as other forms of dementia, including frontotemporal lobar degeneration (FTLD). The authors used linear SVM models on grey matter-segmented T1-weighted MRI scans of four groups of patients: cases pathologically diagnosed with AD, cases who were clinically diagnosed as having mild AD, cognitively normal controls, and cases who FTLD had pathologically confirmed. The approach had a maximum accuracy of 96 percent in separating pathologically verified AD patients and controls, 88.9 percent in differentiating mild AD and controls using whole-brain and hippocampal region analysis, and 89.2 percent in discriminating AD and FTLD. Notably, the method was robustly generalized to datasets collected at other centers and on different scanners, with 87.5-96.4% classification rates. Despite promising findings, the study admitted a number of limitations. First, its sample size, especially the FTLD patients, was small and research could not analyze pathological subtypes of FTLD. Second, the mild AD group was not pathologically confirmed, which induces possible misclassification caused by clinical diagnosis inaccuracy. Third, despite its automation, the preprocessing pipeline had to be checked manually on misregistration, which could also be automated. Lastly, although the approach generalized across scanners, specificity was lower when training with younger AD patients and testing with older cohorts, suggesting issues in dealing with heterogeneity between the populations. The authors concluded that their SVM-based MRI classification is a robust, automated, and generalizable tool and can be used in the clinical practice. It has an advantage over manual volumetric measurements since it overcomes observer bias and presents whole-brain data, making it promising aids in early and differential diagnosis of dementia.

Moradi et al. [14] proposed a machine learning framework to address prediction of Alzheimer disease (AD) conversion in mild cognitive impairment (MCI) subjects based on MRI and cognitive data (2015). The researchers solved this problem of early detection by introducing a new MRI-based biomarker developed by using semi-supervised learning and combining it with age and cognitive data in an aggregate biomarker. As opposed to earlier studies, based upon primarily supervised methods, the authors employed a low-density separation (LDS) semi-supervised algorithm, feature selection through regularized logistic regression on AD and normal controls (without using MCI subjects) and extraction of age-related effects on MRI data, to prevent confounding. Their MRI biomarker had a 10-fold cross-validated area under the curve (AUC) of 0.7661 when differentiating between progressive MCI (pMCI) and stable MCI (sMCI). The aggregate biomarker, when used with baseline age and cognitive scores in a random forest classifier, greatly enhanced predictive performance with a 10-fold cross-validated AUC of 0.9020. Notably, MRI alone contributed to the predictive accuracy the most, though combination with cognitive measures and age yielded the strongest results. Limitations of the study include reliance on the ADNI dataset, which may not generalize across diverse populations or clinical settings; exclusion of other biomarkers such as PET and CSF due to invasiveness or cost, potentially limiting multimodal insights; a relatively small number of unlabeled MCI samples, which reduced the benefits of semi-supervised learning; and the arbitrary 3-year cutoff for defining progressive versus stable MCI, which may not fully capture the heterogeneous progression of AD. The researchers concluded that a combination of MRI-based semi-supervised biomarkers alongside age and cognitive tests provides an effective and non-invasive approach to

identifying early at-risk MCI patients converting to AD. This method provides insights into how machine learning could help in early intervention strategies and enhance Alzheimer disease prognosis.

Xiao et al. [15] introduced a multifeatured fusion framework to classify Alzheimer disease (AD), mild Cognitive impairment (MCI), and normal controls (NC) using structural MRI images consisting of ADNI database data. To overcome the shortcomings of the single-feature technique, the research combined voxel-based morphometry (VBM) parameters with texture features using gray-level co-occurrence matrix (GLCM) and Gabor filters. The method was used to extract 3D morphometric data as well as 2D texture patterns, which improved the discriminative ability. To enhance robustness further, the authors presented a superior feature selection algorithm using support vector machine recursive feature elimination (SVM-RFE) covariance analysis which was good in reducing redundancy and overfitting. Experiments performed with 170 subjects (54 AD, 58 MCI, 58 NC) confirmed the fact that feature fusion was effective compared to single-feature approaches. In AD vs. NC prediction, the model gives an accuracy of 92.86, sensitivity of 87.04 and specificity of 98.28. In MCI vs. NC, accuracy increased to 97.22%, sensitivity to 95.23% and specificity of 100%. In the AD vs. MCI comparison, sensitivity was perfect (100%) and the accuracy was 91.18 %. When using three-way classification (AD, MCI, NC), the framework obtained the accuracy of 85.59%, outperforming the PCA- and multi-kernel-SVM-based baselines. It is important to note that the combination of morphometric features and texture features always resulted in an increase in predictive performance compared with either of the two groups. The researchers achieved the conclusion that VBM combined with optimal use of texture features offers a very strong non-invasive model to early identify AD and MCI. The findings demonstrate the promise of multifeatured combination and enhanced SVM-RFE with covariance to achieve robust and efficient brain MRI-based classification in clinical practice. Nevertheless, the authors admitted some limitations such as comparatively small sample size, single imaging modality (structural MRI), and lack of cross-validation on longitudinal and multimodal data to enhance the generalizability and clinical applications.

Baglat et al. [16] employed several machine learning (ML) models to predict early AD based on T1-weighted MRI data in the OASIS database. The paper focused on the problem of early diagnosis of AD by comparing five classifiers, which were Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), and AdaBoost. However, unlike the previous research involving only one algorithm, the authors tested several models in the same experimental setting using 5-fold cross-validation, which allows the direct comparison of their performance. The Random Forest and AdaBoost methods showed superior classification results of 86%, compared to the Logistic Regression (78 percent) and SVM (81 percent) as well as Decision Tree that exhibited 81 percent in their classification results. AdaBoost and Random Forest also performed best in recall and AUC, showing good adaptiveness to differentiating demented or non-demented patients. Features like Mini-Mental State Examination (MMSE) levels and approximate total intracranial volume (eTIV) were also identified as useful predictors of AD. The study is limited using a relatively small sample (150 subjects), which could make its results less generalizable; the exclusive focus on MRI features without the combination of multimodal biomarkers such as PET or genetic information. The authors concluded that Random Forest and AdaBoost classifiers are useful classifiers on MRI-based early AD diagnosis, and both achieved higher accuracy than traditional models. They also recommended future studies to include deep learning, multimodal neuroimaging (e.g., ADNI), to help increase the accuracy of prediction and clinical viability.

Caldas et al. [17] proposed a machine learning system to identify the premature onset of Alzheimer disease (AD) via inertial measurement unit (IMU)-based gait measurement. The goal of the study was to address the issue of non-invasive and accessible diagnosis of AD based on wearable and non-invasive signals by using gait as a source of digital biomarkers. In contrast to previous methods based on costly neuroimaging or invasive biomarkers, the researchers employed low cost, wearable IMUs and machine learning classification to detect gait disturbances related to AD progression. The frame included IMU measurements of gait on the lower limbs/torso, pre-processing, and feature extraction and classification. Random Forest (RF) yielded the highest accuracy of 91.3% in classifying AD patients and healthy controls compared to Support Vector Machines (SVM) and k-Nearest Neighbors (k-NN). Stride variability, walking speed, and symmetry were observed to be among the most discriminative features. Relatively, the study pointed to the potential of IMU-based gait biomarkers as indicators of early and non-invasive detection of AD. The authors concluded that machine learning on IMU-based gait analysis has the potential to provide a cheap and scalable method of early AD detection. The technique offers a non-invasive alternative to currently established methods of diagnosis, with the possibility of incorporation into a normal healthcare and home monitoring systems.

Yue et al. [18] proposed an explainable machine-learning framework to predict Alzheimer disease (AD) and mild cognitive impairment (MCI) based on the China Longitudinal Aging Study (CLAS) data. Addressing the problem of early and accurate detection, this study used an ensemble learning method with a feature selection and oversampling technique on a significantly large data set of 2,658 older individuals, including demographic, lifestyle, medical, and psychological data. Unlike the previous works that focused on the accuracy of predictive performance and ignored its interpretability, this paper combined feature engineering, classification, and SHapley Additive exPlanations (SHAP) to make any clarification about the features contributions. The framework was tested with nine different classifiers and five feature selection techniques, and AdaBoost with ReliefF showed superior performance. To predict MCI and Normal Control patients (NC), the accuracy, sensitivity, and specificity of the model were 89.2%, 87.7%, 90.7%, respectively, whereas, on predicting AD and NC prediction, the model had an accuracy, sensitivity, and specificity of

99.2%, 99.7%, and 98.7%, respectively. The SHAP analysis has shown education, memory decline, daily life function decline, blood pressure, and lifestyle factor (e.g. fish intake, tea consumption, physical activity and napping habits) as crucial predictors, which are consistent with known clinical factors. Notably, the research provided clinicians with actionable information about lifestyle and physical health interacts with cognitive outcomes. The authors concluded that ensemble learning combined with explainable AI offers both accurate and interpretable AD prediction models, which are potential translational models into the clinical setting.

Alshamlan et al. [19] developed and evaluated various machine-learning (ML) systems predicting Alzheimer disease (AD) based on demographic, cognitive, and structural brain characteristics of the OASIS-2 dataset. The paper tackled the problem of early diagnosis of AD by assessing the performance of Support Vector Machine (SVM), Logistic Regression (LR), and Random Forest (RF) with feature selection techniques that included Minimum Redundancy Maximum Relevance (mRMR) and Mutual Information (MI). In contrast to previous methods that commonly used only classifiers or raw features, the current study thoroughly evaluated data preprocessing (missing data treatment, encoding, and outlier analysis) and explored the effect of feature selection on robustness of the model. LR model with mRMR had an accuracy of 99.08% which was the highest compared with SVM and RF. SVM has shown good results with approximately 98 percent general accuracy, with outliers removed, whereas RF still managed a respectable 96 to 98 percent. The selection of features was shown to enhance efficiency, with both mRMR and MI identifying clinically relevant features (i.e. Clinical Dementia Rating (CDR), Mini-Mental State Examination (MMSE), and brain volume measures). Markers were verified using cross validation procedures (5-fold, 10-fold, and Leave-One-Out) and showed reasonably consistent performance, although not as sensitive to outliers as hold-out validation. Notably, outlier removal was shown to improve performance in hold-out tests, but the authors stated that retaining outliers could increase generalizability in real world settings. The study has some limitations: it utilizes a single dataset (OASIS-2) and thus might not be generalizable to all populations; it does not include multimodal biomarkers (PET, CSF) which might offer more information; and it does not involve more advanced features selection techniques, which might enhance its performance. The authors concluded that ML methods, especially Logistic Regression using mRMR, offer highly accurate and interpretable prediction models of early AD.

Van Steen et al. [20] suggested a machine learning methodology on a large scale to examine the genetic architecture and predictive power of Alzheimer disease in patients and associated risk factors with genome-wide data of 41,686 individuals in the European Alzheimer and Dementia Biobank. To overcome the shortcomings of conventional genome-wide association studies (GWAS) that assume linearity and inadequately represent complex genetic interactions, the study used highly developed models such as Gradient Boosting Machines (GBMs), Neural Networks (NNs), and Model-Based Multifactor Dimensionality Reduction (MB-MDR). This research was based on machine learning to model non-linear relationships and SNP-SNP interactions (contrary to the past works, which principally addressed the effects of additive genetic variation). This allowed future gene discovery as well as accurate prediction. The models were able to replicate all genome-wide significant variants in the training set and found about 22 percent of loci reported in larger meta-analyses and also found six new loci including ARHGAP25, LY6H, COG7, SOD1, and ZNF597, which were also confirmed in external datasets. GBMs showed the best predictive performance (AUC of about 0.692), which was very similar to polygenic risk scores (PRS), but had good correlations with conventional methods. Among other things, the models also pointed out biological pathways that are of prime importance in AD, such as amyloid-beta regulation and microglial activity that have been established to be relevant in AD. It is worth noting that the study revealed that machine learning can be used to complement classical statistical techniques by enhancing genetic discoveries and impound complex biological interactions. The authors found that ML-based methods represent a promising and scalable framework to further develop AD genetics studies and are useful in the future when using larger datasets and federated learning methods.

Govindarajan et al. [21] proposed an explainable machine-learning model to predict early development of Alzheimer's disease (AD) based on clinical and behavioural measures. To deal with the issue of early diagnosis and non-interpretability of conventional machine learning models, the research used various classifiers such as Gradient Boosting, K-Nearest Neighbors, Logistic Regression, Random Forest, and Decision Tree on structured electronic health record (EHR) databases. In contrast to earlier studies that were dependent on neuroimaging or black-box models, the present study concentrated on non-invasive clinical characteristics and incorporated SHapley Additive exPlanations (SHAP) to have a global and individual-level explanation of model. Gradient Boosting model did the best job with accuracy of 93.9 and F1-score of 91.8 and Mini-Mental State Examination (MMSE), Activities of Daily Living (ADL), cholesterol levels, and behavioural indicators were found to be significant predictive features. Moreover, to make the work more clinical-friendly, a streamlit-based web application was created that would allow real-time feature contributions prediction and visualization of its contribution. The authors concluded that a combination of high-performance machine learning with explainable AI and deployable interfaces offers a practical and interpretable solution to the early AD detection, which would be effective in a real-world healthcare environment, where imaging resources might be scarce.

Aslan & Ozupak [22] suggested a comparative machine learning framework of the automatic prediction of the Alzheimer disease (AD) on the basis of clinical and neuroimaging data. To overcome the necessity of precise and timely diagnosis, the paper tested various machine learning models, including K-Nearest Neighbor (KNN), Support

Vector Machines (SVM), Random Forest (RF), and Logistic Regression, on MRI images and related clinical characteristics of the OASIS dataset and cross-validated them on the ADNI dataset so that they could be robust and durable. Contrary to the literature that involved the research of one of the models, this study highlighted the comparative study among algorithms and combining imaging and clinical variables to enhance predictive performance. The grid search was used to optimize the models and the metrics applied to evaluate the models included accuracy, precision, recall, F1-score, and ROC-AUC. On the OASIS data, KNN had the highest performance with an accuracy of 97.33 and F1-score of 97.72, and good generalization on the ADNI data. Age, Mini-Mental State Examination (MMSE), Clinical Dementia Rating (CDR) and measures of brain volume were critical predictive factors. The authors established that machine learning models and in particular the KNN are very effective in the detection of early AD and that a combination of clinical and imaging data improve the diagnostic performance and hence these strategies are ideal in clinical decision support systems.

3. Methodology

The workflow is structured into five principal stages: dataset acquisition, data preprocessing, feature selection, dataset partitioning, and the implementation and evaluation of machine learning algorithms. These sequential steps, depicted in Figure 2, were meticulously designed to facilitate the systematic construction, refinement, and comprehensive assessment of the predictive models. Each stage contributes to ensuring the robustness, accuracy, and clinical relevance of the final classification system.

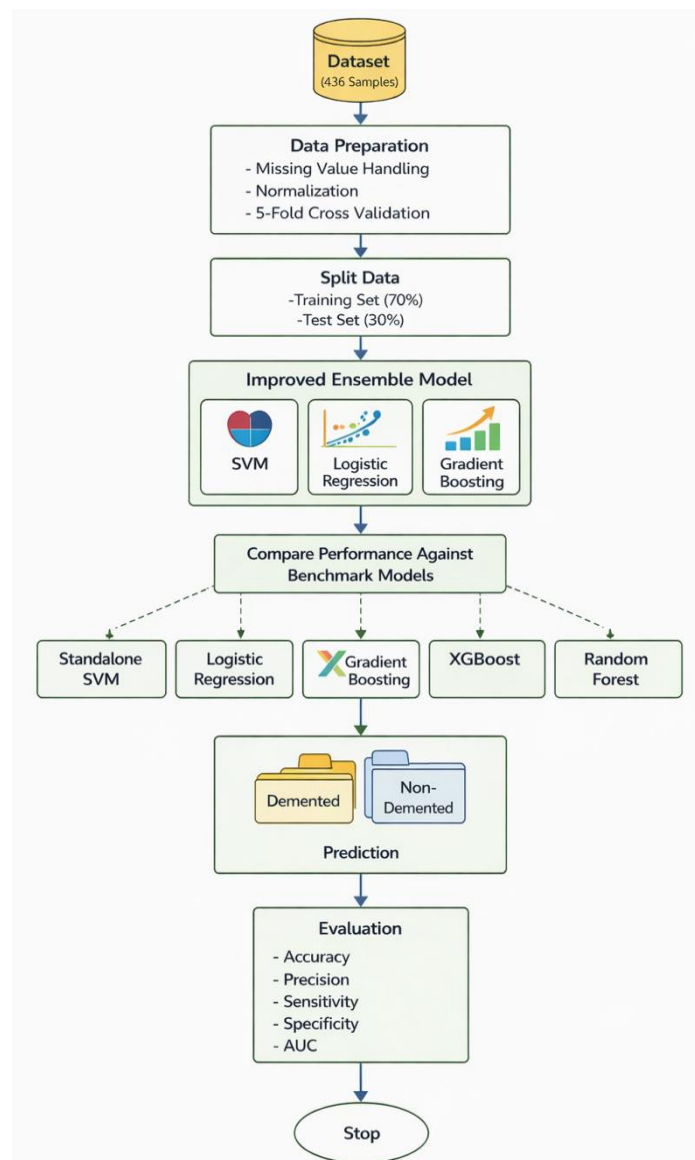


Fig. 2. Model architecture.

3.1. Data Description

The data used in this paper are based on the repository of the cross-sectional dataset of the OASIS and includes 436 records of patients. It contains clinical and demographic factors that are important in respectively cognitive testing and dementia. The data set consists of seven (7) numerical variables which are Age, Mini-Mental State Examination (MMSE), Socioeconomic Status (SES), Estimated Total Intracranial Volume (eTIV), Normalized Whole Brain Volume (nWBV), Atlas Scaling Factor (ASF) and Clinical Dementia Rating (CDR) and two categorical variables Sex and Handedness. The dependent variable (Clinical Dementia Rating (CDR)) has been categorized into two categories; demented and non-demented.

3.1.1. Descriptive Statistics

All numerical variables were calculated using descriptive statistics measures (mean, standard deviation, minimum, and maximum) in order to get a better understanding of the dataset. These statistics can give the central tendency and variability of the data necessary in identifying the possible outliers and a difference in the scale of the features.

Table 1. Summary Statistics of Numerical Features.

	Age	Educ	SES	MMSE	CDR	eTIV	nWBV	ASF	Is_Male	Is_Right_Handed
count	436.000	436.000	436.000	436.000	436.000	436.000	436.000	436.000	436.000	436.000
mean	51.358	3.096	2.243	27.956	0.229	1481.920	0.792	1.199	0.385	1.000
std	25.270	0.966	0.825	2.878	0.421	158.741	0.060	0.129	0.487	0.000
min	18.000	1.000	1.000	14.000	0.000	1123.000	0.644	0.881	0.000	1.000
25%	23.000	3.000	2.000	28.000	0.000	1367.750	0.743	1.112	0.000	1.000
50%	54.000	3.000	2.000	29.000	0.000	1475.500	0.809	1.190	0.000	1.000
75%	74.000	3.000	2.000	29.000	0.000	1579.250	0.842	1.284	1.000	1.000

3.1.2. Distribution Analysis

Histograms and density plots were plotted to investigate the normality and skewness of each numerical feature and also whether there were any outliers or not. As an example, Age and MMSE scores were slightly skewed, whereas eTIV and nWBV measures of brain volume were more or less normally distributed. The knowledge of features distributions is important in the choice of preprocessing methods, including normalization or transformation.

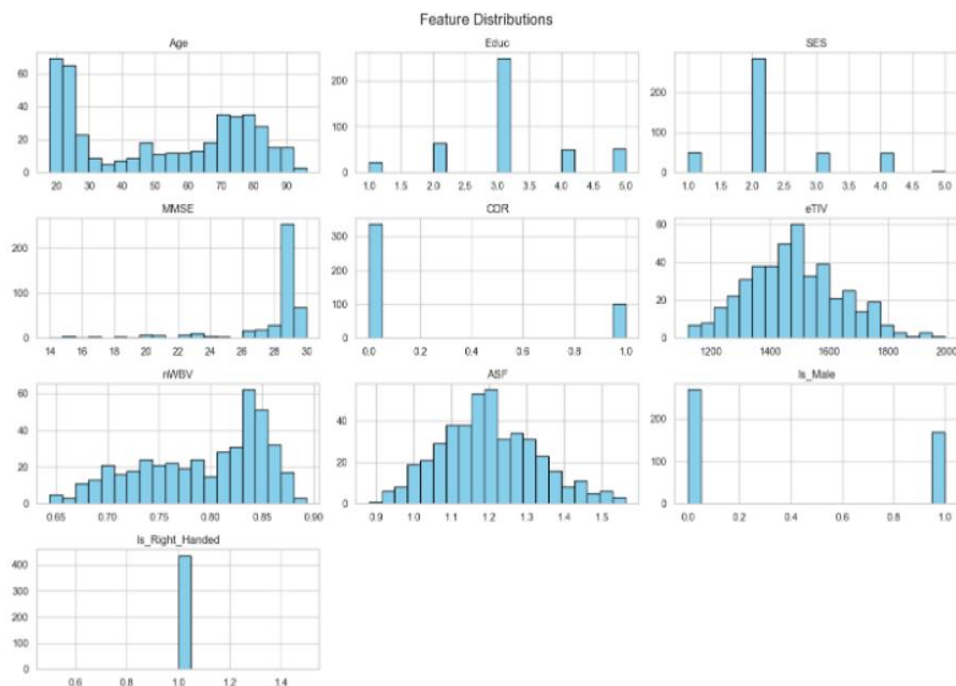


Fig. 3. Histograms of Numerical Features.

3.1.3. Correlation Analysis

A correlation test was done to assess the relationships between the numerical features and their relation with the target variable (CDR). Heatmap was used to compute and visualize Pearson correlation coefficients. The initial results demonstrate that both MMSE and nWBV are correlated with CDR and would be essential to predict the status of dementia. The features with high correlation were also considered to prevent the problem of multicollinearity during development of the model.

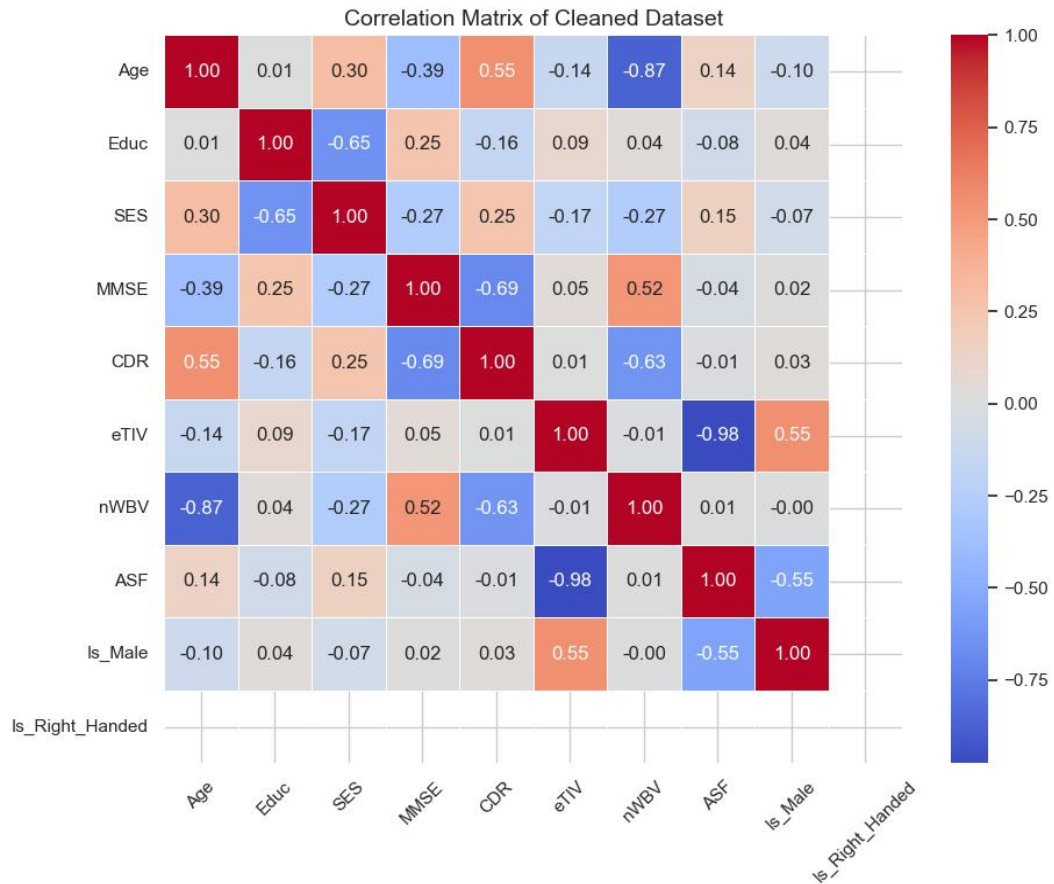


Fig. 4. Correlation Heatmap of Numerical Features.

3.1.4. Relevance and Justification of features.

According to the statistical and correlation analysis, the chosen features have significant relationships with the cognitive deterioration and development of dementia. For instance:

- MMSE is an indicator of cognitive performance,
- nWBV and eTIV are used to record the structural changes of the brain.
- Age is a risk factor with respect to dementia.

These results support the fact that the selected variables should be incorporated in the predictive modeling process.

3.1.5. Class Distribution Analysis.

The target variable (CDR) was binarized and analyzed to determine the class balance between demented and non-demented groups. This is necessary in deciding whether resampling methods are necessary in model training.

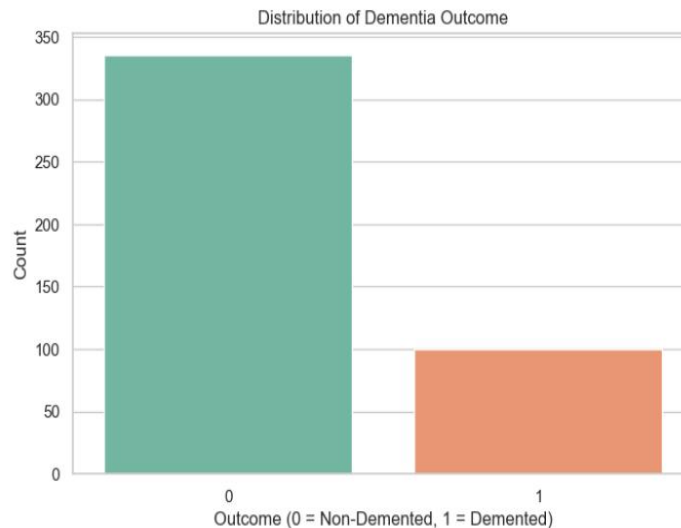


Fig. 5. Class Distribution.

3.2. Data Cleaning and Preprocessing

A preprocessing pipeline was designed such that the data would be structured appropriately and adequate to build a model. It was initiated by the initial investigation of the dataset to determine irrelevant variables, missing data, and inconsistency. Some of the variables, including the Delay column were omitted because they had no clinical importance in dementia prediction and were of little importance to model performance. The elimination of these features assists in the minimization of noise, avoidance of overfitting, and a better computational performance. Missing values were addressed with the help of proper imputation tactics according to the type of variable. In the case of the numerical variables, such as Education (Educ), the Socioeconomic Status (SES), and Mini-Mental State Examination (MMSE), the missing values were substituted with the median of each of the variables. The median was used because it is resistant to outliers and skewed distribution as observed in the exploratory analysis phase, and missing values of the target variable (Clinical Dementia Rating- CDR) were also imputed with the mode (most frequent value) to ensure categorical consistency .To guarantee the reproduction of the results, all preprocessing steps were implemented in a systematic pipeline, which includes the removal of irrelevant features, missing values imputation, target transformation, feature encoding, data partitioning and class balancing (only implemented in the training data).

3.3. Transformation of Target Variable and Encoding of Features

Clinical Dementia Rating (CDR) that is originally an ordinal variable (0, 0.5, 1, 2) was converted into a binary classification variable to simplify the task of predictive modeling and to make it easier to interpret the clinical meanings.

- CDR = 0 $\rightarrow$ 0 (non-demented)
- CDR = 0.5, 1, 2 $\rightarrow$ 1 (demented)

The distribution of the dataset after such transformation was: 336 non-demented samples & 100 demented samples. Categorical variables were coded into binary: Sex $\rightarrow$ Is_Male (Male = 1, Female = 0) & Handedness = Is_Right Handed (Right = 1, Left = 0). The initial categorical variables were then dropped to prevent overlapping and so compatibility with machine learning algorithms.

3.4. Handling of Class Imbalance, Data Splitting and Model Validation.

The dataset showed a class imbalance with the demented class being grossly underrepresented in comparison to the non-demented one. In order to fix this problem, Synthetic Minority Oversampling Technique (SMOTE) was used on the training data. SMOTE uses synthetic samples to interpolate between the existing samples of the minority class, and thus enhances the capacity of the model to learn meaningful decision boundaries of the underrepresented group. In this experiment, the distribution of the classes was balanced in terms of 1:1 ratio and k was taken as 5 nearest neighbors. In order to avoid data leakage and to guarantee an impartial assessment, SMOTE was used on the training data strictly and only after the dataset was split. The stratified sampling was applied to the data to maintain the original distribution of classes by dividing the dataset into training and testing datasets. The 70:30 ratio was taken (70 percent of the data was to be used to train and 30 percent to test). The stratification was done in such a way to provide the representative proportions of both subsets in the sample of demented and non-demented, which enhanced the reliability and validity of the evaluation measures. Stratified 5-fold cross-validation on the training data was also used to further increase model

robustness and generalization. The training set was divided into five equal folds, in each training, four folds were utilized in the training and one fold utilized in the validation. The repetition of this process was made five times so that each fold was used as a validation set at any time. Stratification was used to ensure consistency and prevent bias because the class balance in all folds was preserved and SMOTE was used restricted to each training fold and never exposed the validation folds to synthetics. Lastly, the overall performance of the five folds was averaged to give a stronger and more reliable estimate of the performance as opposed to a train-test analysis.

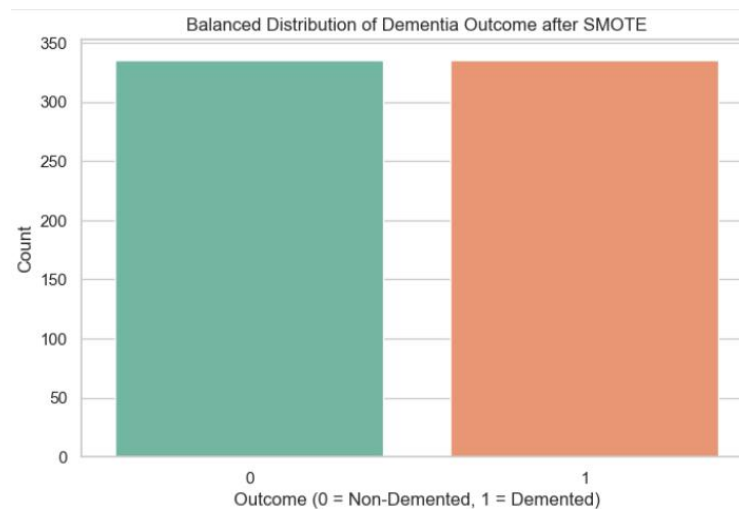


Fig. 6. Class Distribution After SMOTE.

3.5 Machine Learning Models

3.5.1. Logistic Regression

Logistic Regression is an easy but effective linear classification algorithm, which estimates the likelihood of a binary outcome by a logistic (sigmoid) function. It can be used as a baseline model in this study because it is easy to interpret and efficient. The structure of the model is improved to achieve convergence by performing feature scaling with a pipeline and using a greater maximum number of iterations. Logistic Regression is effective when the association between the characteristics and the objective variable is near linear in nature thus it is a strong starting point to compare with more complicated models.

3.5.2. Support Vector Machine (SVM)

The Support Vector Machine is a powerful classification algorithm which operates by identifying the best hyperplane that can separate the data points to different classes. The SVM in this implementation is found to work with non-linear relationships employing a radial basis function (RBF) kernel. Scaling of features is done in order to enhance performance because SVMs are sensitive to input data scale. The model is also able to be tested with ROC-AUC measures with probability estimation turned on. Its main strength is that it has a good generalization capability particularly in high dimensional spaces.

3.5.3. Random Forest

Random Forest is an ensemble classifier algorithm that predicts the data by constructing numerous decision trees and combining the results to enhance the classification rate and decrease the overfitting. The trees are trained on random subsets of data and features, and this brings diversity to the model. The number of trees to use in this work is 100 and this is set to create a tradeoff between the computational capacity and the predictive capability of the random forests. It is especially useful when dealing with complicated data sets with non-linear trends and relations among variables.

3.5.4. Gradient Boosting:

Another ensemble approach is Gradient Boosting which generates models one after another with each model trying to remedy the mistakes already made by the earlier models. Gradient Boosting unlike a random forest tries to minimize the loss function by refining the trees built at each iteration. Parameters that are adjusted in order to improve the performance include the number of estimators, the learning rate, and the depth of the tree in this implementation. This model is said to have very high predictive accuracy, particularly once well regularised, but can be quite overfitting without control.

3.5.5. XGBoost

XGBoost (Extreme Gradient Boosting) uses regularization methods, effective management of missing data and parallel processing which makes it very effective with large and complex data. XGBoost is used in this study using a log-loss evaluation measure and random state to make it reproducible. It tends to work better than other traditional boosting algorithms because of its scalability and the capability of complexities in the data.

3.5.6. Ensemble Model (Voting Classifier)

The enhanced ensemble model uses a combination of various powerful classifiers, namely; Logistic Regression, Support Vector machine and Gradient Boosting, which is combined through soft voting approach. In soft voting, forecasts are made using the average of these forecasted probabilities, and this gives the more certain models a larger weight in determining the final decision. The stronger models like SVM and Gradient Boosting are given weights to make them come first. The method builds on the merits of each model and minimizes the flaws, leading to a better overall performance and strengths. Also, cross-validation is used in order to guarantee the ability of the model to generalize to unknown data and offer more trusted evaluation metrics.

3.6. Model Development and Evaluation Metrics

Classification was done by implementing two models. These two were a Support Vector machine (SVM) using a Radial Basis Function (RBF) kernel and a Logistic Regression (LR) model that was trained on the same data. Performance of the models was evaluated based on the common evaluation measures, such as accuracy, precision, sensitivity, specificity, and AUC-ROC, as described below:

Accuracy was employed to measure the total correct out of both positive and negative predictions.

$$Accuracy = \frac{TP + FP}{TP + TN + FP + FN} \quad (1)$$

Precision was used to measure the proportion of predicted positives that were positive.

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

Sensitivity was employed to measure the proportion of actual positive cases that were correctly identified by the model.

$$Sensitivity = \frac{TP}{TP + FN} \quad (3)$$

Specificity measures the proportion of actual negatives that are correctly identified as negatives by the model.

$$Specificity = \frac{TN}{TN + FP} \quad (4)$$

AUC (Area Under the Curve) was employed to measure the developed models ability to differentiate between positive and negative predictions.

$$FPR = \frac{FP}{FP + TN} \quad (5)$$

Where *FRP* is False Positive Rate, TP is True Positive, TN is True Negative, FP is False Positive, FPR is False Positive Rate and FN is False negative for the equations above.

4. Results, Deployment and Discussion

4.1. Model Performance

The Support Vector Machine (SVM), Logistic Regression (LR), Random Forest (RF), Gradient Boosting (GB), XGBoost (XGB) and the Ensemble Model with Cross-Validation were tested on all the extracted features. Standard evaluation metrics, such as accuracy, precision, sensitivity, specificity, and AUC-ROC, were used to evaluate the performance of these machine learning models, as shown in Table 2.

Table 2. Model performance.

Logistic Regression		Support Vector Machine	Random Forest	Gradient Boosting	XGBoost	Ensemble Model with CV
Metrics	Percentage(%)	Percentage(%)	Percentage(%)	Percentage(%)	Percentage(%)	Percentage(%)
Accuracy	88	85	92	91	91	92
Precision	85	83	89	90	89	90
Sensitivity	91	89	95	92	94	94
Specificity	84	81	88	90	88	89
AUC-ROC	95	94	98	98	98	97

Based on the findings, the Random Forest and Ensemble Model scored the highest accuracy of 92% followed by Gradient Boosting and XGBoost with 91 and Logistic Regression and SVM with 88% and 85%, respectively. Gradient Boosting and the Ensemble Model ranked highest in terms of precision with 90% and 85%, respectively, compared to Logistic Regression and SVM with 83% and 85%. Random Forest, XGBoost, and the Ensemble Model were also the most sensitive at 95%, 94%, and 94% respectively, signifying high sensitivity across the three. The specificity values revealed that Gradient Boosting is the best at 90%, and Logistic Regression and SVM had 84% and 81% respectively.

Moreover, AUC-ROC scores showed excellent model performance among all classifiers as shown in Fig.7 to 12, with Random Forest, Gradient, and XGBoost having the highest score of 98%, Ensemble Model of 97%, and Logistic Regression and SVM with a score of 95% and 94% respectively. In general, the ensemble-based and tree-based models were superior to the individual linear models, but Logistic Regression still showed competitive results with a high sensitivity, which makes it useful in identifying the presence of dementia cases.

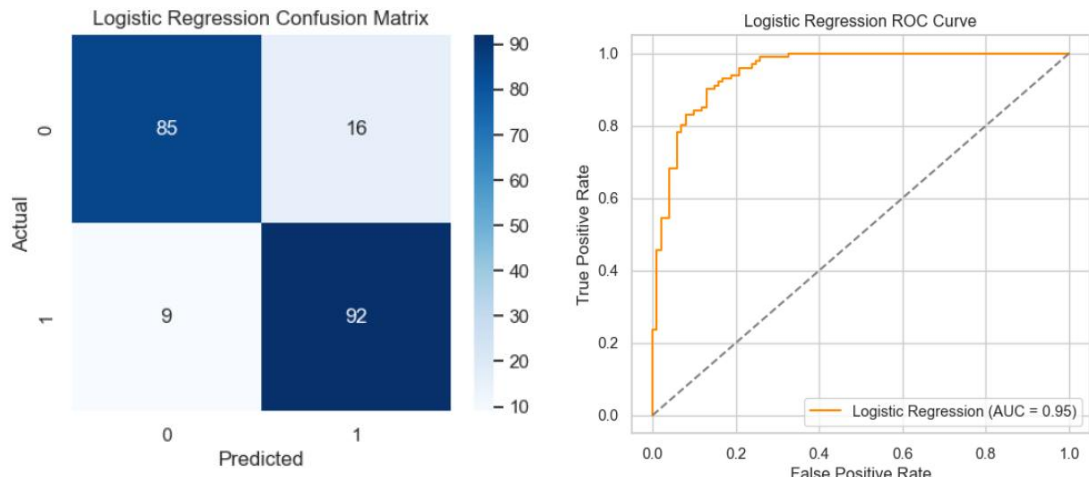


Fig. 7. Confusion Matrix and ROC Curve for Logistic Regression.

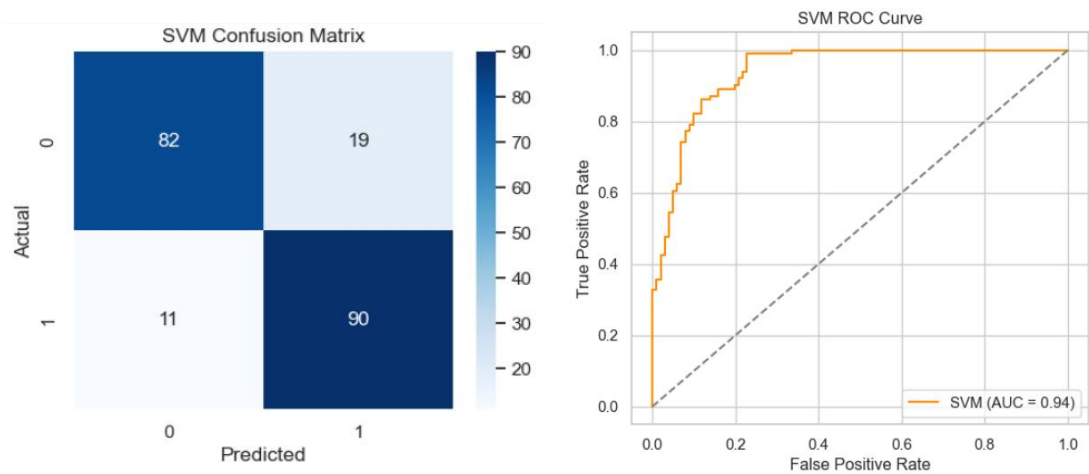


Fig. 8. Confusion Matrix and ROC Curve for Support Vector Machine.

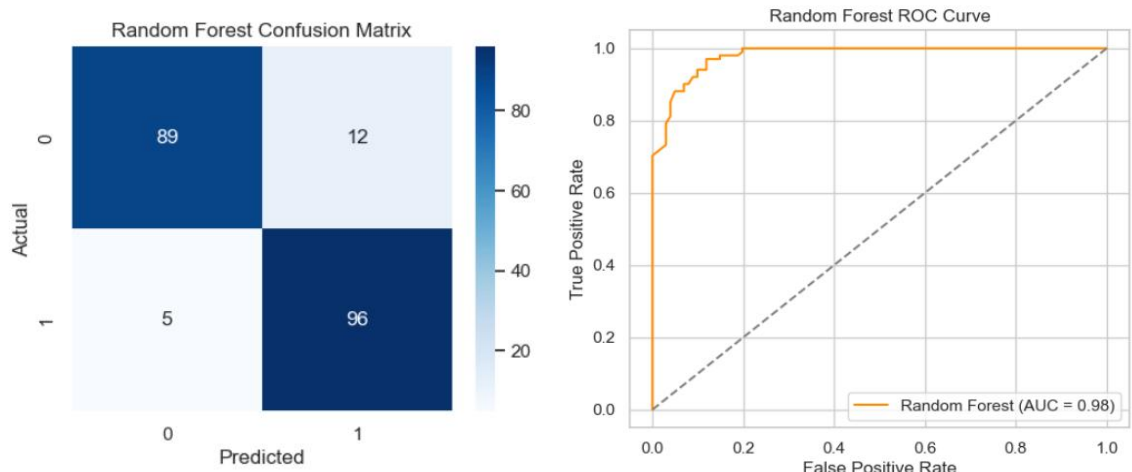


Fig. 9. Confusion Matrix and ROC Curve for Support Vector Machine.

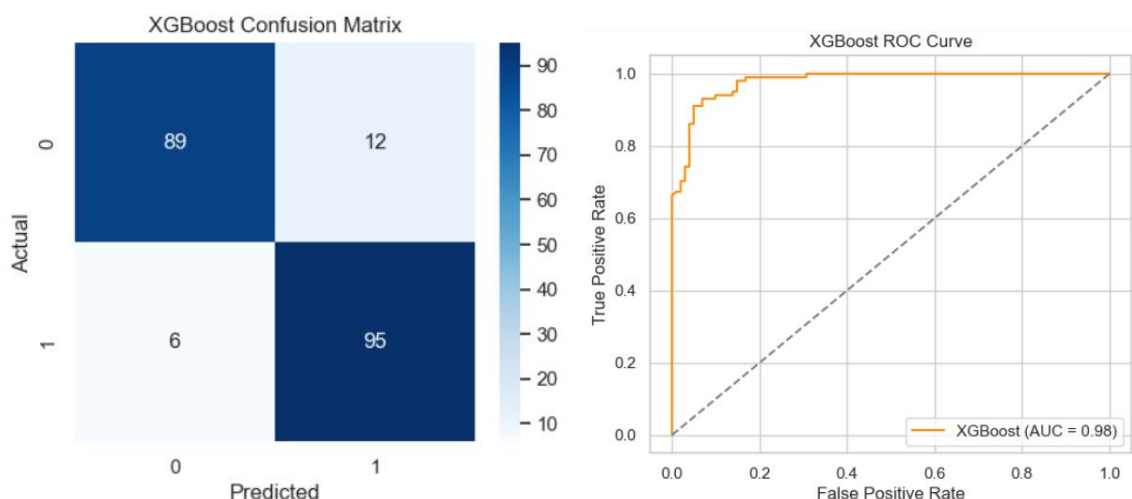


Fig. 10. Confusion Matrix and ROC Curve for XGBoost.

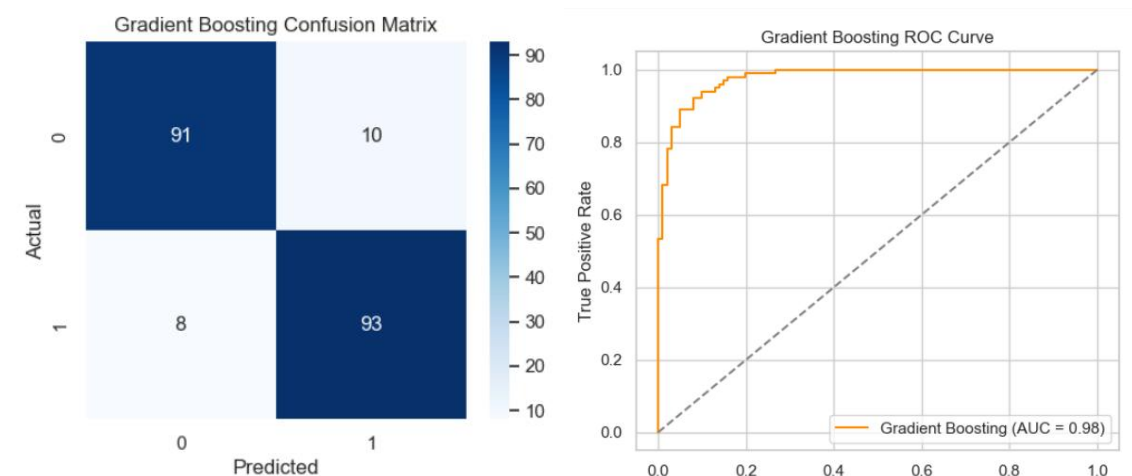


Fig. 11. Confusion Matrix and ROC Curve for XGBoost.

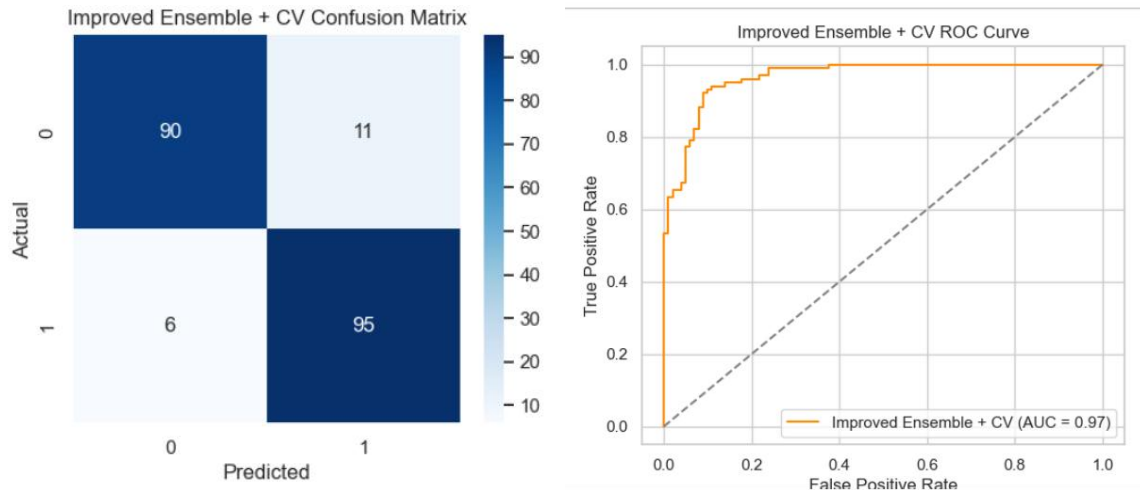


Fig. 12. Confusion Matrix and ROC Curve for Ensemble Model.

4.2. Model Deployment

A lightweight web application was developed, with Flask on the backend. The trained logistic regression model was saved using joblib and made accessible as a RESTful API endpoint at the /predict point. The frontend was built using HTML, CSS, and JavaScript, giving a user-friendly interface to feed patient features into the program. A sign up/log in process was created with an in-memory store providing user authentication. The results of the prediction are displayed in real-time and showed in green for non-demented cases and red for Demented cases. Such deployment makes the model accessible in a clinical decision support environment, as it integrates an efficient preprocessing pipeline, a validated predictive model, and a user-friendly interface as shown in Figures 13 to 16.

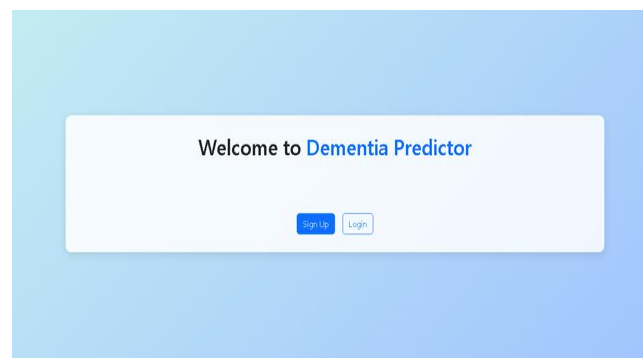


Fig. 13. Welcome page.

Login to Dementia Predictor

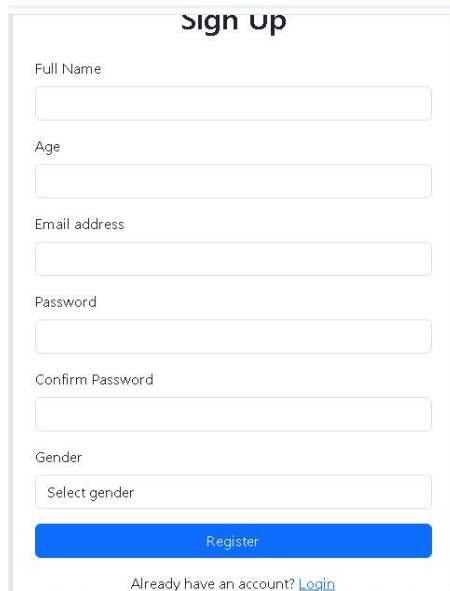
Email address

Password

Login

Don't have an account? [Sign Up](#)

Fig. 14. Login page.



Sign up

Full Name

Age

Email address

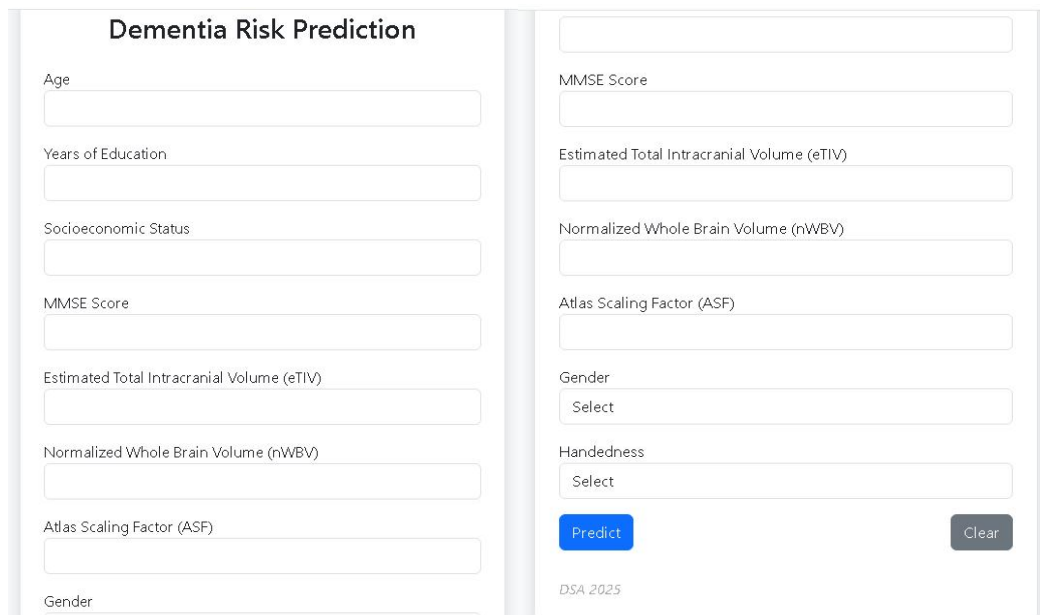
Password

Confirm Password

Gender

Already have an account? [Login](#)

Fig. 15. Registration/login page.



Dementia Risk Prediction

Age

Years of Education

Socioeconomic Status

MMSE Score

Estimated Total Intracranial Volume (eTIV)

Normalized Whole Brain Volume (nWBV)

Atlas Scaling Factor (ASF)

Gender

Handedness

DSA 2025

Fig. 16. Prediction page.

4.3. Discussion

This paper shows that machine learning techniques, such as: Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), Gradient Boosting (GB), XGBoost (XGB), and an ensemble model consisting of all these classifiers, can be successfully used to detect the initial symptoms of Alzheimer disease with the help of the OASIS cross-sectional dataset. Clinical and demographic features were trained on all models and would be a viable, low-cost alternative to neuroimaging biomarkers that are frequently not available in resource-limited healthcare settings. The findings indicate that ensemble and tree based models as compared to the individual linear based models perform best in majority of measures. The ensemble model provided the best accuracy of 92% and good precision (90%), sensitivity (94%), specificity (89%), and AUC-ROC (97%), which is a scientific contribution to a combination of multiple complementary classifiers. Random Forest had the highest sensitivity of 95% and Gradient Boosting and XGBoost had high scores of 90% and 98% on specificity and AUC, respectively. Although simpler, the Logistic Regression and SVM were also capable of competing performance with LR having 88% accuracy, 85% precision, 84% specificity, 89% sensitivity and 95% AUC and the SVM with 85% accuracy, 83% precision, 81% specificity, 89% sensitivity and 94% AUC. It is also worth noting that high sensitivity of the ensemble and Random Forest models is also clinically

important, because by effectively detecting cases of dementia, false negative outcomes are minimized that is important to early intervention and proper planning of treatment.

Nevertheless, the applicability of prediction error, in a clinical context, needs to be reexamined. False negatives and false positives in medical diagnostics may have severe implications on patient care and decision-making. Although the sensitivity and specificity are reported in this study, the analysis of how these errors might affect the clinical outcomes is not provided, which is a significant point of focus in future research.

The study also helps in filling the gap between developing algorithms and their actual implementation by creating a lightweight web application on Flask. The system is based on a RESTful interface and user-friendly interface, indicating that real-time predictions are possible in healthcare settings. However, usability, scalability, security and integration into clinical workflow were not carefully assessed. Also, the web application does not have specific measures of authentication, encryption and adherence to the regulations of healthcare data which is essential in protecting patient privacy.

Future research should hence aim at increasing the number of data to work with larger and possibly multimodal data (e.g., neuroimaging, genetic, or lifestyles) and enhance the predictive power. In addition, research ought to examine the clinical consequences of forecast mistakes, officially assess the usability and scalability of the installed system in actual medical care, and enforce stringent security and privacy controls to fulfill control standards on medical data.

5. Conclusion

This paper proposed, tested, and applied interpretable machine learning models to detect Alzheimer disease in its early stages. There was a good performance of the Logistic Regression (LR) and Support Vector Machine (SVM), but LR was found to be more sensitive and had a greater overall predictive accuracy, hence a single model that can be deployed is LR. In addition to the single models, an ensemble model with LR, SVM, and Gradient Boosting (GB) showed better results in all measures, indicating the highest accuracy, sensitivity and AUC-ROC values. This observation brings out the scientific benefit of ensemble learning whereby the complementary weaknesses of a linear and a non-linear model enhance the robustness and reliability in the identification of early cases of dementia.

These predictive models were further incorporated into the study as a Flask-based web application developed in a lightweight and interpretable AI system that could give immediate, accessible, and color-coded predictions. This type of integration further improves access to dementia screening instruments, especially in low-resource environments, and promotes early intervention by empowering clinicians and community healthcare specialists to inform decisions based on basic data on patients.

All the Declarations and Statements

Author Contributions Statement

Yetunde D. Otun led the development of this study by conceptualizing the research idea, designing the methodology, and carrying out the data curation and preprocessing. She also implemented the machine learning models and prepared the original draft of the manuscript.

Oluwale B. Olajide contributed to the methodological design, performed the formal analysis, interpreted the results, and assisted in the preparation of tables and visualizations.

Abosede O. Oguntunde and Samson A. Arekete were actively involved in reviewing and editing the manuscript, helping to improve its clarity, structure, and overall academic quality.

Benjamin S. Aribisala provided supervision throughout the research process, offering expert guidance and contributing to the critical review and refinement of the manuscript.

All authors have read and agreed to the published version of the manuscript.

Conflict of Interest Statement

The authors declare that there is no conflict of interest regarding the publication of this paper.

Funding Declaration

This work is sponsored by the National Institute of Health (NIH) which is the United States' premier medical research agency. Committed to groundbreaking discoveries, NIH leads efforts in advancing health and medical innovation worldwide. Through collaboration and research excellence, NIH transforms scientific findings into life-saving solutions, addressing global health challenges and improving quality of life for all. NIH's vision focuses on building partnerships and driving innovation in healthcare, research, and human development to foster a healthier future.

Data Availability Statement

The dataset used in this study is publicly available. The data can be accessed from the Open Access Series of Imaging Studies (OASIS) database at https://sites.wustl.edu/oasisbrains/files/2024/04/oasis_cross-sectional-5708aa0a98d82080.xlsx. The processed data and results generated during the study are available from the corresponding author upon reasonable request.

Ethical Declarations

This study uses publicly available and anonymized datasets. Therefore, no ethical approval or informed consent was required. The study complies with all applicable ethical standards for research involving human data.

Acknowledgment

We the authors wish to thank the staff of the faculty of computing at the Redeemer's University and University of Ibadan for their unwavering support and guidance throughout this project. We would like to also express my deep appreciation for the insightful discussions and collaborative efforts of my mentors, supervisors and colleagues. Furthermore, we are indebted to the medical professionals, including hematologists, laboratory technicians, and clinicians, whose expertise supported the study's medical context. We also want to appreciate DATICAN (www.datican.org) for their financial assistance, that made this study was possible. Finally, we would like to thank the Kaggle repository for providing access to the dataset used in this study.

Declaration of Generative AI in Scholarly Writing

During the preparation of this manuscript, the authors used artificial intelligence tools to assist with language editing and improving the clarity of the text. The authors carefully reviewed and validated all content to ensure accuracy and originality. The authors take full responsibility for the final content of the manuscript.

Abbreviations

The following abbreviations are used in this manuscript:

SVM – Support Vector Machine

LR – Logistic Regression

XG – XGBoost

GB – Gradient Boosting

AD – Alzheimer's Disease

RF – Random Forest

SMOTE - Synthetic Minority Oversampling Technique

References

- [1] H. T. Khan, K. M. Addo, and H. Findlay, "Public health challenges and responses to the growing ageing populations," *Public Health Challenges*, vol. 3, no. 3, e213, 2024. doi:10.1002/puh2.213.
- [2] R. K. Chaudhary, U. V. Mateti, P. Khanal, K. B. Rawal, P. Jain, V. S. Patil, *et al.*, "Alzheimer's disease: Epidemiology, neuropathology, and neurochemistry," in *Computational and Experimental Studies in Alzheimer's Disease*, CRC Press, pp. 1–14, 2024.
- [3] Alzheimer's Association, "2019 Alzheimer's disease facts and figures," *Alzheimer's & Dementia*, vol. 15, no. 3, pp. 321–387, 2019. doi:10.1016/j.jalz.2019.01.010.
- [4] H. Hampel, R. Au, S. Mattke, W. M. van der Flier, P. Aisen, L. Apostolova, *et al.*, "Designing the next-generation clinical care pathway for Alzheimer's disease," *Nat. Aging*, vol. 2, no. 8, pp. 692–703, 2022. doi:10.1038/s43587-022-00269-x.
- [5] A. G. Shubar, K. Ramakrishnan, and C. K. Ho, "Optimizing machine learning models for accessible early cognitive impairment prediction: A novel cost-effective model selection algorithm," *IEEE Access*, 2024. doi:10.1109/ACCESS.2024.10763483.
- [6] E. Barbierato and A. Gatti, "The challenges of machine learning: A critical review," *Electronics*, vol. 13, no. 2, p. 2, 2024. doi:10.3390/electronics13020416.
- [7] J. Mateo, J. M. Rius-Peris, A. I. Marañón-Pérez, A. Valiente-Armero, and A. M. Torres, "Extreme gradient boosting machine learning method for predicting medical treatment in patients with acute bronchiolitis," *Biocybern. Biomed. Eng.*, vol. 41, no. 2, pp. 792–801, 2021. doi:10.1016/j.bbe.2021.04.015.
- [8] S. Ng, S. Masarone, D. Watson, and M. R. Barnes, "The benefits and pitfalls of machine learning for biomarker discovery," *Cell Tissue Res.*, vol. 394, no. 1, pp. 17–31, 2023. doi:10.1007/s00441-023-03816-z.
- [9] S. Hanke, F. Mangialasche, M. Bödenler, B. Neumayer, T. Ngandu, P. Mecocci, *et al.*, "AI-based predictive modelling of the onset and progression of dementia," *Smart Cities*, vol. 5, no. 2, pp. 700–714, 2022. doi:10.3390/smartcities5020036.
- [10] L. Bloch, C. M. Friedrich, and Alzheimer's Disease Neuroimaging Initiative, "Machine learning workflow to explain black-box models for early Alzheimer's disease classification evaluated for multiple datasets," *SN Comput. Sci.*, vol. 3, no. 6, p. 509, 2022. doi:10.1007/s42979-022-01371-y.
- [11] V. Vimbi, N. Shaffi, and M. Mahmud, "Interpreting artificial intelligence models: A systematic review on the application of

- LIME and SHAP in Alzheimer's disease detection," *Brain Informatics*, vol. 11, no. 1, p. 10, 2024. doi:10.1186/s40708-024-00222-1.
- [12] M. Li, Y. Jiang, Y. Zhang, and H. Zhu, "Medical image analysis using deep learning algorithms," *Front. Public Health*, vol. 11, 1273253, 2023. doi:10.3389/fpubh.2023.1273253.
- [13] S. Klöppel, C. M. Stonnington, C. Chu, B. Draganski, R. I. Scahill, J. D. Rohrer, N. C. Fox, C. R. Jack, J. Ashburner, and R. S. J. Frackowiak, "Automatic classification of MR scans in Alzheimer's disease," *Brain*, vol. 131, no. 3, pp. 681–689, 2008. doi:10.1093/brain/awn319.
- [14] E. Moradi, A. Pepe, C. Gaser, H. Huttunen, and J. Tohka, "Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects," *NeuroImage*, vol. 104, pp. 398–412, 2015. doi:10.1016/j.neuroimage.2014.10.002.
- [15] Z. Xiao, Y. Ding, T. Lan, C. Zhang, C. Luo, and Z. Qin, "Brain MR image classification for Alzheimer's disease diagnosis based on multifeatured fusion," *Comput. Math. Methods Med.*, vol. 2017, Article 1952373, 2017. doi:10.1155/2017/1952373.
- [16] P. Baglat, A. W. Salehi, A. Gupta, and G. Gupta, "Multiple machine learning models for detection of Alzheimer's disease using OASIS dataset," in *Proc. Int. Working Conf. Transfer and Diffusion of IT*, Springer, pp. 614–622, 2020. doi:10.1007/978-3-030-64849-7_54.
- [17] R. Caldas, N. Costa, D. A. Santos, M. Bessa, and A. Sousa, "Alzheimer's disease detection using inertial measurement units and machine learning algorithms," *IEEE Access*, vol. 9, pp. 134483–134494, 2021. doi:10.1109/ACCESS.2021.3115396.
- [18] L. Yue, W.-g. Chen, S.-c. Liu, S.-b. Chen, and S.-f. Xiao, "An explainable machine learning based prediction model for Alzheimer's disease in China longitudinal aging study," *Front. Aging Neurosci.*, vol. 15, 1267020, 2023. doi:10.3389/fnagi.2023.1267020.
- [19] H. Alshamlan, A. Alwassel, A. Banafa, and L. Alsaleem, "Improving Alzheimer's disease prediction with different machine learning approaches and feature selection techniques," *Diagnostics*, vol. 14, no. 19, p. 2237, 2024. doi:10.3390/diagnostics14192237.
- [20] K. Van Steen, C. M. Van Duijn, V. Escott-Price, et al., "Machine learning in Alzheimer's disease genetics," *Nature Communications*, vol. 16, p. 6726, 2025.
- [21] R. Govindarajan, K. Thirunadanasikamani, K. K. Napa, S. Sathya, J. Senthil Murugan, and K. G. Chandi Priya, "Development of an explainable machine learning model for Alzheimer's disease prediction using clinical and behavioural features," *MethodsX*, vol. 15, p. 103491, 2025.
- [22] E. Aslan and Y. Özüpak, "Comparison of machine learning algorithms for automatic prediction of Alzheimer disease," *Journal of Chinese Medical Association*, vol. 88, no. 2, pp. 98–107, 2025.

Authors' Profiles



Yetunde D. Otun received her B.Sc. degree in computer science from Afe Babalola University, Ado-Ekiti, Nigeria, in 2015, and her M.Sc. degree in computer science from Redeemer's University, Ede, Nigeria, in 2019. Her major field of study is computer science with a focus on data science and medical image analysis. She served as an Assistant IT System Administrator during her National Youth Service Corps (NYSC) program at the Ministry of Youth and Sports, Ibadan, Oyo State, from 2015 to 2016. She also gained industry experience as an IT Trainee at GIS Konsults Nigeria in 2014, IAFD Engineering & Computers in 2013, and Alusoft Technologies Nigeria Limited in 2019. She has participated in professional training programs, including Data Science Africa Summer School, Software Carpentry Workshops, and multiple DATICAN Hackathons on data science and image analysis. She currently conducts research in data science and medical image analysis, with a particular focus on Artificial Intelligence (AI) for medical imaging applications. Ms. Otun is a Ph.D. Scholar under the Data Science and Medical Imaging Analysis Training in Nigeria (DATICAN).



Abosede O. Oguntunde holds a M.Tech (Hons) degree in Computer Engineering from LAUTECH, Ogbomosho, Nigeria. She obtained her M.Sc. and PhD degree in Computer Science from the University of Ibadan, Nigeria. She is an Associate Professor in Department of Computer Science at the Redeemer's University, Ede, Nigeria. Her research interests are in Data Communication, Networking, Mobile Agent Technology, Data mining and Machine Learning. She has published articles in learned journals both at local and international levels and presented papers at conferences.



Samson A. Arekete received his B.Tech. degree in Engineering Physics (Electronics) in 1987 and his M.Tech. degree in Computer Science in 1995, both from the Federal University of Technology, Akure, Nigeria. He obtained his Ph.D. in Computer Science from the same institution in 2013. He has over 20 years of teaching and research experience, in addition to over a decade of professional practice in the marketing research industry. Prof. Arekete began his academic career at the Federal University of Technology, Akure, and later served as Lecturer I at Bells University of Technology, Ota, Nigeria. He joined Redeemer's University in 2007, where he has risen through the ranks and is currently a Professor in the Department of Computer Science, Faculty of Computing and Digital Technologies. His research interests include mobile agents, artificial intelligence, intelligent systems and data science. In addition to his academic contributions, Prof. Arekete held senior positions in industry, including Data Processing Manager at Research & Marketing Services, Lagos; Head of the Data Engineering Unit at Infosense Pty, Cape Town, South Africa; and Assistant General Manager, IT & Analysis Division at Market

Research Consultancy Ltd., Lagos (1996–2005). He has published extensively in peer-reviewed journals and conference proceedings and continues to be an active contributor to the advancement of computer science in Nigeria and beyond.



Oluwole B. Olajide is a scholar, researcher, and technology mentor specializing in Artificial Intelligence, Machine Learning, and Deep Learning, with applications in medical imaging and Natural Language Processing. He is currently pursuing a PhD in Computer Science at the University of Ibadan, Nigeria, where his research focuses on AI-driven frameworks for the detection and severity classification of Sickle Cell Anemia using imaging techniques, as well as Named Entity Recognition for under-resourced languages. As a DATICAN Scholar, he is passionate about advancing AI for social good, particularly in healthcare and education. His doctoral research employs deep convolutional neural networks to analyze chest X-rays in pediatric Sickle Cell patients, supporting evidence-based clinical decisions and addressing diagnostic challenges in low-resource settings. His work also covers hemoglobin classification, pneumonia detection in Sickle Cell Anemia, and the development of NLP models to support African languages. Oluwole earned an MSc in Information Technology from Sheffield Hallam University, United Kingdom, where he gained expertise in computing, systems analysis, and applied research. He also holds Microsoft certifications including Microsoft Certified Professional, Microsoft Technology Associate, and Microsoft Certified Educator, which strengthen his technical expertise. With more than eight years of professional experience, he has held senior roles such as Senior Executive Assistant, showcasing leadership, organizational, and project management skills. As a Tech Mentor at Wecancode, he trains and guides youths and early-career professionals in data science, programming, and AI. He has presented his work at the LASU Research Fair and serves as a Science Journey Ambassador, promoting innovation and the transformative power of scientific research in Africa.



Benjamin S. Aribisala is a Professor of Computer Science in Lagos State University and a Fulbright Scholar. He was the HOD of Computer Science from 2014 to 2018, Dean of the Faculty of Science from 2018 to 2020 and the Director of ICT from July 2020 to December 2020. He was a member of the Governing Council of LASU and a member of the Board of Governors of Lagos State University College of Medicine. His current research interest is in the field of data science and medical application. He is also interested in Artificial Intelligence, Machine Learning, Deep Learning and medical image analysis. Before joining LASU, Professor Benjamin Aribisala had worked in Newcastle University (England, UK) and University of Edinburgh (Scotland, UK) where he conducted research on brain imaging, medical image analysis and machine learning. He is a member of the Nigerian Computer Society, British Computer Society and the Society for brain mapping and therapeutics. He is a great teacher, an excellent mentor, a leader, an administrator and a researcher.

How to cite this paper:

Yetunde D. Otun, Abosede O. Oguntunde, Samson A. Arekete, Oluwole B. Olajide, Benjamin S. Aribisala, "Smart Diagnosis: An Ensemble Machine Learning Web Application for Early Detection of Alzheimer's Disease", *International Journal of Education and Management Engineering (IJEME)*, Vol.16, No.3, pp. 62-80, 2026. DOI:10.5815/ijeme.2026.03.05