INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RESEARCH ARTICLE

A Hybrid Approach for Weak Learners Utilizing Ensemble Technique for Alzheimer's Disease Prognosis

Saima Parvez[1], **Swaleha Zubair**[2]*, **Afreen Khan**[3]

1 M.Sc. Student, Department of Computer Science, Aligarh Muslim University, Aligarh, India
2 Associate Professor, Department of Computer Science, Aligarh Muslim University, Aligarh, India
3 Assistant Professor, Department of Computer Application, Integral University, Lucknow, India

OPEN ACCESS

Received: 29-04-2023
Accepted: 13-07-2023
Published: 26-08-2023

Citation: Parvez S, Zubair S, Khan A (2023) A Hybrid Approach for Weak Learners Utilizing Ensemble Technique for Alzheimer's Disease Prognosis. Indian Journal of Science and Technology 16(32): 2518-2533. https://doi.org/10.17485/IJST/v16i32.1007

* **Corresponding author**.

swalehazubair@yahoo.com

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Abstract

Objectives: To develop a hybrid machine learning (ML) model that predicts Alzheimer's disease (AD) accurately. **Methods :** This study used the Open Access Series of Imaging Studies (OASIS) dataset to develop a hybrid ML model. Given this data, we utilized five algorithms i.e., Logistic Regression, Gaussian Naive Bayes, K Nearest Neighbor, Support Vector Machine, and Decision Tree. An ensemble technique was employed to construct an ML-based hybrid model with 343 observations, 40% of which were used for training and 60% for testing. **Findings**: Using the voting classifier technique, the hybrid Machine learning model obtained an accuracy of 89.28%. Following hyperparameter tuning, the model's accuracy was increased to 90.62%. The effectiveness of AD classification was assessed using Accuracy, Precision, Recall, and F1-score. **Novelty:** The results demonstrate that, even with a limited amount of training data, the Hybrid ML modelling approach can reliably predict Alzheimer's disease in real-world community settings.

Keywords: Alzheimer's Disease; Classification; Machine Learning; OASIS; Prognosis

1 Introduction

Improvements in medical technology and practices in recent years have made a considerable impact on people's standard of living. Our increasingly pampered and lazy culture is connected to degenerative neurological illnesses like Alzheimer's disease (AD). AD was recognized for the first time in 1906 by Alois Alzheimer in a woman named Auguste D who displayed neurofibrillary tangles and senile plaques[(1)]. AD is a neurological disease that gradually decreases memory and cognitive function (recall, adjusting, remembering, and ability to learn). Along with these, AD individuals also have confusion, disorientation, aphesis, hallucinations, and delusions and cannot do simple tasks. It is a terrible sickness for seniors that require cognitive decontamination. The data provided by the World Health Organization (WHO) further emphasizes the importance of this work.

WHO predicts that 90 million people have AD or associated dementia, and by 2050, that figure will quadruple[2]. AD is the most common disease among those aged 65 and older, but head injuries or genetic changes can cause it in younger people. According to the WHO, 7.7 million Indians will have AD or associated dementia by 2030, up from 4.1 million today. These startling statistics highlight the seriousness of the problem and highlight the need for additional research in this area.

To properly handle the issue, the best strategy is to respond before it is too late. Early detection allows patients to participate in clinical trials for the development of new drugs, manage symptoms, and prevent the disease's progression. Numerous medical procedures are used to identify AD, yielding massive amounts of multivariate, heterogeneous data. Due to the variety of medical tests available, it may be challenging to manually compare, interpret, and visualize such data. Machine learning (ML) has the potential to revolutionize healthcare by providing a more efficient means of processing large amounts of data necessary for precise illness detection. ML is a branch of artificial intelligence and a widely studied technological instrument in healthcare practices for the validation and verification of manual diagnosis because of its distinctive ability to handle large datasets, which in turn allows for the creation of accurate predictive models[3]. By utilizing ML techniques, healthcare professionals can not only identify potential risk factors but also predict the course of the disease; this can result in earlier identification and more precise diagnosis[4].

To construct a framework to predict AD, researchers used a variety of data mining approaches such as association rules, regression, clustering, and classification. Numerous published works on this topic are discussed below.

Demidova, L., Klyueva, I., & Pylkin, A. (2019) demonstrated improvements in SVM Classification results and created a hybrid SVM classifier by using a Random Forest classifier as an auxiliary. The authors of this study confirmed the efficiency and improved the SVM classification quality using two auxiliary classifiers, the Random Forest classifier and the KNN classifier, in which parameter values were determined at random by the classifier. Finally, the authors discovered that the Random Forest classifier outperformed the KNN classifier[3]. The shortcoming of this study is that it does not provide a thorough evaluation of any potential downsides or restrictions related to the hybrid technique or the RF classifier itself.

Another study by, Bae, J.B., Lee, S., Jung, W. (2020) performed the work for the identification of AD based on TI-weighted MRI. For the identification, they used deep learning techniques which produced the best results. The authors proposed a convolutional neural network (CNN) using MRI that was based on an algorithm for classifying AD and scans from AD patients and two age- and gender-specific populations, SNUBH and ADNI, which varied in terms of ethnicity and level of education. The average area under the curve for within-dataset validation was 0.91 to 0.94, and the average area under the curve for between-dataset[5]. This study excluded high-risk dementia patients with moderate cognitive impairment (MCI). MCI is a varied group for many reasons, and about 50% of MCI patients have AD pathology. MCI patients were excluded from the model since this study could not demonstrate AD pathology using amyloid PET scans.

Then in 2021, Roobini M and Lakshmi M. established a connection between diabetes and AD. For this study, they used Pearson's recursive graph to predict AD and diabetes. To predict AD with great accuracy and aid in the prevention of AD, the researchers discovered a parameter that is common in both diabetes and AD. In this study, the graph convolutional neural network was used to diagnose diabetes, and the most pertinent features were chosen using a feature reduction technique based on Pearson correlation (GCN). The sensitivity measure, recall, accuracy, and F1- Score measure are some of the factors that were used to generate the performance measures for the proposed job. The proposals' top scores were 98.91%, 97.01%, 98.62%, and 98.98% respectively[6]. However, the limitation of this work is that it used a large amount of data for training and less data used for testing.

Then in 2022, Khan, A., & Zubair, S. (2022) proposed a multi-model based on ML algorithm for the detection of AD. The author performed their research on the OASIS dataset of MRI brain data using a five-stage ML process that they built. Different scales, such as the Mini-Mental State Examination (MMSE), the Clinical Dementia Rating (CDR), and Atlas Scaling Factor (ASF) were used by the authors. On the MRI test dataset that was utilized for this research, 17 different algorithms were used and one of those algorithms the Random Forest Classifier obtained a high level of accuracy, i.e., 86.84%[7].The limitation of this study, it used 17 algorithms of ML and find the best prediction.

Thus, in the present study, we proposed a hybrid ML model using the ensemble technique which identified demented, non-demented, and converted AD patients. We used the OASIS longitudinal dataset, which has 343 subjects and contains a demented, non-demented, and converted group of people with an additional refinement that incorporates the disease's preclinical characteristics. In this work, we used five ML-supervised classification techniques[8]. To obtain a higher accuracy of the hybrid model, the data was divided into two parts: 40% for training, and 60% for testing which was the main focus of our work. To further improve our model, a grid search technique was used to obtain the best parameters of the algorithms and by using these parameters we optimized the model using hyperparameter tuning of the built ML model[9]. Later, we performed a comparative study with state-of-the-art techniques.

This paper is divided into the following parts: In Section 2, the Methodology is shown. In Section 3, the Results and discussion are described. Section 4, ends with a conversation, followed by a conclusion and suggestions for the future direction.

2 Methodology

2.1 Data Source

The proposed hybrid model was built on a longitudinal pool of MRI data which was collected by the Open Access Series of Imaging Studies (OASIS). There are 343 observations in the OASIS longitudinal dataset, with 15 attributes (features). The OASIS dataset contains data pertaining to elderly adults and includes individuals belonging to three categories i.e., demented, non-demented and converted [10]. We used the complete dataset for this analysis. The dataset distribution, classified by dementia categorization group is shown in Table 1 .

Table 1. Details of the dataset Attribute.

S.No.	Attribute Name	Description of Attributes		
1.	Subject ID	Patient's Identification Number		
2.	MRI ID	Patient's Imaging Identification Number		
3.	Group	Patient's either Demented, Non-Demented or Converted		
	Group	Demented	Converted	**Non-demented**
	Subjects	143	4	189
	Age (years)	76.09 ± 7.03	79.43 ± 7.18	76.80 ± 7.84
	Sex (M/F)	83/60	16/25	61/128
	Hand (L/R)	0/143	0/41	0/189
4.	Visit	Number of Visits of each Patient's		
5.	MR Delay	Magnetic Resonance Delay, this time is given before the image collection in real time		
6.	M/F	Patient's Gender		
7.	Hand	Patient's Handedness - right-Handed or left-Handed		
8.	Age	Patient's Age at the time of collecting data		
9.	EDUC	Education, level of patient's education		
10.	SES	Patient's Socio-economic status 1 (highest status) or 0 (lowest status)		
11.	MMSE	Mini-Mental State Examination Score (0-30)		
12.	CDR	Clinical Dementia Rating factor 0 (none), 0.5 (very mild), 1 (mild), 2 (moderate), 3 (extreme dementia)		
13.	eTIV	Estimated Total Intracranial Volume Result		
14.	nWBV	Normalized Whole brain Volume Result		

2.2 Proposed Workflow

In this study, we used an ML algorithm to categorize the data into healthy, diseased or converted patients. For our experimental setup, we used the Jupyter platform of Anaconda, which offered us access to Python modules like pandas, Numpy, seaborne, Matplotlib, etc. Python is a widely used, interpretable, object-oriented programming language with a high level of usability.

The proposed work is a sequential model with seven levels. Each level's sub-levels are maintained in a linear order. The steps proposed are shown in Figure 1 . The first level contained information about the OASIS dataset and the second level of the procedure required data pre-processing, which included cleaning, filling in missing values, feature selection, feature scaling, and category encoding. Also, we explored partitioning the dataset for training and testing. In the third level, we examined the modelling section and applied hyperparameter tuning to the model. The hybrid model is the next level, which has five sub-levels: modelling of the hybrid model, ensemble technique used to generate the hybrid model, voting classifier, hyperparameter tuning, and cross-validation on the hybrid model. The prediction level, which is the final one, assesses the hybrid model created in the previous stage. By classifying the group as AD, non-AD, or converted patients, it predicts the model on the test set. The above-mentioned technique is explored in depth in the following sections.

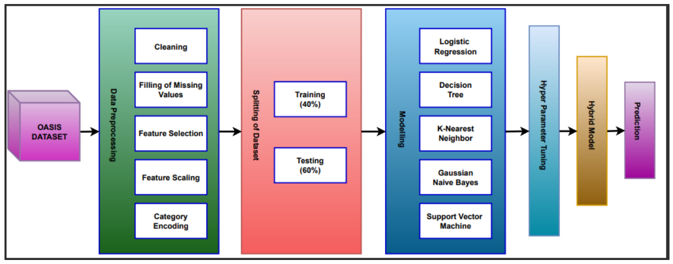

Fig 1. Proposed model

2.2.1 Pre-processing of OASIS Dataset

Most of the data we have at our disposal is unstructured because it is riddled with inaccuracies, and fails to account for key trends and behaviours. As a result, we must do data pre-processing on the full dataset to get rid of these contaminants and get a well-structured model. In order to begin using a hybrid ML model, data pre-processing is the first and most crucial step [11]. To prepare data for analysis, data pre-processing tasks include data cleansing, missing value management, feature selection, feature scaling, and category encoding. The proposed process depicted SES features were lacking many values (i.e., 8 missing values).

In this dataset, we found both relevant and irrelevant features. To find the important features, the feature selection technique has been utilized. The relevant feature set determined for the ML model training is comprised of the following subjects: Subject ID, MRI ID, Group, visit, MR Delay, Age, EDUC, SES, MMSE, eTIV, nWBV, and ASF. In order to prevent the use of unclean data, the dataset is cleansed which primarily consists of irrelevant outliers, features, and duplicates. To address this issue, we used the median approach to impute missing values in the dataset.

Some irrelevant values that were not essential to the analysis were eliminated. Hand and M/F features were eliminated for the model development. For continuous numeric features, we processed the data in the preliminary stage. The dataset also includes categorical features including Group, Subject ID, and MRI ID.

As ML models are based on mathematical computations, it is necessary to translate the string values to numeric form. Thus, we converted both categorical features to numeric format using the label encoding method. The target variable (dependent variable) in this study is the categorical variable i.e., 'Group' which classifies into three categories - demented, non-demented, and converted. Next, we did the scaling of the data values by employing the 'Standard Scalar' method, a data pre-processing component, to scale the feature between zero mean and unit variance. The definition of the standardization process is:

$$scaled_{feature} = (feature - mean(feature))/std(feature) \tag{1}$$

where 'std' is the standard deviation

In this technique, all the features are scaled to the same level. Before training the model, the data were normalized using a standard scale, which enhanced the model's precision and created consistent features. In this study, a standard scalar instance was constructed, fitted to the training data, and then applied to transform the training and testing sets. The next phase is dataset splitting, which is explained in the following section.

2.2.2 Splitting of Dataset

The c lean d ata t hat w as a cquired d uring t he d ata p reparation p rocess i s f urther s eparated. A voiding o verfitting, which emphasizes ancillary details and noise, is the main goal of this step. They just increase the accuracy of the training dataset. As a result, we require a model that functions well when applied to test data, which is a dataset that it has never seen before. This is referred to as generalization. This is accomplished using the strategy, known as data splitting. Employing the sklearn model selection module (sklearn.model_selection), we divided the dataset into two parts: 40% used for training and 60% used for testing, as depicted in Figure 1.

2.2.3 ML Modeling

In machine learning, a "model" is a mathematical simplification of "procedure"; To develop an ML model, the algorithm was trained on the for the label prediction, followed by twining and finally, tested on the holdout data. In order to develop such a model, an ML algorithm uses training from which it learns. When the real model is created using the training set of data obtained from the data splitting stage, the result of the second stage (Section 2.2.2), i.e., a clean collection of split data serves as the input for this level. A trained model that may be used for the analysis and forecasting of more recent data values is the result of the phase.

2.2.3.1 Defining ML algorithms.

In our dataset, the target variable is a multi-class variable. To predict these values from the independent set of features, we used the five supervised ML classifiers viz. Logistic Regression, Decision Tree, Support Vector Machine, K-Nearest Neighbor, and Gaussian Neighbor. These algorithms are discussed below in detail.

i. Logistic Regression (LR): LR is a fundamental ML classifier that is used for classification. A binary dependent variable and a numeric independent variable are used in this method. This method determines the most appropriate coefficient for each model feature. Hence, the model predicts binary outcomes with ease. Gradient descent is utilized to determine the best model optimization process coefficient. The sigmoid or logistic functions scale input values from 0 to 1[12]. The steps involved in the LR algorithm are:

1. Set the coefficient to a random value range
2. Calculate the predicted probability using the sigmoid function It is calculated as Sigmoid

$$x = 1/(1+exp)(-x) \tag{2}$$

3. Calculate the loss utilizing the cross-entropy loss function It is defined as:

$$loss = -(y*log(p)+(1-y)*log(1-p)) \tag{3}$$

Where 'p' represents the predicted level and 'y' is the true label.

4. Estimate the gradient of the loss of the coefficient
5. Update the coefficient using the gradient
6. Repeat steps 'a' to 'e' until the loss is at its smallest

ii.Decision Tree Classifier (DT) : A decision tree classifier is a prominent ML technique used for classification and regression issues. It provides a tree-like structure and all possible outcomes, such as the likelihood of specific events occurring, the cost of resources, and the utility. A DT can deal with both categorical and multidimensional data[13].

The DT algorithm consists of the following steps:

1. Choose the best attribute and use ASM (attribute selection measure) to separate the data
2. After breaking down the data collection into its constituent elements
3. Repeat the process until every branch gets the leaf node

The following formula is used to calculate Gini and Entropy impurity.

In Gini Impurity:

$$\sum_{i=1}^{c} f(1-f) \tag{4}$$

In Entropy impurity:

$$\sum_{i=1}^{c} -flog(f) \tag{5}$$

Where 'f' represents the frequency of label, 'x' at a node and 'n' represents the number of distinct labels.

iii.Support Vector Machine (SVM): SVM is a supervised ML algorithm that may be applied to both classification and regression. In this algorithm, the hyperplane is used to optimize the separation between the two classes. The SVM algorithm's purpose is to expand the margin that offers the best generalization and the lowest error rate. The kernel allows it to manage non-linearly separable data. Common kernels include the linear kernel, radial basic function (RBF) kernel, and polynomial kernel [14]. SVM is more successful with unstructured or semi-structured data. Equation 4 specifies the hyperplane equations utilized in the SVM.

$$W^t X + b \geq 0 \, fordi = \pm 1 \tag{6}$$

$$W^t X + b < 0 \, fordi = -1 \tag{7}$$

Where 'W' denotes the weight vector, 'X' denotes the input vector, 'b' denotes bias, and 't' denotes the margin of separation.

iv.K-Nearest Neighbor (KNN) : KNN classification algorithm is instance-based learning. It collects examples of train data rather than constructing a generic internal model. It performs well with high-dimensional data. The algorithm is sluggish and non-parametric. It generates new classes and stores all existing classes based on similarity measurements. The distance function computes KNN. These distance functions will use Euclidean, Manhattan, Minkowski, or Hamming distances [15]. The Euclidean distance between two data points (x, y) is calculated using the formula:

$$D = \sqrt{\sum_{i=1}^{k} (xi - yi)} \tag{8}$$

Where 'K' represents clusters in the model.

v.Gaussian Naive Bayes: GNB is a classification technique which is based on Bayes Theorem. The Bayes theorem is a mathematical formula that calculates the likelihood of an event depending on prior knowledge [16].

$$p\left(\frac{A}{B}\right) = p\left(\frac{B}{A}\right) * p(A)/p(B) \tag{9}$$

Where 'P (A/B)' shows the probability that event A will happen if event B has previously happened, 'P (B/A)' shows the probability that event B will happen following occurrence A, 'P (A)' shows the prior probability and 'P/B' represents the marginal probability.

2.2.4 Hyperparameter Tuning

The parameter is an essential component of each algorithm in the field of ML. Nevertheless, different types of parameters are required by different algorithms. The parameter and the hyperparameter are the two types of parameters in the ML method. The first category is known as a parameter, and it refers to any value that is determined in advance of the learning process. The value of the parameter, which is an internal variable, can be calculated. The second category is known as a hyperparameter, and it refers to a value that is not only established during the learning process but also controls it. Hyperparameter tuning is the process of selecting hyperparameters in order to improve the model's performance. To select the ideal hyperparameter that provides the best performance on unviewed data, we used several types of hyperparameters in this work (detailed below).

Tuning Hyperparameters can be achieved using a variety of methodologies, of which we employed Grid Search CV. It's a simple brute-force strategy, but it comes at a high computational cost. Cross-validation (CV) training is an important aspect of this procedure. CV was used to modify and evaluate the ML model and it requires data partitioning into training and testing sets. Additionally, train the model several times using various combinations of the training data. Cross-validation used a number of different types k-fold cross-validation approach was employed in this study. In k-fold CV, the dataset was divided into k equal-sized folds, following which the model was trained and evaluated k times. In this type, a certain set of folds was employed as a testing dataset and the remaining k-1 folds were employed as a training dataset. The performance of the model is then displayed as the average of all k iterations. The model was evaluated for each conceivable hyperparameter option. Only a portion of the training data was used to train the model for each conceivable value, while the rest was used for evaluation.

2.2.5 Hybrid ML Modeling

In this study, we employed hybrid modelling to improve the model performance. Because tackling a complicated problem with a single ML approach is challenging, it is ineffective [17]. Instead, a hybrid technique is used to address the difficult problem.

Overfitting and other model faults are resistant in hybrid models. Hybrid models are adaptable since they are tailored to specific issues or datasets. Based on our requirements, we selected a hybrid basic model. It is faster than a single ML model. A hybrid model of learning algorithms can be built in a variety of ways.

We employed the ensemble technique. Ensemble classification is an ML approach that integrates the outcomes of numerous statistical models in order to improve overall performance. In this work, several homogeneous ML models were chosen as weak learners and grouped. When applied to the dataset, each of the weak learners produced its output, either on the entire training set or on a subset of it. The ultimate result was obtained by combining the findings of each weak learner. We combined the predictions of various ML algorithms to determine the average of each algorithm's predictions or to select the forecast with the highest level of confidence.

In this study, we selected five different types of ML algorithms (as explained in Section 2.2.3.1) and created a hybrid ensemble learning model that used all of these ML techniques. The LR, DT, SVM, KNN, and GNB algorithms are among these models. According to Figure 1, which depicts the flow diagram of the hybrid model, the term "hybrid" is employed in this study to describe a collection of different weak learners. On the dataset, all five machine learning algorithms have now been trained. We created 25 weak learners in all, each with its own set of five sub-models. Following training, the prediction and accuracy on the test dataset were compared using the confusion matrices of each of the five models to which we applied 10-fold cross-validation. The five models' accuracy will next be evaluated. We combined the previously mentioned sub-models in this manner to get a more accurate prediction. Finally, the Max Voting Classifier approach was utilized, with the class predicted predominantly by the weak learners serving as the ensemble model's final class prediction.

2.2.5.1 Voting Classifier

. A voting classifier, also known as an ensemble classifier, is an effective approach to improving the predictions of a hybrid model. It is one of the most basic methods for merging predictions from several learning algorithms. Classifiers based on voting are not true classifiers. We can use different algorithms and ensembles to train datasets and then predict the outcome[(18)]. A majority vote on a forecast can be obtained in three ways: hard voting, soft voting, and weighted voting classifier. In this paper, we use hard voting to integrate several ML algorithms. In this form of classifier, the predictions of distinct classifiers are combined using majority voting. The final forecast is the one with the most votes from the separate classifiers. Following developing the hybrid ML model, the ensemble model was predicted using test data in the next stage.

2.2.6 Prediction

In order to determine the accuracy of the prediction, the model's performance was evaluated at this stage using test data. To find a model that produced a high-performing model, ML classifiers were evaluated iteratively. Using the results of hyperparameter tuning, we evaluated our model on the test set after cross-validation. By predicting the target variable using new data, we analyzed the created ML models. Following training, the ML model was applied to held-out data with known target values. Afterwards, ML classifier predictions were compared to target scores. Lastly, the performance measures were evaluated, exhibiting how well predicted and true values matched and were used to assess the learning models' performance.

2.2.7 Performance metrics

The performance of the built model was evaluated using a confusion matrix (CM). A CM describes how well the classification method performs[(19)]. Two rows and two columns make up the CM. The rows represent the predicted class, and the columns represent the actual class. The CM for multi-classes, which has three classes, was represented in this study by rows for the predicted class and columns for the actual class. Entries in the CM represent instances that were expected to belong to a particular class but belong to another class. The elements of the off-diagonal of the matrix reflect those instances that have been incorrectly classified, whereas diagonal entries represent instances that have been correctly classified. A 3*3 CM was used to compute the measurements of accuracy, precision, recall, and F1-Score.

The best AD patient diagnosis model for the MRI dataset was obtained by using these performance evaluation measures. A CM is utilized for accurate prediction; the outcomes concerning the classifier system are described by a CM: true positive (TP), false positive (FP), true negative (TN), and false negative (FN). The measures whose definitions are provided below were constructed using the CM.

- **Accuracy :** This measure shows the overall accuracy of the algorithm used for classification. It is the proportion of instances that have been correctly categorised to all instances. It is defined as:

$$Accuracy = \frac{TP + PN}{TP + TN + FP + FN} \tag{10}$$

- **Precision :** This metric shows the proportion of true positive prediction among all positive predictions. The ratio of true positive events to the total number of both true and false positive occurrences is measured. It is described as follows:

$$Precision = \frac{TP}{TP + FP} \tag{11}$$

- **Recall :** It is also referred to as sensitivity or true positive rate. The recall metric displays the percentage of true positive predictions among all occurrences of real positive outcomes. It is defined as:

$$Recall = \frac{TP}{TP + FN} \tag{12}$$

- **F1- Score :** It is the harmonic mean of precision and recall. It assesses the dataset's test accuracy. It is defined as:

$$F1 - Score = 2 * (Precision * Recall) / (Precision + Recall) \tag{13}$$

- **ROC curve :** The ROC curve graphically depicts the binary/multi-classifier's performance. It depicts the trade-off between true positive and false positive rates by graphing the true positive rate (sensitivity or recall) with the false positive rate (l-specificity) at various thresholds.
- **AUC score :** The AUC score expresses the likelihood that a randomly chosen positive case would be ranked higher than a randomly chosen negative instance. It ranges from 0 to 1. Where 0.5 represents classifiers that perform no better than random guessing and 1 represents the ideal classifier.

3 Results and Discussion

In this study, longitudinal MRI data from the OASIS dataset was employed. This section examines the early detection of AD. For this, predictive analytics and ML modelling were carried out, as was mentioned in Section 2.2.

3.1 Modeling

Table 2 presents the results of the initial ML modelling that was carried out for five ML classifiers, including their performance accuracy and other related metrics. As a result, Gaussian Naive Bayes and Logistic Regression outperformed all other classifiers among the employed algorithms. The most notable improvements were made after the imputation of all the missing values using the median method, as can be seen in Section 2.5.1.

The classification accuracy of the ML modelling before and after hyperparameter tuning is depicted in Table 2 . By employing the 'accuracy score' function, Logistic Regression and Gaussian Naive Bayes achieved 89.73% accuracy, SVM reached 87.5%, Decision Tree achieved 84.3%, and KNN achieved 78.57% accuracy respectively. This ML model used hyperparameters to improve the accuracy which has been discussed in section 2.2.4.

Using all of these algorithms, we employed the ensemble technique to create the ML hybrid model. Our goal was to achieve a high level of accuracy for AD prediction using a hybrid ML model. As explained in Section 4.5, the ensembling methodology was chosen for this study. This has been further discussed in the following section.

3.2 Hybrid ML Modeling

In this paper, the Ensemble technique was used to build the hybrid model. For this, we first created 5 sub-models of each algorithm and then defined an ensemble technique. Then, we constructed an LR model with an L2 penalty which has been discussed in Section 2.2.4. The penalty zeroed out the model coefficient. L1 and L2 penalties were applied to LR. The random

Table 2. Performance results for ML modelling.

Classifier		Accuracy	Precision	Recall	F1 Score		Accuracy	Precision	Recall	F1 Score
Logistic Regression	Before Hyperparameter tuning	0.8973	0.7695	0.6880	0.6892	After Hypterparameter Tunning	0.9017	0.9339	0.6811	0.6597
Decision Tree		0.8437	0.6414	0.6501	0.6431		0.8660	0.5903	0.6418	0.6149
Support Vector Machine		0.875	0.9190	0.6584	0.6406		0.9107	0.9402	0.7101	0.7122
K-Nearest Neighbor		0.7857	0.5899	0.5893	0.5785		0.8035	0.8810	0.6011	0.5899
Gaussian Naïve Bayes		0.8973	0.7788	0.7201	0.7301		0.9017	0.7976	0.7239	0.7353
Hybrid Ensemble Model		0.8928	0.7360	0.6824	0.6755		0.9062	0.9370	0.6956	0.6869

state was used with L2 = 0. The logistic approach was then used to include the previous model in the subsequent model. The Decision Tree classifier was then utilized by employing the fit method. Maximum depth, with a range of 2 to 5, was also used by the DT classifier.

Then, using the kernel parameter, we created the five support vector classifiers and connected the previous sub-model to the subsequent sub-model. We employed the linear kernel, polynomial kernel, and radial basis function; the first and fifth support vector classifiers used the linear kernel. A polynomial kernel was used in the second classifier. The rbf was utilized in the third and fourth classifiers. Furthermore, in the KNN model, we chose 4-6 neighbors using the Minkowski distance function. The p values were set between 1 and 2, with p=2 being the distance metric used in Euclidean distance, and the KNN model employing neighbors from this range. GNB was the final algorithm used to generate a hybrid model. We defined five Naive Bayes classifiers and used the Naive Bayes technique to add each classifier.

After applying all of these parameters, we employed the voting classifier technique. In order to produce a more accurate prediction, we merged the previously stated sub-models in this method. The ML model's hyperparameters were tuned using the CV technique, and its effectiveness was assessed. To accomplish this, we imported the dataset for CV, selected k=10, and then utilized 10-fold CV, which trained and tested the model ten times. We selected a random state as 3 and took its shuffle value as true. The model's performance and prediction accuracy were then assessed using the test data set. Grid search CV was employed to discover which parameters would improve model performance and efficiency. In order to boost performance, we adjusted the model's hyperparameters. To optimize the hyperparameters further, the model was trained. Several hyperparameters settings were tested against a validation set. Table 3 lists the parameters for the built-in hybrid predictive models.

Following that, we integrated all five models (LR, DT, SVM, KNN, and GNB) into a single ensemble model and improved the accuracy of the hybrid model by up to 90%, as shown in Table 2 . Estimators took into account all of these models. The voting classifier class instance was given to the estimator. A majority vote was predicted by this ensemble model.

Table 3. ML model parameters.

Classifier	Parameter Grid	Best Parameters
Logistic Regression	{'penalty' : ['l1', 'l2', 'elasticnet', 'none'], 'C' : np.logspace(-4, 4, 20), 'solver' : ['lbfgs','newton-cg','liblinear','sag','saga'], 'max_iter' : [10, 30,60, 100] [9]}	C=1.281332398719396, max_iter=30, penalty='l2', solver='liblinear'
Decision Tree	{'max_depth': [1], [2], [3], [4], [20], 'max_features': [1], [2], [3], 'random_state':[0, 1, 2, 3, 4, 5, 10, 15,20,35,50], 'criterion':['gini','entropy']}	criterion= 'gini',max_depth= 3,max_features= 2,random_state=50
Support Vector Machine	{'C':[1,10,20,50] [1] ,[9], [19] ,'gamma':[1,0.1,0.02,0.05], 'kernel':['linear','rbf']}	C= 7, gamma= 0.01, kernel='rbf'
K-Nearest Neighbor	{'n_neighbors': [1], [2], [3], [4], [20], [5], [6], [7], [8]}	n_neighbors= 7, weights='distance'
Gaussian Naïve Bayes	{'var_smoothing': np.logspace(0,-7, num=10)}	var_smoothing=0.00181005372000587

3.3 Performance Evaluation

3.3.1 Confusion Matrix

The confusion matrix for each method and hybrid model is shown in Figure 2 shows the confusion matrix after hyperparameter tuning.

The performance of our model was evaluated using a confusion matrix, which summarizes the total number of true positive (TP), false positive (FP), true negative (TN), and false negative (FN) classifications for the AD prediction task, as explained in Section 2.7. Before hyperparameter tuning, the built hybrid model correctly classified 2 patients who were truly demented as positive and 85 patients who were truly non-demented as negative. However, there were 21 false-positive cases where patients were diagnosed with dementia but turned out not to have it or did not have it at all, and no false-negative cases where patients who were predicted as non-demented in reality had dementia.

After hyperparameter tuning, the hybrid model correctly classified the same number of patients as true positive and false negative, but showed an improvement in the number of true negative classifications by 3, correctly classifying 88 patients as non-demented. However, the false positive and false negative rates remained the same, with 21 patients being diagnosed with dementia but not, in reality, having it and 3 patients who were expected to not have dementia but in reality, had it.

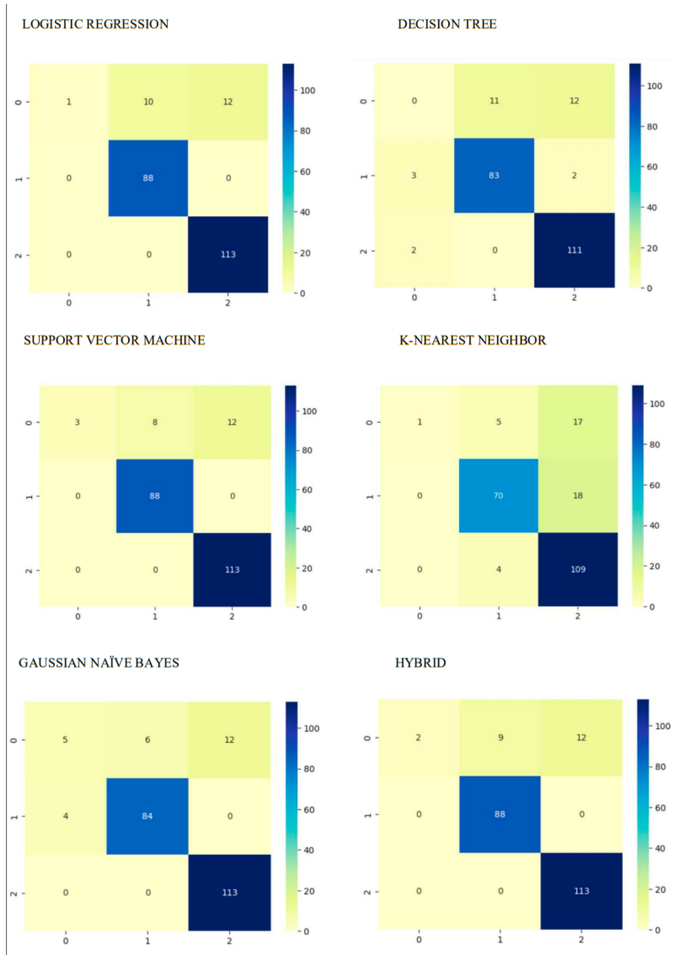

Fig 2. Confusion matrix after hyperparameter tuning

3.3.2 Roc Curve and AUC score

A ROC plot shows a multi-class classification algorithm's performance. It plots against true positives and false positives at various thresholds[20]. The model correctly identifies dementia patients with a high true positive rate (tpr). The model does not identify many demented individuals due to its low false positive rate. The model correctly recognized non-demented patients if the tpr is high and the fpr is low. For the converted class, a high tpr and a low fpr indicate that the model does not identify a large number of converted patients. Figure 3 shows the ROC curves for each algorithm. The ROC curve, which was in the upper left corner, was the most accurate AD test categorization curve. Figure 3 showed that the SVM, LR, and DT ROC curves lean right, giving the Alzheimer detection model higher accuracy. The area under the curve, or AUC, is a single-number assessment of classifier performance. As we can see in Figure 3 , LR, SVM, and DT perform better for the classification of AD, Non-AD and Converted patients.

3.3.3 Precision–Recall Curve

The Precision-Recall Curve (PR-Curve), which is depicted in Figure 4 , demonstrates the link between precision and recall[11]. We plotted the PR curve for multiclass classification, in which case the PR curve will be plotted for each class and the area under the curve (AUPRC) may be determined for each class but only the average of overall AUPRC can be computed in Figure 4 . Area for every algorithm, the SVM was 0.94, LR was 0.94, and the DT was 0.94. Precision shows how many data instances our model claimed were relevant, whereas recall shows how well all ML algorithms can discover relevant data in the OASIS dataset.

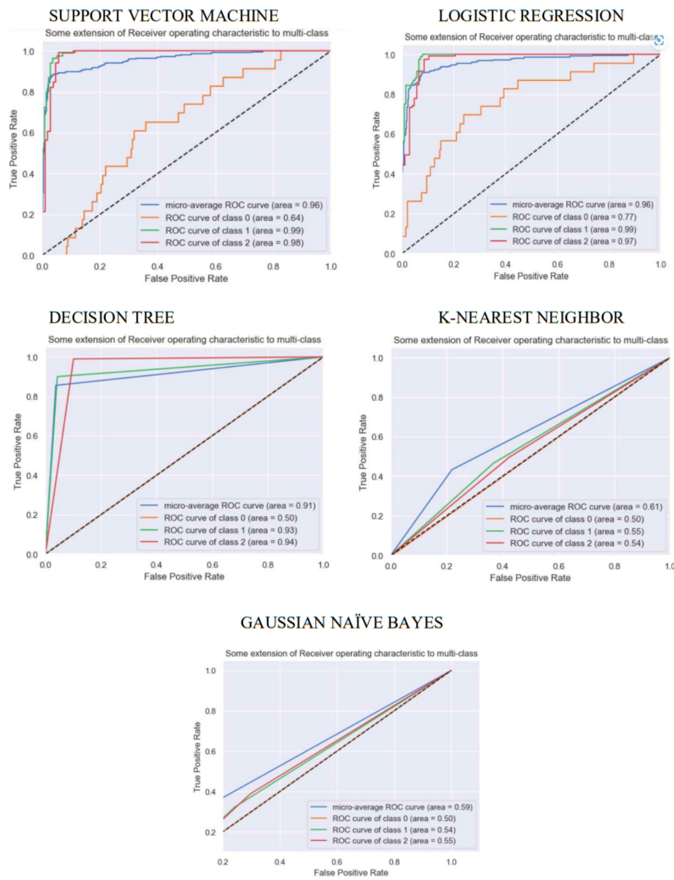

Fig 3. ROC Curve

3.3.4 Classification Report

The Classification Report (CR) summarizes the performance of the multiclass classification algorithm. There are several requirements, including precision, recall, the f1-score, and support for each class [21]. The Scikit Learn library's categorization report function is used to generate this report. The CR for all algorithms and hybrid models for each target variable group (0, 1 and 2 where 0 indicated Demented people, 1 indicated non-demented people and 2 indicated converted people) have been shown in Figure 5 . The hybrid model's classification accuracy was almost 90%. It is the proportion of accurate predictions produced out of all predictions. The classification accuracy of each algorithm used to create a hybrid model was also displayed in Figure 5 . In this study, GNB and LR attained approximately 90% accuracy after applying Hyperparameter tuning, the Bayes algorithm had an accuracy of about 90%, the SVM method had an accuracy of about 91%, the KNN and decision tree had an accuracy of around 80% and 87%, respectively. Our hybrid model was effective at identifying AD-related patients, but high classification accuracy alone was insufficient to conclude because the hybrid model was not perfect; in our case, it has a 10% error rate. Figure 5 depicts the hybrid model's precision as being around 93%.

3.4 State-of-the-Art technique

In the lack of a reliable treatment, early detection of Alzheimer's disease is essential since it enables clinicians to adopt preventive measures. Neuroimaging Studies have demonstrated that the use of ML in neuroimaging data can assist in early AD detection before clinical signs occur [20]. As a result, several research organizations have dedicated themselves to identifying AD and MCI using neuroimaging data and a variety of neuroimaging techniques, including sMRI, fMRI, and PET [22]. Furthermore, considerable advances in the field of ML for AD classification have been made, as demonstrated by the work described in the introduction [5, 6, 7, and 8] However, several issues still present, such as precision, limited availability of data, F1 score, accuracy, and recall. So, we propose a hybrid model that will work significantly better than previous approaches in terms of limited data

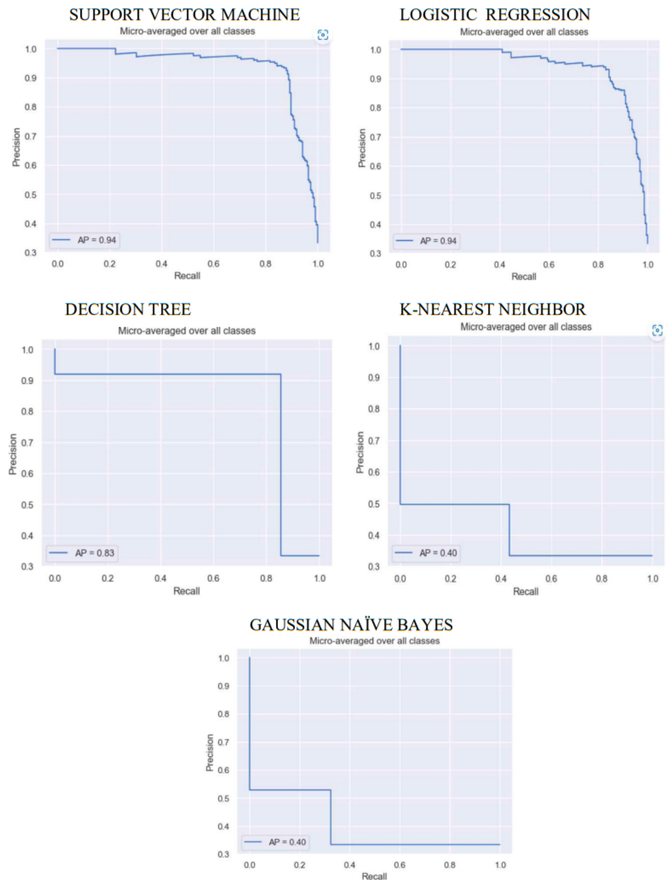

Fig 4. Precision-Recall Curve

supply, precision, F1 score, accuracy, and recall [11]. This is the objective of this research, to create a hybrid ML model that would outperform prior results and reach the best accuracy possible.

So in this study, we successfully created a hybrid model for AD prediction using longitudinal brain MRI data obtained from the OASIS dataset and different ML models like LR, DT, SVM, KNN, and GNB were employed which can recognize AD patients and classify them into demented, non-demented, and converted states using a variety of weak learners, as discussed in section 2.2.5. The ensemble technique was utilised to build the high-performing hybrid ML model. When dealing with this model, a significant result was obtained with an accuracy of 89.28%. Our model outperforms all previous models, which is especially significant given the limited availability of data when using computer-based approaches for diagnosis [8]. In light of this constraint, we chose a train-test ratio of 40-60% throughout the model building. This innovative approach is significant because it indicates the efficacy of such a ratio when data availability is limited, which has not previously been investigated in similar models. We attained the maximum accuracy for our model by reducing the values of three parameters: C values, n_ neighbors, and var_smoothing, which all affected the hybrid model in a better way than the previous works.

We were able to correctly identify Alzheimer's patients by training our model on a small amount of data and testing it on a larger amount [5]. One of the objectives of this study was to develop a hybrid ML model that would outperform previous results and achieve the highest possible accuracy. Furthermore, a notable finding is that we accurately identified individuals as healthy or unhealthy using only twelve MRI measures while employing the test's highest achievable accuracy. Our study's findings may not generalize to the full population because they were based on a small number of populations [6].To deal with this issue in future we can use a big sample size with more diversity of population collected with stratified sampling techniques to increase the generalizability of our finding and acceptability of our model to the broader population. Many MRI studies have been included in ML models for AD prediction. However, no complete model exists that can increase model accuracy in cognitive tests. As a result, we developed a hybrid model of neuropsychiatric testing to increase the identification of AD. Table 4 shows that the proposed technique outperforms other models. As demonstrated in Table 4 , there is significant variety in the

LOGISTIC REGRESSION

	precision	recall	f1-score	support
0	1.00	0.04	0.08	23
1	0.90	1.00	0.95	88
2	0.90	1.00	0.95	113
accuracy			0.90	224
macro avg	0.93	0.68	0.66	224
weighted avg	0.91	0.90	0.86	224

DECISION TREE

	precision	recall	f1-score	support
0	0.00	0.00	0.00	23
1	0.88	0.94	0.91	88
2	0.89	0.98	0.93	113
accuracy			0.87	224
macro avg	0.59	0.64	0.61	224
weighted avg	0.79	0.87	0.83	224

SUPPORT VECTOR MACHINE

	precision	recall	f1-score	support
0	1.00	0.13	0.23	23
1	0.92	1.00	0.96	88
2	0.90	1.00	0.95	113
accuracy			0.91	224
macro avg	0.94	0.71	0.71	224
weighted avg	0.92	0.91	0.88	224

K-NEAREST NEIGHBOR

	precision	recall	f1-score	support
0	1.00	0.04	0.08	23
1	0.89	0.80	0.84	88
2	0.76	0.96	0.85	113
accuracy			0.80	224
macro avg	0.88	0.60	0.59	224
weighted avg	0.83	0.80	0.77	224

GAUSSIAN NAÏVE BAYES

	precision	recall	f1-score	support
0	0.56	0.22	0.31	23
1	0.93	0.95	0.94	88
2	0.90	1.00	0.95	113
accuracy			0.90	224
macro avg	0.80	0.72	0.74	224
weighted avg	0.88	0.90	0.88	224

HYBRID OF MACHINE LEARINING

	precision	recall	f1-score	support
0	1.00	0.09	0.16	23
1	0.91	1.00	0.95	88
2	0.90	1.00	0.95	113
accuracy			0.91	224
macro avg	0.94	0.70	0.69	224
weighted avg	0.92	0.91	0.87	224

Fig 5. Classification Report

datasets, including the number of subjects, train-test ratio and methodologies. In[7] 150 subjects were used whereas we used 343 subjects.

3.5 Importance of the model and its limitation

According to the performed study, a small number of biomarkers combined with ML algorithms can generate outcomes that are nearly as excellent as those attained from more expensive imaging methods like MRI and PET, which ought to be employed as a final validation. This study demonstrated the significance of hybrid ML modelling in predicting AD at an early stage. To identify the AD group, we proposed a hybrid ML model that incorporated conventional diagnostic techniques and a large number of independent cognitive features. Furthermore, five prediction models were employed, and the outcomes demonstrate that each model's forecast accuracy improves over time. For predicting AD, this hybrid model is employed and will be useful when data is scarce and obtaining new data is difficult or expensive. We employed 40% of the available data for training to successfully extract the relevant features from the small datasets, and the resulting accuracy on the hybrid model is 90.28%. It would be highly beneficial in cases where the data is not independently and identically distributed showing that the data does not come from the same distribution and is constantly changing. Another advantage is that this model can be used as a standard for other models. These benefits demonstrate that the model is a viable solution to the data cost and that it would be extremely valuable in the real world, where data is limited.

The proposed hybrid model developed in the study performs well compared to other existing programs. However, the program has a scope for further improvement. For example, as the method uses less training data, it fails to work on fresh unobserved data. However, it can be trained to make a generalized prediction that would be inferred from the small amount of data that it has been exposed to. The future use of our model is likely to involve a similar approach and dataset.

Table 4. Comparison of our suggested approach with other comparable studies utilizing the OASIS dataset.

Authors	Dataset and Distribution	Highlight	Algorithms	Accuracy /AUC
Morshedul Bari Antor, A. H. M. Shafayet Jamil, Maliha Mamtaz, Mohammad Monirujjaman Khan, Sultan Aljahdali, Manjit Kaur, Parminder Singh, Mehedi Masud (2021) [20]	OASIS-2 80%-20% ratio for training and testing respectively	Develop a model to represent the results and analysis for the detection of AD by using various machine learning models.	Random Forest classifier(RFC), SVC, LR and DT	Best accuracy on SVM 74.66%
Khan, Afreen & Zubair, Swaleha. (2019) [7]	OASIS-2 150 subjects used for testing	Based on the imputation of missing variables, a Random Forest Ensemble Classifier is implemented to diagnose AD.	Random Forest Ensemble technique	87% accuracy after imputation of missing values.
Khan, Afreen & Zubair, Swaleha. (2020) [21]	1. OASIS cross-sectional MRI data 2. 416 Subjects, (100 Subjects with mild to moderate dementia)	To develop a K-means discretization approach and also implemented the encoding technique for the prognosis of dementia.	K means Discretization, Non-Discritization, three encoding technique and 18 ML classifier has been used for the analysis.	By using Non-discretization DT achieve high accuracy of 89% while by using discretization SVC(One hot encoding) achieve 86%.
Salehi, Waleed & Baglat, Preety & Gupta, Gaurav. (2020) [22]	OASIS library of MRI patients scans And taking 150 subjects	Trained an automated model which is based on the different ML algorithms for the detection of AD and the best performance achieved on Random Forest.	SVC, LR, DT, RFC and AdaBoost.	Accuracy of different ML model LR(Without imputation)=78.94% LR(W/Imputation)=95.00% DT=81.57% RFC= 86.84%
Kavitha, C., Mani, V., Srividhya, S. R., Khalaf, O. I., & Tavera Romero, C. A. (2022) [23]	OASIS -2 dataset taking 140 subjects and 80%-20% ratio for training and testing respectively	Identify best perimeters by using several ML algorithms for the early stage AD prediction.	DT, RFC, SVM, Gradient Boosting, and Voting classifiers	83% accuracy achieved with RFC.
Proposed work	OASIS -2 dataset 40%-60% ratio for training and testing respectively. Analysis done on 343 subjects	Created a Hybrid Model of ML and Tune the model with the best Hyperparameters which was obtained by using the Grid Search Technique	LR, DT, SVM, KNN, GNB and ensemble techniques were used for building the hybrid model of ML	90.62% accuracy was achieved on the Hybrid model

4 Conclusion

In this study, our objective was to develop a hybrid ML-based model for the accurate prediction of Alzheimer's disease (AD) using OASIS data. We employed a generic framework to classify the OASIS-2 dataset into demented, non-demented, and converted categories. The model development process involved data preparation, splitting, modelling, hyperparameter tuning, and the utilization of grid search for achieving high prediction accuracy. By leveraging ensemble techniques and grid search, we aimed to combine weaker learners with a strong model, thereby enhancing performance. We successfully trained the model on minimal data, 40% for training and 60% for testing, demonstrating its efficiency and potential for wider application. Our results showed that the hybrid ML model, with the optimal collection of hyperparameters identified through grid search, achieved a remarkable prediction accuracy of 90% on the OASIS dataset for AD. This finding highlights the reliability and effectiveness of the developed model.

Importantly, This work provides a noteworthy illustration of a hybrid ML method that can be applied to the OASIS-3 dataset, which comprises 19 joined datasets with more than 800 features. By incorporating effective deep-learning techniques, future research can harness the potential of this model to produce even more detailed and optimal results. In conclusion, our study provides a novel contribution by developing a hybrid ML model capable of accurately predicting AD using OASIS data. The model's high accuracy, combined with the potential for further refinement through deep learning techniques, presents promising prospects for advancing AD diagnosis and treatment in the future.

5 Acknowledgement

The data used in this study were obtained from the Open Access Series of Brain Imaging (OASIS) database (https://www.oasisbrains.org/), made available by the Washington University ADRC. Longitudinal MRI data was retrieved from the following published NIH grants: P50 AG05681, P01 AG03991, R01 AG021910, P50 MH071616, U24 RR021382, R01 MH56584.

References

1) Diogo VS, Ferreira HA, Prata D. Early diagnosis of Alzheimer's disease using machine learning: a multi-diagnostic, generalizable approach. *Alzheimer's Research & Therapy*. 2022;14(1):107. Available from: https://doi.org/10.1186/s13195-022-01047-y.
2) Altay F, Sanchez GR, James Y, Faraone SV, Velipasalar S, Salekin A. Preclinical Stage Alzheimer's Disease Detection Using Magnetic Resonance Image Scans. 2020. Available from: https://doi.org/10.48550/arXiv.2011.14139.
3) Demidova LA, Klyueva IA, Pylkin AN. Hybrid Approach to Improving the Results of the SVM Classification Using the Random Forest Algorithm. *Procedia Computer Science*. 2019;150:455–461. Available from: https://doi.org/10.1016/j.procs.2019.02.077.
4) Khan A, Zubair S. Development of a three tiered cognitive hybrid machine learning algorithm for effective diagnosis of Alzheimer's disease. *Journal of King Saud University - Computer and Information Sciences*. 2022;34(10):8000–8018. Available from: https://doi.org/10.1016/j.jksuci.2022.07.016.
5) Bae JB, Lee S, Jung W, Park S, Kim W, Oh H, et al. Identification of Alzheimer's disease using a convolutional neural network model based on T1-weighted magnetic resonance imaging. *Scientific Reports*. 2020;10(1). Available from: https://doi.org/10.1038/s41598-020-79243-9.
6) Roobini MS, Lakshmi M. Prediction of Alzheimer Disease Using Pearson Recursive Graph Convolutional Neural Network. Research Square Platform LLC. 2021. Available from: https://doi.org/10.21203/rs.3.rs-935323/v1.
7) Khan A, Zubair S. An Improved Multi-Modal based Machine Learning Approach for the Prognosis of Alzheimer's disease. *Journal of King Saud University - Computer and Information Sciences*. 2022;34(6):2688–2706. Available from: https://doi.org/10.1016/j.jksuci.2020.04.004.
8) Shim SO. Multi-Class Classification based on Relative Distribution of Class. In: 2020 2nd International Conference on Computer and Information Sciences (ICCIS). IEEE. 2020;p. 1–4. Available from: https://doi.org/10.1109/ICCIS49240.2020.9257679.
9) Loddo A, Buttau S, Ruberto CD. Deep learning based pipelines for Alzheimer's disease diagnosis: A comparative study and a novel deep-ensemble method. *Computers in Biology and Medicine*. 2022;141:105032. Available from: https://doi.org/10.1016/j.compbiomed.2021.105032.
10) Khan A, Zubair S. Longitudinal Magnetic Resonance Imaging as a Potential Correlate in the Diagnosis of Alzheimer Disease: Exploratory Data Analysis. *JMIR Biomedical Engineering*. 2020;5(1):e14389. Available from: https://doi.org/10.2196/14389.
11) Kruthika KR, Rajeswari, Maheshappa HD. Multistage classifier-based approach for Alzheimer's disease prediction and retrieval. *Informatics in Medicine Unlocked*. 2019;14:34–42. Available from: https://doi.org/10.1016/j.imu.2018.12.003.
12) Xiao R, Cui X, Qiao H, Zheng X, Zhang Y, Zhang C, et al. Early diagnosis model of Alzheimer's disease based on sparse logistic regression with the generalized elastic net. *Biomedical Signal Processing and Control*. 2021;66:102362. Available from: https://doi.org/10.1016/j.bspc.2020.102362.
13) Naganandhini S, Shanmugavadivu P. Effective Diagnosis of Alzheimer's Disease using Modified Decision Tree Classifier. *Procedia Computer Science*. 2019;165:548–555. Available from: https://doi.org/10.1016/j.procs.2020.01.049.
14) Richhariya B, Tanveer M, Rashid AH. Diagnosis of Alzheimer's disease using universum support vector machine based recursive feature elimination (USVM-RFE). *Biomedical Signal Processing and Control*. 2020;59:101903. Available from: https://doi.org/10.1016/j.bspc.2020.101903.
15) Battineni G, Chintalapudi N, Amenta F, Traini E. A Comprehensive Machine-Learning Model Applied to Magnetic Resonance Imaging (MRI) to Predict Alzheimer's Disease (AD) in Older Subjects. *Journal of Clinical Medicine*. 2020;9(7):2146. Available from: https://doi.org/10.3390/jcm9072146.
16) Sharma S, Gupta S, Gupta D, Altameem A, Saudagar AKJ, Poonia RC, et al. HTLML: Hybrid AI Based Model for Detection of Alzheimer's Disease. *Diagnostics*. 2022;12(8):1833. Available from: https://doi.org/10.3390/diagnostics12081833.
17) Li F, Liu M. A hybrid Convolutional and Recurrent Neural Network for Hippocampus Analysis in Alzheimer's Disease. *Journal of Neuroscience Methods*. 2019;323:108–118. Available from: https://doi.org/10.1016/j.jneumeth.2019.05.006.
18) Shanmuga E, Shahina A, Khan N. Dementia Prediction on OASIS Dataset using Supervised and Ensemble Learning Techniques. *International Journal of Engineering and Advanced Technology*. 2020;9(5):244–254. Available from: https://doi.org/10.35940/ijeat.A1827.1010120.
19) De Diego IM, Redondo AR, Fernández RR, Navarro J, Moguerza JM. General Performance Score for classification problems. *Applied Intelligence*. 2022;52(10):12049–12063. Available from: https://doi.org/10.1007/s10489-021-03041-7.
20) Antor MB, Jamil AHMS, Mamtaz M, Khan MM, Aljahdali S, Kaur M, et al. A Comparative Analysis of Machine Learning Algorithms to Predict Alzheimer's Disease. *Journal of Healthcare Engineering*. 2021;2021:1–12. Available from: https://doi.org/10.1155/2021/9917919.
21) Khan A, Zubair S. Expansion of Regularized Kmeans Discretization Machine Learning Approach in Prognosis of Dementia Progression. *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*. 2020;p. 1–6. Available from: https://doi.org/10.1109/ICCCNT49239.2020.9225397.
22) Salehi W, Baglat P, Gupta G. Multiple Machine Learning Models for Detection of Alzheimer's Disease Using OASIS Dataset. *Advances in Information and Communication Technology*. 2020;p. 614–622. Available from: https://doi.org/10.1007/978-3-030-64849-7_54.
23) Kavitha C, Mani V, Srividhya SR, Khalaf OI, Romero CAT. Early-Stage Alzheimer's Disease Prediction Using Machine Learning Models. *Frontiers in Public Health*. 2022;10. Available from: https://doi.org/10.3389/fpubh.2022.853294.