

Research Article

Early MCI-to-AD Conversion Prediction Using Future Value Forecasting of Multimodal Features

Sidra Minhas ¹, Aasia Khanum ¹, Atif Alvi ², Farhan Riaz ³, Shoab A. Khan ³,
Fawaz Alsolami ⁴ and Muazzam A Khan ⁵

¹Department of Computer Science, Forman Christian College University, Lahore, Pakistan

²Department of Computer Science, University of Management and Technology, Lahore, Pakistan

³Department of Computer Engineering, National University of Sciences & Technology, EME College, Rawalpindi, Pakistan

⁴Department of Computer Science, King Abdulaziz University, Jeddah, Saudi Arabia

⁵Department of Computer Sciences, Quaid I Azam University, Islamabad, Pakistan

Correspondence should be addressed to Sidra Minhas; sidraminhas@fccollege.edu.pk

Received 9 December 2020; Revised 27 July 2021; Accepted 12 August 2021; Published 25 September 2021

Academic Editor: Jussi Tohka

Copyright © 2021 Sidra Minhas et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In Alzheimer's disease (AD) progression, it is imperative to identify the subjects with mild cognitive impairment before clinical symptoms of AD appear. This work proposes a technique for decision support in identifying subjects who will show transition from mild cognitive impairment (MCI) to Alzheimer's disease (AD) in the future. We used robust predictors from multivariate MRI-derived biomarkers and neuropsychological measures and tracked their longitudinal trajectories to predict signs of AD in the MCI population. Assuming piecewise linear progression of the disease, we designed a novel weighted gradient offset-based technique to forecast the future marker value using readings from at least two previous follow-up visits. Later, the complete predictor trajectories are used as features for a standard support vector machine classifier to identify MCI-to-AD progressors amongst the MCI patients enrolled in the Alzheimer's disease neuroimaging initiative (ADNI) cohort. We explored the performance of both unimodal and multimodal models in a 5-fold cross-validation setup. The proposed technique resulted in a high classification AUC of 91.2% and 95.7% for 6-month- and 1-year-ahead AD prediction, respectively, using multimodal markers. In the end, we discuss the efficacy of MRI markers as compared to NM for MCI-to-AD conversion prediction.

1. Introduction

Alzheimer's disease (AD) is currently the most prevalent and deadliest form of dementia amongst the elderly—its fatality has increased by 68%, while other major disorders have decreased (Alzheimer's Association 2014). Several serious attempts have been made at disease staging and early diagnosis over the past half-century to curb this high rate. In 2011, the National Institute of Ageing and Alzheimer's Association presented revised guidelines for diagnosing AD [1]. They divided the disease into three stages: (1) preclinical AD, in which measurable changes in biological and pathological markers occur; however, no outward changes are observed, (2) mild cognitive impairment (MCI), during which slight memory and cognitive complaints are observed

and measured, and (3) dementia due to AD, when the patient cannot perform daily life activities due to memory and cognition problems. It is also established that not all MCI patients necessarily develop AD in the future [2]. Generally, there are two kinds of clinical changes for MCI patients: (1) MCI stables (MCIs) are those who retain MCI diagnosis at future time points and (2) MCI progressors (MCIp) are those who show symptoms of AD in the future. Hence, it is imperative to develop reliable techniques to foresee the future clinical trends in MCI patients which can later be used to categorize them as MCIp or MCIs.

A variety of predictors have been proposed for early diagnosis of AD ranging from noninvasive neuropsychological measures (NM) [3] and structural magnetic resonance imaging (MRI) to others like positron emission

tomography (PET), chemical content of cerebrospinal fluid (CSF), and genetic sampling [4]. Recent research is focused on the efficacy of these markers for detecting MCI progression to AD. Numerous studies examined the role of cognitive markers for the conversion prediction task [5–7], while various other projects resorted to scrutinizing the biological markers [8–11]. Similarly, a plethora of studies reported enhanced predictive power of strategies involving multimodal information [12–16], while many authors proposed composite predictors of the latent disease [17, 18]. Among several studies, the authors of [9, 19] presented effective feature selection strategies incorporating regularized linear regression and deep neural networks, respectively. However, a presentation by Sperling et al. [1] indicates that cognitive function and brain structure deteriorate more sharply during MCI to AD transition than other factors like protein deposition. Similarly, NM and MRI measures are most prevalent in the literature as they are comparatively cheaper and easier to obtain than CSF measures which require lumbar puncture. Most of these studies resort to using the baseline data only, whereas, in practice, data from multiple time points may contain helpful information regarding the disease’s underlying pathology.

Identification of AD as a slowly progressing disease has gained interest of many researchers. AD must be monitored over time to capture the predictor variability effectively and understand the disease progression. Hence, to predict MCI conversion to AD, the researchers started taking into account the longitudinal data instead of one time-point data [8, 9, 20, 21]. Huang et al. [22] quantified the 6-month-apart longitudinal change observed in significant voxels of MRI images using linear regression and, later, employed a hierarchical classifier to categorize MCIP and MCIs. Their longitudinal method achieved a classification accuracy of 79.4% compared to 71.7% accuracy obtained when they used baseline data only, hence reinforcing the abilities of longitudinal data. Lee et al. [23] investigated the effects of longitudinal callosal atrophy to predict MCI to AD conversion. Their method was found to be more accurate (84%) for female patients than male patients (61%). Another powerful method for MCI progression was presented by Runtti et al. [24], who employed neuropsychological measures and heterogeneous biomarkers to calculate the Disease State Index (DSI) of all patients. Previous DSI measures were then used to estimate the next DSI using linear regression. The classification accuracy achieved by this method was 76.9%. Guo et al. [25] proposed a robust, novel method using three-year follow-up data to mine a dynamic MRI morphological feature and based MCI-to-AD conversion prediction on it. The classification accuracy obtained by Guo et al. was 84.62%. A few other significant studies employing longitudinal predictor information are summarized in Table 1. Lawrence et al. [28], apart from advocating the superior power of time-series data for early AD prediction, provided a detailed review of studies to gather longitudinal data for this domain. However, the researchers are still faced with the sparsity of data on CSF and PET scans for the said task.

A variety of machine learning and pattern recognition techniques have been validated and tested in the MCI-to-AD

conversion prediction domain, which includes supervised [9, 16], unsupervised, and semisupervised [11, 29] methods. Further stratifications involve classification and regression [15, 21] techniques. While SVM and its variants have been a classical choice for early diagnosis of AD [26], several researchers opted to use random forests [11], hierarchical classifiers [22], and maximum uncertainty linear discriminant analysis [30]. Despite many efforts, the desired performance in terms of AUC and accuracy for a stable clinical decision support technique has not been achieved. In many studies, a tradeoff between sensitivity and specificity of the system is observed. From most of the earlier studies, it is extracted that the quality of predictors is more crucial in detecting MCI progression than the classifier’s sophistication. Recognition of an effective feature set for classification is essential. Another noticeable fact that resulted in low performance of systems for early diagnosis of AD is the use of one time-point data only. Peterson et al. [31] reinforce that the rates of clinical changes in MCI subjects are meaningfully rapid than the controls, hence enhancing our interest in quantifying those changes.

In this paper, we propose a multistep method to forecast next-in-sequence predictor readings engaging the previous two readings and later use the time-series information for MCIP and MCIs classification employing an SVM classifier. Our experiments incorporate nonbrain neuropsychological measures (NM) and brain MRI-derived measures as predictors of MCI progression to AD. We designed a novel weighted gradient offset method, which calculates the future predictor reading employing two consecutive, previously available, time-domain readings of multimodal predictors. The future value forecasting system is governed by quantifying previous gradient changes onto the next interval. The novelty of the proposed method lies in the effective employment of several bimodal predictors to forecast future predictor readings. We validate the proposed method on Alzheimer’s Disease Neuroimaging Initiative (ADNI) data set. The efficacy of our system is judged on nonbrain and brain predictors in individual and combined settings.

2. Materials

2.1. ADNI. Data used in this work are obtained from Alzheimer’s Disease Neuroimaging Initiative (ADNI) database. ADNI is a large-scale multisite study that aims at analyzing markers from cognitive tests, blood tests, tests of CSF, and MRI/PET imaging concerning their ability to characterize the progression of Alzheimer’s disease. ADNI results from the efforts of many co-investigators from a broad range of academic institutions and private corporations. Subjects were recruited from over 50 sites across the US and Canada. ADNI data used in this study were retrieved in January 2019. Till then, ADNI, in its three studies (ADNI-1, -GO, and -2), had recruited over 1500 adults aged between 55 and 90 years to participate in this study. For up-to-date information, see <http://www.adni-info.org>.

Here, for technique validation, we consider the subset of MCI subjects enrolled in ADNI-1 with at least three consecutive follow-up readings available. The subjects with MCI

TABLE 1: Previously performed longitudinal studies.

Author	Follow-up interval (months)	Follow-up duration (months)	Predictors	Dataset (MCIp/MCIs)	Accuracy (%)	AUC (%)
Guo [25]	6	36	MRI	43/35	84.6	92.5
Teng [26]	6	48	PET, NM	33/46	88.6	93.1
Lee [27]	6	24	MRI, CSF, NM		81.0	86.0
Huang [22]	6	48	MRI	70/61	79.4	81.2
Lee [23]	12	24	MRI	78/54	84.0 (female)	—
Arco [28]	6	12	MRI, NM	73/61	74.0	79.3
Runtti [24]	24	24	MRI, NM, CSF, genetic	126/174	76.9	82.3
Zhang [9]	6	24	MRI, NM, PET	35/50	78.4	76.8
Davatzikos [17]	12	36	MRI, CSF	69/170	55.8	73.4
Misra [8]	8	15	MRI	27/76	—	77.0

NM: neuropsychological measures, PET: positron emission tomography, MRI: magnetic resonance imaging, and CSF: cerebrospinal fluid.

converting to AD any time before the last follow-up time, i.e., 36th month, are called MCI progressors (MCIp), and those retaining MCI diagnosis at all times are known as MCI stables (MCIs). Subjects with missing values at designated follow-up times are ignored. For 6-month-ahead prediction, the dataset consists of subjects with three consecutive 6 monthly follow-up readings. Therefore, we have 49 MCIp and 70 MCIs subjects whose IDs are given in the supplementary files (available here). Similarly, for 1-year-ahead prediction, the dataset contains three consecutive annual readings, resulting in 35 MCIp and 50 MCIs subjects. Subject demographic information at baseline is stated in Table 2.

2.2. Markers. A study conducted by Sperling et al. [1] demonstrates that in contrast to chemical and protein compositions, brain morphometry, clinical function, and cognitive performance are most affected during MCI-to-AD transition. Hence, in this work, we concentrate on longitudinal analysis and use of nonbrain neuropsychological markers (NM) and brain MRI morphometric markers only. Further details about the markers considered in this study are specified as follows.

2.2.1. Neuropsychological Markers (NM). ADNI participants are asked to perform a series of neuropsychological tests and clinical assessments regularly. The scores and values of each test and assessment are recorded. The NM employed in this work are identified in [32]. In [33], Pereira et al. comprehensively studied the impact of various neuropsychological measures on MCI-to-AD conversion prediction, and based on those findings, we select a total of eight markers which are as follows: total scores of AD Assessment Scale (ADAS 13), Rey Auditory Verbal Test (RAVLT), Clock Drawing Test (CDT), Clock Copying Test (CCT), Immediate Recall Total Score (LIMM), Mini-Mental State Examination (MMSE), Trail Making Test A (TRAA), and Trail Making Test B (TRAB). Table 3 describes the mean and standard deviations observed in these values. The test procedures and scoring criteria for the selected cognitive tests are given in the ADNI General Procedures Manual [32].

TABLE 2: Subject demographic information.

		MCIp ($n=49$)	MCIs ($n=70$)
6 months	Age	73.3 ± 7.4	74.7 ± 7.3
	Gender (M/F)	29/20	36/34
	Education	15.8 ± 2.9	15.3 ± 2.9
		MCIp ($n=35$)	MCIs ($n=50$)
1 year	Age	73.5 ± 7.2	74.8 ± 7.3
	Gender (M/F)	2/15	29/21
	Education	16.0 ± 2.7	15.8 ± 2.7

2.2.2. MRI Morphometric Markers. For MCI-to-AD conversion prediction, we also employ the longitudinal MRI morphometric measures provided by the University of California, San Francisco, Memory and Ageing Centre, on the ADNI website. The complete details of extracting the MRI-derived measures are provided in [32]. Briefly, T1-weighted, MPRAGE MR scans from a 1.5 T Siemens scanner (dimensions 1 mm × 1 mm × 1 mm, TR: 20 ms, TE: 5 ms) were downloaded and preprocessed for gradient warping, scaling, B1 correction, and N3 inhomogeneity correction by Mayo Clinic [34–36]. Cortical reconstruction and volumetric segmentation were then performed using Freesurfur image analysis suite version 4.3 based on the framework provided in [37]. In the said framework, a within-subject template space and average image, unbiased towards the chronological scan order, was created using robust, inverse consistent registration [38, 39]. Cortical morphometric measures were obtained after iterative topology correction, nonlinear Atlas registration, and nonlinear spherical surface registration. The segmented images were passed through an intensive quality control (QC) process [40] by Mayo clinic and provided for use on the ADNI website. The current study includes the information only from those images which passed the overall QC process. The MRI biomarkers consist of volumes of brain regions obtained after cortical parcellation and white matter parcellation, the surface area of the brain regions, and cortical thickness of the brain regions provided in the UCSFFSL file on the ADNI website. We consider a total of 249 brain MRI features in the current study.

3. Methods

Our work focuses on utilizing the embedded predictive power in longitudinal features for MCI progression prediction. We use multimodal (nonbrain NM and brain MRI)

TABLE 3: Neuropsychological measures used in this study ($M \pm SD$: mean \pm standard deviation).

Neuropsychological measure	MCIp	MCIs	p values
ADAS-cog total 11	19.9 \pm 3.83	14.4 \pm 5.22	<0.001
RAVLT	6.26 \pm 1.71	9.13 \pm 2.75	<0.001
CLOCKSCOR	4.12 \pm 1.12	4.49 \pm 0.64	0.069
COPYSCOR	4.76 \pm 0.49	4.68 \pm 0.65	0.463
LIMMTOT	7.00 \pm 2.87	8.19 \pm 2.82	<0.05
MMSE	26.82 \pm 1.64	27.80 \pm 1.52	<0.05
TRAASCOR	49.58 \pm 28.11	37.01 \pm 9.96	<0.05
TRABSCOR	125.6 \pm 62	104.4 \pm 49.4	0.061

markers recorded at two consecutive follow-ups for MCI subjects and pass them through a pipeline to forecast the next marker value. The time-domain readings (three in number) are then collectively considered for MCIp or MCIs label prediction. The detailed process overview is shown in Figure 1, which shows nested cross-validation loops with a 5-fold outer loop for classification and an inner leave-one-out loop for average weights calculation. The process is repeated for all feature subsets.

The proposed approach for longitudinal trajectory modeling is three-phased: (1) estimating gradient offset and its weights from the training data, (2) forecasting future marker values for the test data using models generated by the training data, and (3) classifying the time-domain trajectories of test instances as one of the two classes. The system is implemented on MATLAB R2018a. Details of each module are in the following sections.

3.1. Preprocessing

3.1.1. Marker Normalization. Feature normalization is an essential step in a multivariate, multiscaled machine learning environment. Normalization is performed to remove bias due to very small or very large-scaled values. We scale all features to have values between 0 and 1 before the train-test split. For longitudinal dataset normalization, we divide the time-domain trajectories of each biomarker by the maximum value of the respective biomarker to produce values between 0 and 1 while preserving the longitudinal trends.

3.1.2. Marker Ranking and Selection. The biomarker set we consider in this work consists of 249 MRI and 8 NM biomarkers, of which not all contribute effectively towards MCIp vs. MCIs segregation. For optimal biomarker subset determination, we adopt a wrapper-based approach for selecting significant biomarkers. We perform two sampled Student’s T -tests on baseline biomarker readings and then sift them according to their ranks in impact towards MCIp vs. MCIs discrimination. The p values of the T -test indicate the significance of a particular biomarker towards effective diagnostics. Later, we form biomarker subsets by incrementally adding one biomarker at a time according to its rank. Each biomarker

subset is modeled and evaluated for its efficacy through our proposed system.

3.2. Longitudinal Trajectory Modeling. To forecast future clinical changes of an MCI subject, i.e., at $t + 2\Delta t$, a reliable predictor model must capture the effects of past feature readings (at t and $t + \Delta t$) and project them onto the future clinical readings. Assuming longitudinal trajectories to be piecewise, we select a linear progression model for MCI-to-AD progression to avoid overfitting to a small dataset [41, 42].

For weights calculation and model validation, we divide the dataset into a 5-fold cross-validation setup. The training set is further stratified into training and validation set according to the leave-one-out (LOO) scheme for hyperparameter tuning. Let $X_c = \{X_1, X_2, \dots, X_N\}$ be the training set of all N longitudinal predictors of n MCI subjects belonging to class c ($c = 1$ for MCIp and $c = 0$ for MCIs). Then, $X_m = [x_{mi}(t) \ x_{mi}(t + \Delta t) \ x_{mi}(t + 2\Delta t)]$, for $m = 1, \dots, N, i = 1, \dots, n, \Delta t = 0.5, 1$, where Δt is 0.5 for 6-month-ahead prediction and 1 for one-year-ahead prediction. The base gradient, $\delta_m^c(\Delta t)$, quantifies the average piecewise change for the m^{th} predictor’s trajectory belonging to class c . To obtain $\delta_m^c(\Delta t)$, we subtract $x_{mi}(t)$ from $x_{mi}(t + \Delta t)$, divide it by Δt , and average it over the total number of instances in the respective class.

Similarly, we calculate the piecewise gradient between consecutive longitudinal values of the validation instance, $\delta_{mi}(\Delta t)$, using Euclidean geometry as $\delta_{mi}(\Delta t) = x_{mi}(t + \Delta t) - (x_{mi}(t)/\Delta t)$ for $m = 1, \dots, N, i = 1, \dots, n$. The validation subject’s piecewise gradient, $\delta_{mi}(\Delta t)$, is subtracted from the respective base gradient $\delta_m^c(\Delta t)$ to obtain the gradient offsets $OS_{mi}(\Delta t)$.

We aim to measure the extent to which the gradient offsets between interval t and $t + \Delta t$, i.e., $OS_{mi}(\Delta t)$, affect the future gradient offset $fOS_{mi}(\Delta t)$, over the interval $t + \Delta t$ and $t + 2\Delta t$ of a particular marker m . The process of obtaining gradient offsets is graphically depicted in Figure 2. We model the gradient offset relationship between the current and next follow-up interval by an over-determined system of linear equations modulated by the vector of linear prediction coefficients μ . For a particular marker m , this relationship is represented by

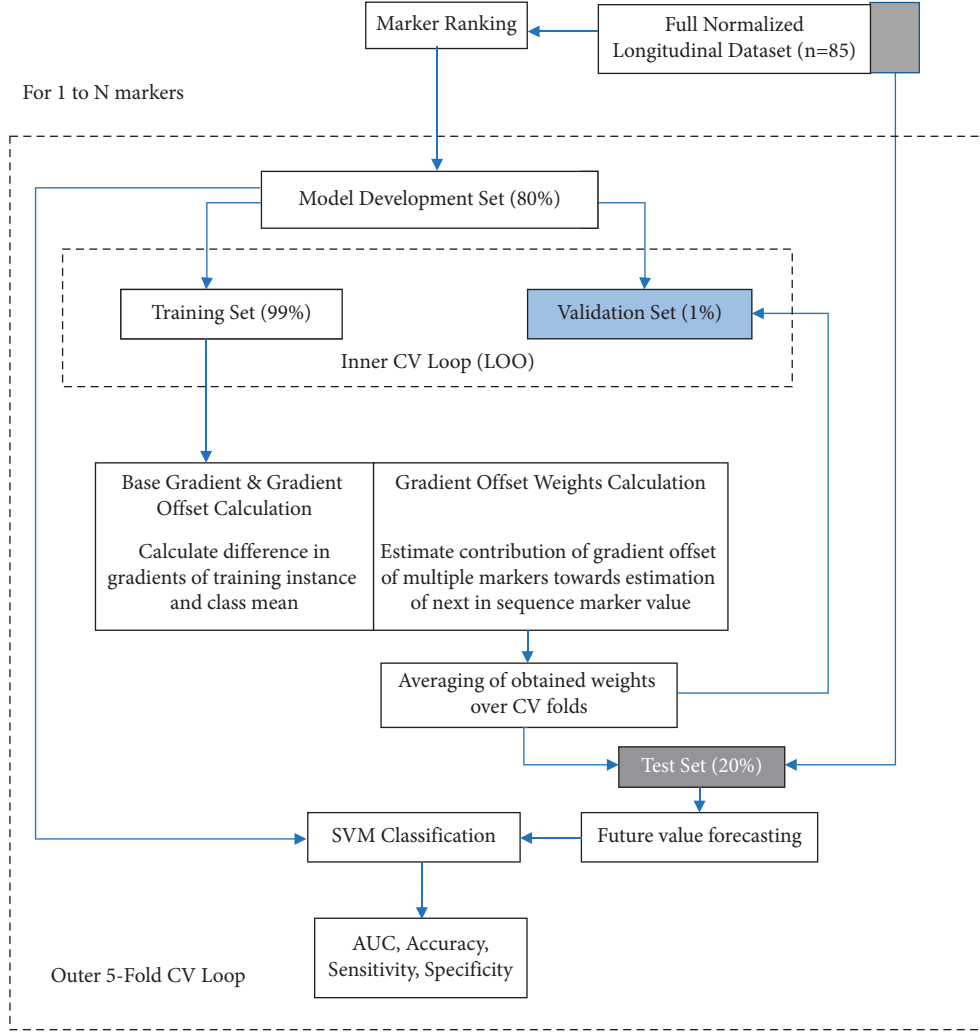
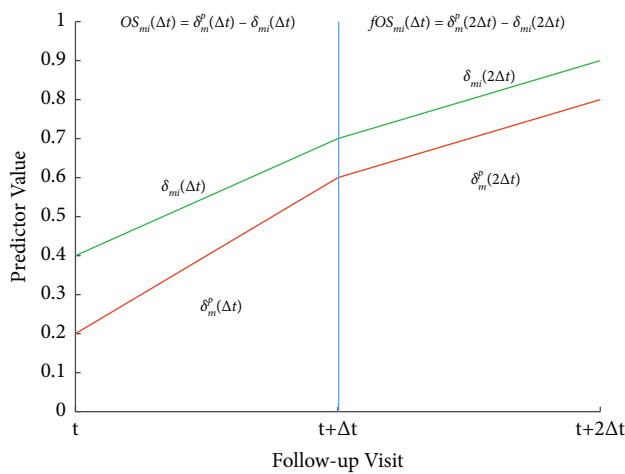


FIGURE 1: System overview.



— Average MCIP Trajectory
— m^{th} Predictor Trajectory of an MCIP subject

FIGURE 2: Gradient offset calculation.

$$A \times \mu = B_m, \quad \text{where } A = \begin{bmatrix} OS_{11}(\Delta t) & \cdots & OS_{N1}(\Delta t) \\ \vdots & \ddots & \vdots \\ OS_{1n}(\Delta t) & \cdots & OS_{Nn}(\Delta t) \end{bmatrix},$$

$$\mu = \begin{bmatrix} \mu_{11} \\ \vdots \\ \mu_{Nn} \end{bmatrix} \text{ and } B = \begin{bmatrix} fOS_{m1}(\Delta t) \\ \vdots \\ fOS_{mn}(\Delta t) \end{bmatrix}, \text{ for } m = 1, \dots, N. \quad (1)$$

We solve equation (1) for μ such that the squared error between predicted and actual offsets is minimized. The derivative of the squared error is equated to 0, and the final formulation of μ is obtained by $\mu = (A^T A)^{-1} A^T B_m$. μ is the vector containing the values of linear prediction weights quantifying the effects of heterogeneous, multivariate markers on the future values of the marker under consideration. For an optimal model generation, we average the obtained weights over LOO validation folds.

3.3. Future Value Forecasting. In this work, we forecast the future clinical changes of an MCI subject using the previous two consecutive follow-up readings of heterogeneous markers to enhance classification performance for MCIp vs. MCIs. Let y be a test instance with two known predictor readings for a predictor m as $y_m(t)$ and $y_m(t + \Delta t)$ and piecewise linear relation between them. We calculate the future values, $y_m(t + 2\Delta t)$, using Euclidean geometry according to

$$\begin{aligned} y_m(t + 2\Delta t) &= y_m(t + \Delta t) + \omega_m * t, \quad \text{where } \omega_m \\ &= \alpha_m + \beta_m \text{ for } m = 1, \dots, N. \end{aligned} \quad (2)$$

Here, ω_m is the piecewise gradient between $(t + \Delta t)$ and $(t + 2\Delta t)$ of a marker m . We calculate ω_m by adjusting the weighted gradient offset β_m to the base gradient α_m . The methods to obtain α_m and β_m are detailed in the following.

3.3.1. Base Gradient α_m . We evaluate the points $y_m(t)$ and $y_m(t + \Delta t)$ for their belongingness to one of the groups, MCIp or MCIs, using the known time-point marker values and annual change in those values as features. These features are fed to a linear support vector machine (SVM) classifier. SVM is selected for binary classification based on limited dataset availability and the robustness of the said classifier [43]. Upon the suggestion of the base group by the SVM classifier, we substitute the average annual change, $\delta_m^c(\Delta t)$, of the respective class for α_m .

3.3.2. Weighted Gradient Offset β_m . It is important to incorporate the effects of multivariate markers in forecasting future marker readings. In this study, we integrate the multivariate effects into the gradient offset measure which is added in α_m . The gradient between the recorded follow-up readings $\delta_{y_m}(\Delta t)$ of y_m is calculated by subtracting $y_m(t)$ from $y_m(t + \Delta t)$ and dividing it by Δt . Gradient offset for the marker OS_{y_m} is calculated by $OS_{y_m} = \alpha_m - \delta_{y_m}$, which quantifies the difference between the base gradient and the actual predictor gradient.

To incorporate the effects of multiple predictors onto the future value of the predictor under consideration, we multiply the vector of linear prediction weights μ by the offset vector OS_{y_m} to obtain β_m according to

$$\beta_m = \mu \times OS_{y_m}^T, \quad \text{for } m = 1, \dots, N. \quad (3)$$

We substitute the obtained values of α_m and β_m in equation (2) to forecast the following value of the predictor m of the test MCI subject. This method is repeated for all N markers being considered in the marker subset to obtain complete time-domain trajectories for the multivariate markers.

3.4. Trajectory Classification. Once the complete time-point biomarker trajectories for individual MCI patients are obtained, we classify them as MCIp and MCIs using a standard linear SVM classifier. After classification, we assign a class

label and likelihood values for the respective class to each MCI individual.

3.5. Performance Evaluation. To analyze the effectiveness of future biomarker values forecasting, we record the mean absolute error (MAE) between the predicted and observed values. Later, to quantify the performance of MCIp vs. MCIs classification, we document the area under ROC (AUC), accuracy, sensitivity, and specificity of the results.

4. Results and Discussion

To fully assess the efficiency of our methods, we design three different types of experiments using the following data:

- (i) NM only
- (ii) MRI markers only
- (iii) NM and MRI combined

For in-depth analysis, we calculate the ground truth (GT) performance scores also in which actual longitudinal marker values, as recorded by ADNI, are used to classify MCIp vs. MCIs, thus overriding the future value-forecasting step. Hence, the GT performance metrics can be used as the benchmark metrics. The aggregate observations from these experiments are described in the following.

4.1. Observations about Marker Ranking. Table 4 lists the top five features noted at 6-month- and 1-year-ahead of conversion predictions. The two NM RAVLT and ADAS display the lowest p values, indicating their valuable significance in MCI-to-AD conversion prediction in both long and short periods. We further note that brain volumes are more significant when performing short-term predictions, whereas average cortical thickness proves to be more significant for 1-year-ahead predictions.

4.2. Observations about Future Value Forecasting. The primary objective of this paper is to accurately forecast the marker readings at a future time point based on preceding follow-up data to aid in better MCI-to-AD conversion prediction. We expect that accurate forecasting of a marker reading will decrease the difference between the GT and the proposed system's classification performance. Mean absolute error (MAE) is the metric used to quantify the error between actual and predicted marker values. In Figure 3, we present the trends observed in MAE with the increasing number of predictors for 6-month- and 1-year-ahead value forecasting. In all of the experiments designed (NM, MRI, and NM + MRI), we observe that initially, the forecasting errors increase and are discontinuous but decrease and stabilize later as the number of predictors crosses a certain threshold. Of 257, 98 features present a p value of less than 0.05, indicating that the addition of significant markers in the feature set for future value forecasting will enhance the system's classification performance. The minimal effect of low-significance features on MAE when the feature set

TABLE 4: Top ranked markers and their p values.

Rank	6 months		1 year	
	Marker	p values	Marker	p values
1	RAVLT	8.6×10^{-12}	RAVLT	4.45×10^{-9}
2	ADAS	2.1×10^{-10}	ADAS	2.6×10^{-8}
3	Volume of right amygdala	9.1×10^{-07}	Cortical thickness of left isthmus cingulate	8.3×10^{-6}
4	Volume of right entorhinal	1.7×10^{-06}	Cortical thickness average of left fusiform	1.78×10^{-5}
5	Volume of left amygdala.	2.0×10^{-06}	Cortical thickness average of right isthmus cingulate	2.7×10^{-05}

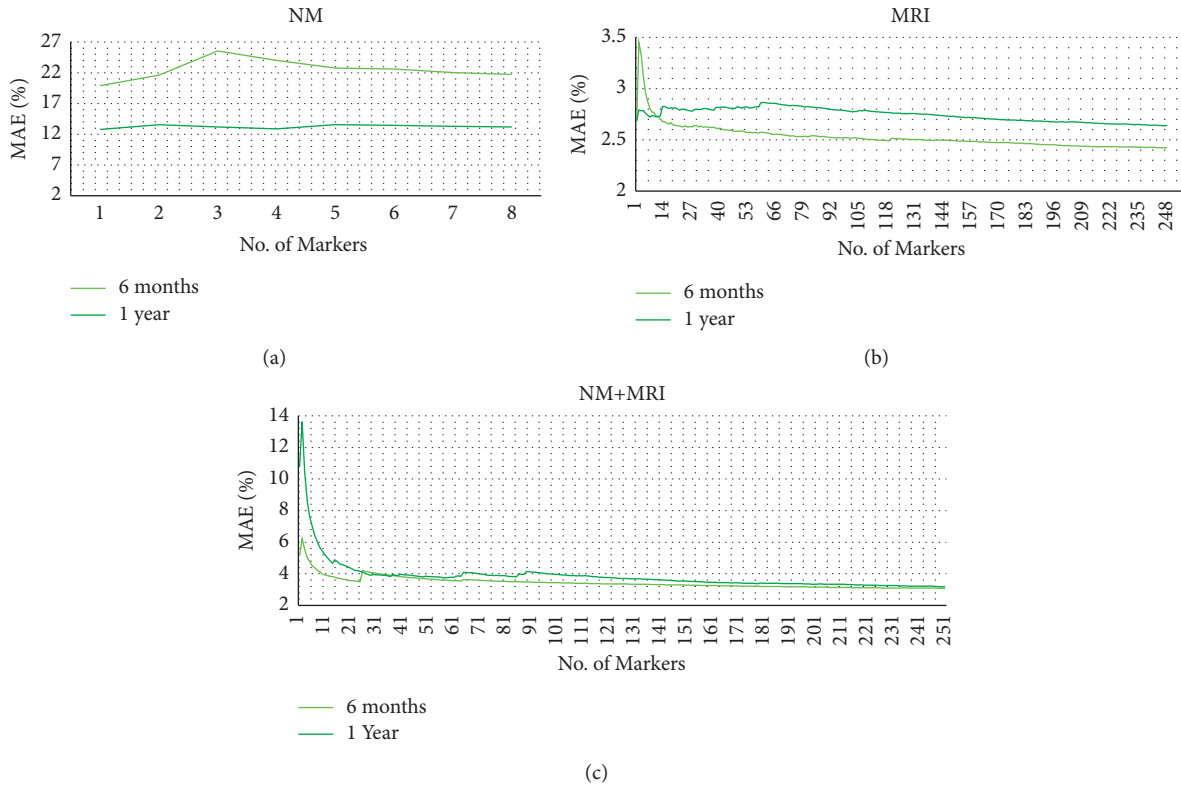


FIGURE 3: Mean absolute error (MAE) versus number of predictors in the predictor subset for 6-month- and 1-year-ahead future value forecasting using (a) NM only, (b) MRI only, and (c) NM and MRI combined.

crosses a certain number of predictors is visible through the tailing of the MAE graphs.

In Figure 3(a), when NM alone is considered for forecasting future time-point readings, the forecasting error reaches the highest (MAE > 12%) amongst all experiments. It reinforces that human behavior is unpredictable, and modeling it can be erroneous, especially over a short (6 months) follow-up duration. On the contrary, in Figure 3(b), experiments considering MRI measures alone report the lowest MAE (<4%) values. The low MAE in MRI measure alone highlights the linear predictability and irreversibility of brain atrophy. 6-month-ahead prediction of MRI markers is more accurate than 1-year-ahead prediction as brain atrophy is a relatively slow process, and minor change occurs between a closer follow-up duration than a longer follow-up duration. In Figure 3(c), we show that the error in forecasting NMs can be reduced by leveraging the power of more predictable MRI measures. The proposed gradient offset method for future value forecasting positively aids in

predicting the future course of a marker by employing differential information from multimodal features.

In Figure 4, we present a comparative summary of MAE of forecasted measures against the classification model, delivering the highest classification AUC. For 6-month-ahead prediction, the NM classification model with 2 markers results in MAE of 21.6% (p value 2.15×10^{-5}). The MRI classification model with 13 markers presents the lowest MAE of 2.67% (p value 0.0001), while the bimodal setup delivers MAE of 3.64% (p value 0.001) with 19 markers. On the other hand, for 1-year-ahead conversion prediction, we record an MAE of 12.8% (p value 0.0001) using the NM classification model with 4 features, while the best MRI model displays an MAE of 2.6% (p value 0.001) with 13 features. Likewise, the bimodal model with 9 biomarkers delivers an MAE of 3.2% (p value 0.002). We observe lower forecasting errors over a comparatively longer follow-up duration, hence strengthening the fact that longer follow-up durations enhance the predictability of a marker.

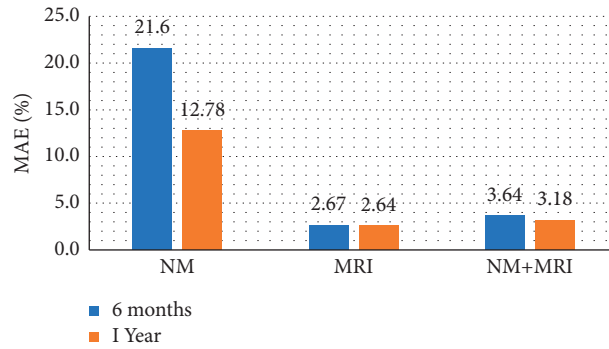


FIGURE 4: Mean absolute error (MAE) recorded for best performing models in experimental setups for 6-month- and 1-year-ahead marker forecasting.

4.3. Observations about MCIp vs. MCIs Classification Using Longitudinal Data. Figure 5 plots the trends in average AUC (blue lines) and accuracy (yellow lines) metrics of both GT (dotted lines) and proposed (solid lines) classification method for 6-month-ahead (first column) and 1-year-ahead (second column) prediction. From all graphs of Figure 5, we can collectively note that initially, classification performance augments as the number of predictors increases but later, we observe no significant change due to the addition of less-significant features. An interesting observation is made that MRI measures, despite being accurately forecasted, deliver the least classification AUC. This observation is as per previous reports by Brooks et al., which state that longitudinal biological markers are less efficient in capturing AD dynamics than cognitive scores. The same was concluded by Gomar et al. [5, 6] and Ewers et al. [44]. The hypothetical predictor model presented by Sperling et al. [1] also depicts a sharper decline in cognitive performance as compared to structural atrophy of the brain. However, the classification power of our system is the highest when NM measures are used in conjunction with MRI measures which is mentioned in Table 5. In 6-month- and 1-year-ahead predictions, the proposed method delivers a maximum AUC of 91.2% (p value 4.7×10^{-6}) and 95.7% (p value 4.1×10^{-7}) using 19 and 9 bimodal features, respectively. In Figure 6, we display the difference between GT AUC and AUC obtained through our system, which is maximum in the case of NM and least in the case of MRI measures. This observation stems from the trends noted earlier that NM are forecasted with the highest error while MRI measures are more accurately forecasted. Table 6 entails the accuracy, sensitivity, and specificity of the proposed method using the optimal model for both 6-month- and 1-year-ahead prediction. The bimodal setup delivers improved sensitivity and specificity as compared to unimodal models.

The above-stated performance metrics conclude that longer follow-up time and multimodal predictors are favorable for future value forecasting and predicting early AD onset. A combination of biomarkers and cognitive scores augment the confidence in the progressive nature of the underlying disorder. ADAS and RAVLT prove to be the two

most informative markers for the said conversion task. These findings have been advocated by many others, including [9, 14, 15, 44].

Table 7 shows a brief comparison of current results with state-of-the-art results obtained on the longitudinal ADNI dataset mainly using NM and MRI data. AUC is considered for comparison to account for imbalance in the data. Misra et al. [8] and Davatzikos et al. [18] employed MRI images to assign abnormality scores to MCIp and MCIs patients. High-dimensional classification methods were applied to the scores to detect MCI progression to AD. They presented respective AUCs of 77% and 73.4% only.

Similarly, Hinrichs et al. [45] proposed a multimodal disease marker using various modalities of data (MRI, PET, and cognitive scores) coupled with SVM classification to identify MCIp patients. The experiments revealed that the proposed technique using longitudinal MRI data produced a maximum MCIp classification AUC of 79%. Moradi et al. [11] also promoted a framework for an aggregate biomarker for MCI conversion prediction using MR imaging data and cognitive scores adjusted with age. The aggregate biomarker devised using random forests presented an AUC of 90.2%. However, the use of high-dimensional imaging data lowers the efficacy of systems.

On the contrary, Zhang et al. [9] attempted to predict future clinical change in a biomarker leveraging MRI and PET data and identified MCIp with an AUC of 76.8% using multikernel SVM. Other attempts of predicting longitudinal trajectories of biomarkers have also been recently reported [49]. Spasov et al. [46] applied the deep learning approach to MRI and clinical data to perform the said classification with an AUC of 92.5%.

While most studies focus on combining multimodal data and projecting them into a single space, we believe that considering predictors independently can preserve more information required for MCIp identification. Owing to the “lag” between brain atrophy and cognitive decline, when discriminating predictors are given independently to the classifier, enhanced classification performance is obtained as recorded. The method proposed in this paper has recorded the highest classification AUC of 95.7% predominantly because of involving the forecasted marker value at the time of conversion as a classification feature. Our approach is

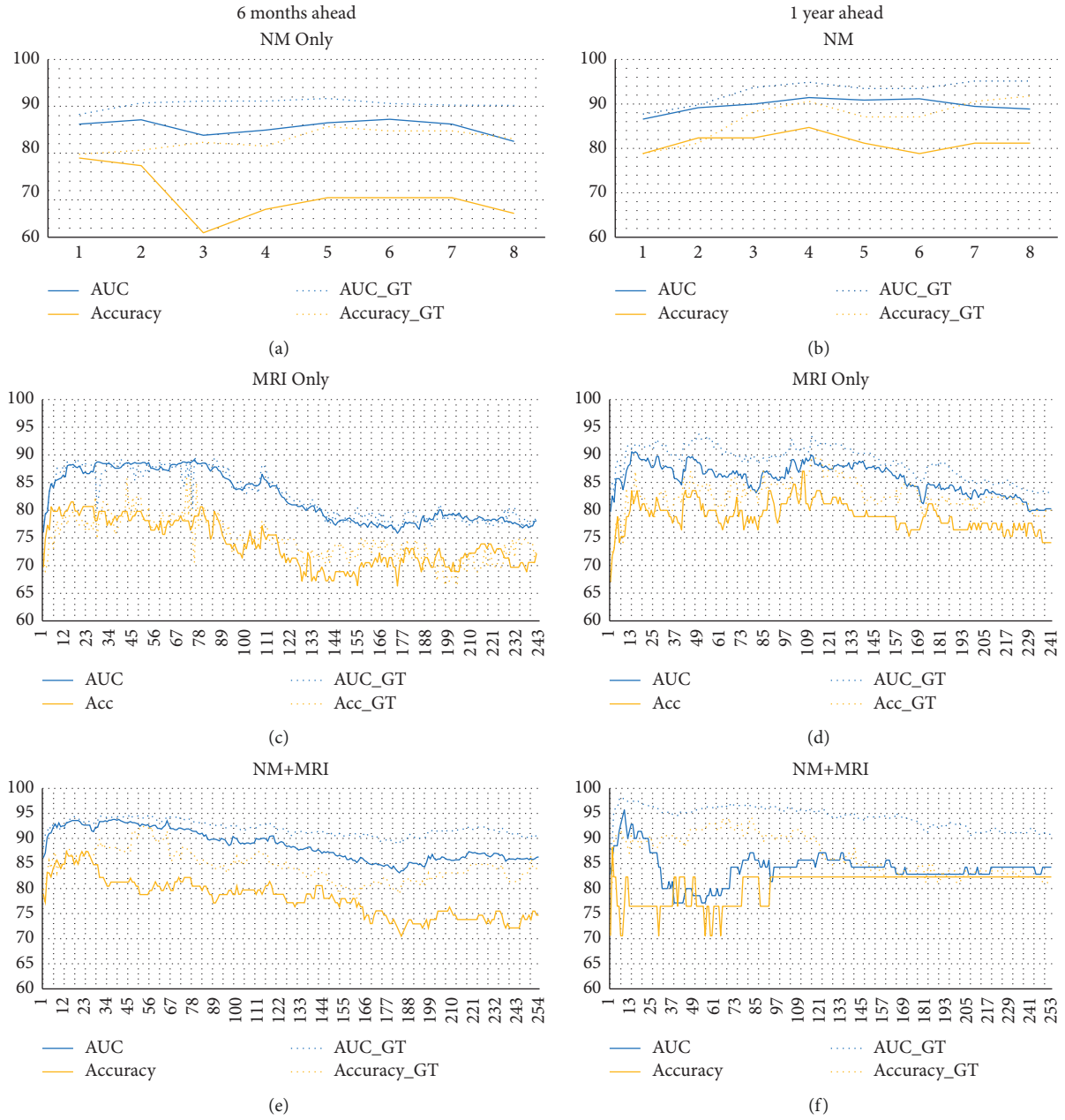


FIGURE 5: Trends noted in GT and obtained AUC and accuracy for 6-month-ahead (Column 1) and 1-year-ahead (Column 2) conversion prediction.

TABLE 5: Comparison of ground truth (GT) AUC and obtained AUC using the proposed method.

	Size of marker set	6 months		1 year		
		GT AUC (p value)	Proposed AUC (p value)	GT AUC (p value)	Proposed AUC (p value)	
NM	2	90.7% (1.7×10^{-5})	87.1% (1.9×10^{-5})	4	93.7% (4.1×10^{-6})	91.5% (5.1×10^{-6})
MRI	13	88.71% (4×10^{-6})	88% (1.3×10^{-6})	13	91.4% (8.7×10^{-7})	90.5% (1.02×10^{-6})
NM + MRI	19	94.2% (4.8×10^{-6})	91.2% (4.6×10^{-6})	9	98% (8.5×10^{-8})	95.7% (4.05×10^{-7})

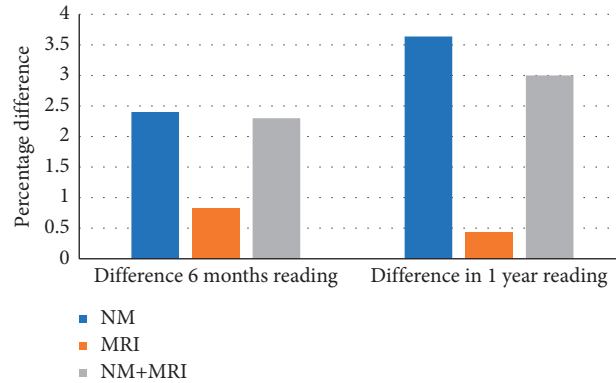


FIGURE 6: Difference in ground truth and obtained classification AUC for 6-month- and 1-year-ahead conversion prediction.

TABLE 6: Average MCIp vs. MCIs performance metrics.

	6 months ahead			1 year ahead		
	Accuracy (%) (p value)	Sensitivity (%)	Specificity (%)	Accuracy (%) (p value)	Sensitivity (%)	Specificity (%)
NM	77.0 (9×10^{-5})	73.5	80	84.7 (9.3×10^{-6})	82.9	86
MRI	79.8 (1.2×10^{-5})	71.8	71.8	83.5 (1.69×10^{-5})	74.28	90
NM + MRI	84.9 (3.9×10^{-5})	81.7	87.1	81 (1.43×10^{-5})	85.7	70

TABLE 7: Comparison with state-of-the-art results based on the ADNI dataset.

Author	Predictors	Dataset size	Validation	AUC (%)
Misra [8]	MR images	27 MCIp, 76 MCIs	LOO	77
Davatzikos [17]	MR images	69 MCIp, 170 MCIs	5-fold	73.4
Hinrichs [45]	MRI	48 AD, 66 CN, 119 MCI	LOO	79
Zhang [9]	MR images	38 MCIp, 50 MCIs	10-fold	76.8
Moradi [11]	MRI, NM, age	164 MCIp, 100 MCIs	10-fold	90
Arco [30]	MRI, NM	73 MCIp, 61 MCIs	LOO	79.23
Spasov [46]	MRI, NM	16 MCIp, 16 MCIs	10-fold	92.5
Guo [47]	MRI	43 MCIp, 35 MCIs	LOO	92.31
Platero [48]	MRI, NM	206 MCIp, 215 MCIs	10-fold	85.5
Proposed	MRI, NM	35 MCIp, 50 MCIs	5-fold	95.7

flexible and can be extended to incorporate other longitudinal predictors very easily.

5. Conclusion

In the current study, we presented a pipeline to identify the MCI patients who will develop AD in the future. Cognitive status and MRI measures were separately and collectively employed. The longitudinal trends observed in the predictors are utilized to forecast the future variability in the predictor values. The time-series data of markers are used to classify a subject as a progressor or a stable feature. Our best bimodal model forecasted the future values with a high precision (MAE: 3.64%) and yielded an excellent classification performance (AUC: 91.2%, accuracy: 84%) for 6-month-ahead conversion prediction. More stable results were obtained for 1-year-ahead AD diagnosis (MAE: 3.18%, AUC: 95.7%, accuracy: 81%). We conclude that it will be interesting to incorporate other predictors like PET and CSF measures in this framework and repeat the process for longer-term-ahead prediction. The use of a genetic algorithm or neural network for weight learning is

also an explorable avenue in this domain to improve the results.

Data Availability

In this paper, we have used data provided by Alzheimer’s Disease Neuroimaging Initiative (ADNI). ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

Data collection and sharing for this project were funded by Alzheimer’s Disease Neuroimaging Initiative (ADNI) (National Institutes of Health under Grant U01 AG024904) and DOD ADNI (Department of Defense, under award number W81XWH-12-2-0012). ADNI is funded by the National

Institute on Aging and the National Institute of Biomedical Imaging and Bioengineering and through generous contributions from the following companies: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private-sector contributions are facilitated by the Foundation for the National Institutes of Health (<https://ida.loni.usc.edu/collaboration/access/www.fnih.org>). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California.

Supplementary Materials

In ADNI, every subject is identified with a unique Roster ID (RID). The RIDs of the subjects considered in this study as MCIp and MCIs are available as supplementary file S1. (*Supplementary Materials*)

References

- [1] R. A. Sperling, P. S. Aisen, L. A. Beckett et al., "Toward defining the preclinical stages of alzheimer's disease: recommendations from the national institute on aging and the alzheimer's association workgroup," *Alzheimer's and Dementia*, vol. 7, no. 3, pp. 1–13, 2011.
- [2] J. J. Manly, M.-X. Tang, N. Schupf, Y. Stern, J.-P. G. Vonsattel, and R. Mayeux, "Frequency and course of mild cognitive impairment in a multiethnic community," *Annals of Neurology*, vol. 63, no. 4, pp. 494–506, 2008.
- [3] G. McKhann, D. Drachman, M. Folstein, R. Katzman, D. Price, and E. M. Stadlan, "Clinical diagnosis of alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of department of health and human services task force on alzheimer's disease," *Neurology*, vol. 34, no. 7, pp. 939–944, 1984.
- [4] M. Riverol and O. L. López, "Biomarkers in alzheimer's disease," *Frontiers of Neurology*, vol. 2, no. 46, 2011.
- [5] J. J. Gomar, M. T. Bobes-Bascaran, C. Conejero-Goldberg, and P. Davies, "Utility of combinations of biomarkers, cognitive markers and risk factors to predict conversion from MCI to AD in patients in the ADNI," *Archives of General Psychiatry*, vol. 68, no. 9, pp. 961–969, 2011.
- [6] J. J. Gomar, C. C. Goldberg, P. Davies, and T. E. Goldberg, "Extension and refinement of the predictive value of different classes of markers in ADNI: four year follow-up data," *Alzheimer's and Dementia*, vol. 10, no. 6, pp. 704–712, 2014.
- [7] B. C. Dickerson, R. A. Sperling, B. T. Hyman, M. S. Albert, and D. Blacker, "Clinical prediction of alzheimer disease dementia across the spectrum of mild cognitive impairment," *Archives of General Psychiatry*, vol. 64, no. 12, pp. 1443–1450, 2007.
- [8] C. Misra, Y. Fan, and C. Davatzikos, "Baseline and longitudinal patterns of brain atrophy in MCI patients, and their use in prediction of short-term conversion to AD: results from ADNI," *NeuroImage*, vol. 44, no. 4, pp. 1415–1422, 2009.
- [9] D. Zhang and D. Shen, "Predicting future clinical changes of MCI patients," *PLoS One*, vol. 7, no. 3, Article ID 33182, 2012.
- [10] L. M. Shaw, H. Vanderstichele, M. Knapik-Czajka et al., "Cerebrospinal fluid biomarker signature in alzheimer's disease neuroimaging," *Annals of Neurology*, vol. 65, no. 4, pp. 403–413, 2009.
- [11] E. Moradi, A. Pepe, C. Gaser, H. Huttunen, and J. Tohka, "Machine learning framework for early MRI-based alzheimer's conversion," *NeuroImage*, vol. 104, pp. 398–412, 2015.
- [12] D. Zhang, D. Shen, and The Alzheimer's Disease Neuroimaging Initiative, "Multimodal multi-task learning for joint prediction of multiple regression and classification variables in alzheimer's disease," *NeuroImage*, vol. 59, no. 2, pp. 895–907, 2012.
- [13] D. Zhang, Y. Wang, L. Zhou, H. Yuan, D. Shen, and The Alzheimer's Disease Neuroimaging Initiative, "Multimodal classification of alzheimer's disease and mild cognitive impairment," *NeuroImage*, vol. 55, no. 3, pp. 856–867, 2011.
- [14] L. G. Brooks and D. A. Loewenstein, "Assessing the progression of mild cognitive impairment to alzheimer's disease: current trends and future directions," *Alzheimer's Research & Therapy*, vol. 2, no. 5, p. 28, 2010.
- [15] R. Casanova, F. C. Hsu, K. M. Sink et al., "Alzheimer's disease risk assessment using large-scale machine learning methods," *PLoS One*, vol. 8, no. 11, Article ID e77949, 2013.
- [16] Y. Cui, B. Liu, S. Luo et al., "Identification of conversion from mild cognitive impairment to alzheimer's disease using multivariate predictors," *PLoS One*, vol. 6, no. 7, 2011.
- [17] C. Davatzikos, P. Bhatt, L. M. Shaw, K. N. Batmanghelich, and J. Q. Trojanowski, "Prediction of MCI to AD conversion, via MRI, CSF biomarkers, and pattern classification," *Neurobiology of Aging*, vol. 32, no. 12, pp. 2322–2327, 2011.
- [18] C. Gaser, K. Franke, S. Klöppel, N. Koutsouleris, and H. Sauer, "BrainAGE in mild cognitive impaired patients: predicting the conversion to alzheimer's disease," *PLoS One*, vol. 8, no. 6, Article ID e67346, 2013.
- [19] H.-I. Suk and D. Shen, "Deep learning-based feature representation for AD/MCI classification," *Medical Image Computing and Computer Assisted Intervention*, vol. 16, no. 2, pp. 583–590, 2013.
- [20] C. Davatzikos, F. Xu, Y. An, Y. Fan, and S. M. Resnick, "Longitudinal progression of alzheimer's-like patterns of atrophy in normal older adults: the SPARE-AD index," *Brain*, vol. 132, no. 8, pp. 2026–2035, 2009.
- [21] B. C. Dickerson and D. A. Wolk, "Biomarker-based prediction of progression in MCI: comparison of AD signature and hippocampal volume with spinal fluid amyloid- β and tau," *Frontiers in Ageing Neuroscience*, vol. 5, no. 55, 2013.
- [22] M. Huang, W. Yang, Q. Feng, and W. Chen, "Longitudinal measurement and hierarchical classification framework for the prediction of AD," *Nature, Scientific Reports*, vol. 7, Article ID 39880, 2017.
- [23] S. H. Lee, A. H. Bachman, D. Yu, J. Lim, and B. A. Ardekani, "Predicting progression from mild cognitive impairment to

- alzheimer's disease using longitudinal callosal atrophy," *NeuroImaging*, vol. 2, pp. 1–7, 2016.
- [24] H. Runtti, J. Mattila, M. van Gils, J. Koikkalainen, H. Soininen, and J. Lötjönen, "Quantitative evaluation of disease progression in a longitudinal mild cognitive impairment cohort," *Journal of Alzheimer's Disease*, vol. 39, no. 1, pp. 49–61, 2014.
- [25] M. Guo, Y. Li, W. Zheng et al., "A novel conversion prediction method of MCI to AD based on longitudinal dynamic morphological features using ADNI structural MRIs," *Journal of Neurology*, vol. 267, no. 10, pp. 2983–2997, 2020.
- [26] L. Teng, Y. Li, Y. Zhao et al., "Predicting MCI progression with FDG-PET and cognitive scores: a longitudinal study," *BMC Neurology*, vol. 20, no. 148, 2020.
- [27] S. F. Cook and R. R. Bies, "Disease progression modeling: key concepts and recent developments," *Current Pharmacology Reports*, vol. 2, no. 5, pp. 221–230, 2016.
- [28] E. Lawrence, C. Vegvari, A. Ower, C. Hadjichrysanthou, F. De Wolf, and R. M. Anderson, "A systematic review of longitudinal studies which measure alzheimer's disease biomarkers," *Journal of Alzheimer's Disease*, vol. 59, no. 4, pp. 1359–1379, 2017.
- [29] D. Zhang and D. Shen, "Semi-supervised multimodal classification of alzheimer's disease," in *Proceedings of the IEEE International Symposium on Biomedical Imaging*, pp. 1628–1631, Chicago, IL, USA, March 2011.
- [30] J. E. Arco, J. Ramírez, J. M. Górriz, C. G. Puntonet, and M. Ruz, "Short-term prediction of MCI to AD Conversion based on longitudinal MRI analysis and neuropsychological tests," in *Innovation in Medicine and Healthcare 2015*, pp. 385–394, Springer, Berlin, Germany, 2016.
- [31] R. C. Petersen, G. E. Smith, S. C. Waring, R. J. Ivnik, E. G. Tangalos, and E. Kokmen, "Mild cognitive impairment clinical characterization and outcome," *JAMA Neurology*, vol. 56, no. 3, pp. 303–308, 1999.
- [32] ADNI, "ADNI general procedures manual," 2010, https://adni.loni.usc.edu/wp-content/uploads/2010/09/ADNI_GeneralProceduresManual.pdf.
- [33] T. Pereira, F. L. Ferreira, D. S. Cardoso et al., "Neuropsychological predictors of conversion from mild cognitive impairment to alzheimer's disease: a feature selection ensemble combining stability and predictability," *BMC Medical Informatics and Decision Making*, vol. 18, no. 137, p. 137, 2018.
- [34] F. Ségonne, A. M. Dale, E. Busa et al., "A hybrid approach to the skull stripping problem in MRI," *NeuroImage*, vol. 22, no. 3, pp. 1060–1075, 2004.
- [35] B. Fischl, D. H. Salat, E. Busa et al., "Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain," *Neuron*, vol. 33, no. 3, pp. 341–355, 2002.
- [36] R. S. Desikan, F. Ségonne, B. Fischl et al., "An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest," *NeuroImage*, vol. 31, no. 3, pp. 968–980, 2006.
- [37] M. Reuter, N. J. Schmansky, H. D. Rosas, and B. Fischl, "Within-subject template estimation for unbiased longitudinal image analysis," *NeuroImage*, vol. 61, no. 4, pp. 1402–1418, 2012.
- [38] M. Reuter and B. Fischl, "Avoiding asymmetry-induced bias in longitudinal image processing," *NeuroImage*, vol. 57, no. 1, pp. 19–21, 2011.
- [39] M. Reuter, H. D. Rosas, and B. Fischl, "Highly accurate inverse consistent registration: a robust approach," *NeuroImage*, vol. 53, no. 4, pp. 1181–1196, 2010.
- [40] C. R. Jack, M. A. Bernstein, N. C. Fox et al., "The alzheimer's disease neuroimaging initiative: MRI methods," *Journal of Magnetic Resonance Imaging*, vol. 27, no. 4, pp. 685–691, 2008.
- [41] K. Ito, B. Corrigan, Q. Zhao et al., "Disease progression model for cognitive deterioration from alzheimer's disease neuroimaging initiative database," *Alzheimer's and Dementia*, vol. 7, no. 2, pp. 151–160, 2011.
- [42] S. G. Popescu, A. Whittington, R. N. Gunn et al., "Nonlinear biomarker interactions in conversion from mild cognitive impairment to alzheimer's disease," *Human Brain Mapping*, vol. 41, no. 15, pp. 4406–4418, 2020.
- [43] L. Xu, G. Liang, C. Liao, G.-D. Chen, and C.-C. Chang, "An efficient classifier for alzheimer's disease genes identification," *Molecules*, vol. 23, no. 12, p. 3140, 2018.
- [44] M. Ewers, C. Walsh, J. Q. Trojanowski et al., "Prediction of conversion from mild cognitive impairment to alzheimer's disease dementia based upon biomarkers and neuropsychological test performance," *Neurobiology of Aging*, vol. 33, no. 7, pp. 1203–1214, 2012.
- [45] C. Hinrichs, V. Singh, G. Xu, and S. C. Johnson, "Predictive markers for AD in a multi-modality framework: an analysis of mci progression in ADNI population," *NeuroImage*, vol. 55, no. 2, pp. 574–589, 2011.
- [46] S. Spasov, L. Passamonti, A. Duggento, P. Liò, and N. Toschi, "A parameter-efficient deep learning approach to predict conversion from mild cognitive impairment to alzheimer's disease," *NeuroImage*, vol. 189, pp. 276–287, 2019.
- [47] C. Platero and M. C. Tobar, "Predicting alzheimer's conversion in mild cognitive impairment patients using longitudinal neuroimaging and clinical markers," *Brain Imaging and Behavior*, 2020.
- [48] L. L. Rackett, "Statistical disease progression modeling in alzheimer disease," *Frontiers in Big Data*, vol. 3, 2020.
- [49] H. Shi, D. Ma, Y. Nie et al., "Early diagnosis of alzheimer's disease on ADNI data using novel longitudinal score based on functional principal component analysis," *Journal of Medical Imaging*, vol. 8, no. 2, 2021.