ORIGINAL ARTICLE

HEALTH CARE SCIENCE

Machine learning in neurological disorders: A multivariate LSTM and AdaBoost approach to Alzheimer's disease time series analysis

Muhammad Irfan[1] | **Seyed Shahrestani[1]** | **Mahmoud Elkhodr[2]**

[1]School of Computer, Data and Mathematical Sciences, Western Sydney University, Sydney, Australia

[2]School of Engineering and Technology, Central Queensland University, Sydney, Australia

Correspondence
Muhammad Irfan, Western Sydney University, 56 Second Ave, Bldg XB Level 2, Room 15 Kingswood, Sydney, NSW 2747, Australia.
Email: 19918600@student.westernsydney.edu.au

Abstract

Introduction: Alzheimer's disease (AD) is a progressive brain disorder that impairs cognitive functions, behavior, and memory. Early detection is crucial as it can slow down the progression of AD. However, early diagnosis and monitoring of AD's advancement pose significant challenges due to the necessity for complex cognitive assessments and medical tests.

Methods: This study introduces a data acquisition technique and a preprocessing pipeline, combined with multivariate long short-term memory (M-LSTM) and AdaBoost models. These models utilize biomarkers from cognitive assessments and neuroimaging scans to detect the progression of AD in patients, using The AD Prediction of Longitudinal Evolution challenge cohort from the Alzheimer's Disease Neuroimaging Initiative database.

Results: The methodology proposed in this study significantly improved performance metrics. The testing accuracy reached 80% with the AdaBoost model, while the M-LSTM model achieved an accuracy of 82%. This represents a 20% increase in accuracy compared to a recent similar study.

Discussion: The findings indicate that the multivariate model, specifically the M-LSTM, is more effective in identifying the progression of AD compared to the AdaBoost model and methodologies used in recent research.

KEYWORDS
Alzheimer's disease, AdaBoost, cognitive data, multivariate LSTM, neuroimaging data

---

- -

1 | INTRODUCTION

Alzheimer's disease (AD) is an irreversible neurological condition that impairs patients' memory, cognition, and behavior [1]. The complexity of analyzing AD biomarkers and the variability of measures obtained from different imaging modalities challenges the effective early identification and treatment of AD [2]. Given the lack of significant advancements in developing modalities to treat AD, previous research has looked at providing feasible and affordable options to offer AD patients the care and treatment they need. For instance, the studies reported in Gavidia-Bovadilla and colleagues [3–5] aim at helping AD patients in their day-to-day activities and offer some forms of neural training.

However, detecting how AD is gradually progressing in a patient is crucial to implement early intervention plans [6]. Traditional time series-based methods and machine learning (ML) algorithms have been widely employed in the literature to categorize and predict AD severity, such as the works described in Gavidia-Bovadilla and colleagues [3, 4, 7–11]. Despite these encouraging developments, several complex modeling issues remain unexplored [12]. Most of the literature has focused on modeling the course of diseases rather than predicting their progression. This entails estimating the state of disease progression at known times based on the data available or in the event of limited observations, concentrating primarily on identifying the stages of its progression.

Various ML approaches have been used in the literature to find biomarkers for AD progression and optimize detection performance. A widely used approach to address image classification problems is the support vector machine (SVM). SVM was used in several studies to predict AD progression, such as in Zhang and colleagues [13–18]. In Zhang et al. [13], a multitask learning approach combined with SVM was employed to detect AD progression with 73.9% accuracy and 68.6% sensitivity. The study reported in Cheng et al. [14] achieved an accuracy of 79.4% and a sensitivity of 84.5%.

A linear discriminant analysis based on the cortex thickness data was utilized, which showed 63% sensitivity and 76% specificity [19]—the results for AD progression detection improved by integrating multimodality data and removing redundant AD-related biomarkers from modalities. Using a combination of cerebral spinal fluid, magnetic resonance imaging (MRI), and cognitive performance biomarkers, 68.5% accuracy, 53.4% sensitivity, and 77% specificity were noted in the studies reported in Ewers et al. [20] and Kim et al. [21]. These studies also used deep convolutional autoencoders to examine the data analysis process of AD progression using MRI data. While previous studies have investigated modeling AD, the AD progression

issues using recurrent models remain unexplored. Given that AD is a progressive and irreversible disease, it is challenging to detect its progression over time using only the cognitive features of a patient. However, the timely structural changes from cognitive normal (CN) to mild cognitive impairment (MCI) and from MCI to AD can help trace the features affecting the progression of the disease [22].

To this end, as far as our knowledge is concerned, a single study has worked on the detection of patients with a progressive nature [23], which has utilized a univariate recurrent neural network (RNN). In contrast, our study proposes a multivariate long short-term memory (M-LSTM) for the detection of AD subjects with a progressive nature. The study has focused on two types of experiments. The first experiment involves patients who are in a state of progressive disease. Detection is then carried out by confirming their progressive and nonprogressive states. In the second experiment, the study involved only patients whose disease status was updated with time. It then observes the spread of the disease over some time. The performance of the proposed M-LSTM model is evaluated using The AD Prediction of Longitudinal Evolution (TADPOLE) challenge [24]. The results show that the proposed model outperformed the competing models (BiPro and AdaBoost) regarding detection.

The remainder of this paper is structured as follows. Section 2 briefly describes the related studies on predictive modeling. Section 3 presents the details of the proposed methodology. Sections 4 and 5 report the experiments that compare the proposed model's performance with other published approaches. Lastly, Section 6 presents the conclusions of this study.

2 | BACKGROUND

RNNs have been widely used in the literature to identify patterns in time series data [25]. Recent studies have shown the success of RNNs in disease detection and prediction. Examples include using gated recurrent units (GRUs) as a diagnostic tool [26]. Moreover, GRU was also used to predict the medication needed for patients' subsequent visits in the early onset of cardiac failure [27]. LSTM models are also used to predict multiple conditions' onset [28]. In addition, LSTM is used for the diagnostic classification of patients in the pediatric intensive care unit (PICU) [29]. The relevant studies have applied different RNN models for medical detection by using longitudinal data on a patient over time.

Research into the detection of AD progression can be classified into three subareas: AD progression categorization [30], AD progression modeling [31, 32], and AD progression estimation [33]. In the categorization framework, studies

strive to determine the patient's status regarding non-conversion or conversion over time. Those with progressive MCI are distinguished from patients with stable MCI by learning deep nonlinear representations of the neuro-imaging data as the output of a convolutional neural network [34]. In terms of the modeling, a study used a many-to-one extended RNN architecture to facilitate the temporal shift to mimic AD progression in patients on successive visits [35]. The method covered a variety of visit periods of AD subjects. In addition, the Bayesian progression score model is suggested to predict the biomarker trends from normal to MCI and AD [31]. In the context of estimation, a study presented a multisource multitask learner to estimate cognitive scores, such as mini-mental state examination (MMSE) and AD assessment scale-cognitive subscale [2]. These scores may be used to measure the degree and progression of cognitive impairment caused by AD. With the help of recurrent components, valid observations, and static relations considered, the temporal relation for imputation for AD [36].

Contrary to a standard bidirectional RNN, the timing of the inputs to the salient layers is reversed in a forward-and-backward direction. These imputed values were dynamically updated during training until they were optimal (dynamic relation). The study used an RNN for missing value imputation in time series data and then fed the imputed values to an LSTM cell to forecast both the diagnosis of AD and MRI biomarkers [37]. The model inspects the temporal and multivariate relations of measurements (dynamic relations) for missing value imputation (and is called LSTM-I in our comparisons below). Another study proposed a minimal RNN called MinRNN to predict patients' clinical diagnoses, cognition, and ventricular volume [38]. Their model has fewer parameters than other RNN models, such as the LSTM, and therefore is less prone to overfitting [38]. A study modified the LSTM cell by using a progressive module to compute the progression score of each biomarker between the given time point and the baseline through a negative exponential function [39].

By the analysis of recent research, this research proposed a data requisition followed by preprocessing pipeline and M-LSTM model. The main contributions of this study can be summarized as follows:

(a) The proposal of an end-to-end approach uses missing value restoration, data filtration, and normalization to model AD disease progression in a patient.

(b) A novel M-LSTM is utilized with the proposed preprocessing. The model detects the patients in a progressive disease state and classifies whether their AD is in a progression or nonprogression state. In addition, the progression of the disease was examined only in patients whose disease state was updated.

(c) The performance of the model was evaluated using a public data set. The proposed model was also generalized by testing it using cross-validation. The model encompassed several new methods to analyze the disease progression, utilizing cognitive and neuroimaging data. Compared to the state-of-the-art literature [23], our proposed method has almost 20% better classification accuracy.

3 | METHODS

This study examined the detection of AD progression using AdaBoost, an ensemble ML model, and M-LSTM, a deep recurrent network. Two experiments are performed on the TADPOLE challenge data set [24] cohort from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database, which contains ADNI-1, ADNI-2, and ADNI-GO. The first experiment involved patients with AD progression and detected their progressive or nonprogressive states with the help of AdaBoost. In the second experiment, the study considered only the patients whose disease status was updated with time and observed the disease's progression over time. The data set is prepared and preprocessed, followed by progressive nature detection using M-LSTM. Finally, the performance of the proposed methodology is evaluated on the data set, and the results are compared with the state-of-the-art literature.

3.1 | Materials and settings

This study used a data set called the TADPOLE challenge, which includes information from 13,915 visits made by 2155 patients. Of these patients, 409 had just one visit, while the remaining 1746 had multiple visits, totaling 13,506 visits with an average of 7.74 per patient. We observed 535 patients whose disease progressed over time, that is, from CN to MCI or from MCI to AD. These 535 patients were labeled as having progressive disease, and together, they made 5254 visits, averaging 9.8 visits per patient. For the remaining 1211 patients, their disease state stayed the same as it was during their first visit, and they were classified as nonprogressive, meaning their disease did not get worse. Only 142 patients had consulted a doctor in a year.

In comparison, 1069 subjects with a nonprogressive nature have seen a consultant physician in 1 year. In more than 2 years, 894 patients with a nonprogressive nature have seen a consultant. The number of patients with a nonprogressive nature who have visited a specialist doctor in over 3 years is 554. Patients' visits are summarized in Table 1. Given that AD takes about 3 years to develop, a research study has considered nonprogressive patients

who have seen a consultant for over 3 years [35]. A patient is deemed nonprogressive if their disease status has not changed during the 3 years of the study. Table 2 shows some comprehensive statistics about the patients with and without AD. TADPOLE data has been preprocessed and sorted based on the patients' visits. The TADPOLE data set contains vital information such as a patient's initial status, the number of visits per month, the updated status after each visit, and all the cognitive and neuroimaging indicators. Unfortunately, the labels for some visits are missing.

Additionally, some patients' visits had no record, such as no cognitive scores or neuroimaging information were recorded. Discarding the missing data is one of the techniques used to address the problem of missing records. Moving average filling is an alternative technique used to address this problem. This is where missing data is filled from previously labeled data. The moving average filling method has been used in this research to address the missing data problem.

3.2 | Data preprocessing for Experiment 1

The TADPOLE data set contains missing information, necessitating a series of preprocessing steps. Initially, a data filtration process was employed to eliminate less informative entries. Subsequently, a missing data handling technique was applied to impute values for the absent labels, ensuring a comprehensive data set. Lastly, data normalization procedures were executed to standardize the information, facilitating consistent and comparable features across the data set.

3.2.1 | Filtration

Some neuroimaging or cognitive features had several missing values. Therefore, features with more than 50% missing values were removed from the first experiment. Additionally, the records collected from some patients did not provide conclusive neuroimaging or cognitive traits. Hence, visits without a record were removed from the data to prevent the model from overfitting.

3.2.2 | Handle missing data

After filtering, the remaining data were filled using a moving average approach of the previously accessed data. When the MMSE of a patient x from a visit n was noted as missing, the missing data were filled in using

$$MMSE = \begin{cases} MMSE \rightarrow \text{if available} \\ Fill \rightarrow \psi \end{cases}. \quad (1)$$

TABLE 2 Statistics of progressive and nonprogressive AD patients.

Features	Progressive patients	Nonprogressive patients
Total	535	554
Gender		
Male	315	307
Female	220	247
Marital status		
Married	424	396
Never married	13	20
Widowed/divorced	98	132
Race		
Asian	8	10
Black	17	21
White	505	517
Ethnicity		
Hispanic	14	15
Not Hispanic	518	536
Age average	76.06	70.64

Abbreviation: AD, Alzheimer's disease.

TABLE 1 Number of patients and visits in TADPOLE.

Patient group	Number of patients	Number of visits	Average visits per patient	Disease progression
Progressive	535	5254	9.8	Yes
Nonprogressive	1211	8252	6.8	No
One visit	409	409	1	N/A
Total	2155	13,915	6.5	N/A

Abbreviation: TADPOLE, The Alzheimer's Disease Prediction of Longitudinal Evolution; N/A, not applicable.

If the MMSE score had already been recorded, it was not required to fill in the data. However, when MMSE was not available, it was replaced by the mean of the score from all other visits using

$$\psi = \frac{1}{n_x} \sum_{i=1}^{n} \partial_i, \tag{2}$$

where ∂ shows the available MMSE score of the previous visits, and n shows the total number of visits made by patient x.

3.2.3 | Data normalization

After preparing the data, the data were normalized using

$$z = \frac{x - \mu}{\sigma}, \tag{3}$$

where x is the original feature value, μ is the mean, and σ is the standard deviation calculated as

$$\mu = \frac{1}{N} \sum_{i=1}^{x} x_i, \quad \sigma = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)^2}. \tag{4}$$

3.2.4 | Model training for Experiment 1

In the first experiment, an ensembled ML classifier, AdaBoost, was trained on the preprocessed data. To improve the classifier accuracy, AdaBoost combines other classifiers. The cleaned, filtered, and preprocessed data were split into 80:20, and fivefold cross-validation was performed during the model training and testing.

3.3 | Methodology for Experiment 2

In Experiment 2, this study included all the cognitive and neuroimaging features in a time-series pattern to study patients' AD progression. There were subjects in the data set that had single or double visits. Therefore, subjects with more than 3 years of consecutive visits were considered, while the remaining patients were removed. Given that every ML or DL model requires a fixed input length, the total number of visits of all the selected patients was equalized using

$$V = \sum_{(i=1)}^{n} W_i (\mathrm{label}_{i\mu} + \mathrm{label}_{i\sigma}), \tag{5}$$

where

$$W_i = \frac{\mathrm{label}_i}{\mathrm{Total}},$$

label_i is the current label, and $V \in$ Wholenumbers since the number of visits should belong to the set of whole numbers. Patients with more than V visits are dropped from the processing data. Those patients whose number of visits was less than V are equalized with the help of the simple moving average (SMA) defined as follows:

$$
\begin{aligned}
\mathrm{SMA}_{k,\mathrm{next}} &= \frac{1}{k} \sum_{i=n-k+2}^{n+1} p_i, \\
\mathrm{SMA}_{k,\mathrm{next}} &= \frac{1}{k} (p_{n-k+2} + p_{n-k+3} + p_{n-k+4} + \ldots + p_n \\
&\quad + p_{n+1} + p_{n-k+1} - p_{n-k+1}), \\
\mathrm{SMA}_{k,\mathrm{next}} &= \frac{1}{k} (p_{n-k+1} + p_{n-k+2} + \ldots + p_n) \\
&\quad - \frac{p_{n-k+1}}{k} + \frac{p_{n+1}}{k}, \\
\mathrm{SMA}_{k,\mathrm{next}} &= \mathrm{SMA}_{k,\mathrm{current}} + \frac{1}{k} (p_{n+1} - p_{n-k+1}).
\end{aligned}
\tag{6}
$$

According to Equation (6), the moving average filter can be formulated quickly on real-time data using a circular or first in, first out (FIFO) buffer with three arithmetic operations. First, the sampling window is equal to the size of the data set at the first filling of the FIFO or circular buffer. Thus $k = n$, and the average calculation is done as a cumulative moving average. An M-LSTM is trained on the cleaned data set after data preprocessing. The LSTM structure is provided in the following section.

3.3.1 | M-LSTM

The M-LSTM is a well-known DNN model. The fundamental architecture of the M-LSTM, an RNN [27], is shown in Figure 1. A single time-series data set can be classified, and sorted, and detection can be made using the M-LSTM. The ability of M-LSTM to estimate time-series data has also been explored and utilized previously, such as in Schafer and Graham [25]. The output value was first calculated in the M-LSTM structure to calculate the predictive time-series data. Then the input value from the current time was used as input to the forget gate. The processing results from the forget gate were then calculated using

$$f_t = \sigma(W_{\mathrm{f}} \otimes [h_{t-1}, x_t] + b_{\mathrm{f}}), \tag{7}$$

where W_f represents the weights of the forget gate, b_f shows the bias value applied to the forget gate, x_t is the input value for the current time, and h_{t-1} shows the previous processing time output value. The f_t follows the range $(0, 1)$. Additionally, the output value and state of the candidate cell at the input gate are determined after computation. Both the input value extracted from the most recent time and the output value taken from the last time are fed to the input gate, refer to

$$\tilde{C}_t = \tanh(W_C \otimes [h_{t-1}, x_t] + b_C), \tag{8}$$

where W_C is the weight of the gate input candidate, and b_C is the bias value of the gate input candidate. The next stage in the M-LSTM model is the process of adjusting the cell values or the model parameters shown as

$$C_t = f_t \otimes C_{t-1} + i_t \otimes \tilde{C}_t, \tag{9}$$

where C_t follows the range $(0, 1)$. At processing time t, the output value h_{t-1} and the input value x_t become the input for the output gate, and the output from the gate output is calculated using

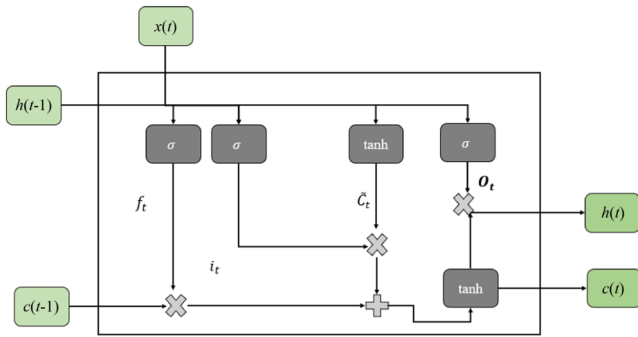$$O_t = \sigma(W_o \otimes [h_{t-1}, x_t] + b_o), \tag{10}$$

FIGURE 1 A multivariate long short-term memory internal structure.

where O_t follows the range $(0, 1)$, W_o is the weight of the gate output, and b_o is the bias value of the gate output. Finally, the final output value of the M-LSTM is generated by the output gate and is the result of the calculation using

$$h_t = O_t \otimes \tanh(C_t), \tag{11}$$

tanh is a hyperbolic tangent function is an activation function that can be tailored to the needs and characteristics of the problem to be resolved. M-LSTM has been used in this work as it can process recorded data in a specific time sequence. Additionally, M-LSTM has been commonly used in the literature for analyzing and modeling time-series data [26, 27].

3.3.2 | Data processing for Experiment 2

Experiment 2 studies the patients who have visited the consultants for over 3 years. The total number of these patients is 1030, who have made 10545 visits to the specialists. On average, each of the 476 patients with a progressive nature has made 10.56 visits to the consultants. The number of patients with a nonprogressive nature was 554, who consulted 5519 times, averaging 9.96 visits per patient. The numbers of patients with more than 3 years of visits are given in Table 3. After applying the SMA, the total number of visits increased from 10,545 to 14,420. The data preprocessing follows the settings described in Experiment 1.

The data set is preprocessed for Experiment 2 with filtration, missing data handling, and normalization like the processes conducted in Experiment 1. Then, the data set was fed to the M-LSTM model using the process detailed in Section 3.3.1. Finally, the M-LSTM model is trained as a classification problem on the preprocessed data set. The model architecture used in this study is given in Table 4.

TABLE 3 Statistics for Experiment 2.

Processing	Type (nature)	Patient	Visits	Average	Standard deviation
Before data preparation	Total	1030	10,545	10.23	3.67
	Progressive	476	5026	10.56	3.98
	Nonprogressive	554	5519	9.96	3.36
After data preparation	Total	1030	14,420	10.23	3.67
	Progressive	476	6664	14	–
	Nonprogressive	554	7756	14	–

Note: – represents not taking standard deviation.

TABLE 4 M-LSTM model architecture.

Layer type	Output shape	Parameters
M-LSTM	(None, 256)	278,528
Fully connected	(None, 128)	32,896
Dropout	(None, 128)	0
Fully connected	(None, 2)	258
Total parameters: 311,682		
Trainable parameters: 311,682		
Nontrainable parameters: 0		

Abbreviation: M-LSTM, multivariate long short-term memory.

4 | RESULTS

4.1 | Experiment 1

After five-cross validation, the AdaBoost model reported an average accuracy, precision, recall, and F1 score of 77.52%, 75.21%, 81.48%, and 78.22%, respectively. After tuning the model's parameters, the accuracy improved to ~80%. Previously, the best accuracy to classify AD as progressive or nonprogressive was 58.47% [23]. The achieved accuracy with AdaBoost with effective preprocessing has improved by 21.34% compared to the literature, reaching 80%. Figure 2 demonstrates the model's performance in Experiment 1 regarding accuracy, precision, recall, and F1 score. AdaBoost models have performed well and better after fine-tuning.

To demonstrate the effectiveness of the proposed model with the baseline work presented in Ho et al. [23], Table 5 provides the accuracy, precision, and recall of both models for AD progression and nonprogression classification, respectively. There is a significant increase in the observed accuracy with AdaBoost. The accuracy and precision of 58.47% and 59.25%, respectively, are increased with the BiPro [23] to 77.52% and 75.21% with AdaBoost. A significant increase in accuracy of 21.34% and precision of 20.57% is achieved with our proposed methodology. The confusion matrix for AD progression and nonprogression is illustrated in Figure 3.

The first-row cell in the first column represents true positives. It indicates that 91 positive instances have been classified correctly. Similarly, the first-row cell in the second column indicates a false positive, meaning 21 negative values were confused with positive labels. The value of the second-row cell in the first column represents a false negative, meaning that 23 of the positive labels were identified as negative. Finally, the values of the second-row cell of the second column
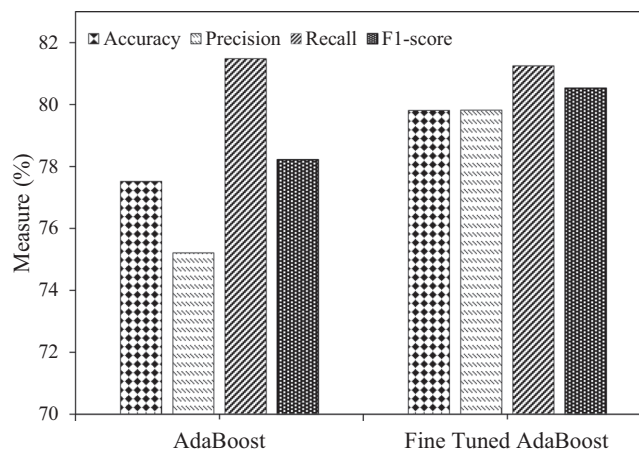
FIGURE 2 Performance measures of Experiment 1 using AdaBoost.

TABLE 5 Performance analysis with baseline [23].

Model	Accuracy (%)	Precision (%)	Recall (%)
BiPro [23]	58.47	59.25	57.78
AdaBoost	77.52	75.21	81.48
Fine Tuned AdaBoost	79.81	79.82	81.25

FIGURE 3 Confusion matrix for Experiment 1. Prog, progression.

show true positives, which means that 83 of the positive samples were correctly classified.

4.2 | Experiment 2

The M-LSTM model for detecting AD progression state is trained and tested on the data set with a split ratio of 80:20, that is, 80% data for training and 20% for testing. Initially, the model achieved an accuracy of ~80%. However, after fine-tuning the parameters and hyperparameters of the M-LSTM model, the model accuracy increased to ~82%. The training and validation accuracy of the proposed model is plotted in Figure 4.
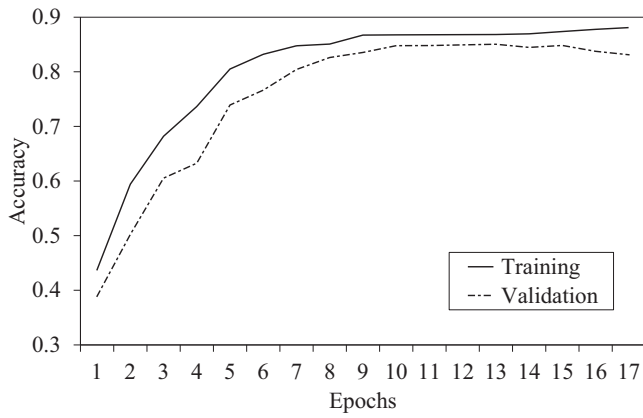
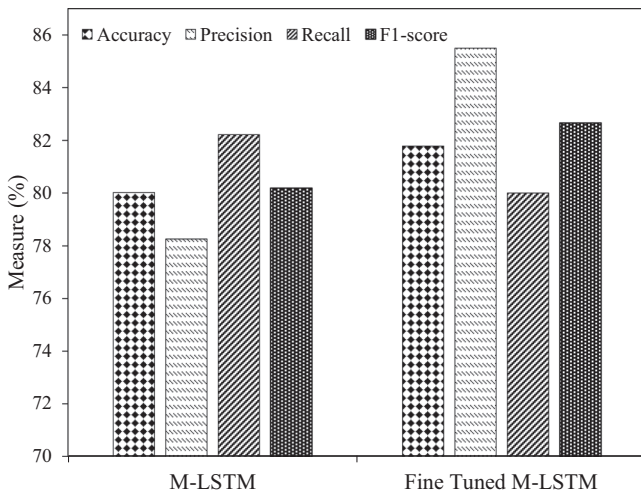**FIGURE 4** Training and validation accuracy curves of multivariate long short-term memory.

FIGURE 5 Performance measures of Experiment 2 using multivariate long short-term memory (M-LSTM).

FIGURE 6 Confusion matrix for Experiment 2. Prog, progression.

The M-LSTM model in this study reported an average accuracy, precision, recall, and $F1$ score of 80.02%, 78.25%, 82.21%, and 80.18%, respectively. After tuning the model's parameters, the accuracy improved to ~82%. Previously in Experiment 1, the best accuracy to classify the AD as progressive or nonprogressive was 79.81% with AdaBoost and 58.47% with BiPro [23]. However, the achieved accuracy with M-LSTM and excellent preprocessing was improved to ~82%. M-LSTM improved accuracy by 23.53% over BiPro and 2.18% with AdaBoost, respectively. In addition, the precision and recall with M-LSTM increased from 78.25% and 82.21% with AdaBoost to 85.5% and 82.66%, respectively. Figure 5 shows the performance of the M-LSTM model for AD progression in Experiment 2 using accuracy, precision, recall, and $F1$ score, respectively.

The confusion matrix for AD progression and non-progression is illustrated in Figure 6. According to the matrix computed, 112 of the positive values were classified correctly. Similarly, the value of the first-row cell and the second column indicates a false positive which means that 19 negative values are confused with positive labels. Likewise, the number of false negatives is 28, meaning the positive labels are identified as negative. Finally, a total of 99 positive samples are correctly classified.

To demonstrate the effectiveness of the M-LSTM model over AdaBoost, Table 6 summarizes the accuracy, precision, recall, and $F1$ score. Better AD progression detection is achieved with M-LSTM compared to AdaBoost with efficient data engineering. Figure 7 demonstrates the achieved performance with both experiments and the baseline to predict AD progression. There is a 2.18% accuracy improvement with M-LSTM, the best classification accuracy for AD progression. The precision and $F1$ scores of 78.25% and 80.53% with the AdaBoost classifier have increased to 85.5% and 82.66%, achieving 7.25% and 2.13% gains, respectively.

5 | DISCUSSION

This study aimed to develop and evaluate two different models, that is, AdaBoost and M-LSTM, for classifying the nature of AD subjects as progressive or nonprogressive.
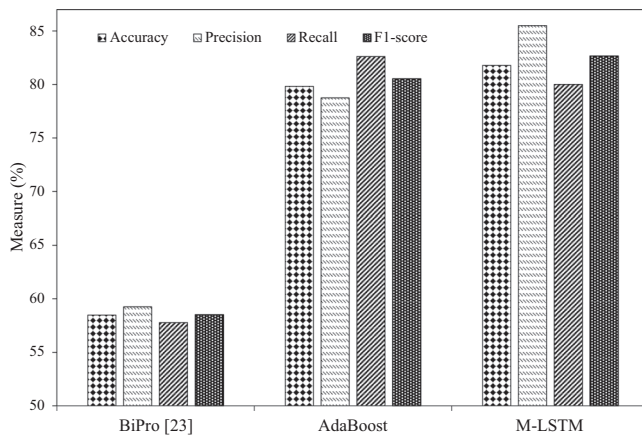
TABLE 6 Performance analysis with AdaBoost.

Model	Experiment	Accuracy	Precision	Recall	$F1$ score (%)
M-LSTM	2	81.78	85.50	80.00	82.66
AdaBoost	1	79.81	79.82	81.25	80.53

Abbreviation: M-LSTM, multivariate long short-term memory.

FIGURE 7 Performance measures comparison between literature and Experiments 1 and 2. M-LSTM, multivariate long short-term memory.

In this study, we conducted two experiments to evaluate the performance of the selected models to classify patients with AD progression. The results of both experiments demonstrate significant improvements in accuracy, precision, recall, and F1 score compared to previous literature and baseline models.

In Experiment 1, the AdaBoost model achieved an average accuracy of 77.52%, precision of 75.21%, recall of 81.48%, and F1 score of 78.22% after five cross-validations. After fine-tuning the model's parameters, the accuracy improved to approximately 80%. This improvement is substantial compared to the best accuracy reported in the literature, which was 58.47% [23]. The achieved accuracy with AdaBoost, coupled with effective preprocessing, represents a remarkable 21.34% improvement over the previous best result, reaching 80%. The performance of the AdaBoost model is further visualized in Figure 2, illustrating its accuracy, precision, recall, and F1 score.

In Experiment 2, the M-LSTM model was trained and tested on the data set. The initial average performance measures achieved by the M-LSTM model were accuracy of 80.02%, precision of 78.25%, recall of 82.21%, and F1 score of 80.18%. After parameter tuning, the accuracy further improved to approximately 82%. Compared to Experiment 1, where the best accuracy was achieved with AdaBoost (79.81%) and the baseline BiPro model (58.47%) [23], the M-LSTM model demonstrated superior performance with an accuracy of approximately 82%. This improvement in accuracy corresponds to a 23.53% gain over the BiPro model and a 2.18% gain over the AdaBoost model. Additionally, the precision and recall of the M-LSTM model increased from 78.25% and 82.21% with AdaBoost to 85.5% and 82.66%, respectively. The performance of the M-LSTM model for AD progression is visualized in Figure 5, showcasing accuracy, precision, recall, and F1 score.

Our findings indicate that the M-LSTM model outperforms AdaBoost in detecting AD progression. The superiority can be attributed to the M-LSTM's ability to effectively process time series data, capturing temporal dependencies crucial in understanding disease progression dynamics. Unlike AdaBoost, which may struggle with the sequential nature of cognitive and neuroimaging indicators, M-LSTM's memory gates enable a more nuanced analysis. The broader implications of these findings in clinical settings are significant. Improved accuracy in classifying AD progression, as demonstrated by M-LSTM, holds promise for enhancing early diagnosis and treatment planning. The ability to capture temporal nuances in patient data opens avenues for more personalized and timely interventions, ultimately improving patient outcomes.

Overall, the results obtained from both experiments highlight the effectiveness of the proposed models, AdaBoost and M-LSTM, in accurately classifying AD progression. The improvements achieved in accuracy, precision, recall, and F1 score compared to the literature and baseline models demonstrate the potential of these models for clinical applications. These findings have important implications for the early diagnosis and treatment of AD, as accurate classification of disease progression can aid in implementing appropriate interventions.

6 | CONCLUSIONS

AD is an irreversible neurological condition that impairs memory, cognition, and behavior. Therefore, early detection is crucial to halt its progression. Timely identification of structural changes from CN to MCI and MCI to AD can aid in detecting progression factors to halt AD progression. Various ML and deep learning models are used for the detection process. This study performed two experiments, that is, one using an ensemble ML model, AdaBoost, and the other using the deep learning model, M-LSTM. First, an efficient preprocessing pipeline for the TADPOLE data set is introduced before digging into the detection part of the progression state. AdaBoost is applied to the prepared data set and results in the detection of subjects involved in AD progression. Different performance measures show better results than a recent state-of-the-art method, BiPro. Furthermore, the second experiment observed the disease's spreading over time using M-LSTM, detected the subject's progressive nature regarding AD, and outperformed the AdaBoost.

AdaBoost is an ensemble ML model that trains the classifier iteratively on the prepared samples, which results in better performance than other classical ML methods. Similarly, M-LSTM is a powerful DL model that uses memory gates to process time series data. However, training

the M-LSTM for a classification problem is more challenging and time-consuming than training AdaBoost. But the results achieved on the prepared data set from M-LSTM are more convincing than AdaBoost and a recent literature method. This research has preprocessed an already existing data set, which is collected for stages detection rather than the detection of AD progressive nature. Furthermore, the preprocessing data pipeline used in this research can be further enhanced by applying custom preprocessing steps rather than standard steps.

Deep learning models are powerful but require excessive computational power and time to process time series data. Most of the datasets have limited information and are unlabeled. Preprocessing the data set and labeling could be useful to avoid the model's overfitting and underfitting problems because the more data available, the better the detection of the AD progression state will be.

AUTHOR CONTRIBUTIONS

Muhammad Irfan: Conceptualization; software, formal analysis; investigation; data curation; writing—original draft preparation; visualization. **Muhammad Irfan** and **Seyed Shahrestani**: Methodology. **Muhammad Irfan**, **Seyed Shahrestani**, and **Mahmoud Elkhodr**: Validation. **Seyed Shahrestani** and **Mahmoud Elkhodr**: Resources; project administration; supervision; writing—review and editing; funding acquisition.

ACKNOWLEDGMENTS

CONFLICT OF INTEREST STATEMENT
The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT
The data set utilized in this study are derived from the TADPOLE challenge, which is available through the Alzheimer's Disease Neuroimaging Initiative, funded by the National Institute of Biomedical Imaging and Bioengineering. The specific data used in this paper can be made available upon reasonable request, subject to the completion of the corresponding PhD student's (Muhammad Irfan) work and in accordance with the conditions and policies governing his PhD program. Requests for access to this data should be directed to the principal supervisor, Seyed Shahrestani.

ETHICS STATEMENT
This study exclusively utilized data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. The ADNI was launched in 2003 as a public–private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging, positron emission tomography, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment and early Alzheimer's disease. For the use of ADNI data, this study adhered to all guidelines and policies set forth by the ADNI and complied with the ethical standards of the relevant institutional and national research committees. Please refer to the acknowledgment statement provided by ADNI. This research involved the use of existing, publicly available data from the ADNI. The ADNI obtained informed consent from all participants involved in the original data collection.

INFORMED CONSENT
Not applicable to this study. This research involved the use of existing, publicly available data from the Alzheimer's Disease Neuroimaging Initiative (ADNI). The ADNI obtained informed consent from all participants involved in the original data collection.

REFERENCES

1. Cuingnet R, Gerardin E, Tessieras J, Auzias G, Lehéricy S, Habert MO, et al. Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database. Neuroimage. 2011;56(2):766–81. https://doi.org/10.1016/j.neuroimage.2010.06.013

2. Nie L, Zhang L, Meng L, Song X, Chang X, Li X. Modeling disease progression via multisource multitask learners: a case study with Alzheimer's disease. IEEE Trans Neural Netw Learn Syst. 2017;28(7):1508–19. https://doi.org/10.1109/TNNLS.2016.2520964

3. Gavidia-Bovadilla G, Kanaan-Izquierdo S, Mataró-Serrat M, Perera-Lluna A, The Alzheimer's Disease Neuroimaging Initiative. Early prediction of Alzheimer's disease using null longitudinal model-based classifiers. PLoS One. 2017;12(1):e0168011. https://doi.org/10.1371/journal.pone.0168011

4. Huang L, Jin Y, Gao Y, Thung KH, Shen D. Longitudinal clinical score prediction in Alzheimer's disease with soft-split sparse regression based random forest. Neurobiol Aging. 2016;46:180–91. https://doi.org/10.1016/j.neurobiolaging.2016.07.005

5. Huang S, Li J, Ye J, Chen K, Wu T, Fleisher A, et al. Identifying Alzheimer's disease-related brain regions from multi-modality neuroimaging data using sparse composite linear discrimination analysis. In: Shawe-Taylor J, Zemel R, Bartlett P, Pereira F, Weinberger KQ, editors. Proceedings of the 24th International Conference on Neural Information Processing Systems, ACM Digital Library; 2011. p. 1431–9. https://proceedings.neurips.cc/paper/2011

6. Alzheimer's Association National Plan Care and Support Milestone Workgroup, Borson S, Boustani MA, Buckwalter KC, Burgio LD, Chodosh J, et al. Report on milestones for care and support under the U.S. National Plan to Address Alzheimer's Disease. Alzheimer's Dementia. 2016;12(3):334–69. https://doi.org/10.1016/j.jalz.2016.01.005

7. Liu W, Zhang B, Zhang Z, Zhou XH. Joint modeling of transitional patterns of Alzheimer's disease. PLoS One. 2013;8(9):e75487. https://doi.org/10.1371/journal.pone.0075487

8. Sukkar R, Katz E, Zhang Y, Raunig D, Wyman BT. Disease progression modeling using hidden Markov models. 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Diego, CA, USA, ACM Digital Library; 2012. p. 2845–8. https://ieeexplore.ieee.org/document/6346556

9. Vu TD, Ho NH, Yang HJ, Kim J, Song HC. Non-white matter tissue extraction and deep convolutional neural network for Alzheimer's disease detection. Soft Computing. 2018;22(20):6825–33. https://doi.org/10.1007/s00500-018-3421-5

10. Zhou J, Liu J, Narayan VA, Ye J. Modeling disease progression via fused sparse group lasso. Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Beijing, China, ACM Digital Library. KDD '12. New York, NY, USA: Association for Computing Machinery; 2012. p. 1095–103. https://doi.org/10.1145/2339530.2339702

11. Zhou J, Liu J, Narayan VA, Ye J. Modeling disease progression via multi-task learning. Neuroimage. 2013;78:233–48. https://doi.org/10.1016/j.neuroimage.2013.03.073

12. Goldstein BA, Navar AM, Pencina MJ, Ioannidis JPA. Opportunities and challenges in developing risk prediction models with electronic health records data: a systematic review. J Am Med Inform Assoc. 2016;24(1):198–208. https://doi.org/10.1093/jamia/ocw042

13. Zhang D, Shen D, The Alzheimer's Disease Neuroimaging Initiative. Multi-modal multi-task learning for joint prediction of multiple regression and classification variables in Alzheimer's disease. Neuroimage. 2012;59(2):895–907. https://doi.org/10.1016/j.neuroimage.2011.09.069

14. Cheng B, Liu M, Zhang D, Munsell BC, Shen D. Domain transfer learning for MCI conversion prediction. IEEE Trans Biomed Eng. 2015;62(7):1805–17. https://doi.org/10.1109/TBME.2015.2404809

15. Zhang D, Shen D, The Alzheimer's Disease Neuroimaging Initiative. Predicting future clinical changes of MCI patients using longitudinal and multimodal biomarkers. PLoS One. 2012;7(3):e33182. https://doi.org/10.1371/journal.pone.0033182

16. Nho K, Shen L, Kim S, Risacher SL, West JD, Foroud T, et al. Automatic prediction of conversion from mild cognitive impairment to probable Alzheimer's disease using structural magnetic resonance imaging. AMIA Annu Symp Proc. 2010;2010:542–6.

17. Wee CY, Yap PT, Shen D, The Alzheimer's Disease Neuroimaging Initiative. Prediction of Alzheimer's disease and mild cognitive impairment using cortical morphological patterns. Hum Brain Mapp. 2013;34(12):3411–25. https://doi.org/10.1002/hbm.22156

18. Wolz R, Julkunen V, Koikkalainen J, Niskanen E, Zhang DP, Rueckert D, et al. Multi-method analysis of MRI images in early diagnostics of Alzheimer's disease. PLoS One. 2011;6(10):e25446. https://doi.org/10.1371/journal.pone.0025446

19. Cho Y, Seong JK, Jeong Y, Shin SY, The Alzheimer's Disease Neuroimaging Initiative. Individual subject classification for Alzheimer's disease based on incremental learning using a spatial frequency representation of cortical thickness data. Neuroimage. 2012;59(3):2217–30. https://doi.org/10.1016/j.neuroimage.2011.09.085

20. Ewers M, Walsh C, Trojanowski JQ, Shaw LM, Petersen RC, Jack CR, et al. Prediction of conversion from mild cognitive impairment to Alzheimer's disease dementia based upon biomarkers and neuropsychological test performance. Neurobiol Aging. 2012;33(7):1203–14. https://doi.org/10.1016/j.neurobiolaging.2010.10.019

21. Kim D, Kim S, Risacher SL, Shen L, Ritchie MD, Weiner MW, et al. A graph-based integration of multimodal brain imaging data for the detection of early mild cognitive impairment (E-MCI). Multimodal Brain Image Anal. 2013;8159:159–69. https://doi.org/10.1007/978-3-319-02126-3_16

22. Veitch DP, Weiner MW, Aisen PS, Beckett LA, Cairns NJ, Green RC, et al. Understanding disease progression and improving Alzheimer's disease clinical trials: recent highlights from the Alzheimer's Disease Neuroimaging Initiative. Alzheimer's Dementia. 2019;15(1):106–52. https://doi.org/10.1016/j.jalz.2018.08.005

23. Ho NH, Yang HJ, Kim J, Dao DP, Park HR, Pant S. Predicting progression of Alzheimer's disease using forward-to-backward bi-directional network with integrative imputation. Neural Net. 2022;150:422–39. https://doi.org/10.1016/j.neunet.2022.03.016

24. Marinescu RV, Oxtoby NP, Young AL, Bron EE, Toga AW, Weiner MW, et al. TADPOLE challenge: accurate Alzheimer's disease prediction through crowdsourced forecasting of future data. Predict Intell Med. 2019;11843:1–10. https://doi.org/10.1007/978-3-030-32281-6_1

25. Schafer JL, Graham JW. Missing data: our view of the state of the art. Psychol Methods. 2002;7(2):147–77. https://doi.org/10.1037/1082-989X.7.2.147

26. Choi E, Bahadori MT, Schuetz A, Stewart WF, Sun J. Doctor AI: predicting clinical events via recurrent neural networks. JMLR Workshop Conf Proc. 2016;56:301–18.

27. Choi E, Schuetz A, Stewart WF, Sun J. Using recurrent neural network models for early detection of heart failure onset. J Am Med Inform Assoc. 2017;24(2):361–70. https://doi.org/10.1093/jamia/ocw112

28. Razavian N, Marcus J, Sontag D. Multi-task prediction of disease onsets from longitudinal lab tests. In: Doshi-Velez F, Fackler J, Kale, D, Wallace, B, Wiens, J, editors. Proceedings of the 1st Machine Learning for Healthcare, Northeastern University, Boston, MA, USA, 18–19 August. Vol. 56. Proceedings of Machine Learning Research. PMLR; 2016. 73–100. https://proceedings.mlr.press/v56/Razavian16.html

29. Lipton ZC, Kale DC, Elkan C, Wetzel R. Learning to diagnose with LSTM recurrent neural networks. arXiv e-prints. 2015. https://doi.org/10.48550/arXiv.1511.03677

30. Lee G, Nho K, Kang B, Sohn KA, Kim D, Weiner MW, et al. Predicting Alzheimer's disease progression using multi-modal deep learning approach. Sci Rep. 2019;9(1):1952. https://doi.org/10.1038/s41598-018-37769-z

31. Bilgel M, Jedynak BM. Predicting time to dementia using a quantitative template of disease progression. Alzheimers Dement. 2019;11:205–15. https://doi.org/10.1016/j.dadm.2019.01.005

32. Venkatraghavan V, Bron EE, Niessen WJ, Klein S, Alzheimer's Disease Neuroimaging Initiative. Disease progression timeline estimation for Alzheimer's disease using discriminative event based modeling. Neuroimage. 2019;186:518–32. https://doi.org/10.1016/j.neuroimage.2018.11.024

33. Lorenzi M, Filippone M, Frisoni GB, Alexander DC, Ourselin S. Probabilistic disease progression modeling to characterize diagnostic uncertainty: application to staging and prediction in Alzheimer's disease. Neuroimage. 2019;190:56–68. https://doi.org/10.1016/j.neuroimage.2017.08.059

34. Fedorov A, Hjelm RD, Abrol A, Fu Z, Du Y, Plis S, et al. Prediction of progression to Alzheimer's disease with deep infomax. 2019 IEEE EMBS International conference on biomedical & health informatics (BHI). IEEE; 2019. 1–5. https://ieeexplore.ieee.org/document/8834630

35. Wang T, Qiu RG, Yu M. Predictive modeling of the progression of Alzheimer's disease with recurrent neural networks. Sci Rep. 2018;8(1):9161. https://doi.org/10.1038/s41598-018-27337-w

36. Jung W, Jun E, Suk HI. Deep recurrent disease progression model for conversion-time prediction of Alzheimer's disease. NeuroImage. 2021;237:118143. https://doi.org/10.1016/j.neuroimage.2021.118143

37. Yoon J, Zame WR, van der Schaar M. Estimating missing data in temporal data streams using multi-directional recurrent neural networks. IEEE Trans Biomed Eng. 2019;66(5):1477–90. https://doi.org/10.1109/TBME.2018.2874712

38. Nguyen M, He T, An L, Alexander DC, Feng J, Yeo BTT, et al. Predicting Alzheimer's disease progression using deep recurrent neural networks. Neuroimage. 2020;222:117203. https://doi.org/10.1016/j.neuroimage.2020.117203

39. Jedynak BM, Lang A, Liu B, Katz E, Zhang Y, Wyman BT, et al. A computational neurodegenerative disease progression score: method and results with the Alzheimer's Disease Neuroimaging Initiative Cohort. Neuroimage. 2012;63(3):1478–86. https://doi.org/10.1016/j.neuroimage.2012.07.059