

Received December 2, 2020, accepted December 5, 2020, date of publication December 9, 2020, date of current version December 23, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3043715

An Ensemble-Learning Based Application to Predict the Earlier Stages of Alzheimer's Disease (AD)

ASIF HASSAN SYED¹, TABREJ KHAN², ATIF HASSAN³, NASHWAN A. ALROMEMA¹, MUHAMMAD BINSAWAD¹, AND ALHUSEEN OMAR ALSAYED⁴

¹Department of Computer Science, Faculty of Computing and Information Technology in Rabigh (FCITR), King Abdulaziz University, Jeddah 21589, Saudi Arabia

²Department of Information Systems, Faculty of Computing and Information Technology in Rabigh (FCITR), King Abdulaziz University, Jeddah 21589, Saudi Arabia

³Department of Computer Science and Engineering, Indian Institute of Technology Kharagpur, Kharagpur 721302, India

⁴Scientific Researcher, Deanship of Scientific Research, King Abdulaziz University, Jeddah 21589, Saudi Arabia

Corresponding author: Asif Hassan Syed (shassan1@kau.edu.sa)

This project was funded by the Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah, under grant No. (DF-676-305-1441). The authors, therefore, gratefully acknowledge DSR technical and financial support.

ABSTRACT The fact that ensemble methods enhance the prediction performance. Therefore, we focused on developing a weighted ensemble method using a novel combination of Cerebrospinal Fluid (CSF) protein biomarkers to predict AD's earlier stages with greater accuracy than the state-of-the-art CSF protein biomarkers. In this regard, two feature selection methods, namely the Recursive Feature Elimination (RFE) and L1 regularization method were used to screen the most important subset of features for building a classification model using the Mild Cognitive Impairment (MCI) dataset. A novel combination of three biomarkers, namely Cystatin C, Matrix metalloproteinases (MMP10), and tau protein, was screened using the linear Support Vector Machine (SVM) and Logistic Regression (LR) classifier based RFE method. Two-tailed unpaired t-test analysis at a 5% significance level showed a significant difference between the mean levels of Cystatin C, MMP10, and tau protein between cognitive normal and cognitively impaired groups. An ensemble model using a weighted average of two best performing classifiers (LR and Linear SVM) was created using a novel subset of three most informative features. Our ensemble model's weighted average results performed significantly better than LR and Linear SVM base classifiers' performance. The Receiver Operating Characteristic Curve (ROC_AUC) and Area under Precision-Recall values (AUPR) of our proposed model were observed to be 0.9799 ± 0.055 0.9108 ± 0.015 , respectively. The performance of our proposed weighted averaged ensemble model built using a novel combination of CSF protein biomarkers was significantly better ($p < 0.001$) than models generated using different combinations of CSF protein biomarkers obtained from recent studies. An ensemble-learning based application was implemented and deployed at Heroku at <https://appsalzheimer.herokuapp.com>.

INDEX TERMS Mild cognitive impairment (MCI), cerebrospinal fluid (CSF) protein biomarkers, classification model, online prediction system.

I. INTRODUCTION

Alzheimer's disease results in a neurodegenerative disorder that causes irreversible and progressive brain cell damage, usually affecting people during their mid-60s [1], [2]. Pre-clinical changes in the brain associated with Alzheimer's begin years before the onset of the disease's typical clinical

The associate editor coordinating the review of this manuscript and approving it for publication was Yiming Tang¹.

symptoms. Though the onset of AD cannot be reversed or stopped, early detection of the disease can allow treatment and spontaneous care of Alzheimer's patients in their earlier stages before irreparable damage to the brain has occurred [3], [4]. Therefore, studies' on the development of new strategies for the earlier diagnosis of AD are among the most active research areas in Alzheimer's science. Researches in the past have shown that individuals with MCI or presymptomatic Alzheimer's have a greater risk of eventually

converting to AD [5]–[10]. The biochemical changes in CSF associated with AD's progression provide a sound and potential source of diagnostic biomarkers to study the disease's preclinical and clinical stages. Thus, for the early detection of AD, identifying specific biomarkers that play a significant role in converting MCI to AD has become of great interest in recent times [11]–[14]. In the current scenario, conventional CSF biomarkers, namely tau, amyloid- β 42 ($A\beta$ 42), and phosphorylated forms of tau (p-tau), have shown more significant potential in the screening of MCI patients who eventually progressed to clinically diagnosable AD [15]–[19].

AD is multifactorial by nature; therefore, a substantial overlap of the clinical biomarkers between the cognitively healthy and cognitively impaired individuals is observed [20]. Thus, considering the present scenario, there is an urgent need to identify novel protein biomarkers from CSF. Identifying new CSF protein biomarkers will enhance our current understanding of the pathophysiology of MCI and the progression of MCI patients to AD. However, herein lies a problem as the analysis and interpretation of multiple factors (clinical biomarkers) involved in converting MCI to AD are difficult and complicated. Therefore, in recent years the applications of pattern recognition and Machine Learning (ML) algorithms in Computer-Aided Diagnostic (CAD) tools have become invaluable. These tools are highly efficient in analyzing and interpreting AD's multifactorial nature, enabling clinicians and scientists to distinguish between healthy controls and patients with MCI who convert to AD. Several machine learning approaches have been used on a different combination of features, namely CSF biomarkers, genotypes, brain imaging, clinical information, and demographics, to predict MCI subjects' conversion to AD with varying levels of precision and accuracy [11], [6], [21]–[27].

Nevertheless, only a few studies have looked for novel CSF protein biomarkers that can augment the accuracy of current, leading CSF protein biomarkers ($A\beta$ 42, Tau, p-tau181) in distinguishing MCI patients from healthy subjects and, at the same time, detecting the earlier stages of AD [11], [13]. Lately, ensembling has gained importance in classifying a progressive form of MCI, which ultimately leads to AD from cognitively normal individuals [23], [28], [29]. Therefore, we propose an ensemble learning model to identify additional CSF protein biomarkers, fitted on attributes (CSF protein biomarker) selected using the RFE method.

The Alzheimer dataset generated by Craig Schapiro *et al.* 2011 [11], comprising of both demographic and CSF protein biomarkers, was employed to build our model for predicting the earlier stages of AD. Our proposed ensemble model can classify MCI patients from cognitively normal subjects with better sensitivity, specificity, and accuracy than state-of-the-art CSF protein biomarkers' (tau, $A\beta$ 42, p-tau) based classification models [11]. Our CSF protein biomarkers' based web application is the first of its kind to predict the earlier stages of AD. The application is written in python, developed in Flask, a micro web framework, and deployed on Heroku at

“<https://appsalzheimer.herokuapp.com>.” The paper onwards is divided into the following sections: 1) The Material and Methods section provides information about Alzheimer's clinical datasets' source and contents. It also provides a detailed description of the data preprocessing methods, supervised learning algorithms, and model evaluation metrics employed for assessing the performance of various models built using Alzheimer's clinical data. 2) The Result section briefly explains the results of various methods employed to generate an efficient predictive model for predicting the earlier stages of AD 3) The Discussion section describes our model efficiency in differentiating cognitively impaired subjects from cognitively normal subjects. 4) Finally, the Conclusion section provides the concluding remarks and scope of our model and its implementation in predicting the earlier stages of AD. A pictorial representation of the methodology adopted in our study to develop an ensemble-learning-based web application for predicting the earlier stages of AD is depicted in Fig. 1.

II. MATERIAL AND METHODS

A. DATA SOURCE AND DESCRIPTION

The MCI clinical dataset was downloaded from Figshare [30]. The cognitive status of the 333 subjects involved in the clinical study was, according to the Clinical Dementia Rating Scale (CDR). The clinical dataset consisted of 91 mildly cognitive impaired subjects (CDR 0.5 and CDR 1) and 242 cognitively normal subjects (CDR 0). The MCI clinical dataset consisted of features including 124 credible CSF protein biomarkers, demographic features, namely age and gender, and a set of non-imaging protein biomarkers, namely, amyloid proteins, native tau, phosphorylated form of Tau (pTau), $A\beta$ 42, β 42, and also allelic variants of Apolipoprotein E genotype (E2, E3, and E4). The Apolipoprotein E variant E4 is the most significant among all other Apolipoprotein E genotype variants because of its association with AD [31], [32].

B. PREPROCESSING OF THE MCI CLINICAL DATASET

In a Machine Learning process, preprocessing of the data is that step in which the data gets cleaned, transformed, or *Encoded*, and reduced to process the data to such a state that the machine algorithms can easily learn and analyze the preprocessed data to build a better predictive model.

The current section discusses the type of data processing steps involved in the present study, as illustrated in Fig. 1. Further, a one-hot encoded technique was used to transform the categorical features into a binary representation. The Z-score normalization was used for performing the standard scaling of the numerical features present in the MCI dataset. The correlation between the numerical-numerical and numerical-categorical features was checked using the Pearson Correlation Coefficient (PCC) to screen out the correlated features from the MCI dataset. In the current study, two hundred copies of the MCI dataset were generated

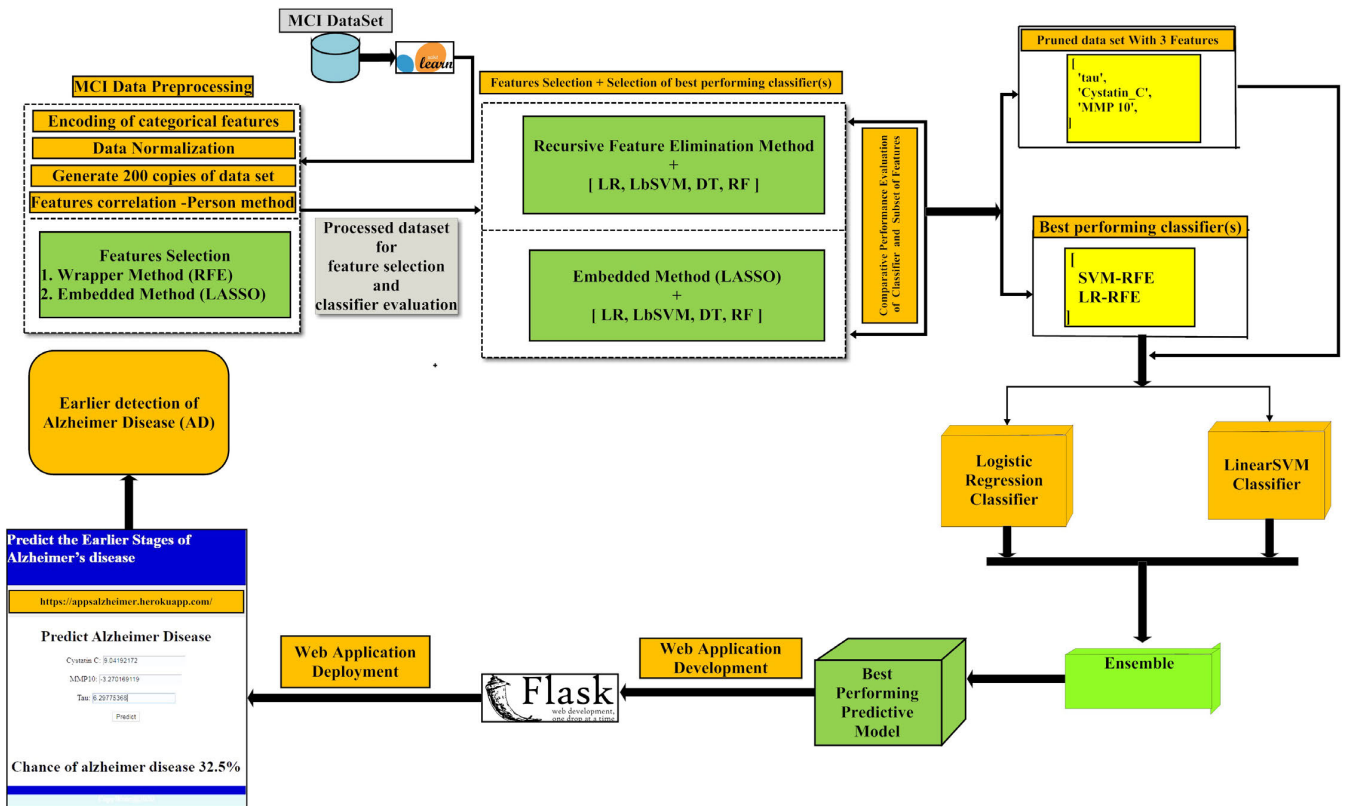


FIGURE 1. A pictorial representation of the methodology implemented in this study for data processing, feature selection, model building & validation, and implementation.

using stratified splitting with the replacement for building a statistically meaningful model. Finally, each copy of the randomly generated two-hundred copies of the dataset was segmented into 80% training-cum-validation data and 20% testing data. Feature selection was performed using the different methods, namely the wrapper method and the embedded method. The feature selection methods mentioned above select the best subset of features based on feature importance/information, thereby selecting the best set of predictors and ignoring the redundant and less important features for building a better predictive model. In the present study, we have focused on the RFE, an important example of a wrapper based feature selection process and L1 regularization with an L1 penalty to screen the most important subset of features, an example of an embedded based feature selection method. A detailed description of the different preprocessing techniques employed in the present study mentioned above is discussed below in the following subsections.

1) ENCODING OF CATEGORICAL FEATURES

The non-continuous categorical features of the MCI dataset were one-hot encoded, which essentially transforms each categorical feature with “n” categories into “n-1” binary categories, a format suitable for downstream estimators.

2) DATA NORMALIZATION

The Z-score normalization was used to normalize the numerical attribute data points. Normalization assists in scaling the

data within a range to avoid training and building incorrect Machine learning-based predictive models. In Z-score normalization, the dataset’s scaling was performed using μ and σ computed on the MCI dataset.

$$Value' = \frac{value - X'}{s} \tag{1}$$

Refer to “equation 1” shown above for calculating the z-score normalization of the numerical variable in the MCI dataset.

3) FEATURES CORRELATION

Features having a Pearson coefficient of correlation >0.99 were checked using Person Correlation. Pearson correlation measures the linear relationship between the two predictors X and Y. A value close to one indicates the highest correlation between the two predictors, while a correlation coefficient of zero indicates no linear correlation between the two predictors. As no predictors were highly correlated (>0.99), therefore none of the predictors were deleted.

4) STRATIFIED FEATURE SAMPLING

As the number of instances in the negative class, i.e., the cognitively normal subjects (CDR 0), is significantly higher than the positive class (CDR > 0), it leads to an imbalance in the dataset. Thus to have an adequate representation of the rarer class during the training cum validation process, a stratified

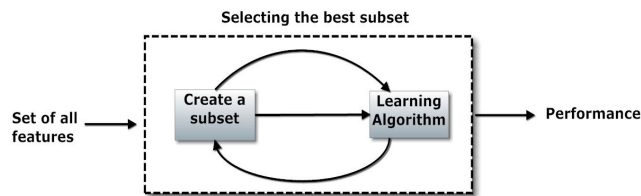


FIGURE 2. Wrapper method of variables selection.

sampling method with replacement was used for feature sampling.

A random 200 copies of the dataset were generated to average out results and get a more realistic estimate for the small MCI dataset. Two hundred copies of the dataset were created, and within each copy, a random train-test split of 80-20% was performed as the MCI dataset has a limited number of samples for building a statistically meaningful model.

5) FEATURE SELECTION

Feature reduction was applied to the training dataset in order to select an optimal subset of the attributes. Feature reduction was performed because the MCI clinical dataset's feature space is quite large, and not all attributes contribute effectively in predicting the dependent variable (class). Removing features having minimal importance reduces model variance while also improving training and inference time.

In this regard, two different types of Feature Selection Methods have been applied in this study, namely, the wrapper-based method and embedded method. Wrapper methods employ a greedy approach to iteratively look through the space of possible feature subsets, evaluating each subset of attributes based on a given machine learning algorithm performance. The wrapper-based feature selection method evaluates the possible interaction between the attributes to look for the best possible combinations of attributes that result in the best performing model, as shown in Fig. 2. Besides, testing all possible combinations of features can be computationally expensive. However, if the dataset is small, particularly when the number of attributes is small, each run would be computationally cheap.

In this study, we have used the RFE method, a popular example of the wrapper method of feature selection. The RFE method [33] was used to automatically prune the least essential attributes from the given set of features in the dataset. As the name suggests, RFE recursively eliminates the feature with the smallest coefficient on every iteration until the desired number of attributes is eventually reached. The RFE based feature selection method removes large correlations among features while also providing the best combination of the desired number of informative attributes for the effective prediction of the dependent variable.

Since there is a class imbalance in the dataset, and the minority class $CDR > 0$ is of importance to us, even at the cost of some misclassification for $CDR 0$ (majority class),

RFE utilizing cost-sensitive estimators with weights, 0.2 for class "0" and 0.8 for class "1" was used. Practically with the Scikit-learn library, we can use any classifier (estimator) to do the feature subset search with RFE. Therefore, the RFE method, combined with LR, DT, RF, and Linear SVM individually, was used to search for the best informative features subset to build a more robust and effective classifying model.

Besides, using a combination of each ML algorithms and the RFE method, we will know whether the feature subset looks the same or different, or their performance in predicting the target variable differs or remains the same. A series of feature sets containing features ranging from 1 to $(1 + n)$ were selected. Here, "n" varies from 0 to 4. Likewise, the number of attributes removed at each iteration in RFE was set to be three features. After each iteration, we removed the three least important features instead of 1 to reduce our model's runtime complexity. Thus after recursive elimination of the three least important features after each loop, we reached our desired set of informative features (i.e., subsets of feature(s) ranging from 1 to 5).

The classifier fit the entire training data at every iteration, and coefficients for each feature were found. The attributes corresponding to the smallest coefficients were removed, and this process continued until only the desired set number of features remained. The performance metrics, namely accuracy, precision, recall, and ROC-AUC, were used to evaluate each machine learning algorithm's performance using several sets of most informative features. A pictorial representation of the classifier-RFE based methodology to screen the best subset of features for predicting the earlier stages of AD is depicted in Fig. 3.

On the other hand, the embedded methods merge the wrapper and filter methods' vital points by benefitting from the built-in feature selection process of certain machine learning algorithms.

Additionally, in the embedded methods, the feature selection and training processes are performed concurrently, as shown in Fig. 4. In the current study, we focus on the Least Absolute Shrinkage and Selection Operator (LASSO) or L1 regularization, an important example of an embedded technique to select the best subset of features significantly associated with the response variable. The LASSO approach shrinks the explanatory variables' coefficients with less or no discriminatory power to zero while selecting a subset of explanatory variables with non-zero coefficients [34], [35]. The selected explanatory variables represent the joint discriminatory power to separate MCI patients from cognitively normal subjects. Before feeding the MCI dataset to the LASSO algorithm, all the categorical variables, namely the explanatory and the response variable, of the MCI dataset, were one-hot encoded.

Moreover, the LASSO is a clear case of the penalized least squares regression with lambda (λ) as an L1-penalty function. The tuning of the hyperparameter (penalty factor lambda) was performed during the cross-validation process. We applied the LASSO regression with 5-fold

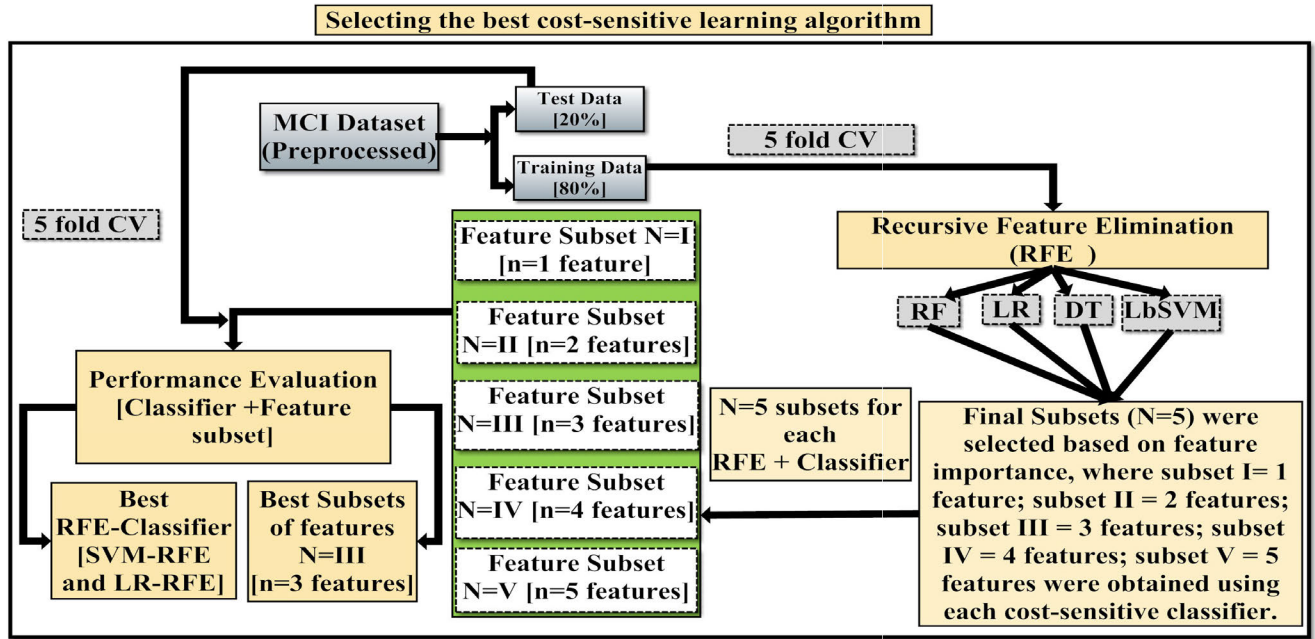


FIGURE 3. Pictorial representation of the classifier-RFE based Feature selection method.

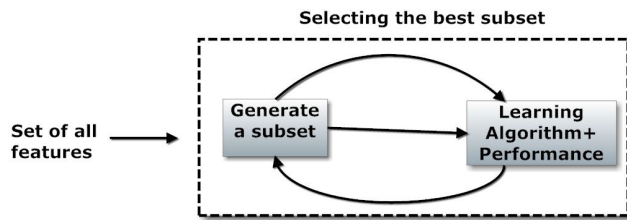


FIGURE 4. Embedded method of variables selection.

cross-validation (CV) 200 times to select the optimal subset of discriminatory attributes. A pictorial representation of the LASSO or L1 regularization-based methodology to screen the best subset of features is depicted in Fig. 5.

The LASSO algorithm can be mathematically defined. Refer to equation “2”:

$$\hat{\beta}^{LASSO} = \arg_{\beta} \min \left\{ \frac{1}{2} \sum_{i=1}^N \left(y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j \right)^2 + \lambda \sum_{j=1}^p |\beta_j| \right\} \quad (2)$$

The above equation can also be rewritten. Refer to “3 and 4,” shown below:

$$\hat{\beta}^{LASSO} = \arg_{\beta} \min \sum_{i=1}^N \left(y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j \right)^2 \quad (3)$$

Subject to the following condition:

$$\sum_{j=1}^p |\beta_j| \leq t, \quad (4)$$

The LASSO method’s advantage is that it is easily interpretable since it shrinks the coefficients of the non-informative and correlated attributes exactly to zero. Moreover, most of the time, the LASSO method is preferred for feature selection when the dataset has a small number of observations and many features. Once the best subset of features is obtained, the pruned MCI dataset with the best subset of features was used to build the four classifiers-based models to classify MCI patients from normal subjects. The classifiers were evaluated based on the various performance measures to screen the best performing model.

C. CLASSIFICATION ALGORITHMS

Four machine learning algorithms, Linear SVM, LR, DT, and RF, were used as base classifiers for the ensembling process.

1) LINEAR SUPPORT VECTOR MACHINE (LSVM)

The non-continuous categorical features of the MCI dataset were one-hot encoded, which essentially transforms each categorical feature with “n” categories into “n-1” binary categories, a format suitable for downstream estimators.

The algorithm can output the best possible separating hyperplane for the provided set of classes given labeled training data. In a dataset, a sample is represented by a p-dimensional vector, and the linear SVM algorithm tends to find a (p-1) dimensional hyperplane that separates the data point into a set of classes. Linear SVM is a fast, discriminative classifier that works on the maximum-margin hyperplane [36]. The maximum margin hyperplane represents the largest distance between the nearest data point of each category or class, and SVM tends to find this maximum-margin hyperplane to classify a sample data.

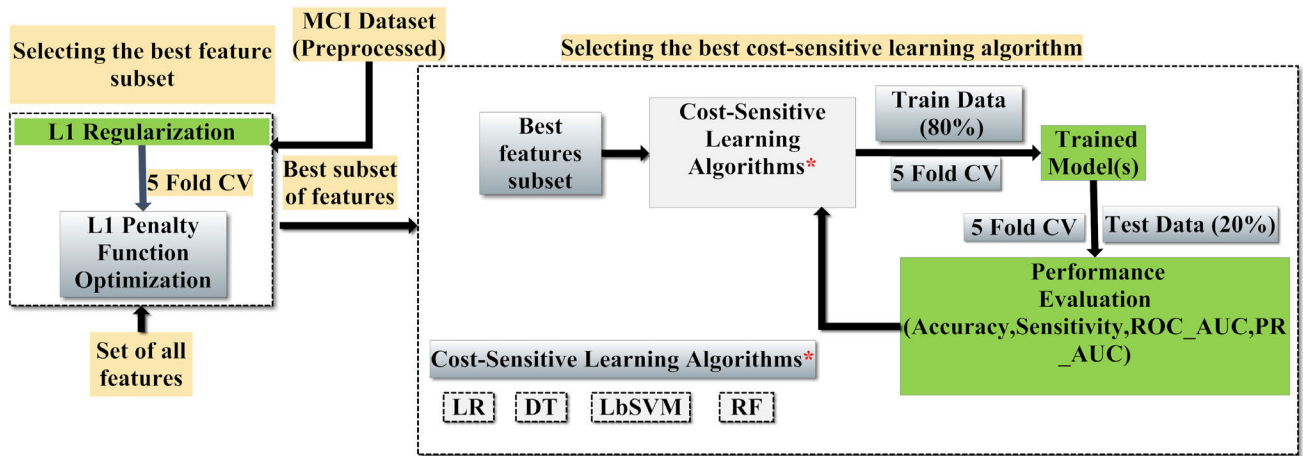


FIGURE 5. Pictorial representation of the L1 regularization based Feature selection method.

Mathematically the maximum-margin hyperplane can be obtained as follows:

Suppose we are given a dataset with a set of data points x_i where $i = 1, 2, \dots, n$. Here “n” is the number of data points. The dataset was categorized into two classes: a positive class denoted by $y_i = 1$ and a negative class represented by $y_i = -1$. We can find

A hyperplane $f(x) = 0$, categorizing the data points in a dataset into the two classes. Refer to equation “5.”

$$f(x) = w^T x + b = \sum_{j=1}^M w_j x_j + b = 0 \quad (5)$$

Here w is an n -dimensional vector, and b is a scalar constant is used to define the hyperplane

In the linear SVM algorithm, the data points associated with the two-classes can be separated linearly. We can do this by two hyperplanes that separate the data points with no data points between the planes. The two-hyperplanes are called separating hyperplanes, and the distance between the nearest data point of each class bounded by the two hyperplanes is called a margin. The linear SVM tends to maximize the margin to improve the classifier’s ability to categorize data points into specific classes. “Equations 6-8 describe the two separating hyperplanes as follows”:

$$f(x_i) = 1 \quad \text{if } y_i = 1 \quad (6)$$

$$f(x_i) = -1 \quad \text{if } y_i = -1 \quad (7)$$

$$y_i f(x_i) = y_i(w^T x_i + b) \geq 1 \quad \text{for } i = 1, 2, \dots, M. \quad (8)$$

The data points used to define the two-hyperplanes are termed as support vectors. According to equation 7, the margin obtained in a multidimensional framework is given by $\|w\|^{-2}$

$$\text{minimize } \frac{1}{2} \|w\|^{-2} \quad (9)$$

$$\text{subject to } y_i(w^T x_i + b) \geq 1 \quad \text{for } i = 1, 2, \dots, M \quad (10)$$

Therefore, the ideal hyperplane separating the data point into two classes can be achieved by solving the following

optimization problem. Refer to the equation “9 and 10”, respectively.

2) LOGISTIC REGRESSION (LR)

The LR is a supervised classification algorithm. The LR algorithm is based on the linear regression model. Refer to equation “11”:

$$P = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_m X_m \quad (11)$$

The LR algorithm fits the training data to a logistic sigmoid function, as shown in equation 9, and predicts the target categorical dependent variable’s probability [37]. The estimated probability of the target variable in LR varies from 0 to 1. Also, a threshold is set to classify a particular instance into a specific target class. Depending on the threshold, the obtained estimated probability is classified into a specific target class. The estimated predictive value for a given x_i value can be interpreted as sample x_i ’s chances to be a member of a target class variable. Let us say, if the predicted value of a sample x_i is >0.5 , then classify the sample under the “CDR > 0” category else under the “CDR = 0” category. “Equation 12, 13, and 14” are the main equations of the LR algorithm are shown below”:

$$\Pr(Y = +1|X) \sim \beta \cdot X \quad \text{and} \quad \Pr(Y = -1|X) = \Pr(Y = +1|X) \quad (12)$$

$$\downarrow \sigma(x) := \frac{1}{1 + e^{-x}} \in [0, 1] \quad (\text{the sigmoid function}) \quad (13)$$

$$\Pr(Y = +1|X) \sim \sigma(\beta \cdot X) \quad \text{and} \quad \Pr(Y = -1|X) = 1 - \Pr(Y = +1|X) \quad (14)$$

This study has two categorical dependent variables, namely CDR > 0 and CDR = 0 groups. Here Y signifies the dependent target variable CDR > 0 group. While X in equation 11 represents the independent explanatory variable in the dataset. Every independent variable, X , is assigned a coefficient value β representing weight. Different weights represent the different correlations between variables X and Y .

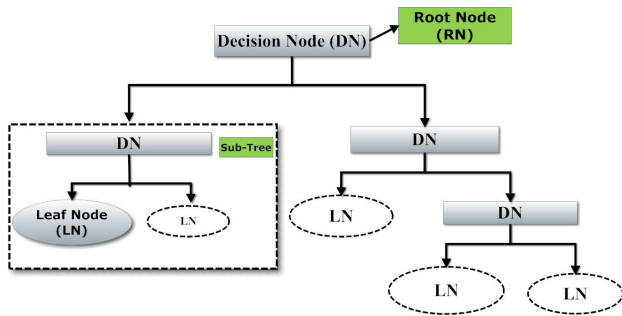


FIGURE 6. The basic structure of a DT algorithm-based model.

3) DECISION TREE (DT)

The DT is a supervised learning algorithm that can be applied to solve both classification and regression problems. Implementing a DT algorithm aims to build a model to predict the target variable’s class label or value. A DT algorithm can be used for both categorical and continuous target variables. The DT classifies the instances in the dataset by segregating them down the tree from the root node to some leaf node, with the leaf node providing a decision point for the labeling test instances with the unknown class variable. The root node of a DT is the best performing predictor node from a set of nodes (attribute) present in the dataset. The representation of an atypical DT structure is shown in Fig. 6. There are various ways to select the root node based on the percentage of heterogeneity obtained by splitting the data using different attributes.

The performance measures, namely Gini indexes, Entropy, classification error, are calculated for each predictor, and a comparison is made to select the best predictor root node to start the splitting of the dataset. The DT algorithm recursively generates a split (decision node) on each subset of data, considering attributes that never have been selected before until it reaches a terminal node (leaf node) corresponding to a subset of data with maximum purity [38].

4) RANDOM FOREST (RF)

The RF is an ensemble learning algorithm that combines relatively uncorrelated tree predictors. Each tree in the RF provides a class prediction, and finally, the most voted class is the final prediction result (class) of the ensemble. A pictorial representation of the working of the algorithm is depicted in Fig. 7.

RF overcomes variance and overfitting by averaging the result of each independent tree predictors [39].

D. ENSEMBLE MODEL GENERATION

Many methods generate an ensemble of models, including stacking, bagging, boosting, voting, and averaging. As the dataset is small and an ensemble of strong learners is required, methods that minimize variances such as voting and averaging need to be selected. Since averaging outperforms voting [40], the trained LR and SVM classifiers were selected

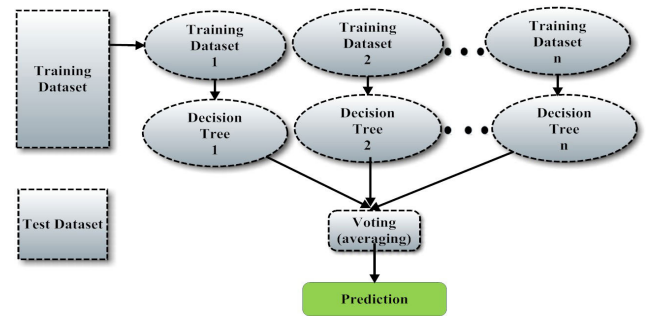


FIGURE 7. Pictorial representation of the working mechanism of the RF algorithm.

for average ensemble while the DT and RF classifiers were left out as the latter two were outperformed by the former.

The idea behind an average ensemble is that among two or more models, some might perform better on certain samples while others might perform better on the remaining ones, and by taking their average, a single classifier is built that can generalize to all samples. The above statement can be empirically understood by checking the PCC of the models’ output. Lower correlation signifies that the models perform well on different subsets of the data, while higher correlations will lead to an ensemble, which is only as good as their base predictor. Giving all models equal weightage in an average ensemble reduces variance and increases bias as poor performing models can have a regularization-like effect on the output. Reducing variance while not affecting a system’s bias, the weighted average ensemble can give more importance/weight to better performing models in the ensemble.

1) WEIGHTED AVERAGE ENSEMBLE LEARNING

The key idea behind ensemble learning is to combine weak learning classifiers to generate a robust classifier. The final ensemble-based model provides better stability by reducing the individual weaker classifiers’ error (bias, noise, and variance) [40]. In this study, an ensemble-learning model was created by training-cum-validating two hundred copies of the pruned dataset (i.e., with three features), and within each copy, a random train-test split of 80-20 is performed. Two-hundred copies of the MCI dataset were generated to average out results and get a more realistic estimate for the small dataset. All training and testing datasets were reduced to three features. Five-fold cross-validation over all of the 200 training datasets was performed, and the final results averaged to find the best classifier. An ensemble of the two best-performing algorithms was created by taking a weighted average of LR and Linear SVM. Where SVM’s weight was set to 0.9 and that of linear regression was set to 0.1.

$$f(x) = \frac{1}{1 + e^{-x}} \quad (15)$$

Since the output of linear SVM is not probability-based, therefore, its soft output was converted to probability via the logistic function. Refer to “15” stated above.

E. TRAINING-CUM-VALIDATION DATASET

The MCI dataset with 333 instances were segmented into 80% training data and 20% independent test data. The 80% data of the MCI dataset was used for performing 5-fold training-cum-cross-validation. The four supervised machine-learning algorithms (LibSVM, LR, Decision tree (DT), and Random Forest (RF)) cost-sensitive estimators (classifiers) with weights, 0.2 for class “0” and 0.8 for class “1” were used to neutralize the bias created by the majority class in the MCI dataset such that the false positive rate would not exceed a threshold of 20%. The cost-sensitive classifiers were trained-cum-validated on various feature subsets obtained from the RFE and L1 regularization feature selection method.

The total number of instances in the training data was 266, where 196 were cognitively normal samples (Negative class), while 70 instances were MCI patients (Positive class). Since binary classification is being performed, we need to set an optimum threshold for our base and ensemble classifiers. A simple binary search was employed with the initial lower, intermediate and upper values set to 0, 0.5, and 1.0, respectively. The objective was to find the threshold for which specificity was per Craig-Schapiro *et al.*, 2012 [11]. Therefore, for each value of the middle variable, the model was evaluated 200 times under 5-fold cross-validation, and the average results were used to check the specificity value. The above-described process was done until specificity was found to be consistent with Craig-Schapiro [11]. The trained-cum-validated prediction model’s performances were evaluated by comparing their sensitivity, accuracy, Precision Area under the curve (PR-AUC), and ROC-AUC values at specificity value in line with the earlier studies conducted by Craig-Schapiro [11].

Finally, after calibrating the model (s) using the best threshold value, we performed predictions on the test dataset. For brevity, other models were also trained and tested at our estimated optimum threshold, and a comparative assessment was made between our results and that of the model developed by Craig-Schapiro *et al.*, 2012 [11].

F. INDEPENDENT TEST DATASET

In the current study, the trained-cum-validated models generated using the various subset of features obtained from RFE and the L1 regularization method were re-evaluated using five-fold cross-validation on the 200 copies of 20% independent test data. The total number of instances in the test data was 67. Here the positive class (MCI patients) consisted of 21 instances, and the negative class (Cognitively normal subjects) consisted of 46 instances. It is recommended to do testing on independent data to eliminate the predictive modeling bias. Five-fold cross-validation over all of the 200 copies of 20% independent test datasets was performed, and the final ensemble model was created by taking a weighted average of LR and Linear SVM.

G. MODEL EVALUATION METRICS

1) ACCURACY (ACC) AND CONFUSION MATRIX

The ACC is calculated as the total number of correct predictions (TP + TN) divided by the total number of instances in a dataset ((Positive + Negative) instances). The equation of accuracy is shown below. Refer to equation “16”:

$$\frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (16)$$

An accuracy of 1.0 is considered the best, while “0.0” is considered the worst. Classification accuracy is an excellent statistical evaluator only when both positive and negative instances in the dataset are equal in number, i.e., the classes are balanced. On the contrary, if the classes are imbalanced, then classification accuracy gives us a false sense of the correctness (or accuracy) of any classification algorithm (classifier) [41]. Under such circumstances, calculating a confusion matrix [42], [43] can give us insight into the errors and the types of errors our classification model is making. It also tells us what predictions the model is getting right. Compared to the confusion matrix’s actual outcomes, this breakdown of correct and incorrect predictions overcomes the limitations of using accuracy alone as a statistical evaluator for assessing a classification model’s performance. Besides, if anyone desires to avoid false-positive more than false negative or vice versa, other statistical evaluators, namely Sensitivity (Recall or True positive Rate (TPR)) and Specificity, are more informative than accuracy [41].

2) SENSITIVITY

Sensitivity (SN) is calculated as the number of correctly predicted positive instances divided by the total number of positive instances in a dataset. The best SN value is 1.0, and the worst is 0.0.

$$\frac{TP}{TP + FN} \quad (17)$$

The equation for determining sensitivity is depicted below. Refer to “17”:

3) SPECIFICITY

Specificity (SP) is calculated as the number of correctly classified negative instances divided by the total number of negative samples in a dataset. The best specificity value is 1.0, and the worst is 0.0.

$$\frac{TN}{TN + FP} \quad (18)$$

The equation for calculating specificity is depicted below. Refer to “18”:

4) RECEIVER OPERATING CHARACTERISTIC CURVE (ROC)

The ROC is a curve generated by plotting the TPR (Sensitivity/Recall) against the True Negative Rate (TNR) (1-Specificity) at various classification or decision thresholds. The Area Under the ROC-AUC Curve value measures a classification model’s quality to distinguish between two classes in a dataset. A random classifier has an AUC value

TABLE 1. List of the subset of features (CSF protein biomarkers) obtained using the RFE method.

Subset	I	II	III	IV	V
Number of features	1	2	3	4	5
SVM-RFE	['tau']	[Cystatin_C, tau]	[Cystatin_C, MMP10, tau]	[Apolipoprotein_A1, Cystatin_C, MMP10, tau]	[Apolipoprotein_A1, Clusterin_Apo_J, Cystatin_C, MMP10, tau]
LR-RFE	['tau']	['VEGF', 'tau']	['Cystatin_C', 'MMP10', 'tau']	['Cystatin_C', 'MMP10', 'VEGF', 'tau']	['Clusterin_Apo_J', 'Cystatin_C', 'MMP10', 'VEGF', 'tau']
DT-RFE	['Ab_42']	['Cystatin_C', 'Ab_42']	['Cystatin_C', 'MMP10', 'Ab_42']	['Cystatin_C', 'MMP10', 'tau', 'Ab_42']	['Cystatin_C', 'MMP10', 'Prolactin', 'tau', 'Ab_42']
RF-RFE	['Ab_42']	['tau', 'Ab_42']	['MMP10', 'tau', 'Ab_42']	['MMP10', 'tau', 'p_tau', 'Ab_42']	['MMP10', 'VEGF', 'tau', 'p_tau', 'Ab_42']

of 0.5, and a perfect classifier that can correctly discriminate two classes in a dataset has an AUC value equal to 1 or vice versa.

5) PRECISION-RECALL (PR) CURVES

Similar to the closely-related ROC curves, the PR curves are an estimation tool for binary classification that helps us visualize various machine learning algorithms' performance at several classification thresholds ranging between 0 and 1. PR curves are used, particularly for imbalanced or skewed data sets, where one class's instances are observed more regularly compared to the other class. Therefore, on these skewed (imbalanced) data sets, PR curves are a suitable alternative to ROC curves that would show quite a massive performance difference between two algorithms that cannot be represented appropriately using the ROC curves [44]. In addition to visual evaluation of a PR curve, the Area under a PR curve (AUPR) is often used for assessing the performance of a machine learning algorithm, regardless of any particular operating point or threshold. The higher the AUPR value, the better the model predicts positive instances as positives (True positive (TP)). The AUPR values range from 0 to 1 with a score of 1 depicting a perfect model, i.e., the model can predict the entire positive instances in the dataset as True Positives with no false negatives and false-positive predictions.

H. STUDENT T-TEST

A two-tailed unpaired student t-test [45] was performed to estimate the significant difference (i.e., p-value < 0.05) between the mean value of the three features (Cystatin C, MMP10, and tau proteins) across two types of population (CDR 0 and CDR > 0). Likewise, a one-tailed unpaired t-test [45] was performed at a significance level of 0.5 to show that our model's performance is better than the model based on traditional biomarkers and features that achieved the best ROC_AUC result Craig-Schapiro *et al.* 2012 [11].

I. IMPLEMENTATION: WEB APPLICATION

The ensemble-learning based predictive model was developed using the pruned MCI dataset with three features.

The model was saved in the python pickle file format (.pk). Flask, a light-weight micro-web framework with Jinja2 templating, was used to develop the web application. Later, the Flask web application was deployed on Heroku, a cloud-based platform, to build and run web applications exclusively in the cloud. The output of the application is probability-based. Individuals with an output probability score greater than 0.5 have a higher probability of being diagnosed with AD than an individual with a score lesser or equal to 0.5.

III. RESULTS

A. BEST INFORMATIVE FEATURES SELECTION AND CLASSIFIER EVALUATION

The RFE employing cost-sensitive classifiers with weights was used to select the most informative set of model building features. The subset of features obtained using RFE based FS method is tabulated in Table 1. The Comparative performance evaluation of different sets of features listed in Table 1 was trained and tested on four different cost-sensitive classifiers (RF, LR, LibSVM, and DT), are shown in Fig. 8(a-d). Five-fold cross-validation over all of the 200 training datasets was performed, and the final results averaged to find the best classifier and the best subset of features. As per the comparative performance evaluation results, as shown in Fig. 8(a-d), the LR and linear SVM model built using a subset of three features (Cystatin C, MMP10, and tau proteins) performed well in discriminating mild cognitive impaired (CDR > 0) from cognitively normal subjects with better sensitivity PR-AUC and ROC-AUC at specificity in line with Craig-Schapiro *et al.*, 2012 [11].

Further, a comparative average performance evaluation of RFE-classifier based models built using a subset of the three most informative features in discriminating cognitive normal (CDR 0) from cognitively impaired (CDR > 0) subjects are shown in Table 2.

The probability distributions of the three proteins (Cystatin C, MMP10, and tau proteins) between the two different populations (CDR 0 and CDR > 0) are shown via histogram plots in Fig. 9(a-c).

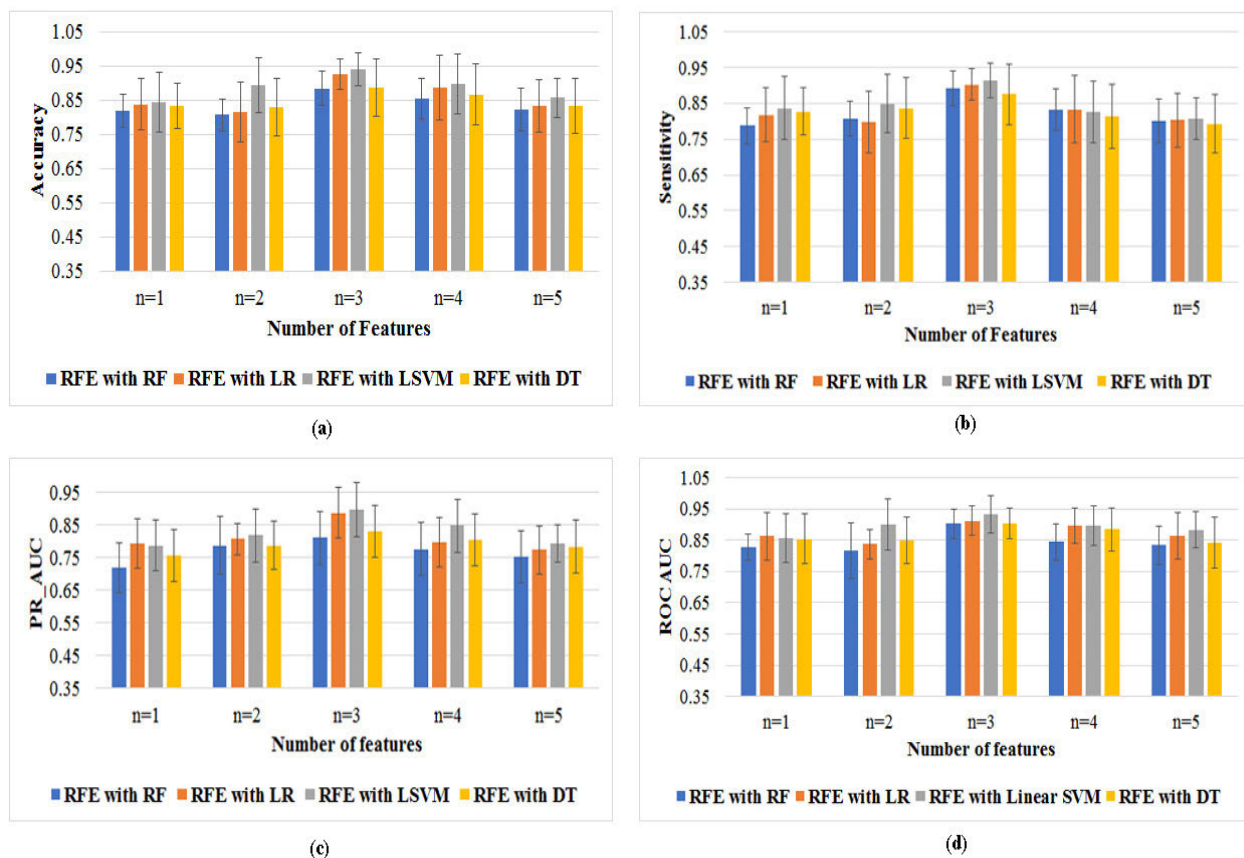


FIGURE 8. (a-d). Comparative evaluation of performance measures. (a) Accuracy. (b) Sensitivity. (c) PR_AUC. and (d) ROC-AUC in models built using a different subset of features.

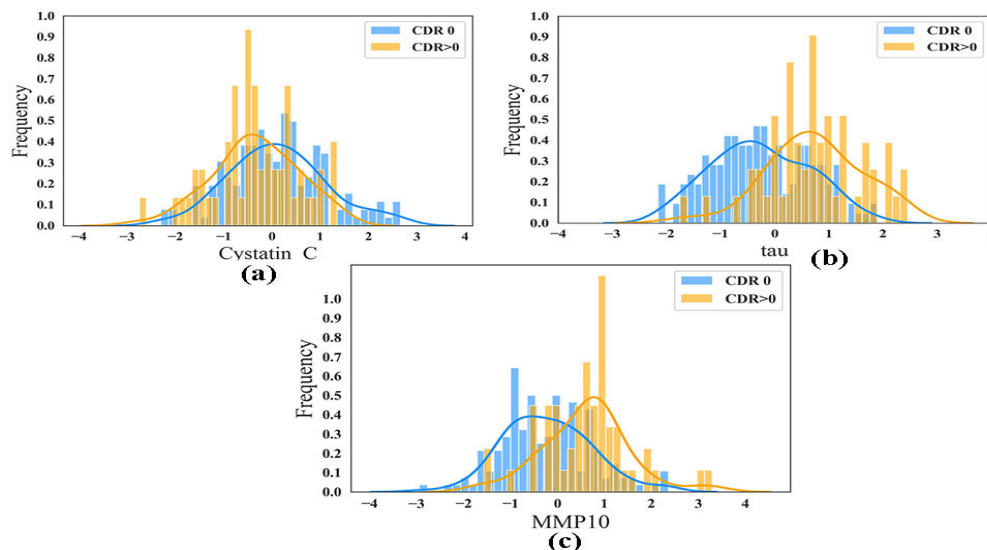


FIGURE 9. (a-c). Histogram representing the frequency distribution of protein biomarker. (a) Cystatin_C, (b) MMP 10, and (c) tau, respectively across two population categories (CDR 0 and CDR > 0).

The p-value of the two-tailed unpaired t-test for Cystatin C, MMP10, and tau proteins selected using the LinearSVM classifier based RFE method is tabulated in Table 3.

The observed p-values of the three selected protein biomarkers were lower than the significance level of 0.05. Therefore, the mean population of selected attributes between cognitively impaired and cognitively normal groups

TABLE 2. Comparative average performance evaluation of RFE-classifier based models built using a subset of the three most informative features.

Classifier	Accuracy	Sensitivity	PR_AUC	ROC_AUC
LR	0.9254 ± 0.045	0.9023 ± 0.035	0.8851 ± 0.078	0.918 ± 0.046
Linear SVM	0.9379 ± 0.048	0.9139 ± 0.028	0.8951 ± 0.082	0.9523 ± 0.060
DT	0.8846 ± 0.084	0.8746 ± 0.072	0.8299 ± 0.082	0.9039 ± 0.051
RF	0.8838 ± 0.049	0.8918 ± 0.341	0.8099 ± 0.063	0.9020 ± 0.048

TABLE 3. List of informative attributes selected from the Mild Cognitive Impaired (MCI) clinical dataset using recursive elimination method (RFE).

Sl. No.	Names of Attributes	Significance level
		p < 0.5 between CDR 0 and CDR 1 subjects
1	Cystatin_C	1.8x10-2 (p < 0.5)
2	MMP10	3.95x10-4 (p < 0.5)
3	Tau	1.43x10-9 (p < 0.5)

TABLE 4. Estimated coefficient and the list of selected features using LASSO or L1 regularization method.

Hyperparameters and Variables	Coefficients
Intercept	0.2734
Tuned Lamda (λ)	0.0055
Mean Square Error (MSE)	0.11501
GRO_alpha	0.000437
Tau	0.106034
Cystatin_C	-0.008370
VEGF	-0.021188
Ab_42	-0.072693

Comparative performance evaluation of the four classifier-based models is represented in Fig. 11. We can observe from Fig. 11 that the cost-sensitive LR classifier based model with an accuracy (accuracy (0.8864 ± 0.098), sensitivity (0.8663 ± 0.048), PR_AUC (0.8266 ± 0.081), and ROC_AUC (0.8966 ± 0.054) outperform the other classifier based models. Comparative performance evaluation of the two best models generated using RFE and L1 regularization at a 5% significance level is tabulated in Table 5.

It can be observed from Table 5 that the Linear SVM-RFE based model’s performance built using a feature subset of three features (Cystatin_C, MMP10, and Tau) performed significantly better than the LR-based model built using a subset of features generated using L1 regularization at a 5% significance level. Finally, the MCI test dataset was pruned based on the selected subset of informative features (Cystatin C, MMP10, and tau proteins).

B. ENSEMBLE-LEARNING MODEL BUILDING AND OPTIMIZATION OF THRESHOLD

The PCC score on the outputs of LR and SVM was found to be 0.63, signifying that an average ensemble would result in a predictor which is better than the individual models. A weighted average of LR and Linear SVM’s soft output obtained from the classifier-RFE method was used to create an ensemble. When running on 5-fold cross-validation overall 200 training dataset, the ensemble model’s weighted-average results show that the threshold value of 0.321 is optimum for correctly classifying instances in the dataset. Finally, our calibrated ensemble model was tested on the 200 copies of 20% independent test datasets using the optimum threshold value, and the weighted-average out results of each class are shown in Table 6. Given the tested model’s overall prediction, the weighted average ensemble model’s accuracy was 0.9552 ± 0.025.

The performance of the ensemble model in terms of the confusion matrix is shown in Fig. 12. A comparative evaluation of the ensemble model with the base classifier is tabulated in Table 7.

The sensitivity, PR_AUC, and ROC_AUC values of the ensemble model were observed to be considerably better than the individual best-performing classifiers (Linear SVM and LR) at a 5% significance level as shown in Table 7.

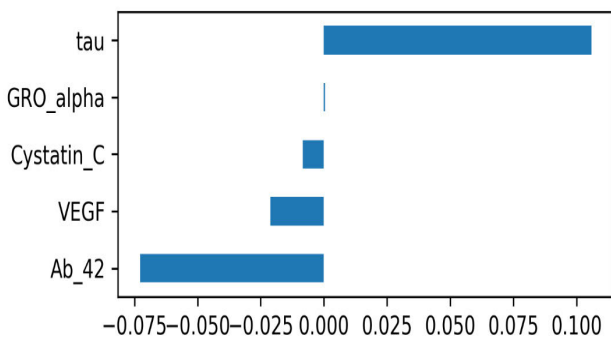
C. COMPARATIVE PERFORMANCE EVALUATION

For brevity, our calibrated ensemble model was trained and tested on datasets generated using the state-of-

FIGURE 10. Feature importance using LASSO model.

is significant. Additionally, the LASSO algorithm was also applied for parameter estimation and feature selection. The selected subset of features with their coefficient and Mean Square Error (MSE) along with the tuned Lambda hyperparameter value is tabulated in Table 4.

The feature importance of the optimal set of features selected using LASSO is represented in Fig. 10. A set of four cost-sensitive classifiers (RF, LR, LibSVM, and DT) based models built using the optimal subset of features GRO_alpha, tau, Cystatin_c, VEGF, and Aβ42) obtained via L1 regularization were evaluated based on the following performance measures: Accuracy, ROC-AUC, PR_AUC, and sensitivity.



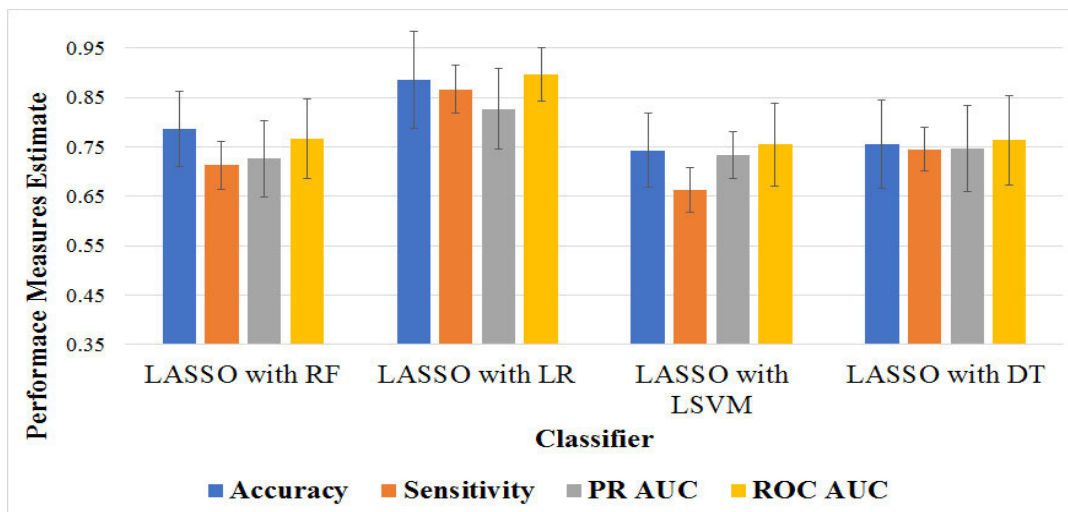


FIGURE 11. Comparative evaluation of performance measures. (a) Accuracy. (b) Sensitivity. (c) PR_AUC and (d) ROC-AUC in models built using a subset of features generated using the LASSO or L1 Regularization method.

TABLE 5. Comparative independent two-sample t-test between model built using SVM-RFE and model built using features selected using LASSO method at a significance level of 5%.

Performance Measures	Models and Feature Selection Method		P-Value (significance level $p < 0.5$)
	Linear SVM with RFE method	LR with L1 Regularization	
Accuracy	0.9379 ± 0.048	0.8864 ± 0.098	0.0001 ($p < 0.5$)
Sensitivity	0.9139 ± 0.053	0.8663 ± 0.048	0.0001 ($p < 0.5$)
ROC_AUC	0.9523 ± 0.060	0.8966 ± 0.054	0.0062 ($p < 0.5$)
PR_AUC	0.8951 ± 0.032	0.8266 ± 0.021	0.0001 ($p < 0.5$)

TABLE 6. Classification report of weighted averaged ensemble method.

Classes	Sensitivity	Precision	ROC_AUC
MCI Patient (Positive class)	0.9565 ± 0.024	0.9777 ± 0.014	0.9799 ± 0.055
Cognitively Normal (Negative Class)	0.9090 ± 0.032	0.9524 ± 0.022	0.9799 ± 0.055

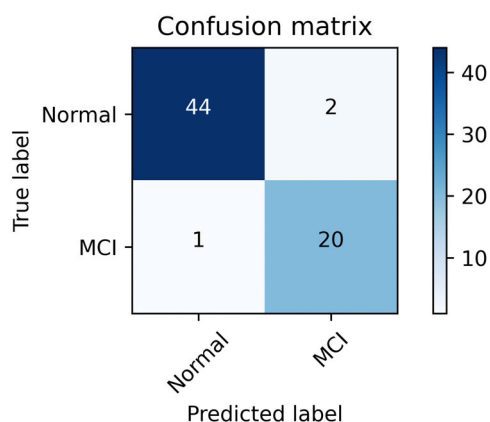


FIGURE 12. Confusion matrix of the final weighted ensemble model tested on the 20% independent test dataset.

the-art features (tau, tau, and $A\beta_{42}$) as well as features that achieved the best ROC_AUC results for

Craig-Schapiro *et al.*, 2012 [11]. A comparative performance assessment of the models, as mentioned above, at a 5% significance level, is tabulated in Table 8.

It can be observed from Table 7 that the calibrated weighted average ensemble model’s performance built using a feature subset of three features (Cystatin_C, MMP10, and Tau) performed significantly better than the model built using the state-of-the-art features (tau, tau, and $A\beta_{42}$). Additionally, our proposed ensemble model also performed significantly better at a 5% significance level than the model built using a combination of CSF protein biomarkers that achieved the best ROC_AUC results for Craig-Schapiro *et al.*, 2012 [11].

Histogram plots of the ROC_AUC scores generated from each of the 200 runs are used to approximately depict the difference in the probability distributions between models built using conventional CSF biomarkers, the best ROC_AUC biomarker proposed by Craig-Schapiro *et al.*, 2012 and our proposed CSF protein biomarkers are shown in Fig. 13 and Fig. 14, respectively.

TABLE 7. Performance evaluation of the calibrated weighted-averaged ensemble model at a 5% significance level in discriminating cognitive normal (CDR 0) from cognitively impaired (CDR > 0) subjects.

Performance Measures	Models and Feature Selection Method		P-Value (significance level $p < 0.5$)
	Weighted Average Ensemble model	Linear SVM with RFE method	
Accuracy	0.9552 ± 0.025	0.9379 ± 0.048	< 0.0001 ($p < 0.5$)
Sensitivity	0.9565 ± 0.024	0.9139 ± 0.053	< 0.0001 ($p < 0.5$)
ROC_AUC	0.9799 ± 0.055	0.9523 ± 0.060	< 0.0001 ($p < 0.5$)
PR_AUC	0.9108 ± 0.015	0.8951 ± 0.032	< 0.0001 ($p < 0.5$)
	Weighted Average Ensemble model	LR with RFE method	P-Value (significance level $p < 0.5$)
Accuracy	0.9552 ± 0.025	0.9254 ± 0.045	< 0.0001 ($p < 0.5$)
Sensitivity	0.9565 ± 0.024	0.9023 ± 0.035	< 0.0001 ($p < 0.5$)
ROC_AUC	0.9799 ± 0.055	0.918 ± 0.046	< 0.0001 ($p < 0.5$)
PR_AUC	0.9108 ± 0.015	0.8851 ± 0.014	< 0.0001 ($p < 0.5$)
	Linear SVM with RFE method	LR with RFE method	P-Value (significance level $p < 0.5$)
Accuracy	0.9379 ± 0.048	0.9254 ± 0.045	0.0075 ($p < 0.5$)
Sensitivity	0.9139 ± 0.053	0.9023 ± 0.035	0.0102 ($p < 0.5$)
ROC_AUC	0.9523 ± 0.060	0.918 ± 0.046	0.0078 ($p < 0.5$)
PR_AUC	0.8951 ± 0.032	0.8851 ± 0.014	< 0.0001 ($p < 0.5$)

TABLE 8. Comparative averaged out performance evaluation of models at a 5% significance level in classifying cognitive normal (class 0) from cognitively impaired (class 1) participants.

Model Evaluation metrics	Models		P-Value (significance level $p < 0.5$)
	Calibrated Weighted Average Ensemble Model built using a novel combination of CSF protein biomarkers (Cystatin C, MMP10, and tau proteins)	State-of-the-art CSF proteins biomarker-based model (tau, pTau, and Aβ42)	
Accuracy	0.9552 ± 0.025	0.8217 ± 0.054	< 0.0001 ($p < 0.5$)
Sensitivity	0.9565 ± 0.024	0.8105 ± 0.094	< 0.0001 ($p < 0.5$)
ROC_AUC	0.9799 ± 0.055	0.6768 ± 0.107	< 0.0001 ($p < 0.5$)
PR_AUC	0.9108 ± 0.015	0.8515 ± 0.054	< 0.0001 ($p < 0.5$)
	Models		P-Value (significance level $p < 0.5$)
	Calibrated Weighted Average Ensemble Model built using a novel combination of CSF protein biomarkers (Cystatin C, MMP10, and tau proteins)	Best ROC_AUC biomarker-based model (Craig-Schapiro <i>et al.</i> 2012)	
Accuracy	0.9552 ± 0.025	0.8146 ± 0.057	< 0.0001 ($p < 0.5$)
Sensitivity	0.9565 ± 0.024	0.8076 ± 0.101	< 0.0001 ($p < 0.5$)
ROC_AUC	0.9799 ± 0.055	0.7191 ± 0.102	< 0.0001 ($p < 0.5$)
PR_AUC	0.9108 ± 0.015	0.8587 ± 0.057	< 0.0001 ($p < 0.5$)

The histogram plot is also used to estimate a statistically significant change between the models.

A one-tailed unpaired t-test was used to estimate how our model is better than models generated using traditional and best ROC_AUC value attributes from studies

conducted by Craig-Schapiro *et al.*, 2012. The p-value of the one-tailed paired t-test for the comparative evaluation of our proposed ensemble model and models generated using traditional CSF protein biomarkers as well as biomarkers which achieved higher ROC_AUC value from Craig-

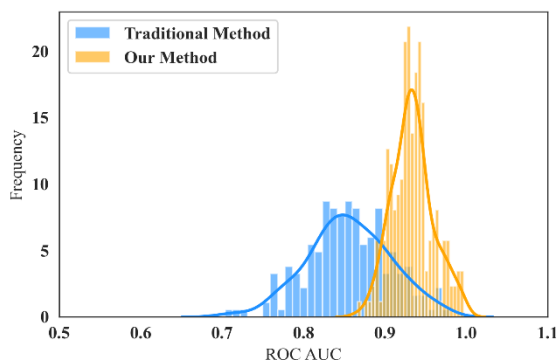


FIGURE 13. Histogram representing ROC_AUC's frequency distribution between the predictive models built using our proposed attributes and traditional CSF protein biomarkers (Tau, p-tau, and A β 42).

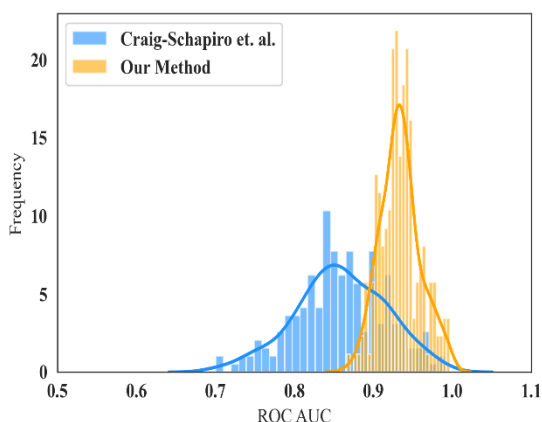


FIGURE 14. Histogram representing the frequency distribution of ROC_AUC between the predictive models built using our proposed attributes and the attributes which show the best ROC_AUC values from studies conducted by Craig-Schapiro *et al.*, 2012.

Schapiro *et al.*, 2012 was observed to be 4.3002×10^{-36} and 2.398×10^{-38} respectively which was much lower than the significance level of 0.05.

D. PREDICTIVE WEB-APPLICATION IMPLEMENTATION

The web-based predictive application based on a novel combination of three CSF protein biomarkers to predict the earlier stages of AD has been made live on Heroku as “<https://appsalzheimer.herokuapp.com>.”

IV. DISCUSSION

The MCI dataset was downloaded from Figshare [30]. The clinical dataset consisted of 91 MCI subjects and 242 cognitively normal subjects. The subjects' cognitive impairment status in the MCI clinical dataset was as per the CDR scale ranging from 0 to 1, where CDR 0 was considered a cognitively normal subject, while CDR 0.5 and CDR 1 were considered very mild and Mildly cognitive subjects, respectively. The raw clinical dataset consisted of 124 features encompassing CSF protein biomarkers, allelic variants of Apolipoprotein E genotype (E2, E3, and E4), and demographic features, namely gender and age. All categorical features were transformed into a binary-encoded format by using the one-hot-encoding technique. All duplicate instances

were removed as well. This study focuses on the two feature selection methods, namely, the RFE method, an important example of the wrapper-based feature selection method, and the LASSO method based on an embedded method of feature selection. Since there is a class imbalance in the dataset, in the RFE based method of feature selection, cost-sensitive classifiers, namely, RF, DT, LR, and LinearSVM, with a series of feature sets containing features ranging from 1 to $(1+n)$ were selected. Here, “ n ” varies from 0 to 4. Likewise, the number of attributes removed at each iteration in RFE was set to be three features.

A subset consisting of three best-performing features, namely, Cystatin C, MMP10, and tau, were generated from the cost-sensitive SVM-RFE and LR-RFE based feature selection process. The Linear SVM and the LR-based model built using three features outperformed other classifier-based models built using the other subset of features generated using the classifier based RFE method. The LR-based model built using the five most important features selected using L1 regularization performed better than the RF, DT, and Linear SVM-based models. However, the LR-model performance built using features screened using L1 regularization was significantly ($p < 0.5$) lower than the LinearSVM based model built using a subset of three most informative features screened using the classifier based RFE feature selection method. Furthermore, we observed a significant difference between the mean of all three protein biomarkers (Cystatin C, MMP10, and tau proteins) derived from the RFE method among subjects belonging to cognitively normal (CDR 0) and Mild cognitive Impaired (CDR > 0) groups at a significance level of 0.05.

The candidate protein biomarker (Cystatin C, MMP10, and tau proteins) screened in this study belongs to different pathways and a functional group whose association with AD pathophysiology has been investigated and documented, most notably, tau protein, a traditional CSF biomarker which has proven useful in the diagnosis and prognosis of AD. In some individuals with a very MCI, CSF tau levels have shown gradual increase years before being ultimately diagnosed with AD [15]–[17]. Therefore, the tau protein has been useful for predicting AD's onset in individuals with very mild or mild cognitive disorders. Cystatin C has a pivotal role in AD pathophysiology as CysC concentration modulates A β amyloidogenesis and oligomerization [46], [47]. MMPs play a vital role as an inflammatory element in AD's disordered physiological process. MMPs transcription is induced through posttranslational modification by the inflammatory mediators (e.g., free radicals or cytokines) and inhibitor proteins (e.g., Metalloproteinase inhibitors (TIMPs)). The activated MMPs remodel the pericellular environment by regulating the extracellular matrix (ECM) and the tight junction's breakdown. The MMPs also interact and alter the properties of growth factors, cell surface components, and signaling molecules, leading to neuroinflammation, cell death, and neurotoxicity [48]. In this context, separate studies conducted by Duits *et al.*, 2015 [49] and Whelan *et al.*, 2019 [50] showed a significant increase in MMP-10 in AD-dementia A β + MCI patients as

compared to cognitively normal individuals. The altered level of MMP-10 in AD and MCI patients suggest the involvement of MMP-10 protein in the pathology of AD and, thus, possible use of MMP-10 as a protein biomarker for earlier diagnosis of AD. So, the implementation of the screened CSF biomarkers, whose association with the pathology of AD is well studied, may improve the reliability of models built to predict AD's earlier stages.

The classifiers' comparative performance evaluation using different subsets of features obtained using the RFE and LASSO method generated a consistently accurate model. However, RFE based LR and Linear SVM model achieved the best results, with a subset consisting of the three most informative subset features. Therefore, an ensemble model of the two best performing classifier was generated. Since the ensemble is probabilistic by nature, it requires a threshold for producing binary output. Thus, the ensemble was calibrated to obtain an optimum threshold for predicting the earlier stages of AD. Our ensemble model's weighted average out results were significantly better than the individual performance of LR and Linear SVM base classifiers tested on the MCI dataset in terms of sensitivity, ROC_AUC, and PR_AUC values. The weighted average result of our ensemble model was compared to models generated using traditional combinations of CSF protein biomarkers (Tau, p-tau, and A β -42) and biomarkers, which achieved the best ROC_AUC result for Craig-Schapiro *et al.*, 2012 [11] at a 5% significance level.

Our novel combination of CSF protein biomarkers based ensemble model performed significantly better in classifying cognitively impaired subjects (Positive instances) from the MCI dataset compared to models generated using the state-of-the-art and best-performing protein biomarkers obtained from a study conducted by Craig-Schapiro *et al.*, 2012 [11]. Lastly, recent studies by Spellman *et al.* [51] and Llano *et al.* [52] examined a clinical Alzheimer's disease Neuroimaging Initiative (ADNI) dataset and proposed a peptide signature analyte-based multivariate model to predict the earlier stages of AD. The study conducted by them found several overlapping potential signature peptides, namely Aldolase A, Fructose-Bisphosphate peptide (ALDOA), Fatty Acid-Binding Protein, Heart (FABPH), Neuronal pentraxin receptor (NPTXR), Peroxiredoxin-1 (PRDX1), and Neurosecretory protein VGF (VGF), for predicting the earlier stages of AD and future disease progression. The multivariate modeling approaches proposed by Spellman *et al.* [51] and Llano *et al.* [52] were able to differentiate AD from Non-AD subjects with a ROC_AUC of 0.74 and 0.89, respectively. Conversely, our ensemble model performance with a ROC_AUC value of 0.9499 ± 0.055 performed better in classifying cognitively impaired subjects from cognitively normal subjects than the traditional combination of CSF biomarkers and signature peptides from Spellman *et al.* [51] and Llano *et al.* [52] studies, respectively.

Additionally, our ensemble model has shown considerably better ROC_AUC performance than the multiple marker studies in the past involving multiple imaging modalities,

neuropsychological testing, and APO_E genotype to predict the earlier stages of AD [24], [53]. Considering the higher AU-PR and the AUC_ROC attained by our model, we can say that the profiling of our novel combination of CSF protein can be recommended for clinical tests for predicting the earlier stages of AD. Our web-based application to predict the earlier stages of AD has been successfully implemented and has been made live on Heroku at "https://appsalzheimer.herokuapp.com."

V. CONCLUSION AND FUTURE SCOPE

The current proposed model uses weighted average ensemble-based learning methods to build a predictive model that can easily discriminate MCI subjects from healthy subjects with higher sensitivity, ROC_AUC, and PR_AUC value. On the other hand, the multiple marker studies are costly and impractical to attain all of these biomarkers from a single patient. Thus, our contemporary ensemble approach's unique benefit is its cost-effectiveness, as the profiling of our novel combination of CSF protein can precisely classify patients with earlier stages of AD. Also, we built a web-based live predictive system built using a novel combination of CSF protein biomarkers, which is the first of its kind online service for predicting the earlier stages of AD. We believe that such our application built using the most informative novel combination of features will benefit researchers and doctors in diagnosing very mild or Mild CI disorders, thereby assisting AD's earlier prediction.

ACKNOWLEDGMENT

This project was funded by the Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah, under grant No. (DF-676-305-1441). The authors, therefore, gratefully acknowledge DSR technical and financial support.

REFERENCES

- [1] G. Livingston, J. Huntley, A. Sommerlad, D. Ames, C. Ballard, S. Banerjee, C. Brayne, A. Burns, J. Cohen-Mansfield, C. Cooper, and S. G. Costafreda, "Dementia prevention, intervention, and care: 2020 report of the lancet commission," *Lancet*, vol. 396, no. 10248, pp. 413–446, Aug. 2020.
- [2] M. A. DeTure and D. W. Dickson, "The neuropathological diagnosis of Alzheimer's disease," *Mol. Neurodegener.*, vol. 14, no. 1, p. 32, 2019.
- [3] M. Pais, L. Martinez, O. Ribeiro, J. Loureiro, R. Fernandez, L. Valiengo, P. Canineu, F. Stella, L. Talib, M. Radanovic, and O. V. Forlenza, "Early diagnosis and treatment of Alzheimer's disease: New definitions and challenges," *Rev. Psychiatry*, vol. 42, no. 4, pp. 431–441, Aug. 2020.
- [4] F. L. Guest, "Early detection and treatment of patients with Alzheimer's disease: Future perspectives," *Adv. Express Med. Biol.*, vol. 1118, pp. 295–317, Dec. 2019.
- [5] C. R. Jack, Jr., D. A. Bennett, K. Blennow, M. C. Carrillo, B. Dunn, S. B. Haeberlein, D. M. Holtzman, W. Jagust, F. Jessen, J. Karlawish, and E. Liu, "NIA-AA research framework: Toward a biological definition of Alzheimer's disease," *Alzheimer's Dementia*, vol. 14, no. 4, pp. 535–562, Apr. 2018.
- [6] F. Gao, H. Yoon, Y. Xu, D. Goradia, J. Luo, T. Wu, and Y. Su, "AD-NET: Age-adjust neural network for improved MCI to AD conversion prediction," *NeuroImage, Clin.*, vol. 27, Jun. 2020, Art. no. 102290.
- [7] K. Huang, Y. Lin, L. Yang, Y. Wang, S. Cai, L. Pang, X. Wu, and L. Huang, "A multipredictor model to predict the conversion of mild cognitive impairment to Alzheimer's disease by using a predictive nomogram," *Neuropsychopharmacology*, vol. 45, no. 2, pp. 358–366, Jan. 2020.

- [8] Z. Sun, M. van de Giessen, B. P. F. Lelieveldt, and M. Staring, "Detection of conversion from mild cognitive impairment to Alzheimer's disease using longitudinal brain MRI," in *Frontiers in Neuroinformatics*, vol. 11. Leiden, The Netherlands: Leiden Univ. Medical Center Leiden, Division of Image Processing, Department of Radiology, 2017, p. 16.
- [9] M. Mitolo, M. Stanzani-Maserati, S. Capellari, C. Testa, P. Rucci, R. Poda, F. Oppi, R. Gallassi, L. Sambati, G. Rizzo, P. Parchi, S. Evangelisti, L. Talozzi, C. Tonon, R. Lodi, and R. Liguori, "Predicting conversion from mild cognitive impairment to Alzheimer's disease using brain 1H-MRS and volumetric changes: A two-year retrospective follow-up study," *NeuroImage, Clin.*, vol. 23, 2019, Art. no. 101843.
- [10] Y. Sun, Q. Bi, X. Wang, X. Hu, H. Li, X. Li, T. Ma, J. Lu, P. Chan, N. Shu, and Y. Han, "Prediction of conversion from amnesic mild cognitive impairment to Alzheimer's disease based on the brain structural connectome," *Frontiers Neurol.*, vol. 9, p. 1178, Jan. 2019.
- [11] R. Craig-Schapiro, M. Kuhn, C. Xiong, E. H. Pickering, J. Liu, T. P. Misko, R. J. Perrin, K. R. Bales, H. Soares, A. M. Fagan, and D. M. Holtzman, "Multiplexed immunoassay panel identifies novel CSF biomarkers for Alzheimer's disease diagnosis and prognosis," *PLoS ONE*, vol. 6, no. 4, Apr. 2011, Art. no. e18850.
- [12] R. Khoury and E. Ghossoub, "Diagnostic biomarkers of Alzheimer's disease: A state-of-the-art review," *Biomarkers Neuropsychiatry*, vol. 1, Dec. 2019, Art. no. 100005.
- [13] J. E. Park, K. Y. Choi, B. C. Kim, S. M. Choi, M. K. Song, J. J. Lee, J. Kim, H. C. Song, H. W. Kim, J. M. Ha, and E. H. Seo, "Cerebrospinal fluid biomarkers for the diagnosis of prodromal Alzheimer's disease in amnesic mild cognitive impairment," *Dementia Geriatric Cogn. Disorders Extra*, vol. 9, no. 1, pp. 100–113, 2019.
- [14] S.-K. Herukka, A. H. Simonsen, N. Andreasen, I. Baldeiras, M. Bjerke, K. Blennow, S. Engelborghs, G. B. Frisoni, T. Gabryelewicz, S. Galluzzi, and R. Handels, "Recommendations for cerebrospinal fluid Alzheimer's disease biomarkers in the diagnostic evaluation of mild cognitive impairment," *Alzheimer's Dement.*, vol. 13, no. 3, pp. 285–295, Mar. 2017.
- [15] S. Lehmann, J. Dumurgier, X. Ayrygnac, C. Marelli, D. Alcolea, J. F. Ormaechea, E. Thouvenot, C. Delaby, C. Hirtz, J. Vialaret, N. Ginestet, E. Bouazziz-Amar, J.-L. Laplanche, P. Labauge, C. Paquet, A. Lleo, and A. Gabelle, "Cerebrospinal fluid a beta 1–40 peptides increase in Alzheimer's disease and are highly correlated with phospho-tau in control individuals," *Alzheimer's Res. Therapy*, vol. 12, no. 1, p. 123, Dec. 2020.
- [16] S. A. Kent, T. L. Spires-Jones, and C. S. Durrant, "The physiological roles of tau and A β : Implications for Alzheimer's disease pathology and therapeutics," *Acta Neuropathol.*, vol. 140, no. 4, pp. 417–447, Oct. 2020.
- [17] C. R. Jack, H. J. Wiste, T. M. Therneau, S. D. Weigand, D. S. Knopman, M. M. Mielke, V. J. Lowe, P. Vemuri, M. M. Machulda, C. G. Schwarz, and J. L. Gunter, "Associations of amyloid, tau, and neurodegeneration biomarker profiles with rates of memory decline among individuals without dementia," *JAMA*, vol. 321, no. 23, pp. 2316–2325, Jun. 2019.
- [18] P. Lewczuk, M. Łukasiewicz-Zajac, P. Mroczko, and J. Kornhuber, "Clinical significance of fluid biomarkers in Alzheimer's Disease," *Pharmacol. Rep.*, vol. 72, no. 3, pp. 528–542, Jun. 2020.
- [19] J. C. Lee, S. J. Kim, S. Hong, and Y. Kim, "Diagnosis of Alzheimer's disease utilizing amyloid and tau as fluid biomarkers," *Express Mol. Med.*, vol. 51, no. 5, pp. 1–10, May 2019.
- [20] K. E. J. Wesenhausen, C. E. Teunissen, P. J. Visser, and B. M. Tijms, "Cerebrospinal fluid proteomics and biological heterogeneity in Alzheimer's disease: A literature review," *Crit. Rev. Clin. Lab. Sci.*, vol. 57, no. 2, pp. 86–98, Feb. 2020.
- [21] Y. Gupta, R. K. Lama, and G.-R. Kwon, "Prediction and classification of Alzheimer's disease based on combined features from Apolipoprotein-E genotype, cerebrospinal fluid, MR, and FDG-PET imaging biomarkers," *Frontiers Comput. Neurosci.*, vol. 13, p. 72, Oct. 2019.
- [22] H. Li, M. Habes, D. A. Wolk, and Y. Fan, "A deep learning model for early prediction of Alzheimer's disease dementia based on hippocampal magnetic resonance imaging data," *Alzheimer's Dementia*, vol. 15, no. 8, pp. 1059–1070, Aug. 2019.
- [23] M. Grassi, N. Rouleaux, D. Caldirola, D. Loewenstein, K. Schruers, G. Perna, M. Dumontier, and A. Disease Neuroimaging Initiative, "A novel ensemble-based machine learning algorithm to predict the conversion from mild cognitive impairment to Alzheimer's disease using socio-demographic characteristics, clinical information, and neuropsychological measures," *Frontiers Neurol.*, vol. 10, p. 756, Jul. 2019.
- [24] G. Lee, K. Nho, B. Kang, K.-A. Sohn, and D. Kim, "Predicting Alzheimer's disease progression using multi-modal deep learning approach," *Sci. Rep.*, vol. 9, no. 1, p. 1952, Dec. 2019.
- [25] F. Li, L. Tran, K.-H. Thung, S. Ji, D. Shen, and J. Li, "A robust deep model for improved classification of AD/MCI patients," *IEEE J. Biomed. Health Informat.*, vol. 19, no. 5, pp. 1610–1616, Sep. 2015.
- [26] Y. Varatharajah, V. K. Ramanan, R. Iyer, and P. Vemuri, "Predicting short-term MCI-to-AD progression using imaging, CSF, genetic factors, cognitive resilience, and demographics," *Sci. Rep.*, vol. 9, no. 1, p. 2235, 2019.
- [27] J. J. Gomar, M. T. Bobes-Bascaran, C. Conejero-Goldberg, P. Davies, and T. E. Goldberg, "Utility of combinations of biomarkers, cognitive markers, and risk factors to predict conversion from mild cognitive impairment to Alzheimer disease in patients in the Alzheimer's disease neuroimaging initiative," *Arch. Gen. Psychiatry*, vol. 68, no. 9, pp. 961–969, Sep. 2011.
- [28] P. J. Moore, T. J. Lyons, J. Gallacher, "Random forest prediction of Alzheimer's disease using pairwise selection from time series data," *PLoS ONE*, vol. 14, no. 2, Feb. 2019, Art. no. e0211558.
- [29] T. Pereira, F. L. Ferreira, S. Cardoso, D. Silva, A. de Mendonça, M. Guerreiro, and S. C. Madeira, "Neuropsychological predictors of conversion from mild cognitive impairment to Alzheimer's disease: A feature selection ensemble combining stability and predictability," *BMC Med. Informat. Decis. Making*, vol. 18, no. 1, p. 137, Dec. 2018.
- [30] R. Craig-Schapiro, M. Kuhn, C. Xiong, E. H. Pickering, J. Liu, T. P. Misko, R. J. Perrin, K. R. Bales, H. Soares, A. M. Fagan, and D. M. Holtzman, "Demographic, clinical, and genotypic characteristics of the 333 study participants," *PLOS ONE*, vol. 6, no. 4, Apr. 2011, Art. no. e18850, doi: 10.1371/journal.pone.0018850.t001.
- [31] E. Konijnberg, B. M. Tijms, J. Gobom, V. Dobricic, I. Bos, S. Vos, M. Tsolaki, F. Verhey, J. Popp, P. Martinez-Lage, and R. Vandenberghe, "APOE ϵ 4 genotype-dependent cerebrospinal fluid proteomic signatures in Alzheimer's disease," *Alzheimer's Res. Therapy*, vol. 12, no. 1, p. 65, Dec. 2020.
- [32] H. Saddiki, A. Fayosse, E. Cognat, S. Sabia, S. Engelborghs, D. Wallon, P. Alexopoulos, K. Blennow, H. Zetterberg, L. Parnetti, and I. Zerr, "Age and the association between apolipoprotein e genotype and alzheimer disease: A cerebrospinal fluid biomarker-based case-control study," *PLOS Med.*, vol. 17, no. 8, Aug. 2020, Art. no. e1003289.
- [33] H. Sanz, C. Valim, E. Vegas, J. M. Oller, and F. Reverter, "SVM-RFE: Selection and visualization of the most relevant features through non-linear kernels," *BMC Bioinf.*, vol. 19, no. 1, p. 432, Dec. 2018.
- [34] V. Fonti, "Feature selection using LASSO," *VU Amsterdam Res. Bus. Anal.*, vol. 30, pp. 1–26, Mar. 2017.
- [35] O. Demir-Kavuk, M. Kamada, T. Akutsu, and E.-W. Knapp, "Prediction using step-wise l1, l2 regularization and feature selection for small data sets with large number of features," *BMC Bioinf.*, vol. 12, no. 1, p. 412, 2011.
- [36] D. A. Pisner and D. Schnyer, "Support vector machine," in *Machine Learning*, A. Mechelli and S. Vieira, Eds. Cambridge, MA, USA: Academic, 2020, ch. 6, pp. 101–121.
- [37] J. Tolles and W. J. Meurer, "Logistic regression: Relating patient characteristics to outcomes," *J. Amer. Med. Assoc.*, vol. 316, no. 5, pp. 533–534, Aug. 2016.
- [38] Y.-Y. Song and Y. Lu, "Decision tree methods: Applications for classification and prediction," *Shanghai Arch. Psychiatry*, vol. 27, no. 2, pp. 130–135, Apr. 2015.
- [39] G. Biau, "Analysis of a random forests model," *J. Mach. Learn. Res.*, vol. 13, pp. 1063–1095, Apr. 2012.
- [40] T. G. Dietterich, "Ensemble methods in machine learning," in *Multiple Classifier Systems—MCS (Lecture Notes in Computer Science)*, vol. 1857. Berlin, Germany: Springer, 2000, doi: 10.1007/3-540-45014-9_1.
- [41] R. Dinga, B. W. J. H. Penninx, D. J. Veltman, L. Schmaal, and A. F. Marquand, "Beyond accuracy: Measures for assessing machine learning models pitfalls and guidelines," 2019, doi: 10.1101/743138.
- [42] D. Zhao, T. Wang, and F. Chu, "Deep convolutional neural network based planet bearing fault classification," *Comput. Ind.*, vol. 107, pp. 59–66, May 2019.
- [43] T. A. Shifat and J.-W. Hur, "EEMD assisted supervised learning for the fault diagnosis of BLDC motor using vibration signal," *J. Mech. Sci. Technol.*, vol. 34, no. 10, pp. 3981–3990, Oct. 2020.
- [44] P. A. Flach and M. Kull, "Precision-recall-gain curves: PR analysis done right," in *Proc. Adv. Neural Inf. Process. Syst.*, Jan. 2015, pp. 838–846.

- [45] P. Mishra, U. Singh, C. M. Pandey, P. Mishra, and G. Pandey, "Application of student's t-test, analysis of variance, and covariance," *Ann. Cardiac Anaesthesia*, vol. 22, no. 4, pp. 407–411, 2019.
- [46] G. Kaur and E. Levy, "Cystatin C in Alzheimer's disease," *Frontiers Mol. Neurosci.*, vol. 5, p. 79, Jul. 2012.
- [47] P. M. Mathews and E. Levy, "Cystatin C in aging and in Alzheimer's disease," *Ageing Res. Rev.*, vol. 32, pp. 38–50, Dec. 2016.
- [48] P. Zipfel, C. Rochais, K. Baranger, S. Rivera, and P. Dallemagne, "Matrix metalloproteinases as new targets in Alzheimer's disease: Opportunities and challenges," *J. Medicinal Chem.*, vol. 63, no. 19, pp. 10705–10725, Oct. 2020.
- [49] F. H. Duits, M. Hernandez-Guillamon, J. Montaner, J. D. C. Goos, A. Montañaola, M. P. Wattjes, F. Barkhof, P. Scheltens, C. E. Teunissen, and W. M. van der Flier, "Matrix metalloproteinases in Alzheimer's disease and concurrent cerebral microbleeds," *J. Alzheimer's Disease*, vol. 48, no. 3, pp. 711–720, Oct. 2015.
- [50] C. D. Whelan, N. Mattsson, M. W. Nagle, S. Vijayaraghavan, C. Hyde, S. Janelidze, E. Stomrud, J. Lee, L. Fitz, T. A. Samad, G. Ramaswamy, R. A. Margolin, A. Malarstig, and O. Hansson, "Multiplex proteomics identifies novel CSF and plasma biomarkers of early Alzheimer's disease," *Acta Neuropathol. Commun.*, vol. 7, no. 1, p. 169, Dec. 2019.
- [51] D. S. Spellman, K. R. Wildsmith, L. A. Honigberg, M. Tuefferd, D. Baker, N. Raghavan, A. C. Nairn, P. Croteau, M. Schirm, R. Allard, J. Lamontagne, D. Chelsky, S. Hoffmann, W. Z. Potter, "Development and evaluation of a multiplexed mass spectrometry based assay for measuring candidate peptide biomarkers in Alzheimer's disease neuroimaging initiative (ADNI) CSF," *PROTEOMICS-Clin. Appl.*, vol. 9, nos. 7–8, pp. 715–731, Aug. 2015.
- [52] D. A. Llano, S. Bundela, R. A. Mudar, and V. Devanarayan, "A multivariate predictive modeling approach reveals a novel CSF peptide signature for both Alzheimer's disease state classification and for predicting future disease progression," *PLoS ONE*, vol. 12, no. 8, Aug. 2017, Art. no. e0182098.
- [53] J. Dukart, F. Sambataro, and A. Bertolino, "Accurate prediction of conversion to Alzheimer's disease using imaging, genetic, and neuropsychological biomarkers," *J. Alzheimer's Disease*, vol. 49, no. 4, pp. 1143–1159, Nov. 2015.



TABREJ KHAN was born in Bokaro, India, in 1981. He received the M.Sc. degree in computer science from Jamia Hamdard University, Delhi, India, in 2008.

In 2008, he started his career as a Software Developer with Sutraa Pvt., Ltd., Delhi, India. He is currently serving as a Lecturer for the Faculty of Computing and Information Technology, King Abdulaziz University, Rabigh, Saudi Arabia. He is also an Excellent Teacher and a Talented

Researcher with more than seven years of teaching and research experience in machine learning, bioinformatics, Web technology, and image processing. He has produced many publications in the *Journal of International Repute* and presented articles at International conferences. His current research interests include deep learning medical informatics and machine learning.

Mr. Khan is a member of the International Association of Engineers (IAENG) and a member of the following societies, The IAENG Society of Bioinformatics, The IAENG Society of Computer Science, and The IAENG Society of Data Mining.



ATIF HASSAN received the master's degree in computer science and engineering from the Indian Institute of Technology Kharagpur, where he is currently pursuing the Ph.D. degree from the Centre of Excellence in Artificial Intelligence. His research interests include natural language processing, data mining, machine learning, and deep learning in general. He loves sinking his teeth into new, real-world problems, and regularly contributes to the AI community through blogs,

research, and python package releases. He is also very active in ML competitions, securing top ten ranks on numerous occasions. In his free time, he works on new ideas and helps all those who reach out to him for guidance in the fields of ML, NLP, and DL.



ASIF HASSAN SYED was born in Jamshedpur, India, in 1982. He received the M.Sc. degree in biotechnology from the Shri Ramachandra Medical College and Deemed University, Chennai, India, in 2005, and the Ph.D. degree from the Indian Institute of Technology Roorkee, India.

He did his Postdoctoral Researcher under the esteemed guidance of Prof. S. E. Hasnain at the University of Hyderabad, from January 2012 to August 2012. In 2012, he started his career as an Assistant Professor with the Faculty of Computing and Information Technology at Rabigh (FCITR), King Abdulaziz University, Jeddah, Saudi Arabia, where he is currently serving as an Associate Professor. He is also an Excellent Teacher and a Talented Researcher with more than eight years of teaching and research experience in bioinformatics, chemoinformatics, genomics and proteomics, and machine learning. He has produced many publications in the *Journal of International Repute* and presented articles at International conferences. His current research interests include genomics and proteomics, medical informatics, and machine learning.

Dr. Syed is a member of the International Association of Engineers (IAENG) and a member of the following societies, The IAENG Society of Bioinformatics, The IAENG Society of Computer Science, and The IAENG Society of Data Mining. He was a recipient of the Deanship of Scientific Research Paper Award from King Abdulaziz University for the article *A qualitative and Quantitative Assay to study DNA/Drug Interaction based on sequence Selective Inhibition of Restriction Endonucleases*, in 2012. Availed DBT Postdoctoral Research Associateship in Biotechnology and Life Sciences 2011 (Under Prof. S. E. Hasnain) funded by DBT-IISC Bengaluru and Bagged the Louis Pasteur Memorial Award for the Best Outgoing Student for the year 2003 by the Department of Microbiology-The New College Chennai-Madras University, India.



NASHWAN A. ALROMEMA received the bachelor's degree in computer science from Zarka'a University, Jordan, in 2004, the M.Sc. degree in computer science from Jordan University, Jordan, in 2009, and the Ph.D. degree in computer science from Universiti Teknologi Malaysia, Malaysia, in 2017. He has worked in several places with different positions, including a Technical Project Team Leader, a Data Service Engineer, a Software Engineer, a Database Designer, and a Programmer.

In 2009, he started his academic career as a Lecturer with the Faculty of Computing and Information Technology-Rabigh (FCITR), King Abdulaziz University, Saudi Arabia, where he is currently serving as an Assistant Professor. He is also an Excellent Teacher and a Talented Researcher with more than 11 years of teaching and research experience in database, computer science, artificial intelligence, and machine learning. He has produced many publications in the *Journal of International Repute* and presented articles at International conferences. His main research interests include database analysis, design and modeling, temporal database models, temporal data mining, time-varying medical data, and image encryption using chaos encryption scheme. He has authored many research articles in data modeling of time-varying data and image encryption in international journals and conferences.



MUHAMMAD BINSAWAD received the master's degree in applied information technology and the post-baccalaureate certificate in information systems management from Towson University, USA, and the Ph.D. degree in information systems from the University of Technology Sydney (UTS), in 2019.

He has professional Information Systems Design, Digital Transformation, Business Analysis, IT Project Management, and Solid knowledge in the SDLC approach. He has taught several subjects in information systems and software engineering at King Abdulaziz University and the University of Technology in Sydney. He is currently an Assistant Professor with the Department of Computer Information Systems, Faculty of Computing and Information Technology, King Abdulaziz University. His research interests include and are not limited to information systems modeling-services, digital transformation human-computer interaction (HCI), and empirical studies. He is also actively involved in international research and events activities and contributing to international conferences and journals.



ALHUSEEN OMAR ALSAYED received the master's degree in information technology from the Faculty of Engineering and Information Technology, University of Technology Sydney, Sydney, NSW, Australia. He is currently working as a Scientific Researcher with the Deanship of Scientific Research, King Abdulaziz University, Jeddah, Saudi Arabia. He is also a certified Trainer and a Professional Practitioner from KAU and Technical and Vocational Training Corporation, Saudi

Arabia. His research interests include E-learning, cloud-based E-learning, collaborative learning, social networking sites, machine learning, and other related topics. He has published many articles in refereed/indexed international journals and conferences /MDPI/ IEEE ACCESS and Taylor & Francis Group. He has been appointed as a Reviewer of IEOM GCC Conference and IEEE. He is also actively involved in international research and events activities and contributing to international conferences.

• • •