Comparative Analysis of Machine Learning and Ensemble Learning Classifiers for Alzheimer's Disease Detection

Palak Goyal
Department of Computer Science and Engineering
Thapar Institute of Engineering and Technology
Patiala, India
pgoyal60_phd18@thapar.edu

Rinkle Rani
Department of Computer Science and Engineering
Thapar Institute of Engineering and Technology
Patiala, India
raggarwal@thapar.edu

Karamjeet Singh
Department of Computer Science and Engineering
Thapar Institute of Engineering and Technology
Patiala, India
karamjeet.singh@thapar.edu

2022 IEEE International Conference on Current Development in Engineering and Technology (CCET) | 978-1-6654-5415-5/22/$31.00 ©2022 IEEE

***Abstract*— Alzheimer's disease (AD), a psychiatric problem, availing between those who are 65 and older. Additionally, the disease's steady development of a variety of visible and invisible symptoms, such as irritability and aggression, has a substantial negative impact on a patient's overall quality of life. Although many treatments have been developed to help reduce its symptoms, AD has no known cure. As a result, the field of AD management is growing, and a comprehensive framework for the early detection of AD must be created. In this study, we created three classification models for predicting AD using machine learning and five models for predicting AD using ensemble learning. SVM, DTs, and RF are the three basis classifiers employed in the current work. Five ensemble classifiers, XGBoost, Voting Classifier, Extra Trees (ETs) Classifier, Gradient Boost, and AdaBoost, are then thoroughly compared. After thoroughly inspecting the dataset for outliers or other noise, a feature selection method known as PCA and several preprocessing techniques are used to lessen the issue of overfitting and performance enhancement. Additionally, this study utilised the longitudinal information from the OASIS website, which included 150 patients overall, 72 of whom were not demented and 78 of whom were. The RF model, which had an accuracy of 83.92% compared to the other two base classifiers, provided the best classification performance, while the ETs Classifier, an ensemble classifier, performed the best when compared to base and ensemble classifiers, with an accuracy of 86.60%.**

Keywords*— Alzheimer's Disease, Ensembling, Machine Learning Classifiers, Principal Component Analysis

I. INTRODUCTION

Two-thirds of people with dementia have AD, a degenerative neurological disorder [1]. As the population ages, it is predicted that this number will increase. For illustration, by 2050, it is predicted that those 60 and older would make up 20% of the population in India suffering with this disease [2]. Moreover, there will be 13.8 million Americans by the year 2050 that are affected with a projected prevalence of AD. [3].

AD's cause is unclear at this time. Aging, apoE 4 gene subtypes, beta-amyloids, gender, head injuries, educational attainment, tension and genetic factors are some of the risk factors for AD that are not seen in the early stages but gets worsen with time [4]. Numerous characteristics and criteria, such as genetic data, neuropsychological assessments, cerebrospinal fluid biomarkers, white matter volume, gray matter volume and brain imaging data, are used to make the diagnosis of AD [5]. The correct preventive treatments be utilized in the early stages of AD in order to accurately limit or stop the illness's progression. Further, AD has not yet been reliably detected by established clinical testing, despite efforts to develop accessible and affordable AD biomarkers. [6]. Further, various tests are used by doctors like the Clinical Dementia Rating (CDR), Mini-Mental State Examination (MMSE), and Estimated Total Intracranial Volume (Etiv) to make the initial diagnosis of AD [7]. Rigid cooperation and a significant amount of time are needed in order to gather all the relevant information [8]. Due to these reasons, X-ray scans, MRI, and computer tomography (CT) are the tests with the greatest accuracy. However, because these tests are extremely pricey, they are rarely used as a preventive measure to identify the condition [8]. Finding and implementing less expensive solutions becomes desirable as a result.

The most crucial finding is that AD patients are less able to recognise their symptoms in the early stages of the disease as compared to those in healthy control (HC) groups [9]. Moreover, communication allows people to share thoughts, facts, and feelings. This process is impeded in cases of Alzheimer disease since this disorder also has a negative impact on cognitive functions, including memory and reasoning ability [10]. Additionally, communication issues have an effect on the social wellbeing of AD patients. Because of the issues with identification of initial symptoms, the clinical characteristics and demographic properties plays a significant part in the prognosis and categorization of AD and the different stages of the disease.

Using longitudinal data from AD and HC patients, this work developed and fully analyzed three ML-based and five EL-based classification models for AD predictions. In this research, three foundation classifiers—RF, DTs, and SVMs well as five hybrid classifiers named XGBoost, Gradient Boost, Extra Trees, AdaBoost, and Voting Classifier are used. Furthermore, a variety of preprocessing techniques, including feature scaling, dataset balance, and the feature selection approach known as PCA, are used to lessen the issue of overfitting and performance enhancement.

Further paper is organized as follows: The literature in this field is detailed in Section II. The workings of the suggested model are shown in Section III. The model's performance evaluation is shown in Section IV. Section V ends the current study by offering conclusions and suggestions for further research.

II. LITERATURE REVIEW

On the basis of a cursory study, some of the numerous issues and challenges surrounding AD have been discussed here.

The ADNI-GO/2 research with PET scans was used by [11] for SVM and regularizeed logistic regression. Further, another study [12] used a variety of ML techniques, including RF, NB, and J48, taking into account a preliminary investigation and demonstrates that J48 is the least effective of all the algorithms.

Some of the studies in this field uses the features that are extracted from MR images. For instance, MRI images from the OASIS dataset were used by [13] for SVM classification using the Gabor Filter, GLCM and ICA features, resulting in extremely high classification accuracy, precision, and recall.

The categorization of AD using MRI analysis has been studied using a variety of machine learning techniques. SVM, CNN and other ensemble classifiers are among the frequently employed classifiers [14]. Since they can handle high-dimensional data and have a reasonable level of accuracy, SVM and its variants have been the most extensively employed among them [14-16].As seen, it can be concluded from above studies that the categorization of MCI and AD participants has been predicted using a variety of machine learning techniques on various combinations of characteristics, including CSF volumes, genetics, images and clinical data. However, only a small number of studies have explored for other biomarkers that can improve the precision of the most widely used CSF protein biomarkers at the moment in classifying CN, MCI and AD [17].

In comparison to individual or base machine learning classifiers, an EL based classifier may offer good results to predict the outcomes. To enhance AD prediction, a study [18]. integrate three separate classifiers using weighted and unweighted approaches. They employ the 11C-PIB PET imaging data, although further thought can be given to the variety of basic classifiers. Thus, the base classifiers may be dependent on one another. Recent studies have explored how to improve prediction accuracy by fusing deep learning and ensemble learning systems [19]. Ensembling has recently become more significant in separating cognitively healthy people from a progressive form of MCI that eventually results in AD [20-22].

For instance, in the study [20], RFE method was employed to find incremented CSF biomarkers in the brain and then the weighted ensemble learning of two classifiers named SVM and Logistic Regression was implemented to classify 4 stages of AD resulting in area under precision-recall curve (AUPR) as 0.91 respectively. In accordance with the studies of the various researches, it can be concluded that the CAD systems can recognize Alzheimer in the early stages. The neural network architecture may therefore be used, for example, in clinical contexts, to enhance AD detection. The quantity of information needed to train machine learning algorithms, however, continues to be one of the fundamental shortcomings of the current approaches. These elements cause the analysis for Alzheimer's detection to differ substantially. In this study, we examined various EL-based and machine learning classifiers for categorizing AD based on different assessment parameters like accuracy and precision.

III. PROPOSED METHOLOGY

The process used in suggested framework is drafted in Fig. 1. It explains how to complete the evaluation task and divide it into smaller tasks. Below is an explanation of each step:

Phase 1: Data Collection: The dataset used for the study was taken from the Open Access Series of Imaging Studies (OASIS) website (https://www.oasis-brains.org/). This set includes longitudinal data of 150 people, separated into two classes: AD and normal patients, having age in range from 60 to 96 years. In addition, the AD class has 78 patients, while the Normal class comprises 72 people. Also, a total of 373 imaging samples were there with each patient scanned over the course of two visits with a gap of a year. Table I provides a summary of the data used in this investigation, along with information such as the number of patients in each class, and the gender distribution, etc.

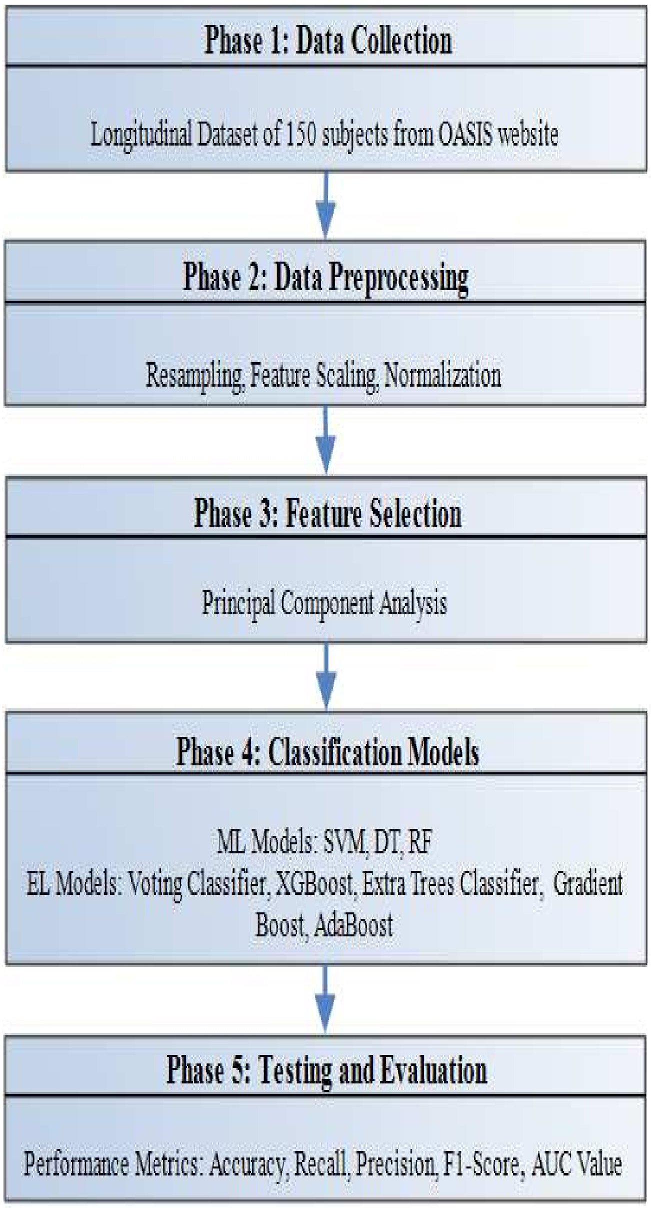

Fig. 1. Proposed Approach

TABLE I. DATASET DESCRIPTION

Characteristic	AD Patients	Normal Patients
Count	78	72
Male/Female	40/38	22/50
Age (Mean ± S.D.)	74.34 ± 5.97	79.68± 9.31

Phase 2: Data Preprocessing: To handle the issue of various problems existing in the dataset such as unbalanced classes, we oversampled the data using the resampling approach, which involves handling instances for the minority classes. As a result, the Normal class is subjected to oversampling utilizing this idea. After being divided into two groups of 130 samples each, the training set has a total of 261 samples, and the testing set has 112 samples. Following the above approach, another method named data normalization is then incorporated to ensure that no characteristic is over-emphasizeed.

Phase 3: Feature Selection: This stage involves choosing characterized features from the full dataset. The study's dataset has a lot of characteristics, which hinders the classifier's performance and introduces overfitting into the dataset. In order to select the final 8 features, PCA is used along with Pearson coefficients.

Phase 4: Classification Models: Using This step involves the classification task which do the binary classification of Normal and AD patients. The current work achieved this by utilising numerous ML and EL classifiers, which are described in more detail below.

ML Based Classifiers: ML play a crucial role for classification, clustering, and other tasks. This idea was applied in this study to categorize AD patients. SVM, DT, and hyper tuned RF are the classifiers employed in this work. Comparing the effectiveness of these classifiers, it was shown that RF outperformed ML classifiers.

EL Based Classifiers: Various methods integrate learning components in EL based models to provide more precise overall forecasts with a specific purpose. The models' performance is classed and contrasted in this study using EL classifiers namely XGBoost, Gradient Boost, Voting Classifier, Extra Trees Classifier, and AdaBoost. As a result, Extra Trees classifier, an ensemble model, performs better than both ML and EL classifiers.

Phase 5: Testing and Evaluation: In this phase involves testing, comparison and analysis of the aforementioned models on the basis of various performance evaluation metrics including recall, accuracy, and precision etc.

IV. RESULTS AND DISCUSSION

Jupyter Notebook was used throughout the study's experimentation, which was conducted on the Anaconda platform. In addition, the classifiers' performance is measured using the following metrics:

Accuracy: It is calculated as follows:

$$Acc = \frac{t_p + t_n}{t_p + t_n + f_n + f_p} \tag{1}$$

Where, variables represent true positives, true negatives, false positives and false negatives respectively.

Precision: It can be framed as the portion of results to a query and recharacterized as follows:

$$Precision = \frac{t_p}{t_p + f_p} \tag{2}$$

Recall: It fits the following description that is successfully retrieved data relevant to a query:

$$Recall = \frac{t_p}{t_p + f_n} \tag{3}$$

F1-Score: It is defined as follows:

$$F1 - Score = \frac{2 * Precision * Recall}{Precision + Recall} \tag{4}$$

AUC Value: It is a measure that tells how good our model is in performing the classification task. This value is calculated by drawing finding the area under ROC curve.

Data must be treated properly to eliminate noise and improve and analyze classifier performance using the aforementioned metrics. First, data analysis is required to determine which pre-processing methods should be used to remove noise from the data. For instance, the plots of CDR vs gender as well as age are displayed in Fig. 2 and Fig. 3. In order to choose the top features that are pertinent for the problem in the next steps, the heatmap, as depicted in Fig. 4, is also generated to indicate the importance of features that are vital in connection to the target value.

Following that, box plots of numeric features are plotted to determine normalization of features. As seen in Fig. 5, the features exhibit high variability, which causes some feature values to become outliers and necessitates the use of feature scaling method, for which, the usual scalar feature scaling method is used in this research..

After pre-processing the data, 8 features are selected based on PCA feature extraction method. This final dataset is then used to evaluate the performance of discussed classifiers. Table II. and 3 display the results for a variety of evaluation measures, including Accuracy, Precision, and others, for ML and EL classifiers.

Fig. 2. Plot of CDR vs gender

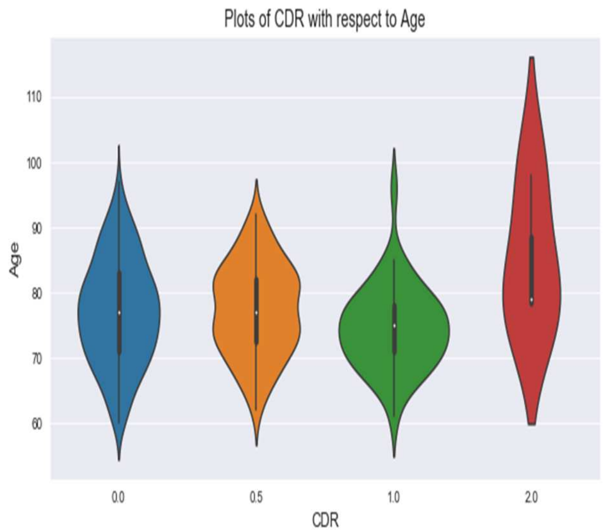

Fig. 3. Plot of CDR vs age

Fig. 4. Pearson Correlation Heatmap

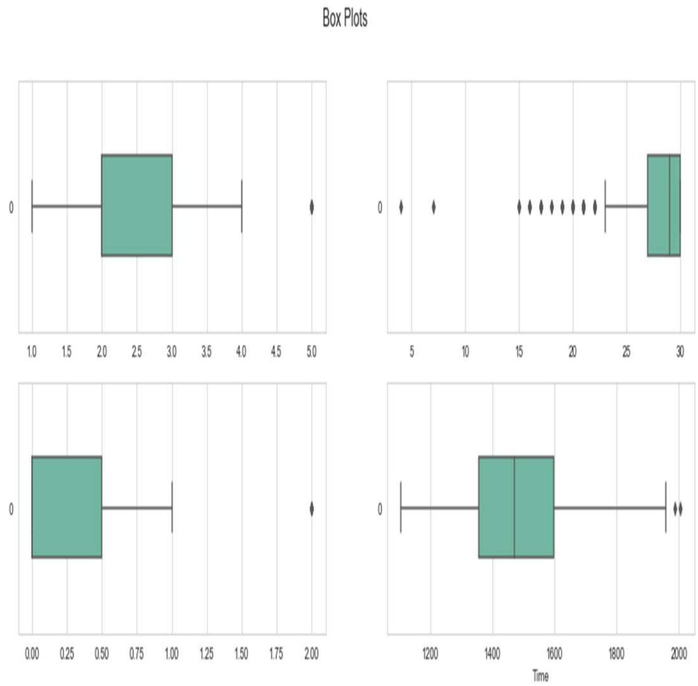

Fig. 5. Box plots for features: SES (Top-Left), MMSE (Top-Right), CDR (Bottom-Left), Etiv (Bottom-Right)

As depicted in Table II and Table III, the best performance produced by Random Forest, an ML-based classifier, and Extra Trees, an EL-based classifier. Figures 6 and 7 display the ROC curves and confusion matrices for the pioneering works. Furthermore, Figures 8 and 9 shows the various plots for all the models. The figures show that EL-based classifiers perform better than ML-based classifiers.

TABLE II. COMPARISON OF THE AD-CN CLASSIFICATION PERFORMANCE OF MACHINE LEARNING CLASSIFIERS

ML Classifiers	Accuracy (%)	Recall (%)	Precision (%)	F1-Score (%)	AUC (%)
Support Vector Machine	77.6	78	79	78	79.0
Decision Trees	79.46	80	80	79	80.0
Random Forest	83.92	84.0	84.0	84.0	84.0

TABLE III. COMPARISON OF AD-CN CLASSIFICATION PERFORMANCE FOF ENSEMBLE LEARNING CLASSIFIERS

EL Model	Accuracy (%)	Recall (%)	Precision (%)	F1-Score (%)	AUC (%)
Voting Classifier	83.9	84.0	84.0	84.0	84.0
XGBoost	83.9	84.0	84.0	84.0	84.0
Gradient Boost	83.03	83.0	83.0	83.0	83.0
ETs Classifier	86.6	87.0	87.0	87.0	87.0
AdaBoost	80.35	80.0	80.0	80.0	80.0

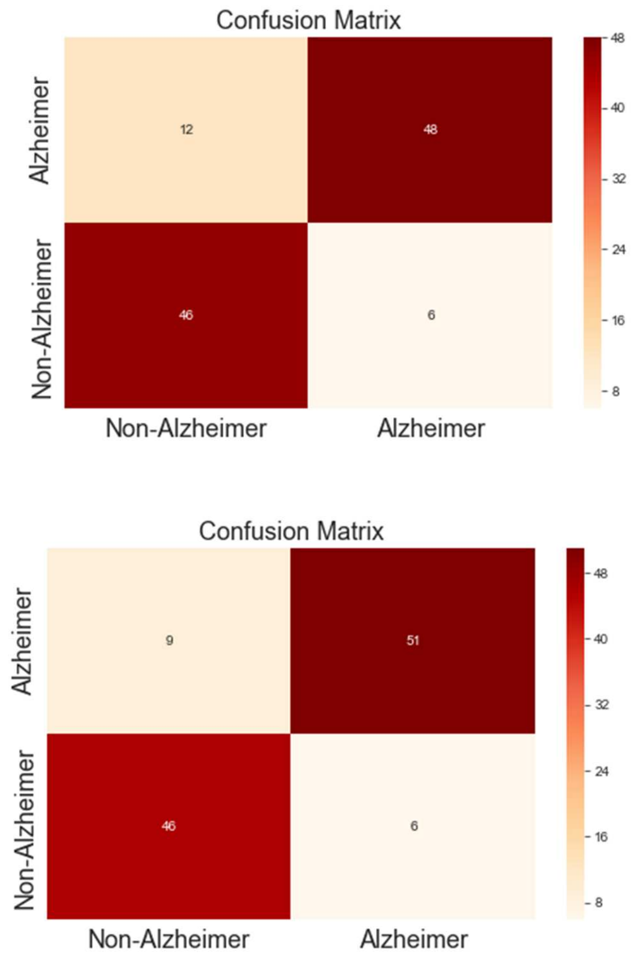

Fig. 6. Confusion Matrix for AD-CN classification: Random Forest (Top), Extra Trees Classifier (Bottom)

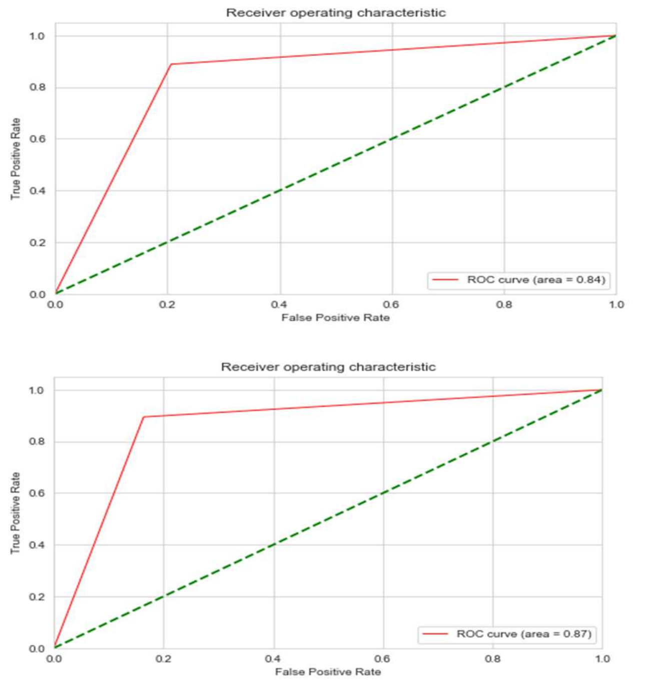

Fig. 7. AD-CN ROC Curves: Random Forest (Top), Extra Trees Classifier (Bottom)

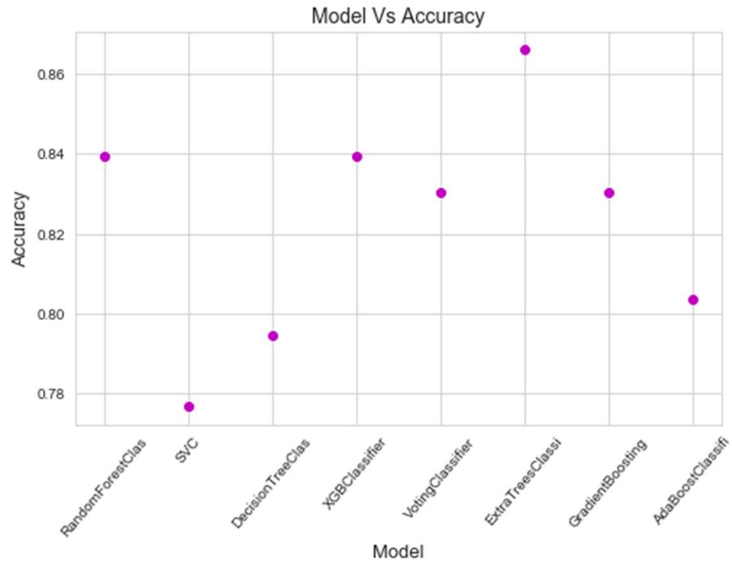

Fig. 8. Accuracy Values Graph of ML and EL Classifiers

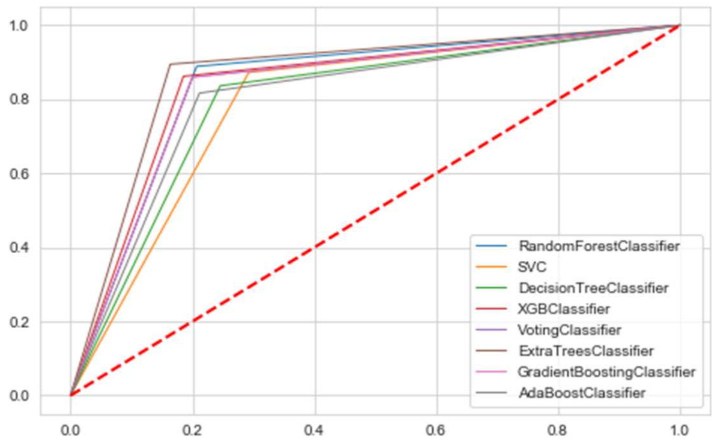

Fig. 9. ROC Curves of ML and EL Classifiers

V. CONCLUSION

AD is an immedicable brain illness affecting primarily the elderly. Despite the fact that AD currently has no viable treatments. The affected individuals' pace of development would be slowed while their recuperation would be accelerated. Recent achievements in the field of medicine using ML and EL approaches have been amazing, and when applied to Alzheimer's disease, a promising accuracy may be reached. The effectiveness of various ML and EL-based classifiers to predict AD is compared in this research. Several preprocessing approaches including a feature selection method known as PCA are used to enhance classifier performance even more and to lessen the issue of overfitting. Extra Trees Classifier outperformed Random Forest in terms of accuracy, coming up at 86.60% as opposed to 83.92% for the best ML model.. The method in this study can be highly helpful if utilised to spot the disease in its early stages.

The classification models used in this study has only been tested for AD dataset, but it can also be applied to other clinical research fields. Further, the performance of the classifiers to predict AD may potentially be assessed and improved in the future using a variety of deep learning techniques.

REFERENCES

[1] J Neelaveni, and MS Geetha Devasana, "Alzheimer disease prediction using machine learning algorithms." In 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), pp. 101-104. IEEE, 2020.

[2] T. Subetha, Rashmita Khilar, and Sarat Kumar Sahoo, "An Early Prediction and Detection of Alzheimer's Disease: A Comparative Analysis on Various Assistive Technologies." In 2020 International Conference on Computational Intelligence for Smart Power System and Sustainable Energy (CISPSSE), pp. 1-5. IEEE, 2020.

[3] Mingxia Liu, Zhang Daoqiang, Adeli Ehsan, and Shen Dinggang, "Inherent structure-based multiview learning with multitemplate feature representation for alzheimer's disease diagnosis." IEEE Transactions on Biomedical Engineering, vol. 63, no. 7 pp:1473-1482, 2015.

[4] Jeet Sukumaran, P. Economo Evan, and L. Lacey Knowles, "Machine learning biogeographic processes from biotic patterns: a new trait-dependent dispersal and diversification model with model choice by simulation-trained discriminant analysis." Systematic Biology, vol. 65, no. 3, pp. 525-545, 2016.

[5] Darko Zibar, Piels Molly, Jones Rasmus, and G. Schaeffer Christian, "Machine learning techniques in optical communication." Journal of Lightwave Technology, vol. 34, no. 6, pp: 1442-1452, 2015.

[6] A. Michael DeTure, and W. Dickson Dennis, "The neuropathological diagnosis of Alzheimer's disease." Molecular neurodegeneration, vol. 14, no. 1 pp:1-18, 2019.

[7] X Zhu, HI Suk, and S Dinggang, "A novel matrix-similarity based loss function for joint regression and classification in AD diagnosis." NeuroImage, vol. 100, pp: 91-105, 2014.

[8] S Kaur, H Aggarwal and R Rani, "Diagnosis of Parkinson's disease using deep CNN with transfer learning and data augmentation", Multimedia Tools and Applications, vol. 80, no. 7, pp. 10113-10139, 2021.

[9] M Liu, D Zhang, E Adeli, and D Shen, "Inherent structure-based multiview learning with multitemplate feature representation for alzheimer's disease diagnosis", IEEE Transactions on Biomedical Engineering, vol. 63, no. 7, pp. 1473-1482, 2015.

[10] S Basheera, and MSS Ram, "Convolution neural network–based Alzheimer's disease classification using hybrid enhanced independent component analysis based segmented gray matter of T2 weighted magnetic resonance imaging with clinical valuation", Alzheimer's & Dementia: Translational Research & Clinical Interventions, vol. 5, pp. 974-986, 2019 .

[11] S Mathotaarachchi, TA Pascoal, M Shin, AL Benedet, MS Kang, T Beaudry, and Alzheimer's Disease Neuroimaging Initiative, "Identifying incipient dementia individuals using machine learning and amyloid imaging", Neurobiology of Aging, vol. 59:, pp. 80-90, 2017.

[12] SIR Bhagyashree, K Nagaraj, M Prince, CH Fall, and M Krishna, "Diagnosis of Dementia by Machine learning methods in Epidemiological studies: a pilot exploratory study from south India", Social Psychiatry and Psychiatric Epidemiology, vol. 53, no. 1, pp: 77-86, 2018.

[13] SK Aruna and S Chitra, "Machine Learning Approach for Identifying Dementia from MRI Images". World Academy of Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering, vol. 9, no. 3, pp: 881-888, 2016.

[14] S Alam, GR Kwon, JI Kim, and CS Park, "Twin SVM-based classification of Alzheimer's disease using complex dual-tree wavelet principal coefficients and LDA", Journal of healthcare engineering, 2017.

[15] RS Azevedo, JR de Sousa, MT Araujo et al, "In situ immune response and mechanisms of cell damage in central nervous system of fatal cases microcephaly by Zika virus", Scientific reports, vol. 8, no. 1, pp: 1-11, 2018.

[16] X Long, L Chen, C Jiang, L Zhang, and Alzheimer's Disease Neuroimaging Initiative, "Prediction and classification of Alzheimer disease based on quantification of MRI deformation", PloS one, vol. 12, no. 3, pp. e0173372, 2017.

[17] JE Park, KY Choi, BC Kim, KH Lee et al., "Cerebrospinal fluid biomarkers for the diagnosis of prodromal Alzheimer's disease in amnestic mild cognitive impairment", Dementia and geriatric cognitive disorders extra, vol. 9, no. 1, pp: 100-113, 2019.

[18] W Wu, J Venugopalan, and MD Wang, "11C-PIB PET image analysis for Alzheimer's diagnosis using weighted voting ensembles", In 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 3914-3917, IEEE, 2017.

[19] X Qiu, L Zhang, Y Ren, PN Suganthan, and G Amaratunga, "Ensemble deep learning for regression and time series forecasting", In 2014 IEEE symposium on computational intelligence in ensemble learning (CIEL), pp. 1-6, IEEE, 2014.

[20] AH Syed, T Khan, A Hassan, NA Alromema, M Binsawad, and AO Alsayed, "An ensemble-learning based application to predict the earlier stages of Alzheimer's disease (AD)", IEEE Access, vol. 8, pp. 222126-222143, 2020.

[21] M Grassi, N Rouleaux, D Caldirola, D Loewenstein, K Schruers, G Perna, G., and Alzheimer's 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 in neurology, vol. 10, p. 756, 2019.

[22] T Pereira, FL Ferreira, S Cardoso, D Silva, A Mendonca, M Guerreiro, and SC Madeira, "Neuropsychological predictors of conversion from mild cognitive impairment to Alzheimer's disease: a feature selection ensemble combining stability and predictability", BMC medical informatics and decision making, vol. 18, no. 1, pp: 1-20, 2018.