



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Special Issue 1, April 2024

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.423

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com

Diagnosis And Classification of Alzheimer's Disease Using Machine Learning Algorithm

Dr. S Xavier Arockiaraj¹, S Janani², S Keerthana Ambikai³, V Keethana⁴, S Mahalakshmi⁵

Associate Professor, Department of Electronics and Communication Engineering, Adhiyamaan College of Engineering, Krishnagiri, Tamil Nadu, India¹

U.G Scholars, Department of Electronics and Communication Engineering, Adhiyamaan College of Engineering, Krishnagiri, Tamil Nadu, India^{2,3,4,5}

ABSTRACT: Alzheimer's disease is a brain disease that first impairs thinking and memory abilities before eventually destroying the capacity to perform even the most basic tasks. Using machine learning techniques, the primary goal of this effort is to diagnose Alzheimer's disease early. The metrics included in the OASIS (Open Access Series of Imaging Studies) longitudinal dataset include Normalized Whole Brain Volume (nWBV), Mini Mental State Examination (MMSE), Clinical Dementia Rating (CDR), Atlas Scaling Factor (ASF), and others. Considering the multitude of factors involved, persons with Alzheimer's disease have a higher chance of surviving if the disease is discovered later. Experts frequently encounter difficulties in assessing each patient, which leads to the use of machine learning techniques. Using methods like Ensemble RUS Boost Tree Classifier, Decision Tree, Logistic Regression, and Naïve Bayes, these algorithms identify pertinent features, preprocess the data, then train. Ultimately, each model's performance indicators are assessed. The performance measures for testing and validation consist of Accuracy, True Positive Rate (TPR), Positive Predictive Value (PPV), Area Under the ROC Curve (AUC), and Confusion Matrix (CM). Determining whether a patient is at risk of Alzheimer's disease is their goal. In comparison to all other machine learning algorithms used, the suggested study achieves an excellent accuracy of 94.6% on Alzheimer's disease test data using the Ensemble RUS Boost Tree classifier. Compared to other cutting-edge machine learning algorithms published in the literature, the suggested method performs better.

KEYWORDS: Alzheimer's Disease, Confusion Matrix, Machine Learning, OASIS Longitudinal Dataset.

I. INTRODUCTION

Alzheimer's disease (AD) poses a significant challenge to healthcare systems worldwide, with its progressive nature severely impacting cognitive function and quality of life. The absence of a cure underscores the importance of early detection for effective management and intervention. Our study focuses on harnessing the power of machine learning algorithms applied to the OASIS Longitudinal dataset, a valuable resource providing insights into both normal aging and AD progression.

Our main objective is to use machine learning and the OASIS longitudinal dataset to precisely build tools for early identification of Alzheimer's disease (AD). By enhancing diagnosis accuracy and implementing individualised patient care plans, we hope to completely revolutionise the healthcare sector. We hope that our work will have a major influence on the area of Alzheimer's disease (AD) identification and, in the process help to shape new approaches to treating this intricate neurological condition. By sharing our approach, findings, and conclusions we hope to draw attention to the potential of machine learning in facilitating early diagnosis and individualised therapy for individuals with Alzheimer's disease.

Our goal is to transform the area of Alzheimer's disease detection and therapy by using machine learning in our research.

To accomplish this, aim our study on Alzheimer's disease detection is essential since it addresses the urgent need for better approaches to controlling this complex disorder. We strive to uncover novel perspectives about the diagnosis and management of Alzheimer's disease. Our goal is to demonstrate through our study how machine learning can be an effective tool for early identification and providing individualised treatment to people with AD. Our goal is to provide cutting-edge solutions that increase diagnostic precision and improve patient outcomes to bring about significant change in the healthcare sector. Through collaborative efforts and knowledge sharing, our goal is to create a future in which individualization, efficacy, and precision define the diagnosis and treatment of Alzheimer's disease.

The paper is structured as follows. A summary of previous studies on machine learning-based Alzheimer's disease prediction is given in Section II emphasising the value of early detection and the contribution of feature engineering to classification accuracy. The study's methodology is described in Section III, with particular attention on how feature selection and machine learning methods were used to predict Alzheimer's disease using the OASIS Longitudinal dataset. The suggested approach, which uses an Ensemble RUS boost tree classifier for early detection, is presented in Section IV. Additionally, an intuitive user interface has been developed to assist researchers and doctors in interpreting data. Section VI discusses the experimental findings, which include a confusion matrix for the Ensemble RUS Boosted Tree model and a clinical features analysis using Power BI. Finally, conclusions and future directions are presented in Section VII and VIII.

II. RELATED WORKS

The increasing impairment of cognitive function and memory associated with Alzheimer's disease presents substantial obstacles. Effective treatment and intervention depend on a timely and precise diagnosis. To help with personalised treatment regimens, machine learning (ML) approaches have become important tools in the identification and prediction of AD. While some authors, such as Scheltens, Blennow, and Breteler (2016), stress the vital necessity for early identification to reduce its impact, Kavitha et al. (2022) emphasise the usefulness of machine learning methods in anticipating AD early [1].

Feature engineering improves classification accuracy and helps with illness categorization, as shown by Aamir et al. (2016) and Maroco et al. (2011) [4][5]. Salehi et al. (2020) and Zhang et al. (2018) investigate how machine learning approaches may be used to accurately diagnose patients using non-imaging data [6][7]. A multi-diagnostic method is used by Vasco Sá Diogo et al. (2022) to contribute to AD prediction, while Nemali et al. (2022) provide an insightful evaluation of Gaussian Process-based prediction models [8][9]. Our work focuses on using machine learning (ML) techniques to detect AD early. We have achieved an impressive accuracy of 94.6% with the Ensemble RUS Boosted Tree classifier, demonstrating the potential of ML-based approaches to improve diagnostic precision and enable proactive healthcare interventions in AD management.

III. EXISTING METHOD

Alzheimer's Disease is predicted using ML algorithms by using a feature selection and extraction technique, and the classification is conducted based on the oasis longitudinal dataset. Late detection reduces chances of survival and poses challenges for clinicians to evaluate each patient accurately. Some existing machine learning approaches reduce importance of ongoing research and collaboration in improving diagnostic accuracy and patient care. Current approaches to diagnose and monitor Alzheimer's disease (AD) involve the analysis of a variety of data sources, such as genetic information, medical records, and patient journals by using machine learning techniques. These techniques use a range of machine learning models, including unsupervised clustering algorithms and supervised classifiers, to extract pertinent information from the data.

To guarantee the precision and dependability of the models, a critical component of these techniques is the meticulous selection and creation of features. Integrating patient journals provides insightful subjective data, frequently using natural language processing techniques to draw conclusions. These systems' main objectives are to support medical professionals in the diagnosis of Alzheimer's disease, tracking the course of the illness, and formulating treatment plans.

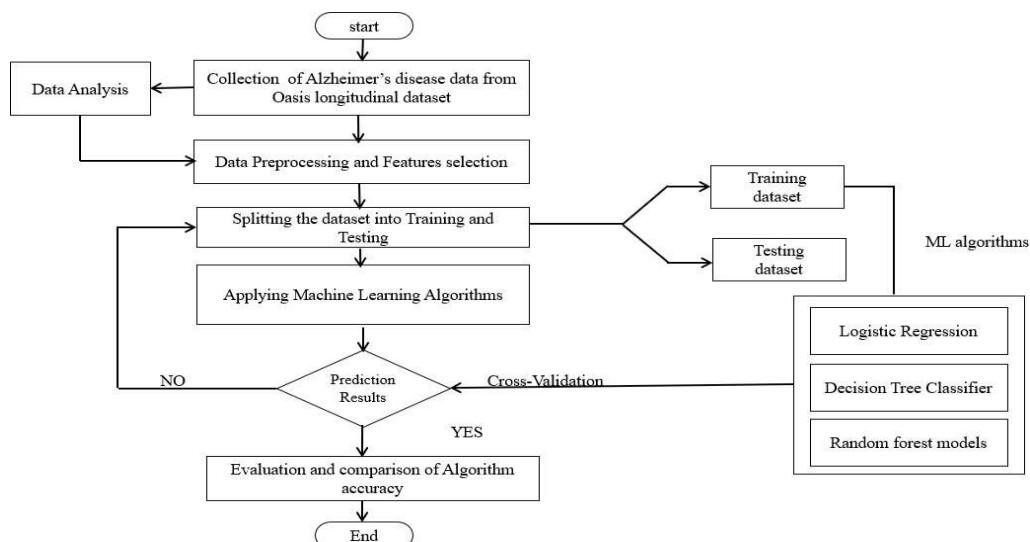
To protect patient data and guarantee compliance with legal obligations, however, the design and deployment of such systems must take ethical issues and privacy norms into account.

IV. PROPOSED SYSTEM

The solution under consideration involves the methodical acquisition of data from the OASIS Longitudinal dataset, with stringent quality assurance protocols in place to guarantee precision and consistency. To prepare the dataset for analysis, data preparation techniques such as imputation of missing values, error correction, and duplication removal will be used. These actions are essential for guaranteeing the consistency and dependability of the data, which paves the way for strong machine learning models. Without doing these first actions, the analysis might yield unreliable data and draw erroneous conclusions.

Sophisticated feature selection techniques that are specific to the properties of the OASIS Longitudinal dataset will be employed in our suggested solution. We will use methods like mutual information, ensemble learning, and significance testing to figure out which factors are most important for predicting the start and course of Alzheimer's disease. The suggested technique seeks to improve the interpretability and performance of prediction models by choosing useful characteristics while reducing noise and overfitting.

Using the supplied dataset, the proposed system would train and assess a variety of machine learning models, including as decision trees, ensemble RUS boosted models, naive Bayes, and logistic regression. To determine how well these models identify Alzheimer's disease, they will be rigorously evaluated using performance measures including True Positive Rate (TPR), Positive Predictive Value (PPV), and Area Under the ROC Curve (AUC). As the main model for Alzheimer's disease diagnosis, the ensemble RUS enhanced model-which has shown to be more accurate in early assessments-will be further refined and assessed. The Ensemble RUS Boosted Tree classifier emerges as the most effective model, achieving an exceptional accuracy of **94.6%** on Alzheimer's disease test data.



The suggested system would perform extensive data analysis to get insights into the clinical features and course of Alzheimer's disease after model training and evaluation. The dataset will be examined for patterns and correlations using visualisation tools like Power BI. This will help identify possible biomarkers and risk factors linked to Alzheimer's disease (AD). The proposed approach intends to contribute to a better understanding of Alzheimer's disease pathology and improve individualised treatment options by illuminating the underlying patterns in the data.

The performance of the ensemble RUS enhanced model will be compared to other machine learning techniques that are often employed in Alzheimer's disease diagnosis using the suggested approach. The efficiency of the suggested technique in precisely identifying Alzheimer's disease will be confirmed by a comparative study of accuracy percentages. To demonstrate the benefits of the ensemble technique in enhancing diagnostic accuracy and reliability, the findings will be given and discussed. To sum up, the suggested system, which makes use of machine learning methods and the rich insights provided by the OASIS Longitudinal dataset, offers a thorough approach to Alzheimer's disease diagnosis and management. Through improving diagnostic precision and deepening our understanding of Alzheimer's disease

pathogenesis, the suggested method hopes to aid in the creation of more potent approaches to Alzheimer's disease identification and individualised patient management.

V. BLOCK DIAGRAM

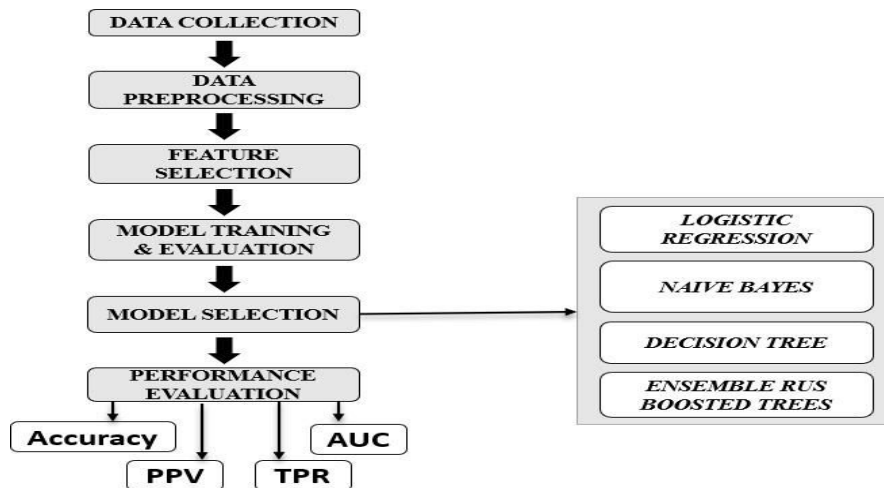


Figure 1: Block diagram for proposed method

This methodology diagram (Figure 1) outlines a structured approach it starts with data collection, preprocessing, and feature selection, followed by model training and evaluation with logistic regression, naive Bayes, decision tree and ensemble RUS boosted models. Performance metrics are then utilized to assess model effectiveness, providing a comprehensive framework for analysis.

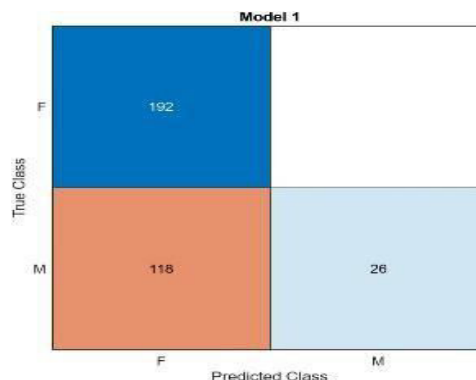


Figure 2: Flowchart for proposed method

Using the OASIS longitudinal dataset, the suggested technique outlines a methodical procedure for the diagnosis of Alzheimer's disease (AD). First, data analysis entails gathering as much information as possible on Alzheimer's disease from the OASIS longitudinal dataset. This dataset includes vital parameters that are necessary to comprehend how the illness progresses. Preprocessing and feature selection are carried out after data collection to guarantee data integrity and find relevant features essential for Alzheimer's disease diagnosis. To make model training and assessment easier, the dataset is then split into training and testing sets. Various machine learning (ML) methods are used to identify patterns and relationships in the data during the training phase. These algorithms include random forest models, decision tree classifiers, logistic regression, and ensemble RUS boosted trees. Based on the chosen characteristics, these algorithms are then used to predict the diagnosis of Alzheimer's disease. In order to identify the most successful method for Alzheimer's disease diagnosis, the prediction results are assessed, and the accuracy of each algorithm is compared in the final step. With significant therapeutic implications, this systematic approach advances our understanding of Alzheimer's disease and establishes a foundation for accurate and reliable diagnostic models.

VI. EXPERIMENTAL RESULTS

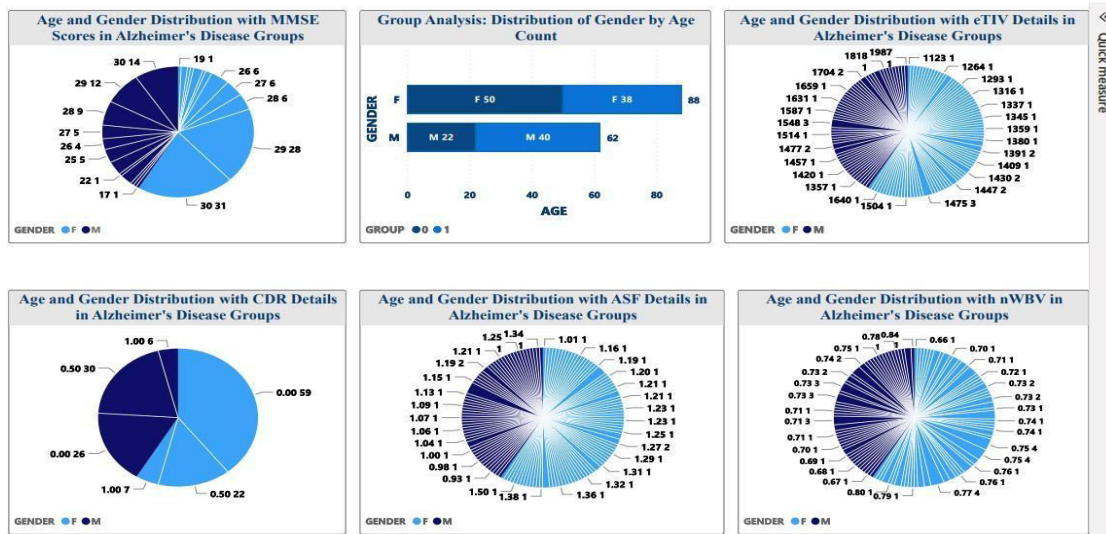


Figure 3: Clinical Characteristics Analysis in Alzheimer's Disease Groups using Power BI

Figure 3 illustrates the data analysis process for comparing clinical features between Alzheimer's disease (AD) groups using Power BI. We looked for trends by using Power BI's interactive capabilities to examine a range of clinical parameters. Preprocessing, visualisation creation, and in-depth analysis were done to understand the characteristics of the Alzheimer's disease subgroup. Power BI simplified comprehensive clinical feature analysis and aided in the identification of critical variables linked to the severity of Alzheimer's disease.

Figure 4 In order to assess the ensemble RUS boosted tree method's performance, we looked closely at the confusion matrix and related performance metrics in this study. The algorithm's classification results were clearly shown visually in the confusion matrix, where rows denoted the actual class (male and female) and columns the expected class. Analysing the matrix showed that 192 examples were correctly categorised as female and 118 as male by the computer. There was some degree of misclassification in the model, though, since it also misclassified 26 instances of females as men.

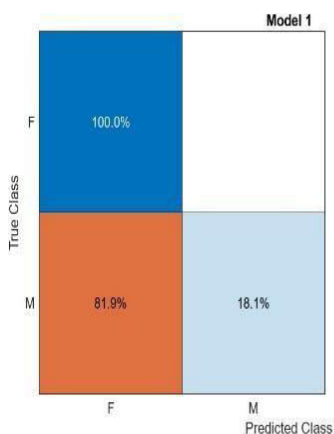


Figure 5: Sensitivity

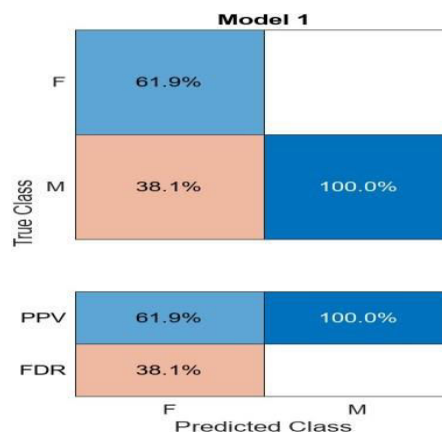
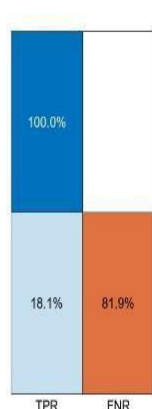


Figure 6: Performance Metrics

Figure 6 displays the performance metrics of the Ensemble RUS Boosted Tree technique. The false discovery rate (FDR) and positive predictive value (PPV) with respect to gender categories are of special interest. Whereas the PPV displays the proportion of correctly diagnosed instances within each gender group, the FDR displays the ratio of false positive predictions. Conversely, Figure 5 illustrates the algorithm's sensitivity, which is also known as the false negative rate (FNR) and true positive rate (TPR). These numbers give insight into how likely the algorithm is to miss positive examples and how accurate it is in finding positive instances, respectively.

VII. FUTURE SCOPE

Further studies aim to enhance the prompt identification and treatment of Alzheimer's disease (AD) by combining the use of proficient imaging methodologies, the integration of genetic and biomarker data, the creation of customised risk assessment models, and the examination of longitudinal data. These models could offer customised riskevaluations and suggestions for treatments aimed at preventing disease. To ensure its efficacy, validation in a variety of populations is essential. Real-time risk assessments, treatment suggestions, and patient monitoring might be offered via clinical decision support systems. Collaboration amongst researchers, physicians, politicians, and industry stakeholders is critical for improving the discipline. Ethical and privacy issues are also critical. We can create more individualised, precise, and approachable strategies to fight Alzheimer's disease (AD) by tackling important issues and investigating new prospects.

VIII. CONCLUSION






In conclusion, we used the OASIS longitudinal dataset to apply machine learning techniques for early detection of Alzheimer's disease in patients. On test data pertaining to Alzheimer's disease, we found that the Ensemble RUS Boosted Tree classifier performed more accurately than previous models, with 94.6% accuracy [9]. The programme had strong predictive abilities, particularly in identifying male occurrences, but it had trouble accurately predicting female occurrences. Our research highlights the role that machine learning plays in early illness detection, which paves the way for timely treatment and improved patient outcomes. We improve our data visualisation with Power BI.

REFERENCES

1. Kavitha et al., "Early-Stage Alzheimer's Disease Prediction Using Machine Learning Models," Frontiers in Public Health, 2022.
2. Available: <https://www.frontiersin.org/journals/public-health/articles/10.3389/fpubh.2022.853294/full>

3. C. Saraiva et al., "Prevention of Alzheimer's disease and dementia," Journal of Alzheimer's Disease, vol. 92, no.1–2, pp. 98-104, 2007.
4. M. Aamir et al., "Computer Aided Diagnosis of Alzheimer's Disease Using Features Selected Based on Particle Swarm Optimization Algorithm," Journal of Medical Imaging and Health Informatics, 2016.
5. P. Scheltens, K. Blennow, et al., "Alzheimer's disease," The Lancet, vol. 388, no. 10043, pp. 505-517, July 2016. Available: <https://pubmed.ncbi.nlm.nih.gov/26921134/>
6. G. Castellazzi, M. G. Cuzzoni, et al., "Machine Learning Approach for the Differential Diagnosis of Alzheimer's Disease," Journal of Alzheimer's Disease, 2011.
7. Available: <https://www.frontiersin.org/articles/10.3389/fninf.2020.00025/full>
8. S. Zhang et al., "Alzheimer's Disease Detection Based on Demographic Features Using Machine Learning Techniques," Brain Informatics, 2018.
9. V. Sá Diogo et al., "Early diagnosis of Alzheimer's disease using machine learning: a multi-diagnostic, generalizable approach," Alzheimer's Disease Neuroimaging Initiative, 2022.
10. Available: <https://pubmed.ncbi.nlm.nih.gov/35922851/>
11. Nemali et al., "Gaussian Process-based prediction of memory performance and biomarker status in ageing and Alzheimer's disease - A systematic model evaluation," Neuroimage, 2022.
12. Aniverthy Amrutesh, Gowtham Bhat C G, Amruthamsh A, Asha Rani KP, Gowrishankar. "Alzheimer's Disease Prediction using Machine Learning and Transfer Learning Models."
13. Hyun Seok Choi, "Detecting Early Alzheimer's," Kaggle.
14. Available: <https://www.kaggle.com/code/hyunseokc/detecting-early-alzheimer-s>

BIOGRAPHY

	Dr. S Xavier Arockiaraj, Associate Professor, Department of Electronics and Communication Engineering Adhiyamaan college of Engineering, Hosur
	Keerthana Ambikai. S, Department of Electronics and Communication Engineering , Adhiyamaan college of Engineering, Hosur.
	Janani. S, Department of Electronics and Communication Engineering, Adhiyamaan college of Engineering, Hosur.
	Keerthana. V, Department of Electronics and Communication Engineering, Adhiyamaan college of Engineering, Hosur.
	Mahalakshmi. S, Department of Electronics and Communication Engineering, Adhiyamaan college of Engineering, Hosur.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details