International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Diagnosis Of Alzheimer Disease Earlier Using SVM Prediction Models

[1] Navaneethan G, [2] Sridhar M, [3] Sujith S, [4] Vinothkumar D, [5] Sindhuri P

[1,2,3,4] Student, CSE Department, Dhirajlal Gandhi College of Technology Salem, India

[1]navaneethanjun10@gmail.com, [2]sridharcseb@gmail.com, [3]ssujith016@gmail.com, [4]vinothslm2003@gmail.com,

[5] Assistant Professor, CSE Department, Dhirajlal Gandhi College of Technology Salem, India

[5]Sindhuri.cse@dgct.ac.in

ABSTRACT:

Alzheimer's disease (AD) is a neurodegenerative condition and the leading cause of dementia in older individuals. The portion of brain that gets influenced in this illness is hippocampus degeneration . Early detection of Alzheimer's is crucial as it can help prevent significant brain damage. Particularly in individuals aged 65 and above, the disease can pose serious risks, sometimes even proving fatal. The primary aim of this study is to utilize machine learning techniques, specifically Random Forest and feature extraction and selection, to predict Alzheimer's disease develop an effective model. The dataset consists of MRI images, enabling the analysis of brain structures. The dataset comprises MRI scans, and the proposed method employs the Random Forest algorithm to classify different stages of Alzheimer's, such as moderate-demented and non-demented.

Keywords: Alzheimer's disease, Neurodegenerative condition, Hippocampus degeneration, Early detection, Machine learning techniques, Random Forest algorithm.

1. INTRODUCTION

Alzheimer's infection is the most common cause of Dementia in individuals of the age 65 a long time and over. It is a dynamic and irreversible neurological infection which takes after a unmistakable design of brain harm as the infection advances. Alzheimer's infection is a exceptionally common sort of dementia. Dementia is an umbrella term depicting a assortment of illnesses and conditions that create when nerve cells in the brain (called neurons) kick the bucket or no longer work in a typical way. The passing or breakdown of neurons causes anomalies in one's memory, conduct, and capacity to think in a clear way. In Alzheimer's infection, these brain conditions in the long run disable an individual's capacity to perform indeed essential capacities such as strolling, talking, and gulping. Improvement of Advertisement can be classified into three stages. To begin with, is the asymptomatic arrange, changes in the brain, blood, or cerebrospinal liquid (CSF) may start to happen without the understanding appearing any specific indications. After the to begin with organize at that point the moment arrange, that is gentle cognitive disability (MCI) arrange, memory complaints and other cognitive conduct may begin to be recognizable for the patients themselves and for near family or companions, which influences day to day exercises but the side effects are gentle. In the last arrange of the malady, or the dementia organize, memory, considering, and behavioral side effects are apparent and noteworthy, and it is recognizable. The neurons of brain begin worsening and the neural connections are gradually broken down. The quiet loses the capacity to react to the environment.

This research will focus on developing an evolving framework to effectively diagnosis and predict AD at a very early stage using the data collected for AD patients. The framework will continuously use large sets of related data to AD patient collected from multiple sources like medical sources, lifestyle and demography. The datasets are taken in the form of MRI images. The detection of Alzheimer's disease using conventional technique is time consuming, so we apply machine learning technique CNN to predict the Alzheimer's disease.

1.1 Existing System

Various statistical methods have been applied to differentiate Alzheimer's disease (AD) from healthy controls (HC). Early and accurate classification of AD and mild cognitive impairment (MCI) is crucial for timely preventive measures. Magnetic resonance imaging (MRI) offers non-invasive insights into brain structure changes. A new method proposes using dual-tree complex wavelet transforms (DTCWT), principal coefficients from MRI slices, and linear discriminant analysis, alongside twin support vector machine. This approach aims to improve the distinction between AD and HC individuals.

1.2 Proposed System

Detecting Alzheimer’s disease in its prodromal stage is crucial to prevent significant brain damage. Traditionally, medical professionals use CT scans for detection, but this process is time-consuming. To address this, we are leveraging machine learning technology to predict Alzheimer’s disease. Our proposed approach employs a CNN algorithm to detect various stages of Alzheimer’s, such as moderate-demented and non-demented, reducing prediction time and enabling real-time predictions. We describe the datasets used and the data preprocessing steps undertaken, including feature extraction through principal component analysis and feature selection techniques. Following preprocessing, we apply the efficient CNN algorithm to predict the disease and categorize it into moderate demented and non-demented stages. Additionally, feature selection aids in analyzing the impact of analytical tests, potentially improving Alzheimer's disease prediction.

2. REVIEW OF LITERATURE

The literature survey on Alzheimer's disease covers various aspects, including early prediction, class imbalance handling, and classification techniques. Researchers aim to identify mild cognitive impairment (MCI) as a precursor to Alzheimer's, using MRI-based biomarkers and machine learning frameworks, achieving notable results such as a 10-fold cross-validation area under the curve (AUC) score of 0.9020 in discriminating between progressive MCI (pMCI) and stable MCI (sMCI). Techniques employed include semi-supervised learning, feature selection with logistic regression, and creating aggregate biomarkers with random forest classifiers. The review also delves into strategies for handling class imbalance, discussing ensembles, K nearest neighbor classifiers, and decision trees, alongside novel algorithms like Jiang et al.'s Genetic algorithm to improve SMOTE. Furthermore, studies emphasize the significance of combining various biomarkers from neuroimaging and neuropsychological tests for accurate Alzheimer's classification. Logistic regression models are highlighted for their utility in differentiating MCI and AD, with a focus on predicting disease progression, while also discussing different types of logistic regression suited for various target variable categories. Overall, the survey underscores the multidimensional approach required for understanding and addressing Alzheimer's disease.

3. SYSTEM IMPLEMENTATION

3.1 Removing Missing Values

Running Random Forest on data with missing values could yield inaccurate error rates and potentially result in overfitting. To address this, missing values are eliminated from the dataset using the complete cases() function

3.2 Creating A Train and Test Sets

Two random samples of data were taken from the original dataset: a training set comprising 80% of the data (5170 wines), and a test set comprising 20% (1293 wines). The training set was used to construct the model, while the test set was reserved to evaluate the model's performance on data that was not involved in its construction.

3.3 Building the Model Using The Train Set

Based on the provided output, the type of random forest employed is classification, as the outcome variable "disease" is categorical. The model utilizes 500 decision trees by default and employs 3 variables (the square root of the total number of variables) for splitting.

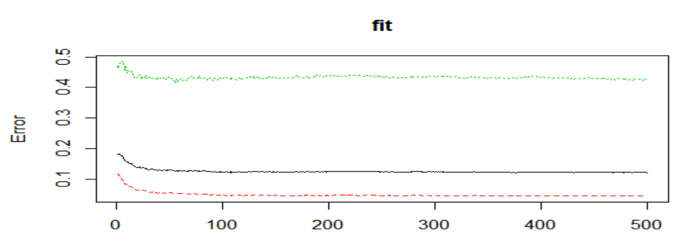

Fig 1 Model Train Set

3.4 Testing the Model Using A Test Set

This step involves evaluating the model's performance using the test data. According to the confusion matrix provided, 1016 wines are classified as having low disease, while 138 are categorized as having high disease. The model achieves an accuracy of 89.25%, indicating that 10.75% of wines are predicted incorrectly (101 + 38). Specifically, if a customer selects a wine categorized as high disease, there is only a 2.9% chance (38/(1016+101+38+138)) that it is actually a low-disease wine. The Random Forest model demonstrates superior accuracy in predicting high-disease wines compared to a single decision tree model.

3.5 Identifying Important Variables to Check While Choosing A Wine

The plots depict the variables crucial for identifying high-disease Alzheimer, ranked from most to least important. On the right, alcohol emerges as pivotal, with a mean decrease in Gini exceeding 300, followed by residual sugar. The left graph illustrates the model's poorer performance without variables of significant mean decrease accuracy. Removing alcohol notably reduces the random forest's accuracy, underlining its importance.

4. SYSTEM DESIGN

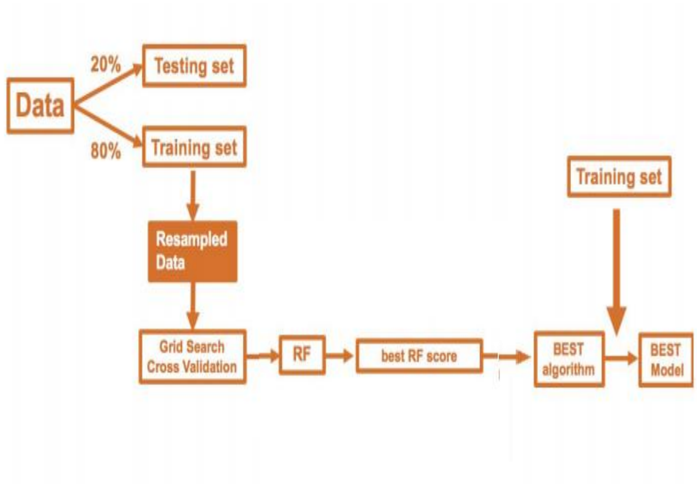

Fig 2 System Architecture

5. CONCLUSION

This study compares and evaluates various machine learning algorithms, alongside pre-processing and Boruta feature selection, for the prognosis and prediction of Alzheimer's disease (AD). Missing values in SES and MMSE columns were addressed by deleting corresponding columns. Experimental results affirm that combining pre-processing with Boruta feature selection yields a reliable technique for early AD prediction. Random Forest (RF) with GridSearchCV (GSCV) and Boruta algorithm emerges as the top-performing classifier, achieving an accuracy of 94.39%. The RF model demonstrates superior feature propagation and classification results, even with a small dataset. The developed GUI prediction tool offers dementia status predictions based on patient demographics, clinical data, and MRI-derived anatomic volume features, addressing limitations highlighted in previous research.

6. FUTURE ENHANCEMENT

The key takeaway is the identification of various factors associated with dementia, highlighting the importance of ongoing monitoring and exploring diverse approaches for intervention. Further research should focus on enhancing our understanding through advanced exploratory data analysis (EDA) with a larger sample size. This may involve not only considering age itself but also grouping it into generations, evaluating brain tissue volume grading, or analyzing exam scores for a more comprehensive investigation.

REFERENCES:

[1] M. Prince, R. Bryce, E. Albanese, A. Wimo, W. Ribeiro, and C. P. Ferri, 'The global prevalence of dementia: A systematic review and metaanalysis', Alzheimer's and Dementia. 2013.

[2] C. Dillon, C. M. Serrano, D. Castro, P. P. Leguizamón, S. L. Heisecke, and F. E. Taragano, 'Behavioral symptoms related to cognitive impairment.', Neuropsychiatr. Dis. Treat., vol. 9, pp. 1443–55, 2013.

[3] A. V. Lebedev et al., 'Random Forest ensembles for detection and prediction of Alzheimer's disease with a good between-cohort robustness', NeuroImageClin., vol. 6, pp. 115–125, 2014.

[4] A. Khan and M. Usman, 'Early Diagnosis of Alzheimer's Disease using Machine Learning Techniques', 2015 7th Int. Jt. Conf. Knowl. Discov.Knowl.Eng. Knowl.Manag. (IC3K)., 2015.

[5] S. Teipel et al., 'Multimodal imaging in Alzheimer's disease: Validity and usefulness for early detection', The Lancet Neurology. 2015.

[6] M. B. Kursa and W. R. Rudnicki, 'Feature Selection with the Boruta Package', J. Stat. Softw., 2015.

[7] T. M. Schouten et al., 'Combining anatomical, diffusion, and resting state functional magnetic International Conference on Biomedical Engineering (ICoBE) Journal of Physics: Conference Series 1372 (2019) 012065 IOP Publishing doi:10.1088/1742-6596/1372/1/012065 8 resonance imaging for individual classification of mild and moderate Alzheimer's disease', NeuroImageClin., 2016.

[8] S. Sarraf, G. Tofighi, and for the A. D. N. Initiativ, 'DeepAD: Alzheimer′s Disease Classification via Deep Convolutional Neural Networks using MRI and fMRI', bioRxiv, 2016.

[9] N. Amoroso et al., 'Deep learning reveals Alzheimer's disease onset in MCI subjects: Results from an international challenge', J. Neurosci. Methods, 2018.

[10] G. Awate, S. Bangare, G. Pradeepini, and S. Patil, 'Detection of Alzheimers Disease from MRI using Convolutional Neural Network with Tensorflow', Jun. 2018.