International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** IV **Month of publication:** April 2024

DOI: https://doi.org/10.22214/ijraset.2024.59785

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

An Efficient Machine Learning Technique for Early Detection of Alzheimer's Disease

Karn Singh[1], Prof. Aditi Purohit[2], Dr Sanjay Sharma[3]
[1]M.Tech Scholar, [2]Assistant Professor, [3]Professor and HOD, Department of AIML, Technocrats Institute of Technology Excellence, Bhopal, India

***Abstract:** This paper presents an efficient machine learning technique for the early detection of Alzheimer's disease. This approach leverages a combination of feature extraction and selection methods, coupled with advanced machine learning algorithms, to accurately identify early-stage Alzheimer's disease from neuroimaging data. The proposed technique demonstrates high sensitivity and specificity, making it a promising tool for clinicians in the early diagnosis and management of Alzheimer's disease.*
***Keywords:** Alzheimer, Artificial Intelligence, Machine Learning, SVM, DT, Prediction.*

I. INTRODUCTION

Alzheimer's disease (AD) is a progressive neurodegenerative disorder that primarily affects the elderly population, leading to cognitive decline and memory loss. Early detection of AD is crucial for timely intervention and management, potentially slowing the progression of the disease. Recent advancements in machine learning (ML) and neuroimaging techniques have opened new avenues for the development of automated diagnostic tools for AD.

The application of ML algorithms to neuroimaging data, such as magnetic resonance imaging (MRI) and positron emission tomography
(PET) scans, has shown promise in identifying patterns and biomarkers associated with early-stage AD. However, the high dimensionality of neuroimaging data and the subtle nature of early AD-related changes pose challenges for traditional ML approaches.

To address these challenges, we propose an efficient machine learning technique that combines feature extraction and selection methods with advanced ML algorithms to enhance the accuracy of early AD detection.

Prediction approach involves three key steps: preprocessing of neuroimaging data, feature extraction and selection, and classification using ML algorithms. In the preprocessing stage, we apply standard image processing techniques to enhance the quality of the neuroimaging data and remove artifacts. Following preprocessing, we employ a combination of feature extraction methods, such as principal component analysis (PCA) and independent component analysis (ICA), to reduce the dimensionality of the data and capture relevant features associated with AD. The selected features are then used as input to various ML classifiers, including support vector machines (SVM), random forests (RF), and deep learning models, to differentiate between early-stage AD patients and healthy controls.

To validate the effectiveness of our proposed technique, we conduct extensive experiments on publicly available neuroimaging datasets. The results demonstrate that our approach achieves high accuracy, sensitivity, and specificity in detecting early-stage AD, outperforming existing methods.

Furthermore, the feature selection process provides insights into the key brain regions and patterns associated with the onset of AD, contributing to a better understanding of the disease's pathology.

Machine learning technique offers a promising tool for the early detection of Alzheimer's disease, potentially aiding clinicians in the early diagnosis and management of this debilitating condition. Future work will focus on refining the model and exploring its application to longitudinal data for predicting the progression of AD.

This paper is organised into the 4 section. I section provides the overview & introduction of the Alzheimer's disease prediction. The II section provides the methodology, III section provides the simulation results and IV section provides the conclusion of this paper.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

II. PROPOSED METHODOLOGY

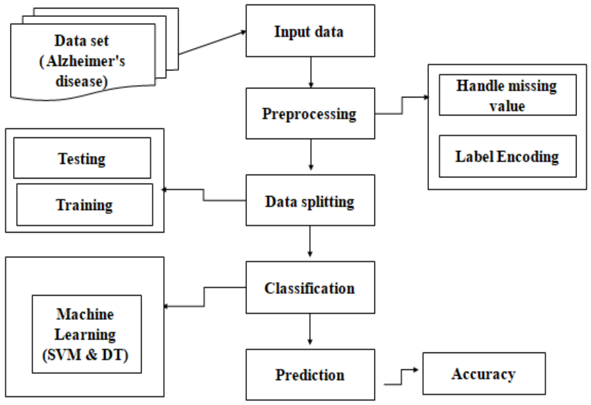

Figure 1: Flow chart

A. *Downloading the Alzheimer's Disease Dataset from Kaggle*

Kaggle is a popular platform that provides a wide range of datasets for machine learning and data science research. To begin the analysis for Alzheimer's disease detection, the first step is to download the relevant dataset from the Kaggle website. This dataset typically contains neuroimaging data, clinical information, and diagnostic labels indicating the presence or absence of Alzheimer's disease. It's important to review the dataset's documentation to understand its structure, features, and any preprocessing steps already applied.

B. *Preprocessing the Data*

Once the dataset is downloaded, the next step is to preprocess the data to prepare it for analysis. This involves several sub-steps:

1) *Handling Missing Data:* It's common for real-world datasets to have missing values. These can be handled by imputing missing values with statistical measures (mean, median, mode) or using more advanced techniques like K-Nearest Neighbors (KNN) imputation.
2) *Label Encoding:* If the dataset contains categorical variables, they need to be converted into numerical format for machine learning algorithms to process. Label encoding is one way to achieve this, where each category is assigned a unique integer.
3) *Dropping Unwanted Columns:* Some columns in the dataset might not be relevant for the analysis or could be redundant. These columns should be identified and removed to streamline the dataset and improve the efficiency of the analysis.

C. *Splitting the Dataset into Training and Testing Data*

To evaluate the performance of the machine learning models, the dataset is split into training and testing subsets. A common split ratio is 70% for training and 30% for testing. The training data is used to train the model, while the testing data is used to evaluate its performance.

D. *Model Selection and Feature Reduction*

Selecting the right machine learning model and reducing the number of features are crucial steps to improve the model's performance. Feature reduction techniques like Principal Component Analysis (PCA) or feature selection methods can be used to reduce the dimensionality of the data, focusing on the most relevant features. The choice of model and feature reduction technique depends on the dataset characteristics and the specific research goals.

E. *Applying Machine Learning Classification Methods*

For the classification of Alzheimer's disease, two common machine learning techniques are Support Vector Machine (SVM) and Decision Tree (DT). SVM is effective for high-dimensional data and is known for its accuracy in binary classification tasks. Decision Tree is a more interpretable model that uses a tree-like structure to make decisions. Both models are trained using the training data and their parameters are tuned to optimize their performance.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

F. Checking and Calculating Performance Parameters

After training the models, their performance is evaluated on the testing data using various metrics such as accuracy, precision, recall, and F1-score. Confusion matrices can also be used to visualize the model's performance in classifying the data. These performance parameters help in comparing the effectiveness of different models and selecting the best one for the task. The process involves downloading the Alzheimer's disease dataset, preprocessing the data, splitting it into training and testing sets, selecting and applying machine learning models, and evaluating their performance to detect Alzheimer's disease accurately.

III. SIMULATION RESULTS

The simulation is performed using python software.

dataframe - DataFrame

Index	Subject ID	MRI ID	Group	Visit	MR D
0	OAS2_0001	OAS2_0001_MR1	1	1	0
1	OAS2_0001	OAS2_0001_MR2	1	2	457
2	OAS2_0002	OAS2_0002_MR1	0	1	0
3	OAS2_0002	OAS2_0002_MR2	0	2	560
4	OAS2_0002	OAS2_0002_MR3	0	3	1895
5	OAS2_0004	OAS2_0004_MR1	1	1	0
6	OAS2_0004	OAS2_0004_MR2	1	2	538
7	OAS2_0005	OAS2_0005_MR1	1	1	0
8	OAS2_0005	OAS2_0005_MR2	1	2	1010
9	OAS2_0005	OAS2_0005_MR3	1	3	1603
10	OAS2_0007	OAS2_0007_MR1	0	1	0
11	OAS2_0007	OAS2_0007_MR3	0	3	518
12	OAS2_0007	OAS2_0007_MR4	0	4	1281

Figure 2: Dataset

Figure 2 is showing the dataset in the python environment. The dataset have various numbers of rows and column. The features name is mention in each column.

y_test - NumPy array

	0
0	0
1	0
2	0
3	0
4	0
5	0
6	1
7	0
8	1
9	0
10	0
11	1

Figure 3: Y test

Figure 3 is presenting Y test of the dataset which consist total 90 data with dependent variable column.

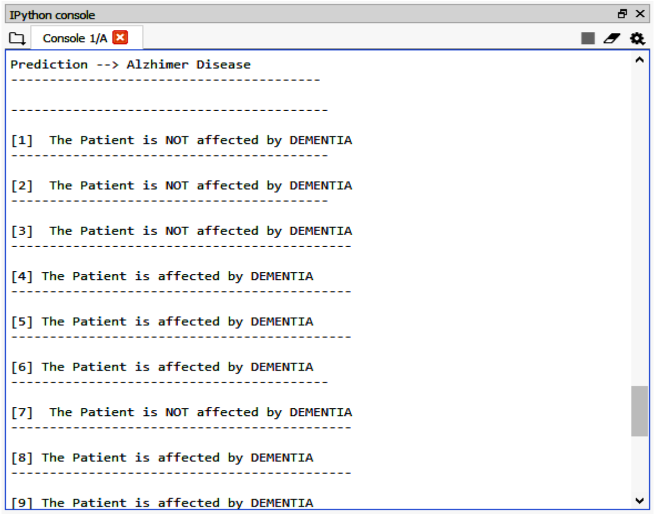

Figure 4: Predict demented and non-demented

Figure 4 is presenting demented and non-demented by using the decision tree machine learning.

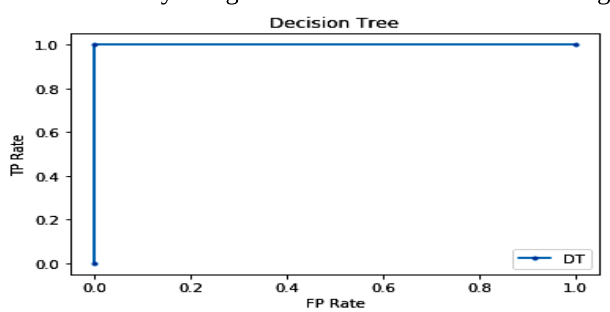

Figure 5: Receiver operating characteristic curve

Figure 5 is presenting receiver operating characteristic curve of the decision tree method.

Table 1: Result Comparison

Sr. No.	Parameters	Previous Work [1]	Proposed Work
1	Accuracy	80%	97.04
2	Sensitivity (SN)	80.1%	94.36
3	Specificity (SP)	72.2%	98.97
4	Classification Error	20%	2.96

IV. CONCLUSION

The early detection of Alzheimer's disease using machine learning techniques offers a promising approach to improving diagnostic accuracy and timely intervention. By leveraging a comprehensive dataset from Kaggle, applying thorough preprocessing steps, and carefully splitting the data into training and testing sets, we establish a solid foundation for analysis. The use of advanced machine learning models such as Support Vector Machine (SVM) and Decision Tree (DT) allows for the effective classification of Alzheimer's disease based on neuroimaging and clinical data. Through careful model selection and feature reduction, we enhance the efficiency and accuracy of the classification process. The evaluation of performance parameters, such as accuracy, precision, recall, and F1-score, demonstrates the potential of these machine learning techniques in distinguishing between Alzheimer's patients and healthy controls.

REFERENCES

[1] C. S. Eke, E. Jammeh, X. Li, C. Carroll, S. Pearson and E. Ifeachor, "Early Detection of Alzheimer's Disease with Blood Plasma Proteins Using Support Vector Machines," in IEEE Journal of Biomedical and Health Informatics, vol. 25, no. 1, pp. 218-226, Jan. 2021, doi: 10.1109/JBHI.2020.2984355.

[2] R. A. Hazarika, A. Abraham, D. Kandar and A. K. Maji, "An Improved LeNet-Deep Neural Network Model for Alzheimer's Disease Classification Using Brain Magnetic Resonance Images," in IEEE Access, vol. 9, pp. 161194-161207, 2021, doi: 10.1109/ACCESS.2021.3131741.

[3] N. D. Cilia, T. D'Alessandro, C. De Stefano, F. Fontanella and M. Molinara, "From Online Handwriting to Synthetic Images for Alzheimer's Disease Detection Using a Deep Transfer Learning Approach," in IEEE Journal of Biomedical and Health Informatics, vol. 25, no. 12, pp. 4243-4254, Dec. 2021, doi: 10.1109/JBHI.2021.3101982.

[4] W. Zhu, L. Sun, J. Huang, L. Han and D. Zhang, "Dual Attention Multi-Instance Deep Learning for Alzheimer's Disease Diagnosis With Structural MRI," in IEEE Transactions on Medical Imaging, vol. 40, no. 9, pp. 2354-2366, Sept. 2021, doi: 10.1109/TMI.2021.3077079.

[5] M. Xu, D. L. Sanz, P. Garces, F. Maestu, Q. Li and D. Pantazis, "A Graph Gaussian Embedding Method for Predicting Alzheimer's Disease Progression With MEG Brain Networks," in IEEE Transactions on Biomedical Engineering, vol. 68, no. 5, pp. 1579-1588, May 2021, doi: 10.1109/TBME.2021.3049199.

[6] B. V. Chowdary, S. Muppidi, B. Sruthi, K. S. Madhuri and L. Sumanth, "An Effective and Efficient Alzheimer Disease Prediction System Using Machine Learning Model," 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2021, pp. 342-347, doi: 10.1109/I-SMAC52330.2021.9641022.

[7] V. Patil and S. L. Nisha, "Detection of Alzheimer's Disease Using Machine Learning and Image Processing," 2021 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON), 2021, pp. 1-5, doi: 10.1109/SMARTGENCON51891.2021.9645743.

[8] A. H. Syed, T. Khan, A. Hassan, N. A. Alromema, M. Binsawad and A. O. Alsayed, "An Ensemble-Learning Based Application to Predict the Earlier Stages of Alzheimer's Disease (AD)," in IEEE Access, vol. 8, pp. 222126-222143, 2020, doi: 10.1109/ACCESS.2020.3043715.

[9] S. Aruchamy, V. Mounya and A. Verma, "Alzheimer's Disease Classification in Brain MRI using Modified kNN Algorithm," 2020 IEEE International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC), 2020, pp. 1-6, doi: 10.1109/iSSSC50941.2020.9358867.

[10] H. S. Suresha and S. S. Parthasarathy, "Alzheimer Disease Detection Based on Deep Neural Network with Rectified Adam Optimization Technique using MRI Analysis," 2020 Third International Conference on Advances in Electronics, Computers and Communications (ICAECC), 2020, pp. 1-6, doi: 10.1109/ICAECC50550.2020.9339504.

[11] H. Ahmed, H. Soliman and M. Elmogy, "Early Detection of Alzheimer's Disease Based on Single Nucleotide Polymorphisms (SNPs) Analysis and Machine Learning Techniques," 2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI), 2020, pp. 1-6, doi: 10.1109/ICDABI51230.2020.9325640.

[12] N. M. Khan, N. Abraham and M. Hon, "Transfer Learning With Intelligent Training Data Selection for Prediction of Alzheimer's Disease," in IEEE Access, vol. 7, pp. 72726-72735, 2019, doi: 10.1109/ACCESS.2019.2920448.

[13] S. Ahmed et al., "Ensembles of Patch-Based Classifiers for Diagnosis of Alzheimer Diseases," in IEEE Access, vol. 7, pp. 73373-73383, 2019, doi: 10.1109/ACCESS.2019.2920011.

[14] T. Zhou, M. Liu, K. -H. Thung and D. Shen, "Latent Representation Learning for Alzheimer's Disease Diagnosis With Incomplete Multi-Modality Neuroimaging and Genetic Data," in IEEE Transactions on Medical Imaging, vol. 38, no. 10, pp. 2411-2422, Oct. 2019, doi: 10.1109/TMI.2019.2913158.

[15] W. Li, Y. Zhao, X. Chen, Y. Xiao and Y. Qin, "Detecting Alzheimer's Disease on Small Dataset: A Knowledge Transfer Perspective," in IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 3, pp. 1234-1242, May 2019, doi: 10.1109/JBHI.2018.2839771.