iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** X **Month of publication:** October 2022

DOI: https://doi.org/10.22214/ijraset.2022.46937

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue X Oct 2022- Available at www.ijraset.com

Alzheimers Disease Detection Using Different Machine Learning Algorithms

S. Harika[1], T. Yamini[2], T. Nagasaikamesh[3], S. H. Basha[4], S. Santhosh Kumar[5], Mrs. S. Sri DurgaKameswari[6]
[1, 2, 3, 4, 5, 6]*ECE Department, GMR Institute of Technology*

Abstract: ***Alzheimer's disease is the most common form of dementia affecting the brain's parts. A broad term used to describe illnesses and conditions that causes a deterioration in memory, language, and other cognitive abilities severe enough to interface with daily life is "dementia". According to estimates, this disease affects 6.2 million Americans and 5 million people in India aged 65and older. In 2019, the most recent year for which data are available, official death certificates reported 121,499 deaths from AD, making Alzheimer's the "sixth leading cause of death in the country and the fifth leading cause of death for people 65 and older". In this paper, we suggest several machine Learning algorithms like Decision trees, SVM, Logistic regression, and Naive Bayes identify AD at an early stage. The Alzheimer's Disease Neuroimaging Initiative (ADNI) and the Open Access Series of Imaging Investigations (OASIS) provide data sets white used to detect the disease in its early stage. The datasets consist of longitudinal MRI data (age, gender, mini mental status, CDR) By taking into account many factors in each method, such as precision, F1 Score, Recall, and specificity are calculated. The results obtained 93.7% of maximum accuracy for the Decision Tree Algorithm.***

I. INTRODUCTION

Alzheimer's disease (AD) is a neurological condition that typically develops slowly and gets worse over time[1]. It is the root cause of 60–70% of dementia (brain damage) cases. The most prevalent initial sign is trouble recalling recent events. As the illness worsens, symptoms may include behavioral problems, linguistic difficulties, disorientation, mood swings, a lack of desire, and self-neglect. Body functions gradually deteriorate(collapses), which eventually results in death. The usual life expectancy upon diagnosis is three to nine years, however, the rate of development might vary. A common problem in the starting stages of Alzheimer's disease is, It can be difficult to recall the correct term or name, Have trouble remembering people's names when you meet unfamiliar persons, and Consistently being in social professional environments can provide a challenge, Having forgotten a passage you just read in a book or an alternative, Having a difficult time locating or misplacing a pricey item, It's getting harder and harder to complete tasks and activities to arrange or plan. An estimation of 55 million individuals worldwide is thought to be affected by Alzheimer's disease, with nearly 10 million new cases being identified each year, according to a World Health Organization poll from 2022. Early diagnosis of this disease is a time-consuming and expensive process since we need to gather a lot of data, apply advanced algorithms for prediction, and include an expert physician. Since automated systems are not susceptible to human mistakes, they may be employed in medical decision support systems and are more accurate than human evaluation. Automating Alzheimer's diagnosis will lessen both the length of the diagnosis process and the amount of human involvement which is essential. Automation also results in lower total expenses and more precise outcomes. Clinical diagnosis of Alzheimer's disease can be challenging, especially early on. Utilizing classification tools, we aim to enhance diagnosis efforts. This study looks into a few methods for categorizing people with Alzheimer's disease using their MRI scans and demographic information [13]. The Alzheimer's Disease Neuroimaging Initiative (ADNI) database and UC Berkeley Biomarkers both provided some of the MRI biomarkers that were used in this study. Multisite research called the Alzheimer's Disease Neuroimaging Initiative (ADNI) seeks to advance Alzheimer's disease clinical trials by enhancing their design and efficiency (AD) in various ways according to the need.[2] This collaborative project examines individuals with AD, those who may develop AD, and controls who do not exhibit any evidence of cognitive impairment using resources from the public and commercial sectors.

II. LITERATURE SURVEY

It was recommended by J.Neelaveni and M. S. Geetha Devasana to use several machine learning algorithms to forecast Alzheimer's disease. As input to the model, they used various psychological parameters including MMSE (Mini-Mental State Examination), age, and Education and the primary factor used is the MMSE. A person is diagnosed with AD if their MMSE score reduces gradually. Using SVM and decision Tree (Machine learning algorithms), they were able to estimate the AD and later compared the algorithms' accuracy results.

After comparing the best accuracy is given by SVM (Support Vector Machine). But in this case, only one parameter—accuracy—was taken into account, and the resultant accuracy was extremely low in comparison to other models. 70% of the training dataset is used to train each algorithm, and 30% of the test dataset is used to test it. This model not only distinguishes between cognitive impairment but also detects the disease in the person.

[4]Sakshi Singh and Komal Gaikwad suggested the method of using "Shallow learning and Deep learning techniques to detect Alzheimer's Disease. With the aid of various shallow learning techniques, it is also possible to study the psychological and socioeconomic effects that the disease may have on those who are affected in addition to clinical insights into the identification of Alzheimer's Disease. When you add additional examples and training data to the network, machine learning approaches like "shallow learning" reach their performance limits. This study employed the Mann- Whitney test and correlation matrix with the Sabon library, which runs the standard uni private Pearson's test, for appropriate characteristics. They tested the dataset by using gradient boosting classifier, XG boost, RFC, Ada boost classifier, Decision tree classifier, SVM Linear, SVM Radial, and Logistic regression. According to this study, linear SVM, and Ada boost classifier closely trails logistic regression in terms of accuracy when based on clinical parameters related to brain volume and the MMSE.[5] Decision tree classifiers are reported to have the highest precision, whereas bagging classifiers excel at both recall and F1 scores. Both the clinical dataset and the MRI dataset have an accuracy of 83%. Additionally, our method makes it abundantly clear that machine learning can be used to categorize electronic health records of patients with Alzheimer's disease or dementia into two categories, with significant room for improvement in both the accuracies of deep learning and shallow learning 2018, [6]Structures Kai Yu, Yingsong Li, and Xiaoguang Liu proposed an approach. But the main disadvantage is this model does, not apply to larger datasets.

[7]Srinivasan Aruchamy and Amrita Haridasan proposed the method using machine learning techniques in 3D MR images for the detection of Alzheimer's Disease. In this, study the author used a different model to detect disease at an early stage by separating grey and white matter using 3D images. The people who are suffering from AD mainly affect the grey and white matter of the brain. The white and grey matter of the brain gets damaged. The 3d image database was collected from OASIS which consists of 460 subjects' male and female 3D image which was taken as input was visualized in three directions Axial, Coronal, and sagittal. The brain consists of grey and white matter separated from 3D images after that single slice extraction was performed. This first feature reduction step was performed to extract prominent features and after that pre-processing of data was performed. In this Four different algorithms are used to detect AD at the early stage. The algorithms used are Logistic Regression, SVM, Naive Bayes, and Ada boost classifier. These are applied to both grey and white matter separately in all three directions. The accuracy of all the algorithms was taken. The input of 460 subjects was taken for the classification and from the classification n t the accuracy logistic regression obtained was 87.8% among all of them the lowest accuracy obtained was for75.3% of Naive Bayes the grey matter. The maximum accuracy obtained was 90.9% for Ada boost algorithms. The drawback of this paper is to study the grey matter and white matter separately effort should be made to study the 3D images. So it takes more time and more effort.

III. METHODOLOGY

This section consists of different methods that are used to predict Alzheimer's Disease in the early stage. In the health care industry, machine learning has a big role. In the past 20 years, the field of artificial intelligence known as machine learning has grown significantly in popularity and importance. Data is fed to an algorithm using machine learning so it can comprehend the relationship between input and output.

The healthcare industry offers a considerable amount of data, and domain to create a cutting-edge, scientific technique to detect illness at an early stage. So, certain machine learning predicts the disease and determines which method has the highest accuracy. Some of the algorithms used are Logistic Regression, Support vector machine, Decision Tree, and Naive Bayes.[8] The dataset which is used for detecting Alzheimer's disease is taken from OASIS. The input required is the dataset consisting of Age, Gender, Years of Education, Mini-mental status, etc. The dataset contains 373 subjects aged from 60 to 96. First Data pre-processing is done [17].

Cleaning and preparing the data for a machine learning model is essential Since it increases the model's precision and efficiency. While doing machine learning projects, we do not always come across clean and formatted data [9]. There are many algorithms in machine learning to predict Alzheimer's disease. The algorithms which are used for implementing the model are Decision tree, support vector machine, Naïve Bayes, and Logistic Regression. The implementation is done in Matlab. The health care industry provided a large amount of data to find and develop a method such that to predict the disease in the early stage which will benefit the people. The block diagram was shown in figure 4.1.1

A. Implementation Of Machine Learning Algorithms

1) *Decision Tree:* The principal applications of decision trees are classification and regression problems. It is a technique of supervised learning. It is a tree-structured classifier, where internal nodes represent the features of a dataset, branches the process of generating decisions, and each leaf node is the classification outcome.
2) *Support Vector Machine:* It is one of the most well-liked techniques for supervised learning, support vector machines are utilized for both difficulties involving classification and regression [5]. However, its main application is for Machine Learning Classification issues. To simply place fresh data points in the appropriate category in the future, The support vector machine algorithm seeks to determine the ideal decision boundary or line for classifying an n-dimensional space. A hyperplane is a name given to this optimal decision boundary. The best decision boundary or line for categorizing- dimensional space is sought after by the support vector machine learning algorithm. These extreme points are called support vectors.[10] These support vectors which are close to the hyperplane affect the position of the hyperplane distance between support vectors and the hyperplane known as margin.
3) *Naïve Bayes:* It is a supervised method that relies on the Bayes theorem to address classification problems. Naive Bayes is one of the quick and easy machine learning approaches for forecasting a class of datasets [16]. It is used for binary as well as multi-class classifications. It will give good predictions in multiclass compared to other classes. In Nave Bayes, three model types exist. One is the Gaussian Naive Bayes model, and two is the Multinomial Naive Bayes model. Third, Bernoulli Naive Bayes model.
4) *Logistic Regression:* One of the Machine Learning algorithms that fall under the Supervised Learning method is logistic regression. It will predict the provided dependent variable from a set of specified independent variables. The output must be categorical or discrete values. It is comparable to linear regression, except logistic regression is used for classification issues whereas linear regression is used for regression issues. It is used for predicting two values (0 or 1). The expression for logistic regression is obtained from the linear regression equation. Three varieties of logistic regression exist 1. Binomial 2. Multinomial 3. Ordinal

B. Model Analysis

An evaluation of the classifier's performance typically involves utilizing a confusion matrix. The True classes and classifier predicted classes are presented in a specific table. In the confusion matrix, four different terminologies are used. False positives and False negatives are present, as well as True positives and True negatives [11]. Further accuracy and parameters were calculated for the classification model. With the help of a confusion matrix, these parameters are calculated for each algorithm.

1) *Precision:* The ratio of the overall number of true positives to the total number of positive forecasts is known as precision. It is equal to the product of true positives and false positives.
2) *F1-Score:* It only quantifies the proportion of accurate predictions a machine learning model has produced.
F-Measure = (2 * Precision * Recall) / (Precision + Recall)
3) *Recall:* The recall is determined as the proportion of Positive samples that were correctly identified as Positive to all Positive samples. The recall is defined as True Positives / (False Negatives + True Positives).
4) *Specificity:* It is the percentage of true negatives that the model properly predicted. Specificity is calculated as (True Negative) / (True Negative + False Positive).

IV. RESULTS AND DISCUSSIONS

The overall objective of this paper is to predict Alzheimer's disease at an early stage. Different Algorithms, including Decision Tree, Support vector machine. Naïve Bayes and logistic regression are compared to determine accuracy. The dataset contains 373 subjects aged from 60 to 96. All the subjects are righthanded of both men and women. By performing the process it was observed the highest accuracy obtained was the decision tree algorithm i.e. 93.7%.

Algorithm	Accuracy
Decision Tree	93.7%
SVM	56.7%
Logistic Regression	57.1%
Naïve Bayes	83.8%

Fig1: Accuracy of algorithms

Algorithm	precision	Recall	F1 score	Specificity
Decision Tree	1	0.9	0.94	1
SVM	1	0.56	0.71	0
Logistic Regression	0	0	0.42	1
Naïve Bayes	0.92	0.81	0.86	0.87

Fig2: Classification of parameters

The above fig1 shows the accuracy of different algorithms and it was observed that decision tree has highest algorithm i.e. 93.7% And the above fig2 show the parameters of different algorithms. These parameters are calculated to know the performance of the model. If all the parameters values are 1 then it is good model. The below fig3 shows the confusion matrix of Decision tree algorithm.

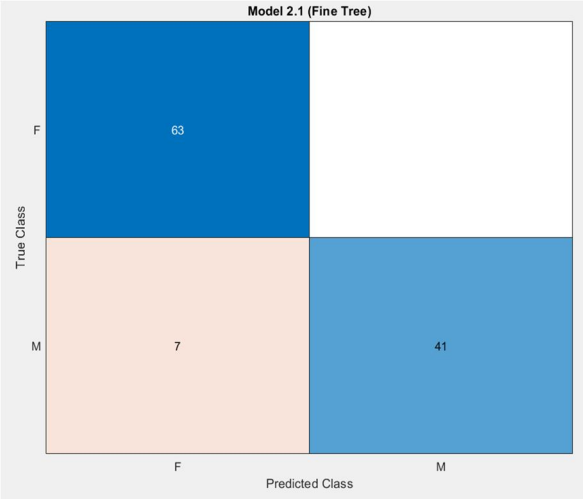


Fig3: Confusion matrix

V. CONCLUSION

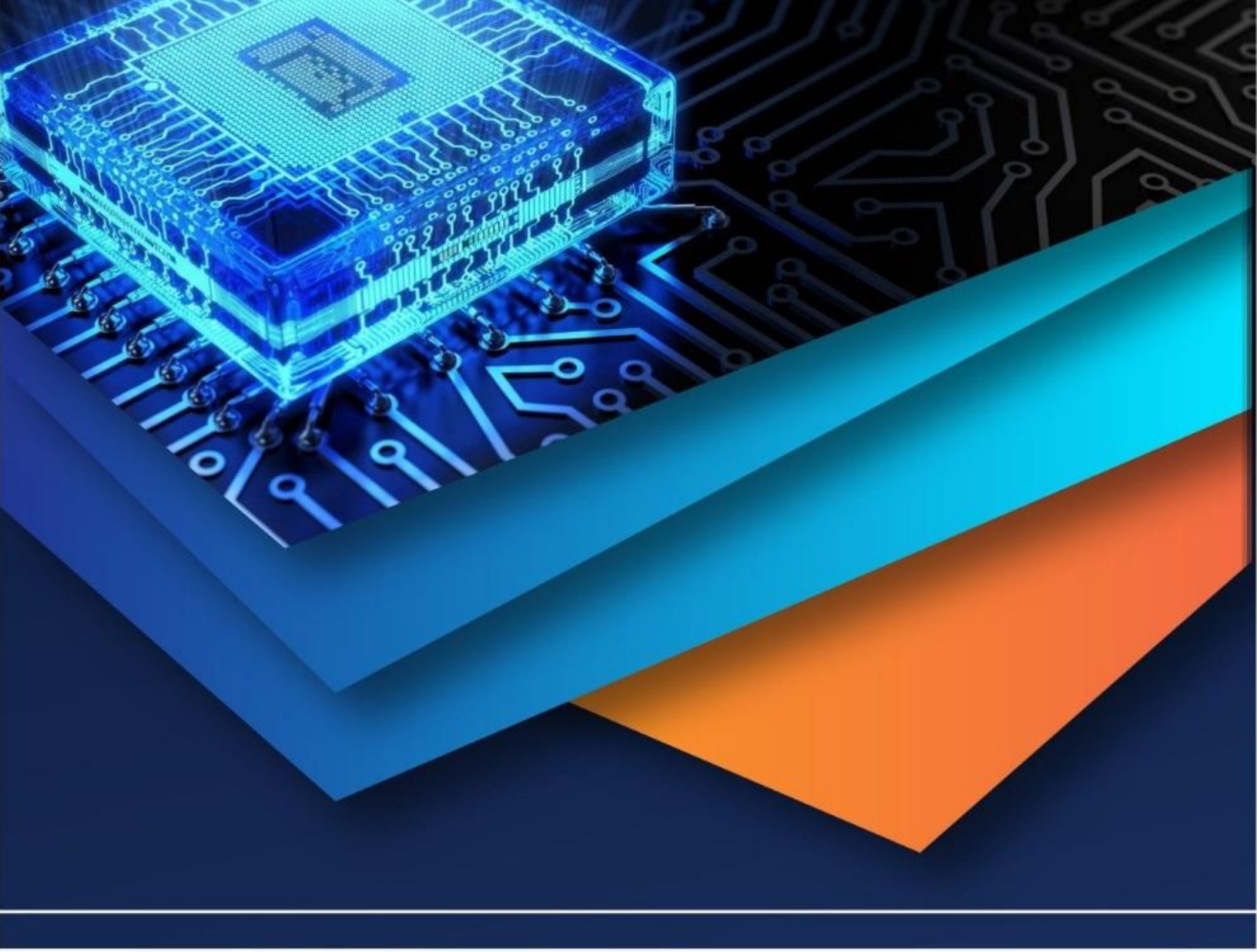
Alzheimer's is also known as Dementia. This disease is affecting most people which makes them lose their memory, thinking ability, and also able to perform their daily functions. Hence to detect this disease in the early stage is very necessary. The results of this project will help us to predict Alzheimer's disease in the early stage by finding the maximum accuracy rate of machine learning algorithms. The results which we obtained from the performance show that the Decision tree is the best method to detect the disease at an early stage when compared to other methods. A better model can be built rather than decision tree algorithms to achieve high accuracy by using Matlab.

VI. ACKNOWLEDGEMENT

I would like to thank our guide SMT.S Durga Kameswari Asst. Professor, Department of Electronics and Communication Engineering of wholehearted and valuable guidance throughout the project, and I would like to sincerely thank our HOD, Principal, all the staff members, and teammates for their direct and indirect support in helping us in the completion of the project.

REFERENCES

[1] P. Kishore, Ch. Usha Kumari, M.N.V.S.S. Kumar, T . Pavani, "Detection and analysis of Alzheimer's disease using various machine learning algorithms."Materials today proceedings, Volume-45.
[2] Wenlu Yanga,f, Ronald L.M. Luib, Tony F. Chand, Shing- Tung Yau, Reisa A. Sperling, and Xudong Huang"Independent component analysis-based classification of Alzheimer's MRI data ."Maritime University, Shanghai, China bDepartment of Mathematics.
[3] J. Ye, T. Wu, J. Li, and K. Chen, "Machine learning approaches for the neuroimaging study of Alzheimer's disease," Computer, volume- 44, no. 4, pp. 99–101.
[4] Grassi M, Loewenstein DA, Caldirola D, Schruers K, DuaraR, Perna G, "A clinically-translatable machine learning algorithm for the prediction of Alzheimer's disease conversion: further evidence of its accuracy via a transfer learning approach.
[5] H.Nishi et al, "Deep-learning derived high-level neuroimaging features predict clinical outcomes for large vessel occlusion."Stroke, Volume-51.
[6] M. Liu, J. Zhang, P.-T. Yap, D. Shen, "View-aligned hypergraph learning for Alzheimer's disease diagnosis with incomplete multi-modality data", Med.ImageAnal., 2017 vol. 36.
[7] Hansson O, Seibyl J, Stomrud E, Zetterberg H, Trojanowski, "CSF biomarkers of Alzheimer's disease concord with amyloid-bPET and predict clinical progression: A study of fully automated immunoassays in Bio-FINDER andADNIcohorts".Alzheimer's Dement 2018.
[8] Malavika, Vanitha "Alzheimer Disease Forecasting using Machine Learning Algorithm", Department of Information Technology, Kumaraguru College of Technology, Coimbatore, India.
[9] Javier Escudero, "Machine Learning-Based Method for Personalized and Cost-Effective Detection of Alzheimer's Disease", Member, IEEE, for the Alzheimer's Disease Neuroimaging Initiative.
[10] Farah Shamoun, Tingting Zhu, and David A.Clifton. "Machine Learning for Clinical outcome predictions."IEEE in Bio-medical Engineering, Volume-14.
[11] Srinivasan Aruchamy, "Alzheimer'sDisease Detection using Machine Learning Techniques in 3D MR Images".Robotics and Automation Group CSIR –CMER Durgapur, India
[12] Godin J, Keefe J, Andrew MK. "Handling missing MiniMental State Examination (MMSE) values:" Results from a cross-sectional long-term care study Journal of epidemiology.
[13] Cao J, Kwong S, Wang R. "A noise detection based Adaboost algorithm for mislabeled data." pattern Recognition, Volume-45.
[14] Dr. CH. Dhanunjaya Rao and G. Hari Charan (2022). "Diagnosis of Alzheimer's Disease using Machine Learning Algorithms."International Journal for Research in Applied Science & Engineering Technology (IJRASET).July 2022.
[15] J.Neelaveni and M.S.Geetha Devasana (2020). Alzheimer's Disease prediction using a machine learning algorithm. 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS). Coimbatore, India.
[16] Sakshi Singh, Komal Gaikwad, Asma Nehal (2022). "Detecting Alzheimer's using Shallow Learning and Deep Learning Techniques." International Journal of Advanced Research in Computer and Communication Engineering IJARCCE. 5 May 2022, Pune, India.
[17] SidraMinhas, Aasia Khanum, Farhan Riaz"Predicting Progression from Mild Cognitive Impairment to Alzheimer's Disease using Autoregressive Modelling of Longitudinal and Multimodal Biomarkers." IEEE Journal of Biomedical and Progress.
[18] Nation, D.A., Sweeney, M.D., Montagne, A., Sagare, A.P., D'Orazio, L.M., Pachicano, M. et al, "Blood-brain barrier breakdown is an early biomarker of human cognitive dysfunction", Nat Med. 2019.
[19] SidraMinhas, Aasia Khanum, Farhan Riaz"Predicting Progression from Mild Cognitive Impairment to Alzheimer's Disease using Autoregressive Modelling of Longitudinal and Multimodal Biomarkers." IEEE Journal of Biomedical and Progress.
[20] Nation, D.A., Sweeney, M.D., Montagne, A., Sagare, A.P., D'Orazio, L.M., Pachicano, M. et al, "Blood-brain barrier breakdown is an early biomarker of human cognitive dysfunction", Nat Med. 2019



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY