Alzheimer Disease Prediction Using Machine Learning Algorithms

T Ganga Bhavani[1], E. Amrutha Sai[2], S. Siva Rama Krishna[3], A.N.D. Aditya[4], N. Ramani Mounika[5]
[1,2,3,4,5]*Pragati Engineering College*

Abstract **- Alzheimer disease is the one amongst neurodegenerative disorders. Though the symptoms are benign initially, they become more severe over time. Alzheimers disease is a prevalent sort of dementia. This disease is challenging one because there is no treatment for the disease. Diagnosis of the disease is done but that too at the later stage only. Thus, if the disease is predicted earlier, the progression or the symptoms of the disease can be slowed down. This paper uses machine learning algorithms to predict Alzheimer disease using psychological parameters.**

I..INTRODUCTION

Machine learning is used to interpret and analyze data. Furthermore, it can classify patterns and model data. It permits decisions to be made that could not be made generally utilizing routine systems while sparing time and endeavors. Machine learning methodologies have been extensively used for computer-aided diagnosis in medical image formation mining and retrieval with wide variety of other applications especially in detection and classifications of brain disease using CRT images and x-rays. It has just been generally late that AD specialists have endeavored to apply machine learning towards AD prediction. As a consequence, the literature in the field of Alzheimer's disease prediction and machine learning is relatively small. However, today's imaging technologies and high throughput diagnostics have lead us overwhelmed with large number (even hundreds) of cellular, clinical and molecular parameters.

1.1 MOTIVATION

In current circumstances, the standard measurements and human instinct don't frequently work. That is the reason we must depend on intensively computational and non-traditional approaches such as machine learning. The custom of using machine learning as a part of disease prediction and visualization is a fragment of an expanding shift towards prescient and customized prescription. This drift is important, not only for the patients in increasing their quality of life and lifestyle, but for physicians in making treatment decisions and also for health economists. It's not the disease of age, it's the disease of the brain and patients may show Symptoms like loss of memory, difficulty in finding the right words or understanding what people are saying, difficulty in performing previously routine tasks and personality and mood changes. When a radiologist views a medical report e.g., a magnetic resonance imaging (MRI) scan of a patient, a biased thinking for a disease would result in missing the chance of detecting other disease conditions. Thus, it leads to considering only a subset of causes and conditions. C S Lee, claimed that approximately 75% of all the medical errors occurred due to diagnostic errors by the radiologists. An increased workload, stress, fatigue, cognitive bias, and poor system are some of the factors behind it. In this situation, smart diagnostic systems provide a safe clinical support to the clinicians. The Alzheimer's association claims that AD is the sixth leading cause of death in United States.

1.2 OBJECTIVE OF PROJECT

The Main objective of the project is: Is to make the diagnosis of the disease easier, to detect the disease in its early stages and use the machine algorithm efficiently

II. METHODOLOGY

Detection of Alzheimer's disease at prodromal stage is very important as it can prevent serious damage to the patient's brain. The detection of the Alzheimer's disease is done the medical professionals through certain steps such as CT scan. The detection of the Alzheimer's disease is time consuming using CT scan. Hence, we are applying the machine learning technology to predict the Alzheimer's disease. The proposed approach detects the various stages of

Alzheimer's Disease such as moderate-demented and non-demented using CNN algorithm. It reduces the time required to predict the output and can be used for real time predictions. We describe the datasets used in this study and how the data was pre- processed before the machine learning task. Feature extraction using principal component analysis and feature selection techniques were also employed. After the data preprocessing is done, the efficient machine learning algorithm that is CNN is applied to predict the disease and categorized it into moderate demented and non-demented. Alzheimer disease is the one amongst neurodegenerative disorders. Though the symptoms are benign initially, they become more severe over time. Alzheimer's disease is a prevalent sort of dementia. This disease is challenging one because there is no treatment for the disease. Diagnosis of the disease is done but that too at the later stage only. Thus if the disease is predicted earlier, the progression or the symptoms of the disease can be slowed down. This paper uses machine learning algorithms to predict Alzheimer disease using psychological parameters like age, number of visits, MMSE and education.

III. SYSTEM ANALYSIS

EXISTING SYSTEM

Method of deep learning along with the brain network and clinical significant information like age, ApoE gene and gender of the subjects for earlier examination of Alzheimer's. Brain network was arranged, calculating functional connections in the brain region by employing the resting-state functional magnetic resonance imaging (R-fMRI) data. To produce a detailed discovery of the early AD, a deep network like auto encoder is used where functional connections of the networks are constructed and are susceptible to AD and MCI.

PROPOSED SYSTEM:

Proposed a method called multistage classifier by using machine learning algorithms like Support Vector Machine, Naïve Bayes and K-nearest neighbor to classify between different subjects. PSO (particle swarm optimization) which is a technique that best selects the features was enforced to obtain best features. Naturally image retrieving process requires two stages: the first stage involves generating features so that it reproduces the query image and then later step correlate those features with those already gathered in the database. The PSO algorithm is used to select the finest biomarkers that show AD or MCI.

IV. SYSTEM STUDY

FEASIBILITY STUDY

The feasibility of the project is analyzed in this phase and business proposal is put forth with a very general plan for the project and some cost estimates. During system analysis the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system is not a burden to the company. For feasibility analysis, some understanding of the major requirements for the system is essential. Three key considerations involved in the feasibility analysis are,

- ECONOMICAL FEASIBILITY

This study is carried out to check the economic impact that the system will have on the organization. The amount of fund that the company can pour into the research and development of the system is limited.

- TECHNICAL FEASIBILITY

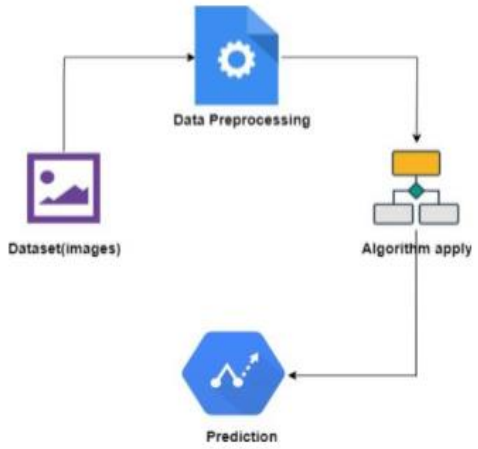
This study is carried out to check the technical feasibility, that is, the technical requirements of the system.

- SOCIAL FEASIBILITY

The aspect of study is to check the level of acceptance of the system by the user. This includes the process of training the user to use the system efficiently.

V. SYSTEM DESIGN

FLOWCHART OF PROJECT



Flowchart of project

SOFTWARE ENVIRONMENT

Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it learn for themselves.

MLPROCESS:

Data Collection: The quantity & quality of your data dictate how accurate our model is. The outcome of this step is generally a representation of data which we will use for training.

Data Preparation: Wrangle data and prepare it for training. Clean that which may require it. Randomize data, which erases the effects of the particular order in which we collected and/or otherwise prepared our data.

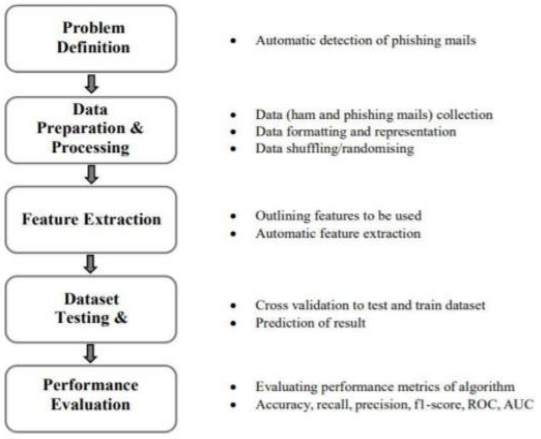
Choose a Model: Different algorithms are for different tasks; choose the right one.

Train the Model: The goal of training is to answer a question or make a prediction correctly as often as possible. Linear regression example: algorithm would need to learn values form and b (x is input, y is output). Each iteration of process is a training step.

Evaluate a Model: Uses some metric or combination of metrics to "measure" objective performance of model.

Parameter Tuning: This step refers to hyper parameter tuning, which is an "art form" as opposed to a science.

Make Predictions: Using further (test set) data which have, until this point, been with held from the model (and for which class labels are known), are used to test the model; a better approximation of how the model will perform in the real world.



VI. IMPLEMENTATION

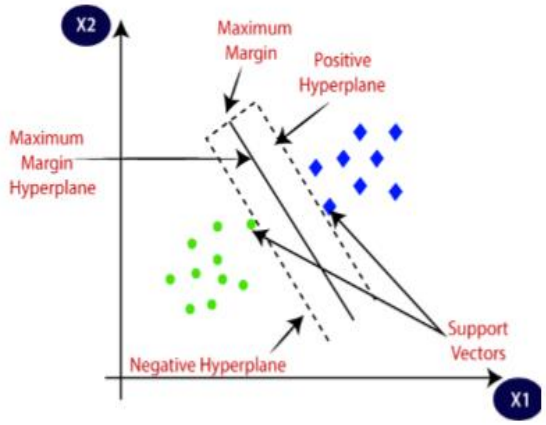
ALGORITHMS:

We have used two algorithms in our project they are:

1. Support Vector Machine
2. Decision tree classifier

SUPPORT VECTOR MACHINE:

Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems. However, primarily, it is used for Classification problems in Machine Learning. The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future. This best decision boundary is called a hyperplane. SVM chooses the extreme points/vectors that help in creating the hyperplane. These extreme cases are called as support vectors, and hence algorithm is termed as Support Vector Machine. Consider the below diagram in which there are two different categories that are classified using a decision boundary or hyperplane.



DECISION TREE CLASSIFIER:

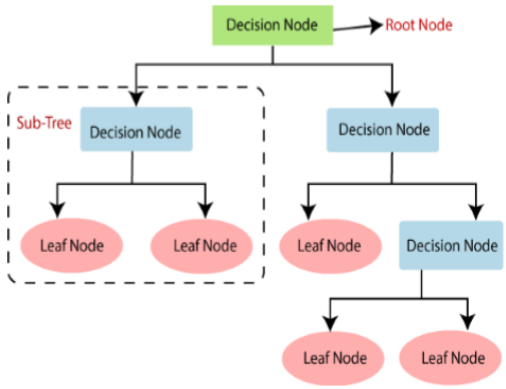
Decision Tree is a Supervised learning technique that can be used for both classification and Regression problems, but mostly it is preferred for solving Classification problems. It is a tree-structured classifier, where internal nodes represent the features of a dataset, branches represent the decision rules and each leaf node represents the outcome.

In a Decision tree, there are two nodes, which are the Decision Node and Leaf Node. Decision nodes are

used to make any decision and have multiple branches, whereas Leaf nodes are the output of those decisions and do not contain any further branches.

The decisions or the test are performed on the basis of features of the given dataset.

It is a graphical representation for getting all the possible solutions to a problem/decision based on given conditions.



VII. CONCLUSION

Machine learning approach to predict the Alzheimer disease using machine learning algorithms is successfully implemented and gives greater prediction accuracy results. The model predicts the disease in the patient and also distinguishes between the cognitive impairment. The future work can be done by combining both brain MRI scans and the psychological parameters to predict the disease with higher accuracy using machine learning algorithms. When they are combined, the disease could be predicted with a higher accuracy in the earlier stage itself.

VIII. ACKNOWLEDGMENT

We would like to express our gratitude to all the people behind the screen who helped us to transform an idea into a real application.

We would like to thank our internal guide Mrs. T. Ganga Bhavani for his technical guidance, constant encouragement and support in carrying out our project. The satisfaction and euphoria that accompany the successful completion of the task would be great but incomplete without the mention of the people who made it possible with their constant guidance and encouragement crowns all the efforts with success.

REFERENCES

[1] K.R.Kruthika, Rajeswari, H.D.Maheshappa, “Multistage classifier-based approach for Alzheimer’s Disease prediction and retrieval”, Informatics in Medicine Unlocked, 2019.

[2] Ronghui Ju , Chenhui Hu, Pan Zhou , and Quanzheng Li, “Early Diagnosis of Alzheimer’s Disease Based on Resting-State Brain Networks and Deep Learning”, IEEE/ACM transactions on computational biology and bioinformatics, vol. 16, no. 1, January/February 2019.

[3] Ruoxuan Cuia, Manhua Liu “RNN-based longitudinal analysis for diagnosis of Alzheimer’s disease”, Informatics in Medicine Unlocked, 2019

[4] Fan Zhang , Zhenzhen Li , Boyan Zhang , Haishun Du , Binjie Wang , Xinhong Zhang, “Multi-modal deep learning model for auxiliary diagnosis of Alzheimer’s disease”, NeuroComputing, 2019

[5] Chenjie Ge , Qixun Qu , Irene Yu-Hua Gu , Asgeir Store Jakola “Multi-stream multi- scale deep convolutional networks for Alzheimer’s disease detection using MR images”, NeuroComputing, 2019

[6] Tesi, N., van der Lee, S.J., Hulsman, M., Jansen, I.E., Stringa, N., van Schoor, N. et al, “Centenarian controls increase variant effect sizes by an average twofold in an extreme caseextreme control analysis of Alzheimer's disease”, Eur J Hum Genet. 2019;27:244–253

[7] J. Shi, X. Zheng, Y. Li, Q. Zhang, S. Ying, "Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of Alzheimer’s disease", IEEE J. Biomed. Health Inform., vol. 22, no. 1, pp. 173-183, Jan. 2018.

[8] M. Liu, J. Zhang, P.-T. Yap, D. Shen, "Viewaligned hypergraph learning for Alzheimer's disease diagnosis with incomplete multi-modality data", Med. Image Anal., 2017 vol. 36, pp. 123- 134.

[9] Hansson O, Seibyl J, Stomrud E, Zetterberg H, Trojanowski JQ,Bittner T, “CSF biomarkers of Alzheimer’s disease concordwith amyloid-bPET and predict clinical progression: A study of fullyautomated immunoassays in BioFINDER

and ADNI cohorts". Alzheimers Dement 2018;14:1470–81.

[10] Van der Lee SJ, Teunissen CE, Pool R, Shipley MJ, Teumer A,Chouraki V, "Circulating metabolites and general cognitive abilityand dementia: Evidence from 11 cohort studies", Alzheimer's Dement2018;14:707–22. [15] *Letter Symbol sfor Quantities*, ANSI StandardY10.5-1968.

[11] Kauppi Karolina, Dale Anders M, "Combining Polygenic Hazard Score With Volumetric MRI and Cognitive Measures Improves Prediction of Progression from Mild Cognitive Impairment to Alzheimer's Disease", Frontiers in Neuroscience, 2018.

[12] Grassi M, Loewenstein DA, Caldirola D, Schruers K, Duara R, 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", Int Psychogeriatr, 2018 14:1–9. doi: 10.1017/S1041610218001618.

[13] 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;25:270–276.