Psychological Parameters Based Alzheimer Disease Prediction

Gundumolu Divya Vani
PG scholar, Department of MCA, CDNR collage, Bhimavaram, Andhra Pradesh.
V.Sarala
(Assistant Professor), Master of Computer Applications, DNR collage, Bhimavaram, Andhra Pradesh.

Abstract: ***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 slow down. This paper uses machine learning algorithms to predict the Alzheimer disease using psychological parameters like age, number of visit, MMSE and education.***

Keywords- Alzheimer disease, mild cognitive impairment, machine learning algorithms, psychological parameters.

I. INTRODUCTION

Alzheimer disease is caused by both genetic and environmental factors, those affects the brain of a person over time. The genetic changes guarantee a person will develop this disease. This disease breaks the brain tissue over time. It occurs to people over age 65. However people live with this disease for about 9 years and about 1 among 8 people of age 65 and over have this disease. MMSE (Mini Mental State Examination) score is the main parameter used for prediction of the disease.

This score reduces periodically if the person is affected. Those people having MCI have a serious risk of growing dementia. When the fundamental MCI results in a loss of memory, the situation expects to develop to dementia due to this kind of disease. There is no treatment to cure Alzheimer's disease. In advanced stages of the disease, complications like dehydration, malnutrition or infection occurs which leads to death. The diagnosis at MCI stage will help the person to focus on healthy approach of life, and good planning to take care of memory loss.

1.1 Alzheimer Disease Predictions

Alzheimer's disease (AD) is the leading cause of dementia in older adults. There is currently a lot of interest in applying machine learning to find out metabolic diseases like Alzheimer's and Diabetes that affect a large population of people around the world. Their incidence rates are increasing at an alarming rate every year. In Alzheimer's disease, the brain is affected by neurodegenerative changes. As our aging population increases, more and more individuals, their families, and healthcare will experience diseases that affect memory and functioning. These effects will be profound on the social, financial, and economic fronts. In its early stages, Alzheimer's disease is hard to predict.

A treatment given at an early stage of AD is more effective, and it causes fewer minor damage than a treatment done at a later stage. Several techniques such as Decision Tree, Random Forest, Support Vector Machine, Gradient Boosting, and Voting classifiers have been employed to identify the best parameters for Alzheimer's disease prediction. Predictions of Alzheimer's disease are based on Open Access Series of Imaging Studies (OASIS) data, and performance is measured with parameters like Precision, Recall, Accuracy, and F1-score for ML models.

The proposed classification scheme can be used by clinicians to make diagnoses of these diseases. It is highly beneficial to lower annual mortality rates of Alzheimer's disease in early diagnosis with these ML algorithms. The proposed work shows better results with the best validation average accuracy of 83% on the test data of AD. This test accuracy score is significantly higher in comparison with existing works.

Alzheimer's Disease (AD) is a progressive neurological condition that leads to short-term memory loss, paranoia, and delusional ideas that are mistaken for the effects of stress or aging. In the United States, this Disease affects about 5.1 million people. AD does not have proper medical treatment. In order to control AD, continuous medication is necessary. AD (1) is chronic so that it can last for years or the rest of your life. Therefore, it is most important to prescribe medication at the appropriate stage so that the brain is not damaged to a great extent. Early detection of this Disease is a tedious and costly process since we must collect a lot of data and use sophisticated tools for prediction and have an experienced doctor involved.

Automated systems are more accurate than human assessment and can be used in medical decision support systems because they are not subject to human errors. Based on previous research on AD, researchers have applied images (MRI scans), biomarkers (chemicals, blood flow), and numerical data extracted from the MRI scans to study this Disease. As such, they were able to determine whether a person was demented or not. In addition to shortening diagnosis time, more human interaction will be reduced by automating Alzheimer's diagnosis. In addition, automation reduces overall costs and provides more accurate results.

For example, we can predict whether a patient is demented by analyzing MRI scans and applying prediction techniques. If a person has early-stage Alzheimer's disease, they are considered demented. By doing so, we can achieve better accuracy. When a person has Alzheimer's disease in the early stages, they can usually function without any assistance. In some cases, the person can still work, drive, and partake in social activities. Although this is the case, the person may still feel uneasy or suffer from memory loss, such as not remembering familiar words and locations.

People close to the individual notice that they have difficulty remembering their names. By conducting a detailed medical interview, a doctor may identify problems with memory and concentration in the patient. Common challenges in early stage of Alzheimer's disease include,

• It's hard to remember the right word or name.

• Having difficulty remembering names when meeting new people.

• Working in social settings or the workplace every day can be challenging.

• Having forgotten something that you have just read in a book or something else.

• Having trouble finding or misplacing a valuable object.

• Tasks and activities are becoming increasingly difficult to plan or organize.

II. LITERATURE SURVEY

Ronghui Ju et.al, suggested 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 [1]. 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 autoencoder is used where functional connections of the networks are constructed and are susceptible to AD and MCI. The dataset is taken from the ADNI database. The classification model consists of the early diagnosis, initially preprocessing of raw R-fMRI is done [1].

Then, the time series data (90 ×130matrix) is obtained and that indicates blood oxygen levels in each and every region of brain and changes a long period. Then, a brain network is built and transformed to a 90 ×90 time series data correlation matrix. The targeted autoencoder model is used which is a three layered model which gives intellectual growth of the nervous system then excerpts brain networks attributes completely [1]. When finite amount of data cases is taken, k-fold cross verification was implemented mainly to avoid the over fitting complication.

K. R. Kruthika et.al, proposed a method called multistage classifier by using machine learning algorithms like Support Vector Machine,

Naive Bayes and K-nearest neighbor to classify between different subjects [2]. 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 correlates those features with already gathered in database [2].

The PSO algorithm is used to select the finest biomarkers that show AD or MCI. The data is taken from Alzheimer's disease Neuroimaging Initiative (ADNI) database. The MRI scans are preprocessed first after taking from the database. The feature selection includes volumetric and thickness measurements. Then the optimum feature lists were obtained from PSO algorithm [2]. The Gaussian Naïve Bayes, K- Nearest Neighbor, Support vector machine was used to distinguish between the subjects.

Here a 2 stage classifier was used where in the initial stage GNB classifier was used to classify the objects between AD, MCI and NC and in later stages SVM and KNN were used to analyze the object based on the performance of the initial one [2]. Control Based Image Retrieval was used for retrieving images from the database. Ruoxuan Cuia et.al, proposed a model where longitudinal analysis is performed on consecutive MRI and is essential to design and compute the evolution of disease with time for the purpose of more precise diagnosis [3]. The actual process uses those features of morphological anomaly of the brain and the longitudinal difference in MRI and constructed classifier for distinguishing between the distinct groups.

The MRI brain images of 6 time points that is for consecutive intervals in a gap of six months are taken as inputs from ADNI database [3]. Then feature learning is done with the 3D Convolutional Neural Network. The CNN is followed by a pooling layer and have many ways for pooling, like collecting mean value otherwise the maximal, or definite sequence of neuron in the section. But for studying the characteristics, the convolutional operation of 2×2×2 is applied so that a linear combination is studied for pooling of neurons [3].

The fully connected layer has neurons that produce output of all neurons in a linear combination, which are taken from preceding layer and then is moved through nonlinearity. Finally for the last fully connected, a softmax layer is particularly used and then tuned finely for back-propagation to predict the class probability [3]. The result of each node varies from 0 to 1, and the total of nodes will always be 1. Finally the classification includes the deep network construction including the 3D CNN training and RNN model training.

Then the results of fully connective layers are directly mapped using a softmax function [3]. The initial parameters that were trained by both 3 dimensional CNN and the RNN network are established and then only the uppermost fully connective layer parameters and the softmax layer that was used for prediction are adjusted so that the dimensional and longitude features were united for distinct identification.

Fan Zhang et.al, proposed a multi-modal model where medical images are used for training. It is done by two separate convolutional neural networks. Here an auxiliary diagnosis using a deep learning model is used [4]. The two separate independent CNN for extracting the characteristics from both the MR scans and also the PET scans is used. It is obtained by a sequence of forward procreation convolution and the downer sampling method.

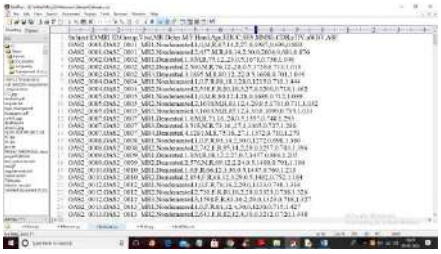
The outputs cohesion is calculated by correlation analysis for the two networks. The structure of CNN consists of sampling layer in the down, a convolution row or a fully connective row, pooling layer and finally the output row [4]. Then it computes the correlation using Pearson correlation coefficient method in between the prognosis of MRI scans and PET images.

The main idea of the correlation search is for regulating the output of both neuroimaging examinations whether they were persistent or not. The purpose of identifying and classifying is finished by using the output layer called softmax logistic regression method [4]. The benefit of this process is, it merges clinical neuropsychological results with neuroimaging results.

III. PROPOSED METHOD

Alzheimer Disease is one of the incurable and dangerous disease and its timely prediction can help in saving patients life. So author of this paper using machine learning algorithms such as SVM and decision tree to predict cured, non-Alzheimer and Alzheimer patients. To trained this algorithms author using dataset which contains patient AGE, number of visits to hospital and many other parameters. After training model we can apply test data on this model to predict patient condition as Normal or Alzheimer.

Below screen showing dataset details used in this project



In above screen first row contains dataset column names and remaining rows contains dataset values and by using above dataset we are training SVM and Decision tree algorithms and then comparing their performance in terms of accuracy.

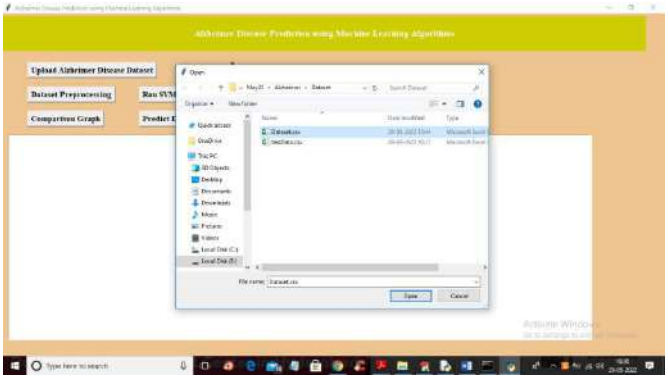
To implement this project we have designed following modules

1) Upload Alzheimer Disease Dataset: using this module we will upload dataset to application
2) Dataset Preprocessing: using this module we will read dataset and then encode non-numeric values in above dataset to numeric and then replace missing values with 0 and then split dataset into train and test part where application used 80% dataset for training and 20% for testing
3) Run SVM Algorithm: using this module we will train SVM with 80% training data and then calculate its accuracy by predicting 20% test data
4) Run Decision Tree Algorithm: using this module we will train decision tree with 80% training data and then calculate its accuracy by predicting 20% test data
5) Comparison Graph: using this module we will plot accuracy comparison graph between all algorithms
6) Predict Disease from Test Data: using this module we will upload test and then machine learning algorithm will predict weather test data is normal or contains Alzheimer symptoms

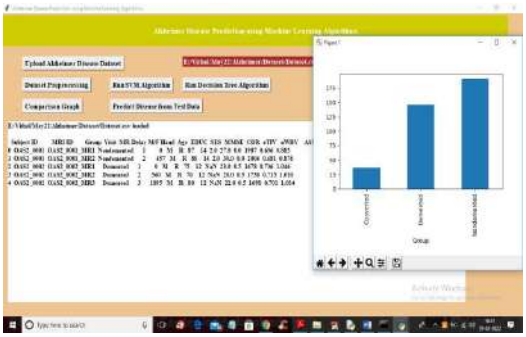
IV. RESULTS ANALYSIS

To run project double click on 'run.bat' file to get below screen

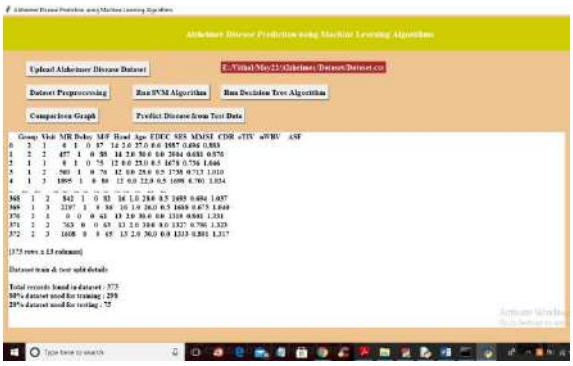
In above screen click on 'Upload Alzheimer Disease Dataset' button to upload dataset and to get below screen



In above screen selecting and uploading 'Dataset.csv' file and then click on 'Open' button to load dataset and get below output



In above screen dataset loaded and we can see dataset contains some NAN or missing values and contains nom-numeric data also so we need to preprocess dataset to remove such values and in above graph x-axis contains LABELS such as 'Converted' mean cured and Demented means presence of Alzheimer and Nondemented means normal and y-axis represents number of records found in that category in dataset. Now close above graph and then click on 'Dataset Preprocessing' button to process dataset and get below output



In above screen we can see all dataset values are converted to numeric and we can see dataset contains 373 records and 298 (80%) are using for training and 75 (20%) records are used for testing algorithm prediction accuracy. Now train and test data is ready and now click on 'Run SVM Algorithm' button to train SVM and get below output

In above screen with SVM we got 57% accuracy and now click on 'Run Decision Tree Algorithm' button to train Decision tree and get below output

In above screen with decision tree we got 82% accuracy and now click on 'Comparison Graph' button to get below output

In above graph x-axis represents algorithm names with different bar in different colour metrics such as accuracy, precision, recall and FSCORE and y-axis represents values. In above graph we can see Decision Tree got high performance and now close above graph and then click on 'Predict Disease from Test Data' button to upload test and get prediction output

In above screen in square bracket we can see patient test data and after =➔ arrow symbol we can see predicted output as NORMAL or CURED or Presence of Alzheimer disease.

V. 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.

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", Neuro Computing, 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", Neuro Computing, 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, "View aligned hyper graph 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 concord with amyloid-bPET and predict clinical progression: A study of fully automated 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 ability and dementia: Evidence from 11 cohort studies", Alzheimer's Dement2018;14:707–22