Alzeimer Disease Prediction Using Machine Learning Algorithms

[1]Prof. Satish Manje, [2]Miss.Chetana Vishe, [3]Miss.Niharika Dhanik, [4]Miss.Namita Sonavane

[1]Asst. Professor,[2,3,4]UG Student,[1,2,3,4]Computer Engg. Dept. Shivajirao S. Jondhle College of Engineering & Technology, Asangaon, Maharashtra, India. *[1]satishmanje93@gmail.com, [2]chetanamvishe@gmail.com, [3]dhanikniharika05@gmail.com, [4]namitasonavane1397@gmail.com*

Abstract- Alzheimer's complaint is the one amongst neurodegenerative diseases. This complaint is a grueling bone because there's no treatment for the complaint. Opinion of the complaint is done but that too at the after stage only. Therefore if the complaint is prognosticated before, the progression of the symptoms of the complaint can be braked down. This paper uses machine literacy algorithms to prognosticate Alzheimer's complaints using cerebral parameters. Civil population- predicated cohort provides a new occasion to make an automated trouble prophecy model grounded on individualities' history of health and healthcare beyond being threat vaticination models. Thus thetested the possibility of machine literacy models to prognosticate the future prevalence of Alzheimer's complaints (Announcement) using large-scale executive health data. These are trained and validated arbitrary timber, support vector machine and logistic retrogression to prognosticate incident Announcement in 1, 2, 3, and 4 posterior times. For prognosticating future prevalence of Announcement in balanced samples (bootstrapping), the machine literacy models showed reasonable performance in 1- time vaticination with AUC of0.775 and0.759, grounded on " definite Announcement" and " probable Announcement" issues, independently; in 2 times,0.730 and0.693; in 3- time,0.677 and0.644; in 4- time,0.725 and0.683. The results were analogous when the entire (unstable) samples were used. Important clinical features named in logistic retrogression included hemoglobin position, age, and urine protein position. This study, in the data-driven machine literacy model, may put a light on the mileage of grounded on large-scale executive health data in Announcement threat vaticination, and a better selection of individualities at threat for Announcement in clinical trials or early discovery in clinical settings will be enabled.

Keywords-vaticination model, deep literacy, machine literacy

I. INTRODUCTION

Alzheimer's complaint is caused by both heritable and environmental factors, which affects the brain of a person over time. The heritable changes guarantee a person will develop this complaint. This complaint breaks the brain kerchief over time. It occurs to people over age 65. However, people live with this complaint about 9 times, and about 1 in 8 people of age 65 and over have this complaint. MMSE (Mini-Mental State Examination) score is the main parameter used for a prophecy of the complaint. This score reduces periodically if the person is impacted. Those people having MCI have serious trouble of growing madness. When the fundamental MCI results in a loss of memory, the situation expects to develop madness due to this kind of complaint. Alzheimer's complaints are not curable. When this reaches higher state, complications like dehydration, malnutrition, or infection occur which leads to death. The opinion at the MCI stage will help the person to concentrate on a healthy approach to life, and good planning to take care of memory loss. It can be regarded to as sophisticated assumption of the whole series of events which can be stipulated to the origin and back to the form of its originality.

II. AIMS AND OBJECTIVE

a) Aim

The project aim is, to find The end of this design is to ameliorate the state of the being medical condition with consideration to this Alzheimer complaint. It's the one amongst neurodegenerative diseases. This complaint is challenging one because there's no treatment for the complaint. Opinion of the complaint is done but that too at the after stage only. Therefore, if the complaint is prognosticated before, the progression or the symptoms of the complaint can be braked down. This paper uses machine literacy algorithms to prognosticate Alzheimer complaint using cerebral parameters The author uses the" Croaker and

 DOI : 10.35291/2454-9150.2022.0153

Case" dataset to distinguish between bones who are diagnosed by the same and also try to train and test how accurate it can be honored by the machine, where it'll be useful to anyone who needs to explore and develop an Architecture of Machine Learning.

b) Objective

To see the best Pre-Trained Architecture model with the highest level of accuracy, and the lowest cost function in the optimal hyperparameter state. True challenge to build an automated system which equals human ability to recognize the disease. Humans are quite good identifying the disease but only after it has crossed the initial phrase. Human face recognition ability helps to develop a non-human face recognition system. The main focus is always to maximize the mutual distance between the hyper-plane and the original margin.

III. LITERATURE SURVEY

Paper 1: A Literature Review on the Heart and Alzheimer the Disease Prediction:

Ronghui Just. al suggested a brand new approach of identifying alzeihmer in the pateints. The idea was to use trained models of autoencoders which then produces a detailed discovery of the early Advertisement, a deep network-suchlike autoencoder is used where connections of various functional of the networks are constructed and are susceptible to Advertisement and MCI. The targeted autoencoder model is used which is a three-concentrated model which gives intellectual growth of the nervous system also excerpts brain network attributes completely (1). When a finite amount of data cases across-verification was executed mainly to avoid the overfitting complication. To distinguish between different subjects, a system called multi stage classifier was proposed by Mr. K. R. Kruthikaet.al, which consists of various algorithms such as K neighbor, Support Vector Machine and Naive Bayes. [4].

Paper 2: ML for diagnosing Early Alzheimer's Disease :

PSO (particle swarm optimization) which is a more selective feature that was forced to find the best features. Naturally the process of retrieving an image requires two stages: the first step involves generating features to reproduce the image in question and the latter step combining those features with those already collected on the website [2]. The PSO algorithm is used to select the best biomarkers showing AD or MCI taken from the Alzheimer's Neuroimaging Initiative (ADNI) data database. MRI scans are screened first after taking them to the website. Feature selection includes volumetric measurements and dimensions. A list of the best features was then obtained in the PSO algorithm [2]. Gaussian Naïve Bayes Machine, K- Close Neighbor, Vector Support Machine was used to distinguish between subjects. Here stage 2 is used where the first GNB phase is used to distinguish items between AD, MCI and NC and in later stages SVM and KNN were used to analyze the object based on the first performance [5].

Paper 3: Machine Learning Algorithms for the diagnosis of Alzheimer's and Parkinson's Disease

Control Based Image Retrieval was used for reacquiring images from the database. Ruoxuan Cuiaet.al, proposed a model where longitudinal analysis is performed on successive MRI and is essential to design and cipher the elaboration of complaint with time for the purpose of more precise opinion (3). The factual 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's for successive intervals in a gap of six months are taken as inputs from ADNI database (3). Also point literacy is done with the 3D Convolutional Neural Network. The CNN is followed by a pooling subcaste and have numerous ways for pooling, like collecting mean value else the minimal, or definite sequence of neuron in the section. But for studying the characteristics, the convolutional operation of $2 \times 2 \times 2$ is applied so that a direct combination is studied for pooling of neurons (3). The completely connected subcaste has neurons that produce affair of all neurons in a direct combination, which are taken from antedating subcaste and also is moved through nonlinearity. Eventually for the last completely connected, a soft max subcaste is particularly used and also tuned finely for back-propagation to prognosticate the class probability (3). The result of each knot varies from 0 to 1, and the aggregate of bumps will always be 1. Eventually the bracket includes the deep network construction including the 3D CNN training and RNN model training. Also, the results of completely connective layers are directly counterplotted using a soft max function (3). The original parameters that were trained by both 3 dimensional CNN and the RNN network are established and also only the upmost completely connective subcaste parameters and the soft max subcaste that was used for vaticination are acclimated so that the dimensional and longitude features were united for distinct identification [7].

IV. EXISTINGSYSTEM

For earlier examination of Alzheimer's disease, deep learning methods with the network-brain and significant information clinical like age,gender of the subjects and ApoE gene was taken into consideration. Autoencoder is used in deep network to produce a detailed discovery of the early AD where functional connections of the networks are constructed and are susceptible to AD and MCI. ADNI dataset is taken is taken into consideration. The classification of model consists of the early diagnosis, initially preprocessing of raw R-fMRI is done. Then, the time series data (90 ×130matrix) is obtained and that indicates blood oxygen levels in each and every region of the brain and changes over a long period. Then, a brain network is built

 DOI : 10.35291/2454-9150.2022.0153

and transformed to a 90 ×90 time series data correlation matrix. The targeted autoencoder model is used that is a three layered model which gives intellectual growth of the nervous system then extracts brain networks attributes completely

When a finite amount of data cases is taken, k-fold cross verification was implemented mainly to avoid the over fitting complication[4].

V. COMPARTIVE STUDY

Sr. No.	Paper Title	Author Name	Method	Advantage	Disadvantage
1.	A Literature Review on the Heart and Alzheimer the Disease Prediction	Jayashri J. Patil, Nilesh Vani, K.R. Kruthika	Concentrated Autoencoder method is used in this paper.	The targeted autoencoder model is used which is a three-concentrated model which gives intellectual growth of the nervous system also excerpts brain network attributes completely.	Although the autoencoder model predicts the disease but it has certain conditions upon which it can give successful result otherwise not. So, this can lead to a serious medical failure.
2.	ML for diagnosing Early Alzheimer's Disease	Aunsia Khan, Muhammad Usman	Particle Swarm Optimisations method is used in this paper.	The PSO algorithm is used to select the best biomarkers showing AD or MCI taken from the Alzheimer's Neuroimaging Initiative (ADNI) data database.	The model based on the Particle swarm optimisation is squite static and can't be used reliably as small particles of the disease can't be swarmed up and optimised
3.	Machine Learning Algorithms for the diagnosis of Alzheimer's and Parkinson's Disease	Nancy Noella, Ruoxuan Cuoia, J. Priyadarshini	CNN and RNN is used here in this paper.	CNN training and RNN model training were used which are some of the advanced algorithms in the machine learning world and so the end results were quite reliable.	Convolutional models are best upon images but they provide false-positive result when the images contain duplicate similar looking sigmas. And RNN can't help to improve this. Hence this model was not adopted.

Table no.1 Comparative Analysis

VI. PROBLEM STATEMENT

The alzeihmer disease has been there for a decades and the main problem of this disease is it does not get identified in the early stage. Here, Thus to find an effective way of identifying the disease using the advancement of science and technology. Using various Machine Learning and Deep Learning algorithms using this techniques, Thus efficiently identify the disease in the early stage and the problem is solved.

VII. PROPOSED SYSTEM

In the projected a way known as time period classifier by victimisation machine learning algorithms to Support Vector Machine, Naive Thomas Bayes and K-nearest neighbor to classify between totally different subjects. PSO (particle swarm optimization) that could be a technique that best selects the options was enforced to get best options. Naturally image retrieving method needs 2 stages: the primary stage involves generating options, so it reproduces the question image and so later step correlate those options with those already gathered within the information. The PSO algorithmic rule is employed to pick the best biomarkers that show AD or MCI. The information is Alzheimer's disease Neuroimaging Initiative (ADNI) information. The magnetic resonance imaging scans area unit preprocessed 1st once taking from the information. The feature choice includes volumetric and thickness measurements. Then the optimum feature lists were obtained from PSO algorithmic rule. The Gaussian Naïve Thomas Bayes, K- Nearest Neighbor, Support vector machine was accustomed distinguish between the topics. Here a pair of stage classifiers was used wherever within the initial stage GNB classifier was accustomed classify the objects between AD, MCI and American state and in later stages SVM and KNN were accustomed analyze the thing supported the performance of the initial one. Management primarily based Image Retrieval was used for retrieving pictures from the information

VIII. ALGORITHM

The general idea of working of proposed system algorithm is given as follow:

Step 1: Start

Step 2: Make functions to calculate correlation and coefficient

Here it checks for the if-else statement block of code for the POST and loggin in purposes. The render request are made

 DOI : 10.35291/2454-9150.2022.0153

to check the password with the already determined database which has the credentials.

def adminloginaction(corr, coeff):

Step 3: Make functions to calculate Gain of Information

The definition here follows GET and POST methods and makes use of return statements.

def (GOI):

patientModel.objects.filter(id=uname).update(status=status)

qs = patientModel.objects.all()

return render(request,
'admin/patientsdetails.html',{"object": qs})

Step 4 : Calculate Entropy

The entropy is widely calculated based on the

Now,

$$\text{Entropy} = \sum_{i=b}^{4} p^{ix} \log(pi)$$

Step 5 : Make Machine Learning model that will predict the alzeihmer disease

The class here similarizes the attention with respect to the doctor's details and the database that contains all the information about the patient.

class doctorForm(forms.ModelForm)

Step 6 : Predict the disease and exit

IX. MATHEMATICAL MODEL

The mathematical model starts with the process of finding correlation andcoefficent between the selected features.

A) Correlation and coefficient :-

For two variables X and Y, the covariance is given by :

$$p(X,Y) = \text{Cov}(X,Y)/\delta X \delta Y$$

B) Gain of Information :-

It is obtained by subtracting the lower node of entripy by upper node.

Gain (D) = I(S1…..Sn)-E(feature D)

C) Entropy = (p(0) * log(P(0)) + p(1) * log(P(1)))

If C = 0, means less complex boundary as classifier would be not penalized by slack, as a result, the optimum hyperplane can use it anywhere and accept all large misclassifications. And as a result, the decision boundary would be linear and under fitted.

$$w^* = arg_w max \frac{1}{||w||_2}, s.t. \min_n y_n \left[w^T \phi(x_n) + b \right] = 1$$

If C = infinitely high, then even small slacks would be highly penalized and the classifier can't afford to misclassify points and therefore overfitting. So parameter C is important.

$$y_n \left[w^T \phi(x) + b \right] = \begin{cases} \geq 0 \; if \; correct \\ < 0 \; if \; incorrect \end{cases}$$

Concept is basically to get rid of Φ and hence rewrite Primal formulation in Dual Formulation known as the dual form of a problem and to solve the obtained constraint optimization problem with the help of Lagrange Multiplier method

And when calculating for weights outside independence, Obtain Primal and Determine Lagrangian form from primal

$$Primal\ Form\ for\ non-perfect\ seperation$$
$$min_{w,b,\{\beta_n\}} \frac{1}{2}||w||_2 + C \sum_n \beta_n$$
$$s.t. \; y_n \left[w^T \phi(x_n) + b \right] \geq 1 - \beta_n \; ; \forall n$$
$$\beta_n \geq 0 \; , \forall n$$

Finding alzeihmer visualisation,

$$\frac{\partial L}{\partial w} = w - \sum_n \alpha_n y_n \phi(x_n) = 0 \Rightarrow w = \sum_n \alpha_n y_n \phi(x_n)$$
$$\frac{\partial L}{\partial b} = \sum_n \alpha_n y_n = 0 \Rightarrow \sum_n \alpha_n y_n = 0$$
$$\frac{\partial L}{\partial \beta_n} = C - \alpha_n - \lambda_n = 0 \Rightarrow C - \alpha_n - \lambda_n = 0$$

Which is gurther given by:

$$Instead\ of\ y_n \left[w^T \phi(x_n) + b \right] > 0 \; ; \forall n$$
$$Equation\ becomes\ y_n \left[w^T \phi(x_n) + b \right] \leq 0 \; , \exists n$$

Finally, considering positive negative preddiction and confusion matrix being the sole autony, the correct outlier can be calculated. And the mathematical formula is given as follows :

Lastly, the final weights for the prediction model can be calculated with the mathematical formula.

X. SYSTEM ARCHITECTURE

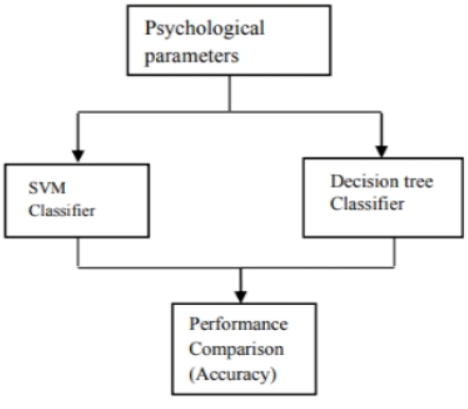

Fig.1: System Architecture

 DOI : 10.35291/2454-9150.2022.0153

Description: The physiological parameters are defined as two types noted as follows:

1. SVM classifier.

2. Decision tree classifier.

SVM Classifier : The particular liner classier which support vectors. They are minimized as to perform structural risk. Mainly they are based on maximize the margin principle. But it becomes more stronger when it works on complex datasets. Decision tree classifier: Internal mode is represented as datasets features decision rules is branch and each node is represented by outcomes. Due to accuracy decision tree is used universally in this scenario stated.

XI. ADVANATGES

1) The main advantage of SVM is that it can distinguish linear and non-linear objects.

2) It detects the data or features without human intelligence.

3) It helps to solve the complex real problems.

4) It can achieve more accuracy in a less quantity of data.

5) The model aids the purpose of detection of the disease in the initial phase itself proving out to be a life saver for multiples.

XII. DESIGN DETAILS

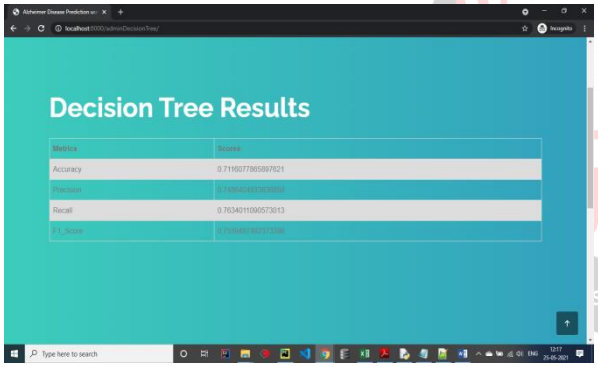

Fig.2 Decision Tree

XIII. CONCLUSION

Thus,we have tried to implement the paper "K.R.Kruthika,Rajeswari, H.D.Maheshappa, "Multistage classifier-based approach for Alzheimer's Disease prediction and retrieval",Informatics in Medicine Unlocked, 2019" is successfully implemented and gives greater prediction accuracy results. The model is acceptably accurate such that it detects the patient sufferring from alzeihmer disease. The model needs refinement though, but then it is readily result giving and can be used in production.

REFERENCE

[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 thinking ability and dementia: Evidence from group 11 studies", Alzheimer's Dement2018; 14: 707- 22.

[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.

 DOI : 10.35291/2454-9150.2022.0153