BBRC
Bioscience Biotechnology
Research Communications

Alzheimer Disease Forecasting using Machine Learning Algorithm

Malavika G[1], Rajathi N[2], Vanitha V[3] and Parameswari P[4]
[1]PG Scholar, Department of Information Technology, Kumaraguru College of Technology, Coimbatore, India.
[2,3]Professor, Department of Information Technology, Kumaraguru College of Technology, Coimbatore, India.
[4]Assistant Professor (SRG) Department of MCA, Kumaraguru College of Technology, Coimbatore, India.

ABSTRACT

Alzheimer disease is a neurodegenerative disease that makes a gradual disorder of human brain cells and it leads to degenerate the cells away and die. In India more than one million cases per year are affected by this disease. The most common in people over the age group of above 65. There is no treatment for this disease to cure, but now a day's medications are available to temporarily decline the process of disease. The primitive detection of this disease may help the doctors, physician, and other family members to treat them in a better way. The objective of the proposed system is to offer a fast, early and cost-efficient method to detect disease in premature period. Machine learning is the blooming field in the healthcare industry, so by using the machine learning techniques the disease will get forecast in the earlier stage. The techniques are K-Nearest Neighbor, Adaboost Classifier, Support Vector Machine, Logistic Regression, Decision Tree Classifier and Random Forest classifier. Among these algorithms, the best prediction accuracy is produced by the Random Forest algorithm.

KEY WORDS: ACCURACY, ALZHEIMER'S DETECTION, MACHINE LEARNING, PRIMITIVE DETECTION.

INTRODUCTION

Alzheimer's disease is a cause of dementia. Dementia leads to memory loss, thinking ability due to some of the brain disease. Alzheimer is one of the brain diseases that causes Dementia. This disease causes the mini strokes in the brain and that occurs the gradual cell destruction and the nerve disorder in the brain. A person who is affected by the disease may not be aware of the strokes due to the minor attacks and that occurs without any perception. It occurs at individual losses.

This disease mainly affects the age of 65, it is not possible to calculate that age nowadays, it can occur as early as 50 unfortunately, but the early 50 cases are rare then the 65 above. The people who are affected early are usually aware of the changes in them. Their new deviations and memory loss affect them deeply, and they always forget the things and they are not able to handle their things as when they are in normal condition. They feel some difficulty to talk and to use the words, while talking with family members, relatives, friends etc.... This leads them to talk less and this advanced stage leads to forgetting the close family members. When they release that   they are not functioning as well as they did formerly, they become depressed.

Overall analysis the maximum Americans suffer from AD's. There are 4.5 million American people suffered by this disease. Research says that these will get expected to increase in the year of 2050 as 14 million. The diagnosis of Alzheimer's disease takes a long process that has an awful effect on the patients with the disease and

15

their families. Analysis of AD's is not simple and easy. It cannot be done without any proper study of brain sample tissues.

There is no proper treatment for this disease to cure, it may reduce the decline but not cure the disease. If the early detection of the disease is done that will be helpful for the physician, family members all other close to them etc. So, that the Machine learning techniques are used to diagnose the disease earlier. There are five techniques used here to find the best accuracy. The techniques are K-Nearest Neighbor, Adaboost Classifier, Support Vector Machine, Logistic Regression, Decision Tree Classifier, Random Forest classifier. Among this the best and high accuracy detector can be determined by using python code implementation.

MATERIAL AND METHODS

Various studies with respect to diagnosis of Alzheimer's disease are discussed. Joshi, S, et.al., uses a various Machine Learning algorithm to categorize the AD's (Alzheimer' disease) and PD's (Parkinson's disease) with high accuracy classifier, by using the major risk factor. By using Fluorodeoxyglucose, Positron Emission Tomography and Pittsburg Compound B imaging techniques Illan, I.A, et.al., compare the forecast accuracy on early AD's (Alzheimer's Disease). Image-based classification method is used by Dong Hye Ye et.al., to classify the brain MRI scans with MCI. A semi-supervised classifier patterns are used to achieve a high sensitivity.
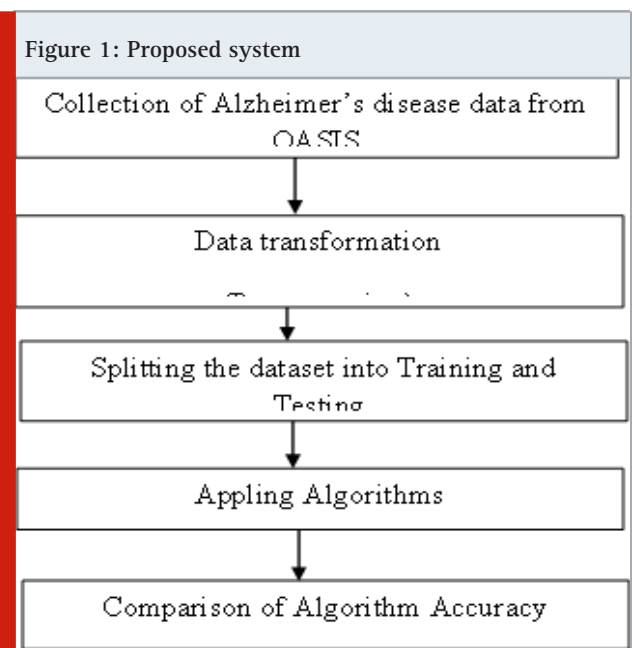
Data mining techniques are used in Aunsia Khan et.al., proposed system. They present an assessment, study and estimation of the current work in the initial detection of Alzheimer using Machine Learning algorithm. Ammarah Farooq et.al., proposed a four-way classifier to mainly classify the AD, MCI, LMCI and healthy persons. Core focus of this work is to classify the stages of disease. By using the deep learning technique. Machine Learning algorithms stand to detect the AD in Arpita Raut et.al.,. This planned method abstracts the surface and figure topographies from the MRI scans. Neural Network is used for detection of various stages of AD's. Karl Backstrom et.al., uses a deep convolutional neural network to offer an effective and simple 3D convolutional network architecture to achieve a high performance of AD detection.

The Deep Convolutional Neural Network is used for detecting Dementia and AD from MRI scan in H. M. Tarek Ullah et.al.,. This paper discussed the fast, costs less and more reliability. Priyanka Lodha et.al., paper mainly focused on using the neuroimaging techniques like CT, MRI, PET, EEG data, to detect Alzheimer in its primitive stage using ML. The assessment of ML Technique is done in Mohamed Mahyoub et.al., proposed system. They rank AD risk factors on clinical assessment data. Mohamed Mahyoub et.al., investigates five different classifiers in the risk factor of AD's data for better accuracy. Gokce UYSAL et.al., evaluate the early forecast of dementia in

AD by using machine learning algorithms. Here, they consider the hippocampus brain region of dementia. This approach can be useful for separating the patients with AD and CN. The age of the brains of individuals can be very useful in many applications. Masoumeh Siar et.al., has greatly paid to forecasting and avoiding early deaths in the medical field. This paper has been done by using (DL).

J. Neelaveni et.al., used machine learning algorithms to guess Alzheimer disease using psychological parameters like age, number of visits, MMSE and education. By using the parameters as input the algorithms are applied. The SVM and Decision Tree algorithm are used and the comparison is done by the accuracy. The best accuracy detector is SVM. Aakash Shah et.al., comprises the complete study and precision of various ML techniques. Voting Classifier Algorithm is used for early discovery of Alzheimer Disease, and to removes the possibility of inaccuracies in the result. Rajathi et. al successfully applied machine learning methods for disease prediction.

Proposed Methodology: Machine learning plays a significant title role in the health care business. There is a large amount of database provided by the healthcare domain to develop an advanced and scientific method to diagnose the disease in an early stage. So, here some of the machine learning algorithms are used to forecast the disease and to find the best accuracy provider among these algorithms. The algorithms are Logistic Regression with null values, Logistic Regression without null values, Support vector machine, Decision Tree, Adaboost. The python code is used for the implementation. The proposed system is pictorially represented in Figure I.

Figure 1: Proposed system

Collection of Alzheimer's disease data from OASIS

Data transformation

Splitting the dataset into Training and Testing

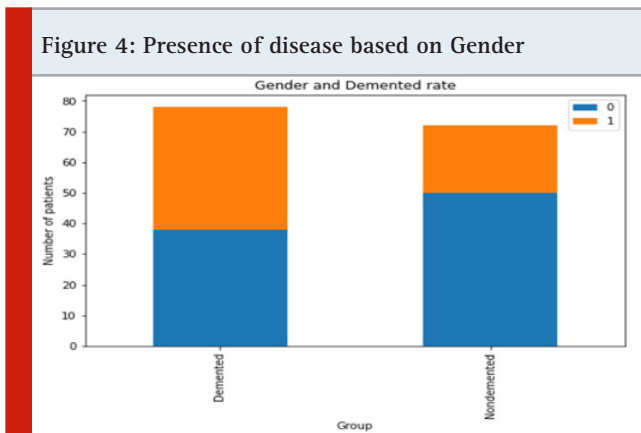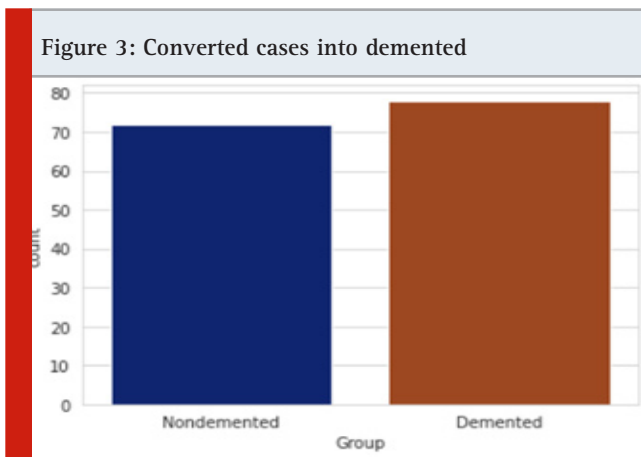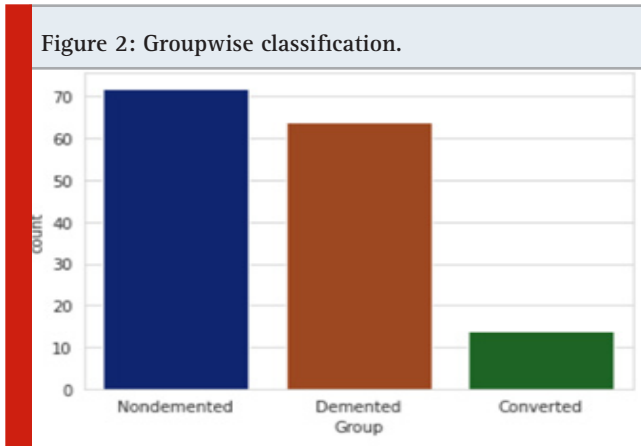Appling Algorithms

Comparison of Algorithm Accuracy

Dataset: The Alzheimer's disease dataset is collected from OASIS, which is offered on their website. It can be applied and used for the purpose of training and

executing the algorithms to identify the disease impact. Here, the longitudinal MRI data are used. The dataset consists of a longitudinal MRI data of 150 subjects aged 60 to 96.

Data pre-processing: Data pre-processing is done to remove the rows with missing values, splitting Training and Testing data and to cross validate the data.

Figure 2: Groupwise classification.

Figure 3: Converted cases into demented

Figure 4: Presence of disease based on Gender

RESULTS AND DISCUSSION

The overall objective of this paper is to predict more accurately the early detection of Alzheimer disease. By using technologies named K-Nearest Neighbor, Adaboost

Classifier, Support Vector Machine, Logistic Regression, Decision Tree Classifier, Random Forest classifier. From the result it's been seen that the Random forest and Adaboost gives more accuracy as compared as other techniques.

Figure 5: Age vs count.

Figure 6: Age vs Group
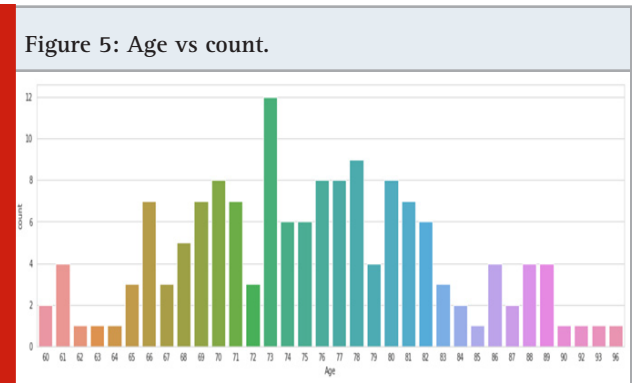
The fig -2 shows that the number of Non demented, Demented and Converted cases. It clearly shows that Nondemented are higher when compared to other, the demented group be less when compared. The converted shows as demented, they are Non demented to Demented case. Fig 3 shows the converted cases into demented. The above Fig- 4 clearly says that the difference between the Demented and Non-Demented patients with respective Gender (0= Female(blue), 1= Male(orange)).

This clearly shows that greater number of are in Demented condition as compared to female. In Non-Demented Condition Females are higher than Male. The fig-5 shows that the maximum age of diseased cases. The age between 68 to 83 be the most affected case in Alzheimer's disease. Fig.6 shows that the gradual increase in 68 age and it's in the peak of mid-70 to 80 and it gradually decreased in the age of after 80.

The performance of the classification models on the test data was represented in Table – 1. The various performance parameters Precision, Recall, F1 scores for both male and female of various models are presented in Table 2.

17

Table 1. Performance of Algorithms

Methodology	Precision		Recall		F1 Score	
	Female(0)	Male(1)	Female(0)	Male(1)	Female(0)	Male(1)
Logistic Regression Decision Tree	0.69	0.79	0.79	0.70	0.74	0.74
Classifier	0.75	0.84	0.83	0.77	0.79	0.80
K-Nearest Neighbor	0.61	0.77	0.81	0.55	0.69	0.64
Support Vector Machine	0.71	0.87	0.88	0.68	0.79	0.77
AdaBoost	0.79	0.82	0.79	0.82	0.79	0.82
Random Forest	0.80	0.88	0.87	0.82	0.83	0.84

Table 2. Classification Performance

Methodology	Classification Accuracy
Logistic Regression	74.1%
Decision Tree Classifier	79.4%
K-Nearest Neighbor	66.9%
Support Vector Machine	77.6%
AdaBoost Classifier	80.3%
Random Forest Classifier	86.8%

From the results obtained, the Random Forest classifier gives the high accuracy than the other models

CONCLUSION

In this paper, various machine learning algorithm were used to predict the Alzheimer disease at early stage. The results obtained shows that the Random forest classifier gives the best performance when compared to other methods. The future work is to apply hybrid approaches and their performances to be studied.

REFERENCES

Aakash Shah, Dhruvi Lalakiya, D., Shekha Desai., Shreya, and Vibha Patel., 2020, June. Early Detection of Alzheimer's Disease Using Various Machine Learning Techniques: A Comparative Study. In 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184) (pp. 522-526). IEEE.

Aunsia Khan. and Muhammad Usman., 2015, November. Early diagnosis of Alzheimer's disease using machine learning techniques: A review paper. In 2015 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K) (Vol. 1, pp. 380-387). IEEE.

Ammarah Farooq., Syed Muhammad Anwar., Muhammad Awais and Saad Rehman., 2017, October. A deep CNN based multi-class classification of Alzheimer's disease using MRI. In 2017 IEEE International Conference on Imaging systems and techniques (IST) (pp. 1-6). IEEE.

Arpita Raut and Vipul Dalal., 2017, July. A machine learning based approach for detection of Alzheimer's disease using analysis of hippocampus region from MRI scan. In 2017 International Conference on Computing Methodologies and Communication (ICCMC) (pp. 236-242). IEEE.

Dong Hye Ye., Pohl, K.M. and Davatzikos, C., 2011, May. Semi-supervised pattern classification: application to structural MRI of Alzheimer's disease. In 2011 International Workshop on Pattern Recognition in NeuroImaging (pp. 1-4). IEEE.

Gokce UYSAL and Mahmut OZTURK., 2019, October. Using Machine Learning Methods for Detecting Alzheimer's Disease through Hippocampal Volume Analysis. In 2019 Medical Technologies Congress (TIPTEKNO) (pp. 1-4). IEEE.

H. M. Tarek Ullah., Dr. Dip Nandi., and Zishan Ahmed Onik., 2018, April. Alzheimer's Disease and Dementia Detection from 3D Brain MRI Data Using Deep Convolutional Neural Networks. In 2018 3rd International Conference for Convergence in Technology (I2CT) (pp. 1-3). IEEE.

Illan, I.A., Gorriz, J.M., Ramirez, J., Chaves, R., Segovia, F., López, M., Salas-Gonzalez, D., Padilla, P. and Puntonet, C.G., 2010. Machine learning for very early Alzheimer's disease diagnosis; a 18 F-FDG and pib PET comparison. In IEEE Nuclear Science Symposuim & Medical Imaging Conference (pp. 2334-2337). IEEE.

Joshi, S., Shenoy, D., GG, V.S., Rrashmi, P.L., Venugopal, K.R. and Patnaik, L.M., 2010, February. Classification of Alzheimer's disease and Parkinson's disease by using machine learning and neural network methods. In 2010 Second International Conference on Machine Learning and Computing (pp. 218-222). IEEE.

Karl Bäckström., Mahmood Nazari., Irene Yu-Hua Gu and Asgeir Store Jakola., 2018, April. An efficient 3D deep convolutional network for Alzheimer's disease

diagnosis using MR images. In 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018) (pp. 149-153). IEEE.

Mahyoub Mohamed., Martin Randles., Thar Baker and Po Yang., 2018, June. Effective Use of Data Science Toward Early Prediction of Alzheimer's Disease. In 2018 IEEE 20th International Conference on High Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/DSS) (pp. 1455-1461). IEEE.

Mohamed Mahyoub., Dr. Martin Randles., Dr. Thar Baker and Dr. Po Yang.,., 2018, September. Comparison analysis of machine learning algorithms to rank alzheimer's disease risk factors by importance. In 2018 11th International Conference on Developments in eSystems Engineering (DeSE) (pp. 1-11). IEEE.

Masoumeh Siar and Mohammad Teshnehlab, M and K.N. Toosi., 2019, October. Age Detection from Brain MRI Images Using the Deep Learning. In 2019 9th International Conference on Computer and Knowledge Engineering (ICCKE) (pp. 369-374). IEEE.

Neelaveni, J. and Geetha Devasana, M.S., 2020, March. Alzheimer Disease Prediction using Machine Learning Algorithms. In 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS) (pp. 101-104). IEEE.

Priyanka Lodha., Ajay Talele and Kishori Degaonkarm, 2018, July. Diagnosis of Alzheimer's Disease using Machine Learning. 2018 Fourth International conference on Computer Communication Control and Automation (ICCUBEA).

Rajathi, N., Kanagaraj, S., Brahmanambika, R. and Manjubarkavi, K., 2018. Early detection of dengue using machine learning algorithms. International Journal of Pure and Applied Mathematics, 118(18), pp.3881-3887.