Machine Learning Models for Alzheimer's Disease Detection Using OASIS Data

Rajesh Kumar Shrivastava, Simar Preet Singh, and Gagandeep Kaur

Abstract Early Prediction of Alzheimer's disease is a challenging task for researchers to contribute. Dementia is the simplest symptom of Alzheimer's disease. Nowadays, most researchers apply Artificial Intelligence to discover mental disorders like Alzheimer's, which mostly affect the old age population worldwide. In Alzheimer's disease, the brain is under neurodegenerative changes. As our population ages, more people will be affected by diseases that impact memory functionalities. These repercussions will profoundly affect the person's social and financial fronts. It is difficult to predict Alzheimer's disease in its early stages. The Medication given early in Alzheimer's disease is more effective and has fewer minor side effects than treatment given later. To find the optimum parameters for Alzheimer's disease prediction, researchers used a variety of algorithms, including Decision Trees, Random Forests, Support Vector Machines, Gradient Boosting, and Voting classifiers. Predictions of Alzheimer's disease are based on data from the Open Access Series of Imaging Studies (OASIS). The performance of machine learning models is tested using measures such as Precision, Recall, Accuracy, and F1-score. Clinicians can use the proposed classification approach to make diagnoses of these disorders. With these ML algorithms, it is extremely beneficial to reduce annual Alzheimer's disease death rates in early diagnosis. On the test data of Alzheimer's disease, the proposed work demonstrates better results, with the best validation average accuracy of 80%.

Keywords Machine learning · Healthcare · Random forest

R. K. Shrivastava · S. P. Singh (✉)
Bennett University, Greater Noida, India
e-mail: dr.simarpreetsingh@gmail.com; simarpreet.singh@bennett.edu.in

R. K. Shrivastava
e-mail: rajesh.shrivastava@bennett.edu.in

G. Kaur
MITS, Gwalior, India
e-mail: gagan873@gmail.com; gagandeep@mitsgwalior.in

D. Koundal et al. (eds.), *Data Analysis for Neurodegenerative Disorders*,
Cognitive Technologies, https://doi.org/10.1007/978-981-99-2154-6_6

1 Introduction

Alzheimer's disease is a degenerative neurologic condition in which the brain shrinks (atrophy) and brain cells die. Alzheimer's disease is the most frequent form of dementia, defined as a progressive loss of cognitive, behavioral, and social abilities that impair a person's capacity to operate independently [1, 2]. Proper treatment helps the patients to reduce the disease symptoms temporarily. Correct medicines can assist patient by preserving their independence and mental functionality. Alzheimer's disease is known as no cure disease it effects patient's brain. In this illness the patient suffers loss of memory that can lead to mortality in the advanced stages of illness.

Alzheimer's disease has its characterization by loss of memory. Early identifications of this disease include recalling recent events or discussions [3]. Memory functionality also get effected as the disease progressed. A person suffering from Alzheimer's may notice difficulties in remembering things and managing thoughts. These changes in the brain are caused following problems:

1. Memory: This illness cause badly memory loss, making the infected person difficult to operate at work or home. An Alzheimer's patient may experience the following symptoms:
   - Replicate statements and queries as needed.
   - Remember about discussions, appointments, or events and remember about them.
   - Frequently misplaced belongings, keeping them in strange places.
   - Family members' names and ordinary objects are eventually forgotten.
   - Facing problems in recognizing items, express thoughts, or in conversations.
2. Reasoning and thinking: The disease of Alzheimer's impairs concentration and thinking, especially when abstract notions like numbers are involved [4]. Multitasking is particularly difficult, including managing finances, paying payments on time, and cheque books balancing can be a difficult task. This all is because a person who is suffering from Alzheimer's disease may eventually lose the ability to recognize and cope with the numbers.
3. Making decisions and judgments: Alzheimer's disease impairs a person's capacity, making sound assessments and decisions in everyday situations becomes very difficult. As a result, a person can make bad or unusual decisions that lead to social encounters or improper dressing that is inappropriate for the present weather. Even responding successfully to ordinary situations like a stove on fire or unexpected driving conditions seems very difficult to a person suffering from this disease.
4. Organizing and carrying out routine activities: Over time, the person with such a disease starts facing problems with daily routine that involve normal tasks, such as playing games, kitchen work driving etc. Extending this, in the advanced stage of Alzheimer's disease patient forget daily work like bathing, dressing, etc.

5. Personality and behaviour change: As Alzheimer's disease causes changes in the brain functionality, this impacts the mood and behaviour of a person. Such a person shows the following issues.

Apathy, Depression, Social isolation, Mood Swings, Others' mistrust, changes in sleeping patterns, Aggressiveness, irritability, Wandering, lost inhibitions, and Delusions.

The daily diagnosis of an Alzheimer's patient is difficult. But the patient needs daily monitoring, and the family members are also curious about the patient's status. Earlier detection of Alzheimer's disease was difficult. In most cases, it is diagnosed after death. But with the help of artificial intelligence, it is possible to detect Alzheimer's disease in its early stage. This chapter compares various methods to detect Alzheimer's using Machine Learning (ML) techniques.

Brain images are not good for Alzheimer's disease. It is mostly used to study strokes, trauma, or tumors to understand cognitive change. This chapter uses Magnetic resonance imaging (MRI) data from the Open Access Series of Imaging Studies (OASIS) [5–7]. MRI provides detailed image of brain. MRI take help of radio waves and a magnetic field to capture brain status.

This work helps in the early detection of Alzheimer's patients. This is very useful for recovering the disease at its initial phase itself. The work in this paper explains various machine learning algorithms [8–13]. In addition, the work evaluates different machine learning algorithms on MRI data to identify Alzheimer patients, and each algorithm shows its accuracy in the identification process. As a result, parameter tuning has been carried out by considering the machine learning approach with the highest accuracy to be the best optimal feasible solution for the dataset.

2 Related Work

Bari et al. [14] described the problem of Alzheimer's disease very well. We also used the same dataset (OASIS) used by the author. The Author used various ML techniques and compared the result. The data set Bari et al. [14] used s small but contains useful information that we reuse in our proposed work.

Moradi et al. [15] Experimented with Moderate Cognitive Impairment (MCI) to use MCI as a bridge between age-related cognitive decline and Alzheimer's disease. The Authors demonstrate the usage of a machine learning-based MRI biomarker. According to the paper, their aggregate biomarker attained a tenfold cross-validation AUC score of 0.9020 in differentiating between progressive MCI (pMCI) and stable MCI (sMCI). The Author used semi-supervised learning to implement sMCI/pMCI classification on data from AD patients and normal controls rather than MCI patients. Regularized logistic regression was used to choose features. To avoid possible confusion between changes owing to AD and those related to normal aging, they eliminated aging effects from MRI data before classifier training. Finally, they created an aggregate biomarker using a random forest classifier by first learning a distinct MRI biomarker and then merging age and cognitive data from MCI participants.

Zhang et al. [16] focused on Eigen brains and machine learning; this research suggests a new computer-aided diagnosis (CAD) approach for MRI brain imaging. They employ critical slices from the 3D volumetric data acquired by the MRI to generate Eigen brain pictures based on EEG data in their method. After that, they used kernel support vector machines with several kernels trained using particle swarm optimization. Their polynomial kernel (92.36%) was more accurate than their linear and radial basis function kernels (91.47% and 86.71%, respectively).

Magnin et al. [17] suggested a method based on support vector machine (SVM) classification of whole-brain anatomical MRI to distinguish patients with Alzheimer's disease from old controls in this publication. The researchers parceled three-dimensional T1-weighted MRI data from 16 patients with Alzheimer's disease and 22 elderly controls into regions of interest (ROIs). The grey matter properties of these ROIs were subsequently used to classify participants using an SVM algorithm. The classifier has a mean accuracy of 94.5% based on their findings. The fact that they should have accounted for age-related changes in the grey matter and worked with a tiny data set could be one of their technique's drawbacks.

Khan et al. [18] reviewed the four most common deep learning and machine learning methods for brain disease identification. Some significant insights into modern ML/DL approaches are revealed in his paper. The most difficult part of the analysis is feature extraction, identification, and classification techniques using ML/DL methods. He also discussed methods to increase classification precision. The author also focuses on the quantity of training data that must be increased. The author used hybrid algorithms and combined supervised learning with ML and DL. Khan et al. [18] also identified the shortcomings of current ML/DL-based methods in identifying different types of brain disorders. The article offers a debate centered on a collection of open research problems to create efficient AI-based medical systems Incorporation of XAI methods is the ultimate goal for applications. This will aid healthcare workers in developing self-assurance, and AI-based solutions will become a treatment for individuals with neurological diseases (Table 1).

Saratxaga et al. [19] proposed a method using balance accuracy calculation. They achieved 88% of accuracy. The author deployed BrainNet 2d and BrainNet3D techniques for classification; both methods belong to deep learning. The author used OASIS-1 and OASIS-2 datasets for training and testing purposes. The authors used 2D and 3D architecture in this paper and applied a subject-level approach using transfer learning. But this approach needs more resources and high computing power. The result could be promising better. There is a scope for modification in the method used.

Sudershan et al. [20] discussed the problem of Alzheimer's disease and the nonavailability of drugs. The authors also discussed the importance of early detection of disease and the problem with the available datasets. The authors experimented with mild cognitive impairment, structural magnetic resonance, import vector machine,

Table 1 Related work

Author	Method	Dataset	Accuracy (%)
Khan et al. [18]	Machine learning and deep learning models	Image modality	85
Saratxaga et al. [19]	Deep learning and image processing technique	OASIS dataset	88
Sudharsan et al. [20]	Machine learning models	ADNI dataset	75
Helaly et al. [21]	Convolutional neural networks	ADNI dataset	93
Shakila Basheer et al. [22]	Deep neural networks	OASIS dataset	92
Martinez-Murcia et al. [23]	Deep learning using convolutional autoencoders	ADNI dataset	84
Prajapati et al. [24]	Deep neural network binary classifier	ADNI dataset	85

and support vector machine. They achieve 75% of accuracy in validation. The experimented models are novel in this dataset but there is a scope for improvement because of poor accuracy.

Helaly et al. [21] discussed the importance of early detection of Alzheimer's disease. The authors used a convolutional neural network (CNN) for their work. They classified Alzheimer's disease into four broad categories. The authors used 2D and 3D brain image data for classification from the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset. Once the classification is done, the authors applied transfer learning models, for example, VGG19. The authors achieved 93.61 and 95.17 for 2D and 3D multi-class classifications and 97% of accuracy in the VGG19 pre-trained model.

Shakila Basheer et al. [22] investigated the MRI images dataset and used a large set of features for processing. As per the research, they conclude that age is the most important feature of Alzheimer's disease. They achieved 92.39% accuracy with the CNN model. There is a scope to improve the result by applying more validation to external data, and researchers also apply the same model to different datasets.

Martinez-Murcia et al. [23] experimented with a deep convolutional autoencoder that helps decompose a large dataset. The author uses conventional features and neuropsychologic variables to get the result. They achieve 84% accuracy, which is acceptable with the image data. But there is still scope for improvement, and researchers can improve accuracy.

Prajapati et al. [24] perform binary classification using a deep neural network. The author used an ADNI image dataset with three hidden layers and a k-fold combination with the fully connected network. They achieved 85% of accuracy with the scope for improvement.

3 Understanding of Data

3.1 Data

In this chapter, we used the Open Access Series of Imaging Studies (OASIS) data [25–27] from Kaggle. With the help of this MRI data, we will find the various categories of dementia. This dataset contains 60–96 years of patients with longitude values. This data is also balanced in nature. It has 72 nondemented and 64 demented patients' information. Each patient was scanned exactly once. Fourteen patients were found nondemented earlier but later were found demented.

3.2 Initial Data Analysis (IDA)

Before starting our analysis, we try to understand the nature of the data with a few statistical analyses. We try to establish correlations between features that help us better understand MRI data and dementia [28].

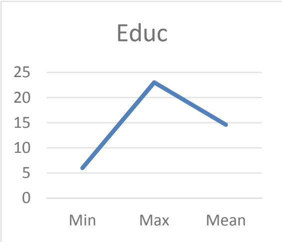
Figure 1 uses the numeric features to represent their minimum, maximum, and average values.

While observing the value of the dataset, we can infer the following.

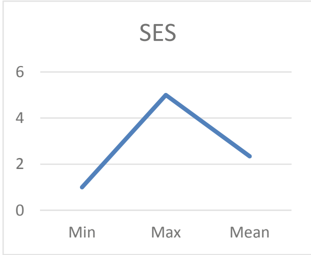
1. Men are more prone than women to be demented, as in Alzheimer's disease.
2. Demented patients had fewer years of education than healthy people.
3. Compared to the Demented group, the Nondemented group had a larger brain volume.
4. Compared to Demented and non-demented patients, demented have a higher proportion of people in their 70 and 80 s.

3.3 Data Pre-Processing

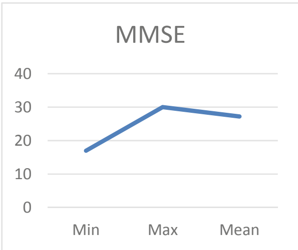
Table 2 represents the sample dataset used in this chapter. This dataset had some values that needed to be added. We have two choices: leave the entire record of the missing values or fill in a value in the missing space. We used Imputation as a method to fill in missing values. After that, we compare our model with our benchmark parameters, such as accuracy, AUC, and recall. We divided the data set into two parts one for training purposes and the second for testing purposes. We divide the dataset into a 7.5:2.5 ratio, i.e., 75% of the data is for training, and 25% is used for testing.



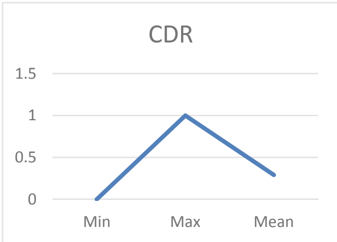
Education Years (EDUC)



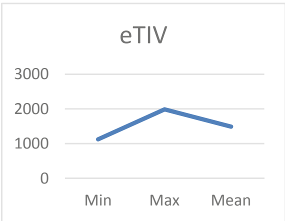
Social and Economic Status (SES)



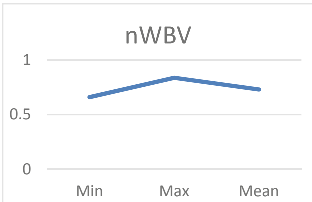
Status of Mini-Mental Examination



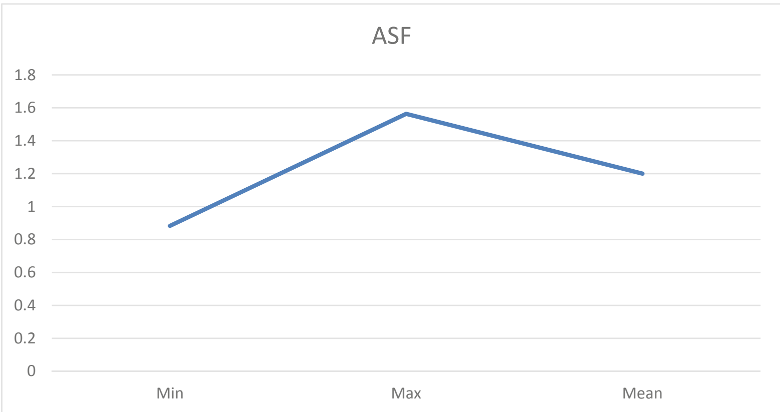
Ratings of Clinical Dementia



Estimated Total Intracranial Volume



Normalize Whole Brain Volume



Atlas Scaling Factor

Fig. 1 Representation of minimum, average, and maximum values of every feature for graph implementation

Table 2 Sample dataset

MRI ID	Group	Visit	MRDelay	M/F	Hand	Age	EDUC	SES	MMSE	CDR	eTIV	nWBV	ASF
OAS2_0001_MR1	Nondemented	1	0	M	R	87	14	2	27	0	1987	0.696	0.883
OAS2_0001_MR2	Nondemented	2	457	M	R	88	14	2	30	0	2004	0.681	0.876
OAS2_0002_MR1	Demented	1	0	M	R	75	12	NaN	23	0.5	1678	0.736	1.046
OAS2_0002_MR2	Demented	2	560	M M	R	76	12	NaN	28	0.5	1738	0.713	1.01

4 Performance Evaluation

4.1 Evaluation Metric

Our focus is on the early detection of Alzheimer's patients. We mainly focused on true positives and false positives. We wish to increase the true positive rate and decrease the false positive rate. So, we effectively detect Alzheimer's patients. For this purpose, we used AUC, Precision, and recall. In the Machine Learning (ML) method, if we have a labeled dataset, then two possibilities occur [29]. The either experimented result matches with the original result, i.e., True event, or not matched i.e., false event. On this basis, we can divide these true or false positive events into the following categories.

Table 3 represents the confusion matrix. Our statistics are based on TP and FP mostly. We try to increase TP and reduce the FP. This confusion matrix helps us to determine precision, recall, and accuracy as follows:

$$Precision = \left(\sum TP\right)/\left(\sum (TP + FP)\right) \tag{1}$$

$$Recall = \left(\sum TP\right)/\left(\sum (TP + FN)\right) \tag{2}$$

$$Accuracy = \sum (TP + TN)/\left(\sum (TP + FP + FN + TN)\right) \tag{3}$$

Equation 1 talks about a condition where we have a ratio of actual Alzheimer's patients and positive patients found by our algorithm. This situation is known as precision. Equation 2 represents a ratio of total positive (patients) out of the total number of people tested correctly. Equation 3 informed us how many correct results we achieved from the total experiments.

Table 3 Confusion matrix (Ma et al. [30])

	Experimental result		
Label in dataset		Positive	Negative
	Positive	TP	FN
	Negative	FP	TN

where *TP* True Positive, i.e., we got the same result as stored in the dataset, correctly identified Alzheimer's patients
FN False Negative, Result is opposite to the dataset result. i.e., we missed an Alzheimer's patient
FP False Positive; we wrongly mention a healthy person as an Alzheimer's patient
TN True Negative; our analysis finds a healthy person as a healthy person. i.e., we worked correctly

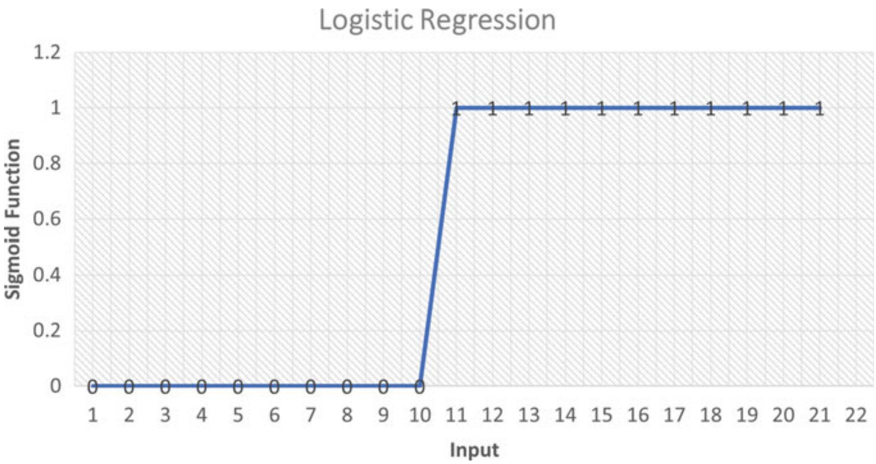


Fig. 2 Logistic regression

4.2 Algorithms

We Applied the following machine learning algorithms.

- Logistic Regression (LR): LR is a binary classification method [31]. It is used with supervised ML models. This model is best suited if we classify data in a yes/no format, such as whether the patient is suffering from Alzheimer's or not. The LR algorithm uses the following mathematical function.

$$Logistic\ function = 1/(1 + e^\wedge(-x)) \tag{4}$$

- This function is also known as the sigmoid function. This LR function uses the likelihood function (conditional probability) to calculate the loss. Figure 2 explains the behavior of the logistic regression function. The values are categorized into two segments. Values move toward positive infinity, represented as 1, and values move toward negative infinity, represents by 0. This model gave 75% of accuracy in a testing environment. We received the following results with LR algorithms.
- The max accuracy at the time of validation is 75%
- The value of the regularization parameter is 10
- Test accuracy is 78%
- Value of recall at the time of test is 75%
- AUC shows 79% score

4.2.1 Support Vector Machine (SVM)

SVM is a supervised classification method [32]. This method uses the kernel function for calculating the distance. This method is best suited for binary classification and

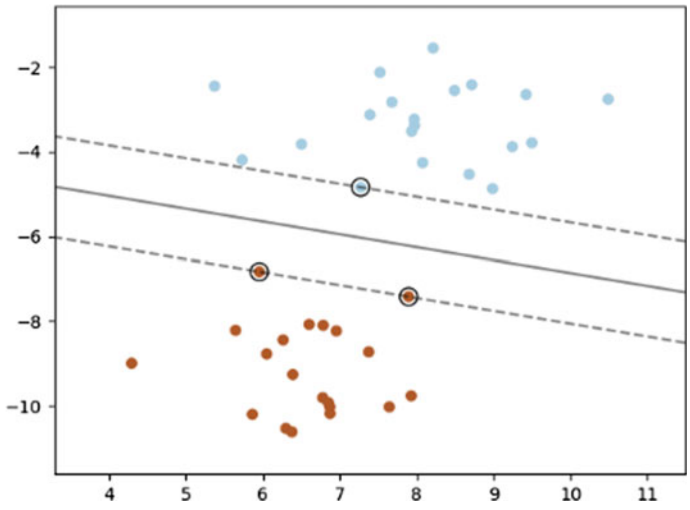


Fig. 3 SVM Model

outlier detection. It uses subset distance (support vector) for classification, making it memory efficient. But, if the number of features exceeds the number of samples, this method must be more balanced. SVM works effectively due to its core mathematical function, i.e., quadratic programming problem solver, but its space and time complexity increases rapidly.

In our experiments, we found the following result with the SVM model.

- We achieve the best accuracy with 77%
- The value for parameter 'C' is 100%
- The value for gamma we achieved 0.1
- Correct prediction with AUC is 81%
- Highest recall value is 70%
- Test recall (with the best parameter) is 82%.

Figure 3 explains the working of the SVM model. The SVM uses a hyper-plane model for classification or regression purposes. The larger separation of two planes is possible due to this hyper-plan function.

4.2.2 Decision Tree

A decision tree is another powerful algorithm in ML, as shown in Fig. 4. It is used for classification and regression. This method returns a true or false decision on an input. As the name suggests, it is a binary tree structure. It may grow internally in a binary tree manner. This method helps us in the question-answering pattern. For example, it takes features like patient age. If the age is greater than 50, then this

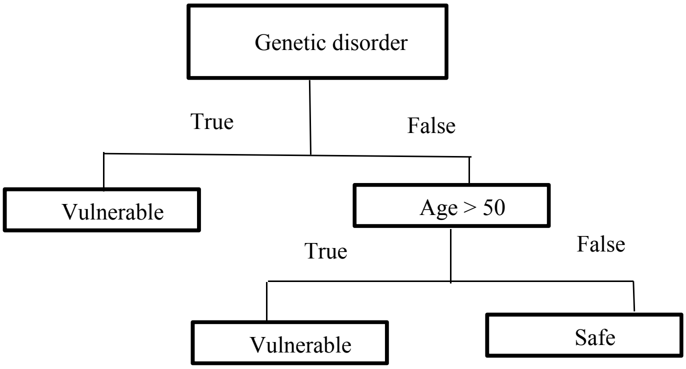


Fig. 4 Decision tree

method gives a positive answer, and the patient may be vulnerable to Alzheimer's disease.

We receive the following result with decision tree algorithms.

- The best accuracy (on validation set) computes as 77%
- Best parameter (for the maximum depth) comes out as 2
- We achieve the best accuracy with 81%
- Highest recall value is 65%
- Correct prediction with AUC is 82%.

4.2.3 Random Forest Classifier

Figure 5 shows the working of the random forest random. A random forest classifier is a collection of various independent trees. It is an ensemble machine-learning technique. As the name suggests, it is a forest of several decision trees.

We received the following result with a random forest classifier.

- Best accuracy (on validation set) computes as 80%
- Best parameters (of M, d, m) comes out as 14 5 7
- We achieve the best accuracy with 84%
- Highest recall value is 80%
- Correct prediction with AUC is 84%.

ML algorithms help us to analyze data and classification. We achieve good accuracy and result with SVM. If data is labeled, supervised learning is the best way to classify the data.

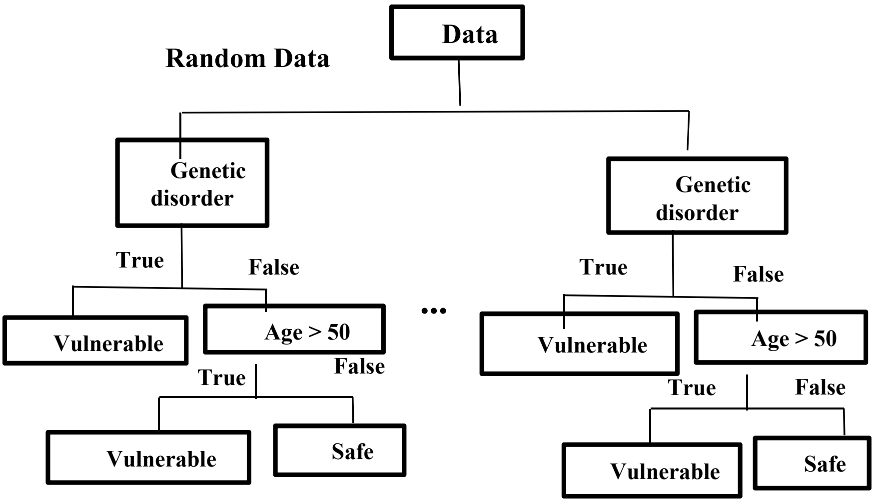


Fig. 5 Random forest

5 Result Analysis

To verify our results, we experimented with different ML methods like SVM, Decision tree, LR, Random forest, etc. The result we get is verified with a confusion matrix. The confusion matrix is the standard method to verify the experimented result. Our primary focus in this research is to avoid a false negative. Because if we start drug with a patient who has not suffered from this disease and we start medicins, it is too dangerous for the patient's health. To avoid this scenario, the AUC is the best option for a performance metric. We also experimented with accuracy and recall to verify our results. We compared the results of all ML algorithms. Figure 6 shows all results. That Random forest is best for AUC. Best accuracy we achieved through SVM and Random forest. But recall is higher in random forest and Logical Regression methods. Since data is labeled and the random forest is an ensemble learning algorithm. So, it gives better results in comparison to all algorithms.

6 Conclusion and Future Directions

In work mentioned, we applied the most common ML techniques and got better results than early researchers. Due to the data size, we are restricting ourselves to employing a more complex model. Since data is labeled then, supervised ML methods are best suitable for this dataset.

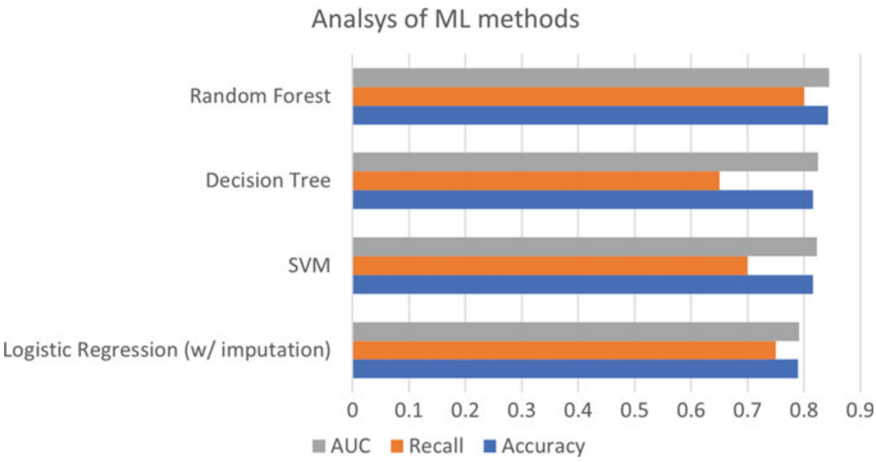


Fig. 6 Result generation and comparison of AI methods

The future direction of this research is to collect more data with a variety such as images and other patterns. With the voluminous data, we can go for deep learning approaches.

References

1. Zhang, D., Shen, D.: Predicting future clinical changes of MCI patients using longitudinal and multimodal biomarkers. PLoS ONE **7**(3), e33182 (2012). https://doi.org/10.1371/journal.pone.0033182
2. Batmanghelich, K.N., Ye, D.H., Pohl, K.M., Taskar, B., Davatzikos, C.: Disease classification and prediction via semi-supervised dimensionality reduction. Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on IEEE, pp. 1086–1090 (2011). https://doi.org/10.1109/ISBI.2011.5872590
3. Ardekani, B.A., Bachman, A.H., Figarsky, K., Sidtis, J.J.: Corpus callosum shape changes in early Alzheimer's disease: an MRI study using the OASISbraindatabase. Brain Struct. Funct. **219**, 343–352 (2014). https://doi.org/10.1007/s00429-013-0503-0
4. https://en.wikipedia.org/wiki/Precision_and_recall
5. Marcus, D.S., Wang, T.H., Parker, J., Csernansky, J.G., Morris, J.C., Buckner, R.L.: Open Access Series of Imaging Studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. J. Cogn. Neurosci. **19**(9), 1498–1507 (2007)
6. Marcus, D.S., Fotenos, A.F., Csernansky, J.G., Morris, J.C., Buckner, R.L.: Open access series of imaging studies: longitudinal MRI data in nondemented and demented older adults. J. Cogn. Neurosci. **22**(12), 2677–2684 (2010)
7. Islam, J., Zhang, Y.: Early diagnosis of Alzheimer's disease: a neuroimaging study with deep learning architectures. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 1881–1883 (2018)
8. Tayal, A., Solanki, A., Singh, S.P.: Integrated frame work for identifying sustainable manufacturing layouts based on big data, machine learning, meta-heuristic and data envelopment analysis. Sustainable Cities and Society (2020)

9. Singh, S.P., Solanki, A., Singh, T., Tayal, A.: Internet of intelligent things: injection of intelligence into IoT devices. Artificial intelligence to solve pervasive internet of …, (2021)
10. Kaur, H., Singh, S.P., Bhatnagar, S., Solanki, A.: Intelligent smart home energy efficiency model using artificial intelligence and internet of things. Artificial Intelligence to Solve Pervasive Internet of …, (2021)
11. Singh, S.P., Sharma, A., Kumar, R.: Designing of fog based FBCMI2E Model using machine learning approaches for intelligent communication systems. Computer Communications (2020)
12. Solanki, A., Kumar, S., Rohan, C., Singh, S.P., Tayal, A.: Prediction of breast and lung cancer, comparative review and analysis using machine learning techniques. Smart Computing and Self-Adaptive Systems (2021)
13. Tiwari, S., Kane, L., Koundal, D., Jain, A., Alhudhaif, A., Polat, K., Zaguia, A., Alenezi, F., Althubiti, S.A.: SPOSDS: A smart Polycystic Ovary Syndrome diagnostic system using machine learning. Expert Syst. Appl. **203**, 117592 (2022)
14. Bari Antor, M., Jamil, A. H. M., Mamtaz, M., Monirujjaman Khan, M., Aljahdali, S., Kaur, M., ... & Masud, M. (2021). A comparative analysis of machine learning algorithms to predict alzheimer's disease. *Journal of Healthcare Engineering*, *2021*.
15. Moradi, E., Pepe, A., Gaser, C., Huttunen, H., Tohka, J., Initiative, A.D.N.: Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects. Neuroimage **104**, 398–412 (2015)
16. Zhang, Y., Dong, Z., Phillips, P., Wang, S., Ji, G., Yang, J., Yuan, T.F.: Detection of subjects and brain regions related to Alzheimer's disease using 3D MRI scans based on eigenbrain and machine learning. Front. Comput. Neurosci. **9**, 66 (2015)
17. Magnin, B., Mesrob, L., Kinkingnéhun, S., Pélégrini-Issac, M., Colliot, O., Sarazin, M., Benali, H., et al.: Support vector machine-based classification of Alzheimer's disease from whole-brain anatomical MRI. Neuroradiology **51**(2), 73–83 (2009)
18. Khan, P., Kader, M.F., Islam, S.R., Rahman, A.B., Kamal, M.S., Toha, M.U., et al.: Machine learning and deep learning approaches for brain disease diagnosis: principles and recent advances. IEEE Access **9**, 37622–37655 (2021)
19. Saratxaga, C.L., Moya, I., Picón, A., Acosta, M., Moreno-Fernandez-de-Leceta, A., Garrote, E., et al.: MRIDeep learning-based solution forAlzheimer's Disease Prediction. J. Pers. Med (2021)
20. Sudharsan, M., Thailambal, G.: Alzheimer's disease prediction using machine learning techniques and principal component analysis (PCA). Materials Today: Proceedings (2021)
21. Helaly, H.A., Badawy, M., Haikal, A.Y.: Deep learning approach for early detection of Alzheimer's disease. Cogn Comput. (2021)
22. Basheer, S., Bhatia, S., Sakri, S.B.: Computational modeling of dementia prediction using deep neural network: analysis on OASIS dataset. IEEE Access **9**, 42449–42462 (2021)
23. Martinez-Murcia, F.J., Ortiz, A., Gorriz, J.M., Ramirez, J., Castillo-Barnes, D.: Studying the manifold structure of Alzheimer's disease: a deep learning approach using convolutional autoencoders. IEEE J Biomed. Health Inform. **24**, 17–26 (2020)
24. Prajapati, R., Khatri, U., Kwon, G.R.: An efficient deep neural network binary classifier for Alzheimer's disease classification. In: International Conference on Artificial Intelligence in Information and Communication (ICAIIC), pp. 231–234, (2021)
25. Marcus, D.S., Fotenos, A.F., Csernansky, J.G., Morris, J.C., Buckner, R.L.: Open Access Series of Imaging Studies (OASIS): longitudinal MRI data in nondemented and demented older adults. J. Cogn. Neurosci. **22**(12), 2677–2684 (2010). https://doi.org/10.1162/jocn.2009.21407
26. Marcus, D.S., Wang, T.H., Parker, J., Csernansky, J.G., Morris, J.C., Buckner, R.L., Open Access Series of Imaging Studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. J. Cogn. Neurosci., **19**, 1498–1507. https://doi.org/10.1162/jocn.2007.19.9.1498
27. Moradi, E., Pepe, A., Gaser, C., Huttunen, H., Tohka, J.: Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects. NeuroImage **104**, 398–412 (2015), ISSN 1053–8119. https://doi.org/10.1016/j.neuroimage.2014.10.002

28. Zhang, Y., Dong, Z., Phillips, P., et al.: Detection of subjects and brain regions related to Alzheimer's disease using 3D MRI scans based on eigenbrain and machine learning. Front. Comput. Neurosci. **9**, 66 (2015). https://doi.org/10.3389/fncom.2015.00066
29. Magnin, B., Mesrob, L., Kinkingnéhun, S., et al.: Support vector machine-based classification of Alzheimer's disease from whole-brain anatomical MRI. Neuroradiology **51**, 73 (2009). https://doi.org/10.1007/s00234-008-0463-x
30. Ma, J., Ding, Y., Cheng, J.C., Tan, Y., Gan, V.J., Zhang, J.: Analyzing the leading causes of traffic fatalities using XGBoost and grid-based analysis: a city management perspective. IEEE Access **7**, 148059–148072 (2019)
31. Filipovych, R., Davatzikos, C.: Semi-supervised pattern classification of medical images: application to mild cognitive impairment (MCI). Neuroimage **55**(3), 1109–1119 (2011). https://doi.org/10.1016/j.neuroimage.2010.12.066
32. http://scikit-learn.org/stable/modules/preprocessing.html#imputation