International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 02 | Feb 2020 www.irjet.net p-ISSN: 2395-0072

Alzheimer's Detection Model Using Machine Learning

Shashi Rekha[1], Gangula Rekha[2], Chinnaka Sudha[3]

[1]Asst.Professor, Dept. of CSE, SR International Institute of Technology, Telangana, INDIA.
[2]Asst.Professor, Dept. of CSE, Kakatiya Institute of Technology and Science,Telangana, INDIA.
[3]Asst.Professor, Dept. of IT, Mahatma Gandhi Institute of Technology, Telangana, INDIA.

Abstract - *Alzheimer's disease (AD), an irreparable brain disease, impairs thinking and memory while the aggregate mind size shrinks which at last prompts demise. Alzheimer's is a neurodegenerative disease and leads to severe memory loss and inability to cope with daily life tasks. Early diagnosis of AD is essential for the progress of more prevailing treatments. Detecting Alzheimer's is a difficult and time consuming task, but requires brain imaging report and human expertise. Needless to say, this conventional approach to detect Alzheimer's is costly and often error prone. In this project an alternative approach has been discussed, that is fast, costs less and more reliable. Artificial intelligence systems can help in providing better health care and medical solutions. The performance of human diagnosis degrades due to fatigue, cognitive biases, systems faults, and distractions. However, artificial intelligence based diagnosis systems are less error prone and give safe support to clinicians in detection and decision making. This work presents a smart and reliable way of diagnosing Alzheimer's disease (AD) and its possible early stage i.e., mild cognitive impairment. The presented framework is based on deep learning and detects Alzheimer's and its initial stages accurately from structural MRI scans. Identifying mild cognitive impairment (MCI) subjects who will progress to Alzheimer's disease is not only crucial in clinical practice, but also has a significant potential to enrich clinical trials. This project proposes to combine MRI data with a neuropsychological test, Mini-Mental State Examination (MMSE), as input to a multi-dimensional space for the classification of Alzheimer's Disease (AD) and it's prodromal stages.*

***Key Words*: Artificial Intelligence, *Alzheimer's disease (AD),* Machine learning, Accuracy, Classifier models.**

1. INTRODUCTION

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

1.1 Motivation

In current circumstances, the standard measurements and human instinct don't frequently work. That is the reason we must depend on intensively computational and non-traditional approaches such as machine learning. The custom of using machine learning as a part of disease prediction and visualization is a fragment of an expanding shift towards prescient and customized prescription. This drift is important, not only for the patients in increasing their quality of life and life style, but for physicians in making treatment decisions and also for health economists.

It's not the disease of age, it's the disease of the brain and patients may show Symptoms like loss of memory, difficulty in finding the right words or understanding what people are saying, difficulty in performing previously routine tasks and personality and mood changes[2].

When a radiologist views a medical report e.g., a magnetic resonance imaging (MRI) scan of a patient, a biased thinking for a disease would result in missing the chance of detecting other disease conditions. Thus, it leads to considering only a subset of causes and conditions. C S Lee, [3] claimed that approximately 75% of all the medical errors occurred due to diagnostic errors by the radiologists. An increased workload, stress, fatigue, cognitive bias, and poor system are some of the factors behind it. In this situation, smart diagnostic systems provide a safe clinical support to the clinicians.

The Alzheimer's association claims that AD is the sixth leading cause of death in United States [4].

1.2 Problem Statement

In the existing system it is difficult to identify if a person is suffering from Alzheimer's disease. It can be only done with the help of clinical history and by knowing if the person has some genetic disorder. Sometimes it is also possible that the doctor may not be able to detect the disease.

Alzheimer's disease (AD), a type of dementia, is characterized by progressive problems with thinking and behavior that starts in the middle or old age. The symptoms usually develop slowly and get serious enough to interfere in daily life. Although the paramount risk factor is oldness but

AD is not just an old age disease. In its early stages, the memory loss is mild while in the later stages, the patient's conversation and their ability to respond degrades dramatically. The current treatments cannot stop Alzheimer's disease (AD) from developing but early diagnosis can aid in precluding the severity of the disease and help the patients to improve the quality life. It has been reported that the number of individuals effected with AD will double in next 20 years, while in 2050, 1 out of 85 individuals will be effected. Thus the accurate diagnosis especially for the early stages of AD is very important.

1.3 Objective

The Main objective of the project is:

Is to make the diagnosis of the disease easier, to detect the disease in its early stages and use the machine algorithm efficiently.

2. Detection Model Design

The class diagram is a static diagram. It represents the static view of an application. Class diagram is not only used for visualizing, describing and documenting different aspects of a system but also for constructing executable code of the software application and is shown in Fig-1 for the Alzheimer's disease detection.

Use Case diagrams identify the functionality provided by the system (use cases), the users who interact with the system (actors), and the association between the users and the functionality. Use Cases are used in the Analysis phase of software development to articulate the high-level requirements of the system and is shown in Fig-2 for the Alzheimer's disease detection.

The Activity Diagram forms effective while modeling the functionality of the system. Hence this diagram reflects the activities, the types of flows between these activities and finally the response of objects to these activities and is shown in Fig-3 for the Alzheimer's disease detection.

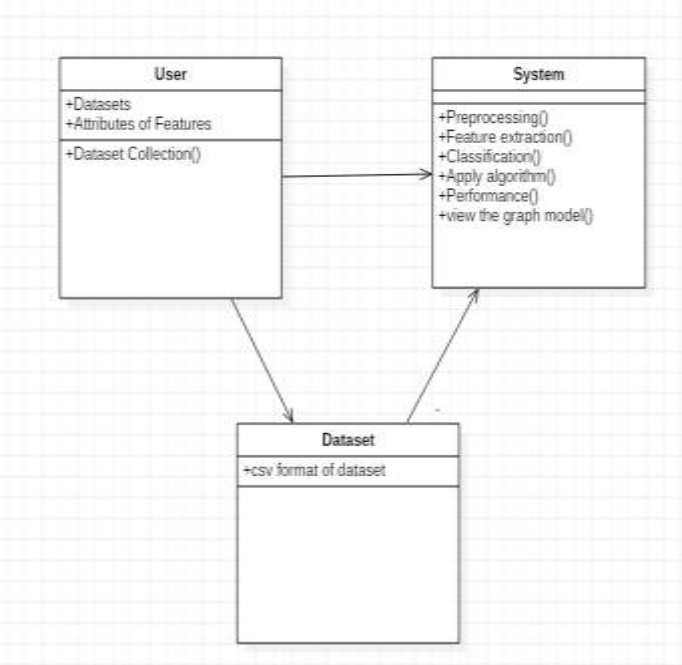

Fig -1: Class Diagram of the model

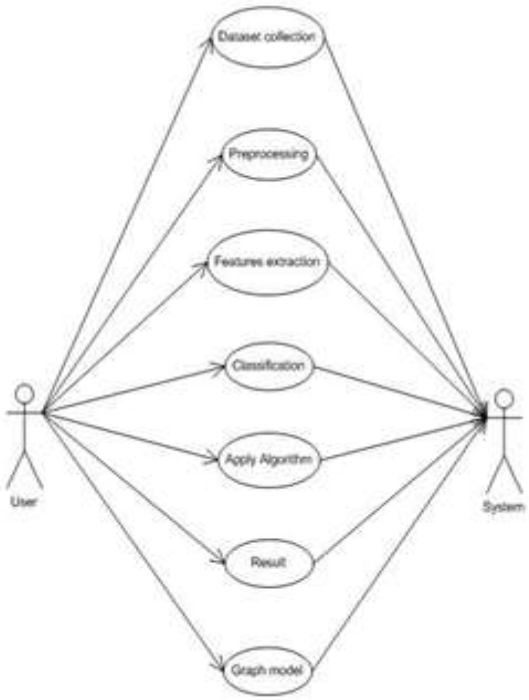

Fig -2: Use Case Diagram of the model

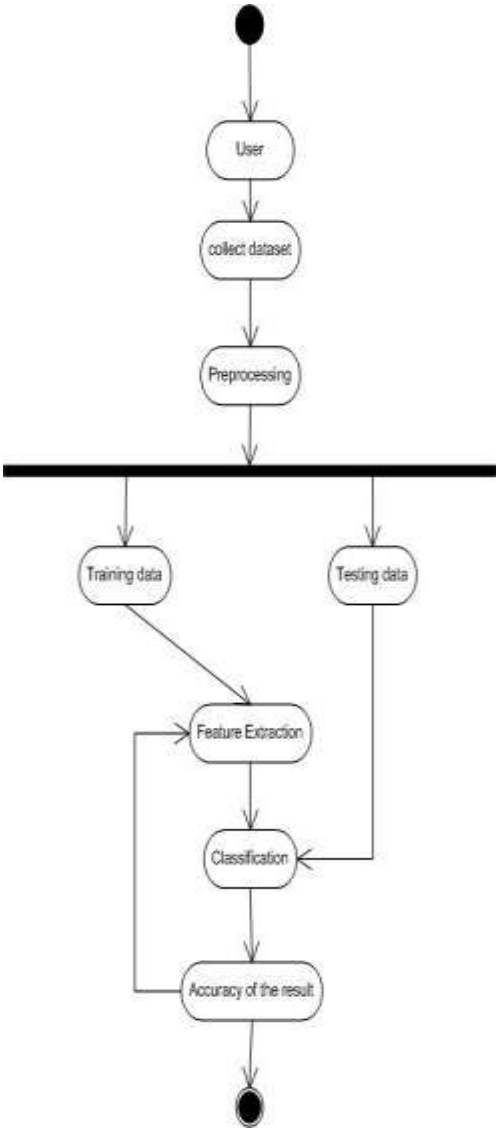

Fig -3: Activity Diagram of the model

3. Implementation

There are four main modules, they are:

3.1 Data preprocessing:

This step contains all the pre-processing functions needed to process the input dataset. First we split data into train, test data files then perform some pre-processing like normalization to avoid the curse of dimensionality. Some exploratory data analysis is performed like response variable distribution and data quality checks like null or missing values etc.

3.2 Feature extraction:

In this step we perform feature extraction and selection methods from sci-kit learn python libraries. For feature selection, we have used methods like simple bag-of-words and n-grams and then term frequency like TF-IDF(team frequency-inverse document frequency) weighting. Feature extraction can be done by finding the correlation among the dataset using heat map.

3.3 Classification:

The extracted features are fed into different classifiers.

We have used Naive-Bayes, KNN (K-Nearest Neighbor), SVM(Support Vector Machine)and Random forest classifiers from sklearn. Each of the extracted features were used in all of the classifiers. Once fitting the model, we compared the f1 score and check the confusion matrix. With the help of the features extracted we would find the best classifier model for giving the result.

3.4 Prediction:

Our finally selected and best performing classifier algorithm is used. It takes a patient health record as input from user then model is used for final classification output that is shown to user along with probability of truth.

The sample code and output from Jupyter notebook was shown below:

Fig -4: Importing packages and functions

```
lon = pd.read_csv('oasis_longitudinal.csv')
lon = lon.fillna(method = 'ffill')
```

```
lon.head()
```

	Subject ID	MRI ID	Group	Visit	MR Delay	M/F	Hand	Age	EDUC	SES	MMSE	CDR	eTIV	nWBV	ASF
0	OAS2_0001	OAS2_0001_MR1	Nondemented	1	0	M	R	87	14	2.0	27.0	0.0	1987	0.696	0.883
1	OAS2_0001	OAS2_0001_MR2	Nondemented	2	457	M	R	88	14	2.0	30.0	0.0	2004	0.681	0.876
2	OAS2_0002	OAS2_0002_MR1	Demented	1	0	M	R	75	12	2.0	23.0	0.5	1678	0.736	1.046
3	OAS2_0002	OAS2_0002_MR2	Demented	2	560	M	R	76	12	2.0	28.0	0.5	1738	0.713	1.010
4	OAS2_0002	OAS2_0002_MR3	Demented	3	1895	M	R	80	12	2.0	22.0	0.5	1698	0.701	1.034

```
lon.tail()
```

	Subject ID	MRI ID	Group	Visit	MR Delay	M/F	Hand	Age	EDUC	SES	MMSE	CDR	eTIV	nWBV	ASF
368	OAS2_0185	OAS2_0185_MR2	Demented	2	842	M	R	82	16	1.0	28.0	0.5	1693	0.694	1.037
369	OAS2_0185	OAS2_0185_MR3	Demented	3	2297	M	R	86	16	1.0	26.0	0.5	1688	0.675	1.040
370	OAS2_0186	OAS2_0186_MR1	Nondemented	1	0	F	R	61	13	2.0	30.0	0.0	1319	0.801	1.331
371	OAS2_0186	OAS2_0186_MR2	Nondemented	2	763	F	R	63	13	2.0	30.0	0.0	1327	0.796	1.323
372	OAS2_0186	OAS2_0186_MR3	Nondemented	3	1608	F	R	65	13	2.0	30.0	0.0	1333	0.801	1.317

Fig -5: Dataset display

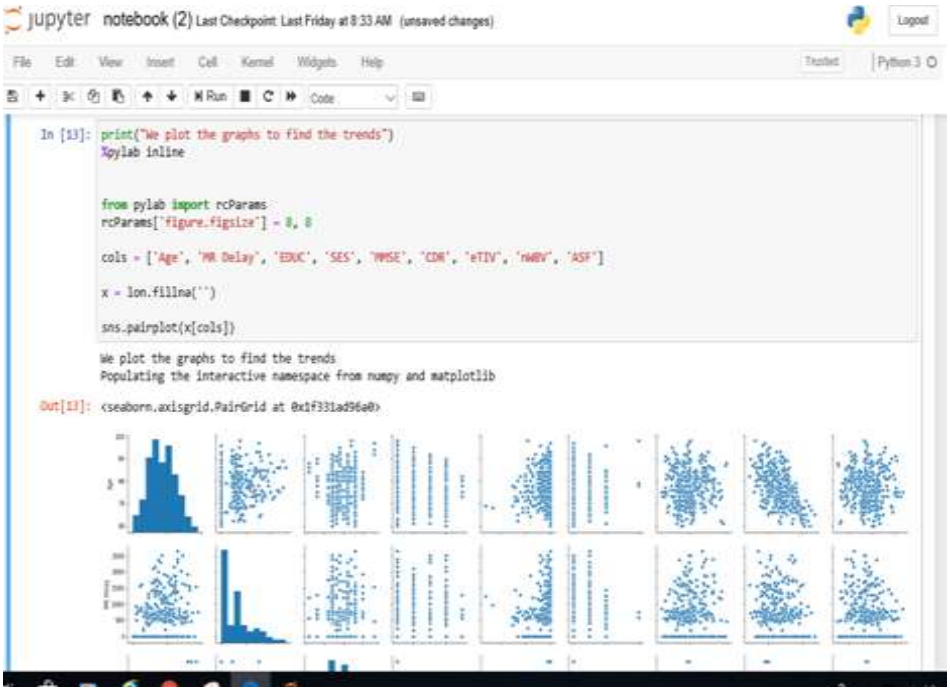

Fig -6: Plotting trends

We locate the correlation heat map for feature extraction

<matplotlib.axes._subplots.AxesSubplot at 0x1f334edd470>

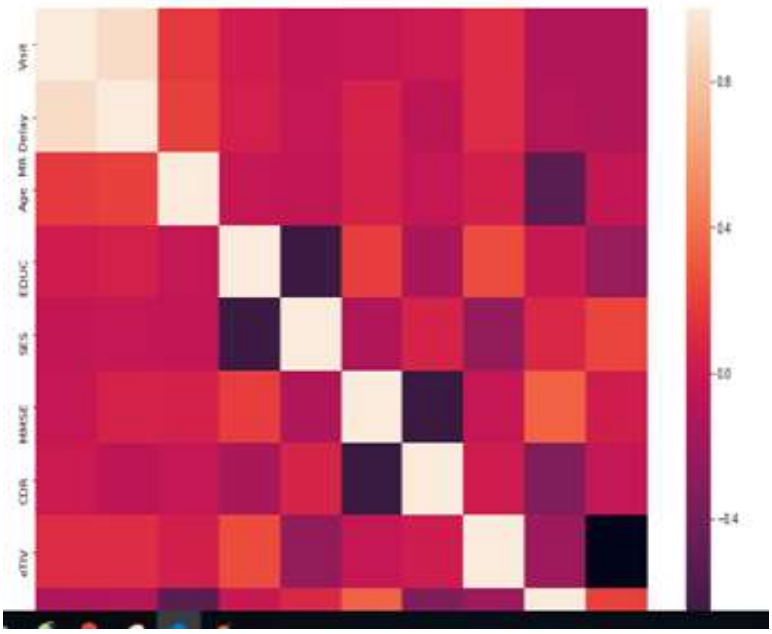

Fig -7: Plotting correlation matrix (heat map)

```
In [18]: from sklearn.preprocessing import LabelEncoder
         from sklearn.model_selection import train_test_split
         col=['Subject ID','Group','M/F','Hand']
         for x in col:
             f = LabelEncoder()
             lon[x] = f.fit_transform(lon[x])

In [19]: lon.head()
```

Out[19]:

	Subject ID	Group	MR Delay	M/F	Hand	Age	EDUC	SES	MMSE	CDR	eTIV	nWBV	ASF
0	0	2	0	1	0	87	14	2.0	27.0	A	1987	0.696	0.883
1	0	2	457	1	0	88	14	2.0	30.0	A	2004	0.681	0.876
2	1	1	0	1	0	75	12	2.0	23.0	B	1678	0.736	1.046
3	1	1	560	1	0	76	12	2.0	28.0	B	1738	0.713	1.010
4	1	1	1895	1	0	80	12	2.0	22.0	B	1698	0.701	1.034

Fig -7: Categorizing values

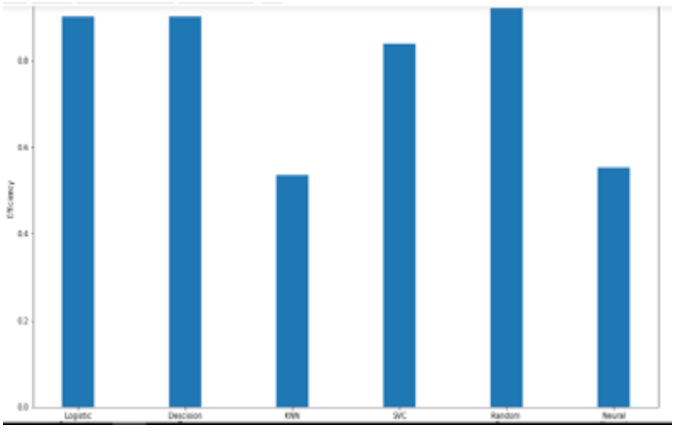

Fig -8: Plotting graphs

```
In [61]: ar1=[[1.461,69,18,2,0,0,29.0,1329,2.0,1202,0.770]]
         ar2=[[1.934,80,12,1,0,1,22.0,1895,2.0,1698,0.701]]
         pred=forest.predict(ar1)
         pred=list(pred)
         print(pred)
         pred1=forest.predict(ar2)
         pred=list(pred1)
         print(pred1)
         ar3=[[0.876,88,14,2,0,1,30.0,457,2.0,2004,0.681]]
         pred2=forest.predict(ar3)
         pred=list(pred)
         print(pred2)

         ['A']
         ['B']
         ['A']

In [62]: if pred1[0] =='A':
          print("Subject is Normal")
         elif pred1[0] =='B':
          print ("Subject has Very Mild Dementia")
         elif pred1[0] =='C':
          print("Subject has Moderate Dementia")
         else :
          print("Subject has Severe Dementia")

         Subject has Very Mild Dementia
```

Fig -9: Giving sample data

4. CONCLUSIONS

This study is based on the comparison and evaluation of recent work done in the prognosis and prediction of Alzheimer's disease using machine learning methods. Explicitly, the recent trends with respect to machine learning has been revealed including the types of data being used and the performance of machine learning methods in predicting early stages of Alzheimer's. It is obvious that machine learning tends to improve the prediction accuracy especially when compared to standard statistical tools. However, based on the review, the clinical diagnosis were not 100% accurate, as pathological verification was not provided which consequently introduce uncertainty in the predicted results

Scope for Future Work

It can be built into a proper user interface application so that it is easily accessible to everyone.

It can be transferred as a tool in hospitals so that it helps the patients with Alzheimer's disease.

REFERENCES

[1] Alzheimer's Disease International(ADI), https://www.alz.co.uk/aboutdementia.

[2] C. S. Lee, G. N. Paul, J. W. Sallie, E. N. David, "Cognitive and system factors contributing to diagnostic errors in radiology," American Journal of Roentgenology, vol. 201(3), pp. 611-617, 2013.

[3] Alzheimer's Association, "2016 Alzheimer's disease facts and figures," Alzheimer's & Dementia, vol. 12(4), pp. 459-509, 2016.