PianalytiX

Alzheimer's disease classification

Members

- **Suyash Ujawane**
- **Krishnendu Bhowmick**
- **Gaurav Insha**
- **Kalyani Pande**

ML Internship TEAM 7.6

Pianalytix

PianalytiX

Abstract

Alzheimer's disease, the most common form of dementia is a neurodegenerative disease. This disease occurs when the neurons in the brain starts degenerating and it generally affects the elderly people. It has no cure if the patient is in the last stage, so the classification of the disease at an early stage is very important for the treatment. Alzheimer's disease accounts for 60 to 80 percent of cases of dementia. In early stages of disease, also known as mild cognitive impairment (MCI), memory loss is mild, but with late-stage Alzheimer's, the patient loses the ability to even carry on a conversation. Nowadays doctors widely use magnetic resonance imaging (MRI) to diagnose various neurodegenerative diseases like Alzheimer's, but this method is prone to human error. Thus, Deep learning plays a vital role in the classification, detection and segmentation of data.

However, early diagnosis of AD can reduce the progression of the disease at the early stages. Deep convolutional Neural Networks (CNN), a class of Artificial Neural Network (ANN) are most popularly used by the researchers for the classification, visualization and segmentation of data. The diagnosis of the disease requires large number of medical tests which leads to multivariate heterogeneous data and it becomes very difficult to classify this huge amount of data and so the chances of error increases.

As the reason behind the success of deep neural networks owes to its three characteristics, i.e., layer-to-layer processing and in-model feature transformation for classification, and sufficient model complexity.

The model can be trained on volumetric and thickness data obtained after preprocessing MRI images using Free Surfer for various classifications tasks. The extracted features include the volume and thickness of different parts of the brain, including cortical and sub-cortical parts of the brain.

Table of contents

1. Purpose

Alzheimer's disease classification using ML and DL algorithms Alzheimer's disease is a progressive and irreversible disorder that slowly destroys memory, thinking skills, and other important mental functions. This disease has three stages early, middle, and late using different ML and DL algorithms we classify these three stages.

2. What is Alzheimer's Disease?

Alzheimer's disease is a progressive and irreversible disorder that slowly destroys memory, thinking skills, and other important mental functions that will eventually remove a person's ability to carry out simple tasks.

Alzheimer's disease is the most common cause of dementia, which is a group of brain disorders that cause the loss of intellectual and social skills. There is no known cure for Alzheimer's disease, but there are medications and management strategies that may temporarily improve symptoms

3. Stages in Alzheimer's Disease:

The symptoms of this disease worsen over time, the rate of progression for each person varies and each person experiences the disease differently. Below are the three stages of Alzheimer's.

A. Mild Alzheimer's Disease (early stage)

People suffering from Alzheimer's disease can appear to be healthy and functioning independently - participating in social activities, driving themselves, or working. Despite looking healthy, they may actually be having trouble making sense of their surroundings and experience memory lapses, like forgetting familiar words or locating everyday objects. This stage is often mistaken for normal aging.

Common difficulties in this stage include:

- Problems remembering the right word or name
- Trouble remembering names when introduced to new people
- Challenges performing tasks in social situations and work settings
- Forgetting material that one has just read
- Losing or misplacing objects
- Increased trouble with planning and organizing

B. Moderate Alzheimer's Disease (middle stage)

This stage is typically the longest stage and can last for several years. In this stage, the degeneration of the brain worsens and spreads to other areas that control language, reasoning, sensory processing and thought. The damage to nerve cells in the brain makes it difficult to express thoughts and perform routine tasks. You may notice in this stage, someone with Alzheimer's will confuse words, get frustrated or angry, or act out.

It is at this point that symptoms will be more noticeable to others and may include:

- Increased risk of wandering and getting lost
- Personality and behavioral changes, like suspiciousness, delusions or compulsive, repetitive behavior
- Changes in sleeping patterns, like sleeping during the day and restlessness at night
- Some individuals experience trouble controlling their bladder and bowels
- Confusion about where they are and what day it is
- Forgetfulness of events or about their own personal history

C. Severe Alzheimer's Disease (late stage)

This is the final stage of the disease where the damage to brain's nerve cells is widespread and people may lose their motor coordination and the ability to walk, speak, feed themselves, and recognize others. Full-time care is usually required during this stage.

In this stage, some may:

- Need around-the-clock assistance with daily activities and personal care
- Become vulnerable to infections, especially pneumonia
- Lose awareness of recent experiences as well as their surroundings

4. Procedure:

4.1. Data Collection and Preprocessing

- Data used in the preparation of this report was obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database **(http://adni.loni.usc.edu).**
- After downloading the images from ADNI database, we need to remove unnecessary details of brain MR images that might cause poor training of our classification model, cortical reconstruction and volumetric segmentation was performed with the Free-Surfer image analysis suite

4.2. Collecting Extracted Features

- After we have pre-processed images using free-surfer software, we need to extract the features from the pre-processed images using the frees-surfer commands.
- **Frees-surfer Command**
  - ***asegstats2table:*** **-** It produces the volumetric data from different cortical structures of the brain segmented.
  - ***aparcstats2table:*** **-** It produces the area or thickness data from different subcortical structures of left part of the brain segmented

4.3. Work On Classification

- After collecting extracted features from the pre-processed images, we prepare 4 csv files which contains subject features of CN , AD and MCI.

  File name as follows:

  I. CN_AD.csv : which contains subject features of CN and AD
  II. AD_MCI.csv : which contains subject features of AD and MCI
  III. MCI_CN.csv : which contains subject features of MCI and CN
  IV. CN_AD_MCI.csv : which contains subject features of CN, AD and MCI

- CN_AD.csv, AD_MCI.csv, and MCI_CN.csv These 3 files will be used for 2-way classification.
- CN_AD_MCI.csv file will be used for 3-way classification.

4.4. CN, AD and MCI

a) CN: (Cognitively Normal)

- ***(early stage)***
- People suffering from Alzheimer's disease can appear to be healthy and functioning independently - participating in social activities, driving themselves, or working. Despite looking healthy, they may actually be having trouble making sense of their surroundings and experience memory lapses, like forgetting familiar words or locating everyday objects. This stage is often mistaken for normal aging.

b) MCI: (Mild Cognitive Impairment)

- ***(middle stage)***
- This stage is typically the longest stage and can last for several years. In this stage, the degeneration of the brain worsens and spreads to other areas that control language, reasoning, sensory processing and thought. The damage to nerve cells in the brain makes it difficult to express thoughts and perform routine tasks. You may

notice in this stage, someone with Alzheimer's will confuse words, get frustrated or angry, or act out.

c) AD:

- ***(late stage)***
- This is the final stage of the disease where the damage to brain's nerve cells is widespread and people may lose their motor coordination and the ability to walk, speak, feed themselves, and recognize others. Full-time care is usually required during this stage.

5. Technology

Now various machine learning algorithms we used like Xgboost, Logistic Regression, SVM, Naïve Bayes, Random Forest, Gradient Boosting, Discriminant Analysis. to train and test our data set. We did the classification with and without PCA.

6. Observation

CN and AD Classification :

Note:- For this we use CN_AD.csv file

❖ Models Without PCA

Without PCA			
Sr. No.	**Algorithm**	**Mean Accuracy**	**Standard Deviation**
1	Random Forest	79.35%	0.041744
2	**Logistic Regression**	**84.39%**	**0.021982**
3	Gradient Boosting	74.94%	0.080145
4	Xgboost	78.72%	0.03351
5	Xgboost with parameter	51.89%	0.07862
6	Linear Discrimenant Analysis	43.92%	0.046358
7	SVM	83.12%	0.030846
8	Naïve Bayes	78.73%	0.094454

❖ Model With PCA

With PCA			
Sr. No.	**Algorithm**	**Mean Accuracy**	**Standard Deviation**
1	Random Forest	79.97%	0.033114
2	Logistic Regression	83.75%	0.023297
3	Gradient Boosting	77.50%	0.013593
4	Xgboost	75.59%	0.054295
5	Xgboost with parameter	75.58%	0.042377
6	**Linear Discrimenant Analysis**	**84.38%**	**0.007412**
7	SVM	79.97%	0.045231

CN and MCI Classification

Note:- For this we use CN_MCI.csv file

❖ Models without PCA

Without PCA			
Sr. No.	**Algorithm**	**Mean Accuracy**	**Standard Deviation**
1	Random Forest	64.15%	0.040759
2	Logistic Regression	62.89%	0.049522
3	Gradient Boosting	61.00%	0.054102
4	Xgboost	66.03%	0.046216
5	Xgboost with parameter	66.03%	0.040759
6	Linear Discrimenant Analysis	58.78%	0.134539
7	SVM	62.89%	0.049522
8	**Naïve Bayes**	**68.55%**	**0.017788**

❖ Model With PCA

With PCA			
Sr. No.	**Algorithm**	**Mean Accuracy**	**Standard Deviation**
1	Random Forest	61.00%	0.023532
2	Logistic Regression	66.03%	0.030811
3	Gradient Boosting	57.23%	0.008894
4	Xgboost	59.74%	0.023532
5	Xgboost with parameter	61.63%	0.044472
6	Linear Discrimenant Analysis	67.29%	0.044472
7	**SVM**	**67.92%**	**0.030811**

MCI and AD Classification

Note:- For this we use MCI_AD.csv file

❖ Model Without PCA

Without PCA			
Sr. No.	**Algorithm**	**Mean Accuracy**	**Standard Deviation**
1	Random Forest	54.71%	0.015405
2	Gradient Boosting	52.20%	0.06226
3	Xgboost	47.79%	0.049522
4	Xgboost with parameter	53.45%	0.06226
5	Linear Discrimenant Analysis	54.36%	0.15382
6	SVM	52.20%	0.023532
7	**Naïve Bayes**	**59.11%**	**0.06226**

❖ Model With PCA

With PCA			
Sr. No.	**Algorithm**	**Mean Accuracy**	**Standard Deviation**
1	Random Forest	53.45%	0.102575
2	Logistic Regression	59.11%	0.104863
3	Gradient Boosting	56.60%	0.096207
4	Xgboost	55.97%	0.06226
5	Xgboost with parameter	54.08%	0.071155
6	Linear Discrimenant Analysis	57.23%	0.124521
7	**SVM**	**59.74%**	**0.023532**

CN, AD and MCI Classification

Note:- For this we use CN_AD_MCI.csv file

❖ Model Without PCA

Without PCA			
Sr. No.	**Algorithm**	**Mean Accuracy**	**Standard Deviation**
1	Random Forest	46.88%	0.045826
3	Gradient Boosting	47.71%	0.046375
4	Xgboost	49.37%	0.061877
5	Xgboost with parameter	50.62%	0.076546
6	Linear Discrimenant Analysis	43.92%	0.046358
7	**SVM**	**52.30%**	**0.036235**
8	Naïve Bayes	49.79%	0.076775

❖ Model With PCA

With PCA			
Sr. No.	**Algorithm**	**Mean Accuracy**	**Standard Deviation**
1	Random Forest	51.87%	0.031858
2	Logistic Regression	54.83%	0.070036
3	Gradient Boosting	52.26%	0.067
4	Xgboost	53.13%	0.02551
5	Xgboost with parameter	46.85%	0.022348
6	Linear Discrimenant Analysis	53.57%	0.060713
7	**SVM**	**56.91%**	**0.049028**

7. Final Observation

- **Algorithm that gives MAX Accuracy Without PCA**

Sr. No	Classification	Algorithm	Mean Accuracy	Standard Deviation
1	CN and AD	Logistic Regression	84.39%	0.021982
2	CN and MCI	Naïve Bayes	68.55%	0.017788
3	MCI and AD	Naïve Bayes	59.11%	0.06226
4	CN, AD and MCI	SVM	52.30%	0.036235

- **Algorithm that gives MAX Accuracy With PCA**

Sr. No	Classification	Algorithm	Mean Accuracy	Standard Deviation
1	CN and AD	Linear Discrimenant Analysis	84.38%	0.007412
2	CN and MCI	SVM	67.92%	0.030811
3	MCI and AD	SVM	59.74%	0.023532
4	CN, AD and MCI	SVM	56.91%	0.049028

8. Reference

- https://www.alz.org/alzheimers-dementia/stages#:~:text=Alzheimer's%20disease%20typically%20progresses%20slowly,progress%20through%20the%20stages%20%E2%80%94%20differently.
- https://en.wikipedia.org/wiki/Machine_learning
- https://en.wikipedia.org/wiki/Deep_learning