Project Report

On

'Alzheimer's Disease prediction using Machine Learning'

Submitted for the requirement of

Project course

BACHELORS OF ENGINEERING

COMPUTER SCIENCE & ENGINEERING

SUBMITTED TO: ***Ms. Neeru Sharma Ma'am***

SUBMITTED BY:

Sr No	*NAME*	*UID*	*SECTION*
1.	*Vikrant Dalvi*	*20BCS1858*	*716/ 'B'*
2.	*Rahul Choudhary*	*20BCS1804*	*710/ 'B'*
3.	*Satyam Kumar*	*20BCS1857*	*719/ 'B'*
4.	*Alok Raj*	*20BCS1958*	*710/ 'B'*
5.	*Deepshikha Kumari*	*20BCS5297*	*710/ 'B'*

BONAFIDE CERTIFICATE

*Certified that this project report **"Alzheimer's Disease prediction using Machine Learning"** is the bonafide work by the aforementioned group members that was carried out under my supervision.*

Signature	***Signature***
(HoD, CSE)	*(Supervisor)*
	Ms. Neeru Sharma

Submitted for the Project Viva-voce examination to be held on May 17, 2023

Internal Examiner ***External Examiner***

ACKNOWLEDGEMENT

*We would like to express our sincere thanks to **Ms Neeru Sharma Ma'am** **(Supervisor)**, for her valuable guidance and support in completing our project. Her continued mentorship & handholding assistance during the entire phase of the completion of this project can't be described in words.*

We would also like to express our gratitude towards our **HOD (CSE Department, Chandigarh University)** ***Mr Sandeep Singh Kang sir*** *for giving us this great opportunity to do a project* **on *"Alzheimer's Disease prediction using Machine Learning"***. *Without his support and suggestions, this project would not have been completed.*

*Finally, we would also like to extend our heartfelt gratitude towards our co-supervisors & guides **Nongmeikapam Thoiba Singh sir & Megha Sharma Ma'am** for their precious mentorship & inputs which have made this project possible.*

Finally, we extend our gratitude towards our friends & family members who have been our constant source of motivation & energy during the development & deployment of this ambitious project.

Vikrant Dalvi- 20BCS1858

Rahul Choudhary- 20BCS1804

Satyam Kumar- 20BCS1857

Alok Raj- 20BCS1948

Deepshikha Kumari- 20BCS5297

TABLE OF CONTENTS

INTRODUCTION

Alzheimer's disease (AD) is a neurodegenerative disease and the most common cause of dementia in older adults. The part of brain that gets affected in this disease is hippocampus degeneration. Detection of Alzheimer’s disease at preliminary stage is very important as it can prevent serious damage to the patient’s brain. It becomes dangerous and sometimes fatal in case of people of 65 years of age or above.

The main objective of this paper is to use machine learning algorithms that is CNN and feature extraction and selection to predict the Alzheimer’s disease and build a useful model. The dataset is taken in the form of MRI images. The proposed approach detects the various stages of Alzheimer’s Disease such as moderate-demented and non-demented using CNN algorithm.

Alzheimer’s Disease (AD) is the most common cause of Dementia in people of the age 65 years and above. It is a progressive and irreversible neurological disease which follows a distinct pattern of brain damage as the disease progresses. Alzheimer’s disease is a very common type of dementia. Dementia is an umbrella term describing a variety of diseases and conditions that develop when nerve cells in the brain (called neurons) die or no longer function in a normal way. The death or malfunction of neurons causes abnormalities in one’s memory, behavior, and ability to think in a clear way.

In Alzheimer’s disease, these brain conditions eventually impair an individual’s ability to perform even basic functions such as walking, speaking, and swallowing. Development of AD can be classified into three stages. First, is the asymptomatic stage, changes in the brain, blood, or cerebrospinal fluid (CSF) may begin to occur without the patient showing any particular symptoms. After the first stage comes the second stage, that is mild cognitive impairment (MCI) stage, memory complaints and other cognitive behavior may start to be noticeable for the patients themselves and for close family or friends, which affects day to day activities but the symptoms are mild. In the final stage of the disease, or the dementia stage, memory, thinking, and behavioral symptoms are evident and significant, and it is noticeable. The neurons of brain start degenerating and the synapses are slowly dissolved. The patient loses the ability to respond to the environment. The hippocampus abnormally shrinks in its size 2.2 to 5.9 percent annually. AD is the 6th leading cause of death in the United States.

According to a report from the Alzheimer’s Association, AD and other dementias are predicted to cost the nation $1.1 trillion in 2050. Currently, 5.8 million Americans are living with AD and by 2050 this number is expected to rise to closely 14 million. It is difficult to manually diagnose AD, or any other types of dementia at an early stage before most of its symptoms are noticeable. Therefore, it is important to use computer analysis to analyze as much patient’s data as possible for a better evaluation and more accurate diagnosis.

The methods detect AD at a very late stage when all of the symptoms appear. However, this research will focus on developing an evolving framework to effectively diagnose and predict AD at a very early stage using the data collected for AD patients. The framework will continuously use large sets of related data to AD patient collected from multiple sources like medical sources, lifestyle and demography.

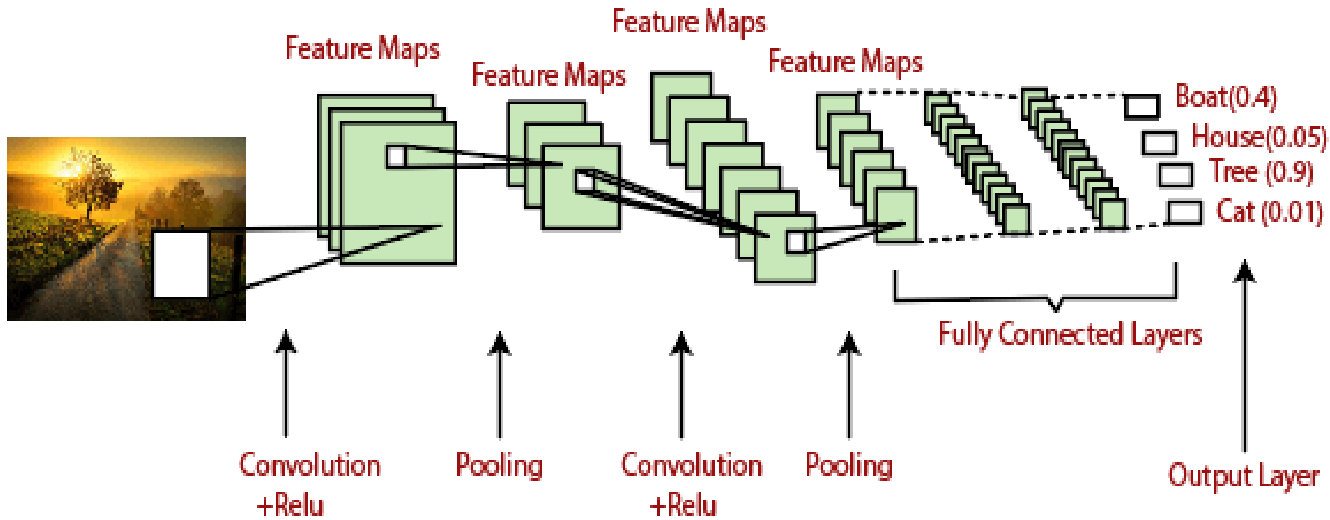
The datasets are taken in the form of MRI images. MRI, which reflects the spontaneous blood-oxygen-level dependent signal fluctuations when a subject is not performing an explicit task. The detection of Alzheimer's disease using conventional technique is time consuming, so we apply machine learning technique CNN to predict the Alzheimer's disease.

CONVOLUTIONAL NEURAL NETWORK

Convolutional Neural Network is one of the main categories to do image classification and image recognition in neural networks. Scene labeling, objects detections, and face recognition, etc., are some of the areas where convolutional neural networks are widely used.

CNN takes an image as input, which is classified and process under a certain category such as dog, cat, lion, tiger, etc. The computer sees an image as an array of pixels and depends on the resolution of the image. Based on image resolution, it will see as h * w * d, where h= height w= width and d= dimension. For example, An RGB image is 6 * 6 * 3 array of the matrix, and the grayscale image is 4 * 4 * 1 array of the matrix.

In CNN, each input image will pass through a sequence of convolution layers along with pooling, fully connected layers, filters (Also known as kernels). After that, we will apply the Soft-max function to classify an object with probabilistic values 0 and 1.



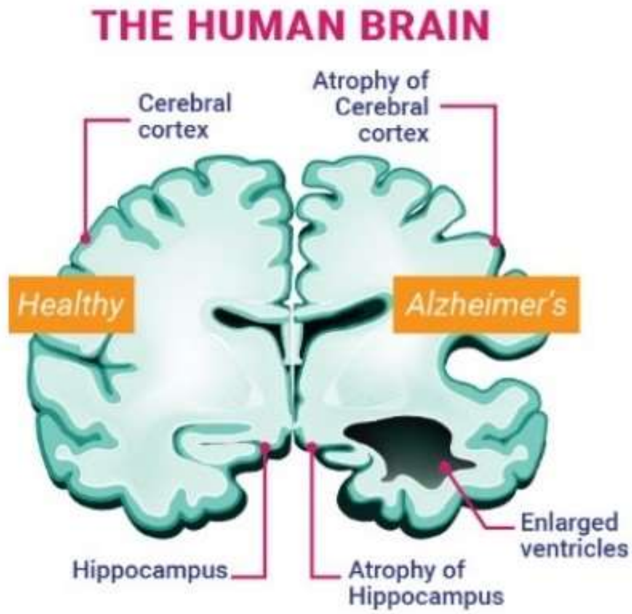
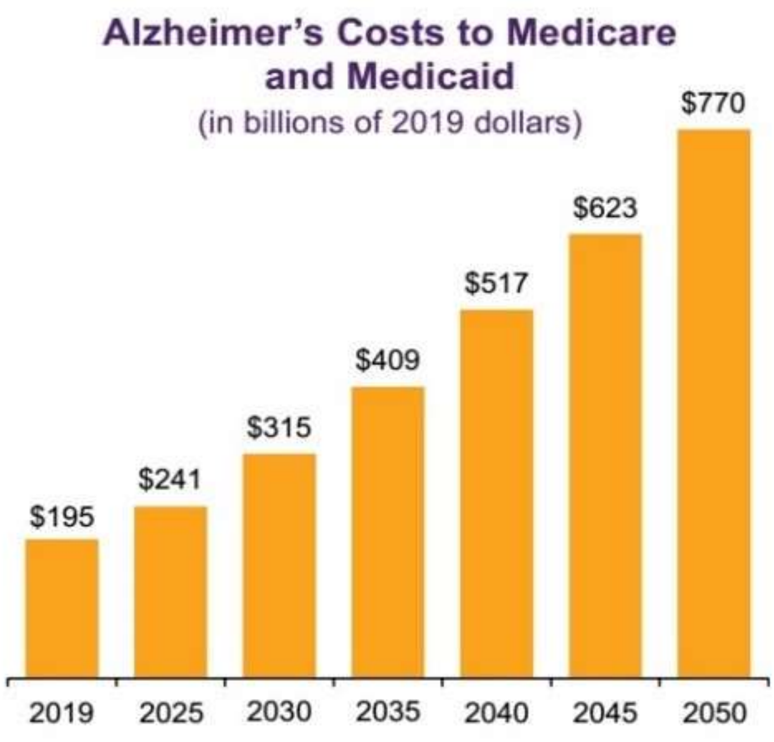


FIGURE 1: Progress of AD from MCI to severe AD [12]



CHAPTER 2

LITERATURE REVIEW AND BACKGROUND STUDY

TIMELINE OF THE REPORTED PROBLEM

Here is a timeline of some significant events and developments related to using deep learning for Alzheimer's detection:

- **2014:**

  Researchers at the University of Toronto develop a deep learning algorithm that can accurately identify Alzheimer's disease using MRI scans. The algorithm achieved a classification accuracy of 82% and outperformed other machine learning methods.

- **2015:**

  A study published in the journal Neuroimage uses a deep learning approach to predict the progression of Alzheimer's disease in patients. The researchers used MRI scans and achieved an accuracy of 84% in predicting whether a patient would progress from mild cognitive impairment to Alzheimer's disease.

- **2016:**

  Researchers at the University of California, Los Angeles develop a deep learning algorithm that can detect Alzheimer's disease in PET scans with an accuracy of 84%. The algorithm also outperformed other machine learning methods.

- **2017:**

  A study published in the journal Radiology uses a deep learning approach to predict the conversion of mild cognitive impairment to Alzheimer's disease using PET scans. The researchers achieved an accuracy of 92% in predicting conversion within three years.

- **2018:**

  Researchers at the University of California, San Francisco develop a deep learning algorithm that can predict Alzheimer's disease from brain images up to six years before clinical diagnosis with an accuracy of 82%.

- **2019:**

  A study published in the journal Neurobiology of Aging uses a deep learning approach to predict Alzheimer's disease from MRI scans. The researchers achieved an accuracy of 96% in identifying individuals with Alzheimer's disease and 94% in identifying individuals with mild cognitive impairment.

- **<u>2020:</u>**

  A team of researchers from the University of Iowa, Columbia University, and the University of California, San Francisco develop a deep learning algorithm that can detect Alzheimer's disease using speech samples. The algorithm achieved an accuracy of 88% in detecting Alzheimer's disease and 72% in detecting mild cognitive impairment.

- **<u>2021:</u>**

  A study published in the journal Nature Medicine uses a deep learning approach to predict Alzheimer's disease from retinal images. The researchers achieved an accuracy of 99% in identifying individuals with Alzheimer's disease and 96% in identifying individuals with mild cognitive impairment.

Overall, these studies demonstrate the potential of deep learning approaches for detecting Alzheimer's disease using a variety of medical and non-medical data sources. However, more research is needed to validate these approaches and develop practical tools for clinical diagnosis and early intervention.

PROPOSED SOLUTIONS

Alzheimer's disease is a complex neurological disorder, and predicting its onset can be a challenging task. However, machine learning techniques have shown promising results in predicting the likelihood of developing Alzheimer's disease based on various factors.

Here are some proposed solutions for Alzheimer's disease prediction using machine learning:

1. ANALYSIS OF NEUROIMAGING DATA:

Machine learning models can be trained on neuroimaging data, such as MRI and PET scans, to identify early changes in brain structure and function that are indicative of Alzheimer's disease. These models can use techniques such as deep learning to extract features from the images and predict the probability of Alzheimer's disease.

2. ANALYSIS OF GENETIC DATA:

Machine learning models can be trained on genetic data to identify genetic markers associated with Alzheimer's disease. These models can use techniques such as random forest, support vector machines, or neural networks to classify individuals based on their genetic profile and predict their risk of developing Alzheimer's disease.

3. ANALYSIS OF CLINICAL DATA:

Machine learning models can be trained on clinical data, such as cognitive tests and medical histories, to identify patterns and risk factors associated with Alzheimer's disease. These models can use techniques such as logistic regression, decision trees, or Bayesian networks to predict the likelihood of developing Alzheimer's disease based on the clinical data.

4. **INTEGRATION OF MULTIPLE DATA SOURCES:**

Machine learning models can integrate data from multiple sources, such as neuroimaging, genetics, and clinical data, to provide a more comprehensive prediction of Alzheimer's disease. These models can use techniques such as ensemble learning to combine the predictions from multiple models and improve their accuracy.

5. **EXPLAINABLE AI:**

Explainable AI techniques can be used to provide insights into how machine learning models are making their predictions. This can help clinicians and researchers understand the factors that contribute to Alzheimer's disease and develop more effective interventions.

These proposed solutions for Alzheimer's disease prediction using machine learning show promising results in identifying individuals at risk of developing Alzheimer's disease and may ultimately lead to earlier detection and more effective treatments.

BIBLIOMETRIC ANALYSIS

As we have seen from our previous modules, Alzheimer's disease (AD) is a progressive neurodegenerative disorder that affects the cognitive abilities of individuals, particularly memory, thinking, and behaviour. Deep learning (DL) algorithms have shown great promise in detecting AD from medical images, particularly magnetic resonance imaging (MRI) scans. Here are some of the key features, effectiveness, and drawbacks of AD detection using DL:

KEY FEATURES:

- Deep learning algorithms can analyse large volumes of data and identify patterns that are difficult for humans to detect.
- DL algorithms can learn to recognize complex relationships between different features in the image data, which is critical in detecting subtle changes in the brain.
- DL algorithms can be trained on large datasets, which helps to improve the accuracy of AD detection.

EFFECTIVENESS:

- Several studies have shown that deep learning algorithms can accurately detect AD from MRI scans, achieving high sensitivity and specificity.
- DL-based methods have been shown to outperform traditional machine learning approaches and human experts in AD detection.
- DL algorithms can detect AD at an early stage, which can help in the development of effective treatments.

DRAWBACKS:

- Deep learning algorithms require large amounts of annotated data for training, which can be time-consuming and expensive to obtain.
- DL-based methods may be sensitive to differences in data acquisition protocols, which can affect the accuracy of AD detection.

- DL algorithms can be prone to overfitting, where the model performs well on the training data but poorly on new data.

In conclusion, deep learning algorithms have shown great promise in detecting AD from MRI scans. While there are some limitations to this approach, the potential benefits of early AD detection using DL-based methods make it a promising area of research for future studies.

REVIEW SUMMARY

Alzheimer's disease (AD) prediction using machine learning (ML) is a rapidly growing field of research. Many studies have been conducted in recent years, which have achieved high accuracy rates in predicting AD using various ML techniques. These studies have used different types of data, including neuroimaging data, genetic data, and clinical data.

One of the most common ML techniques used for AD prediction is deep learning. Deep learning models have shown promising results in accurately predicting AD using brain imaging data, such as magnetic resonance imaging (MRI) and positron emission tomography (PET) scans. Other studies have used machine learning algorithms, such as support vector machines (SVMs) and random forests, to predict AD using a combination of genetic and clinical data.

Despite the promising results achieved by these studies, there are still some challenges that need to be addressed. One major challenge is the lack of large and diverse datasets, which can limit the generalizability of the results. Another challenge is the interpretability of the ML models, as it is often difficult to understand how the models arrive at their predictions.

Overall, AD prediction using ML has the potential to revolutionize the diagnosis and treatment of AD, but further research is needed to address the challenges and improve the accuracy and interpretability of the models.

PROBLEM DEFINITION

The problem definition for Alzheimer's disease (AD) prediction using deep learning involves developing a model that can accurately predict the presence of AD in individuals based on neuroimaging data, such as magnetic resonance imaging (MRI) and positron emission tomography (PET) scans. The goal is to develop a model that can aid in the early diagnosis of AD, which can improve treatment outcomes and potentially slow the progression of the disease.

The input to the deep learning model is typically a set of MRI or PET scans from an individual, which are processed to extract features that are relevant to AD. These features may include measures of brain volume, cortical thickness, and the presence of amyloid plaques. The output of the model is a binary classification indicating whether the individual is likely to have AD or not.

The main challenge in this problem is developing a deep learning model that can accurately distinguish between individuals with AD and those without, given the high degree of variability in neuroimaging data across individuals. This requires careful selection and pre-processing of the data, as well as the development of a robust deep learning architecture that can capture the relevant features of the data.

Another challenge is the interpretability of the deep learning model, as it can be difficult to understand how the model arrives at its predictions. This has implications for clinical practice, where clinicians may need to explain the basis for their diagnoses to patients and their families.

Hence, the problem of AD prediction using deep learning is an important and challenging issue with significant potential to improve diagnosis and treatment of the disease.

GOALS & OBJECTIVES

The primary goal of Alzheimer's disease (AD) prediction using deep learning is to develop an accurate and reliable model that can aid in the early diagnosis of AD, which can improve treatment outcomes and potentially slow the progression of the disease. In addition to this primary goal, there are several related goals and objectives of AD prediction using deep learning:

1. ***To identify the key biomarkers of AD:***
   Deep learning can help identify the most important features or biomarkers in neuroimaging data that are most predictive of AD. By understanding these biomarkers, researchers can better understand the underlying mechanisms of AD and develop more targeted treatments.

2. ***To improve the accuracy of AD diagnosis:***
   Deep learning models can achieve high accuracy rates in predicting AD, which can help improve the accuracy of clinical diagnoses.

3. ***To enable earlier detection of AD:***
   Deep learning models can detect subtle changes in neuroimaging data that may be indicative of AD in its early stages, allowing for earlier detection and intervention.

4. ***To personalize AD treatment:***
   Deep learning models can be used to predict an individual's risk of developing AD, which can help tailor treatment plans to the individual's specific needs and risk factors.

5. ***To provide insights into the progression of AD:***
   Deep learning models can be used to track changes in neuroimaging data over time, which can provide insights into the progression of AD and the efficacy of different treatments.

Effectively, the goals and objectives of AD prediction using deep learning are focused on improving our understanding of the disease and developing more effective strategies for diagnosis and treatment.

CHAPTER 3

DESIGN FLOW & PROCESS

EVALUATION & SELECTION OF FEATURES

In the context of Alzheimer's detection using deep learning, feature selection is a crucial step that can significantly impact the accuracy and generalizability of the model. Here are some steps that can be followed for the evaluation and selection of features:

1. **Data preprocessing**:

   The first step is to preprocess the data by cleaning and normalizing it to ensure that the features are in a consistent format and are suitable for deep learning models. This can involve steps such as removing missing values, standardizing the data, and converting categorical variables to numerical values.

2. **Feature extraction:**

   The next step is to extract relevant features from the preprocessed data. This can be done using techniques such as principal component analysis (PCA), linear discriminant analysis (LDA), or other feature extraction methods specific to the domain.

3. **Feature selection:**

   Once the features are extracted, it is important to select the most informative and relevant features for the deep learning model. This can be done using various feature selection techniques such as correlation-based feature selection (CFS), mutual information-based feature selection, or recursive feature elimination.

4. **Model training and evaluation:**

   After selecting the features, a deep learning model can be trained and evaluated using the selected features. The model can be evaluated using metrics such as accuracy, precision, recall, and F1 score.

5. **Model optimization:**

   Finally, the model can be optimized by tuning the hyperparameters such as the learning rate, number of hidden layers, and number of neurons in each layer.

In conclusion, the selection of relevant features is critical for the accurate and efficient detection of Alzheimer's disease using deep learning models.

DESIGN CONSTRAINTS

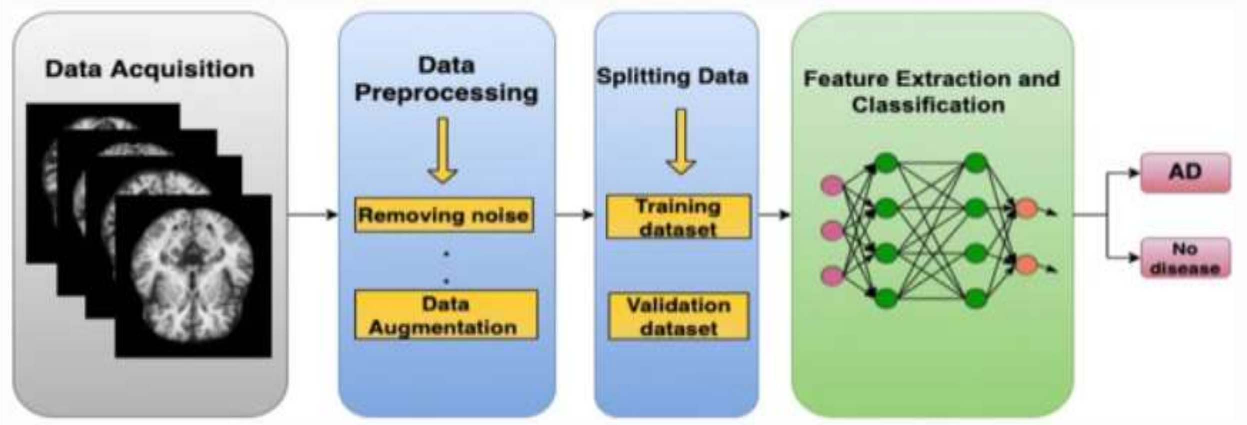
We found a study that developed a deep learning system to predict Alzheimer's disease using images and measurements from multiple ocular imaging modalities (optical coherence tomography, optical coherence tomography-angiography, ultra-widefield retinal photography, and retinal autofluorescence) and patient data.

We trained, validated, and tested a deep learning algorithm to detect Alzheimer's disease based on retinal photographs using data from 648 patients with Alzheimer's disease and 3240 individuals without the disease from 11 multicenter clinical studies in different countries. To the best of our knowledge, our study included the largest sample size and the most comprehensive metadata of patients with Alzheimer's disease for deep learning model development.

We used advanced deep learning techniques (example, unsupervised domain adaptation and feature fusion) to address two challenges:

(1) data distribution discrepancy between training and validation and testing datasets and
(2) individuals who have multiple retinal photographs from each visit, including optic nerve head-centered and macula-centered images of both eyes.

We also proposed deep learning models with different architecture for a practical application.



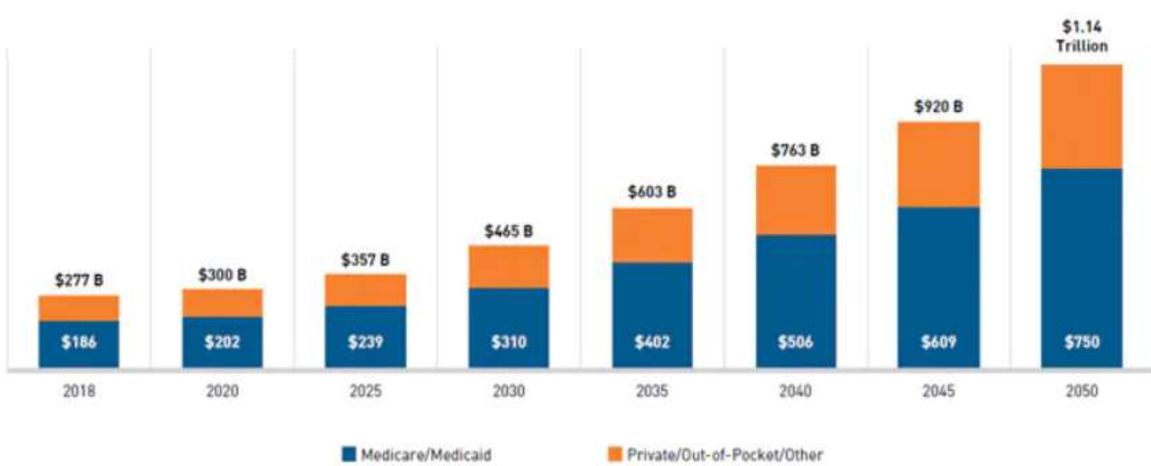
The Process of Diagnosis Alzheimer Disease

We show that deep learning is a promising tool to perform automatic early detection of Alzheimer's disease from MRI data. The proposed model is able to effectively identify CN and AD subjects based on MRI data, clearly outperforming the model based on more traditional features such as ROI volumes and thicknesses.

Alzheimer's affects someone socially because along with memory loss and other problems, increased anxiety is a common symptom of dementia. Someone who feels anxious is less inclined to be social, and may actually dread interacting with other people. Because the family is rarely one voice, family care becomes a juggling act between competing needs, loyalties, responsibilities, and demands. Isolation, unpredictability, fear, fatigue, and overwhelming loss of control are common social issues confronted by most families living with AD.

<u>The Economic and Cost considered of Alzheimer Disease: Managed Care Consideration</u>

As the disease progresses, individuals with AD require increasing levels of medical care, caregiver support, and eventually long-term care that may include home health, assisted living, nursing care, and hospice. It is estimated that approximately 75% of people with AD will be living in a nursing home at age 80 compared with 4% of the general population by this age. Cost of care is usually divided into direct and indirect costs. Direct medical costs associated with AD treatment include physician office visits, hospital admissions, emergency department visits, skilled nursing care, and medications. Long-term care including nursing homes and home healthcare account for the majority of direct costs associated with AD treatment. The average total annual costs in 2021 dollars for Medicare beneficiaries 65 years and older with AD or other dementias have been estimated to be $41,757, which is about 3 times higher than those without AD ($14,026).



<u>Projected Alzheimer diseases cost in billions</u>

<u>Home Safety Checklist for Alzheimer's Disease</u>

1. Display emergency numbers and your home address near all telephones.

2. Use an answering machine when you cannot answer phone calls, and set it to turn on after the fewest number of rings possible. A person with Alzheimer's disease often may be unable to take messages or could become a victim of telephone exploitation. Turn ringers on low to avoid distraction and confusion. Put all portable and cell phones and equipment in a safe place so they will not be easily lost.

3. A person with Alzheimer's disease often may be unable to take messages or could become a victim of telephone exploitation. Turn ringers on low to avoid distraction and confusion. Put all portable and cell phones and equipment in a safe place so they will not be easily lost.

4. Avoid the use of flammable and volatile compounds near gas appliances. Do not store these materials in an area where a gas pilot light is used.

5. Install secure locks on all outside doors and windows.

6. Hide a spare house key outside in case the person with Alzheimer's disease locks you out of the house.

7. Place red tape around floor vents, radiators, and other heating devices to deter the person with Alzheimer's from standing on or touching them when hot.

ANALYSIS & FEATURE FINALIZATION

The analysis and feature finalization phase are a crucial step in the development of deep learning models for Alzheimer's disease prediction. Here are some steps that can be followed during this phase:

1. **Data preprocessing:**

   The first step is to preprocess the data by cleaning and normalizing it to ensure that the features are in a consistent format and are suitable for deep learning models. This can involve steps such as removing missing values, standardizing the data, and converting categorical variables to numerical values.

2. **Feature extraction:**

   The next step is to extract relevant features from the preprocessed data. This can be done using techniques such as principal component analysis (PCA), linear discriminant analysis (LDA), or other feature extraction methods specific to the domain.

3. **Feature selection:**

   Once the features are extracted, it is important to select the most informative and relevant features for the deep learning model. This can be done using various feature selection techniques such as correlation-based feature selection (CFS), mutual information-based feature selection, or recursive feature elimination.

4. **Model training and evaluation:**

   After selecting the features, a deep learning model can be trained and evaluated using the selected features. The model can be evaluated using metrics such as accuracy, precision, recall, and F1 score.

5. **Model optimization:**

   Once the model has been trained, it can be optimized by tuning the hyperparameters such as the learning rate, number of hidden layers, and number of neurons in each layer. The model can also be evaluated using different algorithms and architectures to find the best combination for the problem at hand.

6. **Feature finalization**:

   After optimizing the model, the final set of features can be selected based on their importance in the model. This can be done using techniques such as feature importance or permutation importance.

7. **Validation:**

   Finally, the model can be validated on an independent dataset to ensure that it generalizes well to new data. This is an important step to ensure that the model is not overfitting to the training data.

In summary, the analysis and feature finalization phase involve preprocessing the data, extracting and selecting relevant features, training and optimizing the model, selecting the final set of features, and validating the model. These steps are crucial to ensure that the deep learning model is accurate and generalizes well to new data for Alzheimer's disease prediction.

DESIGN FLOW

The design flow for Alzheimer's disease prediction using deep learning involves several stages. Here is a general outline of the design flow:

1. **Data collection:** The first step is to collect data related to Alzheimer's disease, such as clinical and demographic data, imaging data, and genetic data. This data can be obtained from various sources such as electronic medical records, clinical trials, and public databases.

2. **Data preprocessing:** Once the data has been collected, it needs to be preprocessed to ensure that it is in a format suitable for deep learning models. This involves cleaning the data, removing missing values, and normalizing the data.

3. **Feature engineering:** After preprocessing, relevant features need to be extracted from the data. This can be done using techniques such as principal component analysis (PCA), linear discriminant analysis (LDA), or other feature engineering methods specific to the domain.

4. **Model selection:** The next step is to select a suitable deep learning model for Alzheimer's disease prediction. This can involve experimenting with different algorithms such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and deep belief networks (DBNs).

5. **Model training:** Once a suitable model has been selected, it needs to be trained on the preprocessed data. This involves feeding the data into the model and adjusting the model's weights and biases to minimize the error between the predicted and actual outputs.

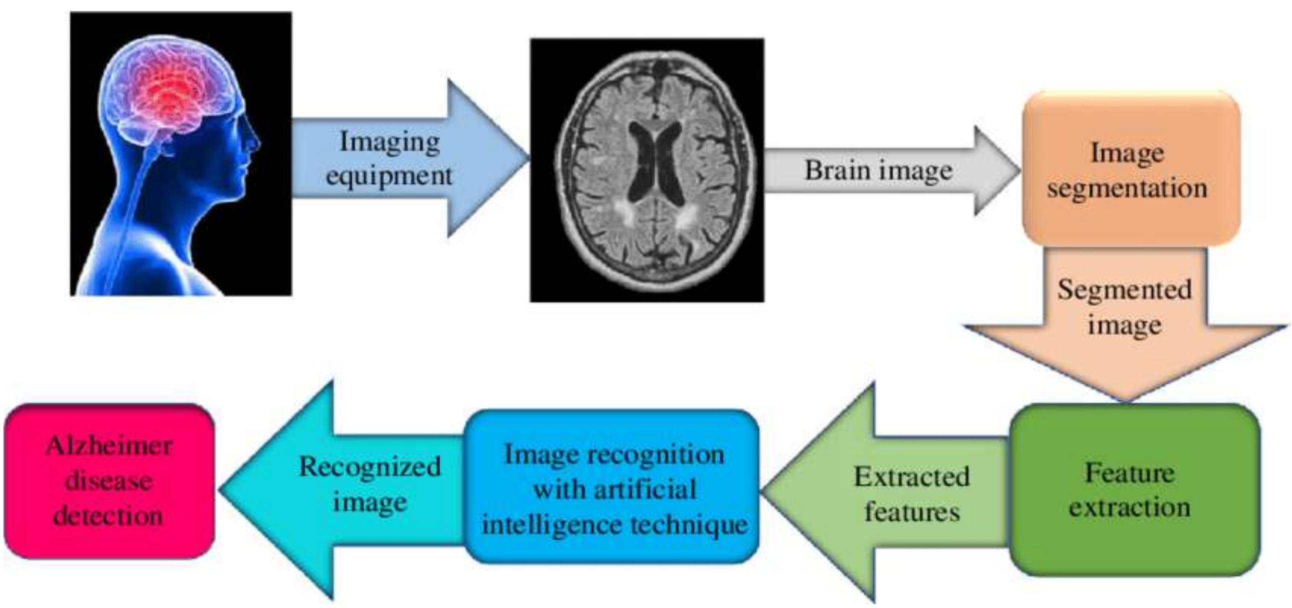
6. **Model evaluation:** After training, the model needs to be evaluated using metrics such as accuracy, precision, recall, and F1 score. This is done to determine the performance of the model and identify areas for improvement.

7. **Model optimization:** If the model's performance is not satisfactory, it can be optimized by tuning hyperparameters such as the learning rate, batch size, and regularization. This can be done using techniques such as grid search or random search.

8. **Model validation:** Once the model has been optimized, it needs to be validated on an independent dataset to ensure that it generalizes well to new data.

9. **Deployment:** Finally, the model can be deployed in a production environment, where it can be used to predict Alzheimer's disease in new patients. This involves integrating the model into a software application or web service and ensuring that it is secure and reliable.

In summary, the design flow for Alzheimer's disease prediction using deep learning involves data collection, data preprocessing, feature engineering, model selection, model training, model evaluation, model optimization, model validation, and deployment.



IMPLEMENTATION PLAN & METHODOLOGY

Implementing an Alzheimer's disease prediction model using deep learning involves several steps. Here is a methodology for implementing such a model:

1. **Define the problem and goals:** The first step is to define the problem and goals of the project. This involves identifying the target population, the features to be used for prediction, and the performance metrics for the model.

2. **Data collection and preprocessing:** The next step is to collect and preprocess the data. This involves cleaning the data, removing missing values, and normalizing the data. Feature engineering techniques can be used to extract relevant features from the data.

3. **Model selection:** Once the data has been preprocessed, a suitable deep learning model needs to be selected for Alzheimer's disease prediction. This can involve experimenting with different algorithms such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and deep belief networks (DBNs).

4. **Model training:** Once a suitable model has been selected, it needs to be trained on the preprocessed data. This involves feeding the data into the model and adjusting the model's weights and biases to minimize the error between the predicted and actual outputs.

5. **Model evaluation:** After training, the model needs to be evaluated using metrics such as accuracy, precision, recall, and F1 score. This is done to determine the performance of the model and identify areas for improvement.

6. **Model optimization:** If the model's performance is not satisfactory, it can be optimized by tuning hyperparameters such as the learning rate, batch size, and regularization. This can be done using techniques such as grid search or random search.

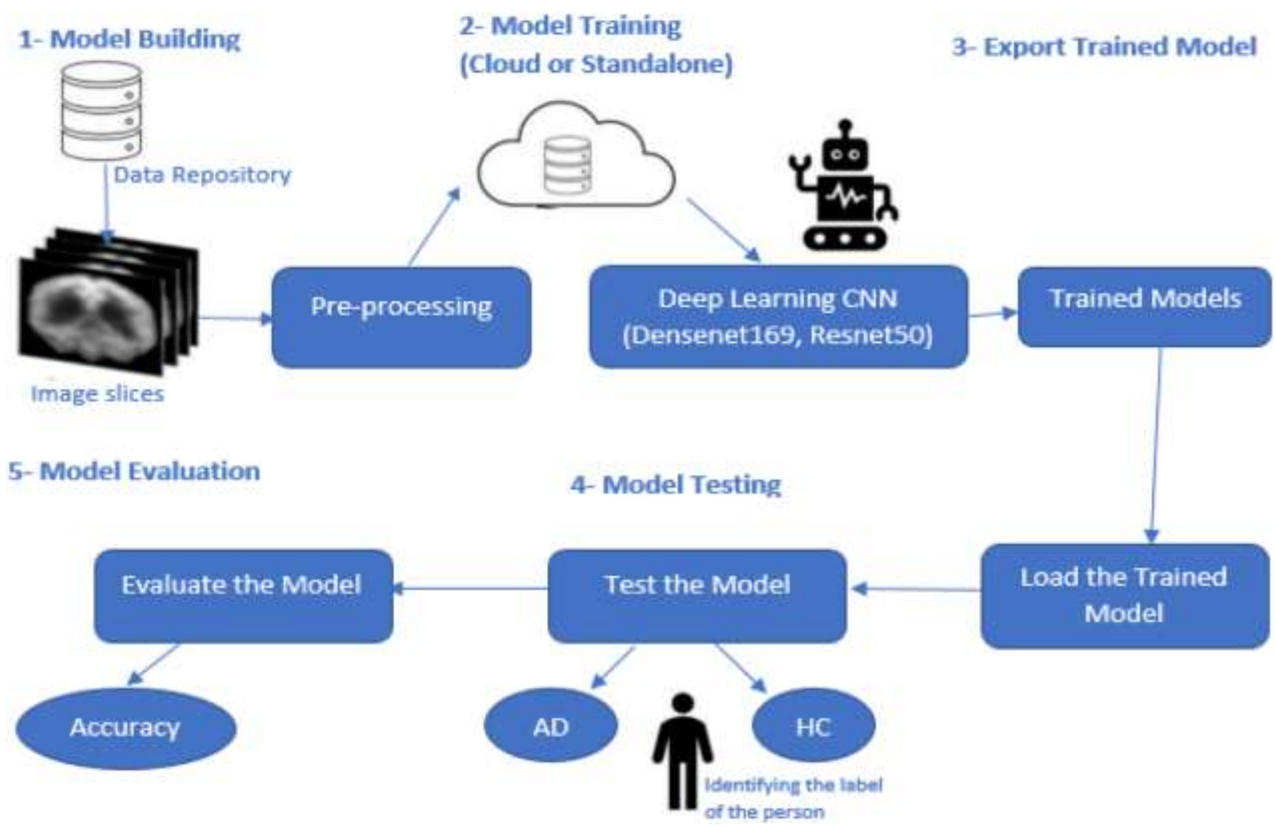
7. **Model validation:** Once the model has been optimized, it needs to be validated on an independent dataset to ensure that it generalizes well to new data.

8. **Deployment:** Finally, the model can be deployed in a production environment, where it can be used to predict Alzheimer's disease in new patients. This involves integrating the model into a software application or web service and ensuring that it is secure and reliable.

In terms of implementation plan, the following steps can be taken:

1. Define the project scope and timeline.
2. Assemble a team with expertise in deep learning, data science, and Alzheimer's disease.
3. Allocate resources such as computing power, software licenses, and data storage.
4. Develop a project plan and set milestones for each stage of the project.
5. Use agile methodology to manage the project, with regular meetings and progress updates.
6. Test the model at each stage of development to ensure that it meets the project goals and performance metrics.
7. Document the code, data, and results to ensure reproducibility and transparency.
8. Publish the results in a peer-reviewed journal or conference and share the model with the Alzheimer's disease research community.

In summary, implementing an Alzheimer's disease prediction model using deep learning involves defining the problem and goals, collecting and preprocessing the data, selecting and training a suitable model, evaluating and optimizing the model, validating the model, and deploying the model. A well-defined implementation plan and methodology can help ensure that the project is completed on time and meets the project goals and performance metrics.


1- Model Building
Data Repository
Image slices
2- Model Training
(Cloud or Standalone)
Pre-processing
Deep Learning CNN
(Densenet169, Resnet50)
3- Export Trained Model
Trained Models
Load the Trained Model
4- Model Testing
Test the Model
AD
HC
Identifying the label of the person
5- Model Evaluation
Evaluate the Model
Accuracy

CHAPTER-4

RESULTS ANALYSIS AND VALIDATION

We evaluate various performance metrics like accuracy, precision, recall and F1 score. To determine the best parameters for each model, we perform 5-fold cross-validation: Decision Tree, SVM, Random Forests, XGBoost and Voting. Finally, we compare accuracy of each model. Several metrics and techniques were used to identify overfitting and parameter tuning issues after the models were developed. Performance evaluations can either be binary or multiclass and are described using the confusion matrix. A learning model was developed to distinguish true Alzheimer's disease affected people from a given population and a novel Machine Learning classifier was developed and validated to predict and separate true Alzheimer's disease affected people.

The following evaluation measures were calculated using these components: precision, recall, accuracy, and F-score.

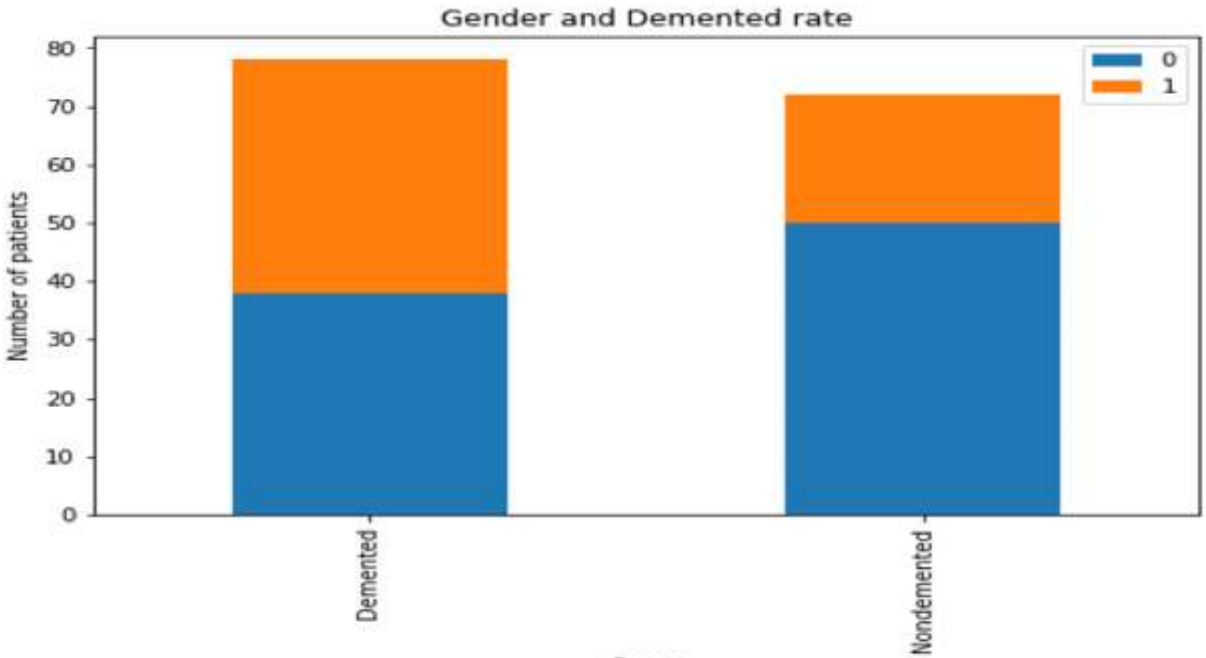
Based on this study, recall (sensitivity) is the proportion of people accurately classified as having Alzheimer's. The precision of Alzheimer's diagnosis is the rate of people correctly classified as not having the disease. Alternatively, F1 represents the weighted average of recall and precision, while accuracy represents the proportion of people correctly classified.

According to the results, the patient receives a report that tells him or her what stage of Alzheimer's Disease he or she is currently in. It is very important to detect the stages because the stages are based on the responses of the patients. In addition, knowing the stage helps doctors better understand how the Disease is affecting them.

This research used these environments, tools, and libraries to conduct its experiments and analysis:

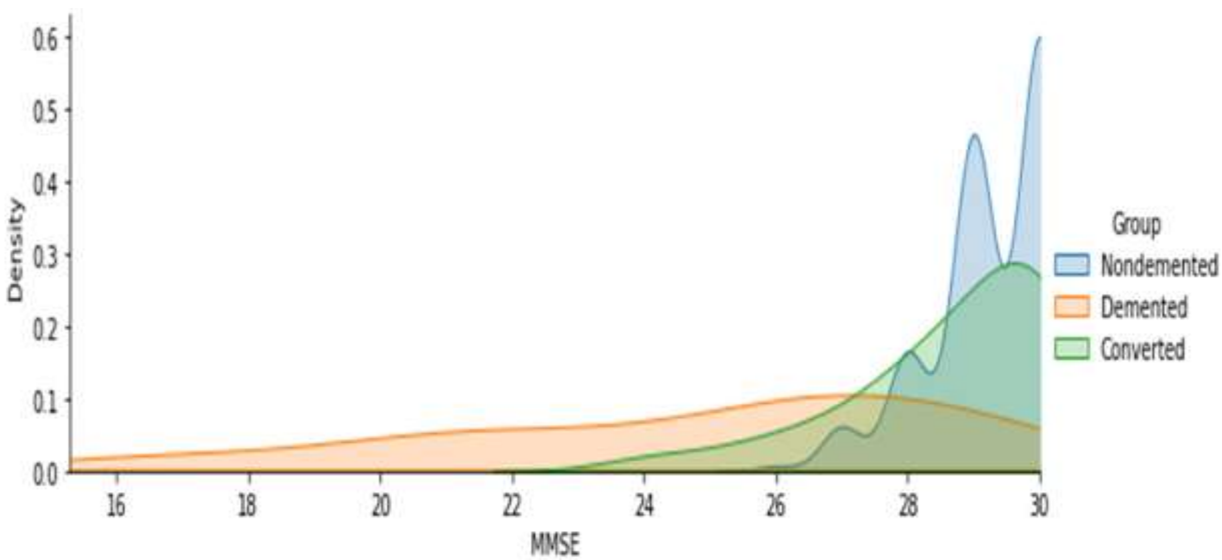
a) Environments Used: Python 3
b) Scikit-learn libraries for machine learning

The Figure 1 indicates that men are more likely than women to have dementia.



[Figure-1]

Figure 2 that the non-demented group had much higher MMSE (Mini-Mental State Examination) scores than those with dementia.

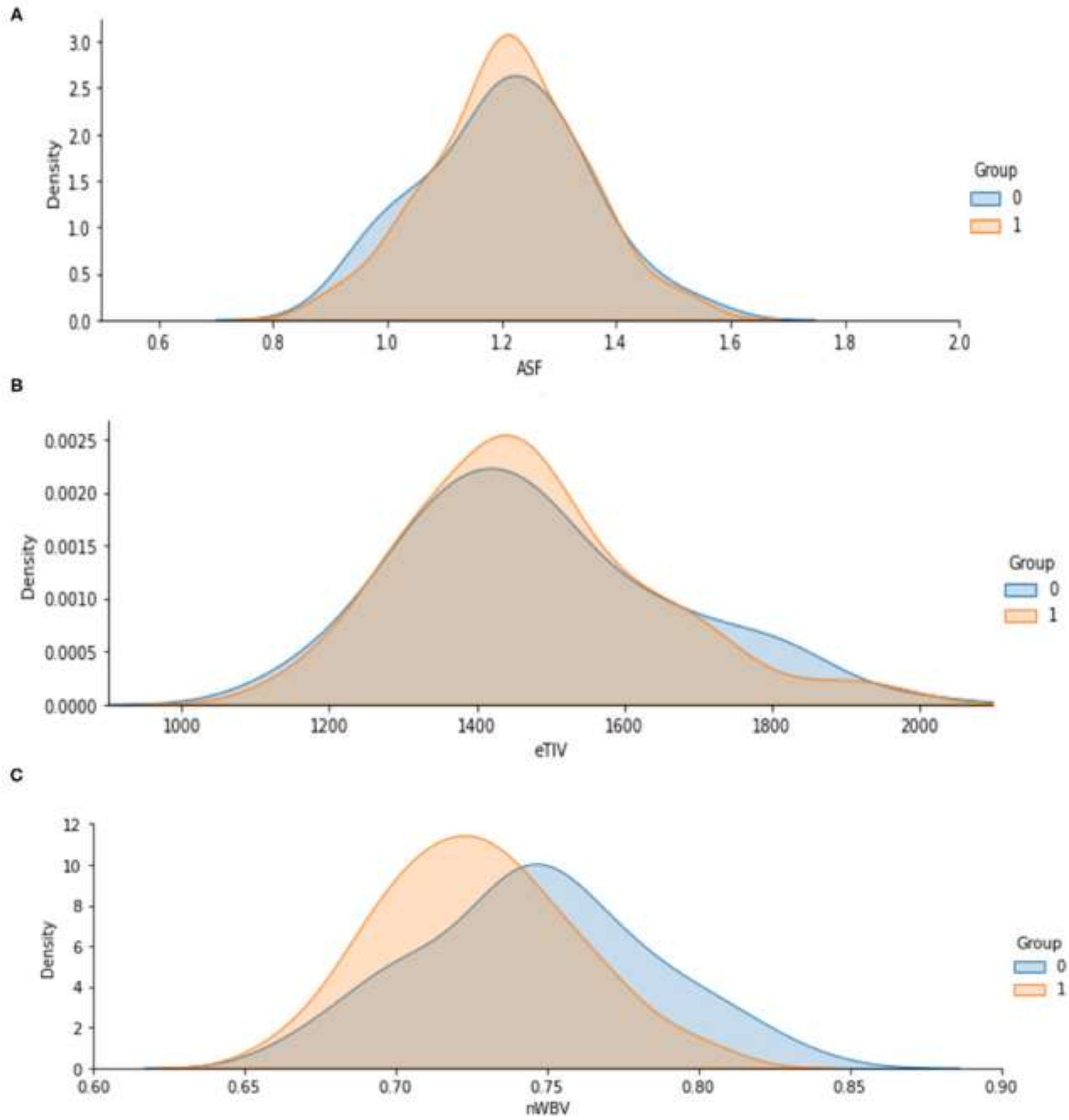


[Figure-2]

[Reference: https://www.frontiersin.org/articles/10.3389/fpubh.2022.853294/full#F5]

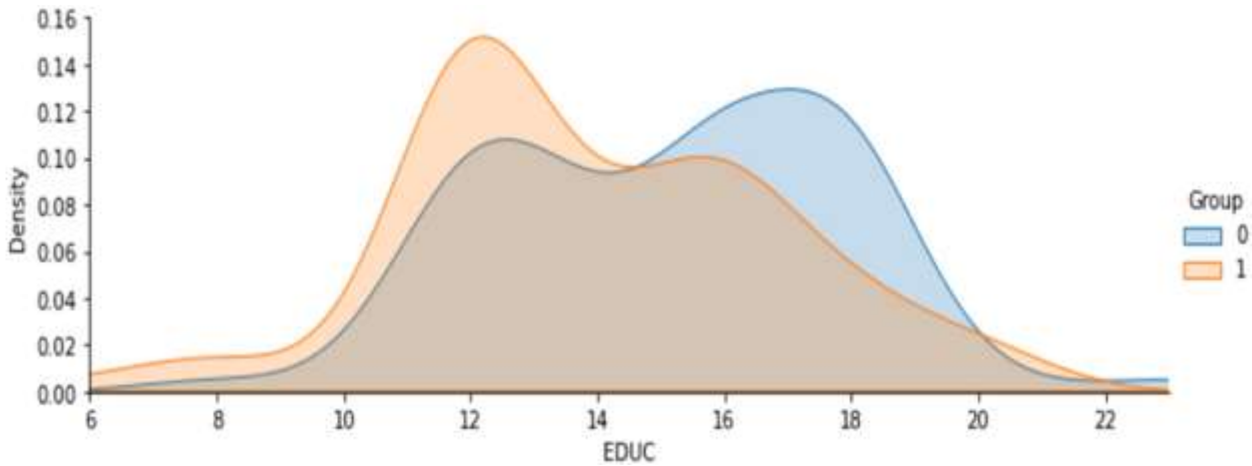
The Figures 3 (A–C) show the analyzed value of ASF, eTIV and nWBV for Demented and Non-demented group of people.

As indicated by the graph in Figure 3, the Non-demented group has a higher brain volume ratio than the Demented group. The reason for this is that the diseases influence the brain tissues causing them to shrink.



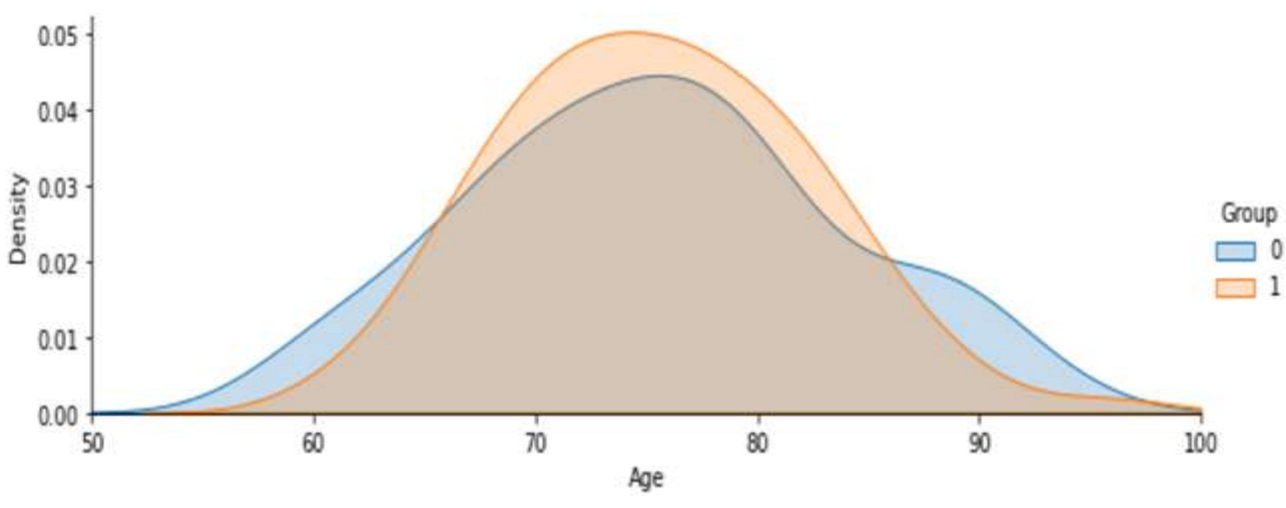
[Figure 3 (A, B, C)]

Figure 4 shows the analyzed results of EDUC for Demented and Non-demented people.



[Figure 4]

Figure 5 shows the analysis on age attribute to find the percentage of people affected based on the demented and non-demented group. It is observed that a higher percentage of Demented patients are 70-80 years old than non-demented patients. It is likely that people with that kind of Disease have a low survival rate. Only a few people are over 90 years old.



[Figure-5]

From the above all analysis on the attributes, the following are the summary on intermediate results:

1. It is more likely for men to have demented, or Alzheimer's Disease, than for women.
2. In terms of years of education, demented patients were less educated.
3. Brain volume in non-demented groups is greater than in demented groups.
4. Among the demented group there is a higher concentration of 70-80-year-olds than in the non-demented patients.

Given table shows the performance comparison of accuracy, precision, recall, and F1 score for different ML models.

Model	Accuracy	Precision	Recall	F1-score
Decision tree classifier	80.46%	0.80	0.79	0.78
Random forest classifier	86.92%	0.85	0.81	0.80
Support vector machine	81.67%	0.77	0.70	0.79
XGBoost	85.92%	0.85	0.83	0.85
Voting classifier	85.12%	0.83	0.83	0.85

The performance measures are defined as:

Accuracy: It is the measure of finding the proportion of correctly classified result from the total instances.

$$Accuracy\,(inPercentage) = \frac{TN+TN}{TP+TN+FP+FN} \times 100$$

Precision: This measures the number of correctly predicted positive rate divided by the total predicted positive rates. If the Precision value is 1, it is meant as a good classifier.

$$Precision = \frac{TP}{TP+FP}$$

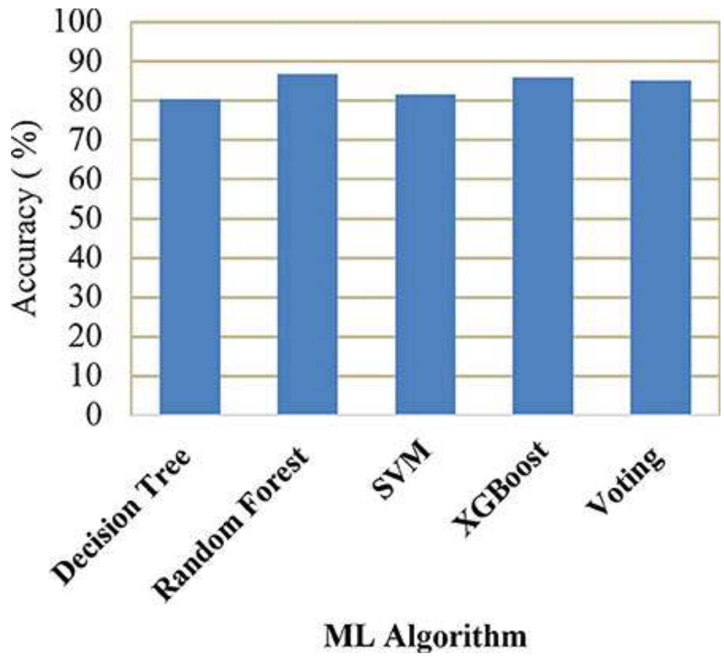
Recall: Recall is a true positive rate. If the recall is 1, it is meant as a good classifier.

$$Recall\,(inPercentage) = \frac{TP}{TP+FN}$$

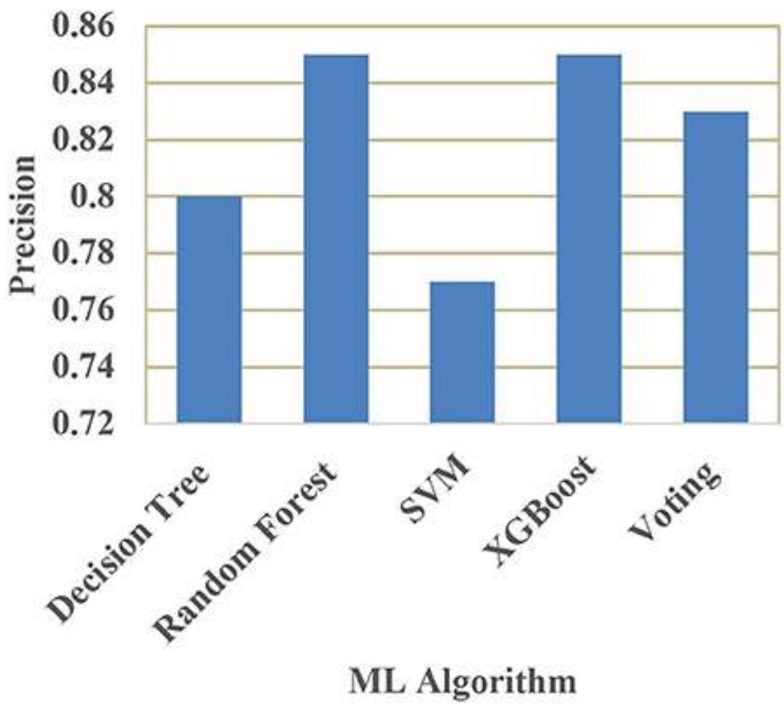
F1 Score: It is a measure which considers both Recall and Precision parameters. F1 score becomes 1 only when both the measure such as Recall and Precision is 1.

$$F1Score\,(inPercentage) = 2^{*}\,\frac{Recall^{*}\,Precision}{Recall+Precision}$$

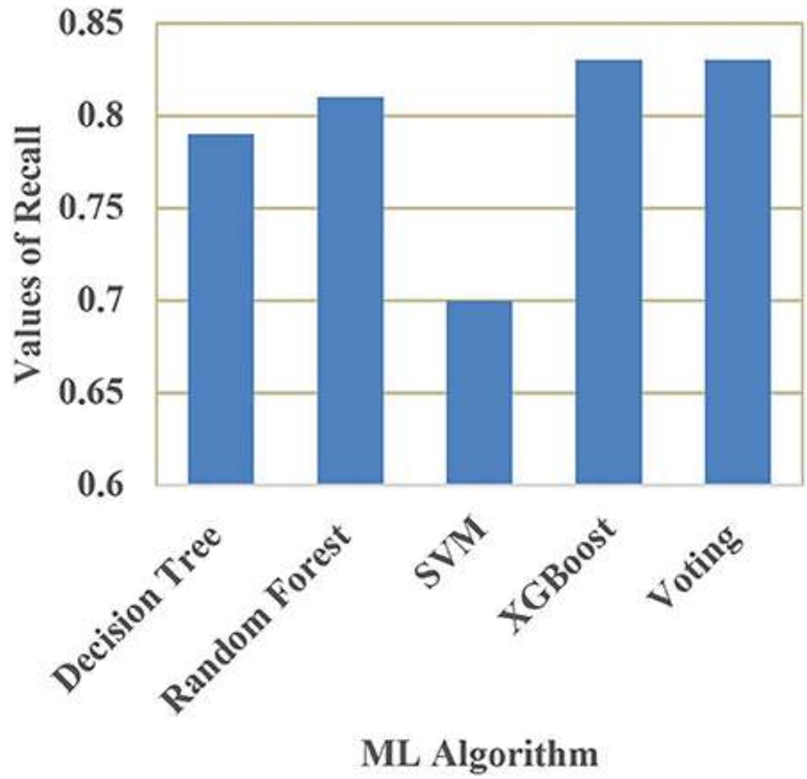
The comparative analyses among all the Machine Learning models in terms of accuracy, precision, recall, and F1 score are presented graphically in Figures 6–9 respectively.



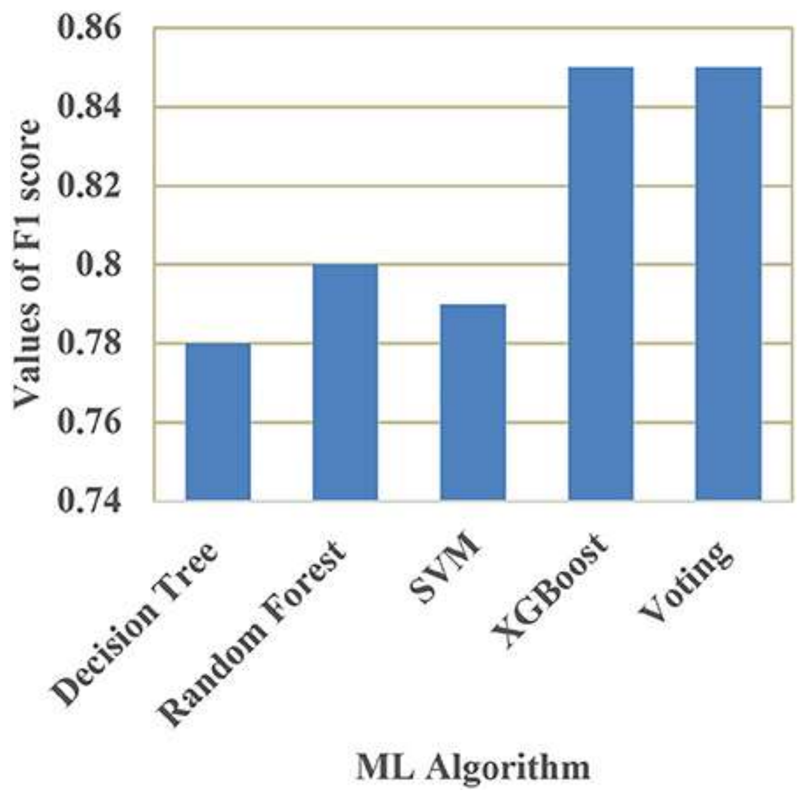
[Figure-6 (Comparison of Accuracy]



[Figure-7 (Comparison of Precision)]



[Figure-8 (Comparison of Recall)]



[Figure-9 (Comparison of F1 score)]

CONCLUSION

Alzheimer's is a major health concern, and rather than offering a cure, it is more important to reduce risk, provide early intervention, and diagnose symptoms early and accurately. As seen in the literature survey there have been a lot of efforts made to detect Alzheimer's Disease with different machine learning algorithms and micro-simulation methods; however, it remains a challenging task to identify relevant attributes that can detect Alzheimer's very early. The future work will focus on the extraction and analysis of new features that will be more likely to aid in the detection of Alzheimer's Disease, and on eliminating redundant and irrelevant features from existing feature sets to improve the accuracy of detection techniques. By adding metrics like MMSE and Education to our model, we'll be able to train it to distinguish between healthy adults and those with Alzheimer's.

REFERENCES

1. Sivakani GA, Ansari R. Machine learning framework for implementing Alzheimer's disease. Int Conferen Commun Signal Process. (2020) 12:588–92. doi: 10.1109/ICCSP48568.2020.9182220

PubMed Abstract | CrossRef Full Text | Google Scholar

2. Khan P, Kader MF, Islam SR, Rahman AB, Kamal MS, Toha MU, et al. Machine learning and deep learning approaches for brain disease diagnosis: principles and recent advances. IEEE Access. (2021) 9:37622–55. doi: 10.1109/ACCESS.2021.3062484

CrossRef Full Text | Google Scholar

3. Martinez-Murcia FJ, Ortiz A, Gorriz JM, 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. (2020) 24:17–26. doi: 10.1109/JBHI.2019.2914970

PubMed Abstract | CrossRef Full Text | Google Scholar

4. Prajapati R, Khatri U, Kwon GR. "An efficient deep neural network binary classifier for alzheimer's disease classification," In: International Conference on Artificial Intelligence in Information and Communication (ICAIIC). (2021), p. 231–234.

Google Scholar

5. Helaly HA, Badawy M, Haikal AY. Deep learning approach for early detection of Alzheimer's disease. Cogn Computing. (2021) 21:1–17. doi: 10.1007/s12559-021-09946-2

PubMed Abstract | CrossRef Full Text | Google Scholar

6. Yaffe K. Modifiable risk factors and prevention of dementia: what is the latest evidence. JAMA Intern Med. (2018) 178:281–2. doi: 10.1001/jamainternmed.2017.7299

PubMed Abstract | CrossRef Full Text | Google Scholar

7. Livingston G, Sommerlad A, Orgeta V, Costafreda SG, Huntley D, et al. Dementia prevention, intervention, and care. The Lancet. (2017) 390:2673–73. doi: 10.1016/S0140-6736<17>31363-6

PubMed Abstract | CrossRef Full Text | Google Scholar

8. O'Donnell CA, Manera V, Köhler S, Irving K. Promoting modifiable risk factors for dementia: is there a role for general practice? British J General Pract. (2015) 65:567–8. doi: 10.3399/bjgp15X687241

PubMed Abstract | CrossRef Full Text | Google Scholar

9. Sulaiman N, Abdulsahib G, Khalaf O, Mohammed MN. "Effect of Using Different Propagations of OLSR and DSDV Routing Protocols", In Proceedings of the IEEE International Conference on Intelligent Systems Structureing and Simulation. (2014), pp. 540–5.

Google Scholar

10. Deckers K, van Boxtel MP, Schiepers OJ, de Vugt M, Muñoz Sánchez JL, Anstey KJ, et al. Target risk factors for dementia prevention: a systematic review and Delphi consensus study on the evidence from observational studies. Int J Geriatric Psychiatry. (2015) 30:234–46. doi: 10.1002/gps.4245

PubMed Abstract | CrossRef Full Text | Google Scholar

11. Schiepers OJ, Köhler S, Deckers K, Irving K, O'donnell CA, Van den Akker, et al. Lifestyle for Brain Health (LIBRA): a new model for dementia prevention. Int J Geriatric Psychiatry. (2018) 33:167–75. doi: 10.1002/gps.4700

PubMed Abstract | CrossRef Full Text | Google Scholar

12. Vos SJ, Van Boxtel MP, Schiepers OJ, Deckers K, De Vugt M, Carrière I, et al. Modifiable risk factors for prevention of dementia in midlife, late life and the oldest-old: validation of the LIBRA Index. J Alzheimer's Dis. (2017) 58:537–47. doi: 10.3233/JAD-161208

PubMed Abstract | CrossRef Full Text | Google Scholar

13. Osamh Khalaf I, Ghaida M, Abdulsahib D. Energy efficient routing and reliable data transmission protocol in WSN. Int J Adv Soft Comput Applicat. (2020) 12:45–53.

Google Scholar