National College of Ireland

Alzheimer's disease can be diagnosed from healthcare information using machine learning

MSc Research Project
Data Analytics

Saikumar Kailasam
Student ID: x20151161

School of Computing
National College of Ireland

Supervisor: Proff. Jorge Basilio

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Saikumar Kailasam
Student ID:	x20151161
Programme:	Data Analytics
Year:	2022
Module:	MSc Research Project
Supervisor:	Proff. Jorge Basilio
Submission Due Date:	16/12/2021
Project Title:	Alzheimer's disease can be diagnosed from healthcare information using machine learning
Word Count:	7496
Page Count:	22

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	
Date:	25th January 2022

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Alzheimer's disease can be diagnosed from healthcare information using machine learning

Saikumar Kailasam
x20151161

Abstract

Alzheimer's disease appears to be one of a number of neurological conditions. Initially, this symptom appears to be normal, but over time it becomes more problematic. Alzheimer's disease is a well-known form of dementia. Alzheimer's is a tough disease to treat because there is no cure, and there is no known treatment for the disease. However, it is only later in the disease progression that the ailment is properly diagnosed. There are many ways in which early detection of an illness might delay its progression or symptoms. Predicting Alzheimer's disease based on parameters such as MMSE scores and the frequency of visits to the doctor is done using fully automated algorithms, such as those used in this study. For the prediction of continuous cognitive variables (AD,CN,LMCI) and categorical classification of three AD phases, machine learning methods including Ada boost, light gradient boosting algorithm, Artificial neural networks (ANN),Ada Boost, Bidirectional LSTM networks and Vote classifiers are utilized. Using data from the ADNI Clinical study, these forecasts have been developed. Metrics that can be used for evaluation include R2, RMSE; Accuracy; Precision; Recall; and F1-score. An approach that uses Vote classifier achieves an accuracy of 93 percent and the more recall values, making it more resilient than other techniques.Using multi-modal ensembles and other cutting-edge methodologies like Bayesian inference can increase model accuracy.

1 Introduction

In AD, the brain disease worsens over time and causes memory loss. Memory loss is just one indication of Alzheimer's disease, which also affects thinking and reasoning, forming judgements and decisions, planning, and carrying out familiar tasks. Even more importantly, it has the capacity to alter one's personality and conduct. As of yet, no one knows what causes Alzheimer's disease, which has no known cure. According to current research, this illness may have a strong hereditary basis. If you have an APOE gene mutation known as APOE e4, your risk of acquiring AD increases. AD is associated with low-level amyloid beta plaques and neurofibrillary tangles.AD is one of the most communicable illnesses on the planet. This disease causes a chronic brain dysfunction that gradually impairs memory, reasoning skills, and ultimately the capacity to do simple tasks. It is one of the key concerns in the healthcare business today and must be studied. Alzheimer's is the most common kind of dementia. Most often. Studies show that Alzheimer's disease is exceedingly contagious, and that current therapies can only slow the progression of symptoms and improve health. The current period is focused on finding

more efficient ways to deal with Alzheimer's disease progression. The major causes of AD are yet unknown. Studies show that hereditary variables such as family history, depression, and lifestyle factors promote the disorders. Alzheimer's disease predominantly damaged brain neurons Uddin and Ashraf (2019) To accurately identify Alzheimer's and other dementia-related infections, machine learning is essential. This saves both patients and doctors both time and resources. Experts will be able to identify potential dangers before the symptoms of Alzheimer's disease worsen. In the process, many lives could be spared, and injuries avoided thanks to the implementation of appropriate therapy. This frees up billions of euros, which can be used for analysis and government spending that can have a positive impact on GDP.

1.1 Motivation and Background

Alzheimer's disease will affect more than 100 million individuals worldwide by the year 2050, according to estimates (Prince et al. (2016)). The number of adults with the condition has increased from around 26 million to over 36 million in the previous ten years. This figure is expected to rise to 106 million in the next 34 years. Experts at the National Institutes of Health attribute the surge in patients to an aging population and the fact that most patients are diagnosed after the age of 65.A 'public health disaster,' according to specialists at the University of California, Los Angeles, is on its way. In the United States, there were 40 million adults 65 and older as of the 2010 census. However, this number is anticipated to nearly treble to 71 million by 2030. There will be 98 million people living there by 2060. The number of adults with the condition has increased from around 26 million to over 36 million in the previous ten years. This figure is expected to rise to 106 million in the next 34 years. Experts at the National Institutes of Health attribute the surge in patients to an aging population and the fact that most patients are diagnosed after the age of 65.A 'public health disaster,' according to specialists at the University of California, Los Angeles, is on its way. In the United States, there were 40 million adults 65 and older as of the 2010 census. However, this number is anticipated to nearly treble to 71 million by 2030. There will be 98 million people living there by 2060. Figure 1 By 2050, over 106 million persons will have Alzheimer's disease, making it a "public health epidemic." So long as humans live longer than ever before, scientists warn.Tadpole (Alzheimer's Disease Prediction of Longitudinal Evolution) competition data was used in this study Alzheimer's is a disease that affects a large number of people, and this challenge aims to create a novel technique of detecting the disease at an early stage.

Another goal is to predict who would acquire Alzheimer's symptoms in a certain age range (1-5 years). This research focuses on "rollover participants" from the Alzheimer's Disease Neuroimaging Initiative study 4. The researchers expect to correctly forecast outcomes in this trial using prior patient visits' data. MMSE, ADAS13, and head-size standardized ventricle volume may all be predicted using Ada Boost, Light Gradient Boosting, and ANN, among other tests. The three continuous characteristics were predicted before classifying persons into NC,MCI and AD.

1.2 Research Question

It is the major goal of TADPOLE to forecast possible clinical test data and diagnoses using previous clinical test measures.

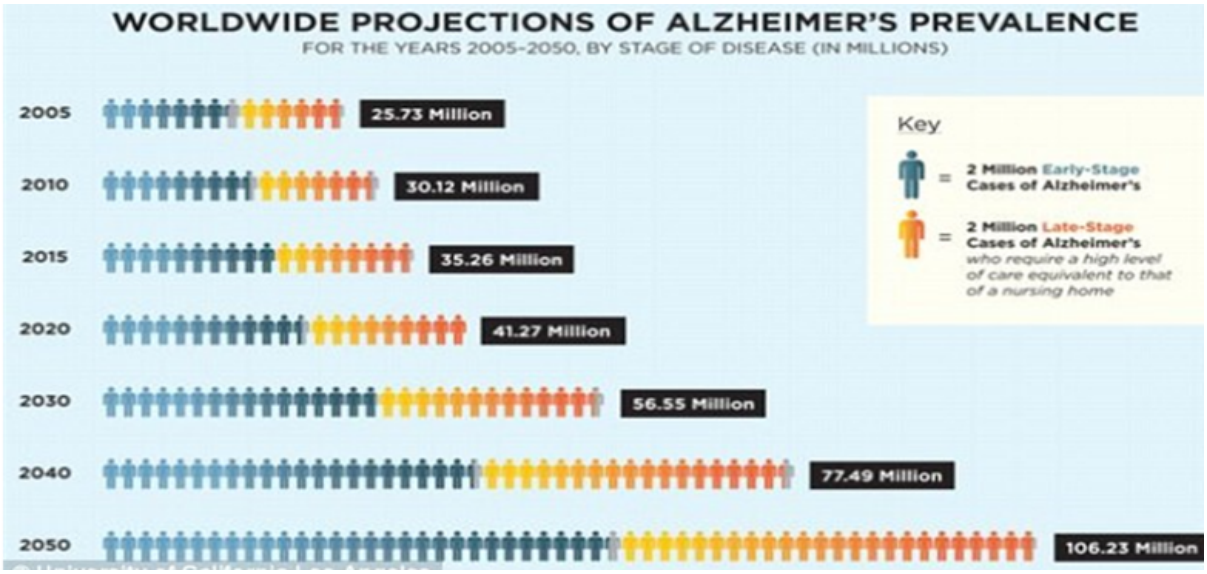

Figure 1: Worldwide Projection Of Alzheimer's Prevalence.

1. It is feasible to appropriately diagnose Alzheimer's disease based on the information provided by a patient.

Is it possible to investigate the model's decision-making process if it doesn't have a neural network design?

If you follow this structure, the entire paper should look like the following. It is stated in Section 3 of Section 2 how this paper's associated literature inquiry came to be; it also explains how the methodology, design specification, and assessment findings of the machine learning models used in part 4 and 5 of part 6 were developed; and it is explained in part 2 how part 6 came to be. 7 concludes with a glance ahead.

2 Related Work

In the related work there are 2 subsections subsection 3.1 and subsection 3.2 footnotes[1].

2.1 Alzheimer's diseases prediction on Machine learning models

In this Minhas et al. (2018) algorithm to predict development from MCI to Alzheimer's disease (AD) over a two- and three-year timeframe has been proposed in this research. These biomarkers include those obtained using magnetic resonance imaging (MRI) and non-magnetic resonance imaging (NM). To begin, we use autoregressive parameters derived from longitudinal training data to estimate previously unmeasured longitudinal biomarker values. The linear prediction coefficients can be calculated using one predictor or several predictors, according to one of three techniques proposed. More accurate predictions can be made when numerous predictors are used together. The SVM classifier is then used to distinguish between MCIp and MCIs based on the marker trajectories. AUCs of 89.93percent and 88.13percent for 2 year and 3 years follow up are achieved using the proposed framework, which is superior than other recently suggested techniques.

[1]URL: `https://purocleanpers.us/mold-alzheimers-disease-relationship`

Current research shows that NM and MRI measurements are better at predicting MCI progression in the short term, whereas NM alone is more powerful in the long run. MRI measures also fail to perform effectively in a longitudinal context, according to this study.

As a result, the study's ability to document changes in brain morphometry caused by slowly advancing diseases like Alzheimer's is limited. An advantage of this design is that extended follow-up periods will provide more data. Features selection strategies and boosting algorithms will be used to improve the performance of this framework in the near future. Other missing value imputation procedures will also be used in order to obtain a larger dataset for more robust modeling, and hence better conversion prediction.

In this particular Neelaveni and Devasana (2020) paper they worked on Machine learning model on Alzheimer disease is the most common neurodegenerative condition. Even if the symptoms are mild at the beginning, they worsen over time. People with Alzheimer's disease, a kind of dementia, are common. Unfortunately, there is no cure for this ailment, making it a difficult one to treat. Only at an advanced stage of the disease is a proper diagnosis possible. A patient's symptoms or condition may improve if they are detected early enough. Machine learning algorithms are used in this study to predict Alzheimer's disease based on factors such as age, frequency of visits, the MMSE, and education. As a result of the effective implementation of machine learning algorithms to forecast Alzheimer's disease, better accuracy results were achieved. It is able to discriminate between mild and severe cognitive impairment in the patient based on the model's predictions. This can be done in the future by merging brain MRI images and psychological characteristics to better forecast the condition using machine learning algorithms. When authors worked together, it is possible to detect the disease at an earlier stage and with more accuracy.

This is the case. Recent research shows that ML technique are reliable in predicting Alzheimer's illness. These algorithms predict AD using tests, demographic information, Positron emission tomography and MRI data. The researchers observed that RF was more accurate with a smaller collection of features than Genetic Algorithm. . The created data set comprises at least one characteristic from each biomarker category, increasing accuracy. In the clinical area, MCI is classed as MCI or AD, which has no impact on the clinical result. In this analysis, the MCI class accuracy was low due to the limited number of subject data points provided, which was a limitation of the study. More data into the prediction class and employing multiple RF feature selection models with variable parameters may be used to improve classification accuracy. A variable-length chromosome (VLC) could be used in the Genetic Algorithm to reduce feature subset size. Using criteria like fitness functions can tell which is the most accurate. In the future, machine learning algorithms can combine brain MRI pictures and psychological aspects to better predict illness progression. Using both procedures together allows for an earlier diagnosis.

In Li and Fan (2019) , the authors examined the possibility of predicting AD using a combination of a Random Forest ROI selection and a GRU AD prediction. However, they also provide a forecast of AD AD's relative cerebral ROI can also be identified.The ADPM accuracy ranged from 0.75 to 0.77, while for AD vs. NC, it ranged from 0.89 to 0.94. When it comes to AD, MCI, and NC prediction, longitudinal MRI delivers the most reliable data.

Analyses provided here consider AD prediction to be a two- or three-class problem, and each type of picture is considered separately.Incorporating multiple visual modalities can increase the number of classes precision. Additionally, their approach can be used or

modified to include a prediction of Alzheimer's disease (AD). The most current research employing ADNI as part of the challenge of neuroimages has achieved a precision of 62 in their classes, which is the highest level of precision ever achieved.

A multi-class symptomatic approach based on a profoundly sparse auto encoder was used in Jabason et al. (2018) to differentiate the clinical condition of Alzheimer's patients from their clinical highlights. By increasing sparsity and regularizing weight rot during the learning calculation, this method obscures the differences between classes. Those components of the analytic order that are most frequently acknowledged are chosen as the most important for the AD that is controlled. Through using early ADNI data set, they tested their proposed technique and discovered that it resulted in significant improvements in the execution of the plans. They used a 5-fold cross-approval procedure to demonstrate these benefits. A multi-nomial strategic layer is supplied to individuals with AD, MCI, and other mental health issues.. An exploratory 5-overlay cross-approval of the proposed order calculation in terms of accuracy, affectability, and explicitness has indicated further developed execution compared to the best-in-class procedures on account of Alzheimer's disease neuroimage drive. Aside from grouping precision (99.08 percent SVM and 98.90 percent MCI, as well as 90percent of CD), it is critical that their proposed technique is extremely sensitive to both SVM and Specificity values (90.01 percent for MCI and 95.77percent of AD). As part of the next phase of AD diagnosis research, researchers will change some procedures to account for missing examples and separate data from other modalities using a profound learning strategy.

2.2 Alzheimer's diseases prediction on Deep learning models

The authors of this work used a deep learning method to define the longitudinal dynamics of cognitive measures and generated prognostic models based on baseline hippocampal MRI measurements and the learnt longitudinal dynamics to predict the development of individual MCI participants to AD. Analyses have shown that the model's predicted development from MCI to AD using 1-year follow-up data has shown good performance in the field of prediction Optimization of the model and greater external data cohort validation will be the focus of upcoming development.

Alzheimer's disease is an incurable, degenerative neurological brain illness. Prevent brain tissue destruction by diagnosing Alzheimer's disease as early as possible. A few statistical and machine learning models have been used by Alzheimer's Disease researchers. Because Alzheimer's Disease Magnetic Resonance Imaging (MRI) data and ordinary healthy MRI data of elderly persons are quite comparable, detecting Alzheimer's Disease is extremely difficult. In recent years, Islam and Zhang (2018) improved deep learning approaches have proven human-level performance in a variety of disciplines, including medical picture processing. They presented a deep convolutional neural network for the diagnosis of Alzheimer's Disease utilizing brain MRI data. Many tests have been carried out in order to show that their suggested method is superior on the OASIS dataset versus comparable baselines.

Alzheimer's disease has been diagnosed using brain MRI data analysis, according to a study published in the jtheirnal Neurology. Multi-class classification is a considerable advance over binary classification, which has been the focus of most prior research efforts. The proposed network can be extremely helpful in the early detection of Alzheimer's disease. Although the suggested model has been evaluated exclusively on the Alzheimer's Disease dataset, we believe it can be successfully applied to other classification issues in

the medical area. Furthermore, the proposed method has the potential to be utilized to adapt CNN to additional areas with limited datasets. their approach will be tested on a variety of datasets, including the ADNI and other neurological illness diagnoses, in the future.

Deep learning algorithms such as Long short-term memory (LSTM) and the Recurrent Neural network (RNN) were used in Hong et al. (2019). 2019 study. The relevance of RNN is to relate prior data to current task. According to this publication, the LSTM was proposed as a method for predicting the model's shape. Instead of categorizing the current diagnosis status, a classification system could be used. Furthermore, the trials show that their model's performance outperforms that of the current methodologies. Furthermore, their solution is table-based, which can accommodate diverse data sizes. Another crucial factor for predicting AD progression was found to be the Cortical Thickness Average (TA).

In this Basher et al. (2021) An increasing number of persons over the age of 65 are developing Alzheimer's disease (AD). Memory function, stress development, and neurological disorders are all investigated in detail in studies focused on the hippocampus as a ROI. Furthermore, Alzheimer's disease and hippocampus volume shrinkage have been connected. Aside from amyloid beta (aß 42) protein and hippocampus volume atrophy (HVA), other biomarkers are being utilized to identify Alzheimer's disease. Researchers used structural magnetic resonance imaging (sMRI) to obtain slice-wise volumetric information from the left and right hippocampi to diagnose Alzheimer's disease. The proposed method combines a CNN and a DNN model. The left and right hippocampi were located using a two-stage Hough-CNN ensemble. A three-dimensional patch extracted from the hippocampus locations. Sagittal, axial, and coronal views distinguish 2D slices from 3D patches. A DVE-CN model extracts volumetric characteristics from each slice. This network was trained and tested using volumetric features. The proposed technique achieved average weighted classification accuracy of 94.82 percent and 94.02 percent using extracted volumetric characteristics from the left and right hippocampi. Furthermore, it has achieved AUC of 92.54 percent for the left hippocampi with 90.62 percent for the right hippocampi, respectively, in the study. Their strategy outperforms the others on the same dataset.

In Fritsch et al. (2019) explored the diagnosis of Alzheimer's disease using neural network language models in an automated approach. These are the most accurate diagnostic and monitoring techniques for patients with Alzheimer's disease that are needed. They put into practice the method that Sebastian Fritsch et al. (2019) . provide for improvising the method. LSTM and neural network language models are augmented utilizing a collection of methods to build N-gram language models (LMs), which are solely mathematical in nature (NNLM). The models are based on an evaluation of the difficulty of transliterating cookie theft images from the Pitt corpus of dementia. LMs and Alzheimer speakers are tested for transliteration. Finally, each transliteration is tested against the fundamentals of both LMs and Alzheimer speakers. In order to categorize patients with an accuracy of 85.6percent at the same error rates, the results provide information about the degree of difference in the data.

Deep learning models such as 1d CNN-LSTM were utilized by Nishikawa et al. (2021) in 2021 to build a Discriminatory dementia system through voice recognition. The accuracy rate is 90.8 percent, and the F-measure is 89.7 percent. As a general rule, today's people have a greater impact on the rest of the world. Their solution was to use the 1D CNN-LSTM. However, it will take an experienced doctor or psychologist about ten to

fifteen minutes to analyze this procedure, according to the author's research. It is usually a strain on the hospitals and the test participants.

Using the three-problems of AD, MCI, and NC, Kotturu and Kumar (2020) 2020 reported their work on the Forecasting of AD by using the scans like MRI which depicts the structural features of the brain to conduct binary classification of AD and we leverage the temporal aspect of the acquisition of clinical pictures to further improve their forecast. In specifically, pipelines for analyzing random collection of MRI, PET, and DTI images are proposed. Random forests, as they learned, are excellent at selecting features, especially when dealing with large-scale health research challenges. After random forests reduce feathers, the LSTM kernel's GRU focuses on the most time series of functions to make Alzheimer's disease predictions. their pipeline outperforms other methods in terms of efficiency.

In Martinez-Murcia et al. (2020) Some of today's most cutting-edge healthcare advances rely on CNN and RNN neural networks as well as deep learning. For the model's input, MRI brain pictures are used, with many hidden layers used to impose and locate them. This jumbled matrix had true, false, and false-positive information. The precision rate of each application is maintained throughout the process. Complex rules for models and MRIs are much easier to understand for experienced researchers than for novice researchers. It is encouraging to see how deep neural networks in healthcare might help solve problems and generate new models. This technique was used to recover features from a large data set (almost 2000 photos) without linear compression (0.63). Their accuracy rate was 84

In Hou et al. (2017) . 2018 on the novel longitudinal framework for the prediction of Alzheimer's disease (AD) clinical score forecasting for diagnosis (AD). In contrast to previous systems that only used data collected at a single point in time for the clinical score prediction, they proposed to employ imaging data from multiple time points instead. It is used in their research to analyze the 445 ADNI individuals' scores from the MRI data obtained at ftheir distinct time points using the Clinical scores, MMSE, and ADAS-Cog models that they have suggested in their research. Correntropy formulations and ensemble tactics have been found to outperform more conventional sparse methods in predicting outcomes, according to testing results. Some of the most important components of their study were addressed, while they used spatial-temporal information to help them estimate AD scores over time. The STGL model's definition of correntropy has proven successful in separating data from preparation information while also eliminating exceptions from the process entirely. AD-related biomarkers will be identified for future relapse efforts using advanced advancement calculations based on the model. Complete knowledge of SVR is employed so that longitudinal DA scores from selected highlights can be more accurately predicted.

3 Methodology

To predict and analyze upcoming clinical tests, many calculations are used in this evaluation for prior clinical test estimates. Many models are expected to be used continuously or in an all-out information consumption mode as classifier or groupings, for example, the light-gradient-boosting grouping of models. Artificial neural networks of brain pathways. It is depicted in the accompanying that the accompanying strategy. Figure 2 represents methodology.

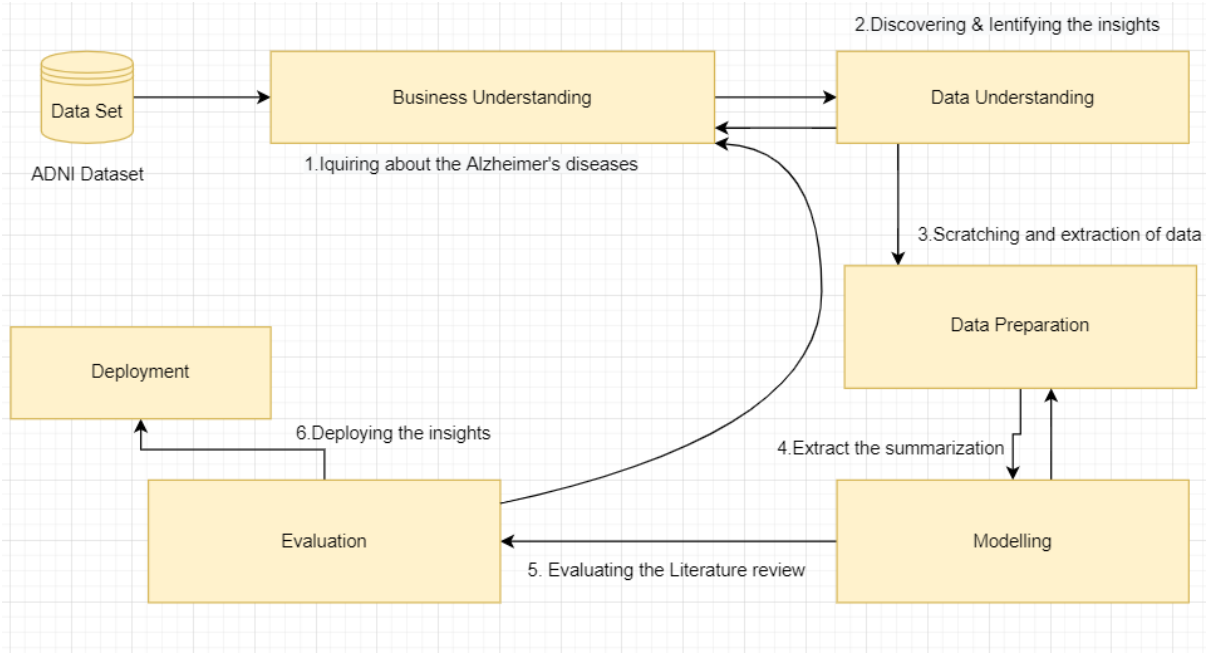

Figure 2: Methods for conducting the study have been proposed

3.1 Business Understanding

Artificial Intelligence (AI) is being used in the healthcare business to help solve some of the most pressing problems, such as diseases that kill millions of people around the world. When it comes to developing new ways to detect disease in its early stages, the utilization of MRI, CT scans, and patient clinical data has gained traction in the global research community. A dementia-related disease known as Alzheimer's is the topic of this research. Detecting Alzheimer's in its early stages is the only approach to prevent or slow down its progression, as there is no treatment for the illness. Alzheimer's disease progresses in three stages: normal cognition, mild cognitive impairment, and Alzheimer's disease (AD). Ada boost, artificial neural networks, Bi-directional LSTM.light gradient boosting, and a neural network were all applied to the TADPOLE ADNI data to help in the prediction and diagnosis of Alzheimer's.

3.2 Data Acquisition

The data collection is based on a freely available database of Alzheimer's disease information. There is an application process that must be completed before one can use the database, as it is only meant for research purposes. For the period of July 2005 through May 2021, all North American patients' examination data was compiled from the ADNI 1 through ADNI 3 databases. It's a database of patient study data, including MRI, PET, genetic, clinical, cognitive, and blood biomarkers, that can be used to track the progression of Alzheimer's. If you are interested in participating in the TADPOLE Challenge, you can get pre-processed, curated data from the Alzheimer Disease Neuroimaging Initiative (ADNI) by downloading the dataset (marinescu2020alzheimers). The TADPOLE dataset contains 1907 variables for 1737 different participants in the study. 8715 records were found.

3.3 Data Preparation and Pre-processing

When a dataset does not fill in uniformly, it is necessary to clean it, which means that data from all of the patients' visits is not included. There are two types of datasets: testing and training validation. Prepare information that includes certain significant features that reflect an individual's central segment data, such as age, gender, and so on, for model planning and model execution assessment.

In Table 1 demographic data

Table 1: demographic data such as RID, PTGENDER and AGE are included

label	**data description**	type of the data	Units
VISCODE	Candidate Visit code	Nominal	NA
PTETHCAT	Candidate Ethnicity	Nominal	NA
PTMARRY	Candidate Marital status	Nominal	NA
RID	Candidate Roaster ID	Numeric	NA
PTGENDER	Candidate Gender	Nominal	NA
PTRACCAT	Candidate Race	Nominal	NA
PTEDUCAT	Candidate Education	Nominal	NA
AGE	Candidate Age	Numeric	NA

In Table 2 an example table is provided.

3.4 Data Cleaning

After implementing.na() to see if there are any missing values in the data set, I discovered that there were no missing values in the data set when I examined the data frames for the presence of duplicates using the method df.duplicates () EDA, feature engineering, and one-hot encoding take place after the df.drop duplicates() function is used to remove any duplicates, and the data set is split into X and Y trains using K-fold CV with the percentages of 0.2 percent and 0.8 percent, respectively. I can also remove the non-essential features in this particular data set, for example, the Dx code for submission has nearly nine columns that are not required for the diagnosis of AD because our classification problem is not impacted by this columns , so I have removed the columns by using the append function. the columns that have been removed include 'directory.id', 'Subject', 'RID', 'Image.Data.ID', 'Modality', 'Visit'.

Figure 3.It is possible to examine and drop rows and columns containing null values in various ways using the dropna() method.

Figure 4. cleaning of data.

3.4.1 Feature Engineering

Date cleaning: To improve the quality of data, many variables in the dataset include a lot of missing data, which can't be used for imputation, as it will degrade its accuracy.

The chance that a prediction is correct is represented by the Threshold or Cut-off. It's a compromise between both the risk of false positives and the risk of false negatives, as the name suggests.The training dataset can be thresholded and customized versions of machine learning algorithms can be developed to solve an imbalanced classification challenge.this is the best way to find the TPR and the FPR.For example, when employing

```
[ ]  dat = dat.dropna()
     dat.isnull().sum().sum()

     0
```

Figure 3: Dropping the Null values.

```
[ ] dat.head(3)
```

	directory.id	Subject	RID	Image.Data.ID	Modality	Visit	Acq.Date	DX.bl	EXAMDATE	AGE	PTGENDER	PTEDUCAT	PTETHCAT	PTRACCAT	APOE4	MMSE	imputed_genotype	APOE Genotype	Dx Codes for Submission
0	4702245ea294ce5d4e9b8a87027dfdf4	011_S_0003	3	32237	MRI	1	9/1/05	AD	9/12/05	81.3	Male	18	Not Hisp/Latino	White	1	20	True	3,4	AD
1	2e89e352af743597b2368c412e0f6de2	022_S_0004	4	64631	MRI	1	9/22/05	LMCI	11/8/05	67.5	Male	10	Hisp/Latino	White	0	27	False	3,3	MCI
2	90419199306997753de8042f1fd55e38	011_S_0005	5	32246	MRI	1	9/2/05	CN	9/7/05	73.7	Male	16	Not Hisp/Latino	White	0	29	True	3,3	CN

```
[ ] #removed unwanted columns
```

```
X = dat
Y = dat['DX.bl']
del dat

#Remove unnecessary columns (features), remove first 9 columns and 'Dx codes for submission'
remove_columns = list(X.columns)[0:9]
remove_columns.append('Dx Codes for Submission')
print('Removing columns:', remove_columns)

X = X.drop(remove_columns, axis=1)

features = list(X.columns)
X.head(5)
```

```
Removing columns: ['directory.id', 'Subject', 'RID', 'Image.Data.ID', 'Modality', 'Visit', 'Acq.Date', 'DX.bl', 'EXAMDATE', 'Dx Codes for Submission']
```

	AGE	PTGENDER	PTEDUCAT	PTETHCAT	PTRACCAT	APOE4	MMSE	imputed_genotype	APOE Genotype
0	81.3	Male	18	Not Hisp/Latino	White	1	20	True	3,4
1	67.5	Male	10	Hisp/Latino	White	0	27	False	3,3
2	73.7	Male	16	Not Hisp/Latino	White	0	29	True	3,3
3	80.4	Female	13	Not Hisp/Latino	White	0	25	True	3,3
4	73.9	Female	12	Not Hisp/Latino	White	1	24	True	3,4

Figure 4: Removing the unnecessary columns

ROC Curves and Accuracy Curves, the greatest or optimum threshold can be determined directly. Other times, a grid search can be used to find the best threshold level and fine-tune it.the major purpose of threshold is to the curve is created by plotting the prediction scores against a series of rising thresholds, allowing us see how accurate and reliable our predictions are in identifying those who will test positive for the positive (minority) category. I need to import the sciPy library in order to use thresholds after the stats procedure is complete. FPR and TPR can be calculated using threshold() method. When calculating percentages, a threshold value of 60 percent was used to compare columns with blank or NaN values. To remove columns with more than the threshold number of missing values, Python code is used. TEMPQC UCSFFSL, for example, had PASS present 99 percent of the time and across multiple diagnostic categories. using panda's library interpolate function, linear interpolation for numerical variables is used to fill in the blanks of missing column values If column X has 5-row values, 3 rows have values, and the remaining 2 vacant rows can be filled by interpolating the 3 values present.

One Hot Encoding: To represent the five stages of Alzheimer's disease (AD, CN, EMCI.LMCI.SMC) using binary numbers 0 and 1, it is applied on category variables such as DX bl and PTGENDER. This is quite helpful for EDA and modeling.

Scaling up and standardizing features: An algorithm like the SVM is quite sensitive to the size of features. With normalization, the range of values in the column will be narrowed from 0 to 1. Making each feature's values bell-shaped and separated by unit variance is known as standardization. Almost all variables, except for PTID Key, EXAM DATE, and DX, are normalized for PCA and later for modeling purposes.

3.5 Exploratory Data Analysis (EDA)

A good EDA is a crucial part of the data science process, much like data wrangling is. To begin to understand the data and patterns within it, as well as to begin to explain the story behind it, this is your opportunity. We often approach datasets with a certain hypothesis or problem in mind, and an insightful EDA will illustrate why that hypothesis is important, but it may also uncover fascinating characteristics linked to the hypothesis/problem that may not be available a priori from our intuition and domain experience.

Figure 5. People over the age of 72 have the most records with 600. A close approximation to a normal distribution can be found.

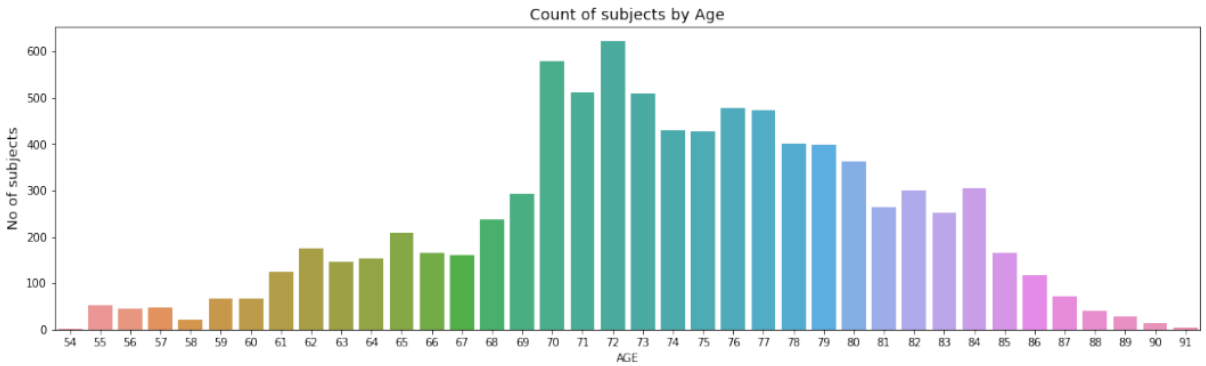

Figure 5: Count of patients by age group

Most of the individuals are in the 68-84 age group, as seen in This bar chart Figure 6 demonstrates how people have Alzheimer's disease at different stages of DX. DX cognitive levels are broken down into three stages, which are: 1) There are records of Alzheimer's Disease under AD. There are two types of cognitive impairment: Early Mild Cognitive Impairment and Late Mild Cognitive Impairment (LMCI). As a group, CN has two groups: CN and SMC. The number of people with AD, CN, and MCI cases.

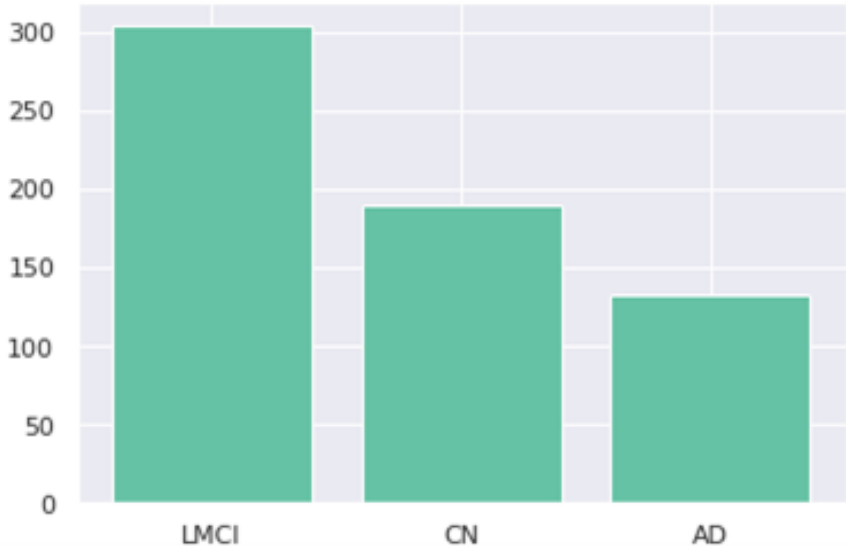

Figure 6: Number of people classified according to their stage of Alzheimer's disease

The Seaborn python module, a mix of box whisker and kernel density plot, is used to make violin plots. As can be seen in Figure 7, there is a strong correlation between gender and age. The median score for men and women is 72 and 74, respectively, as indicated by the white dot in the thick gray bar in the middle of the graph. Between the ages of 68-78, males and females make up half of the population in data sets (for both genders). Using the kernel distribution, we can conclude that most of the data falls between the age range of 70 to 84. It is more likely that data lies in age ranges 55-68 and 85-90 if the kernel area is smaller.

3.6 Principal Component Analysis (PCA)

PCA is a way for reducing the dimensions of a large data collection by translating a large set of variables into a smaller one that retains most of its information. This method is called Principal Component Analysis (PCA). The approach to dimensionality reduction is to exchange a little accuracy for simplicity, which naturally results in a lower number of variables in a dataset. As a result, machine learning algorithms can process data more quickly and with less effort when working with smaller datasets. To summarize, the goal of PCA is simple: decrease the number of variables in a data collection, while retaining as much information as feasible. There are many PCA-specific variables, except for PTID Key, EXAM DATE and DX, those are standard. Using this technique, the number of columns was decreased, In this data set only a small amount of data is loss.

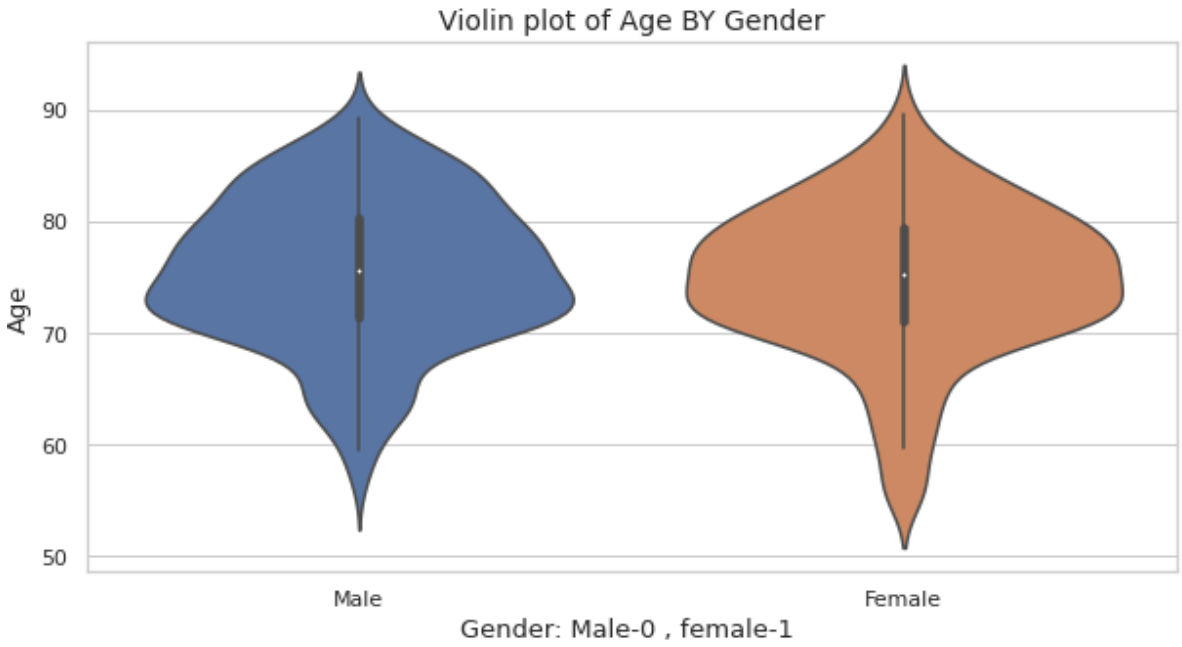

Figure 7: Distribution by age In terms of Gender

4 Design Specification

Data gathering, data pre-processing, building regression and classification models, evaluation of model results, and visualization are the four different phases of our research framework, which can be seen in Figure 8

Downloading TADPOLE ADNI data, which is a zip file that contains input and output CSV files, as well as target and test CSV files (comma separated values), is required for the data collection stage. These files will be needed at various points during data model building and analysis. On Google Colab Pro, the downloaded data is kept in Google drive, which can subsequently be mounted to be fetched and read in the Python script. Because it includes several stages such as data cleaning, data wrangling, and data preparation such as removing variables or using linear interpolation to replace missing values, data pre-processing is essential. Then, using Principal Component Analysis (PCA), the dimensionality of the data is reduced by normalization and standardization (PCA). Finally, the correlation map is used to examine for any relationships between the variables in the dataset. The third stage is modeling several techniques such as Ada Boost, light Gradient Boosting, and Artificial Neural Networks to predict the clinical scores of Alzheimer's Disease Assessment Scale Cognitive Subdomain (ADAS-Cog 13), Ventricle volume normalized scores, and Mini-Mental State Examination scores (MMSE). Ada Boost, light Gradient Boosting classifier, and Ada Boost have been used to accurately categorize the different phases of Alzheimer's disease. To properly evaluate the results of the models applied using K-fold cross-validation, the evaluation step must be carried out first. It has been possible to use a variety of regression and classification measures. Data has been visualized in Colab Notebook using Seaborn, and Matplotlib in Colab Notebook to gain insights.

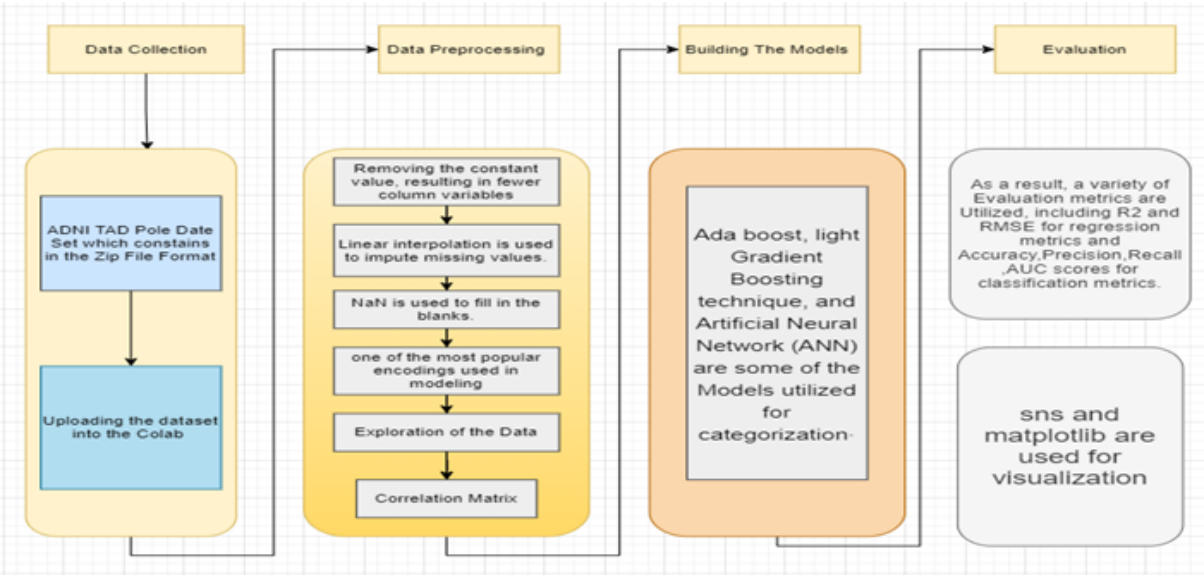

Figure 8: Designed research flow diagram

5 Implementation

Cleaning data in preparation for the models: - Approximately fifteen hundred biomarkers were collected from seventeen hundred thirty-seven persons during twelve thousand seven hundred forty-one timely visits to ADNI. There are eight thousand seven hundred fifteen records and seventeen thirty-seven columns in the incoming data set. Columns with 99 percent constant values are removed from the data, resulting in the deletion of 166 of the original 1737 columns. Missing values and NaN values that are less than 60percent of the total will be filtered out of the data set. There are 8715 rows and 716 columns in all. To do PCA, we import the Standard Scaler preprocessing function from the sklearn python package. The number of 710 columns is decreased to 281 columns and the same 8715 rows with principal component after PCA using sklearn library import PCA is used. Models that have been put into action: - Following are the various machine learning models utilized in this study: Ada Boost, Light Gradient Boosting, and Artificial Neural Network (ANN) to predict, categorize, and justify the research questions

5.1 Experiment with Light Gradient Boosting Machine Algorithm

Once the data has been divided up into test and training segments of 20 percent and 80 percent respectively, import relevant libraries to the lightgbm classifier and define the dataset by supplying the n samples=1000, n featues=20, n informative=15, n redundant=5, and random state=7 to it. the LGBM classifier was afterwards defined and evaluated with the parameters n splits=10, n repeats=3, and random state=1. Finally, the model's accuracy was evaluated. Figure 9 shows the implementation of Light GBM

Treebased learning is used in the Light GBM framework, which boosts performance by using gradients. Light GBM grows trees vertically while other algorithms grow trees horizontally, which means that Light GBM grows tree leaf-wise while other algorithms grow level-wise. It will only grow the leaf that has the greatest delta loss. This algorithm is more efficient at reducing loss when the same leaf is grown repeatedly. As the name

```
# evaluate lightgbm algorithm for classification
from numpy import mean
from numpy import std
from sklearn.datasets import make_classification
from sklearn.model_selection import cross_val_score
from sklearn.model_selection import RepeatedStratifiedKFold
from lightgbm import LGBMClassifier
# define dataset
X_train, y_train,= make_classification(n_samples=1000, n_features=20, n_informative=15, n_redundant=5, random_state=7)
# define the model
model = LGBMClassifier()
# evaluate the model
cv = RepeatedStratifiedKFold(n_splits=10, n_repeats=3, random_state=1)

n_scores = cross_val_score(model, X_train, y_train, scoring='accuracy', cv=cv, n_jobs=-1)
# report performance
print('Accuracy: %.3f (%.3f)' % (mean(n_scores), std(n_scores)))
```

```
Accuracy: 0.925 (0.031)
```

Figure 9: Light GBM with Alzheimer's Diseases

suggests, it's called Light Gradient Boosted Machine. It's an opensource library that makes the gradient boosting algorithm much faster and more effective than it used to be. A type of automatic feature selection and an emphasis on boosting examples with larger gradients are added to the gradient-boosting algorithm by the LightGBM extension. This has the potential to significantly speed up training and improve predictive performance. As a result, LightGBM has become the standard algorithm for machine learning competitions when working with tabular data for regression and classification predictive modeling. As a result, it bears some responsibility for the rise in popularity and widespread adoption of gradient boosting methods in general, as well as Extreme Gradient Boosting (XGBoost).

Light Gradient Boosted Machines are being developed for the categorization of Alzheimer's disease in this project and the accuracy of the model is 92.5percent

5.2 Experiment with Ada Boost

Ada boost classifiers libraries and kfold from sklearn libraries were imported and used to make n splits=10 and the base estimator as the decisionTreeclassifier and generated the classification report with precision, recall, f1 score and support. As part of an Ensemble Method in Machine Learning, AdaBoost or Adaptive Boosting is a technique known as AdaBoost. When it comes to AdaBoost, the most typical method utilized is decision trees with one level, which implies there is only one split. Decision Stumps are another term for these trees . When it came to binary classification, AdaBoost was the first and most successful boosting algorithm ever created. It is abbreviated as AdaBoost, which stands for Adaptive Boosting. AdaBoost is a widely popular boosting strategy that combines numerous "weak classifiers" into a single "strong classifier" that is used to improve classification accuracy. Yoav Freund and Robert Schapire were the ones who came up with the idea. In addition, they were awarded the Gödel Prize in 2003 for their achievements.Figure 10 represents implementation of Ada Boost.The accuracy of the model for Ada Boost is 53.96percent the precision value is 66.86 , Recall is 66.08 ,

f-score is 65.23 are the metrics for Ada boost

```
[ ]  from sklearn.tree import DecisionTreeClassifier
     from sklearn.ensemble import AdaBoostClassifier
     from sklearn.datasets import make_classification
     from sklearn.linear_model import LogisticRegression
     from sklearn.ensemble import RandomForestClassifier
     from sklearn.neighbors import KNeighborsClassifier
     from sklearn.neural_network import MLPClassifier

     from sklearn.model_selection import cross_val_score


     clf = RandomForestClassifier()
     clf2= KNeighborsClassifier(n_neighbors=10)
     #clf3=MLPClassifier()
     clf4=AdaBoostClassifier(base_estimator=DecisionTreeClassifier())
     clf_list=[clf,clf2,clf4]
```

Figure 10: Implementation of Ada Boost

5.3 Experiment with Artificial Neural Networks

Importing the ANN framework into Tensorflow allows us to construct Aritifical neural networks with Tensorflow. It is now possible to construct an ANN, but only in terms of a series of layers, by importing the Keras libraries and packages and setting up the Artificial Neural Network. In the theory section, I explained that ANNs are constructed with completely connected layers. To train the ANN, add the input layer and the first hidden layer after initializing the ANN, and then use the epochs=200,1000,2000,batch size=100,validation split=0.3 command-line parameters. In the end, make a prediction about the results of the test set.

Data must be normalized to the range of zero to one as sigmoid transfer function requires, for training errors caused by a discrepancy in quantity between input and output data to be kept to an acceptable minimum. As a result of the range approach used, all data was normalized to a value between 0 and 1. One set of data was used for training (70 percent of samples) while the other set was used for testing once they had been properly transformed (30 percent of the samples). This ratio is picked from among the other two alternative ratio combinations that have been used previously (Venugopalan et al. (2021)). A 66.7 percent training input and a 33.3percent testing input was used in the first combination. A 75 percent training and a 25 percent testing input was used in the second (testing). In contrast, the best ANN performance was achieved by using only 70 percent of the data for training and the remaining 30 percent for evaluation. In this study, a three-layer neural network was constructed to accommodate for the network's complexity, training time, and "over-fitting" (one input layer, one hidden layer

```
[ ]  from tensorflow.keras.models import Sequential
     from tensorflow.keras.layers import Activation,GRU,LSTM,Bidirectional,Concatenate,Dropout
     batch_size=10
     def get_basic_model():
       model = Sequential()
       model.add(Dense(1000,activation='relu'))
       model.add(Dropout(0.2))
       model.add(Dense(400,activation='relu'))
       model.add(Dropout(0.2))
       model.add(Dense(200,activation='relu'))
       model.add(Dropout(0.2))
       model.add(Dense(200,activation='relu'))
       model.add(Dropout(0.2))
       model.add(Dense(3,activation='softmax'))

       model.compile(optimizer=tf.keras.optimizers.Adam(),
                     loss=tf.keras.losses.SparseCategoricalCrossentropy(),
                     metrics=['accuracy'])
       return model
```

Figure 11: Implementation of ANN

```
Epoch 199/200
117/117 [==============================] - 1s 6ms/step - loss: 0.6609 - accuracy: 0.7086 - val_loss: 0.7240 - val_accuracy: 0.6159
Epoch 200/200
117/117 [==============================] - 1s 6ms/step - loss: 0.6234 - accuracy: 0.6971 - val_loss: 0.8080 - val_accuracy: 0.6093
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense (Dense)               (None, 1000)              23000

 dropout (Dropout)           (None, 1000)              0

 dense_1 (Dense)             (None, 400)               400400

 dropout_1 (Dropout)         (None, 400)               0

 dense_2 (Dense)             (None, 200)               80200

 dropout_2 (Dropout)         (None, 200)               0

 dense_3 (Dense)             (None, 200)               40200

 dropout_3 (Dropout)         (None, 200)               0

 dense_4 (Dense)             (None, 3)                 603

=================================================================
Total params: 544,403
Trainable params: 544,403
Non-trainable params: 0
_________________________________________________________________
```

Figure 12: Implementation of ANN

and one output layer). The network inputs for the input layer consist of 18 neurons (18 variables that differed statistically substantially between the cases and the controls). Weighed summation is carried out by each neuron. In this case, the activation function was a sigmoid function (f(x)=11+ex) with the value 11. The most extensively used Back propagation (BP) algorithm was used to train ANNs. Most experts agree that a sample size of 5–10 times the number of input variables is needed to assure BP-ANN network model reliability and external validity, and our sample size fits this requirement. To determining the correctness of the ANN model, bootstraps with 1000 replicates were used.Figure 11 and Figure 12 shows importing of necessary libraries.The accuracy of the model For Artificial Neural Networks is 60.93percent

5.4 Experiment with Bi-directional LSTM

Imported Bidirectional function from the machine learning sklearn library is used with data that has been preprocessed for both the train and validation data. later added classifier Global Average Pooling1D, Dense, Dropout and the 128 neurons are used to train the models and epochs=200, batch size=100, validation split=0.3. There are a number of ways you may change the number of epochs in a model. For example, you can use 50 epochs, 100 epochs, and so on. However, if you use those numbers, you'll notice a decrease in the model's accuracy and precision and recall values. It is built on the concept of offering two forward as well as back LSTMs for each training sequence, with both of them being connected to the output neuron, which is called Bi-LSTM. As a result, every point in the input pattern of the output layer is given a full historical and future context. The deep learning model of bi-directional LSTM, like backward LSTM, uses forward computation. Contrary to what is shown in this example, the two hidden layers' input sequence is inverted. Before the output information may be updated, both of the concealed layers must have passed. shows the simplest form of the bidirectional LSTM. The bi-directional LSTM attentional process employed to anticipate the patient's state. As soon as the data is entered, it is processed by dropout before being passed to the hidden layer for processing. The data processed by BI-LSTM is separated into two pieces in the hidden layer. The total number of param for bi-directional LSTM 544,403, The model accuracy is around 69.54percent

5.5 Experimenting with Ensemble Vote Classifier

People use Ensemble Vote Classifier to classify data that comes from the sklearn library all the time to figure out what it is.The weight initialization method I choose for your Ann can affect how soon or if it accumulates at all. Even though these weights' input variables are only one of many parameters to alter, they are vital. Their distribution impacts gradients and thus training efficacy.

Weights in Ann represent the strength of inter-layer connections. For the next layer, a nonlinear activation function transforms these inputs and the preceding layer's data. Forward propagation does this layer by layer; reverse propagation finds the optimal values of these weights to produce accurate outputs given an input.

I've tested different weights and the model's accuracy and recall values have changed and constantly dropping.This is how it looks now: I added the weights of 0.25-0.45-0.25-0.45 and added the labels like Random Forest, K-nearest neighbor, Ada boost, light gm, and zipped it up. used the elf to split the data into test and training models, and

then the ensemble vote classifier's accuracy was 93 percent when the data was taken into account when it was used.Figure 13 and Figure 14 shows estimators of vote classifier and separation of classes respectively.

```
VotingClassifier(estimators=[('Randomforsest', RandomForestClassifier()),
                             ('knn', KNeighborsClassifier(n_neighbors=10)),
                             ('Ada Boost',
                              AdaBoostClassifier(base_estimator=DecisionTreeClassifier()))],
                 voting='soft', weights=[2, 1, 2])
```

Figure 13: Estimators used for the Voting classifier

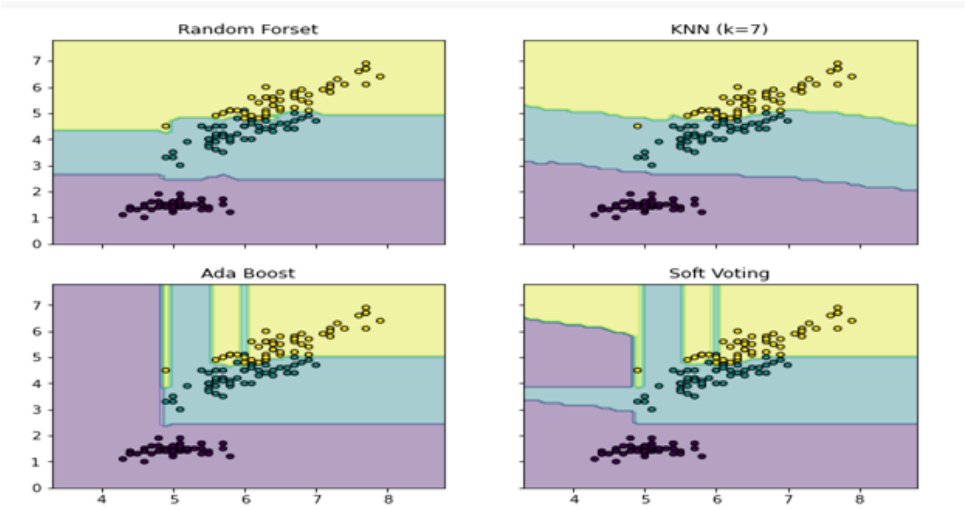

Figure 14: Separation of classes

Voting Classifier decision boundaries for two features of the Iris dataset are plotted. In a toy dataset with three different classifiers and the Voting Classifier averaging the probabilities, plot the first sample's class probability. A soft-voting Voting Classifier with weights [2, 1, 2] is first created, which implies that the predicted probabilities of the Random Forest and Ada Boost each count 2 times as much as the weights of the K Neighbors Classifier when the averaged probability is determined.

5.6 Discussion

Alzheimer's disease prediction was the subject of the proposed machine and deep learning framework for patients afflicted with the AD . Because of this, diseases were studied by analyzing data from patients. Voting classifier is the only one of the five models compared to Ada Boost, Artifical Neural Network, Bi-directional LSTM and Light Gradient Boosting Algorithm to have a significant advantage.

A metric like the F1 score is needed to evaluate, precision and recall to choose the good platform for the categorization challenge of separating three phases of Alzheimer's disease process NC, MCI and AD. By using a combination of models, this study is different from previous studies in the field.

• The Ensemble model has the potential to succeed if the goal is both accuracy and loss.When it comes to True Positive and False Positive rates(FPR)

•With the accessibility of data sets, improved performance and accuracy can be achieved. In Table 2 an example table is provided.

Table 2: A table caption.

Models	**Accuracy**
Ada Boost	53.96%
Artificial Neural Networks	60.93%
bi-directional LSTM	69.54%
Voting classifier	93.3%
Light GBM	92.5%

In Table 2 an example table is provided.

6 Conclusion and Future Work

Machine and deep learning techniques will be used to detect Dementia victims at various phases of CN, LMCI, and AD for early treatment. The results are positive, in keeping with the goal of detection of AD. Classifiers measures include R2 and RMSE, whereas classification metrics include Accuracy, Recall, Precision, and F1-Score for classification. Before modeling, the ADNI Dataset was thoroughly cleaned. In order to better predict the AD ,CN,LMCI, classification analysis was carried out using the following three different approaches: light Gradient Boosting For predicting CN,LMCI and AD the good accuracy and re-call values compared to Ada boost and ANN, the Ensemble Vote Classifier has an accuracy above 93 percent. For the classification of Alzheimer's disease, this anticipated values column can be used as a new input predictor for the new model. There was a k-fold cross-validation.Mis-classification was rather equitable, there were no total misses, and underestimation was probably too high. For a more conservative model, I'm willing to trade accuracy. False Negative: As an example, It would be far worse to predict cognitively normal (CN) when a patient has Alzheimer's (AD) rather than predicting AD when a patient is healthy (CN).

The key focus of this research is on early detection of Alzheimer's which is a form of dementia. Since no known cure has been identified, the feasible method is to early diagnosis.The models namely Light Gradient Boosting, Ada Boost, Artificial Neural Network, Bi-directional LSTM and Ensemble Vote Classifier were utilized and compared inorder to attain best fit classifer for improved diagnosis of Alzheimers. From the interpretation of results I could conclude that Ensemble Vote Classifier Machine was the best fit model which accounted highest accuracy (93.3%) having maximum sensitivity (TPR) and minimum False positive rate. From this analysis, I have been able to solve significant challenges experienced in health care industry by providing outcomes my project.

The log regrssion model's probability forecasts can be evaluated.Find out whether we can remove outliers from the data. Switch to a binary classification system (CN, LMCI equals to Not AD) so that we have two classes (Not AD, AD) and utilize AUC as a statistic.Alternatively, you may lower the log regression. threshold (say, from.5 to.4) to reduce the number of false negatives

7 Acknowledgement

Proff. Jorge Basilio, the supervisor of my research study, deserves special recognition and thanks for his patient monitoring, his tireless assistance, and his accommodating evaluations of my exploration project. The Dr. Jorge Basilio has provided me with excellent guidance and assistance in keeping my progress on track. A great deal has been said about his willingness to put in a lot of effort. For his tolerance throughout the 13-week period, I am grateful.

References

Basher, A., Kim, B. C., Lee, K. H. and Jung, H. Y. (2021). Volumetric feature-based alzheimer's disease diagnosis from smri data using a convolutional neural network and a deep neural network, *IEEE Access* **9**: 29870–29882.

Fritsch, J., Wankerl, S. and Nöth, E. (2019). Automatic diagnosis of alzheimer's disease using neural network language models, *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5841–5845.

Hong, X., Lin, R., Yang, C., Zeng, N., Cai, C., Gou, J. and Yang, J. (2019). Predicting alzheimer's disease using lstm, *IEEE Access* **7**: 80893–80901.

Hou, W., Lei, B., Zou, W., Li, X. and Zhang, C. (2017). Ensemble prediction of longitudinal scores of alzheimer's disease based on ¡inf¿2,1¡/inf¿-norm regularized correntropy with spatial-temporal constraint, *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, pp. 891–894.

Islam, J. and Zhang, Y. (2018). Early diagnosis of alzheimer's disease: A neuroimaging study with deep learning architectures, *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1962–19622.

Jabason, E., Ahmad, M. O. and S Swamy, M. N. (2018). Deep structural and clinical feature learning for semi-supervised multiclass prediction of alzheimer's disease, *2018 IEEE 61st International Midwest Symposium on Circuits and Systems (MWSCAS)*, pp. 791–794.

Kotturu, P. K. and Kumar, A. (2020). Comparative study on machine learning models for early diagnose of alzheimer's disease: Multi correlation method, *2020 5th International Conference on Communication and Electronics Systems (ICCES)*, pp. 778–783.

Li, H. and Fan, Y. (2019). Early prediction of alzheimer's disease dementia based on baseline hippocampal mri and 1-year follow-up cognitive measures using deep recurrent neural networks, *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pp. 368–371.

Martinez-Murcia, F. J., Ortiz, A., Gorriz, J.-M., Ramirez, J. and Castillo-Barnes, D. (2020). Studying the manifold structure of alzheimer's disease: A deep learning approach using convolutional autoencoders, *IEEE Journal of Biomedical and Health Informatics* **24**(1): 17–26.

Minhas, S., Khanum, A., Riaz, F., Khan, S. A. and Alvi, A. (2018). Predicting progression from mild cognitive impairment to alzheimer's disease using autoregressive modelling of longitudinal and multimodal biomarkers, *IEEE Journal of Biomedical and Health Informatics* **22**(3): 818–825.

Neelaveni, J. and Devasana, M. (2020). Alzheimer disease prediction using machine learning algorithms, *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, pp. 101–104.

Nishikawa, K., Hirakawa, R., Kawano, H., Nakashi, K. and Nakatoh, Y. (2021). Detecting system alzheimer's dementia by 1d cnn-lstm in japanese speech, *2021 IEEE International Conference on Consumer Electronics (ICCE)*, pp. 1–3.

Prince, M., Ali, G.-C., Guerchet, M., Prina, A. M., Albanese, E. and Wu, Y.-T. (2016). Recent global trends in the prevalence and incidence of dementia, and survival with dementia, *Alzheimer's Research & Therapy* **8**(1).
URL: *https://doi.org/10.1186/s13195-016-0188-8*

Uddin, M. S. and Ashraf, G. M. (2019). Introductory chapter: Alzheimer's disease—the most common cause of dementia, *Advances in Dementia Research*, IntechOpen.
URL: *https://doi.org/10.5772/intechopen.82196*

Venugopalan, J., Tong, L., Hassanzadeh, H. R. and Wang, M. D. (2021). Multimodal deep learning models for early detection of alzheimer's disease stage, *Scientific Reports* **11**(1).
URL: *https://doi.org/10.1038/s41598-020-74399-w*