National College of Ireland

Alzheimer Disease Detection and Prognosis from Clinical Data using Machine Learning Techniques

Research Project
MSc in Data Analytics (MSCDA - B)

Mubeen Ali Mohammed
Student ID: x18180370

School of Computing
National College of Ireland

Supervisor: Dr. Muhammad Iqbal

National College of Ireland
Project Submission Sheet
School of Computing

National College of Ireland

Student Name:	Mubeen Ali Mohammed
Student ID:	x18180370
Programme:	MSc in Data Analytics (MSCDA - B)
Year:	2019-2020
Module:	Research Project
Supervisor:	Dr. Muhammad Iqbal
Submission Due Date:	28/09/2020
Project Title:	Alzheimer Disease Detection and Prognosis from Clinical Data using Machine Learning Techniques
Word Count:	7056
Page Count:	23

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:	Mubeen Ali Mohammed
Date:	27th September 2020

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 Disease Detection and Prognosis from Clinical Data using Machine Learning Techniques

Mubeen Ali Mohammed
x18180370

Abstract

Alzheimer's disease is the fifth most leading cause of death as per WHO impacting millions globally, 131 million cases projected in 2050. No cure has been found for Alzheimer's disease implying early diagnosing of Alzheimer's stages Normal Cognition (NC), Mild Cognitive Impairment (MCI), and Alzheimer disease (AD) in patients is cost-effective, reduce suffering among the community. In this paper, Machine learning algorithms like ElasticNET, Gradient Boosting, Deep Neural Net, Support Vector Machines, and LSTM networks are used for the prediction of continuous cognitive variables MMSE, ADAS13, Ventricle and categorical classification of 3 AD stages. These predictions are performed on ADNI Clinical data. The evaluation metrics proposed are R^2, RMSE, Accuracy, Precision, Recall, F1-score. Gradient boosted regressor is a robust model compared to other algorithms achieving R^2 of 90% and lowest RMSE scores. Dynamic LSTM obtained an accuracy of 78% outperforming other classifiers showing promising results by detecting Alzheimer patients empowering medical supervisors to initiate appropriate treatment. The model is better at predicting Alzheimer's stages however model accuracy can be enhanced by using a multi-modal, ensemble approach along with other state-of-the-art methods.

Keywords : **Alzheimer disease, Elastic Net, Gradient Boosting, Neural Network, LSTM Network, Support Vector Machines, BioMedical data**

1 Introduction

Alzheimer's has been one of the significantly impacting diseases in the world. It is fifth among the diseases which are also a terminal as per World Health Organisation [1]. Dementia depicts the indications of memory loss and intellectual capacity that affect an individual's everyday life and their families. It is one of the major difficulties in the Healthcare field which has to be looked into in the 21st century. Dementia most commonly occurs in the form of Alzheimer's. Recent studies point out that 8 percent of people over 65 and around 21 percent of people over 80 suffer from dementia (Brookmeyer et al.; 2007a).

Dementia has relatively high costs of social care along with health care as compared to stroke, cancer, and severe heart associated with fatal problems. Projected to costs

[1] `https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death`

around 2 trillion in 2030 [2] A medication which reduces the advancement by 50 percent will lower the cost of yearly care by almost 10 percent i.e 100 billion in 2018

Machine Learning is crucial for efficient recognition of dementia-related infections such as Alzheimer's, which by diagnosing the symptoms more quickly saves patients and specialists part of the time and cash. This will significantly help specialists avoid potential risks before the signs of Alzheimer's disease become more dreadful. It might trigger valid treatment, evading injury, and sparing many individual existences. This recovers billions of euros and time for analytical focus and ultimately government spending that can positively influence the GDP of the nation.

1.1 Motivation and Background

The statistics in Figure 1 indicate the impact of widespread Alzheimer disease.It shows that medical costs are projected to spiral upwards almost going up from 206 billion dollars in 2020 to 777 billion dollars in the U.S and the same increasing trend projections continue until 2050. Also, the number of people who are affected by this dementia disease numbers is expected to increase from 54 million cases by almost three times to 131 million in 2050 worldwide.

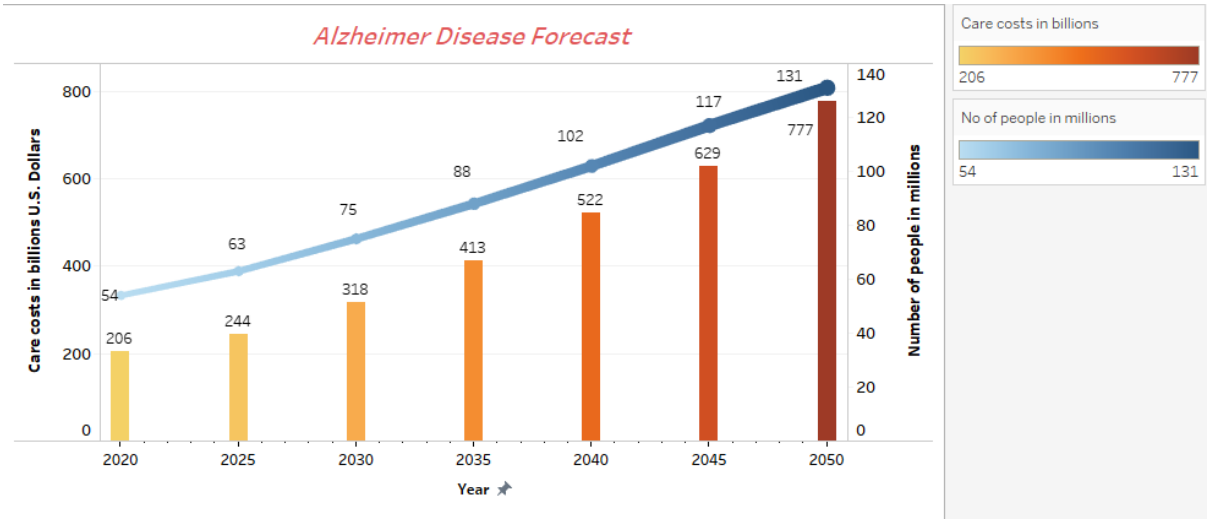

Figure 1: Alzheimer disease worldwide cases and U.S medical costs forecast from 2020-2050, (Statista; 2020)

This research makes use of The Alzheimer's Disease Prediction Of Longitudinal Evolution (TADPOLE) competition data organized in 2017. The need to find an innovative way of early diagnosing this disease has led to this challenge which bears the potential to solve a key widespread disease like Alzheimer's.The framework and rules of the competition have been elaborated in (Marinescu et al.; 2019)

The objective [3] is to recognize which individuals inside an age category in danger of AD (Alzheimer's Disease) will begin to show manifestations over the short to medium term (1-5 years). It centers around "rollover people" (patients originating from a past stage) in the Alzheimer's Disease Neuroimaging Initiative (ADNI) study [4]. In this study, the objective is to accurately predict using clinically significant factors dependent on historic

[2]Statistics `https://www.dementiastatistics.org/statistics/global-prevalence/`
[3]`https://tadpole.grand-challenge.org/`
[4]`http://adni.loni.usc.edu/`

patient visit information. Three major factors to be forecasted are the Mini-Mental State Examination (MMSE) test score, Alzheimer's Disease Assessment Scale Cognitive Subdomain (ADAS13) score, and head-size standardized ventricle volume with the usage of multiple algorithms like Support Vector Machines, Gradient Boosting, Neural Net, Elastic Net. After foreseeing these three continuous factors, the last forecast is classifying the subject into three classes namely Normal Cognition, Mild Cognitive Impairment (MCI), and AD (Alzheimer diagnosed) with the Dynamic LSTM framework.

1.2 Research Question

The main aims proposed in the paper is about Alzheimer's Disease Prediction of Longitudinal Evolution (TADPOLE) challenge is to utilize previous clinical test measurements to predict future clinical test data and diagnosis which are mentioned in detail below

1) Prediction of clinically important continuous variables namely the Mini-Mental State Examination (MMSE), Alzheimer Disease Assessment scale (ADAS13) and Head sized normalised ventricle volume (Ventricles_Norm)- Regression task

2) Final prediction to classify the patient accurately as Healthy NC (Normal cognition), MCI (Mild cognitive impairment) and AD (Alzheimer Disease) by including new input predicted variables MMSE, ADAS13 and Ventricles_Norm from above - Classification task

The further sections in the paper are organized as follows: Section 2 describes the Related Literature review, Section 3 illustrates the CRISP-DM Methodology followed in this paper, Section 4 explains the Design Specification, Section 5 illustrate the Implementation aspects of research, Section 6 examines the evaluation results of the applied Machine learning models, Section 7 describes the Conclusion and Future Work.

2 Related Work

Many individuals experience the ill effects of this Alzheimer's disease (AD) every year. As per an examination, one in every 85 people will experience the painful effects of AD until 2050 (Brookmeyer et al.; 2007b). Mild Cognitive impairment (MCI) diagnosed people are generally rough, however, patients in the last phase of AD endure cardiovascular breakdown and dysfunctionality of the respiratory system leading to death (Beheshti et al.; 2016). Researchers have made a few Computer-Aided Diagnostic Systems (CADS) for the exact location and grouping of AD-related highlights (Hosseini-Asl et al.; 2016). The biomarkers of Alzheimer's illness are neurochemical markers required to analyze the the grave intensity of the condition (Hampel et al.; 2018).

2.1 Machine learning prediction on longitudinal progression of Alzheimer disease

Many classifiers and regressive methods have been applied towards the detection of AD. The paper (Lu; 2019) uses a statistical regression technique and imputes the missing

values on ADNI data using the mean biomarker value. They perform manual feature selection taking known variables from literature which are cognitive variables (ADAS13, RAVLT, MMSE), amyloid-beta, volumetric MRI, etc. Prediction of MCI to the AD stage has been done using the Aalen additive regression model which outputs the likelihood of a patient staying in MCI with time. The drawback was the paper was not able to predict ADAS-Cog13 and Ventricles. On the same data, (Iddi et al.; 2019) used various mixed effect techniques like Joint Mixed Effect modelling, Latent Time Joint Mixed-effect Modelling, and then performing a model average previously mentioned models forecasting ADAS 13, ventricles. For diagnosing Alzheimer's stages Random Forest classifier using forecasted clinical scores achieving an 80% accuracy over 2.5 year period.

In (Oxtoby et al.; 2018) the author made use of features like MRI biomarkers, cognitive scores criteria, etc from ADNI data, etc and each feature used a univariate and multivariate approach of estimating disease stages as part of data-driven illness progression model. For continuous variables in clinical data, an exponential distribution model was used and AD classification of future forecasts was performed using the Linear scale approach. The model was over-fit and a poor R^2 value of 0.49 was achieved. In this study (Mehdipour Ghazi et al.; 2019) the authors have used linear discriminant analysis (LDA) to choose the main features of biomarkers and other MRI measures.ADAS 13 and ventricles were forecasted using a data-driven progression based on disease progression scores (DPS). The clinical diagnosis was anticipated dependent on the DPS scores utilizing both a Bayesian classifier with probabilities displaying utilizing Gaussian blend models. The achieved a very good normalized mean absolute error (NMAE) value 0.991 and AUC value of 0.934.

A new novel method named Data-Driven Inference of Vertexwise Evolution (DIVE) algorithm was introduced by (Marinescu et al.; 2019) to predict ADAS-Cog13, Ventricle volume. Missing data has been automatically been taken care of using this method. Biomarkers that are used as input for DIVE are grouped based on their progression pattern spread over illness time-course. (Marinescu et al.; 2020)For clinical classification task, the posterior likelihood for every class is calculated based on forecasted DPS scores utilizing non-parametric Kernel Density Estimators (KDE), fitted on the DPS scores for every distinguishing class individually.

This research (Raket; 2019) uses a latent time mixed-effects model on forecasting of continuous value ADAS Cog13 utilizing exponential type parameterization to find biomarker trajectory. The above model has latent variables such as the ailment stage, baseline cognition, and time-bound cognitive decline. The Clinical diagnostic classes are distinguished using kernel density approximation of every state throughout the disease time from the previous forecasted model. (Aksman et al.; 2019) Forecasting of ADAS-Cog13, Ventricles variables were done using a multi-task linear regression model on age, some variables used as covariates MMSE, Gender, Education, etc. Kernel coupled models permit the patients' trajectory linkage using multiple biomarker measures to classify AD, MCI, and NC. This model performs better than the usage of a mixed effect model when modelling biomarker trajectories.

2.1.1 Supervised Learning Models

Random Forest has been one of the favourite Machine learning algorithm used in predicting Alzheimer's disease, This study (Moore et al.; 2019) has manually selected cognitive measures, demographics measures, MRI based features for modelling. The objective was

to understand the correlation among data point sequences at various time partitions. It outshines the traditional SVM and linear mixed-effects model. For clinical AD Diagnosis an AUC score of 0.82 and accuracy of 0.73 which is decent. The limitation of this paper is it uses a minimal number of predictors from the ADNI clinical dataset corpus. The focus of this paper (Lei et al.; 2020) has been to consider clinical measures at different time stamps including longitudinal MRI neuroimaging data. The framework proposed is in three parts. Initially feature selection is done using cross-entropy joint learning, encoding the features utilising a polynomial network and score predictions using Support vector regression. Longitudinal clinical scores prediction is performed using baseline data and these scores are recalculated incorporating different time stamps of the patient's cognitive trajectory increasing accuracy. Limitations are that they considered only longitudinal data also demographic measures like age, gender, etc. had been excluded.

(Venkatraghavan et al.; 2019) Event-based modeling approach has been proposed in this paper by considering only those patients who were later diagnosed as AD to ascertain the illness severity of patients. Biomarker forecasts of continuous measures have been performed using linear effect algorithms for features including MMSE, ADAS 13. For AD classification, the same forecasted clinical values have been used along with patient last known disease state is fed into Support Vector machines (SVM) with Radial function. The results are not so great but further data pre-processing improvements.

2.1.2 Regularised Regression and Classification Methods

Using 53 patients' Alzheimer data and 180 health control data from "German resting-state initiative for diagnostic biomarkers", ElasticNet regression used by (Teipel et al.; 2017) retrieved consistent network connectivity leading to better increased diagnostic accuracy and variable selection. The regularised regression method performed superior in comparison to step-wise logistic regression. The accuracy achieved was around 0.80 with the usage of bootstrap cross-validation. Another similar study on (Schouten et al.; 2016) used "Austrian Stroke Prevention Family" data made use of Elasticnet classifier with a combination of multi modalities like anatomical Magnetic resonance imaging (MRI), diffusion dMRI, rsMRI measures. In a stepwise way, if measures are amalgamated from various modalities it leads to enhanced AD stages classification AUC score of 0.952. This proves that AD classification using a multimodal approach outweighs unimodal. Paper (Shen et al.; 2011) emphasizes combined research of structural MRI data, proteomic data to validate the predictions and selection of biomarkers in large MRI/AD group. Elastic net logistic regression has enhanced predictive power while also optimizing feature selection. Predictions HC/AD gave classification accuracy of 0.919, HC/MCI as 0.905, MCI/AD as 0.865. Hence it proved very beneficial to use elastic nets for discovering potential biomarkers including AD diagnostic accuracy.

2.1.3 Ensemble Meta Algorithms

Izquierdo et al. (2017) demonstrated the use of stochastic gradient boosting on 1142 patients' clinical and neuroimaging data available from ADNI. MMSE, CDRS, ADAS Cog-13 cognitive prediction scores modelled using Gradient boosting outperformed all other state-of-the-art methods like Multi layer perceptron, KNN, SVM, Ridge regression, etc in terms of better predicted and actual cognitive measures around 0.9. The paper (Braithwaite et al.; 2020) investigates the viability of gauging feature value with Shapley numbers along with ensemble methods. Their approach was to perform classification

using Gradient Boosting(an amalgamation of weak classifiers) on a controlled study of 70,719 Alzheimer patients in Finland. The outcomes from the added variable recognized new cases for the future examinations on AD hazard factors.

2.2 Deep learning research on Alzheimer disease prediction

This section of Deep learning has seen many applications and has gained attention across all research including the medical and healthcare field.ANN, RNN, and CNN have been some of the widely used deep learning algorithms. A paper (Peña-Bautista et al.; 2019) uses Artificial neural network which promises increased versatility in the clinical diagnosis of AD stages using Lipid peroxidation known to be a key factor in determining Alzheimer's. Plasma and Urine sample data modelled using ANN has proved to show better accuracies and signs of quicker AD prediction. This approach is cost-effective and can significantly increase the detection rate of AD using human urine and plasma sample. The AUC score was 0.882 in plasma and 0.839 in Urine outperforming SVM and PLS models.

2.2.1 Recurrent Neural Nets and Long short term memory Networks

(Pearlmutter; 1989) proposed that RNN are arrangement based learning techniques by configuration which provide ongoing, combined modelling of longitudinal data considering dependencies among readings. This method does not require a patient trajectory arrangement. LSTM systems are the most broadly utilized kind of RNNs created to successfully record long-duration transient dependencies with the erupting and disappearing gradient issue. (Hochreiter and Schmidhuber; 1997) (Gers et al.; 1999) . A stochastic memory cell with refresh components that understands to store past info over a prolonged time duration.

This paper (Greff et al.; 2016) showcased various applications of LSTM, especially in time series problems. In general, the LSTM framework consists of three reset ports including the full gate. As mentioned by (Petersen et al.; 2010) longitudinal data cohorts generally contain missing patient biomarker measures observed in ADNI data. Hence the standard RNNs won't be applicable. The author (Lipton et al.; 2016) suggests methods like imputing dataset, linear interpolation to take care of empty row values. Since these are multi-stage process it gives inadequate results that are inspired by the data cleaning method used.

A shortage of capable methods to inherent handle missing values other than RNN is an issue. The author (Che et al.; 2018) customized the RNN architecture by adding extra parameters to treat the missing data and this generally is biased to a selected portion of group or age group etc thereby introducing complexity.(Petersen et al.; 2010) uses LSTM to Neurogenerative disease progression modelling like Alzheimer's using various MRI biomarkers to perform AD diagnostic classification of Alzheimer patients.

MRI scans are the best as per (Biagioni and Galvin; 2011) when it comes to diagnosing AD in subjects through a lot of changes in brain structure. Volumetric analysis is the best way to evaluate brain decay. (Oxtoby and Alexander; 2017) specifies Alzheimer detection is possible through sturdy, complex disease progression modelling to be done using deep learning neural networks providing a better perspective on a diagnostic longitudinal dataset.

Longitudinal disease trajectory modelling is challenging as other methods don't consider biomarker trajectories, measure interdependence. RNN framework is proposed how-

ever due to the usual missing value nature of longitudinal study data points LSTM is introduced in this paper (Ghazi et al.; 2019) to handle these issues. This LSTM is utilised to predict six MRI biomarkers and compared them with other methods. It has an AUC score of 0.90 for the clinical diagnosis of AD. The method used is novel as RNN in the form of LSTM is applied to an Alzheimer's Disease progression Modelling.

In this paper, how different types of machine learning models have been applied to various Alzheimer datasets is ascertained in the related work section. The dataset used in this project is from The Alzheimer's Disease Prediction Of Longitudinal Evolution (TADPOLE) Challenge data present in the ADNI database. The main objective is to make use of these state-of-the-art machine learning methods like Regularised ElasticNet model, Gradient Boosting (GB) Ensemble model, Support Vector Machines (SVM), Neural network, LSTM to perform prediction of important biomarkers like ADAS cog-13, MMSE and Ventricles Normalised cognitive scores while also exhibiting classification performance of these techniques on three Alzheimer stages Cognitive Normal (CN), Mild Cognitive Impairment (MCI) and AD is proposed.

3 Methodology

In this research, the objective is to apply various algorithms on past clinical test measurements to forecast future clinical test data and diagnosis. Based on tasks like predicting continuous or categorical data usage of multiple machine learning models both as regressors and Classification like Support Vector machines, Elastic Net, Gradient Boosting, Neural Network with 3 hidden layers and Dynamic Long Short Term Memory has been discussed (LSTM). The methodology followed in this study is described below

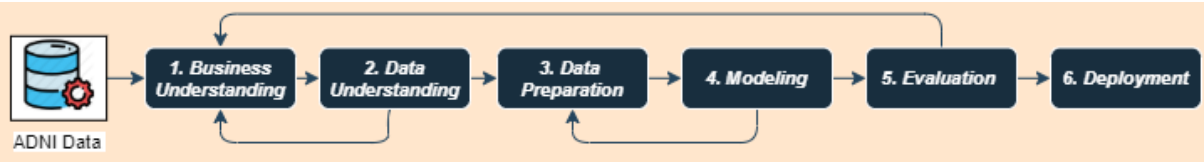

Figure 2: CRISP-DM Methodology

3.1 Business Understanding

In the Healthcare industry, recent research has been on applying Artificial Intelligence to solve the key challenges in the form of many diseases leading to the death of many people around the world. For instance, the usage of MRI, CT scans, and patient's clinical data as part of detecting the disease in advance with the use of the latest state-of-the-art algorithms has gained traction in the world research community. In this paper, the focus is on dementia based disease known as Alzheimer's. There is no absolute cure found yet for Alzheimer and the only way to prevent and slow down the pace is by diagnosing this disease in early stages. There are three critical stages of Alzheimer's disease namely Normal Cognition (NC), Mild Cognitive Impairment (MCI), and Alzheimer's Disease (AD). Five machine learning methods have been implemented on the TADPOLE ADNI data namely ElasticNet, Support Vector Machines, Gradient Boosting, Neural Network, and Dynamic LSTM (Long short term memory) to predict and diagnose the Alzheimer disease.

3.2 Data Acquisition

The data set has been collected from freely accessible large database on Alzheimers disease at [5]. One has to apply for access to the database to gain login credentials as the database is meant for research purposes. This database consists of all North American patients examination data recorded from ADNI 1, ADNI 2, ADNI GO and ADNI 3 at different points, monthly time intervals from July 2005 to May 2017. It is a repository containing various patient study data in the form of MRI, PET, Genetic, Clinical, cognitive and blood biomarkers to determine the progression of the Alzheimer disease. The TADPOLE Challenge dataset can be downloaded from the Alzheimer disease Neuroimaging Initiative (ADNI) containing pre-processed, curated data collected following set of protocols to maintain consistency in data for further research (Weiner et al.; 2012)

Alzheimer Stages	**No. of Participants**	**No. of Records**
CN	418	2764
EMCI	310	1140
LMCI	562	3547
SMC	106	106
AD	342	1161
Total	1737	8715

Table 1: Alzheimer records by stage

The number of attributes in the TADPOLE dataset is 1907 variables for 1737 subject participants. The number of records is 8715

3.3 Data Preparation and Pre-processing

On investigating the clinical data it is observed that the input data is not uniformly filled and requires cleaning i.e the data is not present for all the patient visits. The total dataset is split into Training and Validation datasets. Training data is being used to prepare the model and assessing model performance using the validation data (Shahbaz et al.; 2019). Table 2 indicates some important attributes which describe the basic demographic information of an individual such as Age, Gender, etc.

Label	**Description**	**Data Type**	**Units**
RID	Participant's Roaster ID.	Numeric	NA
AGE	Participant's age.	Numeric	Years
PTGENDER	Participant's gender	Nominal	NA
PTEDUCAT	Participant's education	Numeric	Years
PTETHCAT	Participant's ethnicity	Nominal	NA
PTRACCAT	Participant's race	Nominal	NA
PTMARRY	Participant's marital status.	Nominal	NA
VISCODE	Participant's Visit code.	Nominal	NA

[5]Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at : http://adni.loni.usc.edu/wpcontent/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

Years_bl	Participant's year of examination.	Numeric	Years
SITE	A code indicating the site of a participant's examination	Numeric	NA

Table 2: List of demographic attributes in the data

Table 3 mentions the list of important cognitive ability based tests to determine the Alzheimer's stage in each patient.

Label	**Description**	**Data Type**	**Units**
CDRSB_bl	Clinical Dementia Rating Sum of Boxes (score)	Numeric	NA
ADAS11_bl	11 item-AD Cognitive Scale (score)	Numeric	NA
ADAS13_bl	13 item-AD Cognitive Scale (score)	Numeric	NA
MMSE_bl	Mini-Mental State Examination (score)	Numeric	NA
RAVLT_bl	Rey's Auditory Verbal Learning Test (scores for immediate response, learning, forgetting and percentage forgetting)	Numeric	NA
FAQ_bl	Functional Activities Questionnaire	Numeric	NA

Table 3: List of cognitive assessment attributes in data

Table 4 mentions the list of variables from the clinical examination of each subject after important AD biomarkers have been given.

Label	**Description**	**Data Type**	**Units**
APOE4	APOE4 gene presence	Binary	NA
Hippocampus_bl	Volume of hippocampus	Numeric	mm3
Ventricles_bl	Volume of ventricles	Numeric	mm3
WholeBrain_bl	volume of Brain	Numeric	mm3
Fusiform_bl	The volume of the fusiform gyrus.	Numeric	mm3
Entorhinal_bl	The volume of the entorhinal cortex.	Numeric	mm3
MidTemp_bl	TThe volume of the middle temporal gyrus.	Numeric	mm3
ICV	Intra Cranial Volume	Numeric	mm3

Table 4: List of clinical assessment attributes in the data

3.3.1 Feature Engineering

- Data Cleaning: It is observed many variables in the data containing a lot of missing data which cannot be utilized for imputation as it will worsen the data quality. Columns containing blank or NaN calculation of proportions has been compared to a threshold value to 60%. Hence, this leads to the removal of these columns having more than threshold missing values from the data using Python code. Also removed the same-value column for eg. TEMPQC_UCSFFSL_ had PASS present 99% of times and across different diagnosed categories.
  The imputation of missing column values is performed using linear interpolation for numerical variables using pandas library interpolate function. Ex: if column X has 5-row values, 3 rows contain values the remaining 2 blank rows can be filled

interpolating the 3 values present.

- One Hot Encoding: It is being used on categorical variables like PTGENDER, DX_bl depicting 5 Alzheimer stage values AD, CN, EMCI, LMCI, and SMC, etc. are encoded into binary numbers 0 and 1. This is very useful for EDA and Modelling.

- Feature Scaling and Standardisation: Distance-based algorithms like SVM are sensitive to the scale of features. Normalization is reshaping the column between 0 and 1. Standardization refers to making each feature centered around 0 and values are Bell-shaped type distributed separated by unit variance. A lot of variables except PTID_Key, EXAM DATE, and DX are standardized for PCA and later for modelling.

3.4 Exploratory Data Analysis (EDA)

As seen in Figure 3, most number of the subjects are in the range of 68-84 age bracket. The highest number of records 600 belongs to people aged 72. It almost follows a normal distribution.

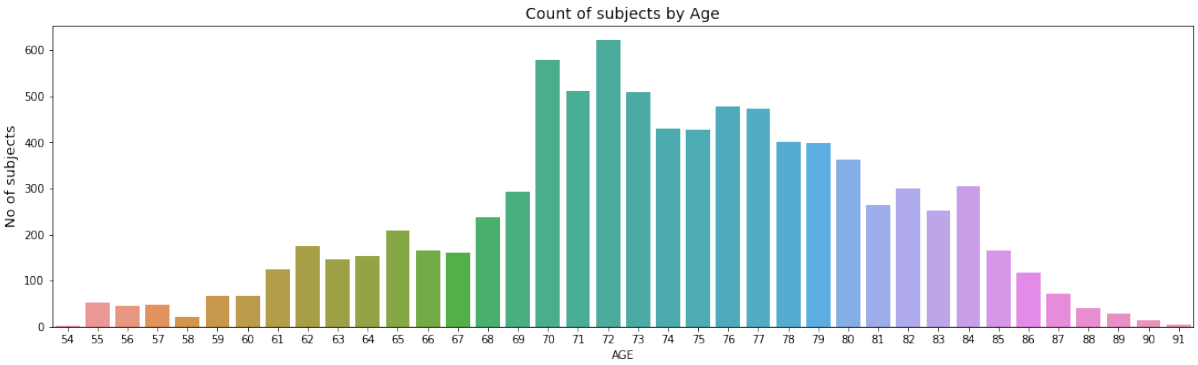

Figure 3: Patient count by Age

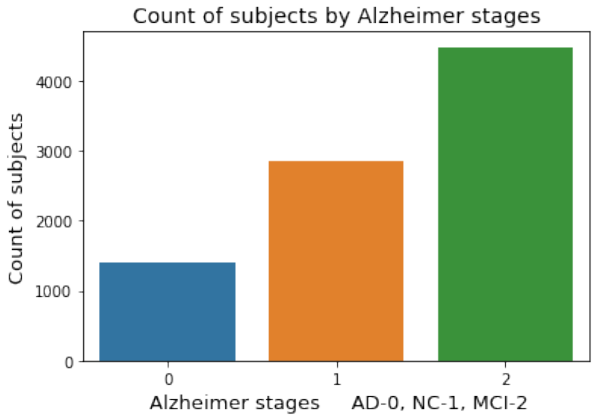

Figure 4: Number of patients by Alzheimer stage

This bar chart Figure 4 depicts the count of patients by various stages of Alzheimer's variable DX. We merge 5 variable DX cognitive levels into 3 stages as follows: 1) Alzheimer's Disease records under AD 2) Early Mild Cognitive Impairment (EMCI),

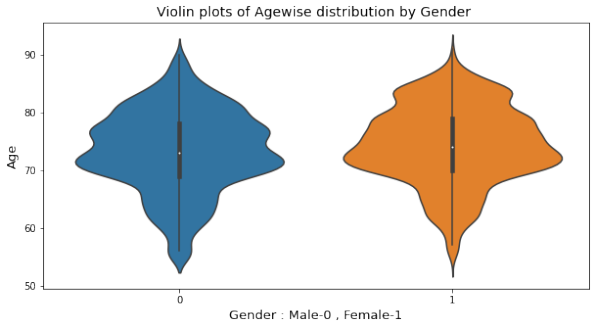

Figure 5: Age-wise distribution by Gender

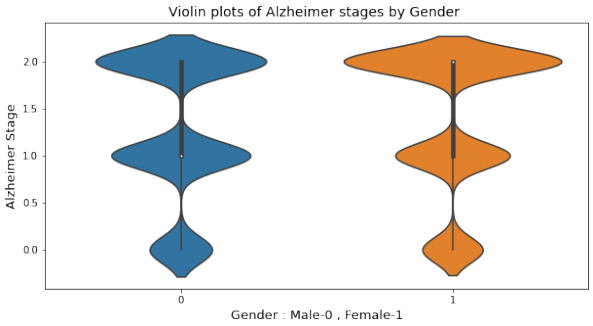

Figure 6: Alzheimer stages by Gender

Late MCI (LMCI) Under MCI 3) Cognitive Normal (CN) and Significant Memory concern (SMC) under CN. We see AD, CN, MCI cases are 4687, 2870, 1161 respectively

Violin plots are created using the Seaborn python package which is a hybrid of box-whisker and Kernel density plot. In figure 5, the relationship between Gender and Age is plotted. The white dot in the thick central gray bar is the median value, for Male & Female it is 72, 74 respectively. 50% of Male age population in data lies between 68-78, for Female it lies between 70-80. The wider bulge in the kernel distribution implies a high probability of data lies in the age group 70-84. Skinnier kernel area implies less probability of data lies in age ranges 55-68 and 85-90. In figure 6, Gender distribution across various Alzheimer stages are plotted. Alzheimer Scale encoded category value: 0.0-AD Stage, 1.0-CN stage, 2.0 MCI stage 50% of the patients' data for both Male and Female falls under MCI, and CN Alzheimer stage. For both genders, the thinner bulge at Y-axis value 0.0 in kernel density shows less probability of data falling in Alzheimer disease (AD) stage, Middle bulge at 1.0 implies a decent probability of data lying in cognitive Normal (CN) stage, Broader bulging towards the top section of the graph makes the maximum probability of data points(patients) lie are under Mild Cognitive impairment (MCI) stage.

3.4.1 Correlation Map

Figure 7 showcases relationship between important variables used in our data. Variables like Age, Married status were having a decent positive correlation coefficient of 0.25. Education and Gender were seen to be correlated with an r-value of 0.19. Variables like DX (Alzheimer's stage) and MMSE (Mini-Mental State Examination scores) were seen positively correlated with a value of 0.22. A strong negative correlation value of -0.85 was found between variables like ADAS13 and MMSE which are both different clinical assessment measures showcasing they negatively influence each other.

3.5 Principal Component Analysis (PCA)

PCA is a Dimensionality reduction technique used to reduce the number of columns after cleaning data. We have hundreds of features in the input dataset. There is a need to reduce them before feeding the data into various Machine learning models. A lot of variables except PTID_Key, EXAM DATE, and DX are standardised for PCA. This step reduced the number of columns from 710 to 281 without much information loss.

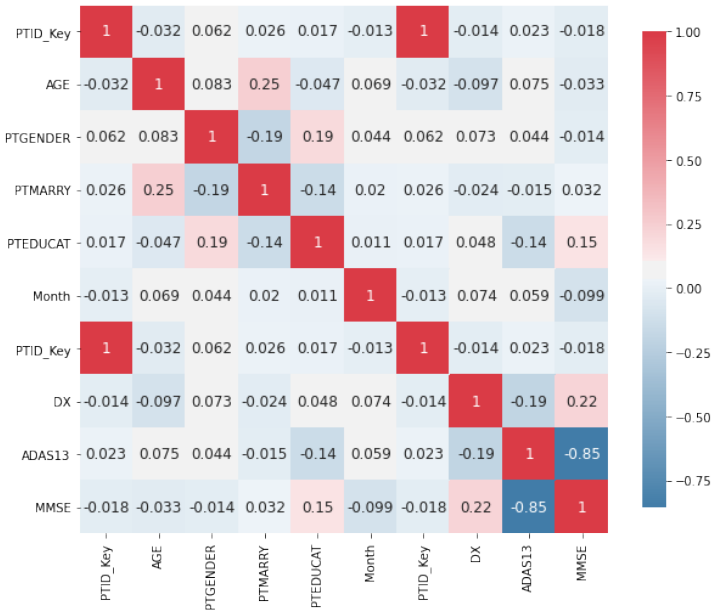

Figure 7: Correlation Matrix

4 Design Specification

The design framework of our research consists of four main stages as mentioned below which are Data Collection, Data Pre-Processing, Building Regression, and Classification Models, Evaluation of Model results and visualization as shown in Figure 8

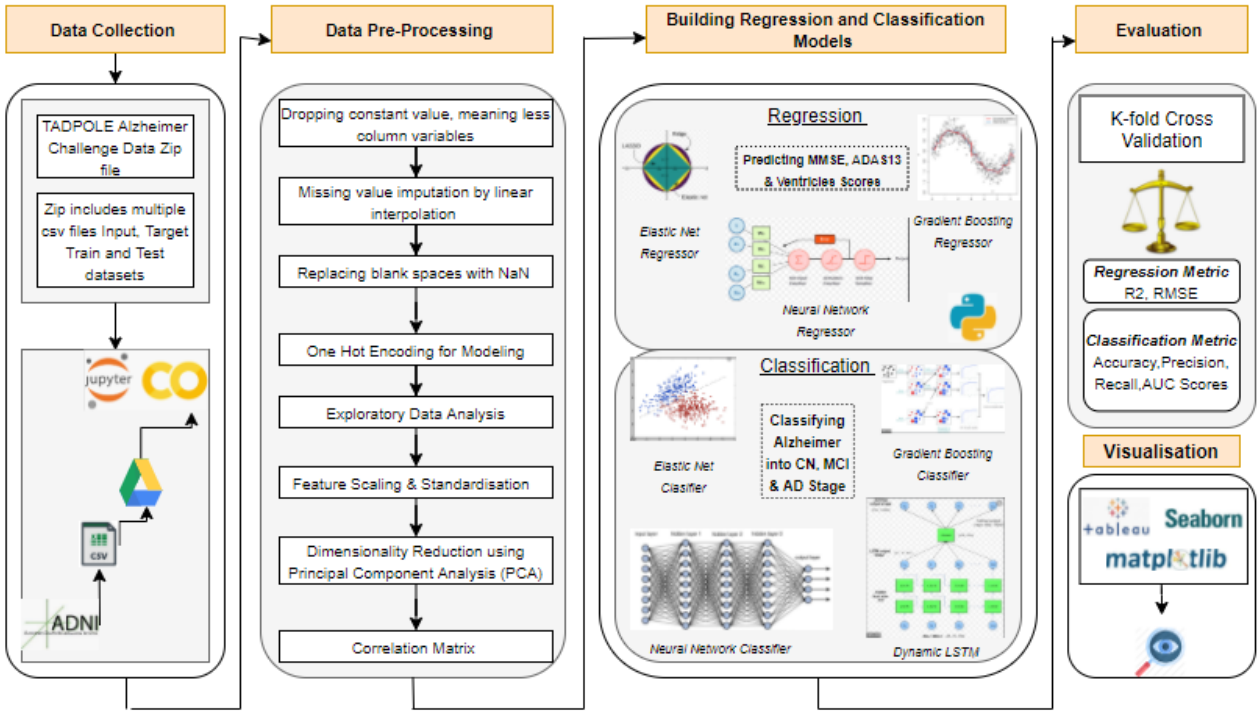

Figure 8: Alzheimer Disease Detection - Design Flow

- Data Collection stage consists of downloading the TADPOLE ADNI data from the database which is a zip file consists of various CSV(comma separated files) containing Input CSV file, Target Train CSV file and Test CSV file which is used

later at different stages of modeling. This downloaded data is stored in Google drive which can be later mounted to be fetched and read in the Python code on Google Colab Pro.

- Data Pre-Processing is vital as it contains many steps of Data Cleaning, Data wrangling, and Data preparation like dropping variables, Missing value imputation using linear interpolation, performing Exploratory Data Analysis. Later, Normalisation and standardization are conducted to perform dimensionality reduction with the help of Principle Component Analysis (PCA). At last, the variables are checked for any correlations based on the correlation map

- Third stage is Modeling various algorithms like Baseline Model, Elastic Net Regressor, Gradient Boosting Regressor, Support Vector Regressor, and Neural Network regressor 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). The classification algorithms like Elastic Net, Gradient Boosting classifier, Neural Network classifier, and Dynamic long short term memory (LSTM) have been utilized to categorize accurately different stages of Alzheimer's disease.

- Evaluation stage is a vital stage to evaluate the results of the models applied with the help of K-fold cross-validation. Various Regression and Classification metrics have been utilized. All the graphs and plots have been visualized to get the insights using Tableau, Seaborn, and Matplotlib in Colab Notebook.

5 Implementation

5.1 Data Cleaning for the models

The ADNI data consists of approximately 1500 biomarkers obtained through 1737 patients during 12741 timely trips to ADNI. Input data contains 8715 records and 1737 columns. After cleaning the data for columns with 99% constant values 166 columns from 1737 columns are dropped. Filtering for missing values and NaN values not crossing the threshold of 60% are 1010 columns dropped. The total number of rows is 8715 and columns are 716. We make use of sklearn python library and import preprocessing function is known as StandardScaler to be used for PCA. After applying PCA with sklearn library importing PCA, the number of 710 columns is reduced to 281 columns with the same 8715 rows with principle components.

5.2 Modelling

Multiple machine learning models like Elastic Net, Support Vector Machines, Gradient Boosting, Neural Network and Dynamic LSTM to predict and classify Alzheimer disease and justifying the research questions are used in this study which are mentioned below

5.2.1 Elastic Net (EN)

EN is a linear regularised regression technique which is a mixture of lasso regression L1 norm and Ridge regression L2 Norm penalizing coefficients.ElasticnetCV is useful

when numerous correlated features are (Shen et al.; 2011) .It is used for the prediction of continuous variables like MMSE, Ventricles, ADAS13 scores. The hyperparameters chosen for ElasticnetCV python function have alpha values: used to tune the model by penalizing this constant multiplied with L1 or L2 term. 7 alpha values are chosen along with mixing parameter l1 ratio (penalty for both L1L2): 0.01, 0.1, 0.5, 0.9, 0.99.All other parameters have default values.

```
ElasticNetCV(alphas=[0.0001, 0.0005, 0.001, 0.01, 0.1, 1, 10], copy_X=True,
             cv=None, eps=0.001, fit_intercept=True,
             l1_ratio=[0.01, 0.1, 0.5, 0.9, 0.99], max_iter=5000, n_alphas=100,
             n_jobs=None, normalize=False, positive=False, precompute='auto',
             random_state=None, selection='cyclic', tol=0.0001, verbose=0)
```

Figure 9: Elastic Net Regressor - selected hyperparameters

5.2.2 Support Vector Machines

SVM is a supervised technique used for both classification regression (Lei et al.; 2020).It tries to find a hyperplane in data separation in a higher dimensional place to see which side the data points fall under. Support vector regressor is a linear kernel method having a decision boundary that is distanced from hyperplane and its needed to make sure data points are lying within this boundary. Parameters selected kernel coefficient gamma=scale, penalty value c=1.0, epsilon=0.2. A support vector classifier is used to distinguish 3 AD stages using SVC default function parameters with a 5,10 fold cross-validation.

```
SVR(C=1.0, cache_size=200, coef0=0.0, degree=3, epsilon=0.2, gamma='scale',
    kernel='rbf', max_iter=-1, shrinking=True, tol=0.001, verbose=False)
```

Figure 10: Support Vector Regressor - selected hyperparameters

5.2.3 Gradient Boosting (GBM)

```
GradientBoostingRegressor(alpha=0.9, ccp_alpha=0.0, criterion='friedman_mse',
                          init=None, learning_rate=0.05, loss='huber',
                          max_depth=3, max_features='sqrt', max_leaf_nodes=None,
                          min_impurity_decrease=0.0, min_impurity_split=None,
                          min_samples_leaf=15, min_samples_split=10,
                          min_weight_fraction_leaf=0.0, n_estimators=3000,
                          n_iter_no_change=None, presort='deprecated',
                          random_state=None, subsample=1.0, tol=0.0001,
                          validation_fraction=0.1, verbose=0, warm_start=False)
```

Figure 11: Gradient Boosting Regressor - selected hyperparameters

GBM is an ensemble meta-algorithm utilized for both regression and classification (Izquierdo et al.; 2017). It balances both bias and variance based on a set of weak learners bundled to improve accuracy. An iterative process of allocating more weight to wrongly predicted observation and less weightage to correct observation in the next step. Gradient boosting regressor python function with estimators=3000 shows the number of trees, learning rate=0.05 shows individual tree effect on result, max_features=' sqrt' to find optimum variable count for the best split. To prevent over-fitting 3 variables used are max_depth=3 tree Depth, min_samples_leaf=15 minimum observations for each terminal node, and min_samples_split=10 is minimum observations number needed by

a node to split. loss='huber' is a robust regression loss function less impacted by outliers. Same hyperparameters are entered into Support vector classification function for diagnosing Alzheimer stages into AD, MCI, and CN except for the regression value loss='huber'

5.2.4 Neural Network and Dynamic LSTM

A neural net is a deep learning technique being used for regression, mainly for Alzheimer's classification. It is an artificial representation of human brain neurons to recognize patterns consisting of input, hidden, and output deep layers. .They understand complex features behaviour by calculating hybrid equations (neuron layer), taking multiple features (Input), and producing desired results for image, audio, or text applications (Output). LSTM consists of main elements such as Forget gate, Input gate, Input Modulation, and Output gates. This paper uses LSTM, a type of Recurrent neural network which is a sequential learning technique with backpropagation to train a time-based model on longitudinal Alzheimer progression modelling of data (Ghazi et al.; 2019). It is inherently capable of handling missing data. The Time interval has been normalized to be uniform as 6 months, the usage of nearest patient history visit data to fill in the training data, and their corresponding label of the disease classification as well as the regression results for ADAS13, Ventricles, and MMSE is seen. For the time series data formation, the time interval between two consecutive visits is 6 months. Prediction of the future 8 visits results are based on current record also these come as hyper-parameters which then would be tuned by cross-validation for classification of Alzheimer's disease. Tensorflow is used to build the LSTM framework. 3 layers are built with a feature size of 278, state size of tensors =64, batch_size=64 which is unit tensor representing batch size. Pred_times= 8 which is the future patient visits, num_class=3 based on Alzheimer's stages. Dynamic RNN is run to get all output sequences by adding dropout to avoid overfitting problems. To process the outputs the softmax activation function is used to get output probabilities for three classes.

6 Evaluation

Input and target predict data files are available from original data zip. This is cleaned and combined based on common columns PTID_key, ExamDate columns in both data files. Train and Validation preprocessed data for modeling is created using a 75:25 split ratio. Train preprocessed contains 6716 rows and 286 columns. Validation preprocessed consists of 2238 rows with 286 columns fed as model input data. Evaluation metric for regression task i.e prediction of continuous cognitive scores like ADAS13, Ventricles, and MMSE are R^2 and RMSE. For the classification of Alzheimer's disease into AD, CN, and MCI evaluation metrics used are Accuracy, Precision, Recall, F1 score.

6.1 Experiment with Gradient Boosted Regressor

		ADAS 13		Ventricles Norm		MMSE	
		R^2	RMSE	R^2	RMSE	R^2	RMSE
Baseline Regressor	Train	60%	5.97	91%	0.003	20%	2.52
	Test	38%	6.89	94%	0.003	25%	2.56
ElasticNet Regressor	Train	76%	0.12	94%	0.83	62%	-0.83
	Test	68%	7.76	83%	0.004	58%	2.82
Gradient Boosted Regressor	Train	91%	3.1	99%	0.03	85%	1.31
	Test	**90%**	**2.93**	**97%**	**0.08**	**79%**	**1.05**

Figure 12: Regression Model Performance

Gradientboosted regressor is robust in performance as higher test R^2 values of 90%,97% and 79% show model is significantly better in predicting continuous response variables ADAS13,Ventricles & MMSE respectively. 5, 10 fold cross validation scores comparison for GBR model shows minimal difference in test R^2 less than 0.1% change. Model reaches optimum performance at lesser learning rate value=0.05 compared to at 0.1. Best loss function 'huber' is selected which has a combined strength of other functions 'least square regression' and 'least absolute deviation'.Larger the estimator better the model performance hence iterations of n_estimators=500, 1000,1500,2000,2500,3000 peaked at '3000'. RMSE scores calculated can be seen in Figure 12

6.2 Experiment with ElasticNet Regressor

Used ElasticNetCV function imported from machine learning sklearn library, preprocessed train, and validation input data is used. Alpha values 0.0001,0.0005,0.001 ,0.01,0.1,1,10 is a penalising array for tuning model. Parameters l1 ratio values are : 0.01, 0.1, 0.5, 0.9, 0.99. Max_iterations were taken as trial and error starting from 1000 (default) changing by 500 and finally, the best regression fit is obtained at 5000. For ADAS13, Ventricle_Norm and MMSE, validation R^2 values obtained are 68%,83% and 58% with 5 fold cross-validation to mitigate overfitting. 10 fold cross-validation was performed and there was negligible change in the model coefficient of determination. For the Ventricles score prediction ElasticNet model can explain 83% of the variability in the data compared to MMSE, ADAS13.RMSE values are also less as seen in Figure 12

6.3 Experiment with Neural Network Regressor

Neural Net regressor is built with 3 hidden layers having neurons 128,512 and 1024 with one output variable. The learning rate=0.001 is selected as lower the rate the best neurons will learn patterns.Epoch values of 25,50,100,125,150,175,200 are tried at 200 value we see the loss function stabilising.Batch size of 32,50 is tried however at 25 model gives enhanced accuracy. Also the regularisation rate (dropout) is taken as least 0.001 for best R^2 .Adam optimiser is used along with activation function Relu(Rectified linear units). In figure 13, loss function is least at Epoch 62 with validation MSE score of 9.5 for Ventricle prediction. Figure 14 at epoch number 199, ADAS13 reaches loss function

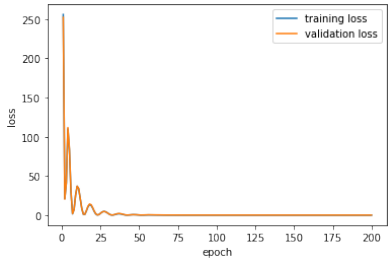

Figure 13: Ventricles loss

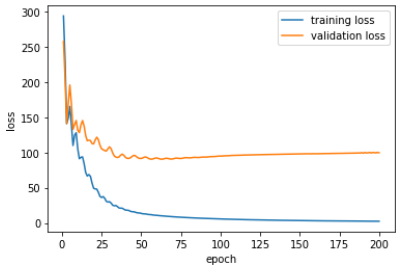

Figure 14: ADAS 13 loss

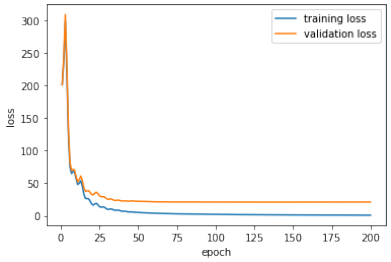

Figure 15: MMSE loss

is minimal with validation MSE=0.345. In figure 15, loss function is best at epoch=132 with Validation MSE=4.58

6.4 Experiment with Support Vector Classifier

SVC is popularly used for classification imported from sklearn library. 5,10 fold cross-validation is performed on the SVC model to observe changes in validation accuracy which is 61%, 60% respectively. Hence, on 5 fold CV individual performance metric values of Alzheimer's stages CN, MCI, and AD are mentioned in figure 16. For SVC, AD and CN accuracy are above 75% showing a good model fit however F1 scores are 66% for all classes implying model was able to identify the patients by stages CN, MCI, and AD.

		Accuracy	Precision	Recall	F1 Score
Support Vector Classifier	CN	74.71%	73%	60.72%	66.30%
	MCI	64.16%	56.15%	80.27%	66.08%
	AD	82.22%	56.51%	68.40%	61.89%
Gradient Boosting Classifier	CN	**84.67%**	84.20%	75.36%	**79.54%**
	MCI	**77.52%**	74.89%	72.21%	**73.53%**
	AD	**87.66%**	64.21%	69.80%	**66.89%**

Figure 16: Classification scores by Alzheimer stages

6.5 Experiment using Gradient Boosting Classifier

GBC is imported as an ensembled classification model from sklearn library. It is a robust modelling technique utilized and similar as used for the regression task above. We use same parameters estimators=3000, learning rate=0.05, max_depth=3 is optimum nodes in tree , max_features='sqrt', min_samples_leaf15, min_samples_split=10 tried other values 4,6,8 etc. The default loss function 'logistic regression=deviance' is being used to distinguish Alzheimer's with various probabilistic outputs. The accuracy in classifying AD, MCI, CN patients separately are 84%, 77%, 87% higher than SVC. Both high Accuracy and F1 score signify that GBC can diagnose the AD patient stages as shown in figure 16.

6.6 Experiment with Neural Network Classifier

Neural Net classifier with 3 hidden layers 128, 512, 1024 neurons, and the output layer has 3 values to classify. The least learning rate=0.001 is selected for the best performance. Epoch values of 100,200 with a various batch size of 128,64,32 were taken. The best

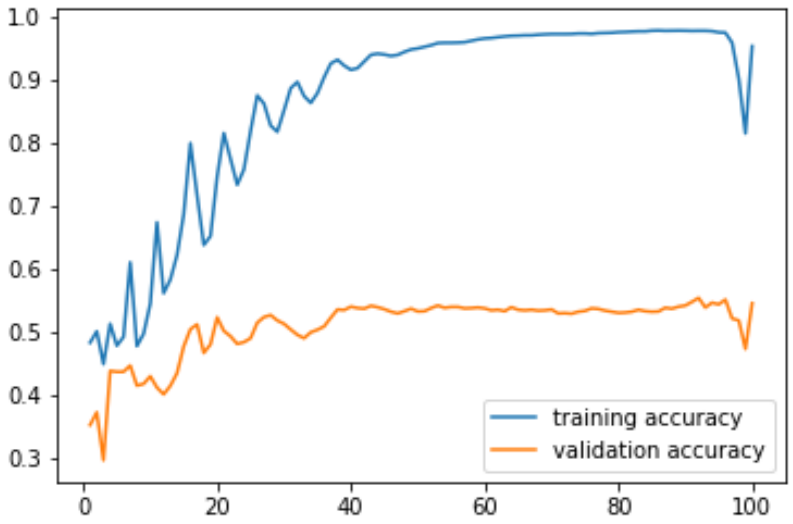

Figure 17: Neural Network Classifier Accuracy vs Epoch

model accuracy value was at Epoch=100 and batch=32. ReLU activation function with Adam optimizer is being used. The highest training accuracy obtained was almost 97% whereas the validation accuracy just around 60% were obtained at Epoch=91 as seen in Figure 17.

6.7 Experiment with Dynamic Long short Term Memory (LSTM)

Tensorflow is being used by importing the RNN framework to build dynamic LSTM. This is a type of dynamic RNN with number of layers as 3, num_classes=3, pred_times=8, batch_size=64, state_size=64, feature_size=278. Using Epoch values of 1000,2000,3000 and 4000 we fix at 4000 as best validation and training accuracy was obtained. Adam Optimiser is used for reducing loss function. The highest training and Validation accuracy was found at Epoch=3487 of 98%, 77%. As seen in figure 18, instead of 4000 iterations we can stop the model training at 500 epoch which takes less time with almost similar accuracy.

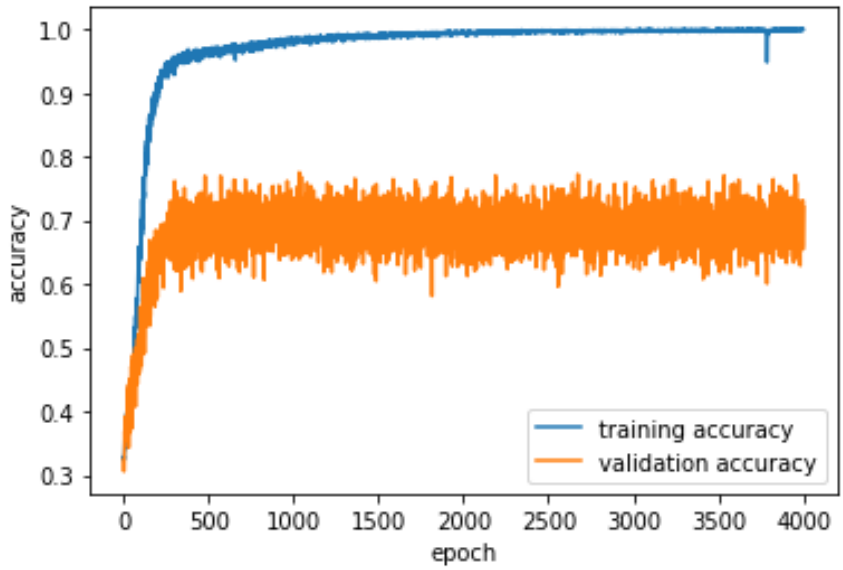

Figure 18: LSTM Accuracy vs Epoch

6.8 Discussion

Prediction using regression methods is performed on vital Clinical cognitive assessment metrics like ADAS13, Ventricles_Norm, and MMSE scores of an Alzheimer patient. Four experiments on the regression task have been performed using Elastic Net Regressor, Gradient Boosting Regressor, Support Vector Regressor, and Neural Net Regressor. The support vector regressor performed very poorly hence it was discarded. As seen in Figure 12, For ADAS 13 prediction test R^2 values of ElasticNet regressor and Gradient boosting methods compared to Baseline model which are 68%, 90% and 38% show Gradient boosting is the best model with 90% of data variability being explained by it when ADAS 13 is the response output variable. For Ventricles Norm as well Gradient Boosting Regressor (GBR) has better test R^2 (97%)(Coefficient of determination) compared to ElasticNet (83%) and Baseline (94%). Also, for MMSE prediction GBR with test R^2=79% has outperformed Elastic Net (58%) and Baseline (25%) showing best-fit model explaining 79% data variability when MMSE is used as a response variable. RMSE plays a vital role also in the evaluation of regression models. For ADAS 13 Neural net regressor, Gradient boosting regressor and ElasticNet regressor have test RMSE scores of 3.08, 2.93, 7.76 which have low values. GBR having the lowest scores indicate lesser standard deviation from variance unexplained. Overall, Gradient boosted Regressor is the best in terms of predicting ADAS 13, Ventricles Norm, and MMSE score. 5,10 fold cross-validation was performed had not shown much difference for each regressor method in terms of R^2.

For the Classification problem of distinguishing three Alzheimer disease progression stages NC, MCI and AD, there is a need to have an evaluation metric like F1 score which can balance the precision and recall to determine the better model. Hence it was calculated. Gradient boosting classifier (GBC) has higher accuracy CN: 84%, MCI: 77% AD: 87% and F1 score CN: 79% MCI: 73% AD: 67% compared to Support Vector Classifier (SVC) accuracy CN: 74%, MCI: 64% AD: 82% and F1 score CN: 66%, MCI: 66% AD: 61% as shown in Figure 16. This implies that GBC has more chances of predicting a patient at different stages CN, MCI, and AD in comparison to SVC. Four experiments are carried for AD classification, the other two methods are deep learning-based Neural Net classifier and RNN based dynamic LSTM network. From Table 5, it is observed that Dynamic LSTM network has got the best validation classification overall accuracy of 78% compared to Gradient Booster classifier (73%), Neural net classifier (67%) and Support Vector Classifier (61%). To avoid overfitting while training process dropout layers are introduced in LSTM and Neural net which leads to enhanced model accuracy. Also, cross-validation of 5,10 fold is performed however very little change in accuracy (<1%) is observed. This value signifies that Dynamic LSTM predicts patients' status rightly 78% of the times on the Alzheimer's clinical data from ADNI compared to an accuracy value of 73% in (Moore et al.; 2019) who performed Random Forest prediction using same data. However, our LSTM underperforms when compared to this paper accuracy of 87% (Vivar et al.; 2018) showing a state-of-the-art multi-modal approach of using Graph convolutional networks output with RNN to classify AD.

Models	**Train Accuracy**	**Validation Accuracy**
Support Vector Classifier	75%	61%
Neural Net Classifier	97%	67%
Gradient Boosting Classifier	84%	73%
Dynamic LSTM	**98%**	**78%**

Table 5: Performance Comparison of Classification Models

7 Conclusion and Future Work

The objective of the entire project is to identify potential Alzheimer patients at various stages CN, MCI, and AD for early treatment using Machine and Deep learning algorithms. The results are promising as it echoes intending to early diagnose of Alzheimer patients. Evaluation metrics used for regression are R^2, RMSE, for classification are Accuracy, Recall, Precision, and F1-Score. ADNI Dataset has cleaned appropriately before modeling.Regression was performed on predicting cognitive assessment metrics ADAS13, Ventricle Norm, and MMSE with 3 methods: Gradient boosting regressor being the best with R^2 above 90% for predicting continuous variable ADAS 13 and Ventricle Norm including lowest RMSE values compared to ElasticNet Regressor and Neural Network Regressor. This predicted values column can be used as a new input predictor for Alzheimer's classification. K fold Cross Validation has been performed.

Dynamic LSTM performed the best in early patient diagnosing of Cognitive normal, Mild Cognitive Impairment, and Alzheimer's disease with a validation accuracy of 78% in comparison to Deep Neural Net, Gradient Boosting Classifier and Support Vector Classifiers and other state-of-the-art methods. The chosen models give promising results however there is a scope of further improvement in solving the problem as we need to increase accuracy for this model to be used in real-time detection. The future scope would be to make use of the Multi-Modal or Ensemble approach of Machine, Deep learning algorithms to enable better classification of AD stages. Also, model tuning can be done by adding more layers in a neural network used. Given time limitations, further improvements would have been possible.

8 Acknowledgement

I am grateful to my supervisor Dr. Muhammad Iqbal for being resourceful and guiding me during the research project. I am thankful for his patience during 13 weeks duration.

References

Aksman, L. M., Scelsi, M. A., Marquand, A. F., Alexander, D. C., Ourselin, S., Altmann, A. and ADNI (2019). Modeling longitudinal imaging biomarkers with parametric bayesian multi-task learning, *Human brain mapping* **40**(13): 3982–4000.

Beheshti, I., Demirel, H., Initiative, A. D. N. et al. (2016). Feature-ranking-based alzheimer's disease classification from structural mri, *Magnetic resonance imaging* **34**(3): 252–263.

Biagioni, M. C. and Galvin, J. E. (2011). Using biomarkers to improve detection of alzheimer's disease, *Neurodegenerative disease management* **1**(2): 127–139.

Braithwaite, B., Paananen, J., Taipale, H., Tanskanen, A., Tiihonen, J., Hartikainen, S. and Tolppanen, A. (2020). Detection of medications associated with alzheimer's disease using ensemble methods and cooperative game theory, *International Journal of Medical Informatics* p. 104142.

Brookmeyer, R., Johnson, E., Ziegler-Graham, K. and Arrighi, H. M. (2007a). Forecasting the global burden of Alzheimer{\textquoteright}s disease, *Alzheimer's {&} Dementia: The Journal of the Alzheimer's Association* .

Brookmeyer, R., Johnson, E., Ziegler-Graham, K. and Arrighi, H. M. (2007b). Forecasting the global burden of alzheimer's disease, *Alzheimer's & dementia* **3**(3): 186–191.

Che, Z., Purushotham, S., Cho, K., Sontag, D. and Liu, Y. (2018). Recurrent neural networks for multivariate time series with missing values, *Scientific reports* **8**(1): 1–12.

Gers, F. A., Schmidhuber, J. and Cummins, F. (1999). Learning to forget: Continual prediction with lstm.

Ghazi, M. M., Nielsen, M., Pai, A., Cardoso, M. J., Modat, M., Ourselin, S., Sørensen, L., Initiative, A. D. N. et al. (2019). Training recurrent neural networks robust to incomplete data: application to alzheimer's disease progression modeling, *Medical image analysis* **53**: 39–46.

Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R. and Schmidhuber, J. (2016). Lstm: A search space odyssey, *IEEE transactions on neural networks and learning systems* **28**(10): 2222–2232.

Hampel, H., O'Bryant, S. E., Molinuevo, J. L., Zetterberg, H., Masters, C. L., Lista, S., Kiddle, S. J., Batrla, R. and Blennow, K. (2018). Blood-based biomarkers for alzheimer disease: mapping the road to the clinic, *Nature Reviews Neurology* **14**(11): 639–652.

Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory, *Neural computation* **9**(8): 1735–1780.

Hosseini-Asl, E., Keynton, R. and El-Baz, A. (2016). Alzheimer's disease diagnostics by adaptation of 3d convolutional network, *2016 IEEE International Conference on Image Processing (ICIP)*, IEEE, pp. 126–130.

Iddi, S., Li, D., Aisen, P. S., Rafii, M. S., Thompson, W. K., Donohue, M. C., Initiative, A. D. N. et al. (2019). Predicting the course of alzheimer's progression, *Brain informatics* **6**(1): 6.

Izquierdo, W., Martin, H., Cabrerizo, M., Barreto, A., Andrian, J., Rishe, N., Gonzalez-Arias, S., Loewenstein, D., Duara, R. and Adjouadi, M. (2017). Robust prediction of cognitive test scores in alzheimer's patients, *2017 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, IEEE, pp. 1–7.

Lei, B., Yang, M., Yang, P., Zhou, F., Hou, W., Zou, W., Li, X., Wang, T., Xiao, X. and Wang, S. (2020). Deep and joint learning of longitudinal data for alzheimer's disease prediction, *Pattern Recognition* **102**: 107247.

Lipton, Z. C., Kale, D. C. and Wetzel, R. (2016). Modeling missing data in clinical time series with rnns, *arXiv preprint arXiv:1606.04130* .

Lu, P. (2019). *Statistical Learning from Multimodal Genetic and Neuroimaging data for prediction of Alzheimer's Disease*, PhD thesis, Sorbonne Université.

Marinescu, R. V., Oxtoby, N. P., Young, A. L., Bron, E. E., Toga, A. W., Weiner, M. W., Barkhof, F., Fox, N. C., Eshaghi, A., Toni, T. et al. (2020). The alzheimer's disease prediction of longitudinal evolution (tadpole) challenge: Results after 1 year follow-up, *arXiv preprint arXiv:2002.03419* .

Marinescu, R. V., Oxtoby, N. P., Young, A. L., Bron, E. E., Toga, A. W., Weiner, M. W., Barkhof, F., Fox, N. C., Golland, P., Klein, S. and Alexander, D. C. (2019). TADPOLE challenge: Accurate alzheimer's disease prediction through crowdsourced forecasting of future data, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*.

Mehdipour Ghazi, M., Nielsen, M., Pai, A., Modat, M., Cardoso, M. J., Ourselin, S. and Sørensen, L. (2019). Robust parametric modeling of alzheimer's disease progression, *arXiv* pp. arXiv–1908.

Moore, P., Lyons, T., Gallacher, J. and Initiative, A. D. N. (2019). Random forest prediction of alzheimer's disease using pairwise selection from time series data, *PloS one* **14**(2): e0211558.

Oxtoby, N. P. and Alexander, D. C. (2017). Imaging plus x: multimodal models of neurodegenerative disease, *Current opinion in neurology* **30**(4): 371.

Oxtoby, N. P., Young, A. L., Cash, D. M., Benzinger, T. L., Fagan, A. M., Morris, J. C., Bateman, R. J., Fox, N. C., Schott, J. M. and Alexander, D. C. (2018). Data-driven models of dominantly-inherited alzheimer's disease progression, *Brain* **141**(5): 1529–1544.

Pearlmutter, B. A. (1989). Learning state space trajectories in recurrent neural networks, *Neural Computation* **1**(2): 263–269.

Peña-Bautista, C., Durand, T., Oger, C., Baquero, M., Vento, M. and Cháfer-Pericás, C. (2019). Assessment of lipid peroxidation and artificial neural network models in early alzheimer disease diagnosis, *Clinical biochemistry* **72**: 64–70.

Petersen, R. C., Aisen, P., Beckett, L. A., Donohue, M., Gamst, A., Harvey, D. J., Jack, C., Jagust, W., Shaw, L., Toga, A. et al. (2010). Alzheimer's disease neuroimaging initiative (adni): clinical characterization, *Neurology* **74**(3): 201–209.

Raket, L. L. (2019). Disease progression modeling in alzheimer's disease: insights from the shape of cognitive decline, *medRxiv* .

Schouten, T. M., Koini, M., de Vos, F., Seiler, S., van der Grond, J., Lechner, A., Hafkemeijer, A., Möller, C., Schmidt, R., de Rooij, M. et al. (2016). Combining anatomical, diffusion, and resting state functional magnetic resonance imaging for individual classification of mild and moderate alzheimer's disease, *NeuroImage: Clinical* **11**: 46–51.

Shahbaz, M., Ali, S., Guergachi, A., Niazi, A. and Umer, A. (2019). Classification of alzheimer's disease using machine learning techniques., *DATA*, pp. 296–303.

Shen, L., Kim, S., Qi, Y., Inlow, M., Swaminathan, S., Nho, K., Wan, J., Risacher, S. L., Shaw, L. M., Trojanowski, J. Q. et al. (2011). Identifying neuroimaging and proteomic biomarkers for mci and ad via the elastic net, *International Workshop on Multimodal Brain Image Analysis*, Springer, pp. 27–34.

Statista (2020). Global forecast of the total number of people with dementia worldwide from 2015 to 2050 (in millions).
URL: *https://www.statista.com/statistics/471423/total-global-dementia-affected-population-forecast/*

Teipel, S. J., Grothe, M. J., Metzger, C. D., Grimmer, T., Sorg, C., Ewers, M., Franzmeier, N., Meisenzahl, E., Klöppel, S., Borchardt, V. et al. (2017). Robust detection of impaired resting state functional connectivity networks in alzheimer's disease using elastic net regularized regression, *Frontiers in aging neuroscience* **8**: 318.

Venkatraghavan, V., Bron, E. E., Niessen, W. J., Klein, S., Initiative, A. D. N. et al. (2019). Disease progression timeline estimation for alzheimer's disease using discriminative event based modeling, *NeuroImage* **186**: 518–532.

Vivar, G., Zwergal, A., Navab, N. and Ahmadi, S.-A. (2018). Multi-modal disease classification in incomplete datasets using geometric matrix completion, *Graphs in Biomedical Image Analysis and Integrating Medical Imaging and Non-Imaging Modalities*, Springer, pp. 24–31.

Weiner, M. W., Veitch, D. P., Aisen, P. S., Beckett, L. A., Cairns, N. J., Green, R. C., Harvey, D., Jack, C. R., Jagust, W., Liu, E., Morris, J. C., Petersen, R. C., Saykin, A. J., Schmidt, M. E., Shaw, L., Siuciak, J. A., Soares, H., Toga, A. W. and Trojanowski, J. Q. (2012). The Alzheimer's disease neuroimaging initiative: A review of papers published since its inception, *Alzheimer's and Dementia* .