

Dementia Prediction Model *Project 4*

Contributors:

**Hansle Goh, Dumiduni
Geeganage, Noelle Watson**

Purpose

What is Dementia?

Dementia is the impaired ability to remember, think, or make decisions that interferes with doing everyday activities.

Alzheimer's disease is the most common type of dementia. Though dementia mostly affects older adults, it is not a part of normal aging.

[Source](#)

Problem statement

Can we use machine learning models such as logistic regression, random forests, and deep learning to predict dementia?



Methodology

Data collection

- Dataset #1 - Dementia Patient Health and Prescription Dataset (*Kaggle*)
- Dataset #2 - Classification and Prediction of Dementia by SVM (*Data.Mendeley*)

Exploratory Data analysis

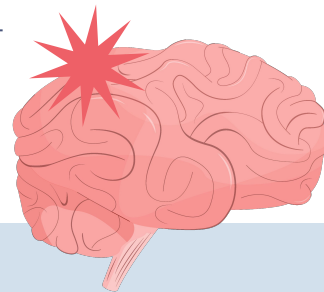
Used Pandas and Spark to explore the Data

Predictive Models

Deep Learning
Random Forest
Logistic Regression

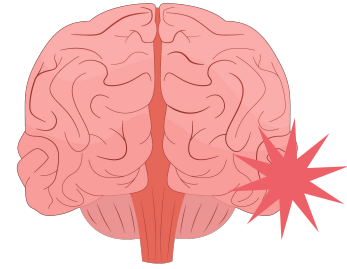
Results validation

Confusion Matrix
Classification Report
SHAP Values
Feature Importances



Overview (Dataset #1):

1,000 records:



Dementia Diagnosis

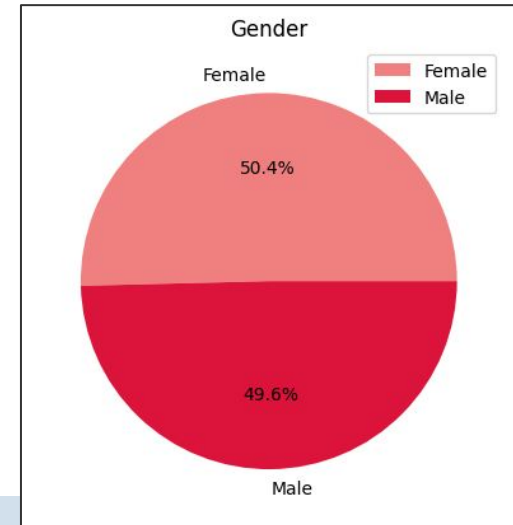
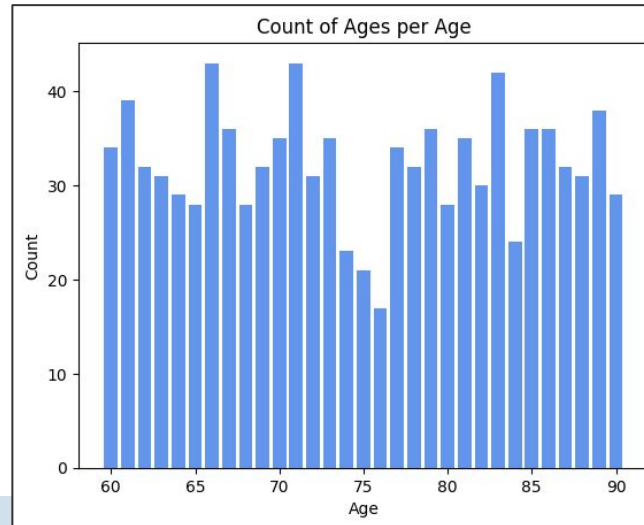
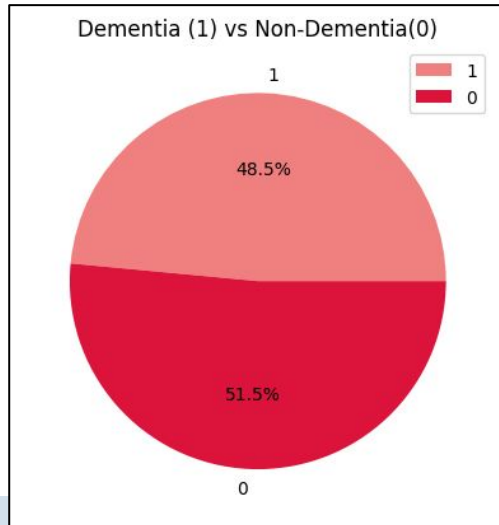
- Does not have dementia: 515
- Has dementia: 485

Age Range

- 60 - 90

Gender

- Female: 504
- Male: 496

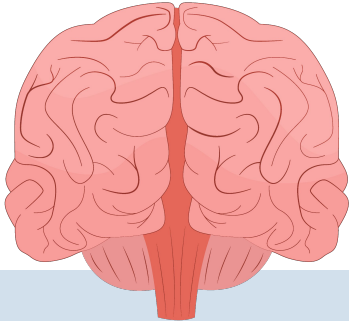


Data Features (Dataset #1)

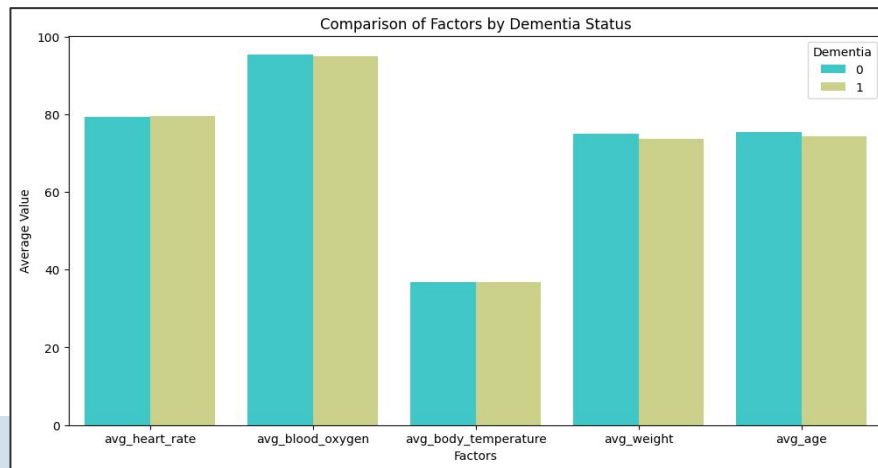
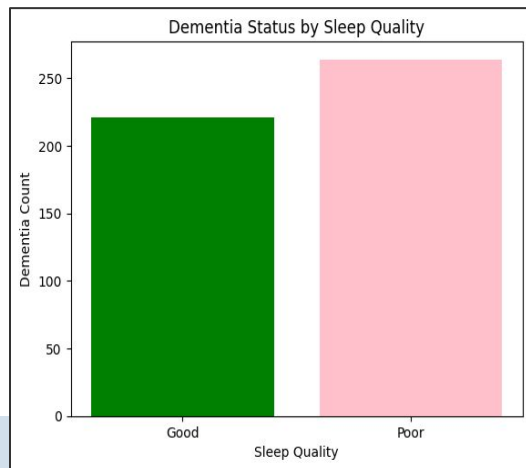
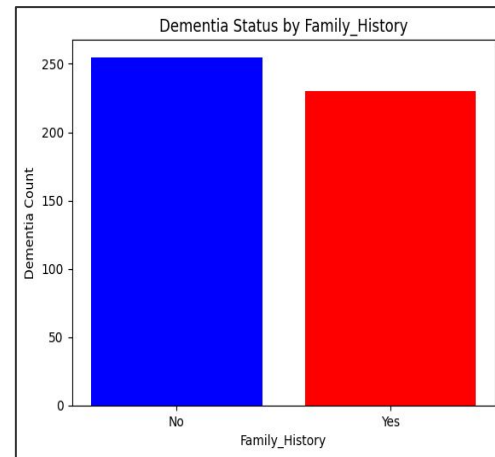
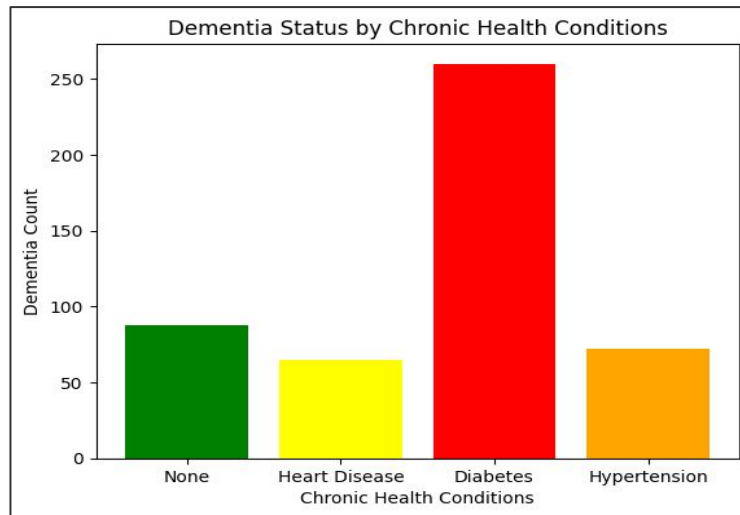
Column Name	Details
Alcohol Level	Alcohol Consumption Level
Heart Rate	Heartbeats per minute
Blood Oxygen Levels	Saturation of oxygen in the blood
Body Temperature	Celcius
Weight	Kilograms
MRI Delay	Time delay in obtaining an MRI Scan
Age	60-90

Column Name	Details
Education Level	Highest level completed
Dominant Hand	Right / Left
Gender	Male / Female
Family History	Family history of dementia
Smoking Status	Smoker / Non-Smoker
APOE e4	Presence of a gene variant associated with Alzheimer's
Physical Activity	Level of physical activity in patient
Depression Status	Depressed / Not Depressed

Column Name	Details
Cognitive Test Scores	Scores from Cognitive Tests
Medication History	History or medication use
Sleep Quality	Patient quality of sleep
Chronic Health Conditions	Lists any chronic health conditions
Dementia	Demented / Not Demented



Data Exploration

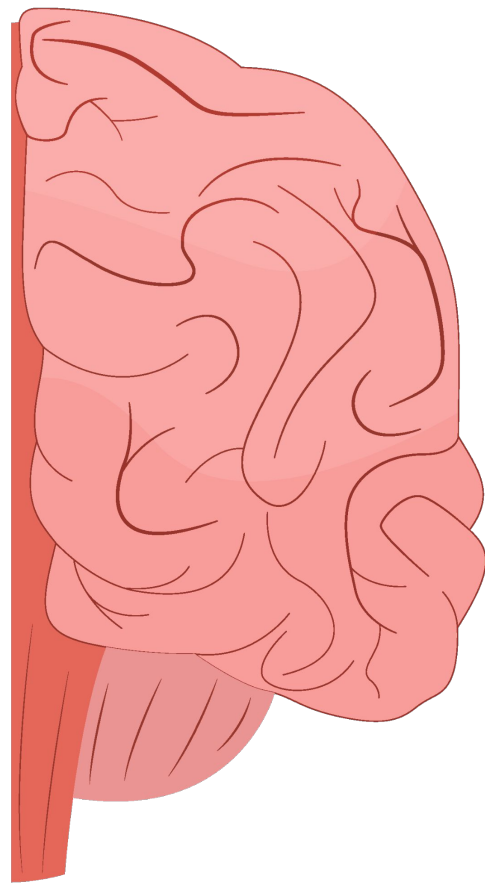


Attempted Models:

Deep Learning Model

Random Forest Model

Logistic Regression Model



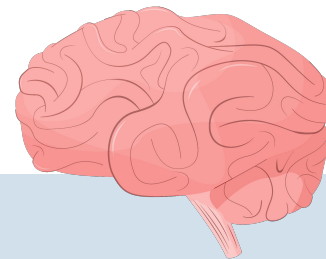
Deep Learning Model

Accuracy: 99.20%

Loss: 0.024

Model: "sequential"

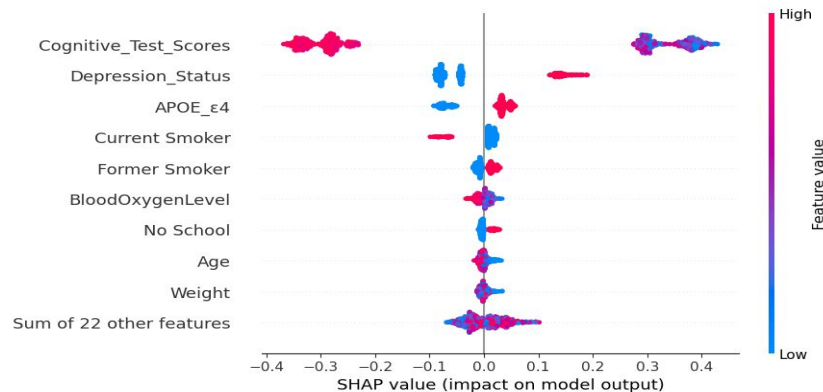
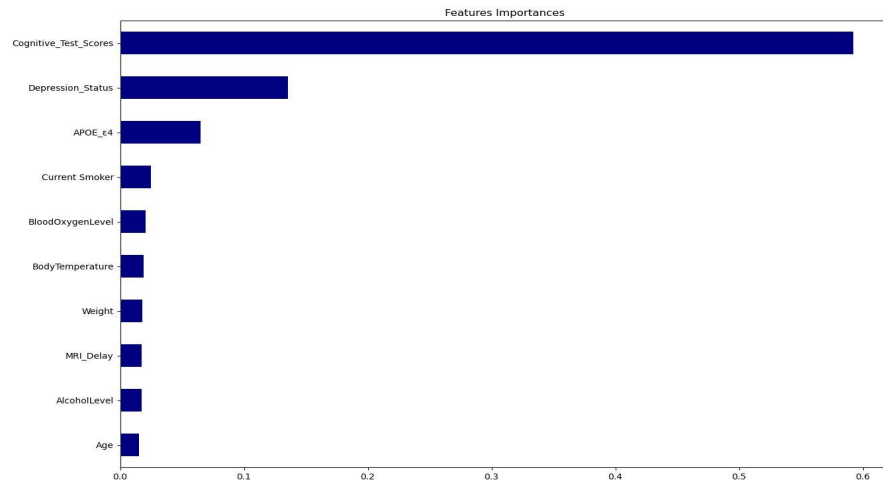
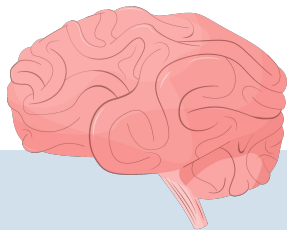
Layer(type)	Output Shape	Param #
dense(Dense)	(None, 436)	13952
dense_1(Dense)	(None, 271)	118427
dense_2(Dense)	(None, 151)	41072
dense_3(Dense)	(None, 66)	10032
dense_4(Dense)	(None, 181)	12127
dense_5(Dense)	(None, 271)	49322
dense_6(Dense)	(None, 121)	32912
dense_7(Dense)	(None, 1)	122
dense_8(Dense)	(None, 1)	2



Random Forest Model

Accuracy: 100%

Highest Feature Importance: 0.592

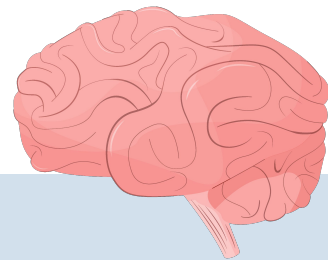


Logistic Regression Model

Accuracy: 98.80%

3 Type II Errors

	Predicted 0	Predicted 1
Actual 0	126	0
Actual 1	3	121



Accuracies Without Score Feature

Deep Learning

Loss: 0.4415

80.80%

Random Forest

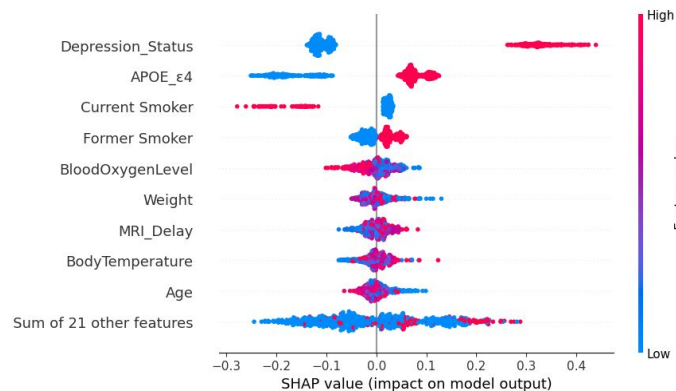
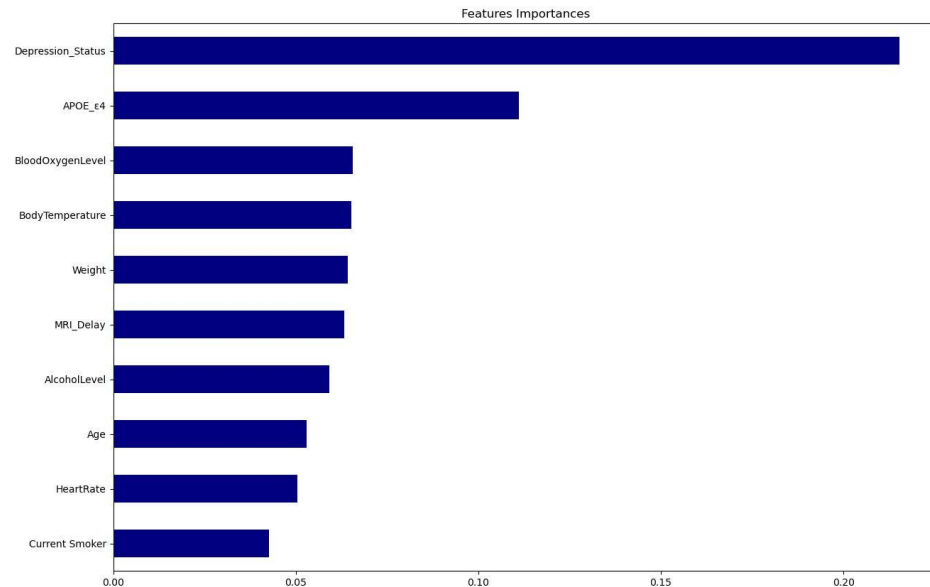
Type I and II Errors

74.00%

Logistic Regression

Type I and II Errors

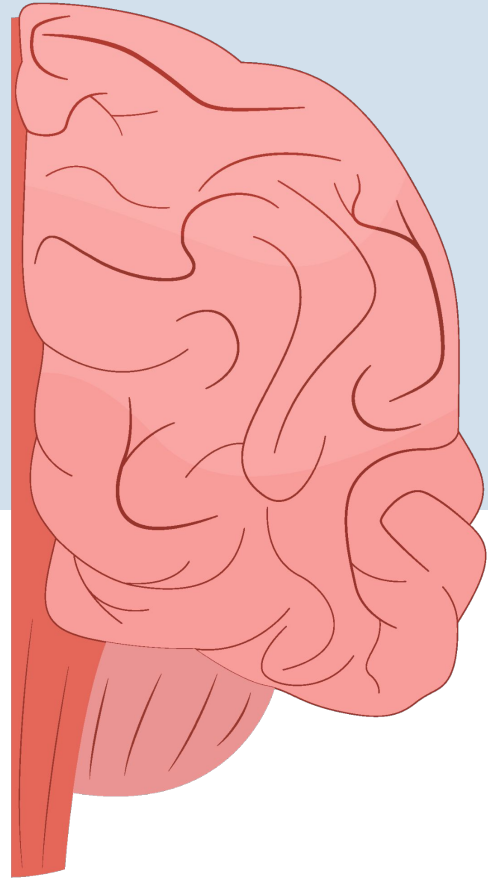
74.40%



Website:

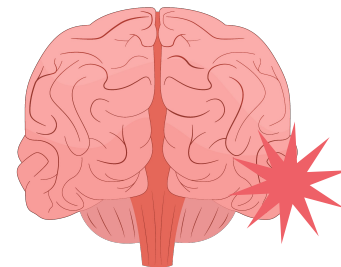
Dataset #1

Dumiduni / Noelle



Overview (Dataset #2):

336 records:



Dementia Diagnosis

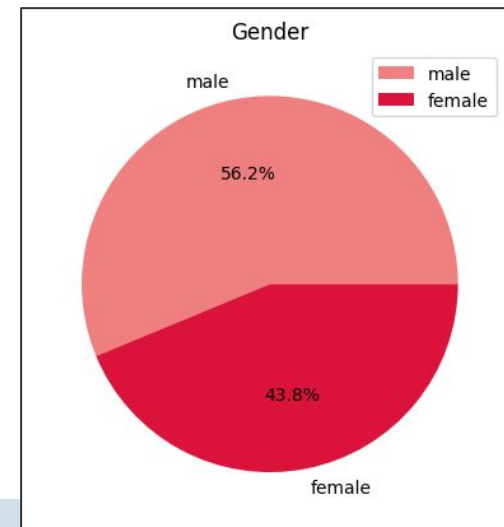
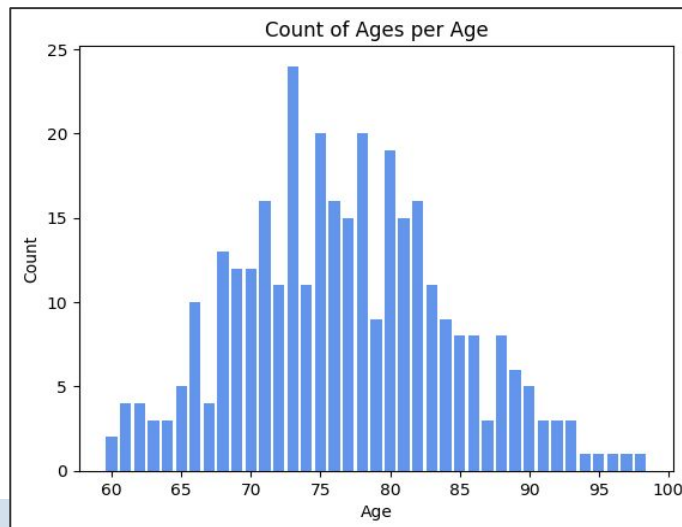
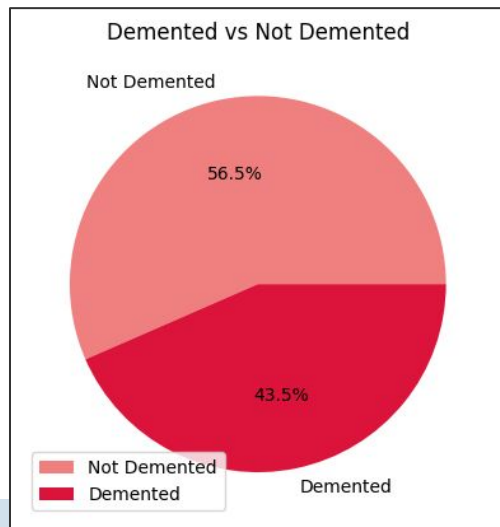
- Does not have dementia: 190
- Has dementia: 146

Age Range

- 60 - 98

Gender

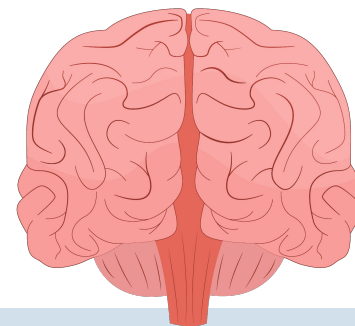
- Female: 147
- Male: 189



Data Features (Dataset #2)

Column Name	Details
Group	Demented/ Non-Demented
Visit	Which visit the subject was scanned on
MR Delay	Magnetic Resonance Delay
M/F	Male / Female
Hand	Right /Left
Age	60-96
EDUC	Highest grade completed

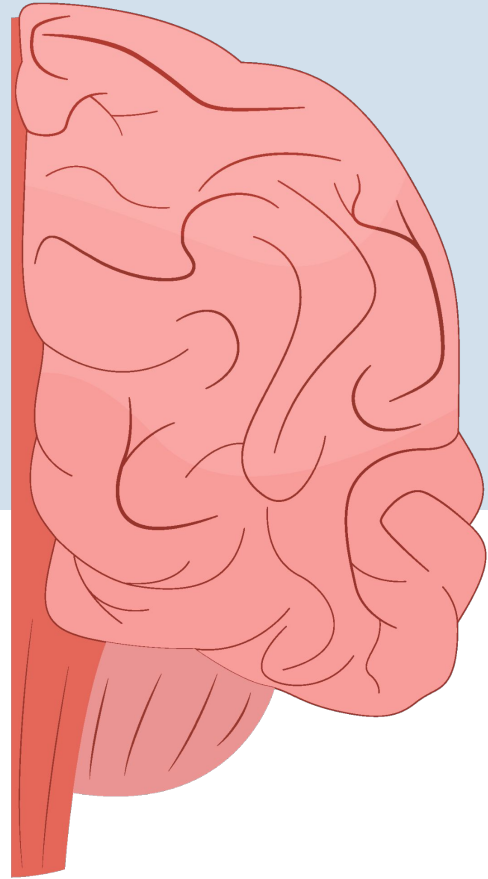
Column Name	Details
SES	Socioeconomic Status
MMSE	Mini-Mental State Exam
CDR	Clinical Dementia Rating
eTIV	Estimated Total Intracranial Volume
nWBV	Normalized Whole Brain Volume
ASF	Atlas Scaling Factor



Predictive Model:

Dataset #2

Dumiduni



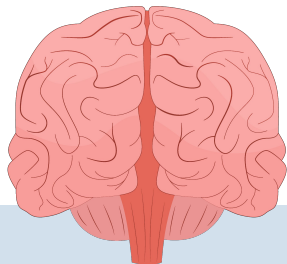
Predictive Model - Data with Rating Scales

Confusion Matrix

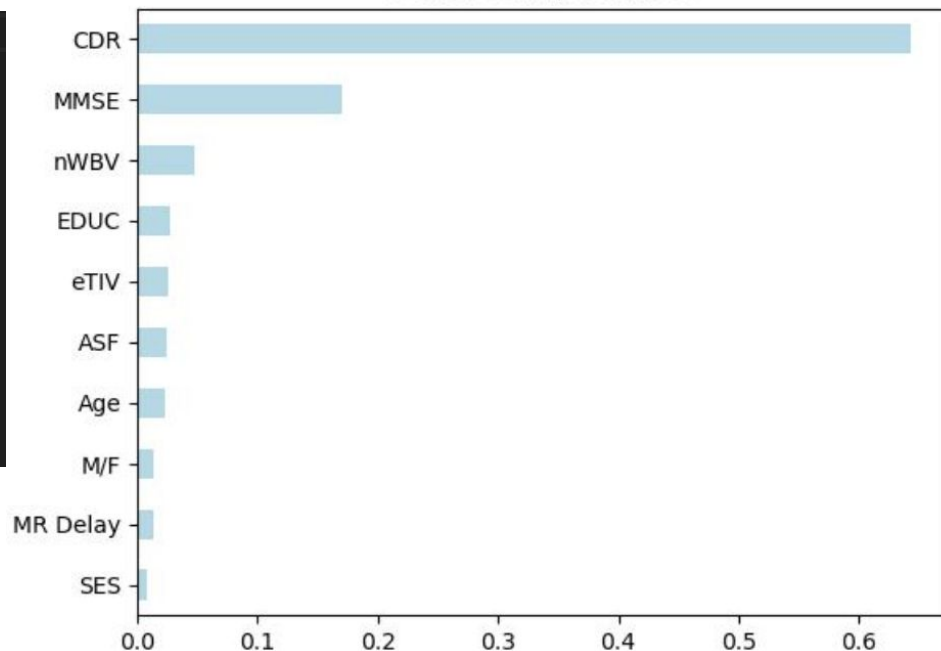
Accuracy Score : 1.0

Classification Report

		precision	recall	f1-score	support
	0	1.00	1.00	1.00	48
	1	1.00	1.00	1.00	36
accuracy				1.00	84
macro avg		1.00	1.00	1.00	84
weighted avg		1.00	1.00	1.00	84



Features Importances



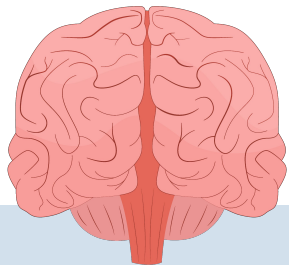
Predictive Model - Data with Rating Scales

Confusion Matrix

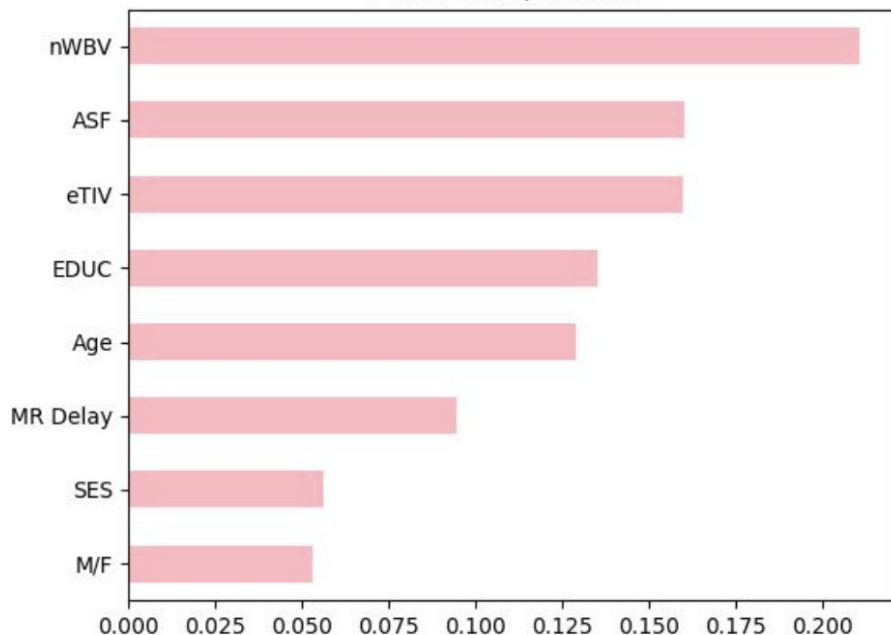
Accuracy Score : 0.8214285714285714

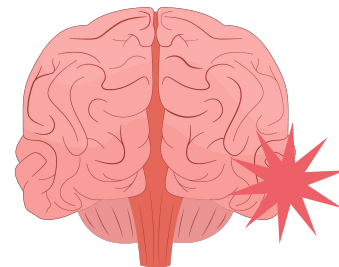
Classification Report

		precision	recall	f1-score	support
	0	0.84	0.85	0.85	48
	1	0.80	0.78	0.79	36
	accuracy			0.82	84
	macro avg	0.82	0.82	0.82	84
	weighted avg	0.82	0.82	0.82	84



Features Importances





Resources

- [Dataset #1](#)
- [Dataset #2](#)
- [Show all column names](#)
- [Save Keras model summary](#)
- [Keras tuner](#)
- [Feature importance in random forests](#)
- [Markdown-Table of contents](#)
- [Template for slides](#)