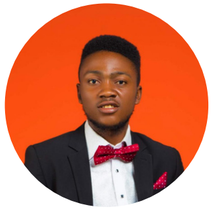
Classification & Prediction of Dementia

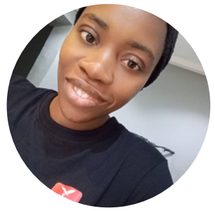
27th August 2022

Our Team

Abdulmalik Adeyemo
Project Team Lead
Presenter 1



Adewumi Solomon
Presenter 2



Glory Eke
Assistant PL

Odedola Solomon
Query Analyst

Kalyani Pusadkar
Data Science

Oluwatosin Osukoya
Data Storytelling

Jesujana Kayode
Data Science

Elizabeth Okon
Data Storytelling

Justus Ilemobayo
DevOps

Godson E.
Data science

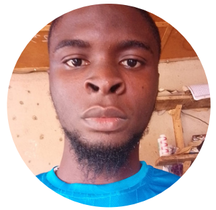
Joseph Michael
Data Science

Teminijesu Jesufemi
Data Science

Udeh Sandra
Data Science

Onaolapo Sunday
Data Science

Chidiebere Nnadiegbulam
DevOps



Akinyemi A. A.
Data Science

Moses O.
Data Science

Oswald Ohiole Ojo
Data Science

Emuejevoke E.
Data Science

Amao Jacobs
Data Science

Problem Statement

Dementia

- The seventh leading cause of mortality
- A major cause of disability among older people worldwide
- Leads to stigmatization and barriers to diagnosis and care

Existing Solution

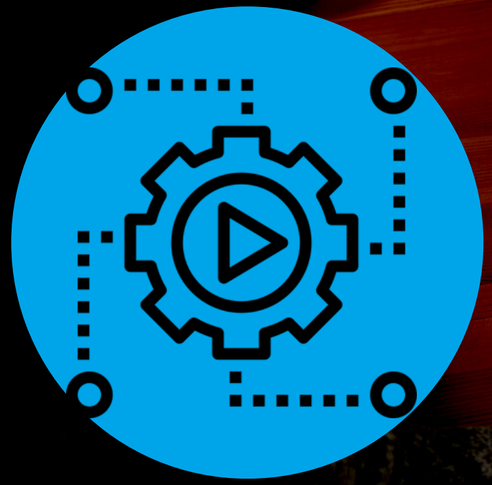
Manual Diagnosis System

- Doctor checks medical history, symptoms and conduct a physical examination

- Doctors would ask someone close to the patient about the symptoms

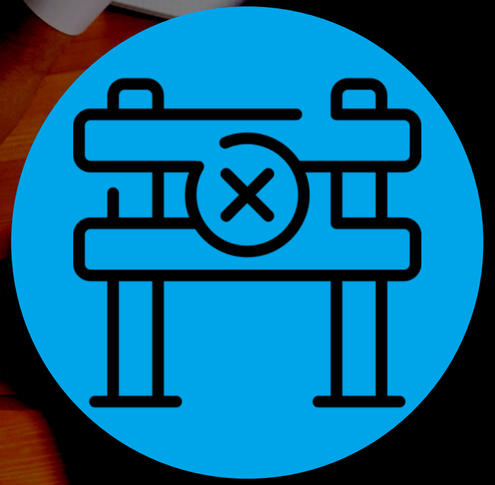
Our Solution

An Artificial Intelligent Diagnostic System



Automated Diagnosis without Specialist

Accurate Result



Breaks Barrier to diagnosis and care

The Dataset

HAMOYE

The data was collected from Kaggle and the description is as follows:

- consists of a longitudinal collection of 150 subjects aged 60 to 96
- Each subject was scanned on two or more visits, for a total of 373 imaging sessions
- 72 of the patients were nondemented and 64 were demented
- 51 patients were diagnosed with mild to moderate Alzheimer's disease
- 14 patients were categorized as converted
- Important features are; "EDUC", "Sex", "CDR", "eTIV", "MMSE", "Age", "nWBV"

The Dataset

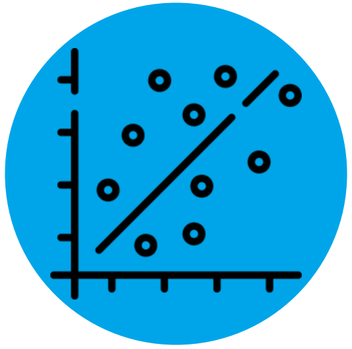
HAMOYE

The data preparation process is as follows:

Dropped default columns

Categorical to Numeric Conversion

Impute missing values



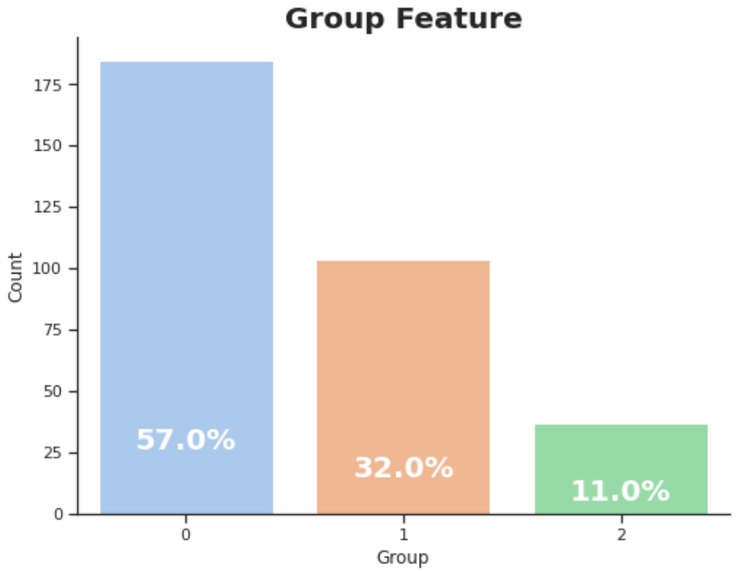
Feature - Output correlation

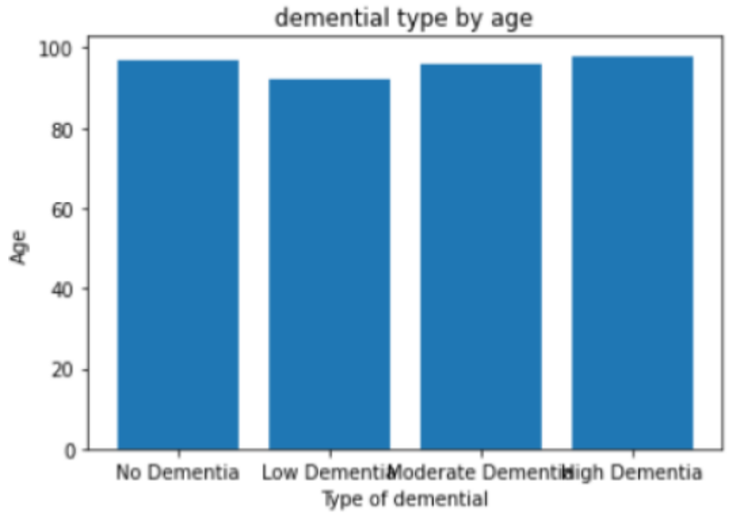
Outliers Detection & Removal

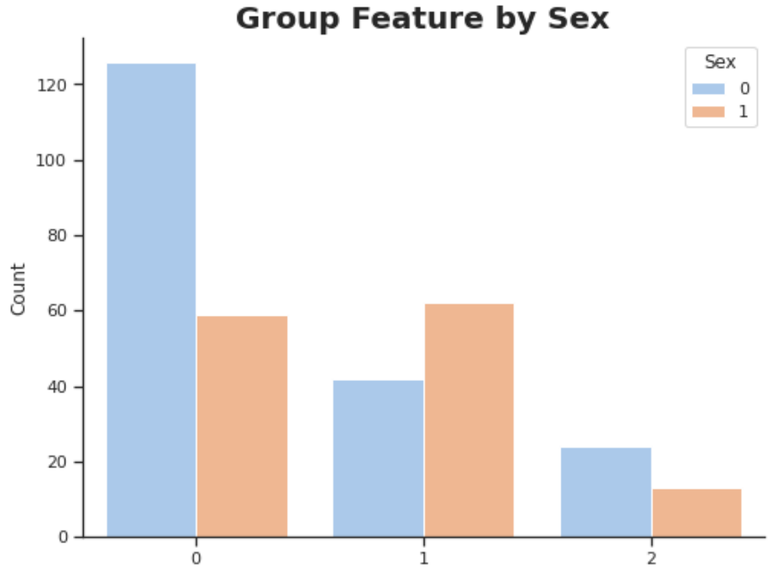
Features Normalization

The Dataset

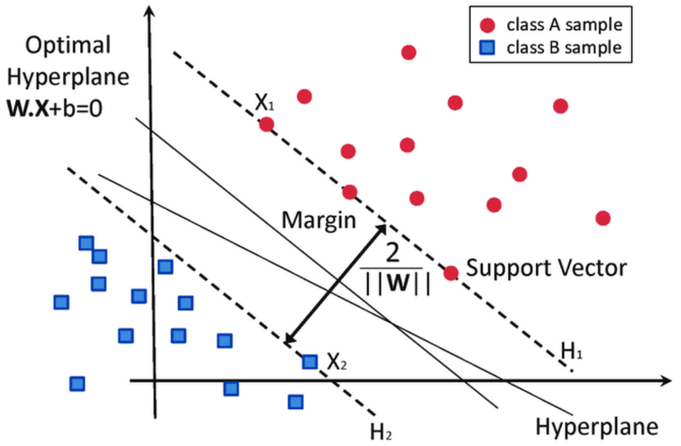
Visualizations



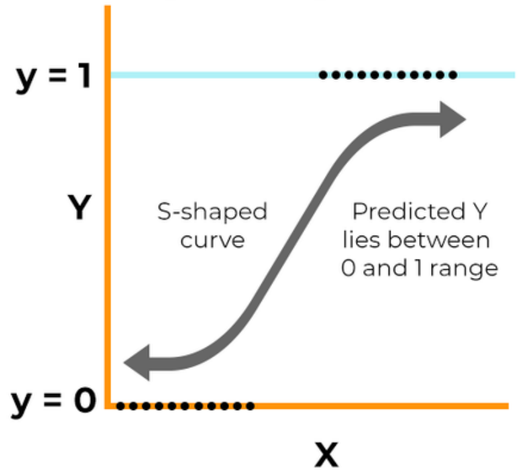




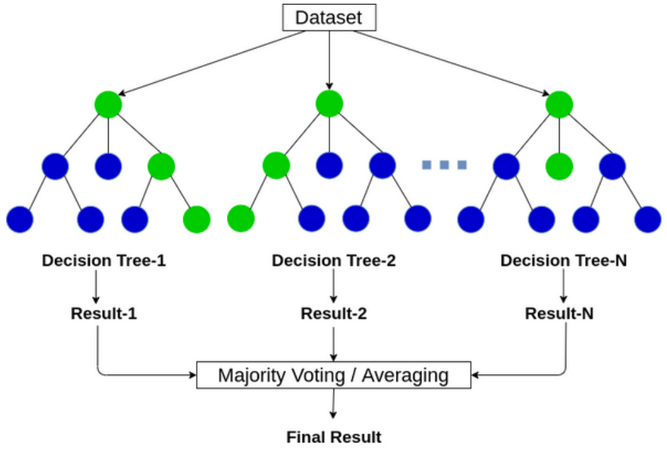
Model



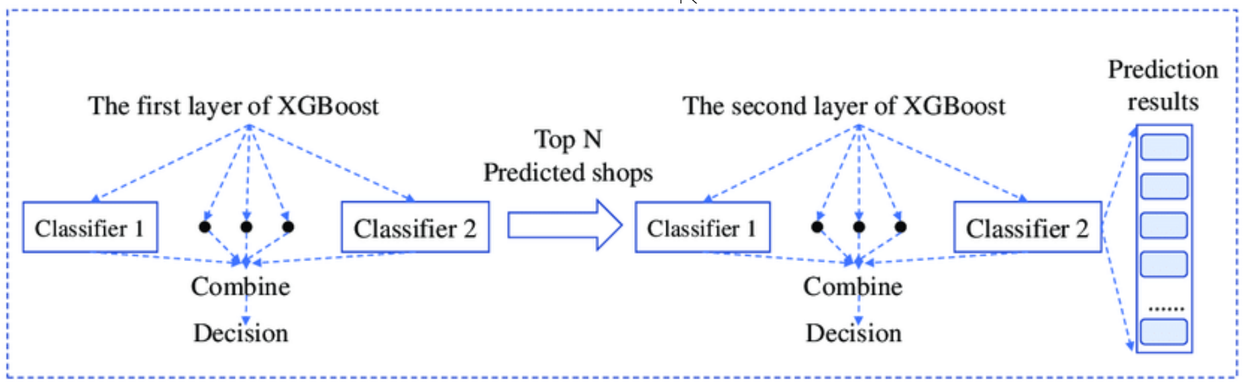
Support Vector Machine



Logistic Regression



Random Forest



Xtreme Gradient Boosting

Model

Logistic Regression

	precision	recall	f1-score	support
0	0.88	1.00	0.94	37
1	0.88	1.00	0.93	21
2	0.00	0.00	0.00	8
accuracy			0.88	66
macro avg	0.59	0.67	0.62	66
weighted avg	0.77	0.88	0.82	66

Random Forest

	precision	recall	f1-score	support
0	0.88	1.00	0.94	37
1	0.95	1.00	0.98	21
2	1.00	0.25	0.40	8
accuracy			0.91	66
macro avg	0.95	0.75	0.77	66
weighted avg	0.92	0.91	0.88	66

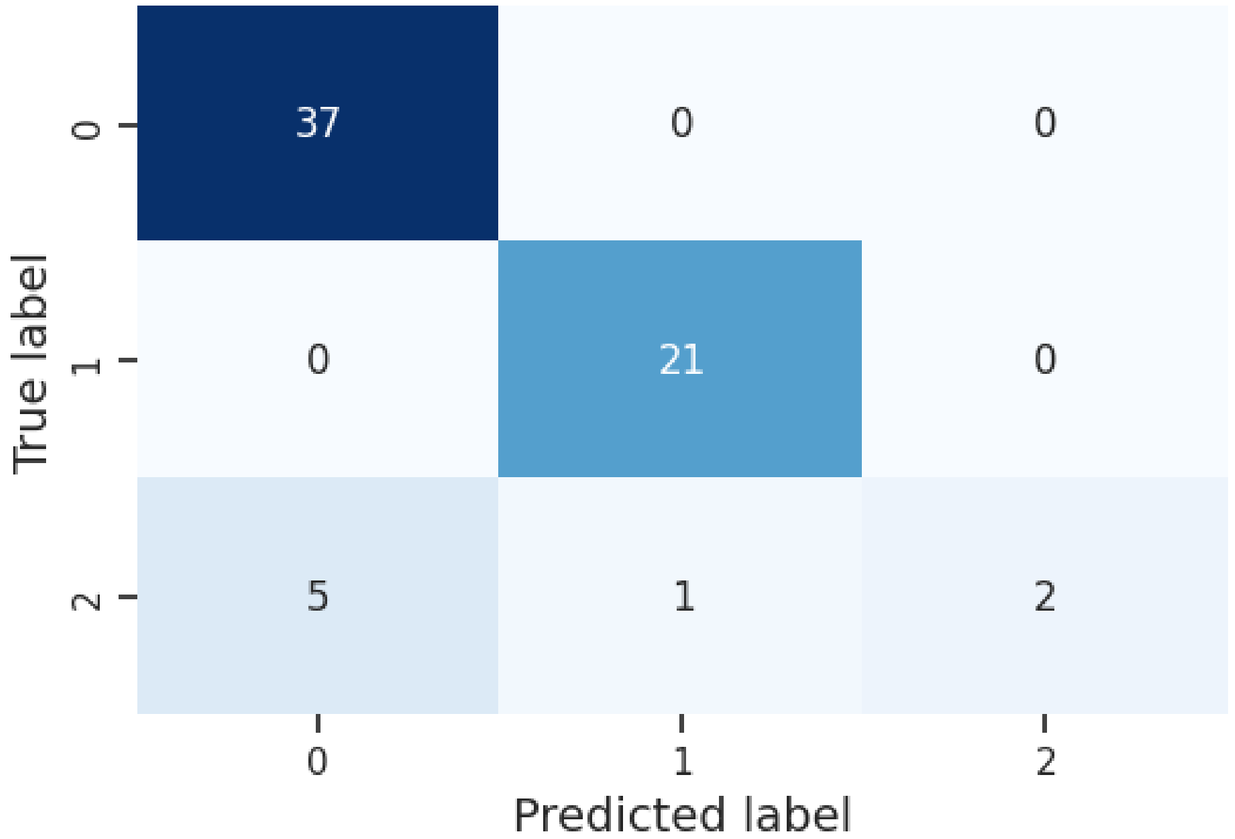
Support vector machine

	precision	recall	f1-score	support
0	0.88	1.00	0.94	37
1	0.91	1.00	0.95	21
2	1.00	0.12	0.22	8
accuracy			0.89	66
macro avg	0.93	0.71	0.70	66
weighted avg	0.91	0.89	0.86	66

XGradient boost

	precision	recall	f1-score	support
0	0.86	1.00	0.92	37
1	0.95	0.90	0.93	21
2	0.67	0.25	0.36	8
accuracy			0.88	66
macro avg	0.83	0.72	0.74	66
weighted avg	0.87	0.88	0.86	66

Model



Random Forest Confusion Matrix

Summary

- We can determine the type of dementia by looking at the CDR value
- Converted category should be changed to demented or nondemented for real-time application
- It was observed that the models performed better without the converted category