Introduction:

Alzheimer's disease (AD) is a progressive neurological disorder that affects memory, thinking, and behavior. Early detection is crucial for improving patient care and management, but it remains a significant challenge due to the complexity of contributing factors and the long progression period of the disease. This project aims to develop a predictive model to assess the likelihood of an individual having Alzheimer's disease based on a range of features, including demographic details, medical history, lifestyle factors, clinical measurements, and cognitive assessments.

Below is a summary of the modeling approach for this project.

Goals:

- To obtain a good baseline model
- Select models that produce high accuracy rates using all the features
- Adapt an approach which improves the accuracy for early detection
    - By excluding the highly correlated features in modeling
    - Restricting to the dataset with no memory complaint

Models Explored:

We will be exploring the following models in this project.

1. Logistic Regression
2. Random forest classifier
3. XGBoost classifier
4. Gradient boosting classifier
5. Adaboost classifier
6. KNN classifier
7. Support Vector machine
8. Decision Tree
9. LDA
10. QDA
11. Naïve Bayes

Metrics Checked:

To evaluate the performance of the models, we will be considering the following metrics.

- Accuracy Score
- F1 Score
- Precision Score
- Recall Score
- Confusion Matrices