

PREDICTING ALZHEIMER'S DISEASE WITHOUT USING IMAGING DATA

Ahana Sil[1], Subhajit Chakrabarty[2]

[1] South Brunswick High School, NJ

[2] Louisiana State University Shreveport

Corresponding Author: Ahana Sil
Email: ahana.sil01@gmail.com
doi: 10.34107/UDUK9890452

ABSTRACT

Alzheimer's Disease (AD) remains one of the top ten causes of death globally for the past several years. Due to increased life expectancy and aging population worldwide, a steady rise in the population affected by AD and other forms of dementia is forecasted by 2050. In recent years, triangulation of multiple sources of reports such as patient cognitive tests, biomarkers and advanced brain imaging, claim to have enabled improved diagnoses and prediction. This study attempts to identify the most weighted determinants of AD and to test whether cost convenient data sources such as medical history, biomarkers, socioeconomic status and general health, are effective to predict Alzheimer's without expensive brain imaging reports. For this study, data from Alzheimer's Disease Neuroimaging Initiative (ADNI) was used for preliminary analysis. Baseline subjects (N=1,798) over the four phases (ADNI 1, ADNI GO, ADNI 2 and ADNI 3) were observed. The main objective was to assess the predictability of imaging versus non-imaging data for Alzheimer's disease. The unit of analysis was the patient. Variables pertaining to socioeconomic indicators, genetic and blood biomarkers, cognitive test scores and medical history (non-imaging) were modelled to predict Alzheimer's, controlling for MRI biomarkers (imaging). Using Singular Value Decomposition (SVD), significant components were identified separately for Medical History, Cognitive Tests and MRI markers. The SVD components, gender, education years, genetic marker and other variables, were all added to the model. A comprehensive comparison of Ordinary Least Squares (OLS), Lasso, Ridge, ElasticNet, Multi-layer Perceptron (MLP), and TabNet regression models was conducted using 5-fold cross-validation. The performance of each model was evaluated based on Mean Squared Error (MSE) and R-squared, both with and without the inclusion of MRI-derived SVD components. The results indicate that OLS and Ridge Regression consistently demonstrated the best performance, achieving the highest R-squared and lowest MSE values, with minimal impact from the inclusion of MRI features. While other models like MLP and TabNet showed some potential, they did not surpass the performance of the linear models with the current feature set and default/initial hyperparameters. The findings suggest that simpler linear models are most suitable for this prediction task with the available features, although further hyperparameter tuning for more complex models could potentially yield improvements. The preliminary findings suggest that statistically, expensive MRI imaging may not be critical to predict Alzheimer's disease. Cognitive tests and blood biomarkers alone may predict the progression of the disease. This may have cost implications for remote rural or financially disadvantaged elderly population.

Keywords: Alzheimer, ADNI, imaging, MRI, prediction

INTRODUCTION

Alzheimer's disease (AD), the most common of all other neurodegenerative diseases, has persistently remained one of the top ten causes of death globally for the past several years. Predictions indicate a forthcoming "tsunami" of individuals who will be affected by Alzheimer's and other forms of dementia in the coming decades [1] due to globally increasing life expectancy coupled with an aging population [2,3]. The population affected by dementia worldwide is projected to increase manifold from the estimated 57 million in 2019 to 158 million by 2050, driving a surge in research and development in recent years with hopes of preventing the early onset of the disease for as long as possible [4]. Reportedly, a staggering 100,000 people die annually in the United States due to dementia. With the baby boomer