Predicting Alzheimer's Disease Using Driving Simulator Data

Ryan Blanchette[1], Anahita Khojandi[1], Daniel Cox[2], Michael Oliver[3], Roberto Fernandez[3]

***Abstract*— Early detection of Alzheimer's Disease (AD) is critical in creating better outcomes for patients. Performance in complex tasks such as vehicular driving may be a sensitive tool for early detection of AD and serve as a good indicator of functional status. In this study, we investigate the classification of AD patients and controls using driving simulator data. Our results show that machine learning algorithms, especially random forest classifier, can accurately discriminate AD patients and controls (AUC = 0.96, Sensitivity = 87%, and Specificity = 93%). The model-identified most important features include Pothole Avoidance, Road Signs Recalled, Inattention Measurements, Reaction Time, and Detection Times, among others, all of which closely align with previous studies about cognitive functions that are affected by AD.**

I. Introduction

Among the top ten leading causes of death in the United States, Alzheimer's Disease (AD) is the only one that cannot be prevented, slowed down, or cured [1]. Already on its way to reaching epidemic proportions, estimates show that AD will affect 14 million individuals in the U.S. by 2060 [2], [3]. Clinically, this neurodegenerative disease is characterized by the insidious decline of cognitive function, with progressive impairment in multiple domains includes memory, visuospatial function, executive function, language, and attention [4]. The pathological hallmark of the disease is the accumulation of abnormal proteins, known as amyloid and tau, in a cascade of events that may begin decades before the onset of symptoms [5]. Various experimental drugs targeting these abnormal proteins have been developed. However, to date, none have achieved the desired outcome of slowing down or halting disease progression [5] presumably because such interventions have been instituted too late in the course of the disease [6]. These results highlight the importance of making a diagnosis in the pre-symptomatic stage, where therapeutic interventions are most likely to have a meaningful impact. Furthermore, early detection allows for prompt initiation of currently available symptomatic treatments as well as non-pharmacological interventions, lifestyle modifications, and supportive care that can significantly improve quality of life [7].

Standardized neuropsychological testing remains the gold standard for clinical diagnosis of AD, but such tests are often not sensitive enough for preclinical detection. The more recent development of brain imaging and spinal fluid biomarkers has opened the possibility of preclinical detection. However, the use of these tools is either too invasive or expensive for screening in the general population. [8]. Therefore, the development of non-invasive, cost-effective, and widely accessible screening tools for early detection remains a priority in the field.

Highly demanding and complex behavioral tasks that are the first to be affected in AD [9] most often have the potential to be used as tools for early detection. Vehicular driving, in particular, requires the interaction of multiple cognitive domains and could be a sensitive measure in the early detection of disease. In fact, driving simulators have been shown to differentiate AD and non-AD individuals [10], [11], [12]. Driving simulation studies are advantageous in that they can be controlled more efficiently than standard road tests. They can also offer a comprehensive view of several critical cognitive functions that are negatively impacted by AD, namely visuospatial attention and response, memory, and complex decision making [12]. However, additional studies are needed to more accurately define essential critical cognitive functions that can help stratify AD and non-AD patients and pave the way for using driving and driving simulators as a reliable diagnostic tool.

In this study, we investigate driving simulator data obtained from AD patients and controls to determine whether subjects can be successfully stratified into their corresponding cohorts. Specifically, we develop machine learning models to examine a large array of driving data and classify the patients into two groups (AD/control) and provide various performance metrics. More importantly, we identify key variables that help differentiate AD patients and controls and provide insights.

II. Methodology

A. Data Description

Driving data were collected on patients age 65 and up with early Alzheimer's Disease (AD) and age-matched control subjects via a mid-range driving simulator located at the University of Virginia. Participants were recruited from a database of diagnosed patients at a memory disorders clinic based on clinical judgement and standard of care. In addition, AD participants met core clinical criteria for probable AD based on the recommendations from the NIA and Alzheimer's Association work groups [13] and MMSE > 20. The control participants were individuals without active neurological, psychiatric, or ophthalmologic illnesses except for the need to wear corrective lenses. All subjects had to have a valid license, and either be actively driving or have ceased driving within less than one year of testing and must have also been fluent in English. Once recruited, a battery of neuropsychological tests was performed, and if they met criteria based upon the results of these tests, subjects who continued to meet inclusion criteria

[1]Ryan Blanchette and Anahita Khojandi are with the Department of Industrial and Systems Engineering, University of Tennessee-Knoxville. (corresponding author: khojandi@utk.edu)

[2]Daniel Cox is with the Department of Psychiatry and Neurobehavioral Sciences, University of Virginia School of Medicine

[3]Michael Oliver and Roberto Fernandez are with Pat Summitt Clinic, University of Tennessee Medical Center

underwent the off-road driving simulator evaluation. A total of 35 subjects (18 AD and 17 Control) participated in the study. In order to assess driver competence, 125 operational driving variables assessing participant visual, motor, cognitive, and executive function abilities were collected. These variables included both time-based and score-based measurements. Driving simulator results were available for 30 of the 35 participants in this study. The five without simulator results had demographic data available but were not used in this analysis. The 30 participants that were used were equally split between AD and Control.

The study used the Driver Guidance System (DGS-78), which uses a 210° field of a view. The simulator is designed to mimic the inside of a vehicle, including seatbelt, dashboard, steering wheel, mirrors, and all other usual controls [14].

B. Data Preprocessing

In order to perform the analysis, we used R programming [15], in particular, the tidyverse [16] and caret [17] packages. We used 85 of the 125 simulator variables, as the remaining 40 variables either contained missing values or were redundant. We note that if the sample was large, the missing data could be reliably imputed; however, because of the small sample size, we opted to remove these 40 variables. The only remaining missing value in the dataset of the 85 variables was a single participant's driving experience. We imputed this value using ordinary least squares (OLS) regression with age as the input variable and driving experience as the target (R^2 = 0.75).

C. Modeling Algorithms

We applied two different algorithms in this study. First, we used logistic regression (LR), which uses a logit link function, specifically,

$$Ln\left(\frac{P}{1-P}\right), \tag{1}$$

to transform a binary target variable.

The primary advantages of logistic regression include its simplicity of implementing and its ease of interpretation. However, logistic regression can suffer from issues of overfitting and poor performance on nonlinear data. Hence, we also used the random forest classifier (RF), which uses many randomly sampled decision trees with randomly chosen variables to make predictions [18]. RF is generally robust against overfitting and is particularly advantageous when dealing with a high ratio of variables to observations, as is the case in our driving simulator dataset. Also, using RF will also allow us to maintain the critical interpretability needed in this study, as it produces a variable importance ranking. The RF models were trained using 500 trees.

D. Variable Selection

When modeling with high dimensional data, it is generally beneficial to use a variable reduction technique to reduce the high ratio of variables to observations [19]. Since a potential value in this study is the explainability of the model, rather than using a transformation algorithm such as Principal Component Analysis (PCA) [19], we used recursive feature elimination. This method applies a series of models, such as random forests, and uses the results to rank variables based on their importance and eliminate the least important one, one at a time. The combination of variables that produced the highest performance was retained.

E. Evaluation Metrics

Each model was fit and evaluated using three commonly used classification metrics: Area under the receiver operating characteristics curve (AUC), sensitivity, and specificity. In addition, instead of using a traditional train/test split such as 80/20, each model was trained and tested using Leave-one-out cross-validation (LOOCV). This cross-validation is preferred when modeling a small sample as it uses all but one observation for model fitting and the one left-out observation for testing in each fold. The process continues until predictions are made for all observations for a total of n folds (n = number of observations). Finally, the results are aggregated over the n folds, and the three metrics are generated to evaluate the performance of the modeling technique.

In order to further verify the stability of each model, stratified bootstrapping was performed [20]. Twenty random balanced samples of size 20 (10 AD/10 control) were taken. For each sample, the models were fitted, cross-validated, and evaluated. Consequently, the mean and standard deviations of the performance metrics across the 20 samples were calculated.

III. RESULTS

Table I provides the mean and standard deviations for both AD and control groups of select variables. For each variable, Wilcoxon Rank Sum tests were performed to measure the difference in means of the AD and Control groups.

TABLE I. DESCRIPTIVE STATISTICS OF SELECT VARIABLES. AVERAGES AND STANDARD DEVIATIONS ARE PROVIDED.

Measurement	***AD***	***Control***
Demographics		
Age	74.93 (7.83)	74.20 (5.39)
Driving Experience	56.27 (6.57)	57.27 (6.53)
Sex	60% Female	40% Female
Dual Processing		
Potholes Avoided	0.12 (0.05)**	0.29 (0.11)
Response Inhibition		
Lane Incursion	0.00 (0.27)	0.00 (0.80)
Working Memory		
Signs Recalled	4.80 (3.53)**	13.30 (4.45)
Coordination		
Braking Duration	2.87 (1.92)*	4.80 (1.01)

*Significant at p-value < 0.01; ** Significant at p-value < 0.001

TABLE II. TOP TEN MOST IMPORTANT CONTRIBUTING VARIABLES TO THE STRATIFICATION OF SUBJECTS INTO AD PATIENTS AND CONTROLS

Measurement	***Description***	***Category***
AbPrSp	Absolute processing speed.	Divided/Selective Attention
AvgLdsDT	The total amount of time it takes to slow down in response to the lead car's brake lights being activated for 3 seconds.	Dual Processing
AvgLdsTRT	The average amount of time between reaction and movement to from the gas to the brake pedals that it takes to slow down in response to the lead car's brake lights being activated for 3 seconds.	Dual Processing
AvgRTB	Pothole reaction time.	Dual Processing
CrctRspn	Total amount of times an individual responded appropriately and accurately (e.g. turned to avoid a pothole and successfully avoided the pothole.)	Dual Processing
IntnMs	Measurement of no movement by participants and subsequent run over of potholes.	Dual Processing
Inttv	Inattention/errors in pothole avoidance and brake light tasks.	Dual Processing
PtAvd	Pothole avoidance steering average.	Dual Processing
InfctTrn	Individual made the appropriate action to avoid the pothole, but were unsuccessful in completely avoiding them.	Response Inhibition
SgnRcd	The number of signs that participants correctly recalled in order of presentation. Sign presentation stopped after three consecutive signs were missed.	Working Memory

A. Variable Selection and Model Performance

Ten variables were retained using recursive feature elimination to use for modeling. These ten variables were considered the "best" subset and were used in both logistic regression and random forest models; however, both models were also fitted with all variables to provide comparisons and insights. Table II gives descriptions of these ten "best" variables.

A total of four models (2 LR and 2 RF) were fitted to the data of 30 participants and cross-validated. Each of the two algorithms was once fitted with all 85 variables and once fitted with the ten best variables. Table III shows the mean and standard deviations of the AUC, sensitivity, and specificity of each of the four models.

TABLE III. PERFORMANCE OF FOUR MODELS EVALUATED USING LEAVE-ONE-OUT CROSS VALIDATION AND STRATIFIED BOOTSTRAPPING. AVERAGES AND STANDARD DEVIATIONS ARE PROVIDED

Model	***Number of Variables***	***AUC***	***Sensitivity (%)***	***Specificity (%)***
LR	All	0.55 (0.15)	53 (18)	52 (14)
LR	10 Best	0.85 (0.11)	75 (17)	90 (9)
RF	All	0.89 (0.05)	79 (6)	92 (4)
RF	**10 Best**	**0.96 (0.02)**	**87 (7)**	**93 (7)**

In terms of AUC, the RF model with the ten best variables performed the best with an average cross-validated AUC of 0.96. Fig. 1 shows the scaled importance of the ten variables used.

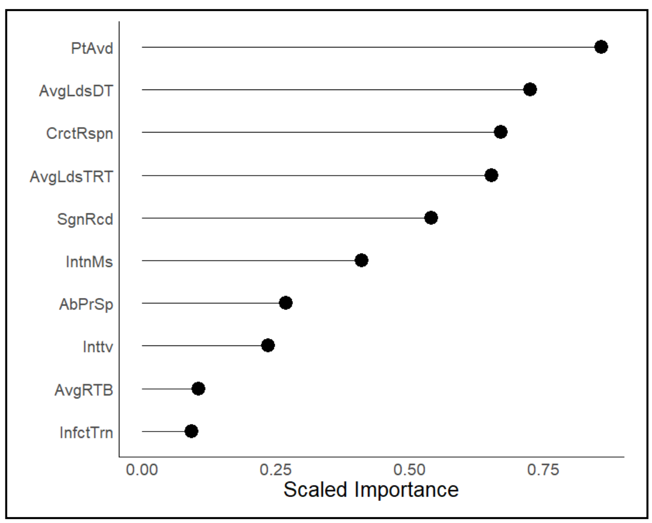

Figure 1. Scaled importance of the top ten most important contributing variables to the stratification of subjects into AD patients and controls

IV. DISCUSSION

As previously noted, driving simulation studies are effective in measuring several critical cognitive functions, including visuospatial attention and response, memory, and complex decision making [12]. The ten most important variables in our model were all under the category of dual processing, working memory, or coordination. The dual processing and coordination variables measured the participants' visuospatial attention and response, and the Signs Recalled variable measured their strength of memory.

Our results are further supported by a meta-analysis conducted on 27 primary studies that examined driving ability for adults with dementia [21]. The study noted that visuospatial processing is critical to safe driving and is at a significant deficit in those observed with AD, which aligns with the most important variable in our model, Pothole Avoidance Steering Average. Other studies have found that reaction times and visuospatial perception and processing are significantly affected in patients with AD [22], [23]. This aligns with the important variables in our model that measured reaction and detection times.

V. Conclusion

In this study, we demonstrated that machine learning algorithms, particularly random forest, could effectively classify subjects into their respective groups based upon characteristics of driving. By applying these algorithms, we can predict who is at risk of AD based upon performance on a driving simulator. These results underscore the utility of driving as a sensitive measure of cognitive function and support the need for further exploration of virtual reality driving as a diagnostic test in AD. Machine learning can play a critical role in determining which variables should be targeted as markers of impaired cognition in future research.

References

[1] 2017 Alzheimer's Disease Facts and Figures. *Alzheimers Dement*. 2017; 13: 325-73.

[2] Murphy SL, Xu JQ, Kochanek KD, Arias E. "Mortality in the United States, 2017." *NCHS Data Brief* no 328 (2018). Web.

[3] Centers for Disease Control and Prevention. (2019). *Alzheimer's Disease.*

[4] Alzheimer's Disease Fact Sheet. (2020). Retrieved from https://www.nia.nih.gov/health/alzheimers-disease-fact-sheet

[5] Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, Iwatsubo T, Jack Jr CR, Kaye J, Montine TJ, Park DC. "Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease." *Alzheimer's & dementia 7(3)* (2011): 280-292.

[6] Cummings JL, Morstorf T, Zhong K. "Alzheimer's disease drug-development pipeline: few candidates, frequent failures." *Alzheimer's research & therapy*. 6(4) (2014): 37.

[7] Borson S, Frank L, Bayley PJ, Boustani M, Dean M, Lin PJ, McCarten JR, Morris JC, Salmon DP, Schmitt FA, Stefanacci RG. "Improving dementia care: the role of screening and detection of cognitive impairment." *Alzheimer's & Dementia.* 9(2) (2013): 151-159.

[8] Bogdanovic N. "The challenges of diagnosis in Alzheimer's disease." *Eur Neurol Rev.* 14 (2018): 15-16.

[9] Roehr S, Riedel-Heller SG, Kaduszkiewicz H, Wagner M, Fuchs A, van der Leeden C, Wiese B, Werle J, Bickel H, König HH, Wolfsgruber S. "Is function in instrumental activities of daily living a useful feature in predicting Alzheimer's disease dementia in subjective cognitive decline?". *International Journal of Geriatric Psychiatry.* 34(1) (2019): 193-203,

[10] Stein, Anthony C, and Dubinsky, Richard M. "Driving Simulator Performance in Patients with Possible and Probable Alzheimer's Disease." *Annals of advances in automotive medicine. Association for the Advancement of Automotive Medicine. Annual Scientific Conference* 55 (2011): 325–34. Print.

[11] Stinchcombe, Arne et al. "Assessment of Drivers with Alzheimer's Disease in High Demand Driving Situations: Coping with Intersections in a Driving Simulator." *Geriatrics* 1.3 (2016): n. pag. Web.

[12] Tomioka, Hiroi et al. "Detection of Hypofrontality in Drivers with Alzheimer's Disease by Near-Infrared Spectroscopy." *Neurosci. Lett.* 451.3 (2009): 252–256. Web.

[13] McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & dementia: the journal of the Alzheimer's Association* 2011;7:263-269.

[14] Stephany M. Cox, Daniel J. Cox, Michael J. Kofler, Matthew A. Moncrief, Ronald J. Johnson, Ann E. Lambert, Sarah A. Cain, et al. "Driving Simulator Performance in Novice Drivers with Autism Spectrum Disorder: The Role of Executive Functions and Basic Motor Skills." *Journal of Autism and Developmental Disorders*. 46(4) (2016): 1379-1391.

[15] R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

[16] Hadley Wickham (2017). tidyverse: Easily Install and Load the 'Tidyverse'. R package version 1.2.1. https://CRAN.R-project.org/package=tidyverse.

[17] Max Kuhn. Contributions from Jed Wing, Steve Weston, Andre Williams, Chris Keefer, Allan Engelhardt, Tony Cooper, Zachary Mayer, Brenton Kenkel, the R Core Team, Michael Benesty, Reynald Lescarbeau, Andrew Ziem, Luca Scrucca, Yuan Tang, Can Candan and Tyler Hunt. (2019). caret: Classification and Regression Training. R package version 6.0-84. https://CRAN.R-project.org/package=caret.

[18] Breiman, Leo "Random Forests." (2001).

[19] James, Gareth. et al. *An Introduction to Statistical Learning with Applications in R.* New York, NY: Springer New York, 2013. Web.

[20] Kim, J. H. "Estimating Classification Error Rate: Repeated Cross-Validation, Repeated Hold-out and Bootstrap." *Comput. Statist. Data Anal.*, vol. 53, no. 11, (2009): 3735-3745.

[21] Reger, Mark A. et al. "The Relationship Between Neuropsychological Functioning and Driving Ability in Dementia: A Meta-Analysis." *Neuropsychology* 18.1 (2004): 85–93. Web.

[22] Kavcic, Voyko et al. "Neurophysiological and Perceptual Correlates of Navigational Impairment in Alzheimer's Disease." *Brain* 129.3 (2006): 736–746. Web.

[23] Fernandez, J., Roberto, Kavcic, J., Voyko, and Duffy, J., Charles. "Neurophysiologic Analyses of Low- and High-Level Visual Processing in Alzheimer Disease." *Neurology* 68.24 (2007): 2066–2076. Web.