Multimodal Predictive Analytics of Alzheimer's Disease Progression

David Senra Costa
Wentworth Institute of Technology

Abstract*—This electronic document is a "live" template and already defines the components of your paper [title, text, heads, etc.] in its style sheet. **CRITICAL: Do Not Use Symbols, Special Characters, Footnotes, or Math in Paper Title or Abstract. (*provide a short abstract*)

I. Introduction

Begin your introduction by clearly presenting your topic and explaining its significance - why it is important or interesting. Instead of listing questions separately, weave them together into a cohesive narrative that naturally connects the topic, its relevance, and its context. Provide an overview of existing research and key findings in this area, incorporating necessary citations to support your discussion. Your goal is to create a compelling introduction that sets the stage for your report.

Alzheimer's disease (AD) represents one of the most significant healthcare challenges of the 21st century, affecting millions worldwide and placing enormous burdens on patients, families, and healthcare systems. This research investigates the multifaceted nature of AD by analyzing the comprehensive OASIS-3 dataset [1], which contains psychometric assessments, clinical measures, demographic information, and volumetric brain data derived from neuroimaging.

The complex pathophysiology of Alzheimer's disease involves multiple interacting factors, necessitating a multimodal approach to understanding and predicting disease trajectory. While significant advances have been made in identifying biomarkers associated with AD, the integration of these biomarkers with cognitive assessments and demographic factors remains an evolving area of research. Recent studies have demonstrated that machine learning approaches can effectively leverage diverse data types to improve predictive accuracy for AD progression [2].

This study addresses key questions through a comprehensive analysis pipeline incorporating data cleaning, imputation techniques, and both supervised and unsupervised machine learning methods. By implementing Hierarchical clustering followed by Decision Trees (DT), Random Forest (RF), and logistic regression (LR) models, this research examines how cognitive decline patterns manifest across different demographic segments and investigates correlations between brain region volumes and cognitive performance. The findings have implications for early detection strategies, personalized treatment approaches, and addressing health disparities in AD care.

II. Datasets

A. Source of dataset

In this section, introduce your dataset by explaining its source—where you obtained it and whether it is from a credible provider. Include details such as when the dataset was generated and how it was created by its original author. If you generated the dataset yourself, describe the methods and processes you used.

OASIS-3 is the third comprehensive dataset included in the encompassing Open Access Series of Imaging Studies catalogue. The dataset - known officially as the *Longitudinal Multimodal Neuroimaging, Clinical, and Cognitive Dataset for Normal Aging and Alzheimer's Disease* - contains observations from 1,378 participants at various stages of cognitive decline. It was compiled retrospectively by the Washington University in Saint Louis School of Medicine Knight Alzheimer's Disease Research Center (KADRC) from assorted studies performed over the course of 30 years [1]. Published in 2019, its recent and extensive longitudinal nature provides great value in predictive and quantitative analysis of AD.

B. Characteristics of OASIS-3

The KADRC accumulated many different forms of data into comma separated value tables. The complete OASIS-3 dataset consists of 20 tables and approximately 24 MB of data (excluding raw imaging files which are multiple GB in size). For this analysis, we focused on a subset containing clinical, demographic, psychometric, and volumetric brain measurements (4 value tables).

The dataset structure includes longitudinal observations from 1,378 participants, with a total of Z visits recorded over time. The key variables used in our analysis are summarized in Table 1.

Table 1. Details regarding key variables involved in analysis.

Variable Name	*Variable Category*	*Data Type*	*Units/Range*	*Description*	*Table of Origin*
Total Hippocampal Volume	Intercranial Volumes	Continuous Numeric	mm^3		
Gender					
Age					

Data preprocessing included:

1. KNN-Imputation in missing fields of incomplete records

2. Normalization of volumetric measurements by total intracranial volume

The clinical data was merged with volumetric data using the participant ID and visit date fields. Additionally, new categorical variable "Cognitive Status" was created based on CDR scores, where CDR=0 was classified as "Normal," CDR=0.5 as "MCI" (Mild Cognitive Impairment), and CDR≥1 as "Dementia."

Describe the dataset's format and size. Additionally, provide an overview of the dataset's characteristics, including its features, size, structure, and any relevant attributes that are

979-8-3315-0993-4/25/$31.00 ©2025 IEEE

important for your analysis. Describe the dataset's format and size, as well as its key features, including the parameters, columns, rows, and character attributes along with their respective units. Using a table to present this information is recommended for clarity. Explain whether you cleaned the data or converted any units, specifying the formulas or rules applied. If multiple datasets were combined, describe how they were merged. Additionally, mention if you created any new categories for analysis, detailing what they are and how they were generated. Providing this background ensures transparency and helps readers understand the reliability and relevance of your data.

Example: XXXX

III. Methodology

In this part, you should give an introduction of the methods/model. First, what's the method/model. What's the assumption of this method/model. What's the advantage/disadvantage of this method/model. Why did you choose it. What Python module or function do you apply to apply this method/model. Any optional input/extra work did you adjust to make the results better. If you have multiple methods, feel free to use subsection A., B. to separate them.

Example: Before you begin to format your paper, first write and save the content as a separate text file. Complete all content and organizational editing before formatting. Please note sections A-D below for more information on proofreading, spelling and grammar.

A. Method A

Example: The equations are an exception to the prescribed specifications of this template. You will need to determine whether or not your equation should be typed using either the Times New Roman or the Symbol font (please no other font). To create multileveled equations, it may be necessary to treat the equation as a graphic and insert it into the text after your paper is styled.

$$a + b = \gamma \quad (1)$$

Note that the equation is centered using a center tab stop. Be sure that the symbols in your equation have been defined before or immediately following the equation. Use "(1)", not "Eq. (1)" or "equation (1)", except at the beginning of a sentence: "Equation (1) is . . ."

B. Method B

- Bulletin 1
- Bulletin 2.
- Bulletin 3

C. Method C

Example: The equations are an exception to the prescribed specifications of this template. You will need to determine whether or not your equation should be typed using either the Times New Roman or the Symbol font (please no other font). To create multileveled equations, it may be necessary to treat the equation as a graphic and insert it into the text after your paper is styled.

$$a + b = \gamma \quad (1)$$

Note that the equation is centered using a center tab stop. Be sure that the symbols in your equation have been defined before or immediately following the equation. Use "(1)", not "Eq. (1)" or "equation (1)", except at the beginning of a sentence: "Equation (1) is . . ."

An excellent style manual for science writers is [7].

IV. Results

In this section, present your findings using an appropriate method, such as equations, numerical summaries, or visualizations like charts and graphs. Clearly explain all results and provide guidance on how to interpret them. If any unexpected results arise, discuss possible reasons or contributing factors. To improve clarity and organization, consider using subsections (e.g., A, B) to separate different aspects of your results.

Example: After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention prescribed by your conference for the name of your paper. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper; use the scroll down window on the left of the MS Word Formatting toolbar.

A. Result A

Example: XXX

1) For papers with more than six authors: Add author names horizontally, moving to a third row if needed for more than 8 authors.

2) For papers with less than six authors: To change the default, adjust the template as follows.

a) Selection: Highlight all author and affiliation lines.

b) Change number of columns: Select the Columns icon from the MS Word Standard toolbar and then select the correct number of columns from the selection palette.

c) Deletion: Delete the author and affiliation lines for the extra authors.

B. Results B

Example: Headings, or heads, are organizational devices that guide the reader through your paper. There are two types: component heads and text heads.

C. Results C

a) Positioning Figures and Tables: Place figures and tables at the top and bottom of columns. Avoid placing them in the middle of columns. Large figures and tables may span across both columns. Figure captions should be below the figures; table heads should appear above the tables. Insert figures and tables after they are cited in the text. Use the abbreviation "Fig. 1", even at the beginning of a sentence.

TABLE I. Table Type Styles

Table Head	**Table Column Head**		
	Table column subhead	***Subhead***	***Subhead***
copy	More table copy[a]		

[a] Sample of a Table footnote. (*Table footnote*)

Fig. 1. Example of a figure caption. (*figure caption*)

Identify applicable funding agency here. If none, delete this text box.

Figure Labels: Use 8 point Times New Roman for Figure labels. Use words rather than symbols or abbreviations when writing Figure axis labels to avoid confusing the reader. As an example, write the quantity "Magnetization", or "Magnetization, M", not just "M". If including units in the label, present them within parentheses. Do not label axes only with units. In the example, write "Magnetization (A/m)" or "Magnetization {A[m(1)]}", not just "A/m". Do not label axes with a ratio of quantities and units. For example, write "Temperature (K)", not "Temperature/K".

V. Discussion

Every method/project has its shortage or weakness. Please discuss the unsatisfied results in your project. And discuss the feasible suggestions of future work to revise/improve your result.

Example: xxx

VI. Conclusion

In this part, you should summarize your project. What important results did you find for your topic and what's the effect of this result on the real-world?

Example: xxx

Acknowledgment *(Heading 5)*

The preferred spelling of the word "acknowledgment" in America is without an "e" after the "g". Avoid the stilted expression "one of us (R. B. G.) thanks ...". Instead, try "R. B. G. thanks...". Put sponsor acknowledgments in the unnumbered footnote on the first page.

We suggest that you use a text box to insert a graphic (which is ideally a 300 dpi TIFF or EPS file, with all fonts embedded) because, in an MSW document, this method is somewhat more stable than directly inserting a picture.

To have non-visible rules on your frame, use the MSWord "Format" pull-down menu, select Text Box > Colors and Lines to choose No Fill and No Line.

References

Use the IEEE format for the citation. The template will number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use "Ref. [3]" or "reference [3]" except at the beginning of a sentence: "Reference [3] was the first ..." Unless there are six authors or more give all authors' names; do not use "et al.". Papers that have not been published, even if they have been submitted for publication, should be cited as "unpublished" [4]. Papers that have been accepted for publication should be cited as "in press" [5]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

[1] OASIS-3: Longitudinal Neuroimaging, Clinical, and Cognitive Dataset for Normal Aging and Alzheimer Disease Pamela J LaMontagne, Tammie L.S. Benzinger, John C. Morris, Sarah Keefe, Russ Hornbeck, Chengjie Xiong, Elizabeth Grant, Jason Hassenstab, Krista Moulder, Andrei Vlassenko, Marcus E. Raichle, Carlos Cruchaga, Daniel Marcus, 2019. medRxiv. doi: 10.1101/2019.12.13.19014902

[2] M. Arnal Segura, G. Bini, A. Krithara, G. Paliouras, and G. Tartaglia, "Machine learning methods for classifying multiple sclerosis and alzheimer's disease using Genomic Data," International Journal of Molecular Sciences, vol. 26, no. 5, Feb. 2025, doi: https://doi.org/10.3390/ijms26052085.

[3]