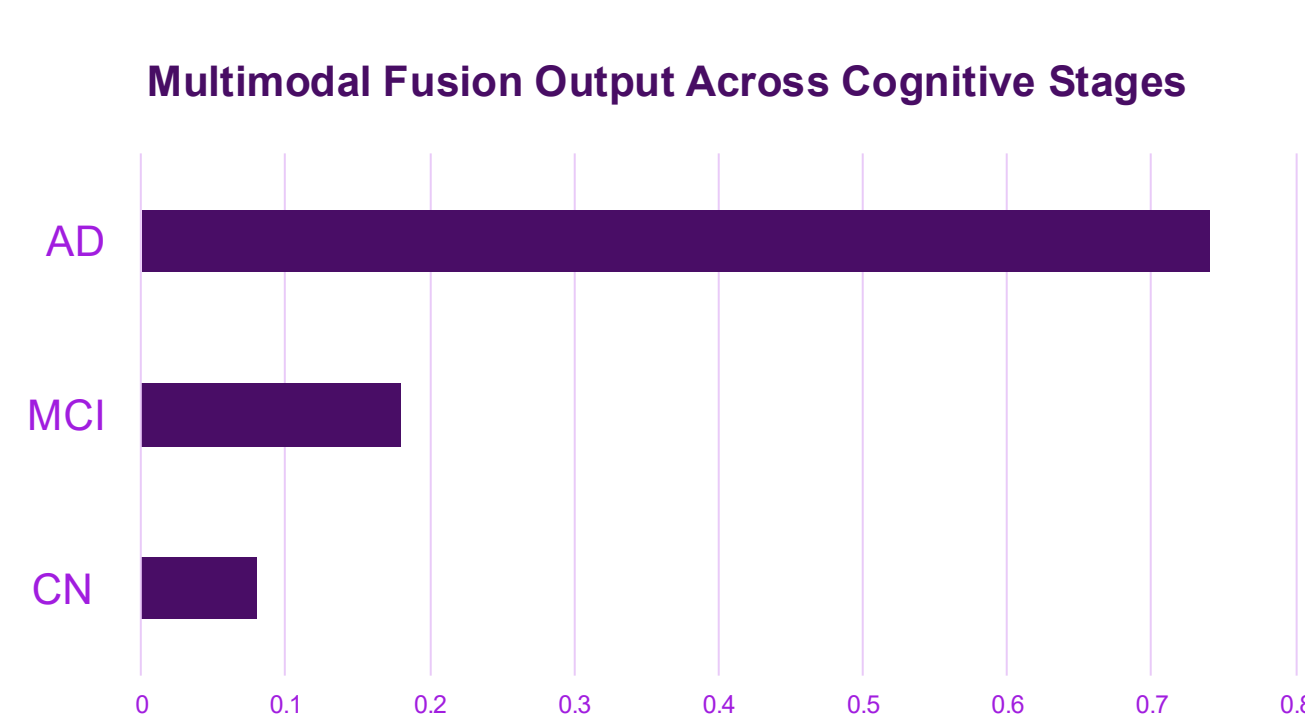


AI-Driven Multimodal Fusion of Speech, Neuropsychiatric Symptoms, and Cognitive Measures for Alzheimer’s Disease Risk Stratification

Clarina Xavier – Undergraduate Sophomore Student at Rutgers University, New Brunswick, New Jersey

- Finding a non-invasive method to diagnose Alzheimer’s Disease earlier and preventing substantial neurological damage for patients.
- Speech, neuropsychiatric symptoms, and cognitive assessments offer promising non-invasive indicators of disease progression which is vital in detecting Alzheimer’s earlier.
- Collecting data of speech features through auditory datasets, Neuropsychiatric Symptoms, and cognitive measures through the Neuropsychiatric Inventory Questionnaire(NPI-Q) on participants.
- Framework to classify patients as Cognitive Normality, Mild Cognitive Impairment, and Alzheimer’s Disease Stage.

INTRODUCTION	METHODS	EXAMPLE (FRAMEWORK PREDICTION)	CONCLUSION	REFERENCES																																																											
- Traditional Alzheimer's diagnosis relies on imaging, biomarkers, and laboratory testing. Diagnosis is often costly and occurs after significant neurological decline. - This study proposes a multimodal AI framework for classifying Cognitively Normal (CN), Mild Cognitive Impairment (MCI), and Alzheimer's Disease (AD) stages. Due to the lack of a single dataset containing all modalities, we can adapt 2 algorithms - Algorithm 1 independently estimates disease-stage probabilities from speech, neuropsychiatric, and cognitive features. - Algorithm 2 combines modality-specific probabilities through weighted fusion to generate a final stage prediction.	ALGORITHM 1: Each modality independently estimates the probability of Cognitively Normal (CN), Mild Cognitive Impairment (MCI), or Alzheimer's Disease (AD) using Random Forest classifiers. ALGORITHM 2: Modality-specific posterior class probabilities, are aggregated using a weighted probabilistic fusion scheme Weights are user-defined parameters that may be determined using prior clinical knowledge and model performance. $\sum_{m=1}^M w_m = 1$. Weighted Log Probability followed by normalization $\log \tilde{P}(Z) = \sum_{m=1}^M w_m \log p_m(Z)$ $P(Z S, N, C) = \frac{\exp(\log \tilde{P}(Z))}{\sum_{Z'} \exp(\log \tilde{P}(Z'))}$ Finding if patient has CN, MCI or AD based on maximum probability $\hat{Z} = \arg \max_Z P(Z S, N, C)$	<table><tr><th colspan="4">INPUT FOR ALGORITHM 1</th></tr><tr><th>MODALITY</th><th>SPEECH</th><th>NEUROPSYCHIATRIC</th><th>COGNITIVE FEATURES</th></tr><tr><td>1</td><td>Speech Rate</td><td>Anxiety</td><td>Round 1: Immediate and Delayed Recall of the East Boston Story Score</td></tr><tr><td>2</td><td>Pause Duration</td><td>Depression</td><td>Round 2: Immediate and Delayed Recall of the East Boston Story Score</td></tr><tr><td>3</td><td>Sentence Length</td><td>Apathy</td><td>Mini-Mental State Examination Score (MMSE)</td></tr><tr><td>4</td><td>Number of Filler words</td><td>Agitation</td><td>Symbol Digit Modalities Test Score (SDMT)</td></tr><tr><td>OUTPUT</td><td>$P_S=[P_S(CN), P_S(MCI), P_S(AD)]$</td><td>$P_N=[P_N(CN), P_N(MCI), P_N(AD)]$</td><td>$P_C= [P_C(CN), P_C(MCI), P_C(AD)]$</td></tr></table> ALGORITHM 2: Let’s consider sample weights based on literature <table><tr><th>Speech</th><th>NPI-Q</th><th>COGNITIVE</th></tr><tr><td>0.3</td><td>0.2</td><td>0.5</td></tr></table> <table><tr><th colspan="4">Independent Model Outputs</th><th>Alg – 2 Fusion</th></tr><tr><th>STAGE</th><th>SPEECH</th><th>NPI-Q</th><th>COGNITIVE</th><th>OUTPUT</th></tr><tr><td>CN</td><td>0.1</td><td>0.05</td><td>0.1</td><td>0.08</td></tr><tr><td>MCI</td><td>0.3</td><td>0.25</td><td>0.2</td><td>0.18</td></tr><tr><td>AD</td><td>0.6</td><td>0.70</td><td>0.7</td><td>0.74</td></tr></table> Multimodal Fusion Output Across Cognitive Stages 	INPUT FOR ALGORITHM 1				MODALITY	SPEECH	NEUROPSYCHIATRIC	COGNITIVE FEATURES	1	Speech Rate	Anxiety	Round 1: Immediate and Delayed Recall of the East Boston Story Score	2	Pause Duration	Depression	Round 2: Immediate and Delayed Recall of the East Boston Story Score	3	Sentence Length	Apathy	Mini-Mental State Examination Score (MMSE)	4	Number of Filler words	Agitation	Symbol Digit Modalities Test Score (SDMT)	OUTPUT	$P_S=[P_S(CN), P_S(MCI), P_S(AD)]$	$P_N=[P_N(CN), P_N(MCI), P_N(AD)]$	$P_C= [P_C(CN), P_C(MCI), P_C(AD)]$	Speech	NPI-Q	COGNITIVE	0.3	0.2	0.5	Independent Model Outputs				Alg – 2 Fusion	STAGE	SPEECH	NPI-Q	COGNITIVE	OUTPUT	CN	0.1	0.05	0.1	0.08	MCI	0.3	0.25	0.2	0.18	AD	0.6	0.70	0.7	0.74	The modalities selected here are significant, as past studies have indicated from population-based studies that up to 50% of individuals with mild cognitive impairment (MCI) and up to 97% of individuals with prodromal AD have at least one NPS, supporting its presence as a core feature of AD. Future framework must determine: - Optimal modality-specific weights using real-world validation datasets and investigate whether speech, neuropsychiatric, or cognitive features contribute differently across Alzheimer’s disease stages. - Feature engineering on comparing individual features for each modality and identify the most predictive combination for disease stage estimation. Validate the framework using real-world longitudinal datasets containing speech, neuropsychiatric, and cognitive assessments to evaluate predictive performance, robustness, and generalizability.	K. B. Rajan, D. A. Wilson, C. Weuve, et al.,“Cognitive Impairment 18 Years Before Clinical Diagnosis of Alzheimer Disease Dementia.”Neurology, vol. 85, no. 10, pp. 898– 904, 2015. M. Ruthirakuhan, D. X. Guan, M. Mortby, J. Gatchel, and G. M. Babulal, “Updates and Future Perspectives on Neuropsychiatric Symptoms in Alzheimer’s Disease.” Alzheimer’s & Dementia, vol. 21, no. 3, e70079, 2025.
INPUT FOR ALGORITHM 1																																																															
MODALITY	SPEECH	NEUROPSYCHIATRIC	COGNITIVE FEATURES																																																												
1	Speech Rate	Anxiety	Round 1: Immediate and Delayed Recall of the East Boston Story Score																																																												
2	Pause Duration	Depression	Round 2: Immediate and Delayed Recall of the East Boston Story Score																																																												
3	Sentence Length	Apathy	Mini-Mental State Examination Score (MMSE)																																																												
4	Number of Filler words	Agitation	Symbol Digit Modalities Test Score (SDMT)																																																												
OUTPUT	$P_S=[P_S(CN), P_S(MCI), P_S(AD)]$	$P_N=[P_N(CN), P_N(MCI), P_N(AD)]$	$P_C= [P_C(CN), P_C(MCI), P_C(AD)]$																																																												
Speech	NPI-Q	COGNITIVE																																																													
0.3	0.2	0.5																																																													
Independent Model Outputs				Alg – 2 Fusion																																																											
STAGE	SPEECH	NPI-Q	COGNITIVE	OUTPUT																																																											
CN	0.1	0.05	0.1	0.08																																																											
MCI	0.3	0.25	0.2	0.18																																																											
AD	0.6	0.70	0.7	0.74																																																											
				CONTACT																																																											
				clarina.xavier@rutgers.edu																																																											
																																																															
																																																															