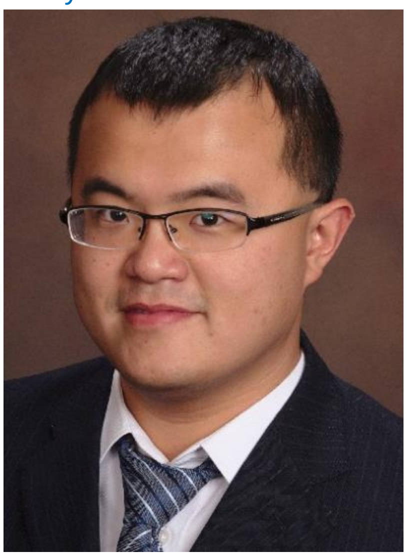
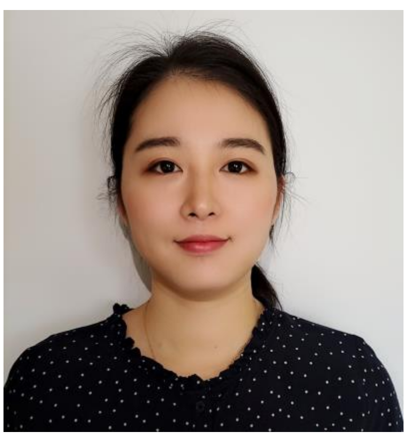
1. Please provide your preferred information for use in announcing the winners of the competition:
    a. Name (first and last name or first name and last initial):
        i. Wenyao Xu
        ii. Wei Bo
        iii. Harry Lin
        iv. Johnson Schwede

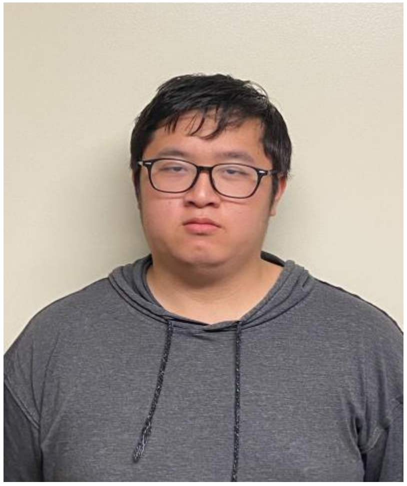
    b. Hometown: The team is from Buffalo, NY, USA.
    c. A recent picture of yourself or digital avatar (feel free to attach separately):
        i. Wenyao Xu



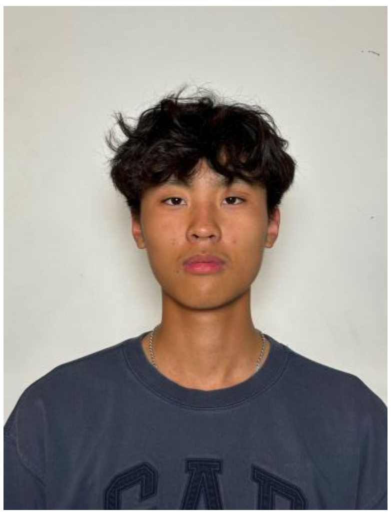
        ii. Wei Bo



        iii. Harry Lin



iv. Johnson Schwede



d. Social handle or URL (optional):

2. Who are you (mini-bio) and what do you do professionally?

Team Captain: Wenyao Xu

Dr. Xu is a faculty of the Computer Science and Engineering Department in the State University of New York (SUNY) at Buffalo, where he founded and directs the ESC (Embedded Sensing and Computing) Group. He has published over 200 technical papers, and is a named inventor on many international and U.S. patents.

He focuses on exploring novel IoT technologies to build up innovative Internet-of-Things (IoT) systems for high-impact real-world applications in the fields of healthcare, security and advanced manufacturing.

Team Leader / PI: Wei Bo

Wei is a third-year Ph.D. student of the Computer Science and Engineering Department in the State University of New York (SUNY) at Buffalo. She received an M.S. degree from Washington University in Saint Louis, US and B.S. degree from Shanghai University of Finance and Economics, China.

She is motivated and combines hands-on analytical experience with research interests on smart health, embedded/mobile sensing, and related intelligent analytics/computing. She has published multiple papers (on IEEE ICHI, JBHI, GSA, IEEE TMI, Elsevier SH) and is professional in how data and analysis inform practical applications to drive overall objective achievements in the engineering field.

Team Member: Harry Lin

Harry is a junior student of the Computer Science and Engineering Department in the State University of New York (SUNY) at Buffalo alongside an interest in pursuing a M.S. degree.

He is at the commencement of his career path and fueled by a resolute determination and a fervent eagerness to acquire new knowledge and embrace novel experiences, he has cultivated a keen interest in research. Beyond his academic pursuits, he dedicates his off-time to self-improvement, demonstrating a commitment to enhancing not only his scholarly abilities but also his overall personal development.

Team Member: Johnson Schwede

Johnson is a first-year undergraduate student majoring in computer science at Middlebury College. He received his high school degree from Nichols School located in Buffalo, New York.

He is a diligent and innovative student who is interested in conducting research in health-oriented subjects. He has previously conducted research in collaboration with the Department of Surgery at the State University of New York (SUNY) at Buffalo, Roswell Park Comprehensive Cancer Center, and Nichols School about the effects of antibiotics on long-term survival and recurrence free survival in non-small cell lung cancer lobectomy patients.

3. What motivated you to compete in this challenge?

   Our motivation to compete in this challenge stems from a deep commitment to advancing the early detection and management of AD/ADRD. As researchers in the fields of machine learning, healthcare, and data science, we recognize the critical need for innovative solutions that can address the growing global burden of these conditions. This challenge provides a unique opportunity to leverage our expertise in voice analysis and machine learning to contribute to the development of accessible, non-invasive diagnostic tools that could make a significant impact on patient outcomes, particularly in underserved communities.

   Furthermore, the challenge aligns with our research goals of developing inclusive and scalable technologies that can be applied across diverse populations. We are motivated by the potential to create language-agnostic models that can transcend linguistic and cultural barriers, ensuring that the benefits of early AD/ADRD detection are available to a broader audience. The chance to collaborate with other leading researchers and apply cutting-edge techniques to a problem of such global importance is both inspiring and aligned with our long-term vision of improving public health through technological innovation.

Finally, we are driven by the opportunity to contribute to a field where early detection can significantly alter the course of the disease, allowing for timely interventions that can improve the quality of life for millions of individuals. This challenge offers a platform to push the boundaries of what is possible in AD/ADRD research, and we are eager to be a part of this transformative effort.

4. High level summary of your dataset: the data source, target, predictors, sample size and use for early, inclusive prediction of AD/ADRD.

   Data Source: The dataset, known as the Global Vocal Biomarkers for AD/ADRD Recognition (GloVoAD), is derived from the DementiaBank (DB) collection, which includes a wide range of speech samples from both individuals diagnosed with AD/ADRD and cognitively healthy controls.

   Target Variable: The primary target variable is the cognitive status of the individuals, categorized into healthy controls, probable AD, AD, mild cognitive impairment (MCI), and primary progressive aphasia (PPA). This classification is determined through interviews, cognitive tests, and language-based tasks.

   Predictors: The key predictor variables are the acoustic features extracted from 30-second audio segments using the eGeMAPS V02 parameter set. These features include pitch, formants, speech rate, and other critical acoustic markers indicative of cognitive decline. Additional predictors include demographic information such as age, gender, and health status.

   Sample Size: The curated dataset comprises 2,058 participants who are diverse in demographic, geographic/language, and cognitive status, including 1,140 healthy controls and 918 individuals with varying stages of AD/ADRD (441 with probable AD, 170 with AD, 268 with MCI, and 39 with PPA).

   Use: This dataset is designed to support the development of language-agnostic machine learning models for the early and inclusive prediction of AD/ADRD, with a particular emphasis on improving detection accuracy across diverse demographic and geographic populations. The dataset is also tailored for use in telehealth platforms, making it more accessible to under-resourced communities.

5. What are two or three unique strengths of this dataset or type of data for early, inclusive prediction of AD/ADRD?

   Comprehensive and Integrated Data: The final GloVoAD dataset is obtained by structuring and unifying the diverse corpora to establish consistency across the board, and thoroughly organizing the demographic, geographic, and typical clinical data it contained. The dataset combines high-quality audio recordings with detailed structural data, including demographic information (e.g., age, gender, race) and clinical assessments (e.g., MMSE, MoCA scores). This integration of acoustic features with rich structural data allows for a more nuanced analysis, enabling researchers to account for various demographic and clinical factors that may influence the onset and progression of AD/ADRD. This comprehensive approach enhances the development of predictive models that are both accurate and fair, making them applicable to diverse populations.

   Support Language-Agnostic ML Models Development: The GloVoAD dataset leverages the established field of voice analysis and is designed to support the development of language-agnostic machine learning (ML) models, which enables the detection of

AD/ADRD-related cognitive decline regardless of the spoken language. This feature makes the dataset particularly valuable for developing machine learning models that are broadly applicable across different linguistic and cultural contexts, enhancing the inclusivity and generalizability of the predictive models. The proposed acoustic features, such as eGeMAPS V02 parameter set, include pitch, formants, speech rate, and other critical acoustic markers indicative of cognitive decline.

Focus on Early Detection: By capturing subtle changes in speech patterns, which are among the earliest indicators of cognitive decline, the dataset is uniquely positioned to support the early detection of AD/ADRD. Early identification of cognitive impairments can lead to timely interventions, potentially slowing the progression of the disease and improving patient outcomes, particularly in underserved communities where access to traditional diagnostic tools may be limited.

6. Did you use any tools or resources for developing your submission (e.g., to find a dataset, or explore the contents of a public dataset)?

   We utilized the DementiaBank (DB) dataset as the primary data source for our submission. This extensive collection of speech samples from individuals diagnosed with AD/ADRD and healthy controls provided the foundation for our analysis.

   We use Python to explore, process, and analyze the dataset. Python, with libraries such as pandas, numpy, librosa, and scikit-learn, was essential for extracting acoustic features, cleaning, and organizing the dataset.

   In addition, we utilized Pyannote and WhisperX for the initial and refined diarization of the audio data. Pyannote helped in identifying the primary speaker in each segment, while WhisperX ensured that the speech data was accurately segmented for analysis. Audacity was also employed to manually correct diarization errors and to ensure the quality of the audio recordings, which was crucial for maintaining the integrity of the data.

   Furthermore, we relied on publicly available documentation, research papers, and guidelines to align our methods with best practices in the field of voice analysis for AD/ADRD detection. This included literature on the use of the eGeMAPS V02 acoustic feature set and other relevant methodologies. These tools and resources were critical to the successful development of our submission.

7. Were there any data types or sources that you explored but didn’t fit for this challenge?

   Yes, during the development of our submission, we explored several data types and sources that ultimately did not fit the specific requirements of this challenge.

   Video Data: We considered incorporating video data that included facial expressions and gestures, which can also provide insights into cognitive decline. However, we decided to exclude video content because our focus was on developing language-agnostic models based on audio data. Video data would have introduced additional complexity and required significantly different analysis techniques that did not align with our primary objectives.

   Small and Incomplete Corpora: We evaluated some smaller corpora and those with incomplete demographic or cognitive data. These sources were ultimately excluded from the final dataset because they lacked the necessary completeness and consistency to contribute meaningfully to our analysis. Including them would have compromised the reliability and generalizability of the models we aimed to develop.

Highly Specialized Variables: Certain datasets included highly specialized variables, such as specific narrative tasks or unique cognitive assessments that were not widely used across other corpora. While these variables could offer detailed insights in specific contexts, they were not generalizable across the entire dataset. To maintain a more standardized and broadly applicable dataset, we chose not to include these specialized data types.

8. How would you improve or enrich this dataset if you had access to a big research team and an unlimited budget?

   Expansion of Demographic Diversity: We would expand the dataset to include a more globally diverse population, ensuring representation from various ethnic, cultural, and linguistic groups. This would involve collecting data from multiple countries and regions, particularly focusing on underrepresented populations in AD/ADRD research. Increasing the range of age groups and balancing gender representation would improve the dataset's ability to detect AD/ADRD across different demographics. Special attention would be given to populations at higher risk, such as older adults and women, who are disproportionately affected by these conditions.

   Longitudinal Data Collection: We would enhance the dataset by collecting longitudinal data, following participants over time to track the progression of cognitive decline. This would allow for the development of models that can predict not just the presence of AD/ADRD, but also its progression, enabling earlier and more personalized interventions. Increasing the frequency of assessments and data collection points would provide a richer temporal resolution, capturing subtle changes in speech and cognitive function over shorter periods. This would lead to more accurate detection of early-stage AD/ADRD.

   Expert-Led Annotations: With a large team, we could involve domain experts, such as neurologists and speech-language pathologists, in the annotation process, ensuring that the data is labeled with the highest accuracy. This would involve more detailed annotations of cognitive and linguistic features, improving the quality of the training data for machine learning models.

   Inclusion of Multimodal Data: We would incorporate multimodal data, including video recordings of participants during speech tasks. This would allow us to analyze facial expressions, gestures, and other non-verbal cues that are also indicative of cognitive decline. Integrating audio and video data would create a more comprehensive dataset, enabling the development of more sophisticated models that can detect AD/ADRD with higher accuracy. Adding neuroimaging data (e.g., MRI, PET scans) and other biological markers (e.g., genetic information, blood biomarkers) would provide a deeper understanding of the physiological changes associated with AD/ADRD. This integration could enhance the predictive power of models by correlating speech patterns with underlying neurological conditions.