AZCARE: A Novel Data-Centric AI Approach for Early Detection of Alzheimer's Disease With Multi-Modal Data and Multi-Model Ensemble

Li, Alexis (School: Hamilton High School)

Afflicting 45 million people worldwide, Alzheimer’s disease (AD) is a neurodegenerative disease that begins 20 years or more before symptoms manifest. Early detection is critical for intervention before it becomes irreversible. This study, AZCARE, proposes a novel Data-Centric AI for early detection of AD (DCAIAD) with a 3-layer prototype constructed in cloud. This contains a multi-modality data layer using the Oasis1 and 3 datasets that includes MRI, PUP, freeSurfer, psychology, and ADRC data with pipeline generated synthetic data, a processing pipeline layer to make standardized data using the reproducible pipeline clinica/clinicadl, and a AI Multi-Modeling Layer (comparing prediction performance in 12 ML Algorithms, 2 VotingClassifier ensembles and Brain3DCNN and other DL models). The results from 100+ experiments show the prediction accuracy on different datasets varied from 14.9 to 86.7% on ML models and 75.5 to 82% on DL models. Prediction performance also fluctuated according to data changes. For example, accuracy on the original freeSurfer data was 84%, but with oversampling the accuracy went up to 86%, undersampling 56%. By dealing with highly correlated fields and data leakage, ADRC accuracy went from 98% to 84%. The study suggests DCAIAD is better approach to addressing issues of data scarcity, neuroimaging data processing complexity, and bias. It helps in the detection of AD 5-10 years before the disease fully develops and make it easier and faster to build AI for AD applications. DCAIAD is a promising approach to building accountable AI, affordable AI, comparable AI, and accessible AI for early detection of AD.

Awards Won:
University of Arizona: Renewal Tuition Scholarship