A Machine Learning Approach to Identifying Blood-Based Biomarkers for Differential Diagnosis of Alzheimer's Disease

Sreenivas, Anjali (School: Nikola Tesla STEM High School)

Alzheimer’s Disease, described by several researchers as the public health crisis of the 21st century, is a devastating neurodegenerative disorder of gradual onset, clinically characterized by cognitive deterioration and memory loss. Despite its prevalence, there is yet to be a globally accessible and affordable method by which the disease can be differentially and timely diagnosed. This study seeks to investigate the potential of blood-based molecules as Alzheimer’s biomarkers through the interpretation of machine learning models that yield the highest predictive power within metabolomics and transcriptomics datasets aggregated from multiple studies. The Mann-Whitney U test with FDR corrections was used to determine the statistical significance of multiomics feature values in differentiating between the two experimental groups (cognitively healthy controls and Alzheimer’s patients), and gene ontology was performed on the top differentially expressed genes identified to determine biological pathways that appear to be altered in Alzheimer’s patients. While serum metabolites demonstrated little potential as effective Alzheimer’s biomarkers, the Random Forest machine learning architecture, when applied to the transcriptomics dataset, had an AUROC value of 0.713. Further analyses revealed that 78 genes were highly differentially expressed with p-values < 0.01 and > 90% increases in transcript abundance within the Alzheimer’s cohort. The identified genes are robust candidate biomarkers shared across multiple study sites that could one day make diagnosing Alzheimer's with a simple blood test a reality. Furthermore, the biological pathways identified corroborate the blood-brain barrier hypothesis and may be effective therapeutic targets in advancing the quest for a cure.

Awards Won:

Fourth Award of $500

Fourth Award of $500

American Statistical Association: In-Kind membership to ASA for all winners, including honorable mentions

American Statistical Association: Honorable Mention