Neuro-Behavioral Convergence: A Multimodal Approach to Stroke Risk Assessment

Yang, Janet (School: Canyon Crest Academy)

Stroke is a critical global health issue, and many cases are preventable if risk factors are identified early. Current risk assessment approaches often separate behavioral and neurophysiological analyses, and physiological screening remains less accessible. In this project, I performed a multimodal retrospective analysis integrating public health and electroencephalography (EEG) data to identify stroke-associated factors and translated these findings into an accessible risk-awareness tool. It could also support physician-guided education and prevention. First, I applied an interpretable machine learning (ML) approach to high-dimensional CDC Behavioral Risk Factor Surveillance System (BRFSS) data using LASSO-regularized logistic regression to identify key risk factors while reducing overfitting. Second, I analyzed a clinical EEG dataset using Elastic Net regression to determine whether brainwave features alone contain stroke-related signals after removing demographic and clinical confounders. Third, I developed a cross-platform application that displays personalized stroke risk and allows users to explore how lifestyle changes may impact their estimated risk. The BRFSS model demonstrated strong discrimination (AUC = 0.8052; sensitivity = 77.61%) with diabetes, smoking, general health, and mobility as the strongest contributors. The EEG model achieved an AUC of 0.8039 and 75.86% accuracy, indicating that neurophysiological patterns contain stroke-related information independent of behavioral data. Convergent performance across datasets suggests that behavioral and physiological signals capture complementary aspects of risk. The system was designed with a scalable, model-agnostic architecture. It can be extended to other neurological diseases.

Awards Won: