NeuraHealth: An Automated Screening Pipeline To Detect Undiagnosed Cognitive Impairment in Electronic Health Records Using Deep Learning and Natural Language Processing

Tyagi, Tanish (School: Phillips Exeter Academy)

Dementia-related Cognitive Impairment (CI) is a neurodegenerative disorder, affecting over 55 million people worldwide and growing rapidly at the rate of one new case every three seconds. The World Health Organization has established early detection as a principal goal for optimal management of CI. However, up to 90% of cases go undetected globally, costing USD 1.3 trillion annually and estimated to reach 2.8 trillion by 2030. With no cure, a recurring failure of clinical trials, and lack of early diagnosis, the mortality rate is 100%. Information in Electronic Health Records (EHR) provides vital clues for early diagnosis, but a manual review by experts is tedious and error-prone. This project presents an automated and scalable EHR screening pipeline, NeuraHealth, capable of diagnosing CI in EHR using deep learning and natural language processing. NeuraHealth consists of two novel components: a bi-directional attention-based model trained to diagnose CI and a two-step framework for disease diagnosis in EHR. Using patient EHR from Mass General Brigham, relevant sequences were extracted, classified, and used as features for a patient-level regularized logistic regression model. The diagnostic models outperformed all existing state-of-the-art computational and clinical methods by 15% due to enhanced understanding of linguistic context in complex language structures of EHR. Further, the models were integrated into a web application to make early diagnosis accessible in medical facilities, regions with scarce health services, and primary care centers. NeuraHealth has the potential to also be effectively applied to detect other chronic diseases in EHR like sleep apnea and diabetes.

Awards Won: