Capstone Two Project Proposal: Predicting Strokes

Problem Statement

Considering the limitations of current stroke risk assessment methods, can a dataset of personal choices achieve an accuracy of at least 75% in predicting stroke risk within a year?

Context

Stroke remains a significant health threat and early detection of potential risk factors is crucial. While some stroke risk factors are unchangeable, this research focuses on the potential of data-driven prediction based on personal choices. This research aims to develop a data-driven model that identifies modifiable behaviors associated with increased stroke risk.

Criteria for Success

Success for this project is the accuracy of at least 75% of the model in predicting stroke risk.

Scope of Solution Space

The initial focus is on identifying core modifiable behaviors with significant associations to stroke risk. Find the solution that analyzes large datasets containing personal choice alongside stroke incidence data.

Constraints

There's a limited amount of personal choice provided, such as, smoking habits, hypertension, work-type, marital status, glucose level, gender, and bmi.

Stakeholders

A few stakeholders in mind are healthcare providers with expertise in stroke risk assessment and patient demographics and public health officials who understand the public health needs and potential intervention strategies.

Data Sources

CSV File of datasets containing information on personal choices linked with stroke incidence data.