Project Proposal

Bobby Sokhi - Talha Ghumra - Cameron Jordan

Kwantlen Polytechnic University

MATH 4250

March 23, 2023

A Stroke occurs when blood stops going to any part of the brain resulting in brain cells being damaged. The effects of stroke can range from mild to severe and among those who survive their stroke, many never fully recover. Studies have shown that stroke is more prevalent in women than compared to men and can happen at any point in your life. Stroke is the leading cause of disability in adults and is the third leading cause of death in the world according to the World Health Organization.

In 2013, it was estimated that there were 405000 individuals experiencing the effects of stroke in Canada, yielding a prevelance of 1.15 %. This value is expected to increase to between 654000 and 726000 by 2038. One way to prevent or reduce the serious impact of stroke is to first know the risk factors that significantly affect it. To overcome these problems, it is necessary to conduct an analysis that aims to determine the significant risk factors that lead to having a stroke. This model can then be used to help predict a person's stroke likelihood.

We propose to conduct an analysis on a Stroke dataset obtained from the website Kaggle that contains clinical and personal variables of patients. In our analysis, we will be using 10 binary predictors to classify the presence of stroke in middle aged adults by implementing classificaton techniques i.e. 'Logistic', 'Decision Tree', 'Random Forest', 'SVM' and 'KNN'. The model that outputs the highest accuracy will then be used to develop an web application which can help users identify their own likelihood of having a stroke.