Overview

In this project, you will step into the shoes of an entry-level health data analyst at a leading health organization, helping to build and deploy a stroke prediction model to enhance clinical decision-making.

Project Scenario

A leading healthcare organization has noticed a trend in an increasing number of patients being diagnosed with strokes. To mitigate this growing problem, the organization has decided to launch a project aimed at predicting the likelihood of a patient getting a stroke based on a variety of health factors. The hospital has access to a vast amount of patient data, including medical history and demographic information, which can be used to build the predictive model.

Once the predictive model is validated and tested, the healthcare organization plans to integrate it into its clinical decision-making process. The model will be used to identify patients who are at high risk of getting a stroke and provide early intervention and prevention measures. Additionally, the model will be used to track the progress of high-risk patients and monitor the impact of preventive measures on reducing the incidence of stroke.

The success of this project will not only help the healthcare organization reduce the number of strokes in its patient population, but it will also position the organization as a leader in the use of advanced analytics and machine learning to improve patient outcomes. The predictive model will be a valuable tool for healthcare providers and patients alike, providing insight into their risk of getting a stroke and the steps they can take to prevent it.

Project Objectives

1. Explore the dataset to identify the most important patient and/or clinical characteristics.
2. Build a well-validated stroke prediction model for clinical use.
3. Deploy the model to enhance the organization's clinical decision-making.

Your Challenge

Your challenge will be to build a well-validated stroke prediction model for clinical use using patient characteristics. To do this, you will load, clean, process, analyze, and visualize data.

Then, you will build and deploy a prediction model using the cleaned and processed dataset.

In this project, we'll use we'll use data containing 11 clinical features for predicting stroke events.

After you perform your analysis, you will share your findings.