

Project Proposal: Stroke Prediction Analysis

Objective:

This project aims to analyze stroke prediction data to determine the likelihood of a person having a stroke based on various factors. By leveraging statistical analysis and machine learning techniques, we seek to uncover patterns and relationships within the dataset and develop predictive models.

Dataset Overview:

The dataset includes key factors that may influence stroke risk, such as:

- **Demographic Information:** Sex, age
- **Health Conditions:** Presence of heart disease, BMI (Body Mass Index)
- **Lifestyle Factors:** Smoking status
- **Other Features:** Additional attributes included in the dataset

Goals:

1. Identify the most significant factors contributing to stroke risk.
2. Develop a robust predictive model to estimate stroke likelihood.
3. Provide actionable insights to inform preventive measures and health interventions.

The source of the data is:

<https://www.kaggle.com/datasets/fedesoriano/stroke-prediction-dataset/data>