

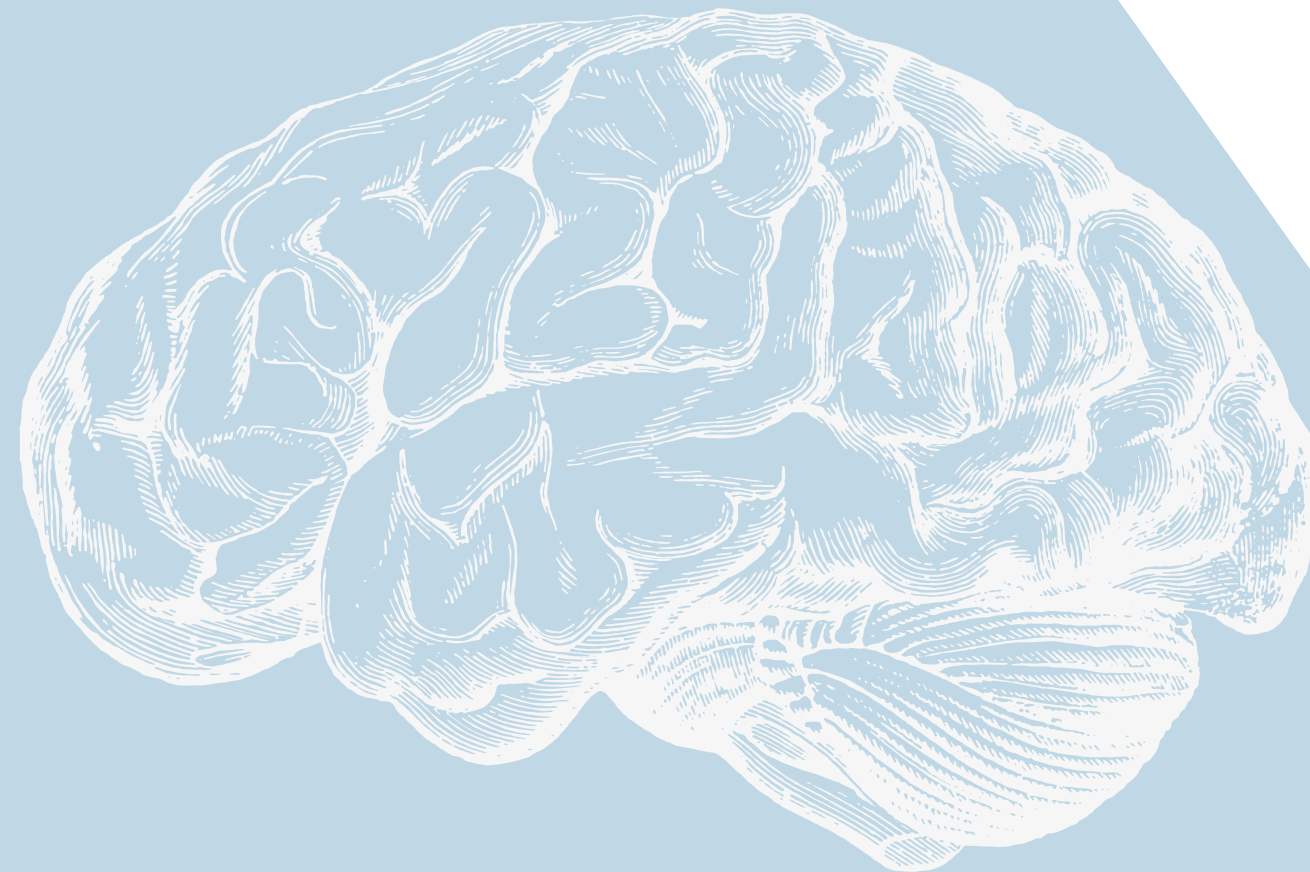
Proposal

STROKE

By:

REHAB ALSHEHRI

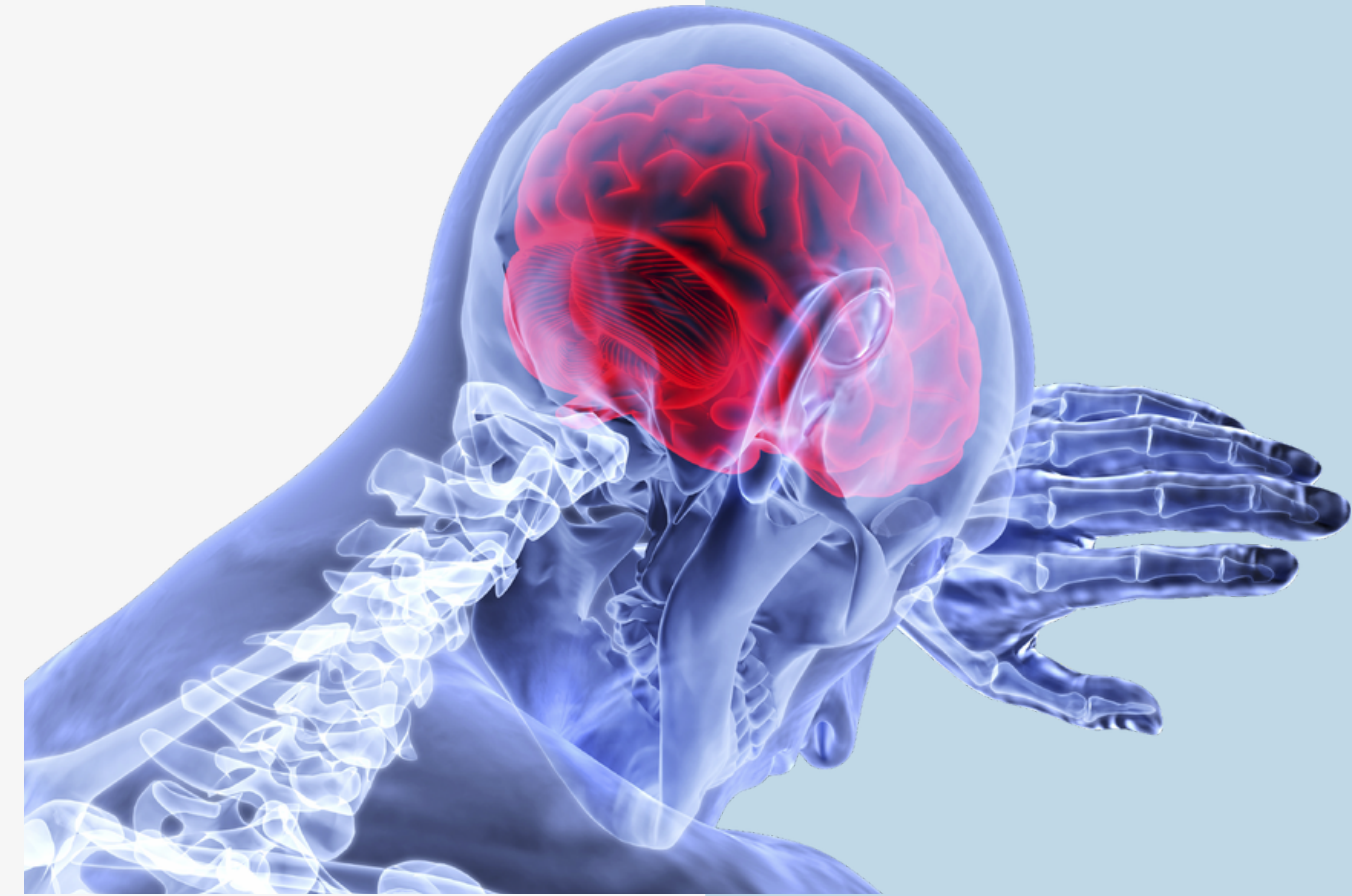
Email: RehabAlshehri9@gmail.com



Comparative study of ML model to predict Stroke

Project Motivation

in my project, We will analyze the most relevant risk factors for stroke as well as predict whether a patient is likely to have a stroke based on input parameters such as gender, age, various diseases and smoking status. **Using the Classification model**, stroke is, according to the World Health Organization (WHO), the second leading cause of death globally, and is responsible for approximately 11% of all deaths.



Question/need:



- Who is more likely to have a stroke then gender, male or female?, and what is the marital status?
- What is the effect of smoking status on the risk of stroke?
- What is the impact of heart patients on the risk of stroke?

Data Description:

Dataset source:

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

Dataset info:

Stroke data for the healthcare dataset from fedesoriano, 5110 Observations of 12 features.

Descriptions of the Dataset:

- id: unique identifier
- gender: "Male", "Female"
- age: age of the patient
- hypertension: 0 if the patient doesn't have hypertension, 1 if the patient has hypertension
- heart_disease: 0 if the patient doesn't have any heart diseases, 1 if the patient has a heart disease
- ever_married: "No" or "Yes"
- work_type: "children", "Govt_jov", "Never_worked", "Private" or "Self-employed"

Data Description:

- Residence_type: "Rural" or "Urban"
- avg_glucose_level: average glucose level in blood
- bmi: body mass index
- smoking_status: "formerly smoked", "never smoked", "smokes" or "Unknown"*
- stroke: 1 if the patient had a stroke or 0 if not

Note: "Unknown" in smoking_status means that the information is unavailable for this patient

- In modeling, The Classification model, Stroke My Target, based on data taken from patients, can help make decisions about lifestyle changes to reduce complications and save high-risk patients.



Tools:

I will conduct the experiment by using:

- **Environment:** Jupyter Notebook
- **Programming language:** Python
- **Libraries:** Pandas, Numpy, Matplotlib, Seaborn, sklearn