Introduction:

According to the World Health Organization stroke is the 2nd cause of death globally, this dataset is used to predict whether a patient is likely to get stroke based on the input parameters like gender, age, various diseases, and smoking .

Question:

is smocking ,weight ,gender ,age, heart disease ,work type will affect on stroke ?

Data Description:

This dataset is used to predict whether a patient is likely to get stroke based on the input parameters like gender, age, various diseases, and smoking status.

Data sources,

the dataset contain contains 5110 observations with 12 attributes.

1- id: unique identifier

2- gender: "Male", "Female" or "Other"

3- age: age of the patient

4- hypertension: 0 if the patient doesn't have hypertension, 1 if the patient has hypertension

5- heart\_disease: 0 if the patient doesn't have any heart diseases, 1 if the patient has a heart disease

6- ever\_married: "No" or "Yes"

7- work\_type: "children", "Govt\_jov", "Never\_worked", "Private" or "Self-employed"

8- Residence\_type: "Rural" or "Urban"

9- avg\_glucose\_level: average glucose level in blood

10- bmi: body mass index

11- smoking\_status: "formerly smoked", "never smoked", "smokes" or "Unknown"\*

12- stroke: 1 if the patient had a stroke or 0 if not

Tools:

The libraries that im going to use in this projecti numpy, pandas, .

MVP Goal:

the goal i will archive is to see what are the most affected to strike deases .

Ex:https://github.com/thisismetis/Metis\_Fundamentals/blob/main/project\_deliverable\_templates/project\_deliverable\_examples/regression/regression\_mvp\_ex.md