Project 2 - Part 5 (Core): Presentation

Source of the dataset:

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

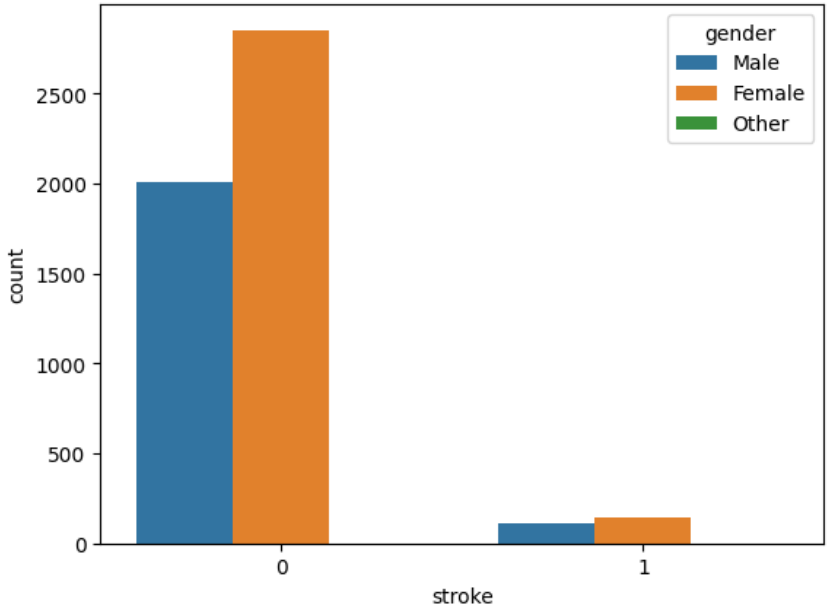
Who is your stakeholder, and what problem are you solving for them.

- Stakeholders are the stroke company, they are trying to understand the data they have gathered at their first glance. Data can be difficult to read especially when it is continuous.

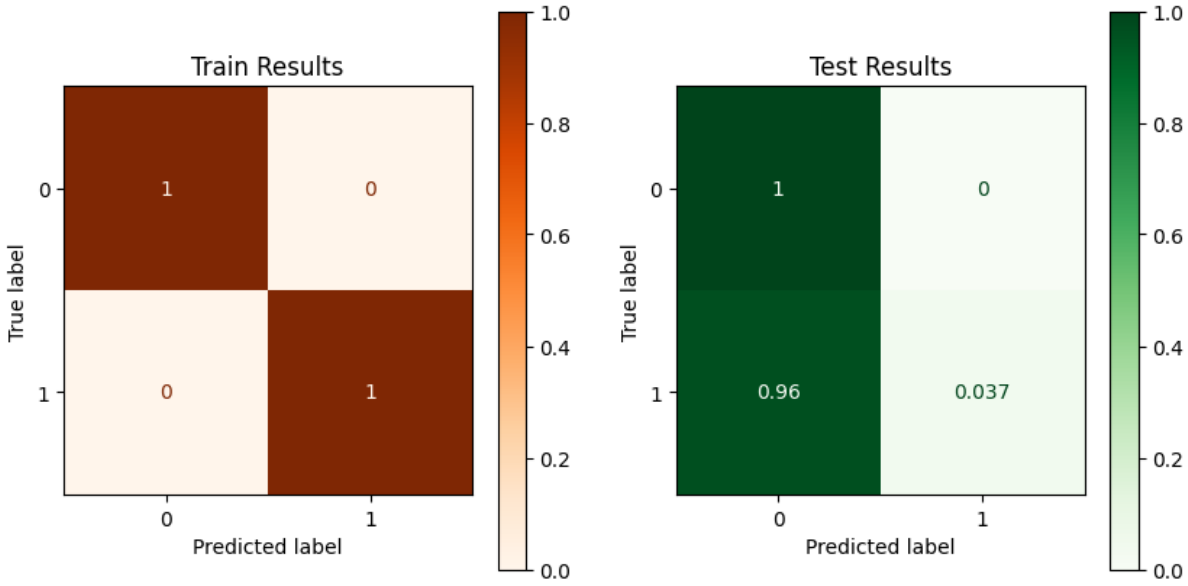
A brief introduction of your data

- The dataset is about people that have been living with stroke and those that have not, the data is giving us information about people's health results and whether or not their health might have been affected by the area that they are living in.

Two visuals that demonstrate key findings of interest to your stakeholder.



- The above graph explains a multivariate relationship between our target (stroke) and feature (gender)
- The count of people with stroke is less that than the ones without it. Males have the lowest number in both categories.



- The above graph indicates that the are no false positives in our dataset, this is good for our data as it would prevent the possibility of prescribing stroke medication to people who are stroke free.

A brief description of the strengths and limitations of your model for your stakeholder.

- This is a classification problem, our major goal would be deceasing false negatives as they tend to be more effective than any other. For this problem, it would be very important to decrease this number as people in need of treatment would not be discovered.
- I would recommend using the precision, accuracy and recall as the model to better the results of our data collection. This would then lead in no misinformation and all people would be attended to their specific needs.