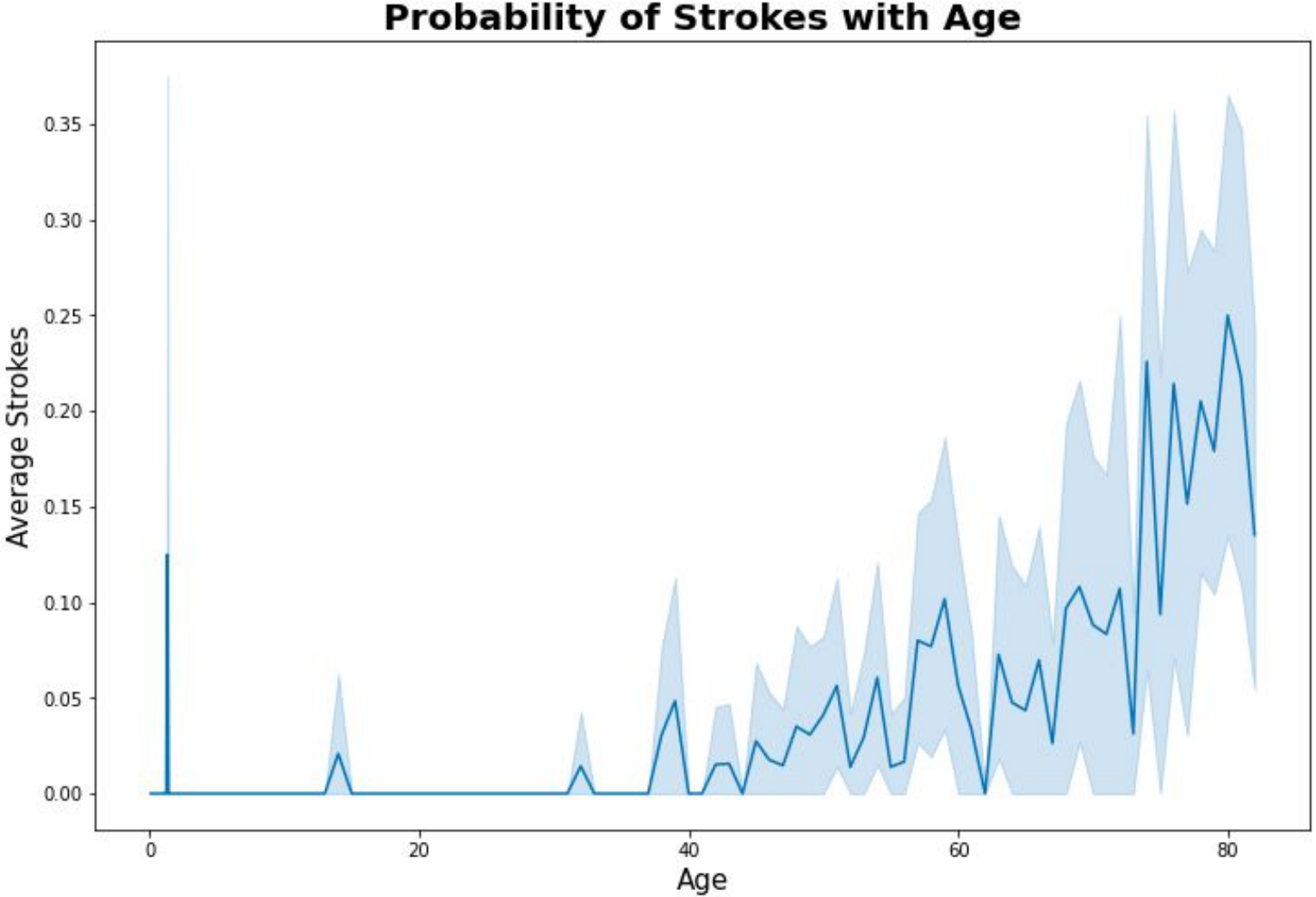
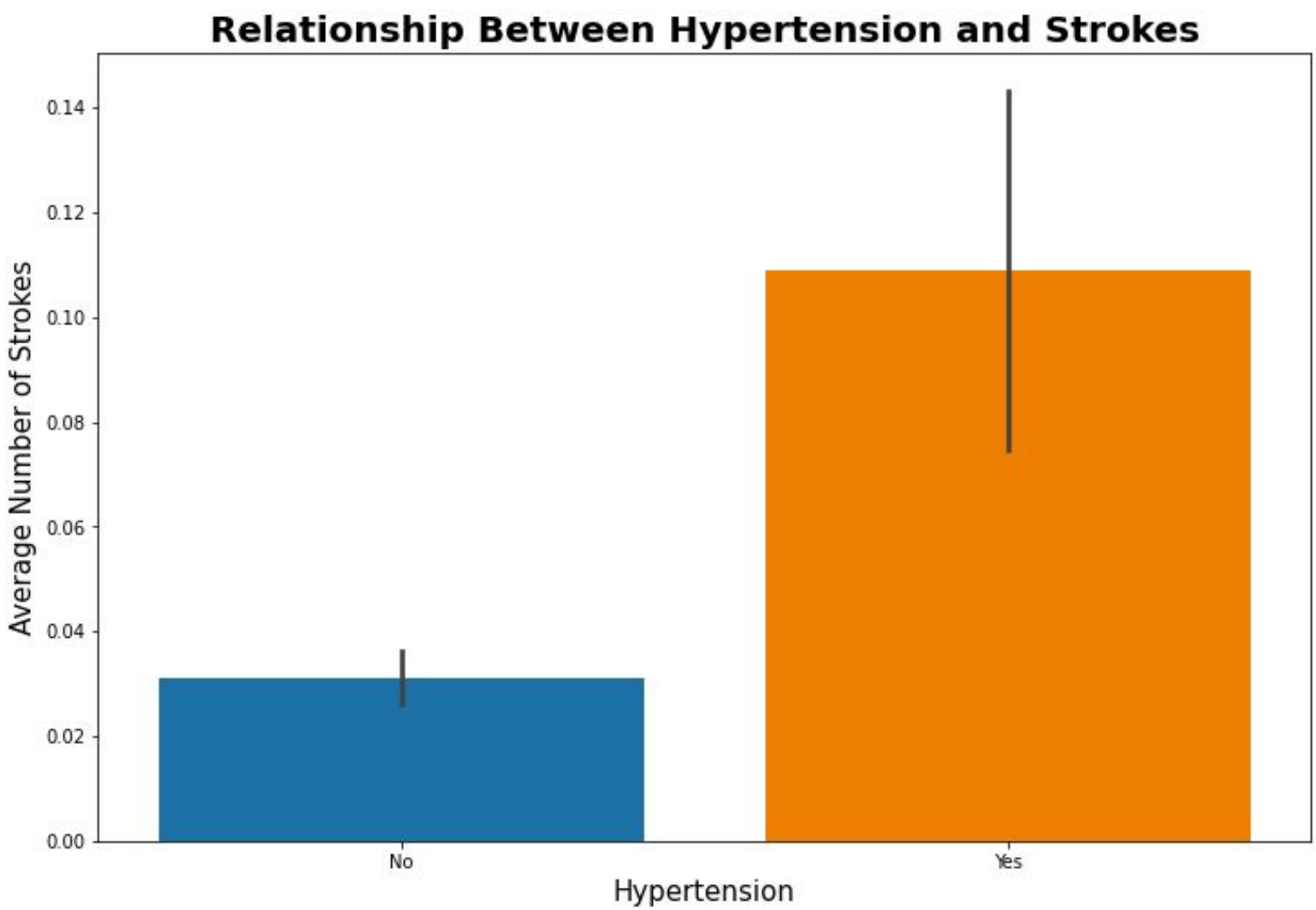
Stroke Prediction Model

Alex Re

About the Dataset

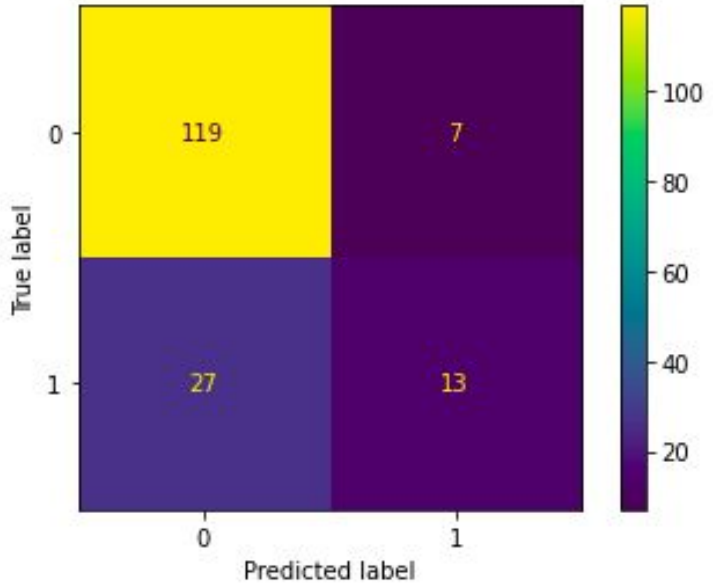
- Stakeholders: Doctors of patients at risk of having a stroke
- Purpose: Using characteristics of patients to predict if they will have a stroke
- Important Features:
  - Age
  - Smoker or Non-Smoker
  - Hypertension
  - Blood-Glucose Levels
  - Marital Status
  - Place of Residence
  - Area of Work


Probability of Strokes with Age
Average Strokes
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0
20
40
60
80
Age


Relationship Between Hypertension and Strokes
Average Number of Strokes
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
No
Yes
Hypertension

How did the model do?

- Did well with predicting true negatives
- Could only predict 32.5% of true positives correctly
  - This could be an issue when it comes to predicting strokes
  - The model will most likely predict that you won't have a stroke, even if you will
- Unbalanced dataset
  - There were many more people who didn't have strokes than those who did



Recommendations

- I would not recommend the use of this model without further improvement
- It will not accurately predict if a patient is likely to have a stroke, which could be deadly
- Moving forward, I will use techniques to balance the data so that the prediction is more accurate
- Doctors should closely monitor patients who have hypertension and/or are over 60 years of age

Any Questions?

Libraries That Were Used

- Scikit Learn
- Matplotlib
- Seaborn
- Pandas