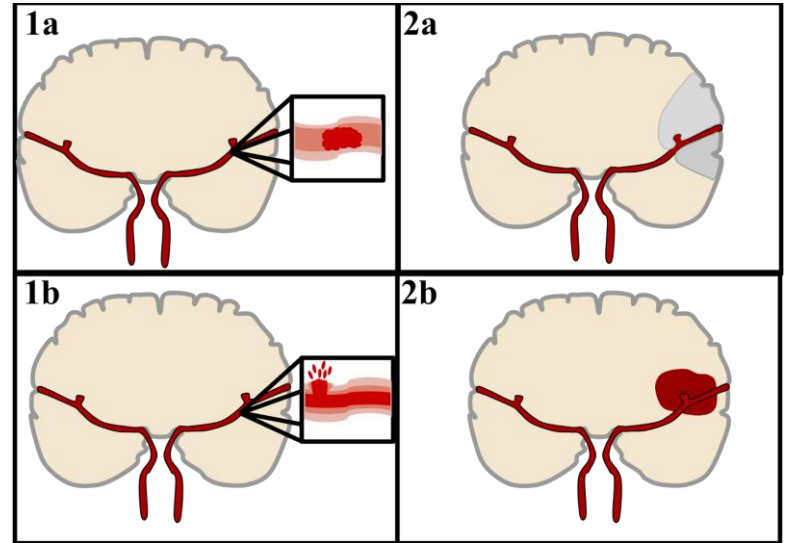


What feature is the best predictor of a stroke?

Beckett Schultz

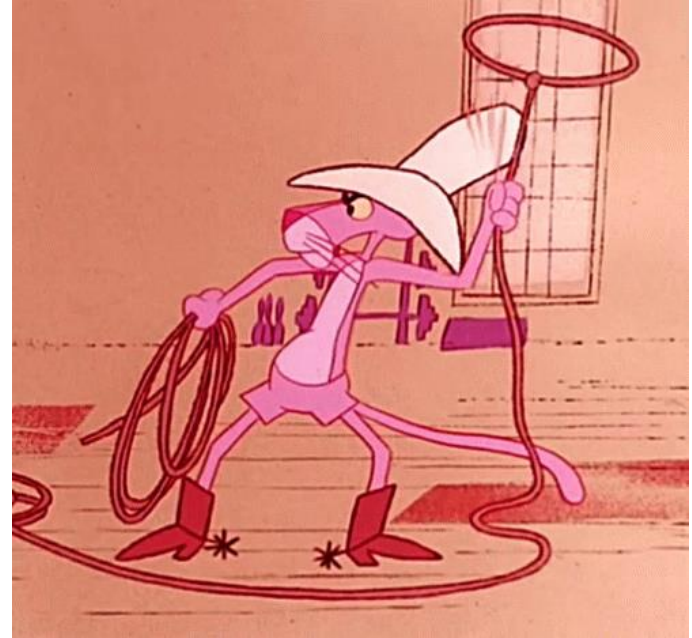
Introduction

- 2nd
- 11%
- Sample data with various inputs



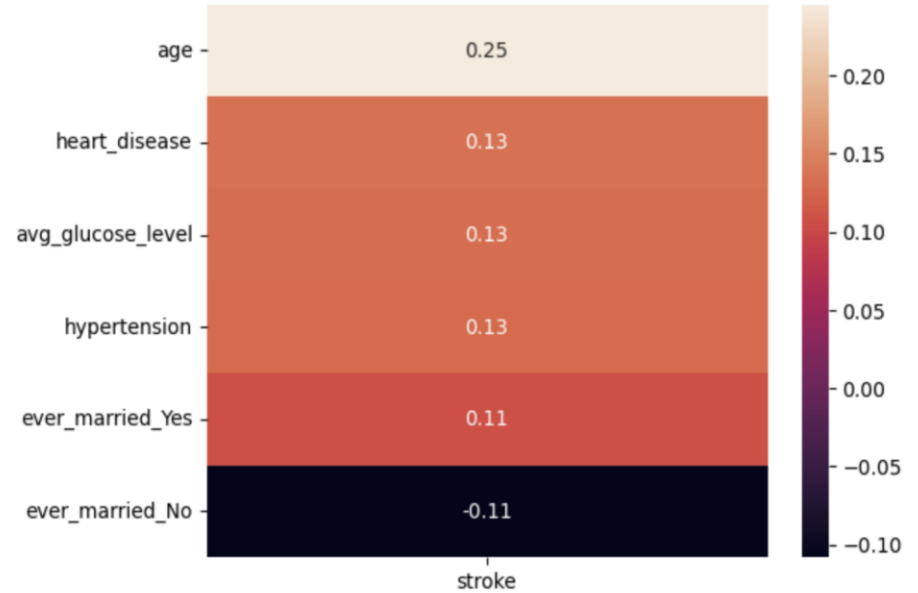
Data Wrangling

- Kaggle
- Initially, ...
-



* EDA *

- Histograms
- 5



Preprocessing

- Encoding
- Scaler
- Train | Test
- Storage



Model 1

- Logistic Regression
- ‘balanced’



Model 2

- Logistic Regression, ...
but better
- SMOTE



Other Models

- 3
- PyCaret



Hyperparameter Tuning

- GridSearchCV
- Randomized SearchCV
- PCA



Conclusions

- Champions:
Pipeline and Age
- Looking forward

