

STROKE PREDICTOR

It's a model to predict if the person is at risk of strokes or not depending on some factors like gender, age, and some other factors.



AGENDA

- **How the data is handled**
- **Training the model**
- **Application and website**



HOW THE DATA IS HANDLED

(PRE-PROCESSING – DATA CLEANING – EXPLANATORY DATA ANALYSIS)

ABOUT THE DATASET

For this project, we are using Cerebral Stroke Prediction-Imbalanced Dataset from Kaggle. The dataset consists of 12 features including the target column which is imbalanced. columns are:

- ID, Gender, Age
- Hypertension
- Heart disease
- Ever married
- Work type
- Residence type
- Avg glucose level
- Bmi
- Smoking status
- stroke



PRE-PROCESSING FOR THE DATA

At first, we split the data into two parts, one for training which is 80% of the data, and the other 20% for testing with a constant random state.

Then, explore the data to know about the loss and None values

After that, convert them into "Unknown" not to confuse the model while processing.



PRE-PROCESSING FOR THE DATA

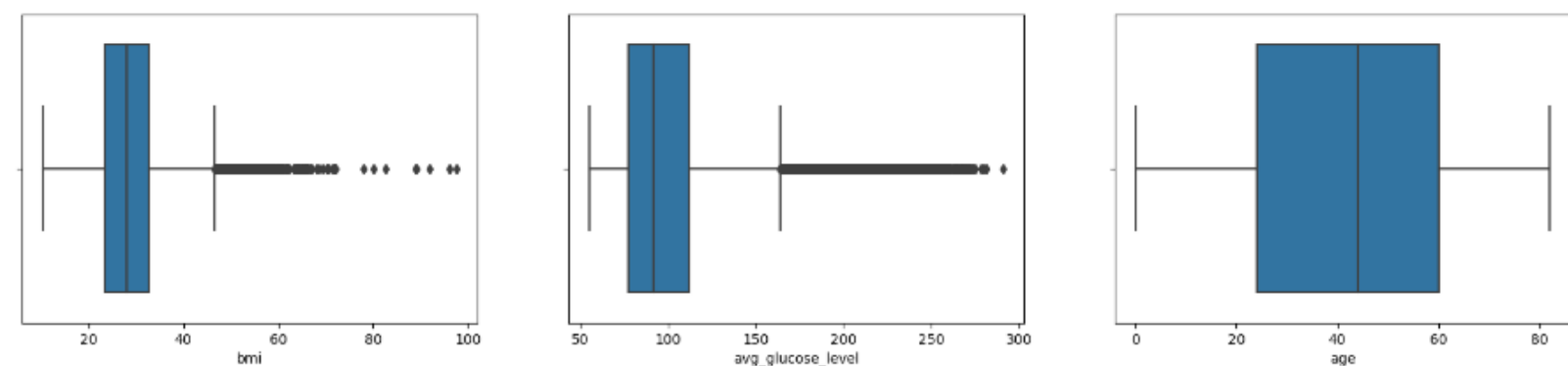
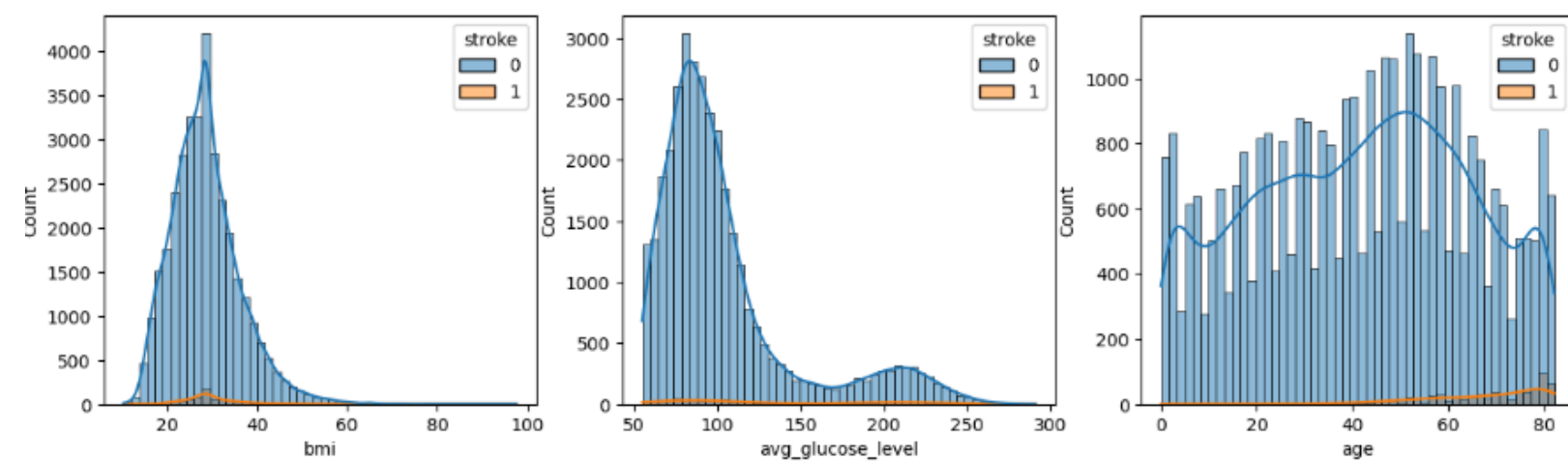


And then, the visualization phase helps to have a better imagination for the data and detect the outliers. and that helps to know more information about it like:

- marriage status for records
- location of living between urban and rural
- working type
- health status and if there is a huge number of them have diseases like hypertension, heart disease, ...etc
- age and gender
- percentage of persons that have strokes

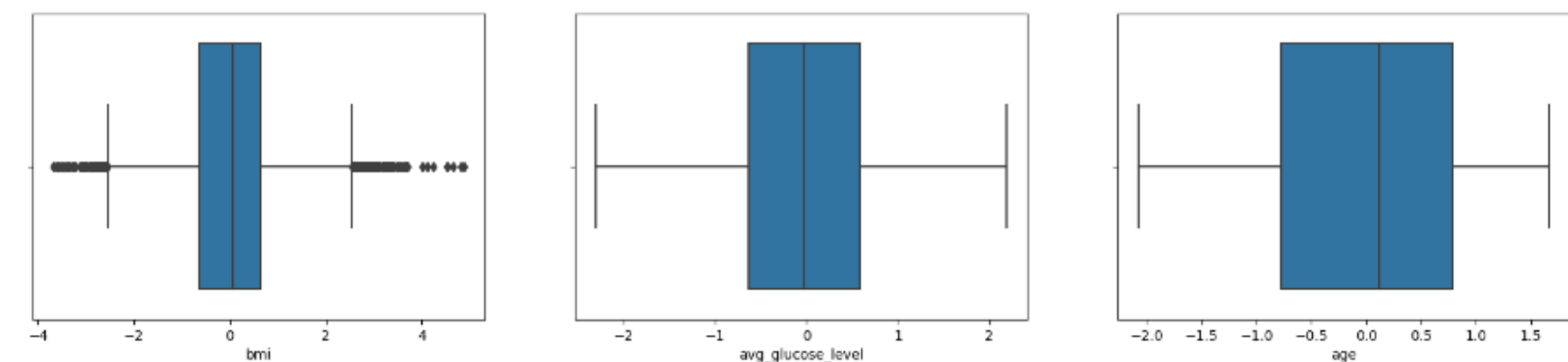
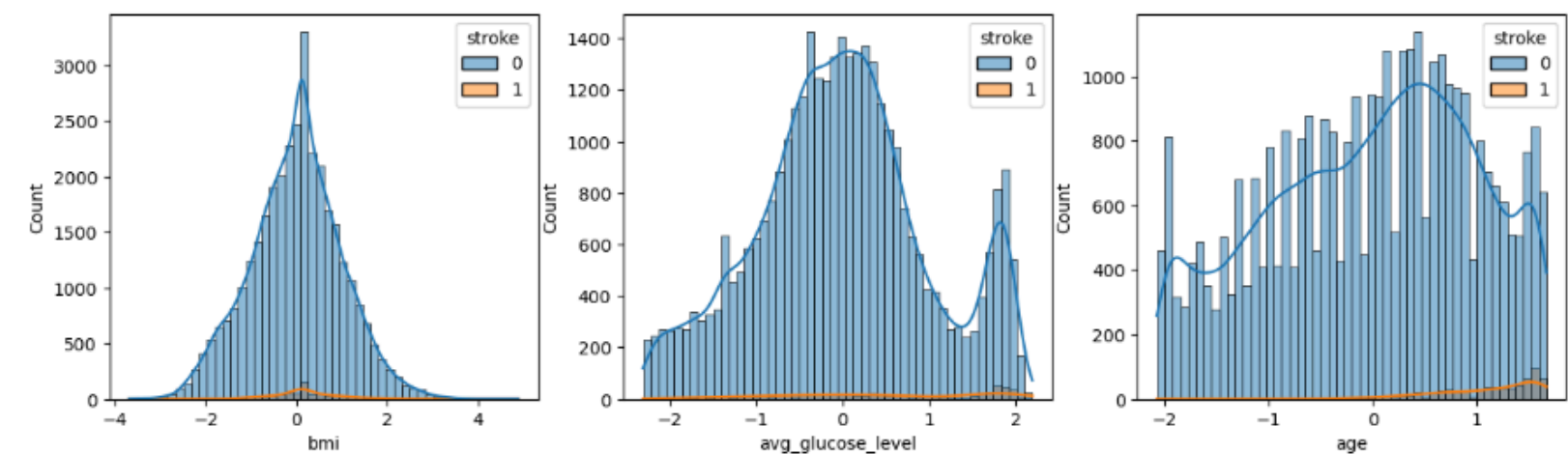
And according to these charts and plots, we can handle the data in a more proper way.

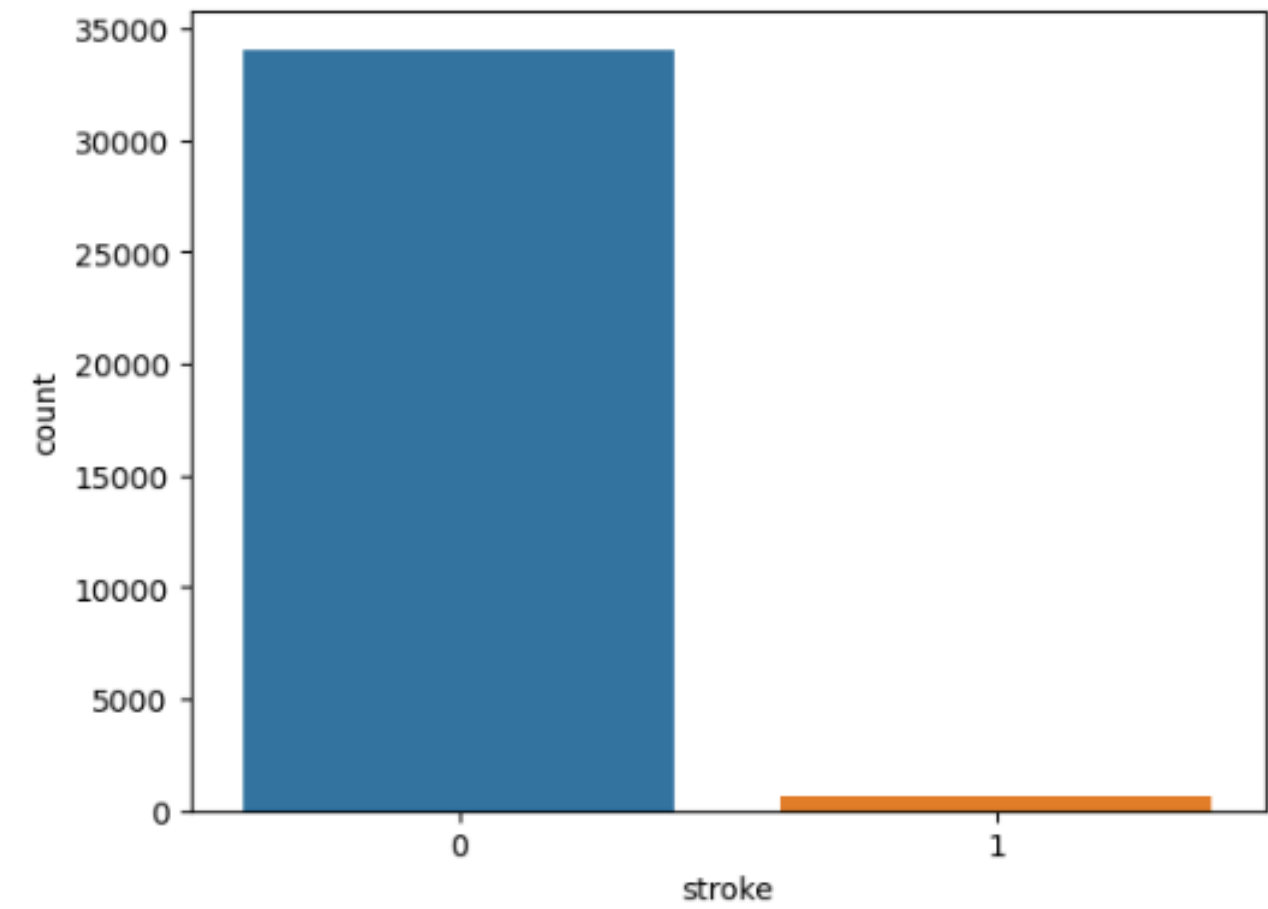
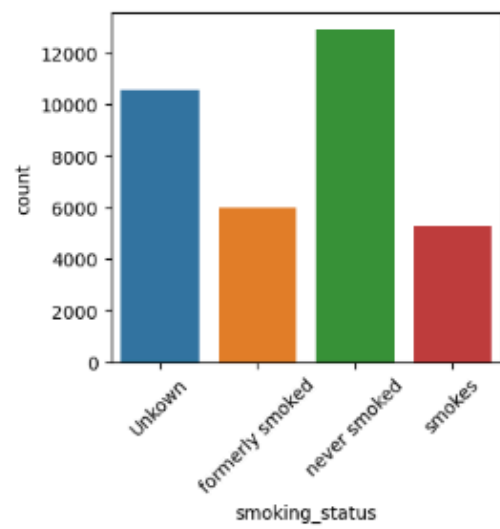
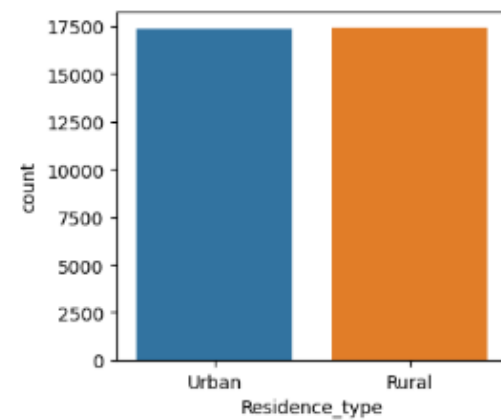
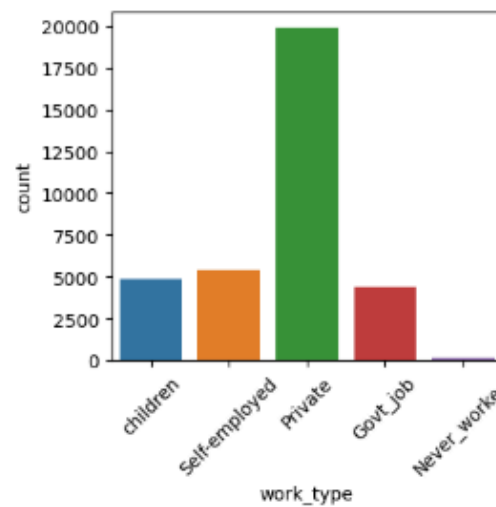
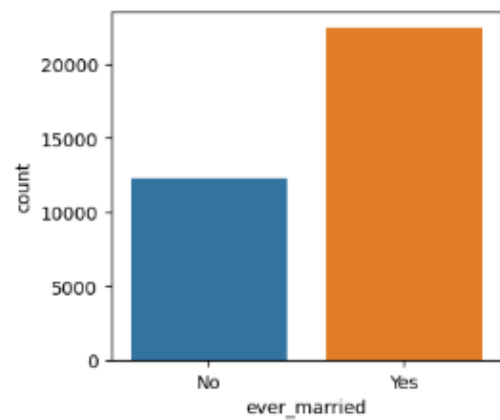
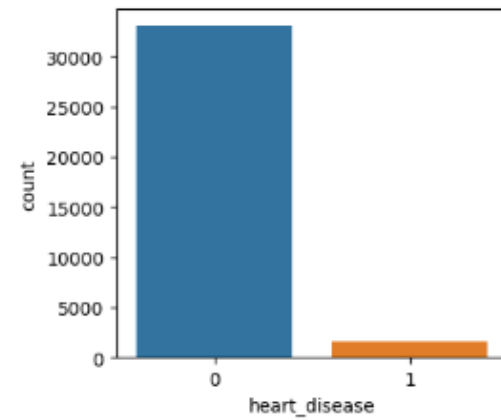
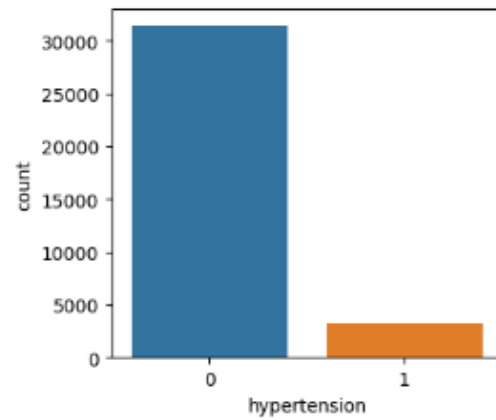
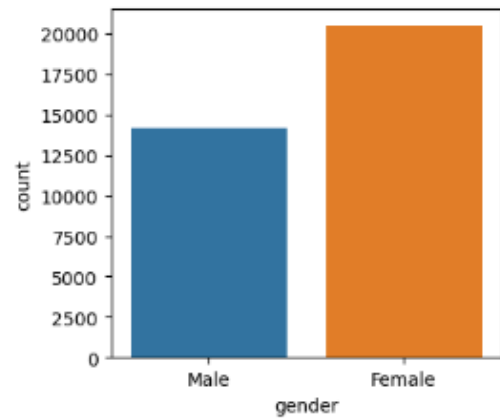




VISUALIZATION OF DATA BEFORE APPLYING THE POWER TRANSFORMER

VISUALIZATION OF DATA AFTER APPLYING THE POWER TRANSFORMER





TRAIN IN THE MODEL

(THE PHASE OF PREDICTION)



MODEL TRAINING



Several machine learning models were trained and evaluated to determine the most suitable for the task. After comparative analysis, it was found that the Logistic Regression and Gradient Boosting Classifier yielded the best performance. These models were then fine-tuned using Randomized Search CV to optimize their hyperparameters.

APPLICATION AND WEBSITE

(DEPLOYMENT THE MODEL ON AN APPLICATION AND WEBSITE)

THE WEBSITE USING REACT.JS



stroke clinic test About

Check if you are at risk of having a stroke

Your full name : Anas

Age : 50 🧑

Average Glucose level : 122

BMI : 48

Gender : Male 🧑

work type : private

Do you smoke ? never smoked

Living area : ☒ urban ☐ rural 🌻

Smokign not found ? ☐ YES ☒ NO

Do you have hyper Tension ? ☐

Do you have heart disease ? ☐

Do you ever Married ? ☐ 🧑

submit

© Powered by anas 2024

Ziad , You are at risk of having a stroke 🧑. It is best to go to the doctor

Be careful 🧑 : These results are based on artificial intelligence and **do not** replace the doctor in one way or another

New Test

© Powered by anas 2024

THE MOBILE APP USING FLUTTER



Stroke Classification

Age: 30.00

BMI: 20.00

Glucose Level: 50

Gender: ☐ Male ☒ Female

Residence Type: ☒ Urban ☐ Rural

Heart Disease: ☒ 0 ☐ 1

Hypertension: ☒ 0 ☐ 1

Ever Married: ☐ No ☒ Yes

Work Type: Private

Smoking Status: smokes

Classify Stroke

Result: Clear

Stroke Classification

Age: 45.83

BMI: 27.60

Glucose Level: 125

Gender: ☐ Male ☒ Female

Residence Type: ☒ Urban ☐ Rural

Heart Disease: ☒ 0 ☐ 1

Hypertension: ☒ 0 ☐ 1

Ever Married: ☐ No ☒ Yes

Work Type: Private

Smoking Status: smokes

Classify Stroke

Result: Potential Stroke

OUR TEAM



■ Ziad Mohamed Elsaied Elmimy

■ Saad Ahmed AbdElGhaffar

■ Mohamed Mahmoud Hamed Komeha

■ Anas mohamed al saieed abd al ghany

■ Mohamed Fathy Kotb

■ Abdulhaleem Osama Mohamed

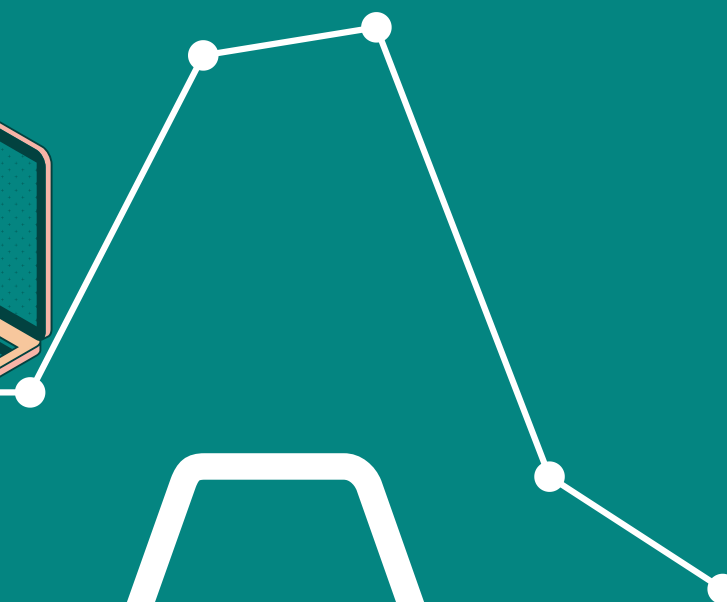
■ محمود دحيح فشيخ

■ Ahmed Mohamed Shetaia

Q



A



All is ok !!

THNAK
YOU

