

Stroke Risk Prediction

Team 4



OUR TEAM MEMBERS

Zixing
Li



Ting-Hsuan
Chen



Rohan
Gupta



Deniz Ipek
Ozakyol



Maria Stella
Vardanega

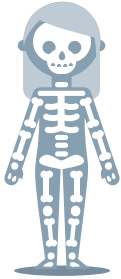




INTRODUCTION

Our goal is to analyze the major risk factors that may be associated with a stroke event.

OUR DATASET



Confidential Source

<https://www.kaggle.com/fedesoriano/stroke-prediction-dataset/metadata>



Columns

12



Rows

5111

The features of our dataset

	heart_disease	avg_glucose_level
gender	ever_married	bmi
age	work_type	smoking_status
hypertension	Residence_type	stroke



TABLE OF CONTENTS

01. Data Cleaning

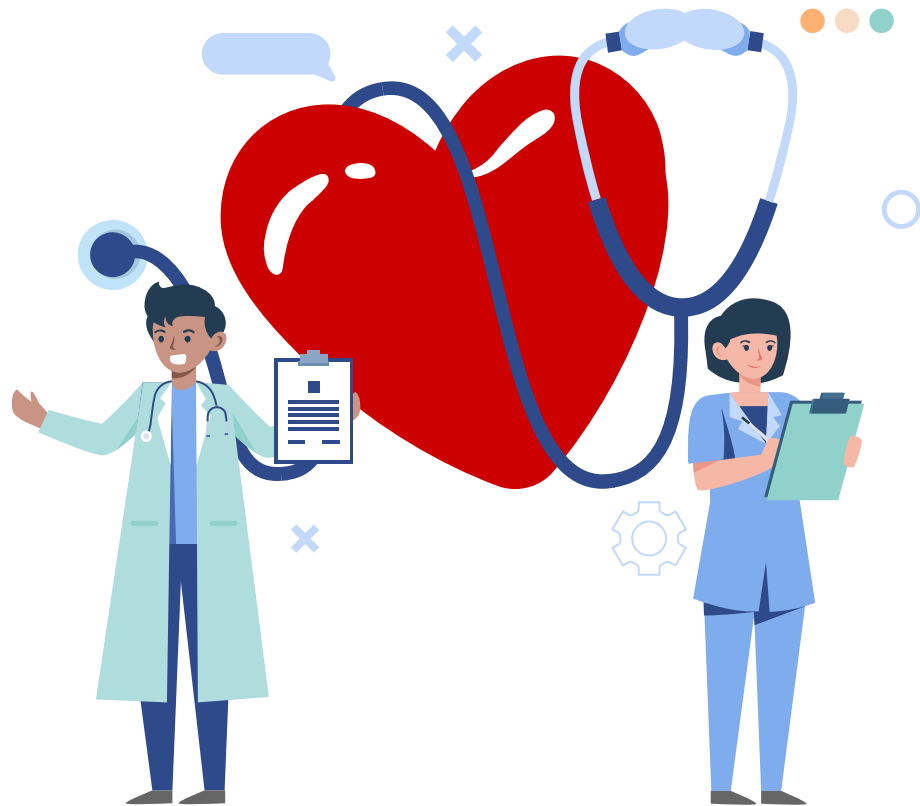
02. Data Exploration

03. Machine Learning

04. Conclusions

01.

Data Cleaning



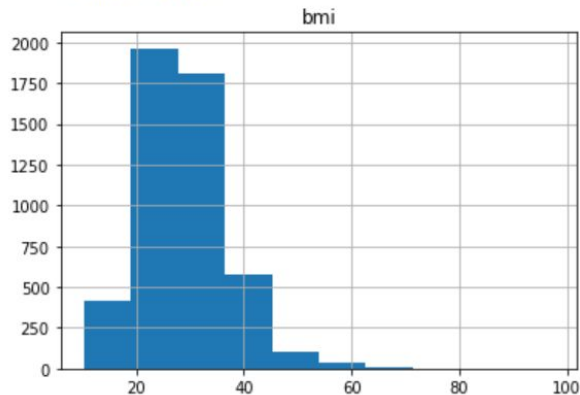
Dealing with missing values

```
1 stroke_data.isna().sum()
```

id	0
gender	0
age	0
hypertension	0
heart_disease	0
ever_married	0
work_type	0
Residence_type	0
avg_glucose_level	0
bmi	201
smoking_status	0
stroke	0
dtype: int64	

```
1 stroke_data.hist(column='bmi')
```

```
array([[<matplotlib.axes._subplots.AxesSubplot object at 0x7efe4d451d0>]],  
      dtype=object)
```



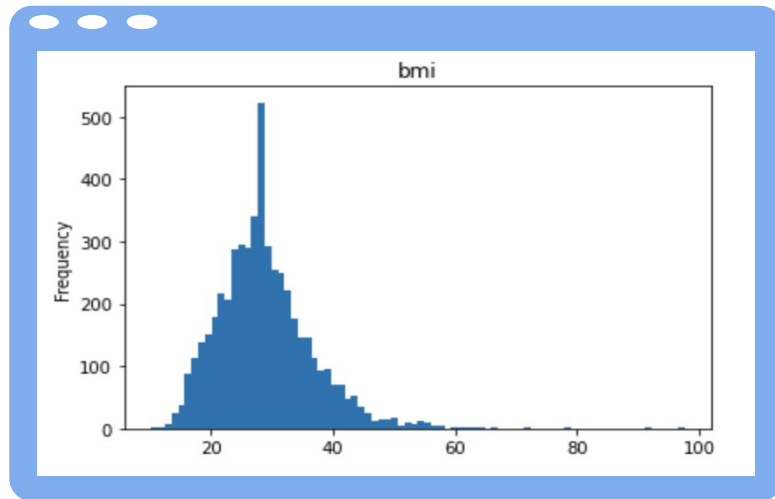
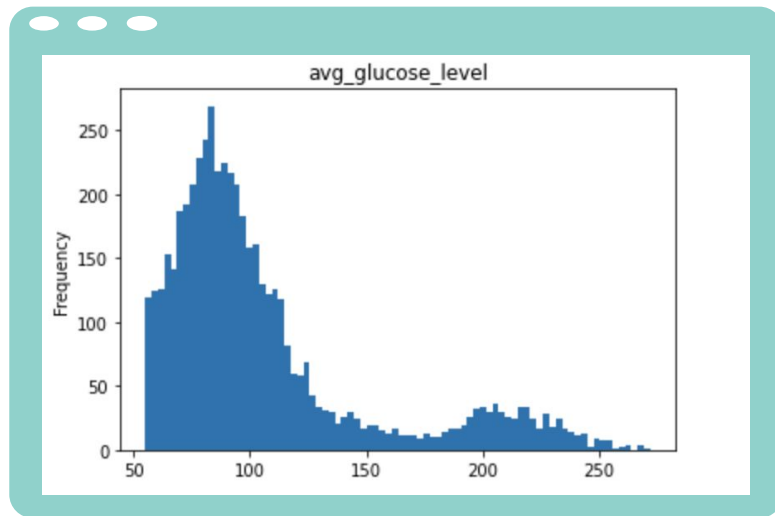
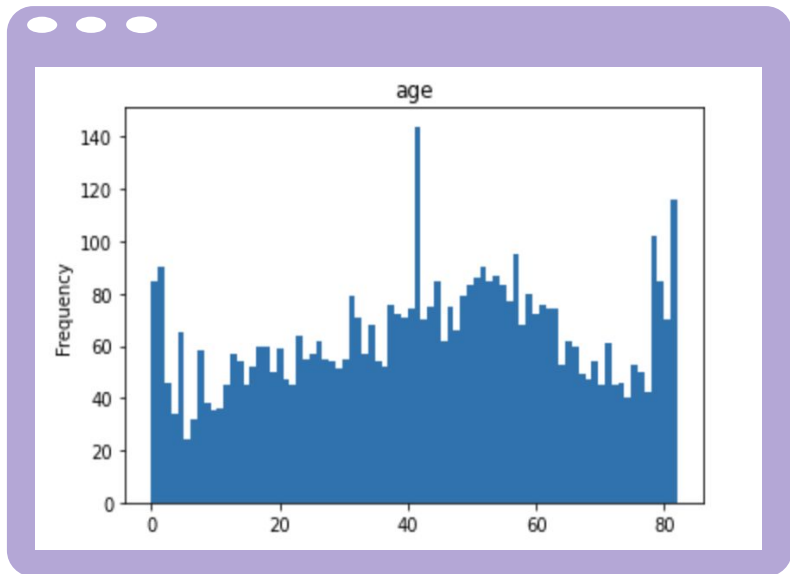
02.

Data Exploration



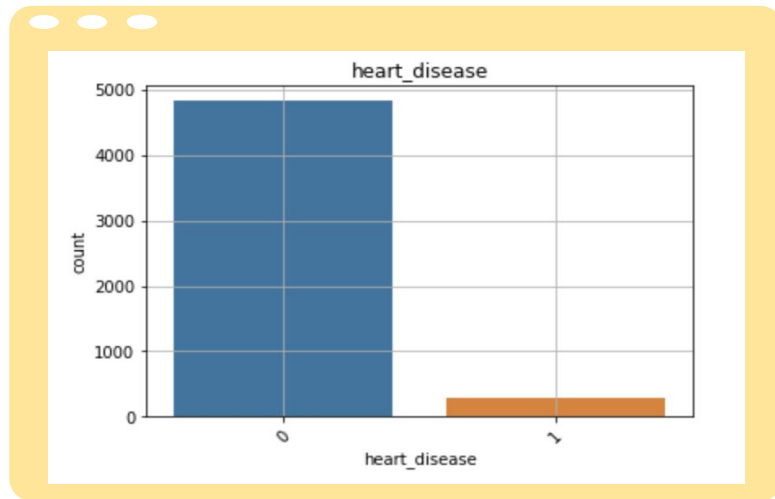
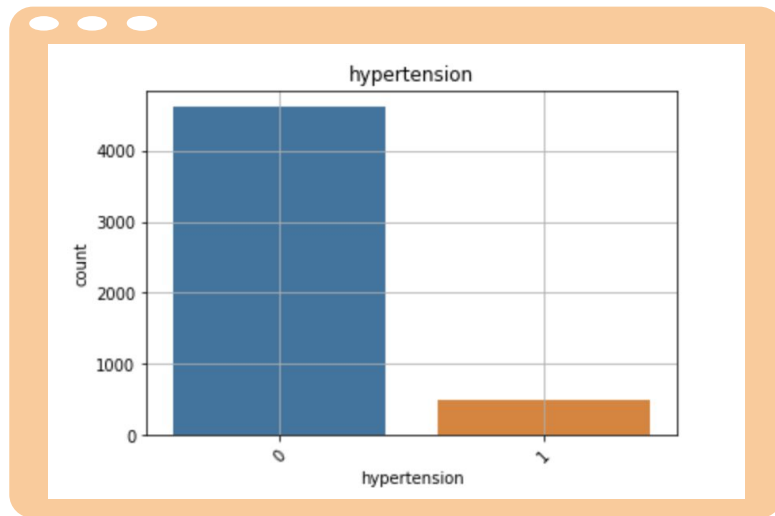
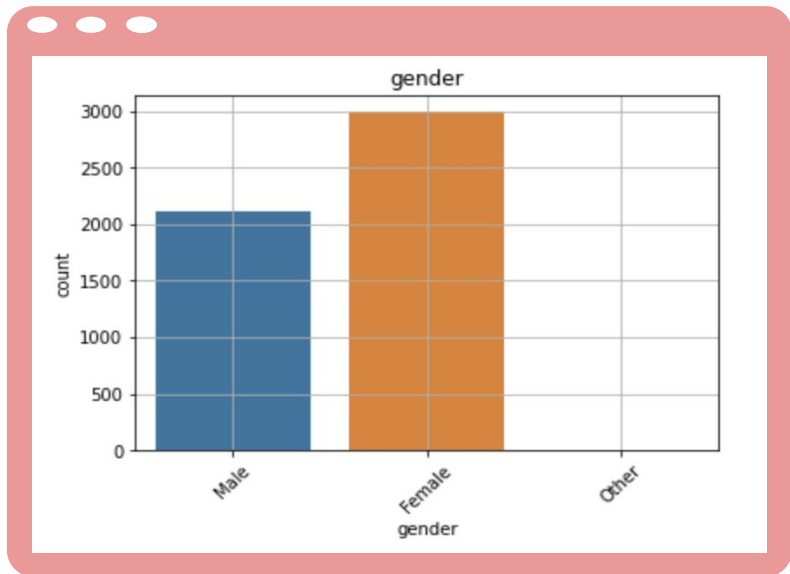
A glance at our numerical and categorical variables

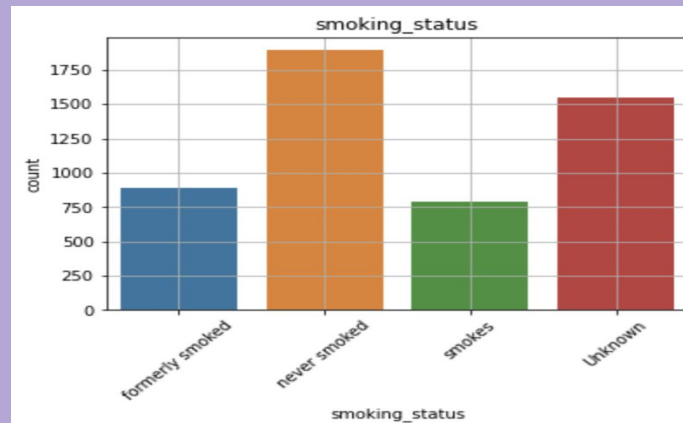
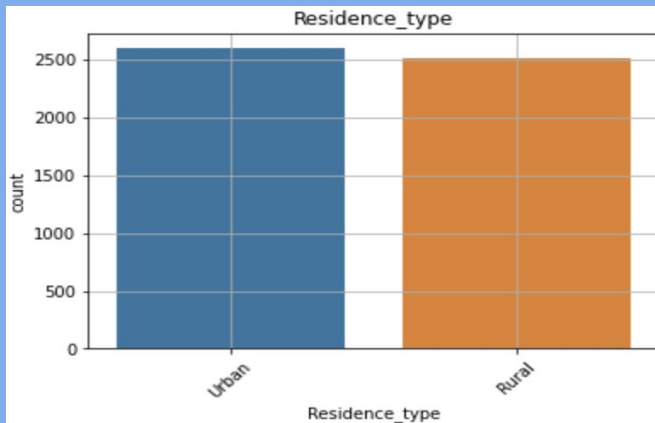
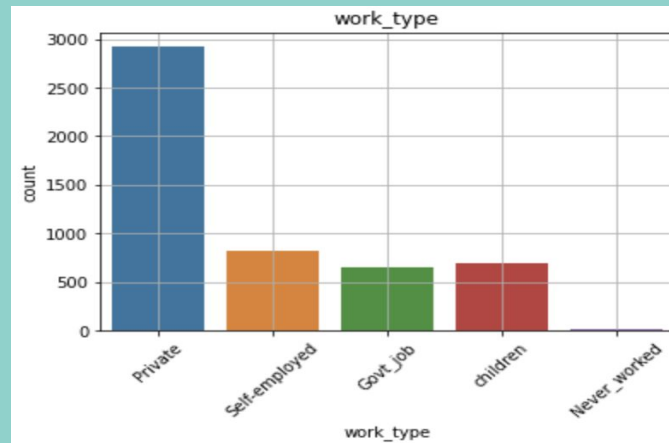
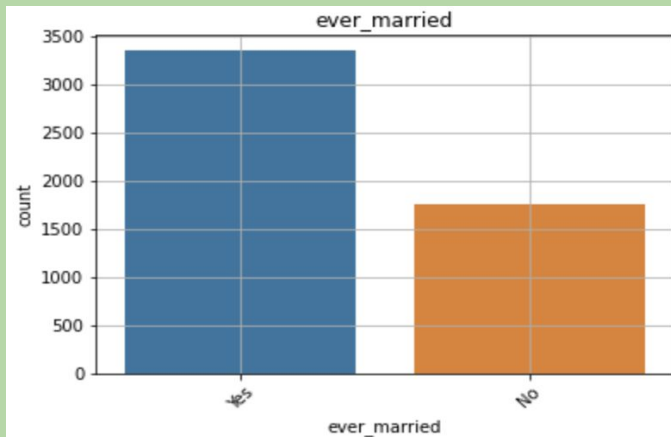
- Numerical variables distribution



A glance at our numerical and categorical variables

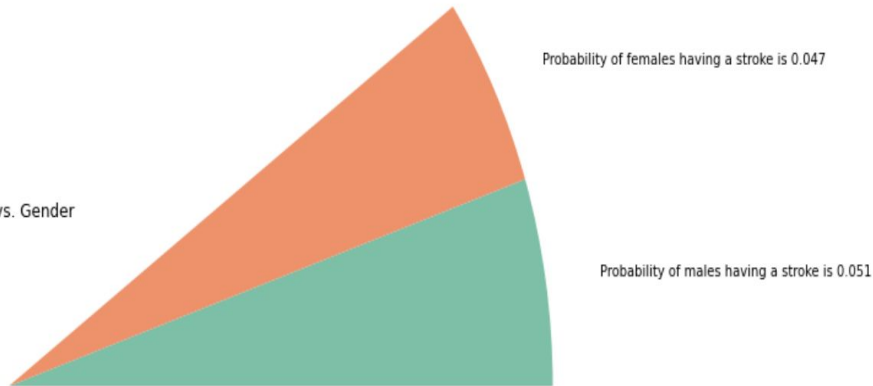
- Categorical variables distribution





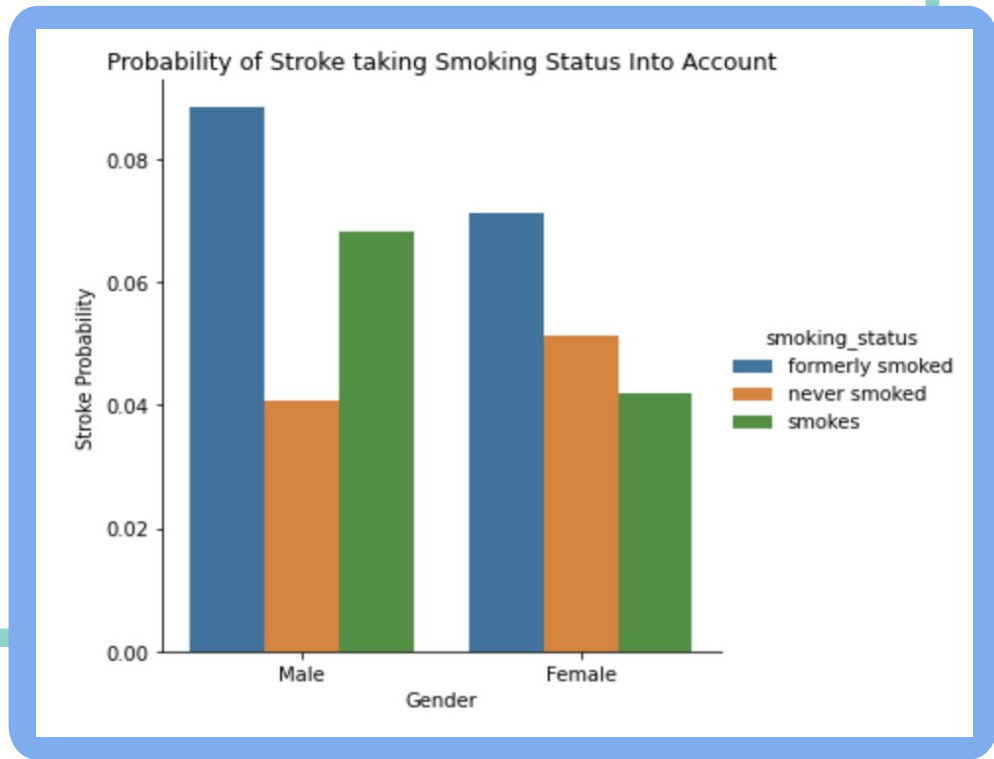
Do males have more stroke events than females?

Stroke vs. Gender



It seems like males face a higher chance of having a stroke but since the difference is small, more factors need to be examined!

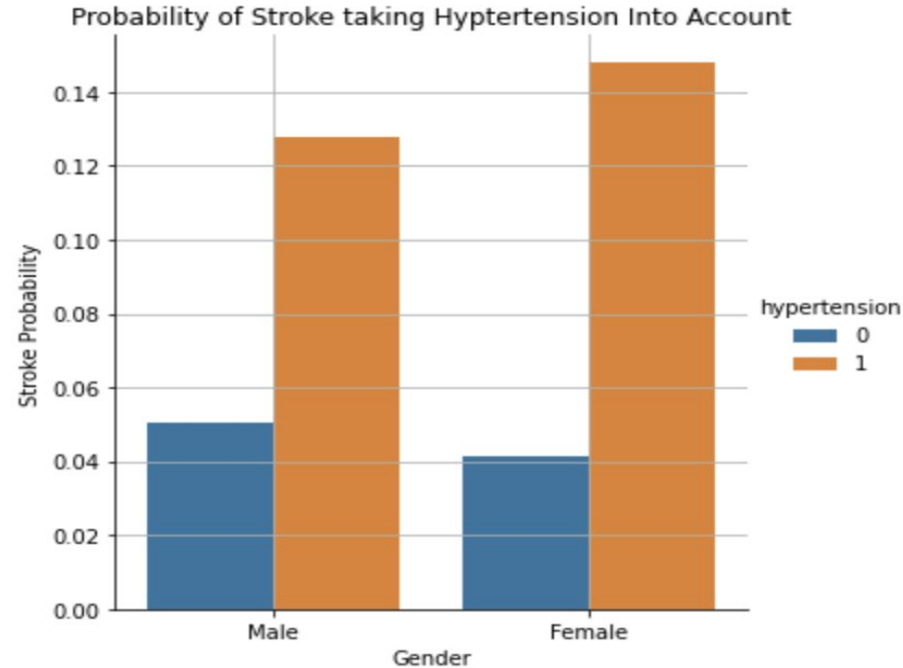




The Smoking Effect



The Hypertension Effect

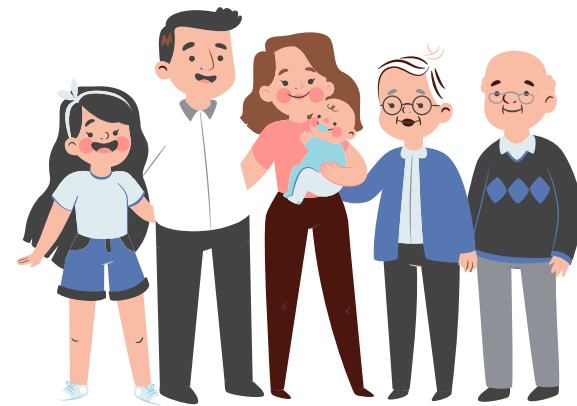
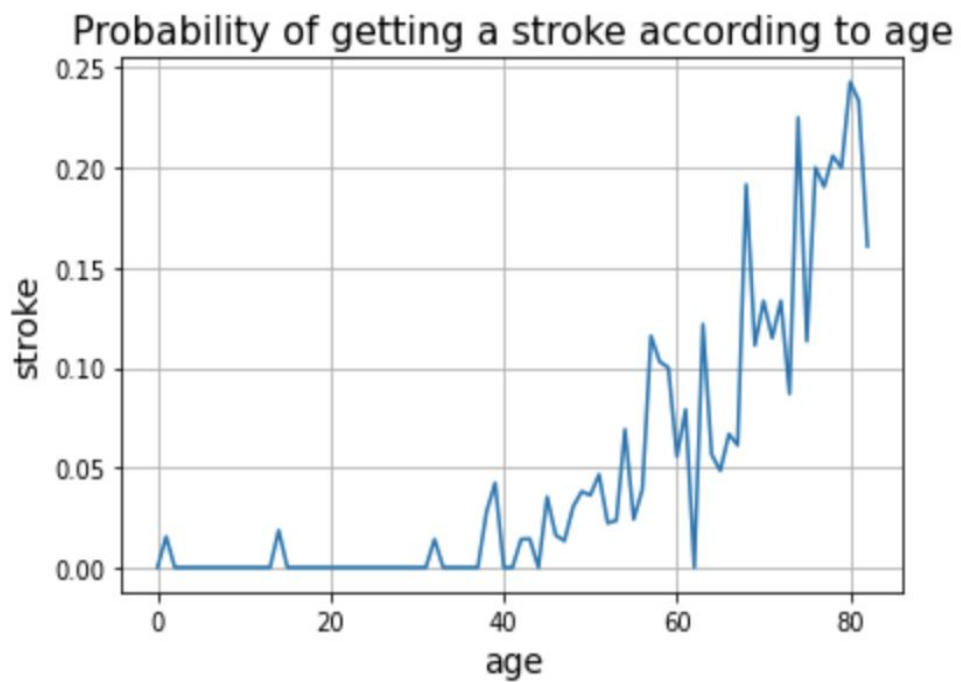


Thus, we can't say if any gender has a higher chance of facing a stroke!



Is there an age where strokes happen more often?

- Average age of getting a stroke : 67.73
- which age has the highest probability of getting stroke : 80



Is an older man more likely to get a stroke than an older woman?

Gender/Age	All	Young	Middle	Old
Female	0.04709	0.00302	0.02673	0.15246
Male	0.05106	0	0.03733	0.15827
All	0.04873	0.00162	0.03091	0.16166



How are physical conditions related to a stroke event?

Hypertension

with
hypertension

13%

without
hypertension

4%



How are physical conditions related to a stroke event?

Heart disease

with
heart disease

17%

without
heart disease

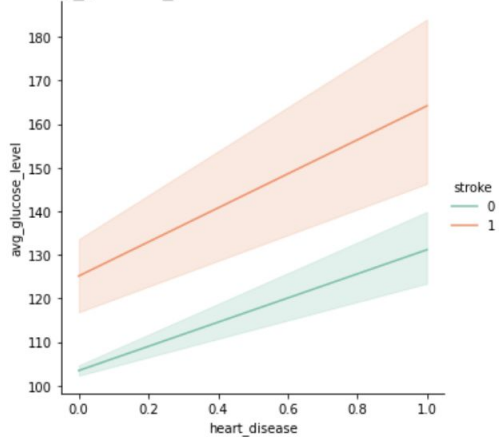
4%



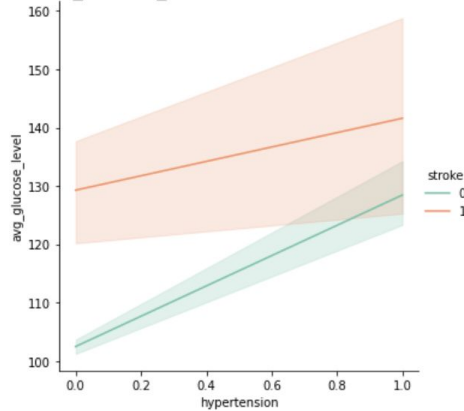
How are physical conditions related to a stroke event?

Average Glucose Level

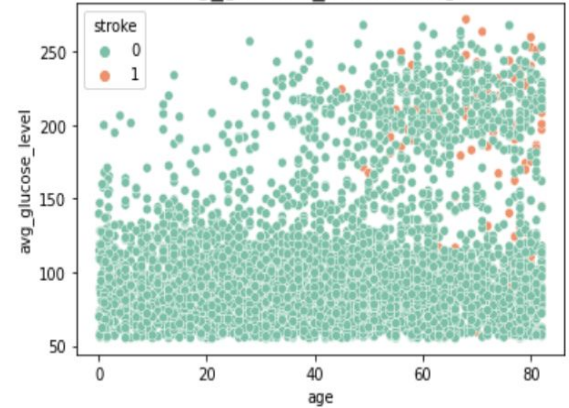
avg_glucose_level vs. heart_disease effect



avg_glucose_level vs. hypertension effect

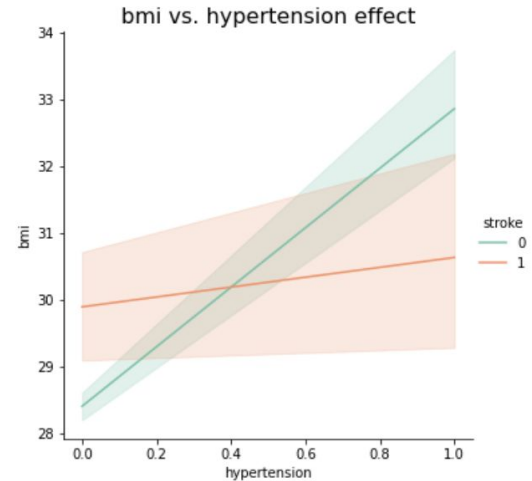
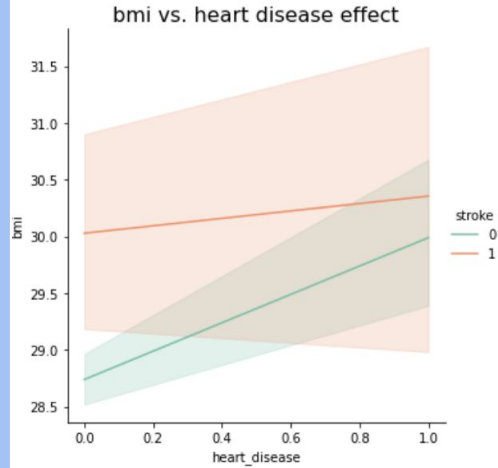
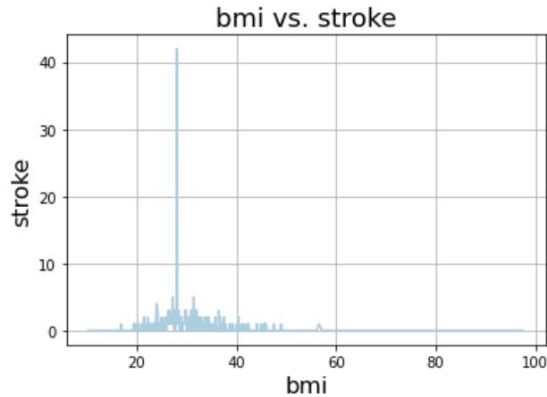


avg_glucose_level vs. age

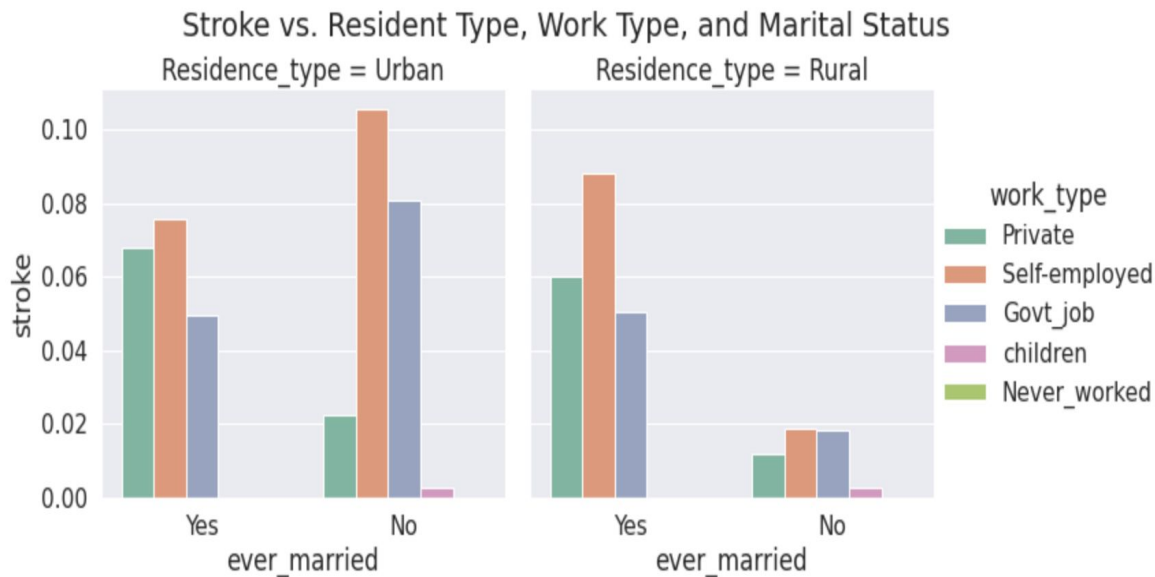


How are physical conditions related to a stroke event?

BMI (Body Mass Index)



Effect of Work Type, Marriage Status and Residence Type on Stroke



Observations

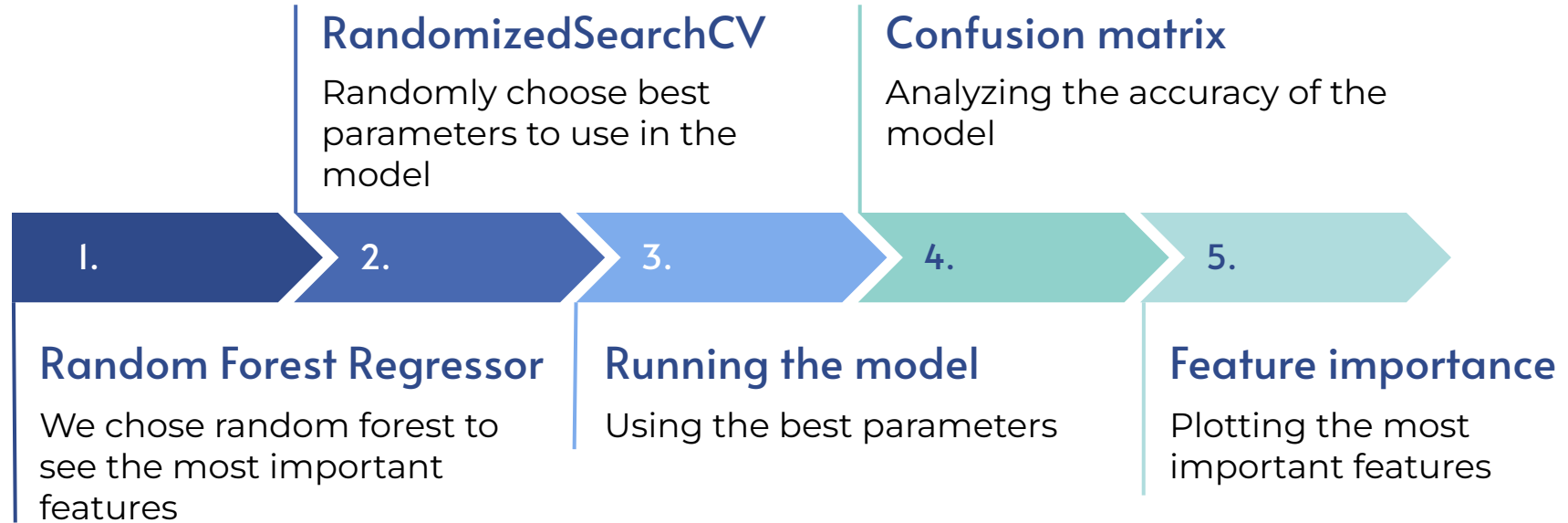
- Self employed population have the highest chance of facing a stroke
- There is a difference in the chance of a stroke of unmarried people who work in rural versus who work in urban areas

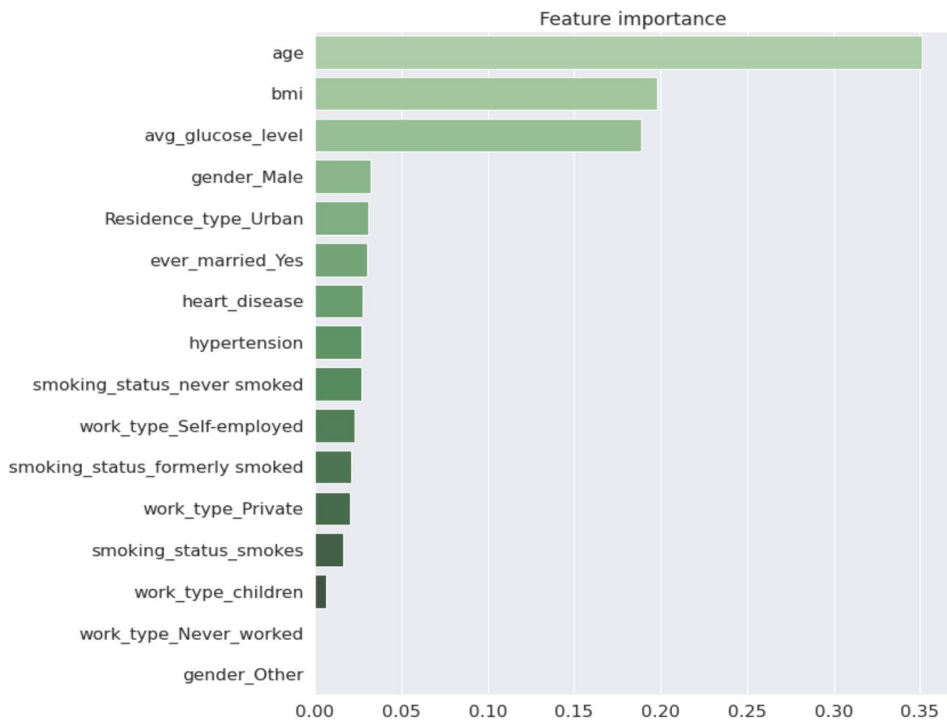
03.

Machine Learning



MACHINE LEARNING MODEL





Plotting the most important features in predicting a stroke event

04.

Conclusions



Gender

Although the difference is not much, males are more likely to have stroke event than females.

Smoking Status

Males who formerly smoked and currently smoke have a higher chance of facing a stroke and females who never smoked face a higher chance. Although it was discovered that underlying hypertension may be behind this relationship.

Age

As the age increases, the incidence of stroke also increases. The average age of having a stroke is 67 years old. Furthermore, age has more influence on the chance of getting a stroke compared to gender.

Hypertension

The probability of a patient with hypertension getting a stroke is 13%, which is three times that of a patient without hypertension.

Heart Disease

The probability of a patient with heart disease getting a stroke is 17%, which is four times that of a patient without heart disease.

Glucose Level

When the average glucose level rises, the number of individuals who get strokes and other physical problems also increases.

BMI (Body Mass Index)

BMI of patients who have experienced a stroke event is concentrated between 20 and 50. Furthermore, patients with the BMI of 28(overweight) are more likely to have a stroke event.

Work Type, Marriage Status, Residence Type

We cannot clearly say that these 3 factors are directly correlated to the possibility of having a stroke.

THANKS!

Do you have any questions?

Citations

For the machine learning section, these websites was used:

To normalize the target variable: [link](#)

To tune the hyperparameters: [link](#)

To create the confusion matrix: [link](#)

To create the important feature visualization this notebook was used:
02-Exploratory-Data-Analysis.ipynb

CREDITS: This presentation template was created by
Slidesgo, including icons by Flaticon, and infographics &
images by Freepik and illustrations

