

# **Cardiac arrests: Risk assessment and Analysis**



**11/13/2023**

**Akanksha Chaturvedi**

**[akanksha.chaturvedi@pace.edu](mailto:akanksha.chaturvedi@pace.edu)**

**Practical Data science CS-667**

**MS in Data science**

# Agenda

- Summary
- Literature review
- Project plan
- Dataset
- Exploratory data analysis
- Modeling methodologies
- Findings
- Result
- Conclusion



# **Problem Statement**

As we know after covid many people are getting cardiac arrests. Not only elderly people but at the young age also people are getting cardiac arrests. So it is very important to predict the possibility of cardiac arrest in people by looking at various factors in their day to day life. So can we predict cardiac arrest at early stages by considering feature comparisons?

# **Solution**

Build a classification model to predict cardiac arrest on early stages by analyzing the various factors that can affect the heart health.



# Project Plan

---

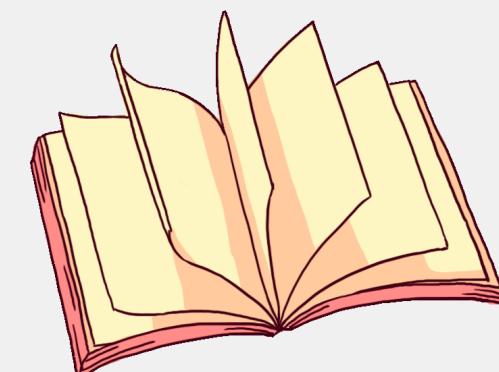
Deliverable	Details	Due Date	Status
Data & EDA	Data visualization and EDA of Ube ride dataset.	31-Oct-2023	COMPLETE
Methods, findings, & Recmmendations	Model building, results, recommendations for improvements	14-Nov-2023	COMPLETE
Final Presentation		5-Dec-2023	Not Started

# Literature review

5

I have done 4 literature review. The four models were built on various algorithms of machine learning, deep learning and data mining. All the four models used the datasets collected from different medical industries. The models trained on algorithms like random forest, decision tree, naive-bayes algorithm, XG boost, logistic regression, classification, PCA and function SMO. The accuracy of all these models were shown using F1 score, confusion matrix or silhoutte's score. These scores showed the accuracies of each model and explained what model should we use for predictions based on past values. The links to these papers are provided as below:

- <https://dl.acm.org/doi/pdf/10.1145/3542954.3543004>
- <https://dl.acm.org/doi/pdf/10.1145/3352411.3352413>
- <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9692977>
- <https://www.sciencedirect.com/science/article/pii/S2772442523000497>



D<sub>2</sub> A<sub>1</sub> T<sub>1</sub> A<sub>1</sub>

# Dataset



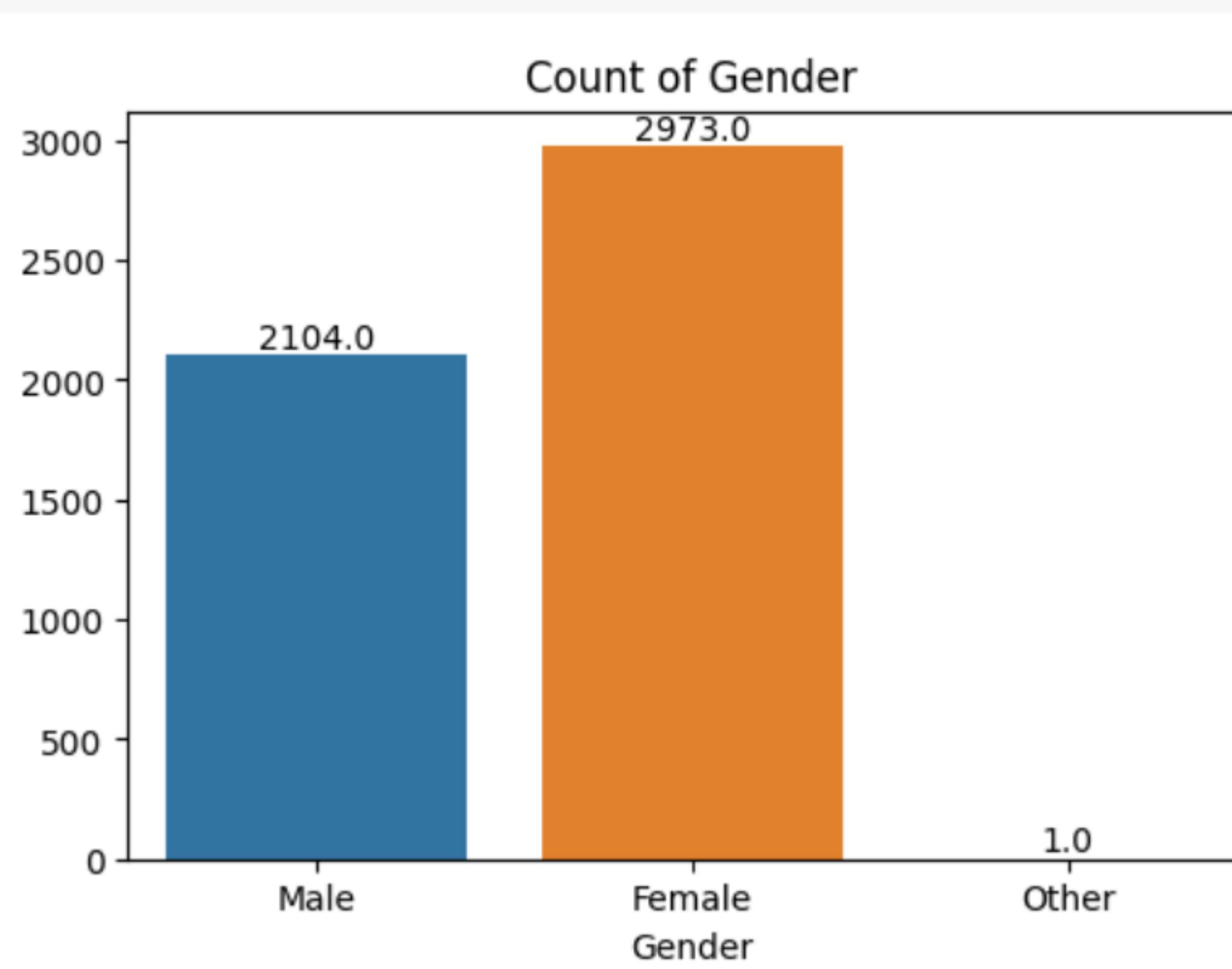
- **Link:**
  - <https://data.world/uci/heart-disease/workspace/file?filename=cleveland.data.csv>;
- **Description:** The dataset consists of information from the year 2008 to 2019. The dataset contains the various features like gender, marital status, glucose level, bmi, work type, residence type, previous heart disease if any to predict the possibility of cardiac arrest in people.
- **Size:** The dataset contains 5110 rows and 12 columns.
- **Time period:** 2008 to 2019
- **Assumption:** I have assumed that the ID column is not going to be used in the prediction so I removed it.
- **Data Sample:**

	<code>id</code>	<code>gender</code>	<code>age</code>	<code>hypertension</code>	<code>heart_disease</code>	<code>ever_married</code>	<code>work_type</code>	<code>Residence_type</code>	<code>avg_glucose_level</code>
0	9046.0	Male	67.0	0	1	Yes	Private	Urban	228.69

# EDA and Data visualization

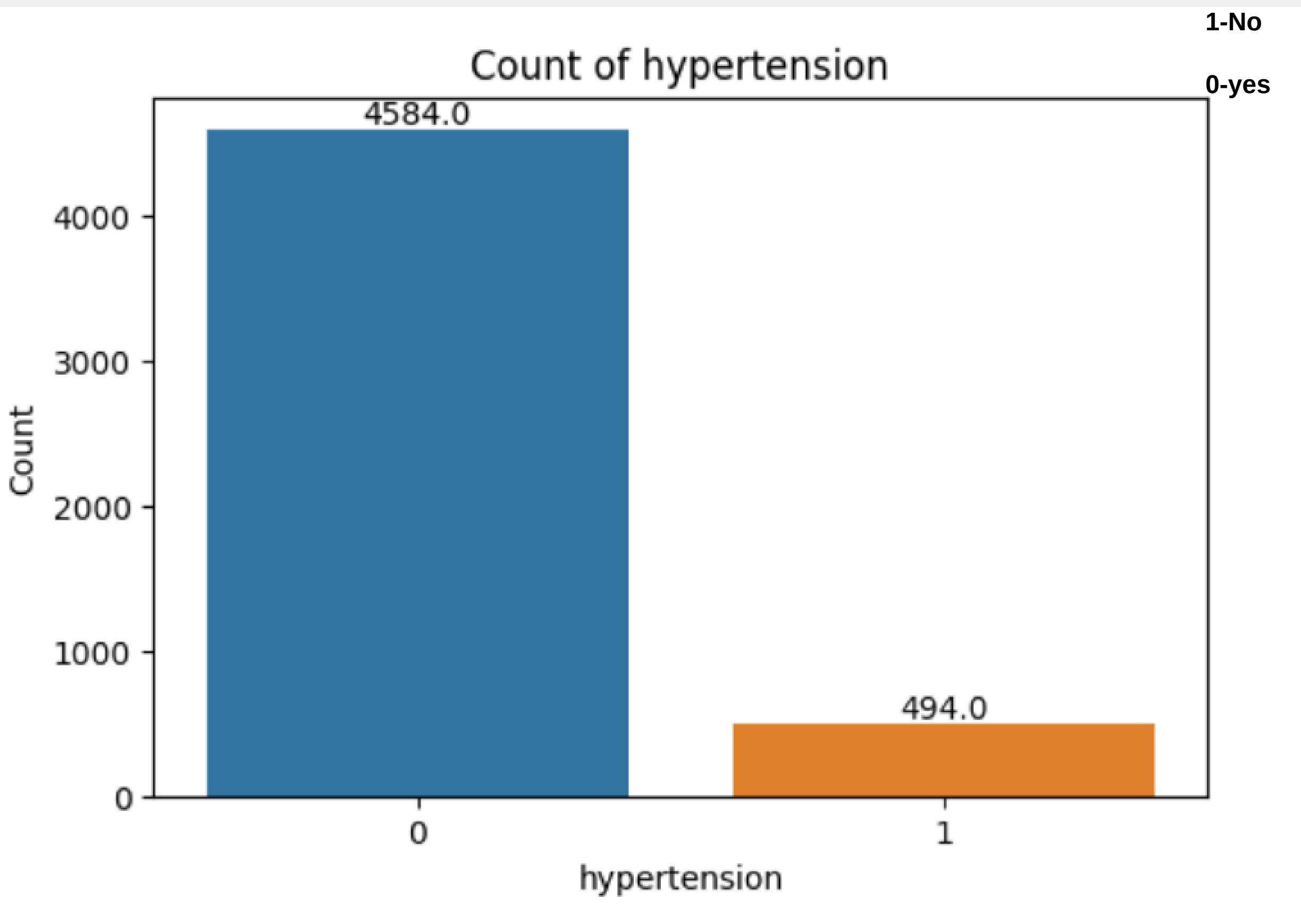


# Gender based analysis



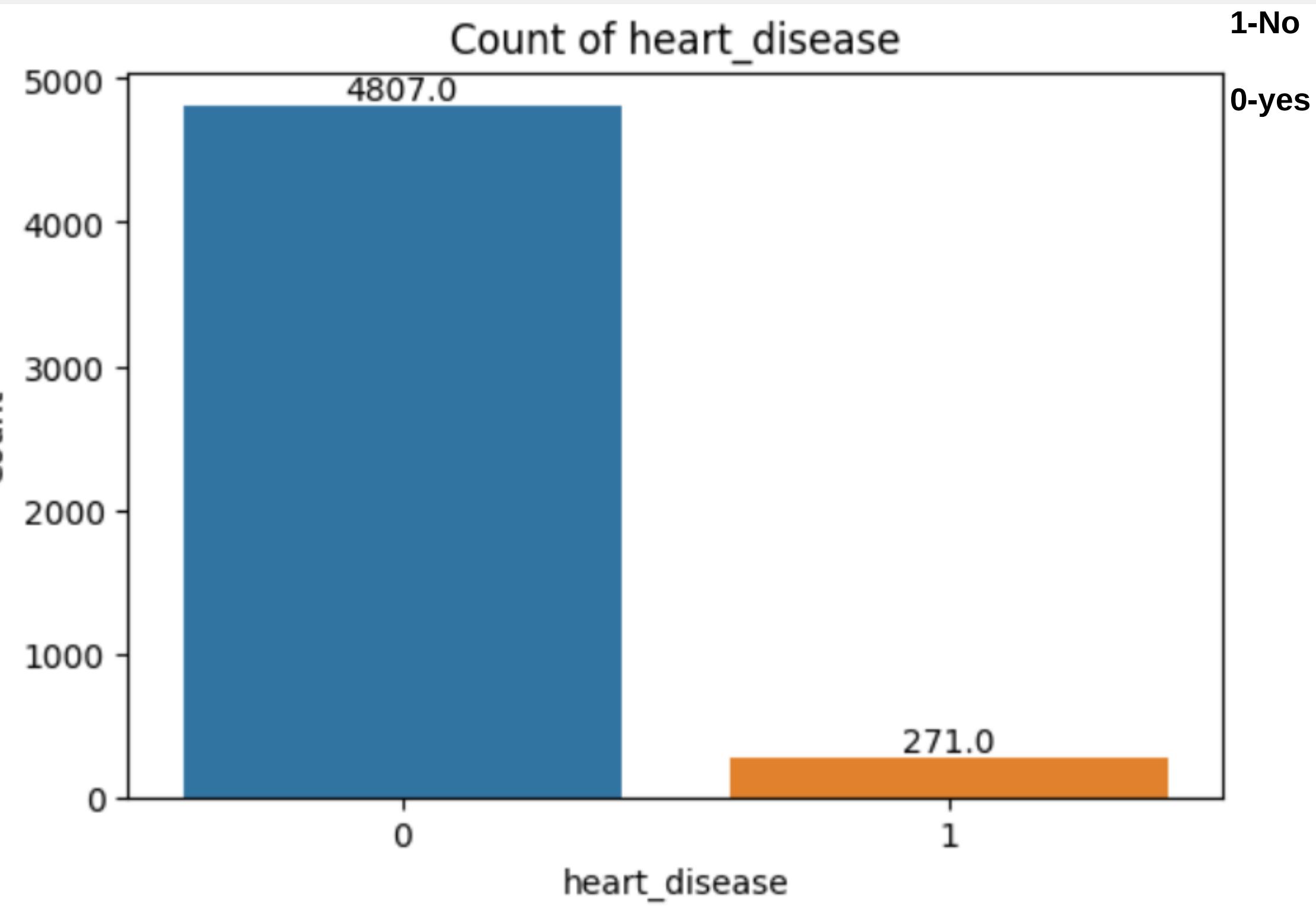
- The graph shows the count of male, female and others having cardiac arrests.
- As we can see the number of females getting heart attack is 2973 and the number of males getting heart attack is 2104.
- The number of females are more prone to cardiac arrest than males.

# Hypertension as a factor affecting heart health



- The graph shows the count of people having hypertension or not.
- As we can see the number of people with hypertension are more affected with a count of 4584 as compared to the people who don't have hypertension that is 494..

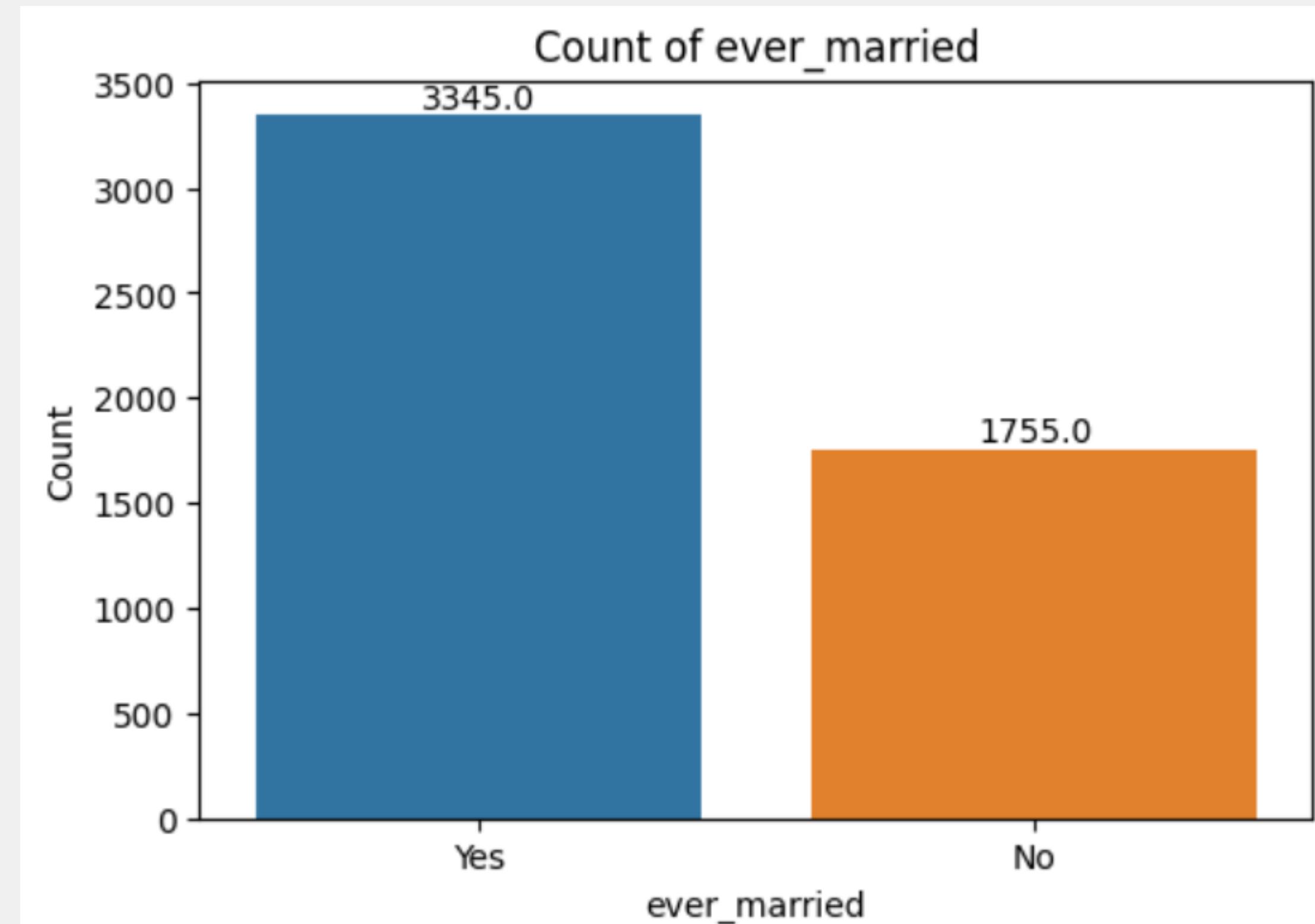
# Affect of heart disease on cardiac arrest



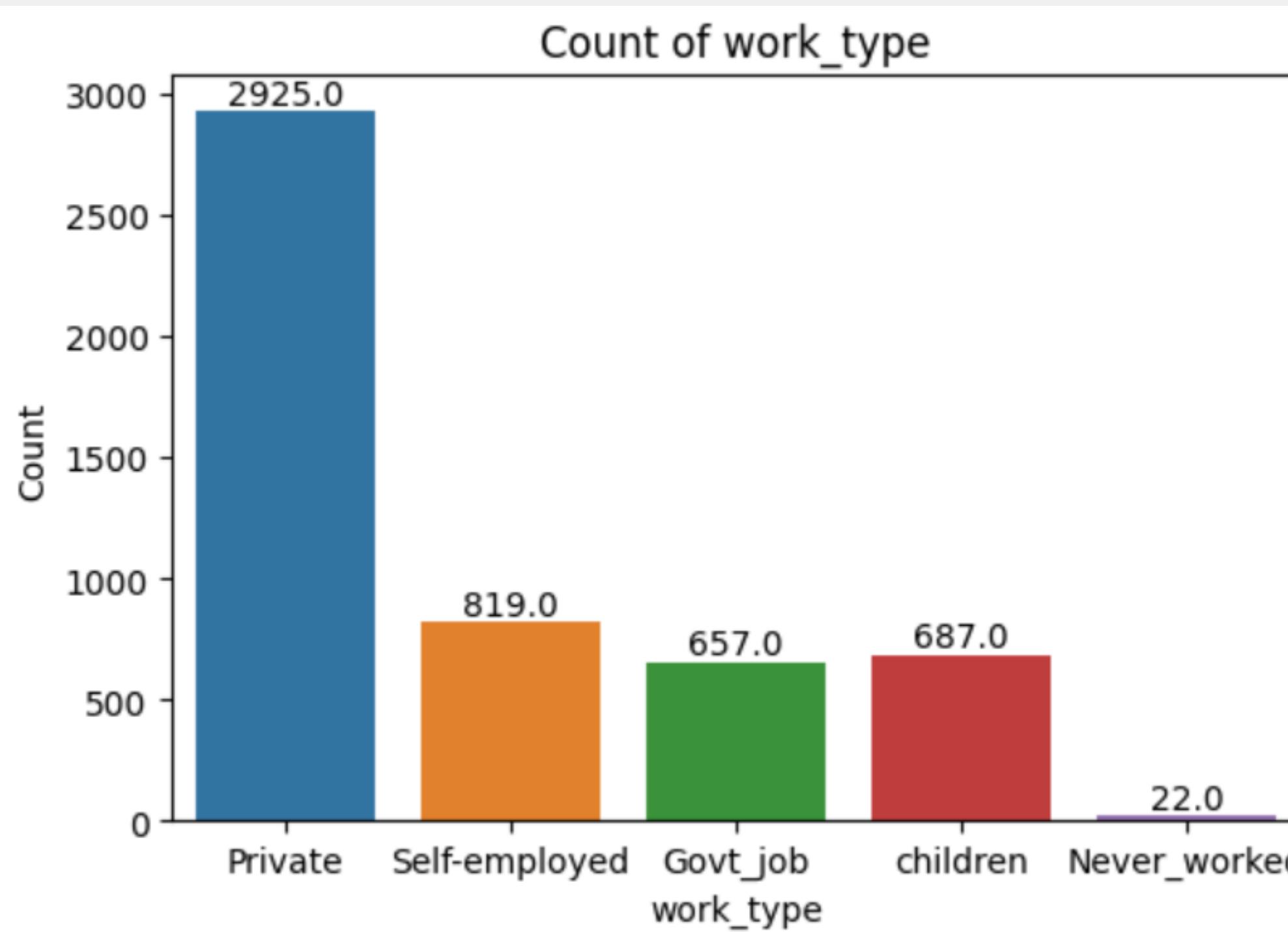
- As we can see people already having any kind of heart disease are more prone to cardiac arrest than a person not having any prior heart disease.

# If marital status can be a factor for cardiac arrest?

As we can see from the graph marital status is highly related with cardiac arrest. Married people are more prone to cardiac arrest than unmarried people.



# If work type can be a factor for cardiac arrest?



As we can see from the graph, the private employees are more prone to a heart attack as compared to other professions. But the interesting thing in this graph is children are equally prone to heart attack almost like a govt. employees.

# Modeling



# Cardiac arrest prediction using XG boost classifier

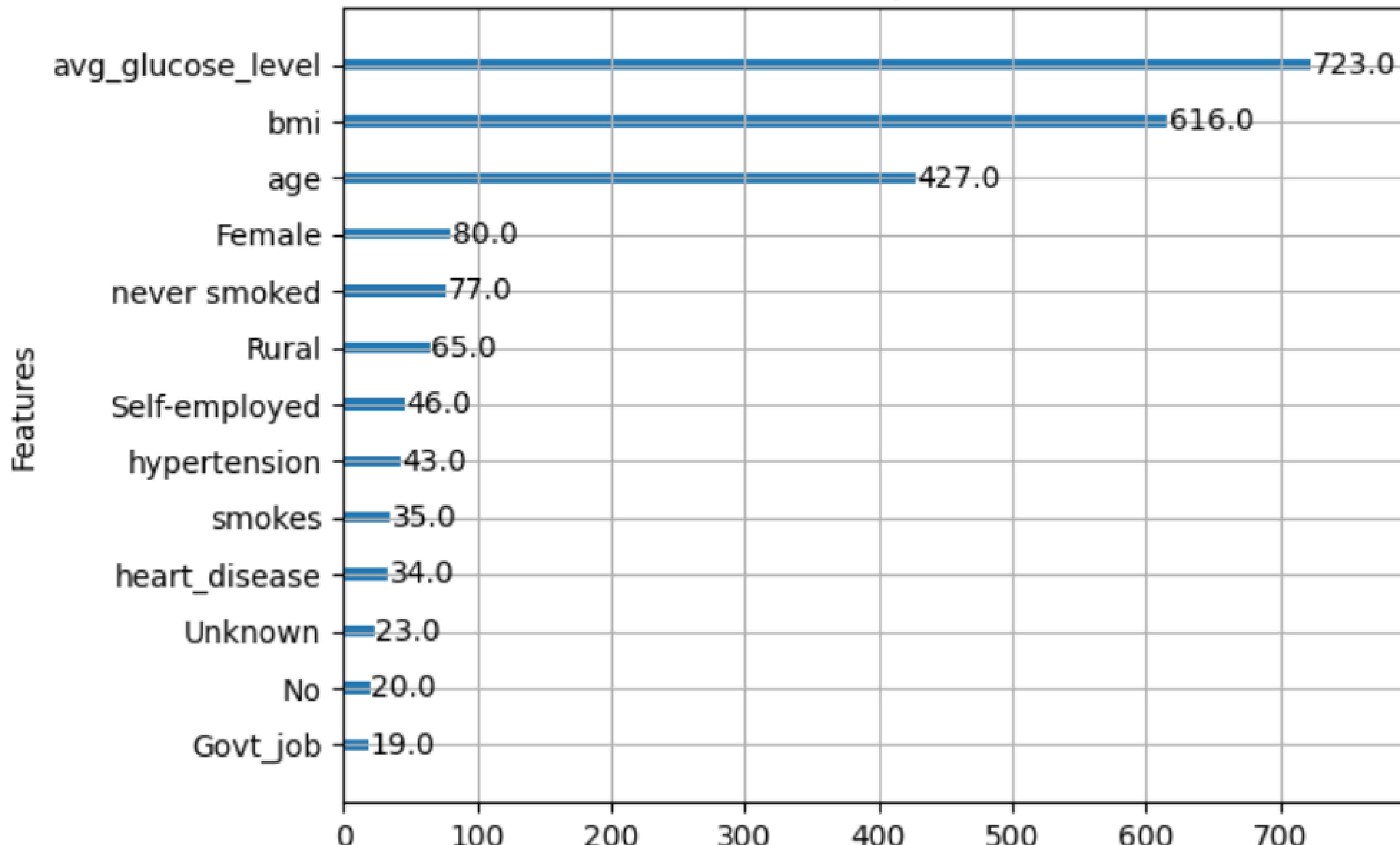
## Introduction:

Heart failure is a very critical problem now a days when a lot of people are getting cardiac arrests in very young age. We have employed a model where we can predict the chances of cardiac arrest based on people's day to day life.

## Features:

We have selected features like work type, gender, marital status, BMI, glucose level to predict the chances of cardiac arrest in people. We can predict the heart failure by analyzing different features

## Feature importance

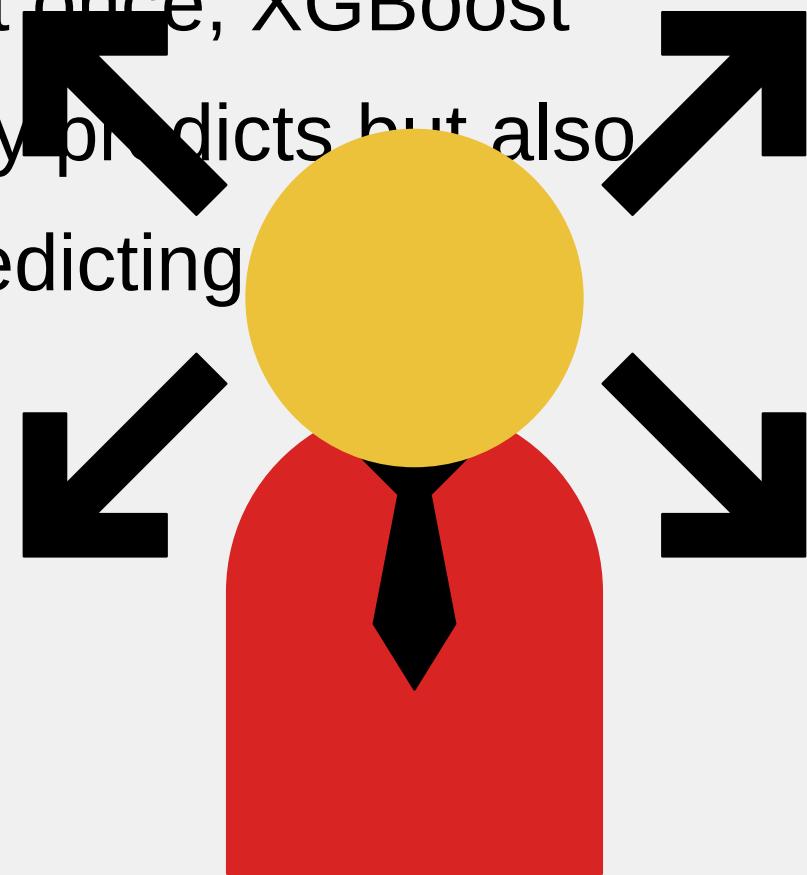


## **Importance of early detection:**

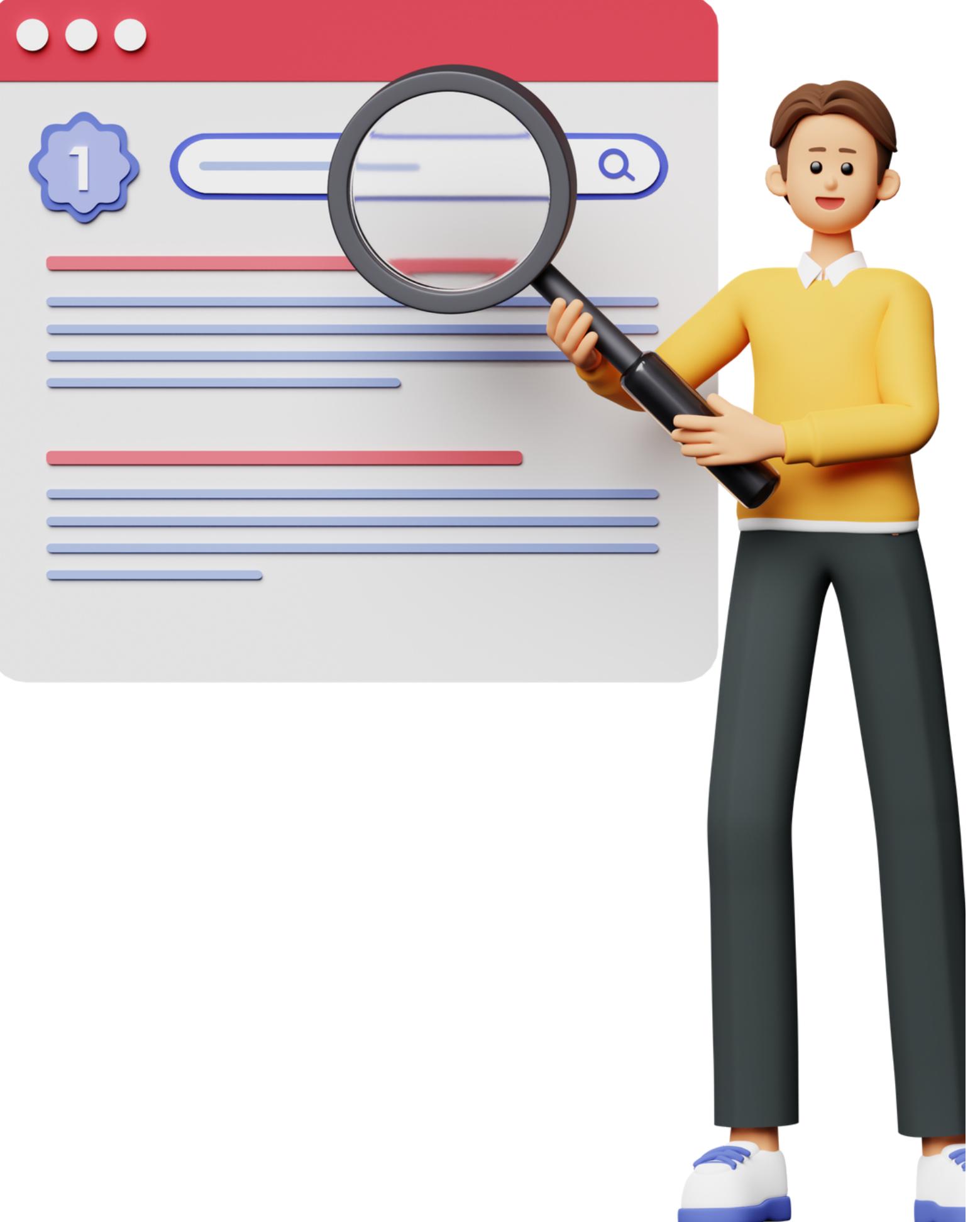
If cardiac arrest is predicted in early stages then there would be a good chance that we can save a person's life. We can analyze the various features which affects heart health.

## **Why XG boost classifier:**

Imagine XGBoost as a super-smart health detective team. This team, known for its accuracy and efficiency, works together to predict health outcomes. They learn from their mistakes, focusing on the big picture without getting bogged down in unnecessary details. XGBoost avoids making up stories, ensuring accurate predictions without guessing. It works quickly and adapts to missing information, making it efficient and adaptable. Just like a detective agency that can handle many cases at once, XGBoost excels in predicting health issues for a large group of people. Importantly, it not only predicts but also explains its reasoning, making it the ideal detective team for understanding and predicting mysteries in a way that makes sense.



# Findings



# Model Accuracy in the prediction of cardiac arrest

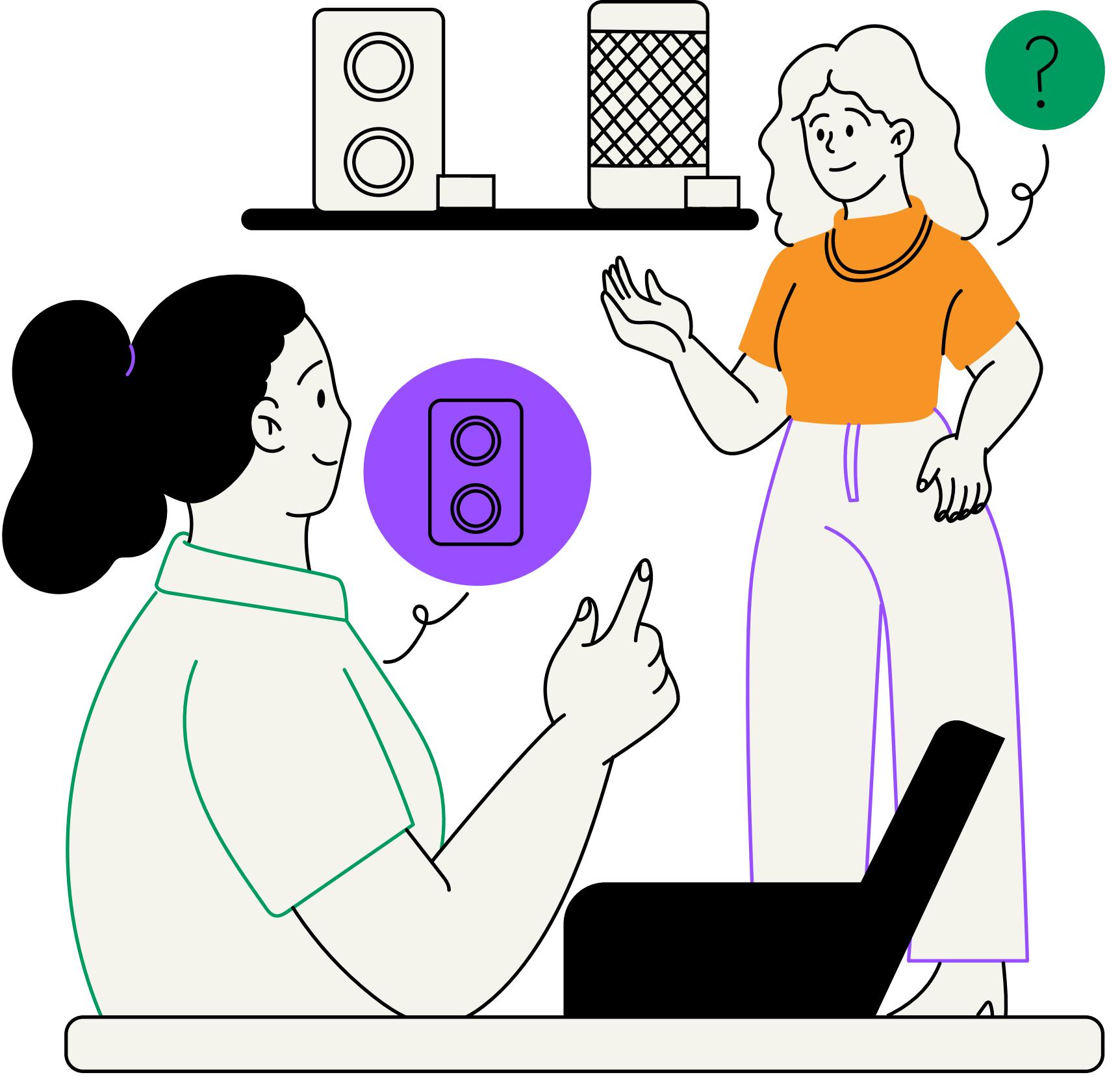
## Methodology to calculate accuracy:

- The accuracy of the model is given by the accuracy score. The accuracy of the testing and prediction model is given and it is 1 for testing and 0.94 for the training model.
- I have checked the authenticity of prediction if it is true positive or negative with the help of the confusion matrix.



## **Findings:**

- **Heart health plays a vital role in cardiac arrest**
- **Gender has significant importance in heartache cases. females are more prone to cardiac arrest than males.**
- **Jobs that do not involve much physical activity can be dangerous for heart health and can lead to cardiac arrest.**



# Recommendation

# How to increase model accuracy

## 1. Get More Data:

Collect a larger variety of information about Uber rides to help the computer learn better.

## 2. Think About Important Details:

Pay attention to details that might affect the price, like what kind of weather it is when you're travelling and does it affects the fare amount.

## 3. Make Sure Data is Clean:

Check that the information you have is accurate and complete. Clean up any mistakes or missing info.

## 4. Adjust Model Settings:

Tweak the computer's settings to make it better at predicting prices. It's like adjusting the settings on your phone for better performance.

## 5. Keep Improving:

Regularly check how well the predictions are working. If you get more information or notice something new, update the computer program to keep making better predictions.

# Next Steps

I will include more features so that I can get more accurate results. Also, I will be using more techniques to see what is the best model to predict the cardiac arrest using feature comparison.

