Heart Disease Prediction Project

NAME :

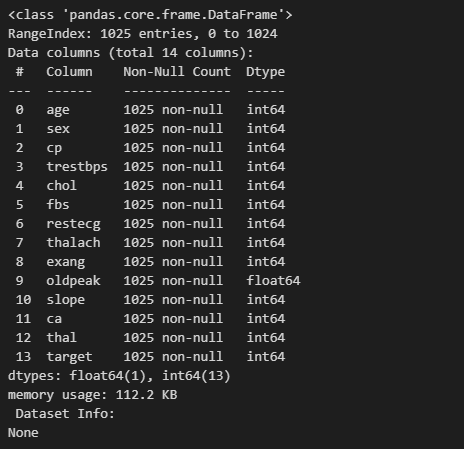
M Nauman Anwar

# 1. Introduction

This project aims to predict the presence of heart disease using machine learning. The dataset was obtained from Kaggle and Support Vector Machine (SVM) algorithm was used to train the model. The project workflow includes data preprocessing, analysis, model training, saving, and deployment via a Streamlit app.

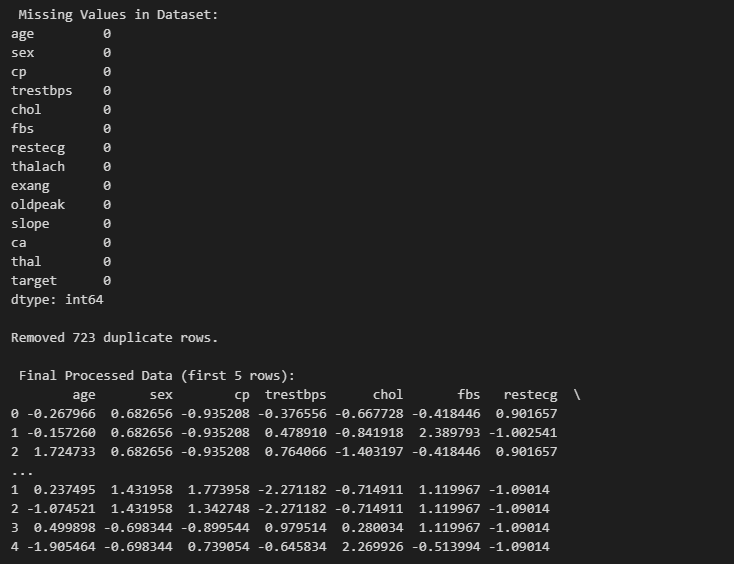
# 2. Dataset

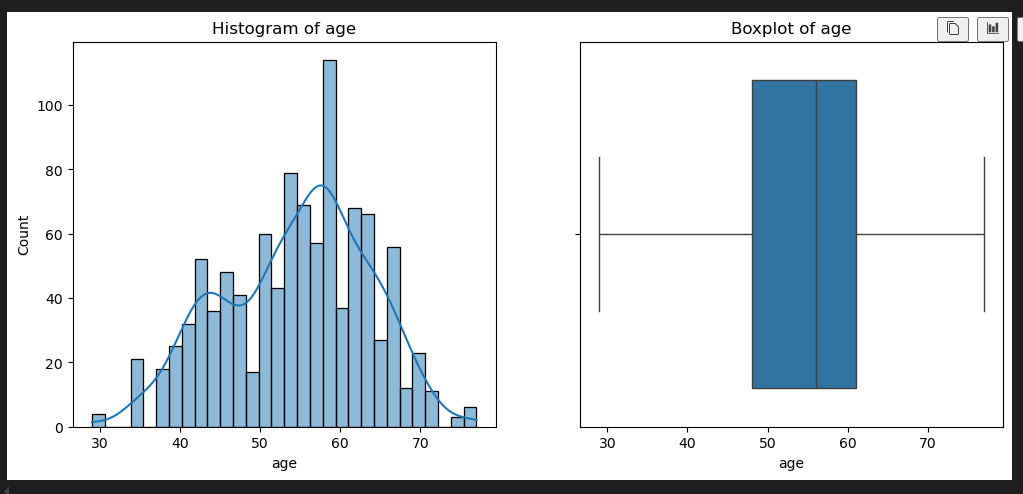
The dataset contains patient health data with features such as age, sex, chest pain type, cholesterol, blood pressure, maximum heart rate, etc. The target variable indicates whether the patient has heart disease (1) or not (0).

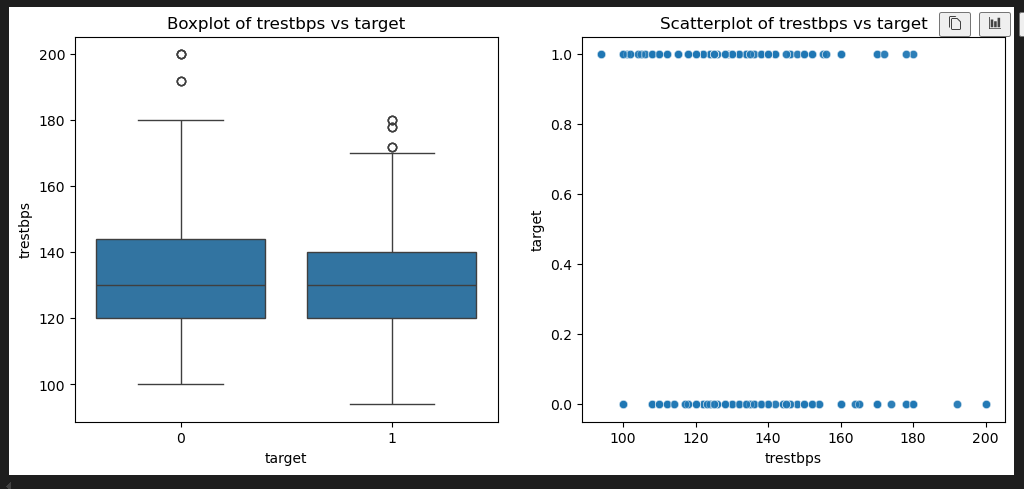


# 3. Data Preprocessing

Data cleaning steps performed:  
- Checked for missing values  
- Removed duplicate rows  
- Encoded categorical variables  
- Scaled numerical features

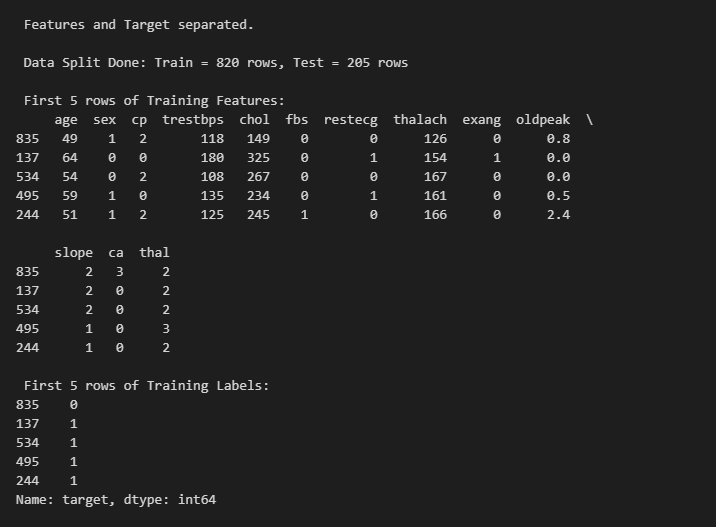






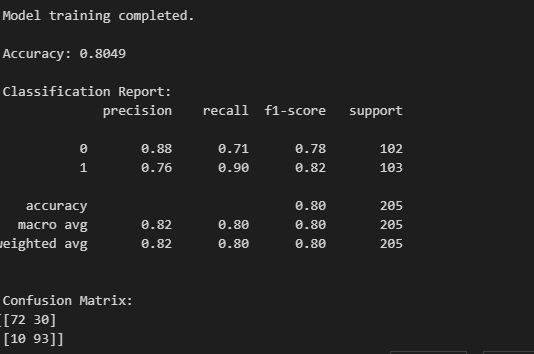
# 4. Data Analysis

Univariate analysis (histograms, boxplots) and bivariate analysis (boxplots, countplots, correlation heatmap) were performed to understand feature distributions and relationships with the target variable.



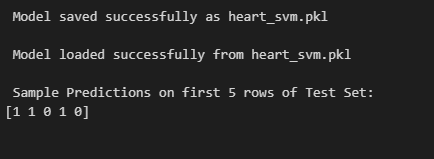
# 5. Model Training

The dataset was split into training and testing sets. A Support Vector Machine (SVM) model was trained on the training set. The model was then evaluated using accuracy, classification report, and confusion matrix.



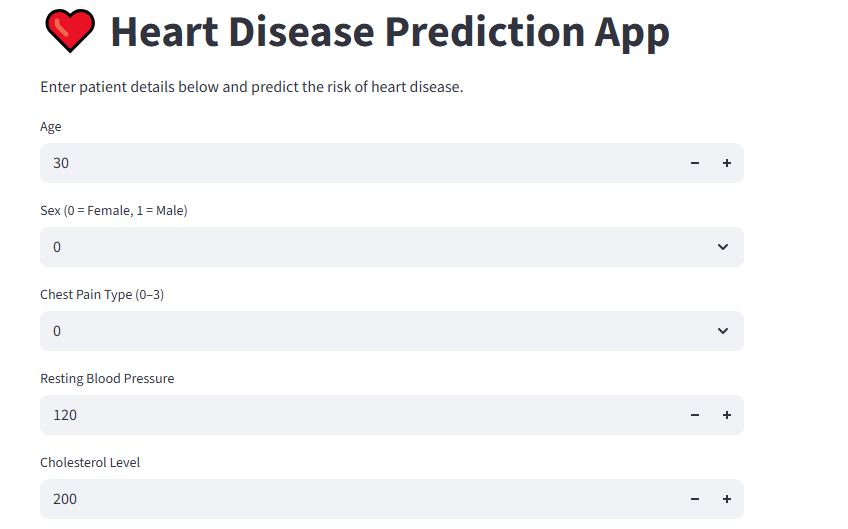
# 6. Model Saving

The trained model was saved using Python's pickle library for future use without retraining.



# 7. Application

A simple Streamlit web application was developed. Users can input patient details and get a prediction whether the patient is likely to have heart disease or not.





# 8. Conclusion

The SVM model successfully predicted the risk of heart disease with good accuracy. The Streamlit app made the model accessible and user-friendly for making predictions.