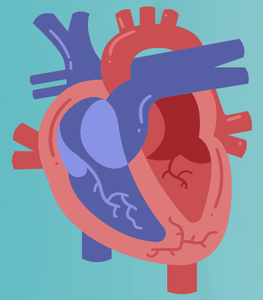


Leveraging Machine Learning for Cardiovascular Disease Prediction

Problem Statement

Cardiovascular diseases remain a leading cause of mortality worldwide, necessitating early detection and prevention strategies.

Research Questions

Key risk factors for cardiovascular diseases

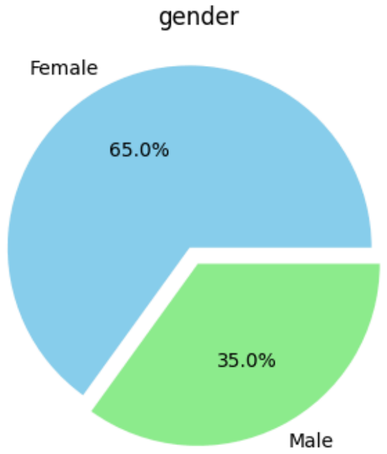
ML's role in predicting CVD likelihood

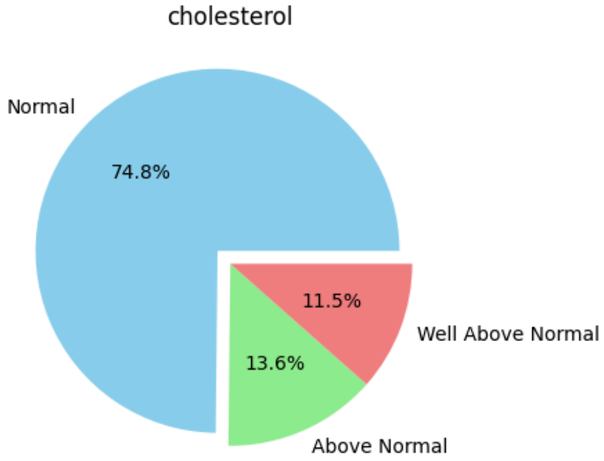
Patient data insights for preventive measures

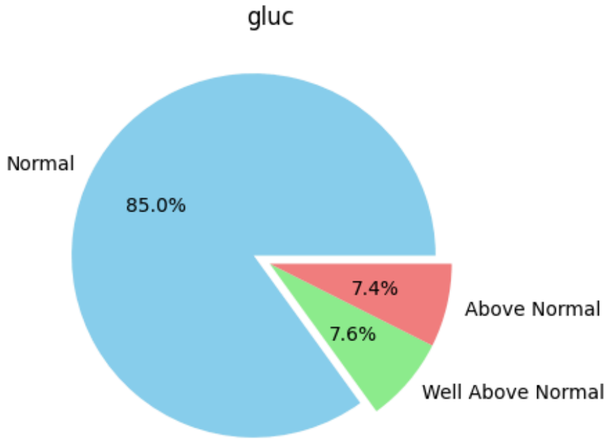
Objectives

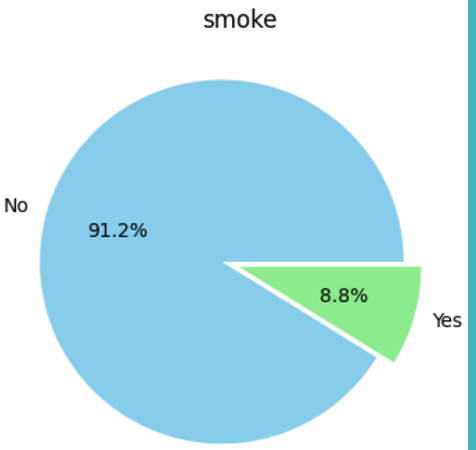
- To analyze data to prioritize key CVD risk factors.
- To use machine learning to predict CVD occurrence.
- To offer actionable insights to end users for early intervention based on patient segmentation.
- To create a user friendly app, scalable for various populations, predicting CVD risk and providing health recommendations

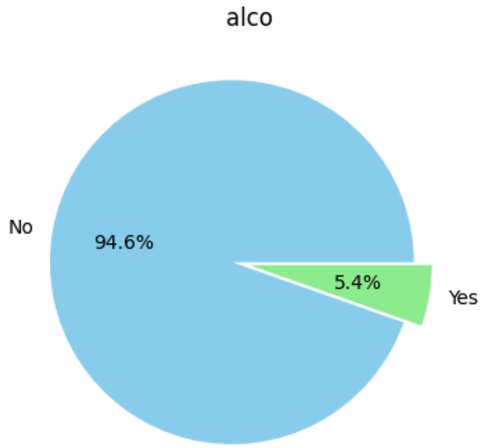
DATA DESCRIPTION

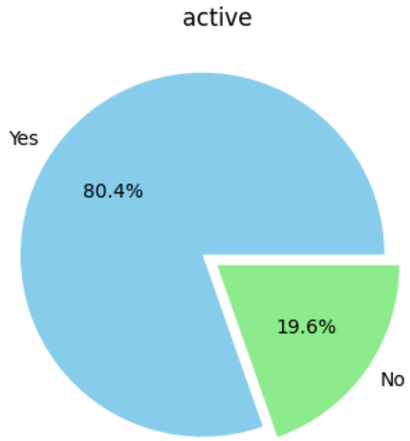


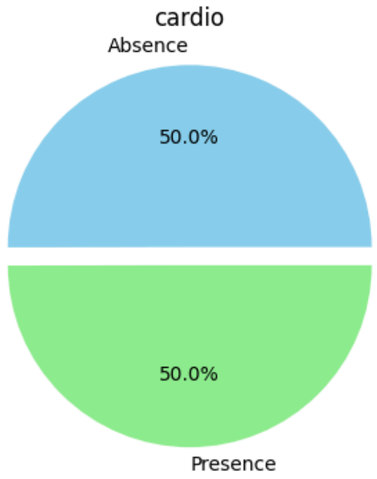




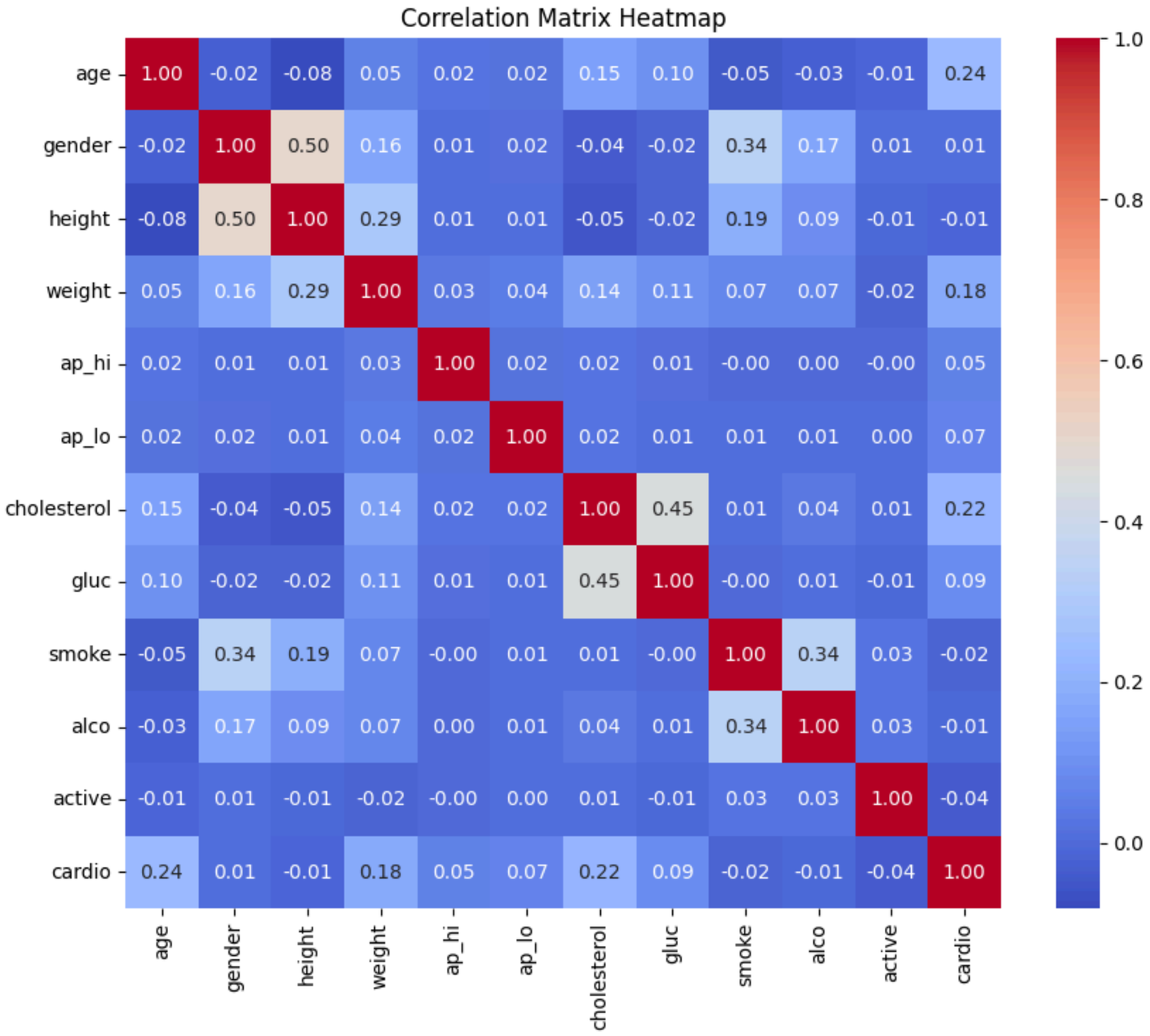




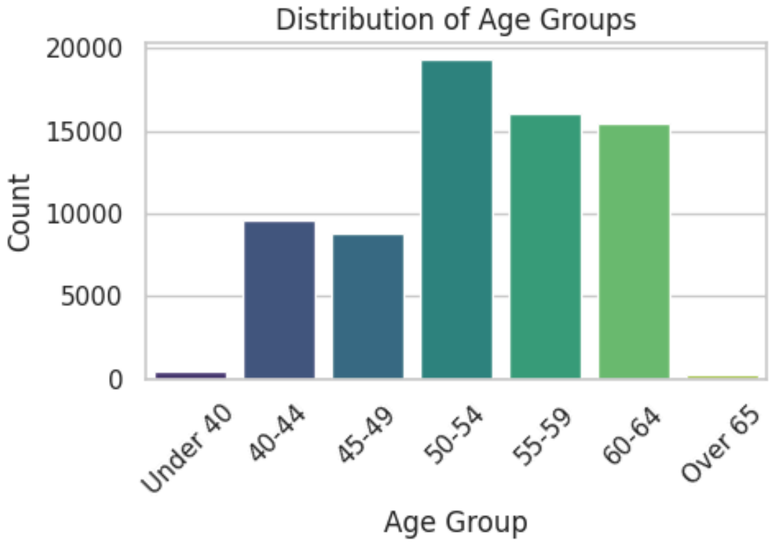


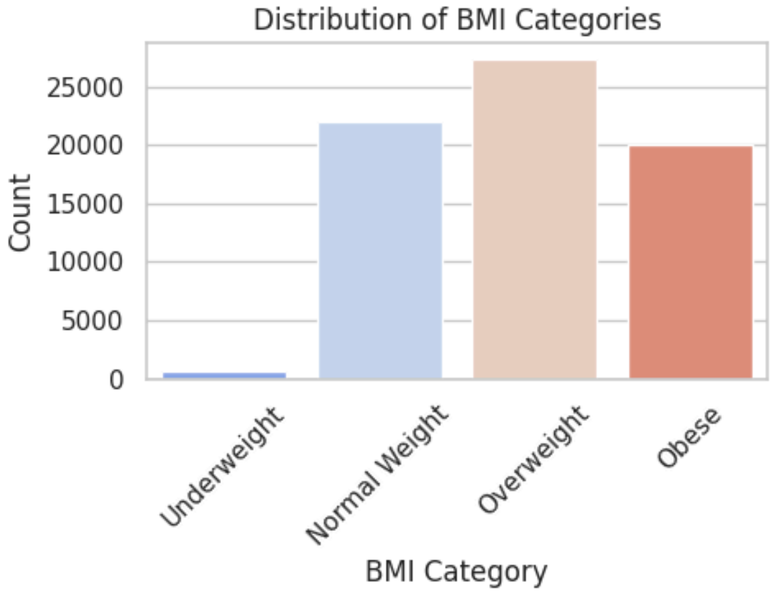


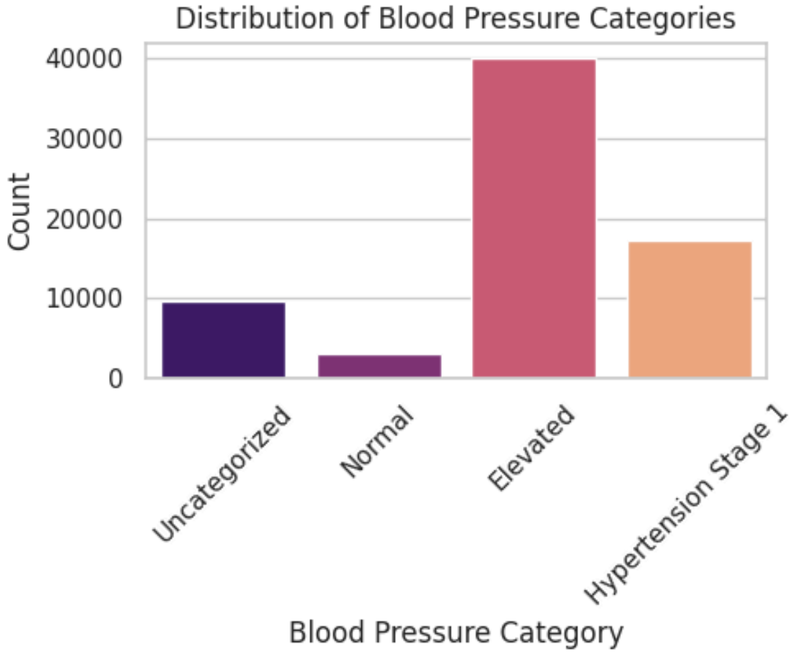
Key Risk Factors Identification

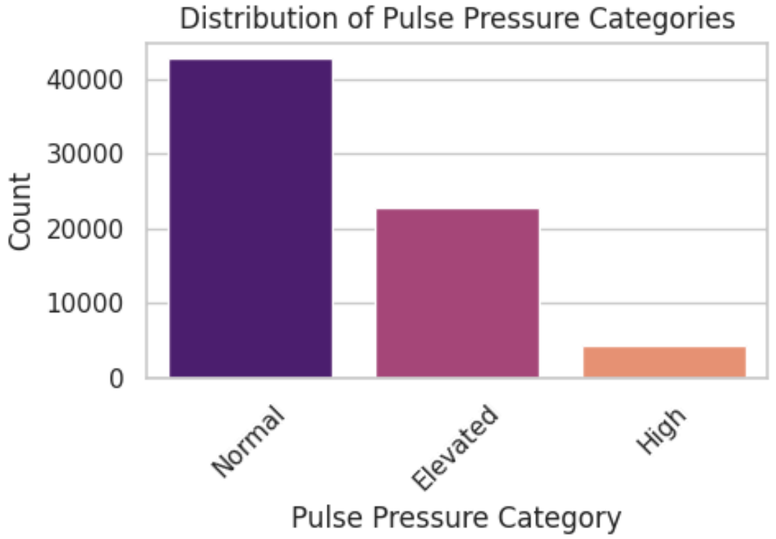


Relationship between Cardio and other variables





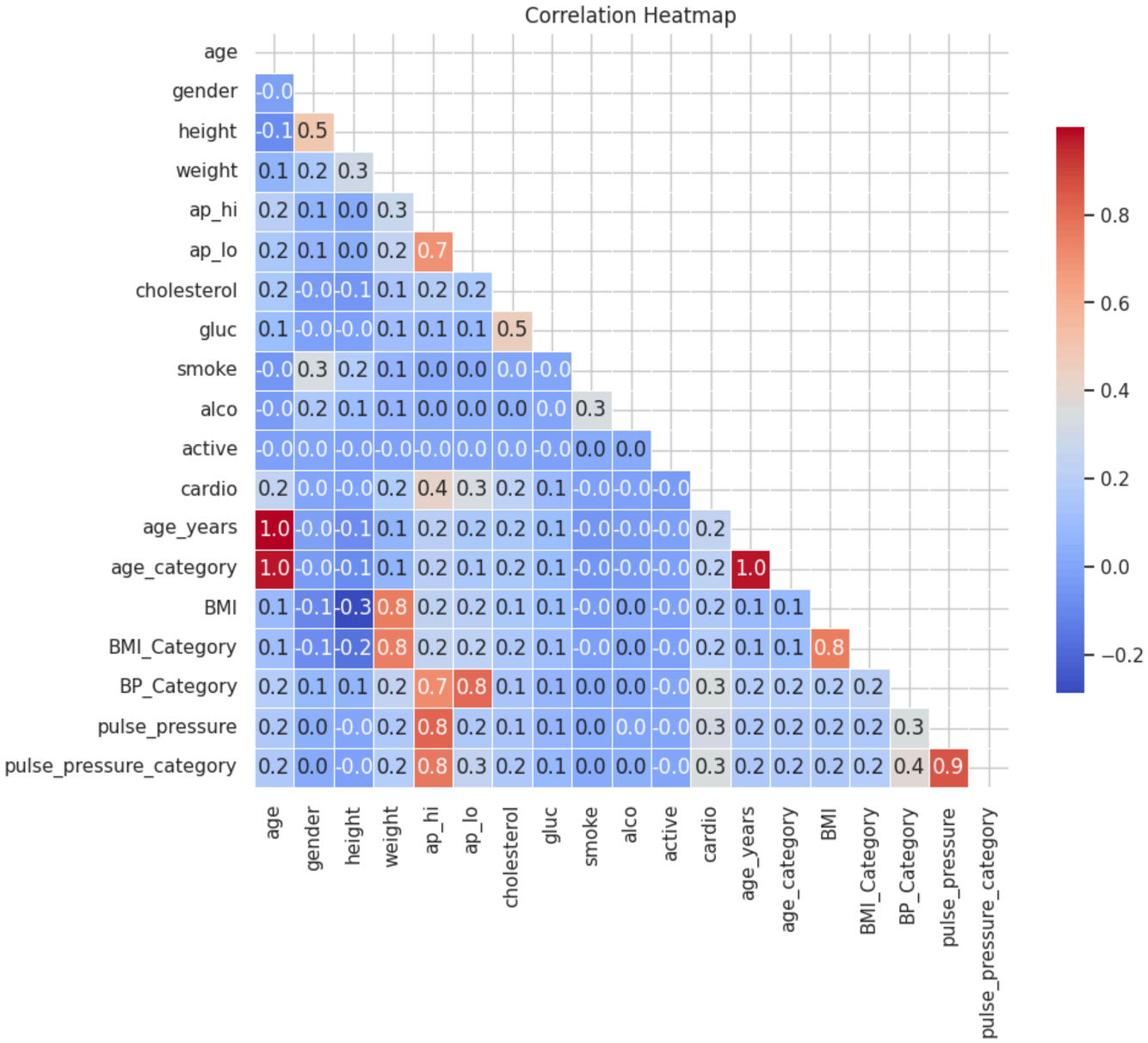




Feature Engineering

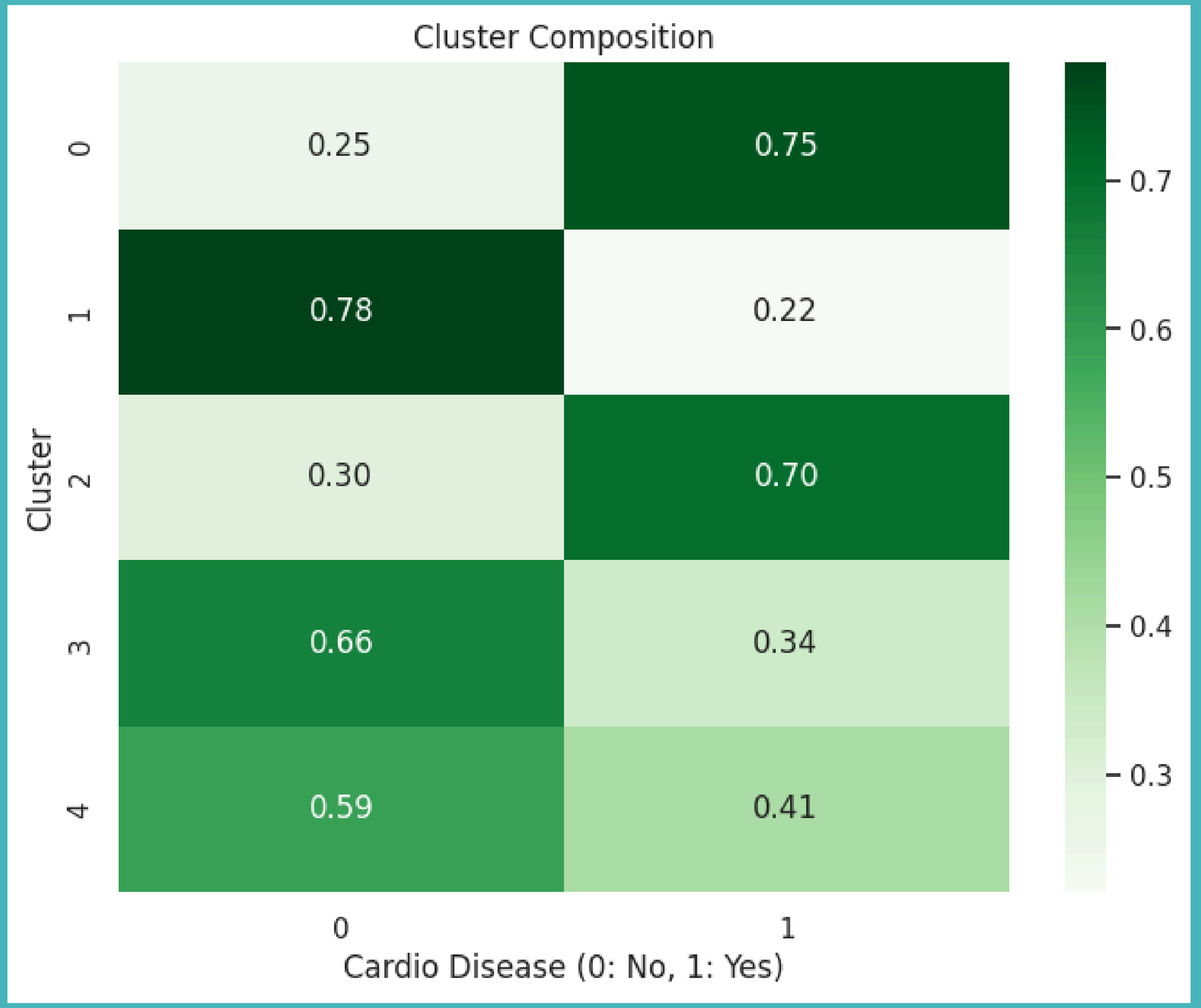
Created and categorized the features as shown

Key Risk Factors Identification



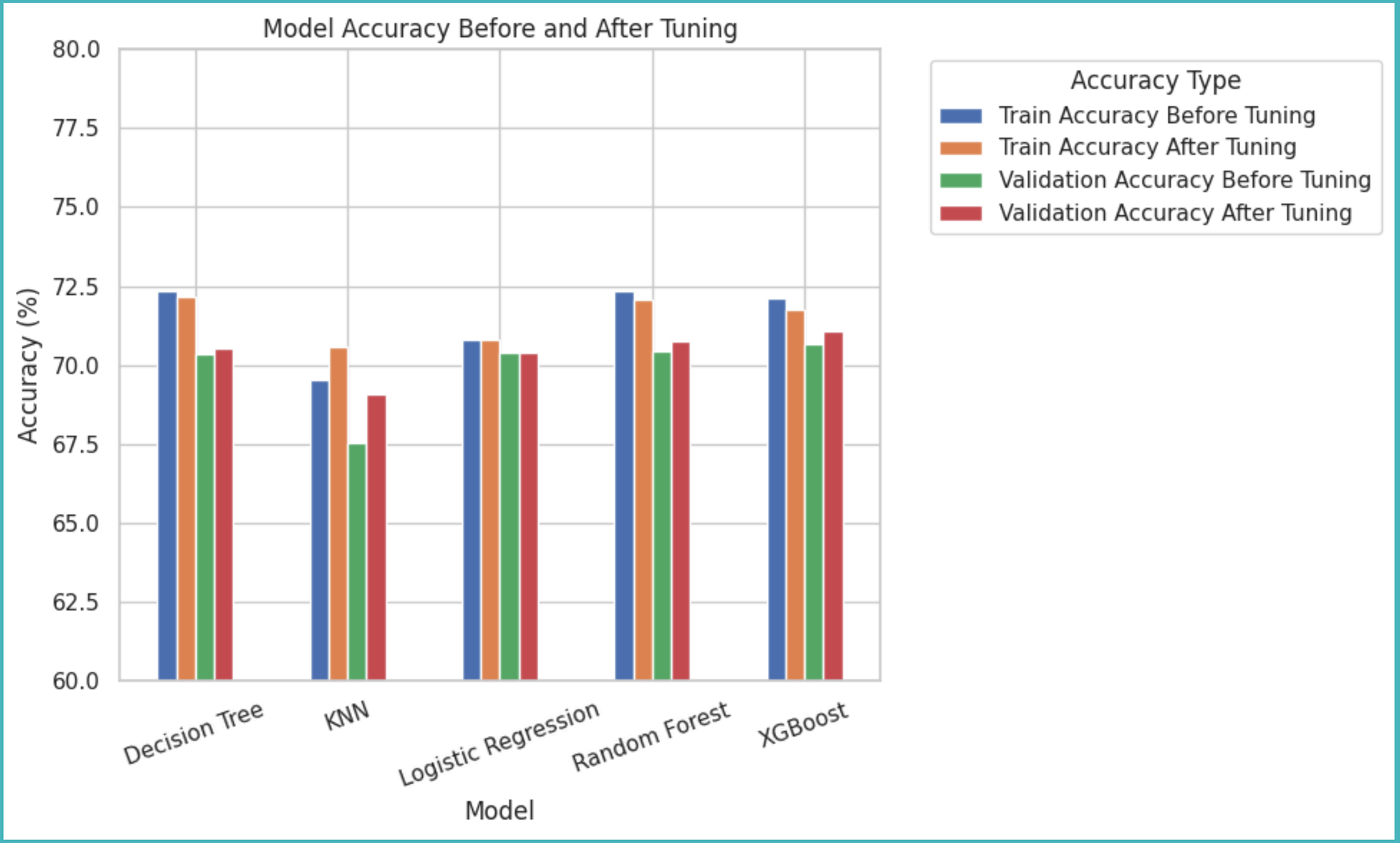
Relationship between Cardio and other variables after Feature engineering

Clustering & Segmentation



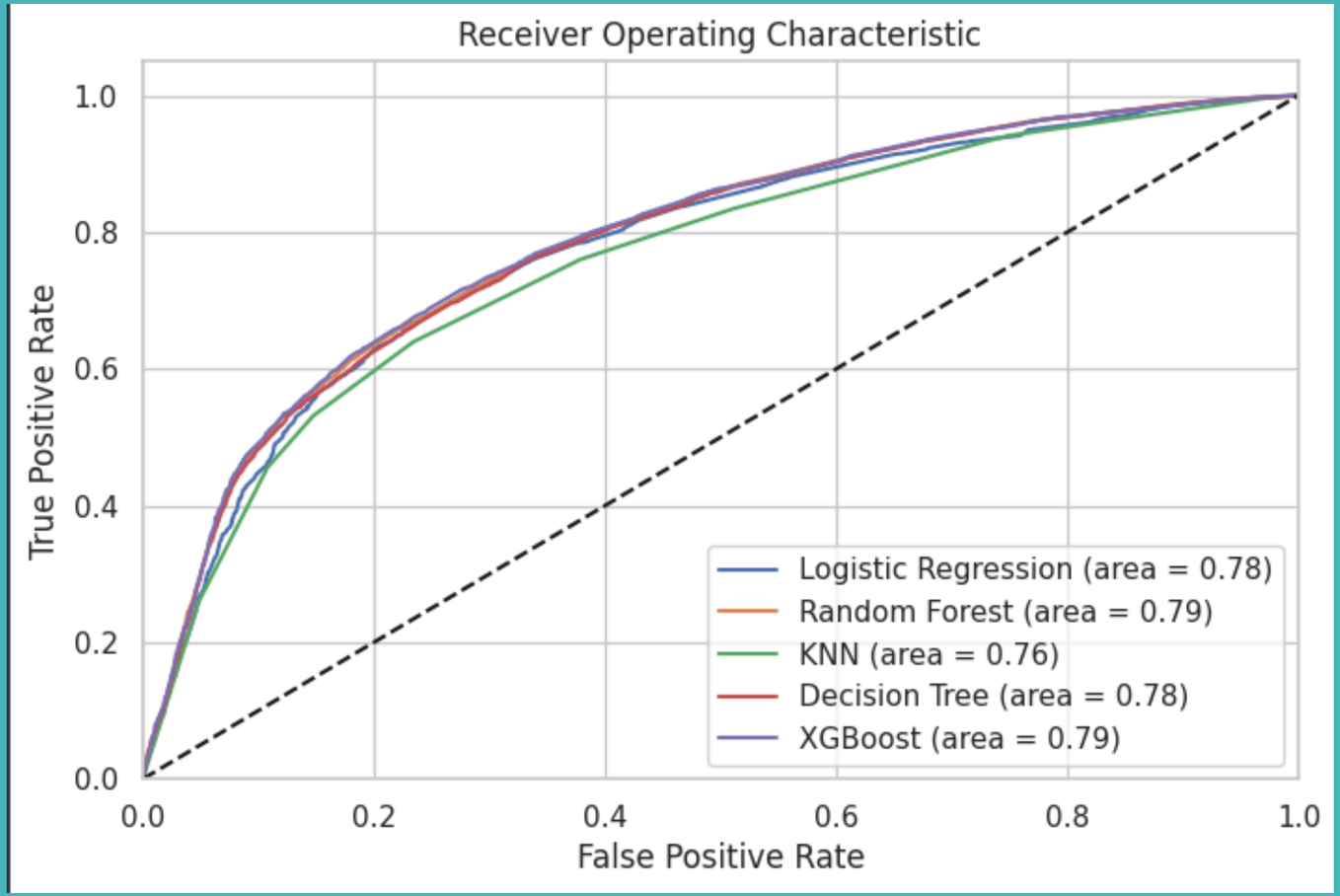
The figure above shows the distribution of CVD risk for each cluster

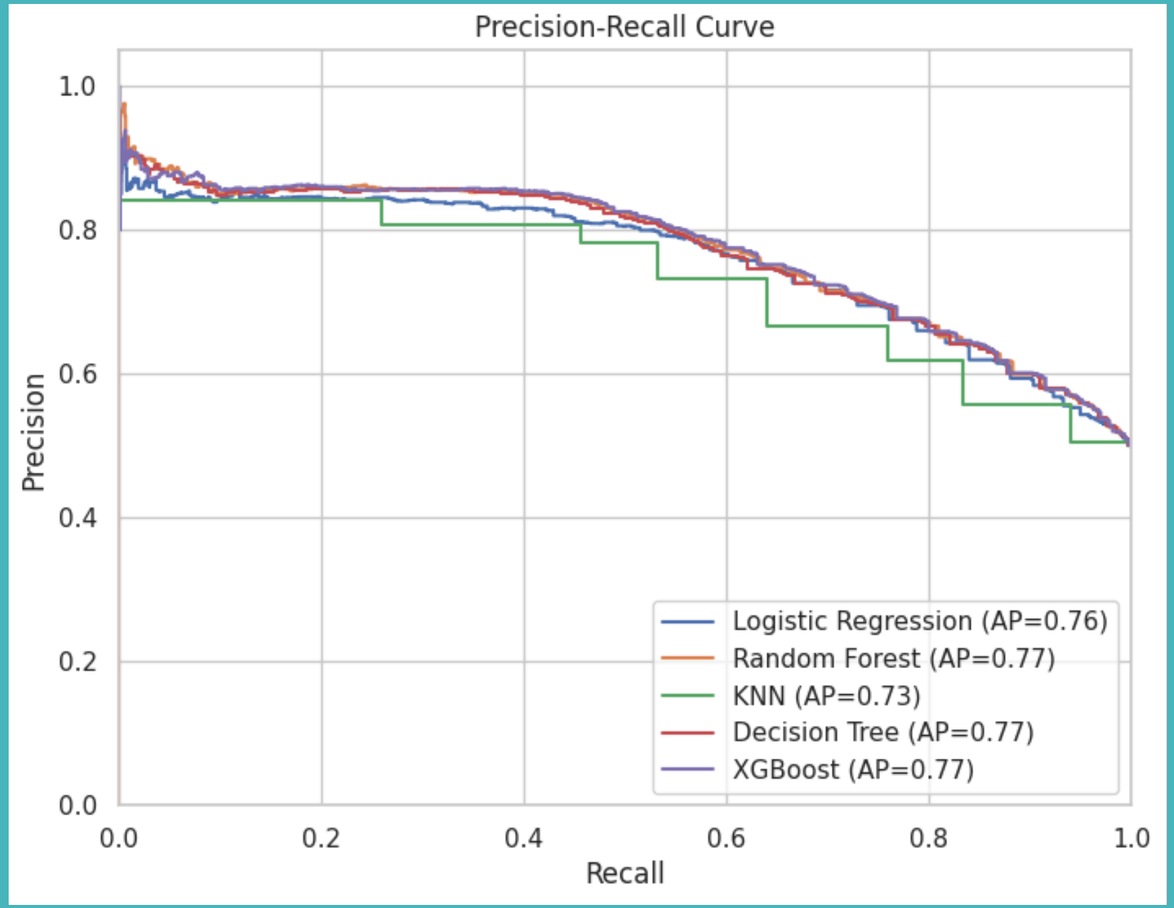
Modelling



Different modelling techniques used before and after tuning

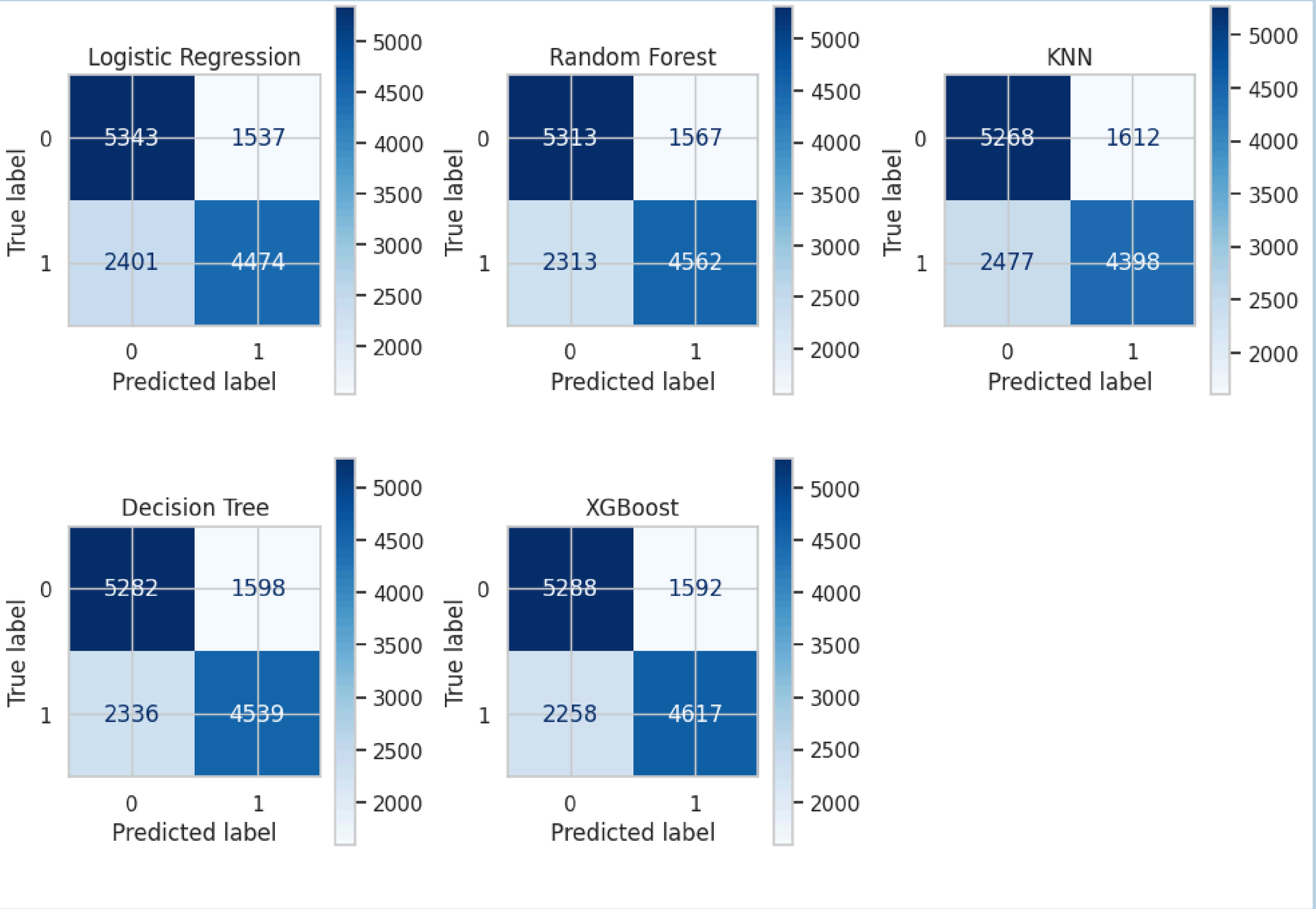
Model Evaluation



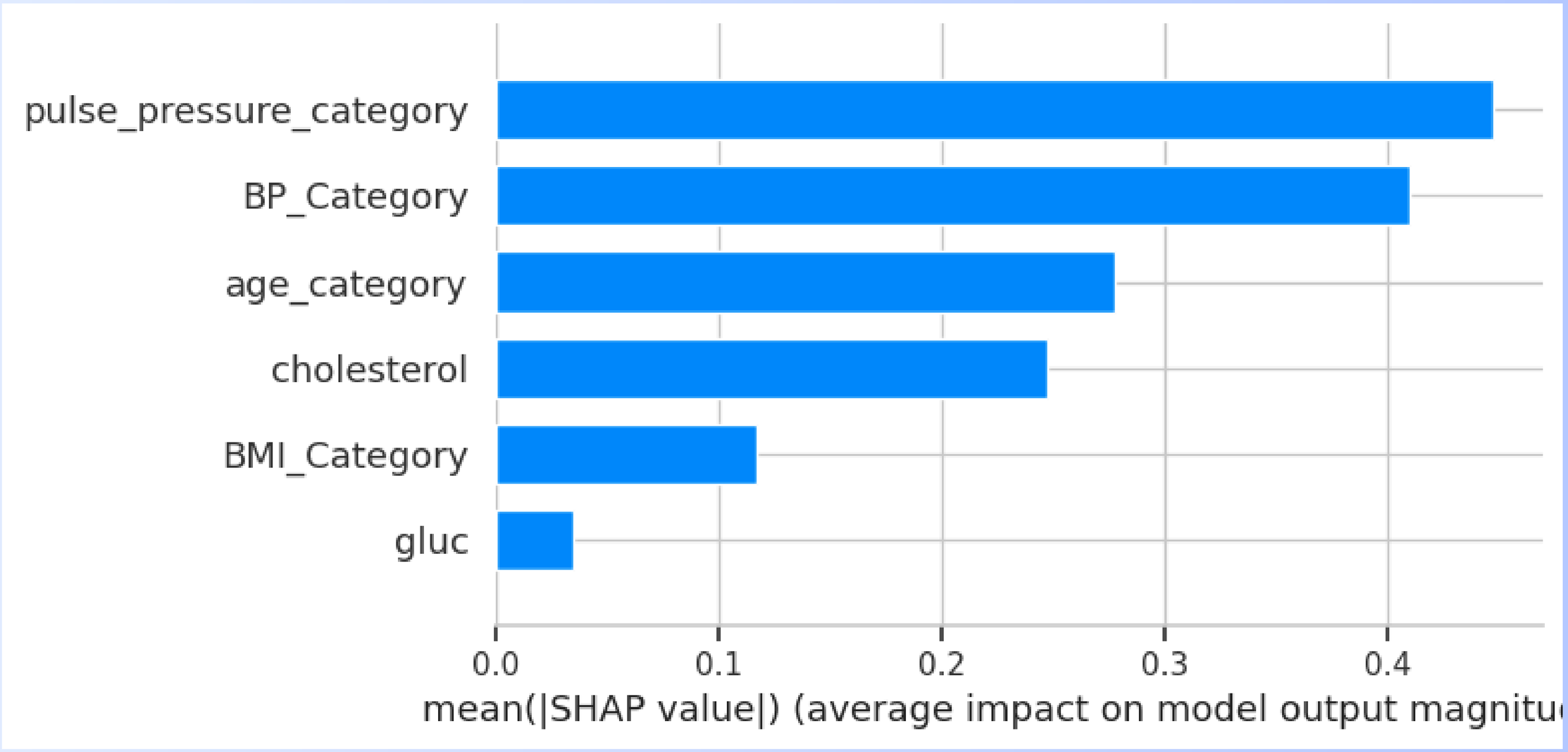


The models performed competitively with the highest performers being Random forest and XG boost

Confusion Matrices



Feature Impact Analysis



Recommendations

01 Enhancing Data Collection

02 Incorporating More Features

03 Addressing Bias and Fairness

04 Balancing Interpretability and Accuracy

05 Regular Validation and Updates

06 Risk Communication

Conclusions

- Improving Heart Disease Predictions with Combined Techniques
- End users and accuracy should be considered when choosing tools to apply in a model
- The predictions could be improved further by adding more data types
- Being Careful with Simple Methods
- Fine-Tuning Our Tools
- The Power of Good Information
- Choosing the Best Tools
- Ready to Help Patients

Thank you very much!

Moses Kigo	moseskigo@gmail.com
Josephine Wanjiru	josephinewanjiru28@gmail.com
Mercy Chepkemoi	chepkemoi465@gmail.com
Eunita Nyengo	nyengookoth@gmail.com
Eric Lekishon	ericlemiso@gmail.com