Predictive AI modeling on Coronary Heart Disease

BACKGROUND

Coronary heart disease is the most commonly diagnosed heart disease worldwide. Around 1 in 6 deaths globally are caused by coronary heart disease. The disease kills an estimated nine million people each year and it is the world's single biggest killer. Predicting coronary heart disease can assist in implementing preventive measures to reduce its severity and lower mortality rates.

METHODS

In my study I have used AI predictive algorithms to build a model that predicts the presence of Coronary Heart Disease.

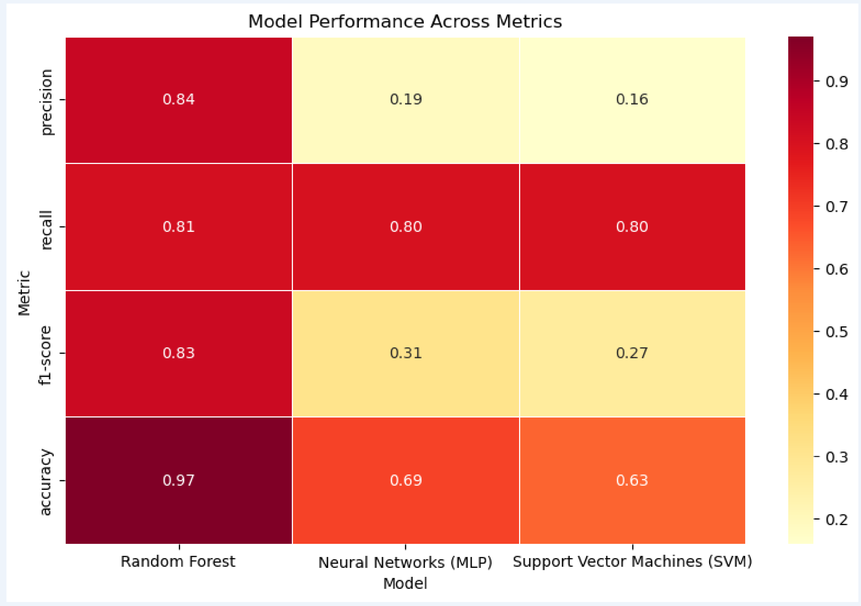
The best performing model predicts the disease with an accuracy of 97%, a recall of 81%, a precision of 84 % and a f1-score at 83%.

The validation dataset contains 15 selected features and more than 100 thousand selected datapoints. The data was preprocessed, investigated trough an exploratory data analysis, applied to different AI models, evaluated and visualised.

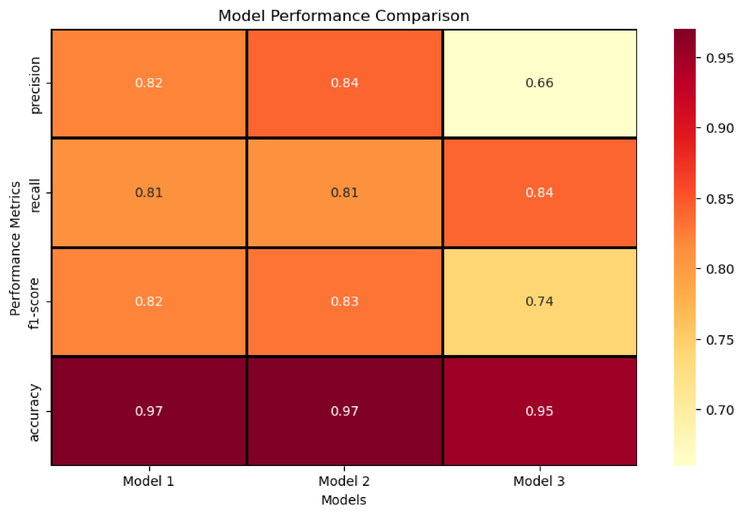
All data was taken from the behavioural risk factor surveillance system conducted by the centers for disease control and prevention in the USA. Each year more than 400 thousand adults are questioned, making it the largest continuously conducted health survey system in the world.

Further Research

Concerning further research on this topic I am interested in predicting Coronary Heart Disease with a smaller and different set of features and dig deeper into understanding the complex affect of features in the dataset. My code is open sourced and available on Github. I highly recommend you to take a look! It can be accessed trough the QR-Code down below.



This heatmap shows performance evaluation metrics of different tested machine learning algorithms, where a darker colour indicates a better performance. The Random Forest model had the best performance.



This heatmap shows performance evaluation metrics of different Random Forest algorithms, where a darker colour indicates a better performance. Model 2 had the best performance.