Project 4 Proposal.

I am very interested in the use of data for predicting disease in patients when the correct data is input into a machine learning model. My plan for this project is to use a cardiovascular dataset which takes a persons age, sex chest pain type, resting blood pressure, cholesterol, fasting blood sugar, resting electrocardiogram results, maximum heart rate achieved, exercised induced angina, oldpeak and the slope of the peak exercise. The data set if form Kaggle, and is compiled of over 900 individual datasets from patients from Hungarian Institute of Cardiology, University Hospital Zurich, University Hospital Basel and the V.A. Medical Centre Long Beach. The dataset can be found at the following address: <https://www.kaggle.com/datasets/fedesoriano/heart-failure-prediction>

I intend to first get familiar with the dataset, then clean and filter it so I have all the values presented in a way which will make it possible to apply machine learning to. The next step will include applying multiple machine learning methods to try and get the best accuracy score. This will allow me to then check again the data manually. Once I am happy with the machine learning model, I will start building a web app, along with a data base. This will allow users to input data and see if their risk of heart disease is evident. All the data will then be stored on a database, which will allow for future modelling, in which we can then make the model more accurate.