Heart disease prediction

The dataset I found is related to heart health with a total of 918 rows and 12 columns. It has the following columns:

* Age: age of Individual
* Sex: gender of individual (M, F)
* Chest Pain type: Type of chest pain experienced by the individual such as ATA (Atypical Angina), NAP (Non-Angina Pain), ASY (Asymptomatic) or other types.
* Resting Blood Pressure
* Cholesterol: Serum cholesterol in mg/dl
* Fasting Blood Sugar: Fasting blood sugar > 120 mg/dl. (1= True; 0=false)
* Resting ECG: Resting Electrocardiographic results (Values include ‘Normal’, ‘ST’ and other abnormalities)
* MaxHR: Maximum heart rate.
* Exercise Angina: Exercise-induced angina (Y= yes; N= no)
* Old Peak: ST depression induced by exercise relative to rest.
* ST\_Slope: The slope of the peak exercise ST segment (Values include ‘Up, ‘Flat’, ‘Down’)
* Heart disease: The presence of heart disease (1=yes; 0=no)

The columns don’t have any missing value. There are 5 object data types that I will convert to binary or int first. I can use this dataset to analyze or predict heart disease based on various factors and test results. I will use either classification or regression models after comparing the results such as accuracy for my dataset.

Resource

https://www.kaggle.com/code/pythonafroz/eda-heart-disease-prediction-roc-pr-curve/input