



# Cardio Guide

By AYUSH

## About Cardio Guide

**Cardio Guide is an application which uses Machine Learning Model to predict the chances of Heart Disease with an accuracy of 81.967%. With this it also provide you with tips to improve your health status which directly benefits your heart.**

**Heart diseases is a term covering any disorder of the heart. Heart diseases have become a major concern to deal with as studies show that the number of deaths due to heart diseases have increased significantly over the past few decades in India, in fact it has become the leading cause of death in India. Heart disease describes a range of conditions that affect your heart. Diseases under the heart disease umbrella include blood vessel diseases, such as coronary artery disease, heart rhythm problems (arrhythmias) and heart defects you're born with (congenital heart defects), among others.**

**Thus preventing Heart diseases has become more than necessary. Good data-driven systems for predicting heart diseases can improve the entire research and prevention process, making sure that more people can live healthy lives. This is where Machine Learning comes into play. Machine Learning helps in predicting the Heart diseases, and the predictions made are quite accurate.**

## Dataset

The dataset consists of 779 individuals data. There are 15 columns in the dataset, however the first column name is not a good parameter as far as machine learning is considered so, there are effectively 14 columns.

**Age :** displays the age of the individual.

**Sex :** displays the gender of the individual using the following format : 1 = male 0 = female.

**Chest-pain type :** displays the type of chest-pain experienced by the individual using the following format : 1 = typical angina 2 = atypical angina 3 = non - anginal pain 4 = asymptotic

**Resting Blood Pressure :** displays the resting blood pressure value of an individual in mmHg (unit)

**Serum Cholestrol :** displays the serum cholestrol in mg/dl (unit)

**Fasting Blood Sugar :** compares the fasting blood sugar value of an individual with 120mg/dl. If fasting blood sugar > 120mg/dl then : 1 (true) else : 0 (false)

**Resting ECG :** 0 = normal 1 = having ST-T wave abnormality 2 = left ventricular hyperthrophy

**Max heart rate achieved :** displays the max heart rate achieved by an individual.

**Exercise induced angina :** 1 = yes 0 = no

**ST depression induced by exercise relative to rest :** displays the value which is integer or float.

**Peak exercise ST segment :** 1 = upsloping 2 = flat 3 = downsloping

**Number of major vessels (0-3) colored by flourosopy :** displays the value as integer or float.

**Thal :** displays the thalassemia : 3 = normal 6 = fixed defect 7 = reversable defect

**Diagnosis of heart disease :** Displays whether the individual is suffering from heart disease or not : 0 = absence 1,2,3,4 = present.

## **Model Training & Prediction**

We can train our prediction model by analyzing existing data because we already know whether each patient has heart disease. This process is also known as supervision and learning. The trained model is then used to predict if users suffer from heart disease. The training and prediction process is described as follows:

First, data is divided into two parts using component splitting. In this experiment, data is split based on a ratio of 80:20 for the training set and the prediction set. The training set data is used in the logistic regression component for model training, while the prediction set data is used in the prediction component.

Logistic Regression classification model is used.

The two inputs of the prediction component are the model and the prediction set. The prediction result shows the predicted data, actual data, and the probability of different results in each group.

The confusion matrix, also known as the error matrix, is used to evaluate the accuracy of the model.