**Modules:**

**User module:**

User need to collect data set from kaggle website of 4 different counties which has supervised and unsupervised data. This dataset has features and labels which are used for prediction. User can load data using flask web framework and enter all symptoms related to heart disease and predict stages of heart disease.

**Preprocessing Module:**

This is preprocessing module where four datasets are converted to supervised learning data and then converted to single combined dataset. This dataset is used as input for application in the next for creating model.

**Model training Module:**

In this stage combined dataset is taken as in put and model is created using random forest classifier in three steps .

First data is dividing in to testing and training set and features and labels are extracted from these datasets and then data is trained and fitting is done. Then a pkl file is created which is model for this application.

This pkl file is used as model for predicting final results.

**Algorithm Random Forest Classifier:**

Step1: load dataset

Setp2: preprocessing data

Step3: load combined dataset

Step 4:split test and train set from dataset

Step 5: divide features and labels

Step 6: fit train and test set

Step 7: collect input data and give to predict function to get (0,1,2,3,4) values