HEART DISEASE PREDICTION

In order to make the dataset we are working with easier to understand, we will first take a look at it. It would enable us to make better use of the data. We have to import pandas, matplotlib and seaborn. We can now conduct exploratory data analysis after finishing data wrangling. The main tasks we will complete in this phase of our heart disease prediction project are listed below: A correlation search-To see the relationships between various variables, we will construct a correlation matrix. The goal of this research was to decrease the number of attributes that were used to diagnose heart disease. In the past, 13 attributes were used to make this prediction, but this research work used a genetic algorithm and feature subset selection to reduce the number of attributes to just six.

Natural evolution theory is incorporated into the Genetic Algorithm . The beginning population of the genetic search has randomly generated rules and no starting traits.

A new population was created to match the fittest rules in the existing population and any progeny of these rules based on the concept of survival of the fittest. By using the genetic operators of cross over and mutation, offspring were produced. The generation process lasted until a population P developed that satisfied all of its rules. In addition to the genetic algorithm, the CFS Evaluator is also used. Weka 3.6.0 tool is used to conduct the observations.909 records with 13 attributes made up the first data set .Consistencies were eliminated by categorising all attributes, and resolved for ease of use. after being reduced from 13 to 6 qualities. On the dataset, different classifiers are applied. matching these 6 characteristics of heart disease prediction. There is a performance analysis of these classifiers. It can be perceived from the table that Decision Tree has outperformed with highest accuracy and least mean absolute error.

DM Techniques	Accuracy	Model Construction Time	Mean Absolute Error
Naive Bayes	96.5%	0.02s	0.044
Decision Tree	99.2%	0.09s	0.00016
Classification via Clustering	88.3%	0.06s	0.117

For the purpose of predicting cardiac disease, various classifiers are used in conjunction with various data mining approaches. Observations show that the same classifier can sometimes provide

Varying accuracy for different data mining technique.