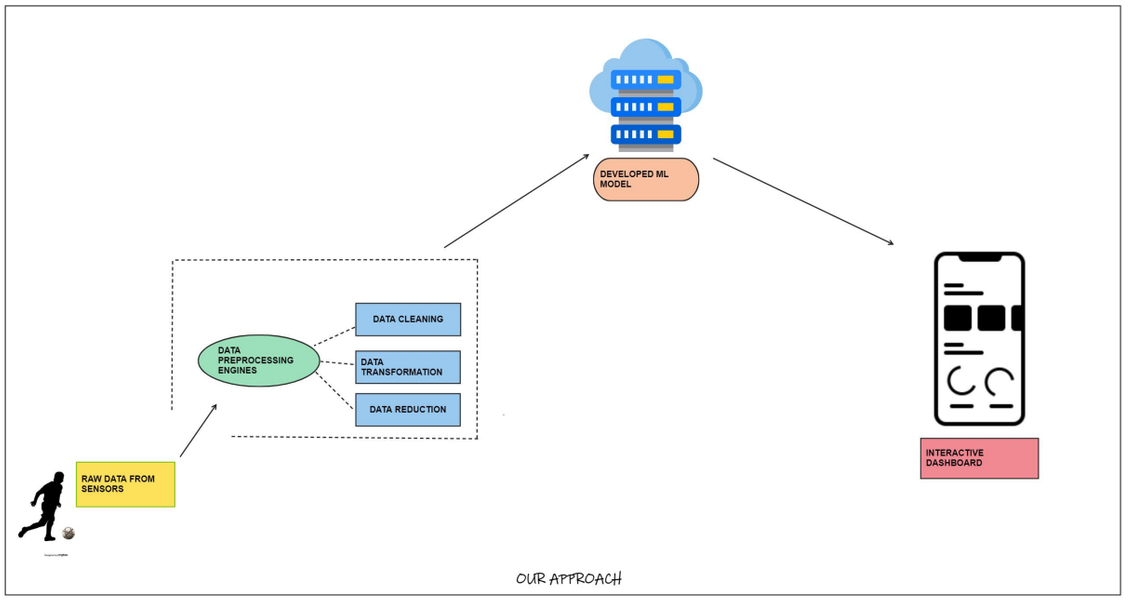
Visualizing and Predicting Heart Diseases with an Interactive Dash Board

Novelty of the solution:

Machine learning is used to diagnose, detect, and forecast many disorders in the medical industry. The primary purpose of this study is to give clinicians a tool to detect cardiac problems at an early stage. As a result, it will be easier to deliver appropriate treatment to patients while avoiding serious effects.In the system of the human heart, the heart's electrical activity is recorded by ECG with various wave forms through skin electrodes. Our approach begins with acquiring raw data through various sensors attached to the human body, let's say a fitness tracker watch ,the raw data is sent to the preprocessing engine where the data is structured and ready for further processing.The processed data is sent to the developed model for prediction ,after which the results are displayed in a interactive dashboard where the users can keep track of their body status.



Feasibility of idea:

I. Existing System:

Heart disease is even being highlighted as a silent killer which leads to the death of a person without obvious symptoms. The nature of the disease is the cause of growing anxiety about the disease & its consequences. Hence continued efforts are being done to predict the possibility of this deadly disease in prior. So that various tools & techniques are regularly being experimented with to suit the present-day health needs. Even though heart disease can occur in different forms, there is a

common set of core risk factors that influence whether someone will ultimately be at risk for heart disease or not.

II. Proposed Solution

The working of the system starts with the collection of data and selecting the important attributes. Then the required data is pre-processed into the required format. The data is then divided into two parts training and testing data. The algorithms are applied and the model is trained using the training data. The accuracy of the system is obtained by testing the system using the testing data.
This system is implemented using the following modules.

1.Collection of Dataset :

Initially, we collect a dataset for our heart disease prediction system. After the collection of the dataset, we split the dataset into training data and testing data. The training dataset is used for prediction model learning and testing data is used for evaluating the prediction model. For this project, 70% of training data is used and 30% of data is used for testing.

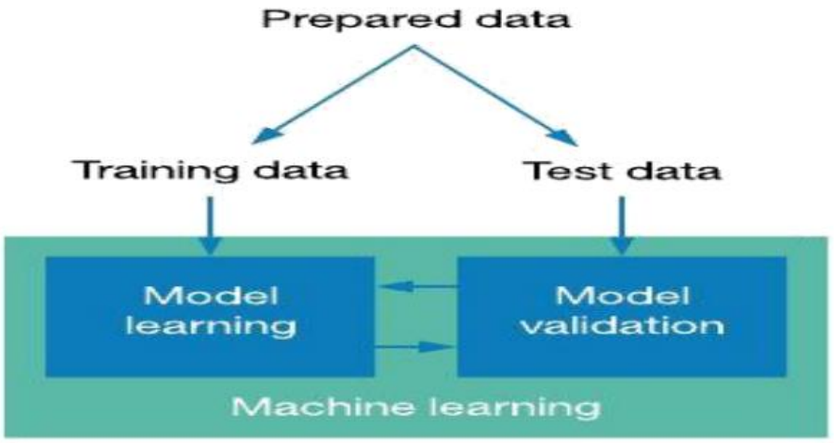


Figure: Collection of Data

2.Selection of attributes :

Attribute or Feature selection includes the selection of appropriate attributes for the prediction system. This is used to increase the efficiency of the system. Various attributes of the patient like gender, chest pain type, fasting blood pressure, serum cholesterol, exang, etc are selected for the prediction. The Correlation matrix is used for attribute selection for this model.

Figure: Correlation matrix

3.Data pre-processing:

Data pre-processing is an important step for the creation of a machine learning model. Initially, data may not be clean or in the required format for the model which can cause misleading outcomes. In pre-processing of data, we transform data into our required format.

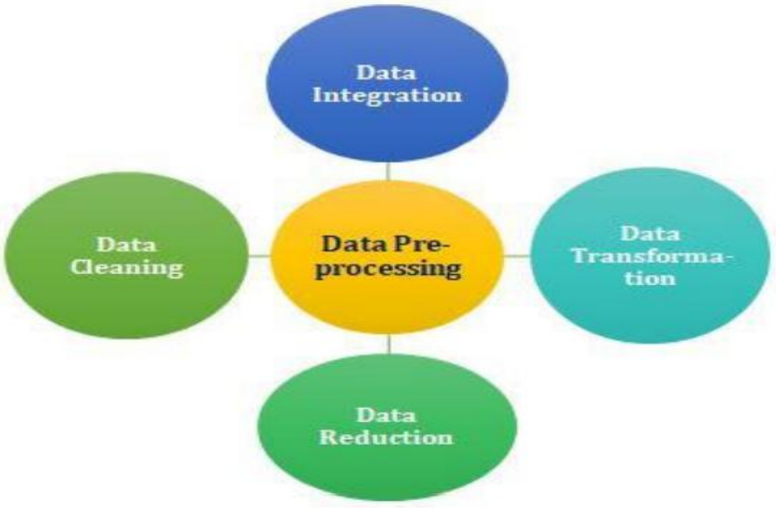


Figure: Data Pre-processing

4.Balancing of Data:

Imbalanced datasets can be balanced in two ways.

(a) Under Sampling: In Under Sampling, dataset balance is done by the reduction of the size of the ample class.This process is considered when the amount of data is adequate.

(b) Over Sampling: In Over Sampling, dataset balance is done by increasing the size of the scarce samples.This process is considered when the amount of data is inadequate.

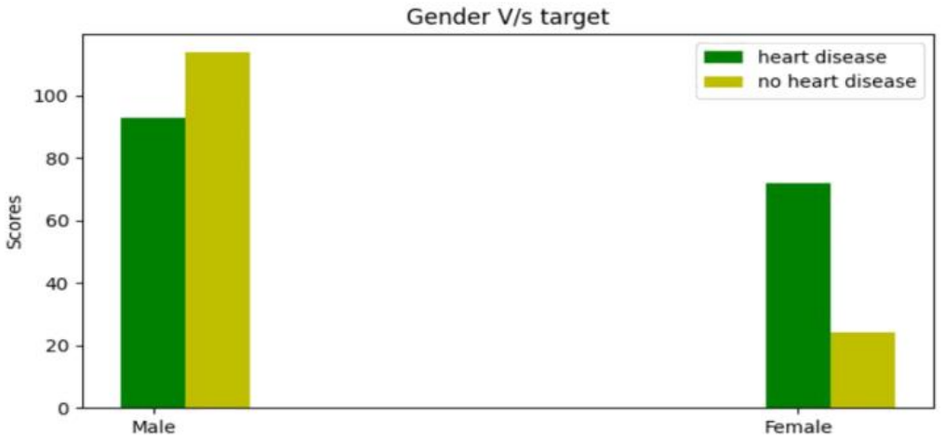


Figure: Data Balancing

5.Disease Prediction:

Various machine learning algorithms like SVM, Naive Bayes, Decision Tree, Random Tree, Logistic Regression, Ada-boost, Xg-boost are used for classification. Comparative analysis is performed among algorithms and the algorithm that gives the highest accuracy is used for heart disease prediction.

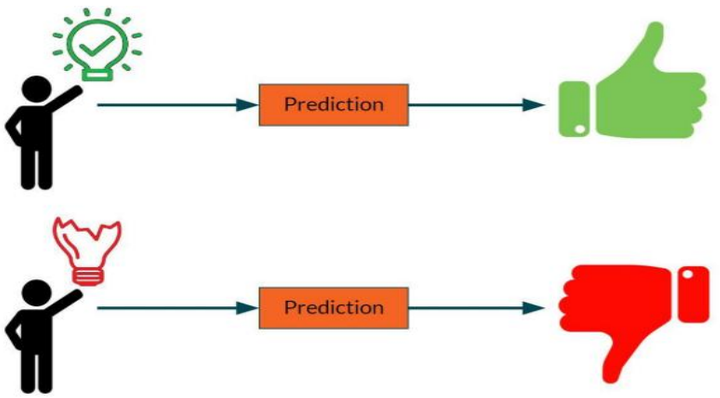
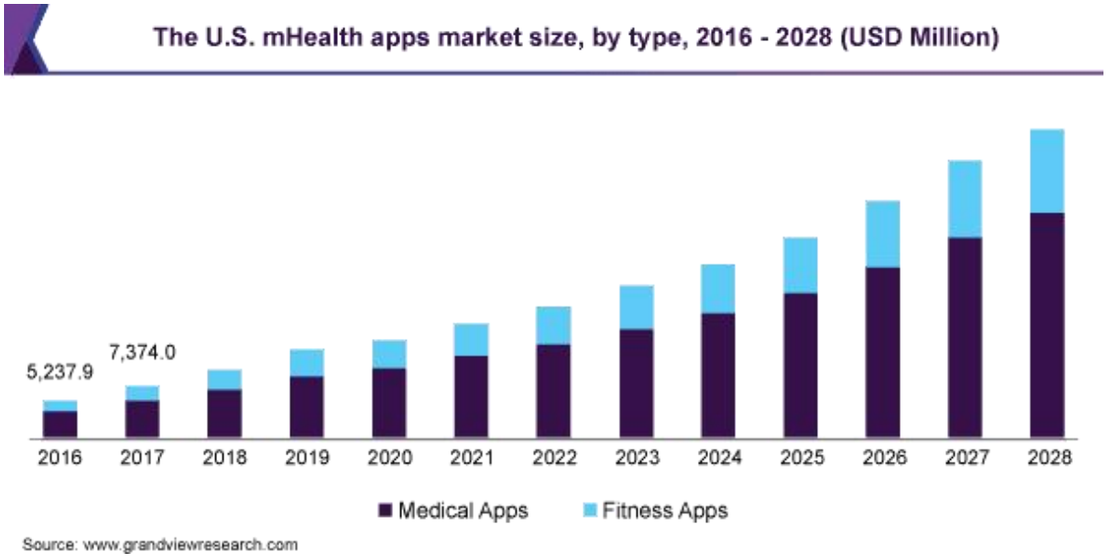


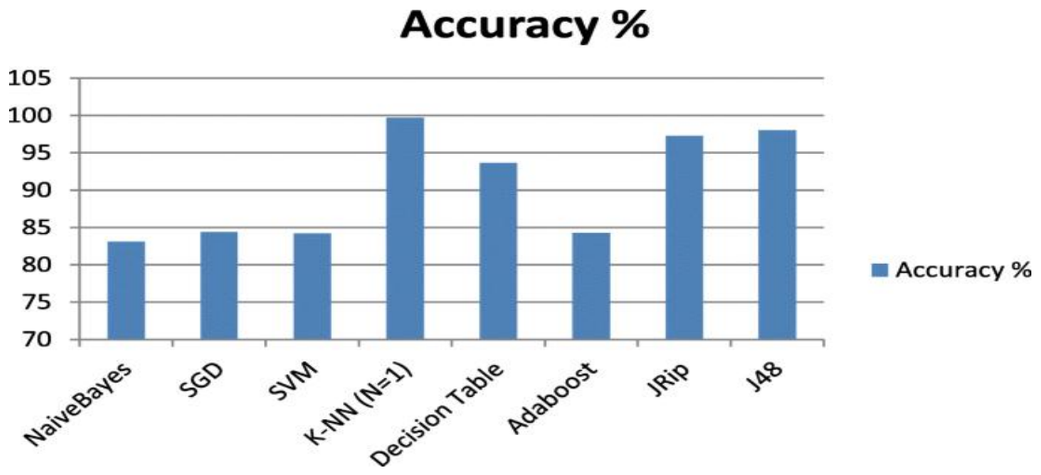
Figure: Prediction of Disease

Business Model:

In today's world we come across a lot of instances where large number of people are affected by heart diseases. Since there are large number of people who are affected by these heart disease it is very important to predict these diseases and inform them or caution them. The boon of machine learning has paved the way to predict these different types of heart diseases and segregate them based on the disease types and alert the user. Over the years, the usage of health application or different gadgets which focuses on health of the users have been promoted in a tremendous rate. The success rate of these application or gadgets have also increase tremendously.

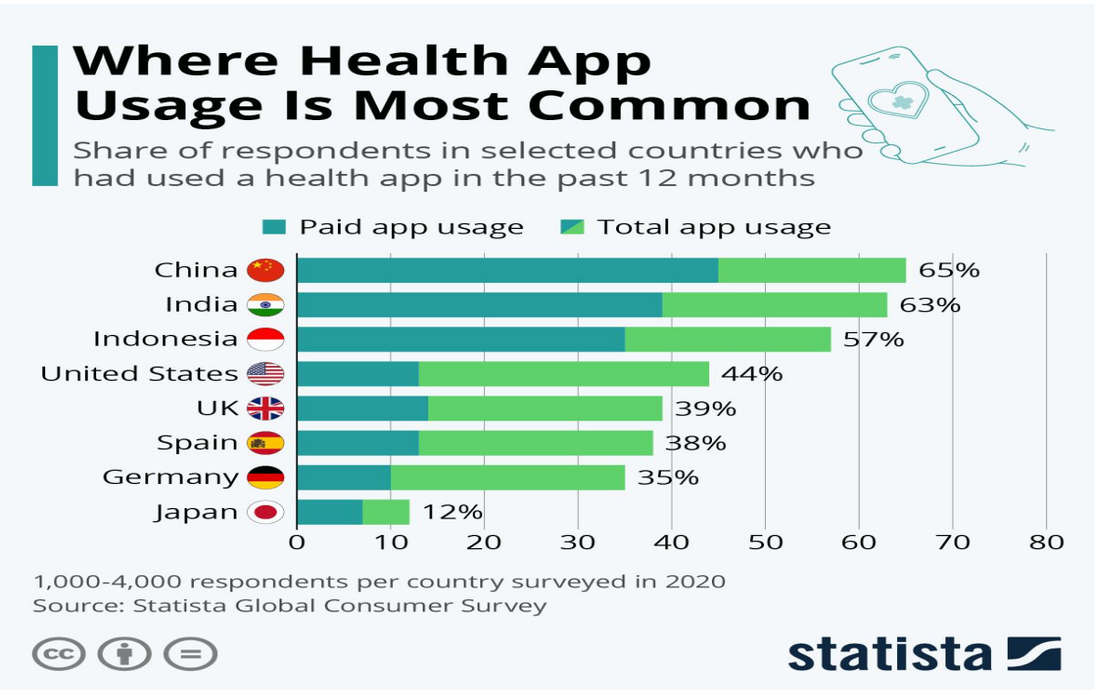


Nowadays, diseases are unpredictable in nature therefore it is important to know the current body condition. Therefore it is nowadays essential to keep a gadget or an application to keep in check of your medical status.This application will be widely accepted by the users as we are finding the relationship between variables which contribute to the efficiency of the prediction. We find the variables which contribute for efficient prediction of heart disease by classifying their types and their impact on the users body.



One of the major drawbacks of these prediction of this heart disease is the accuracy. Hence over the years developers used various algorithms to acquire maximum accuracy. Hence a good algorithm with good accuracy percentage is used. Our application uses the best algorithm to achieve maximum efficiency and gain user trust.

Over the years another major drawback is the usage of these health application. The usage of the health application or gadgets are low as the users don't completely trust the application or devices as its based on low efficiency algorithm in the past. But our application uses the best algorithm to predict the heart diseases and increases the efficiency by a large margin. Our application provides with trustable information and increases trust among the users.



Social Impact:

Heart diseases have become more and more frequent among people including our Country. Therefore, predicting the disease before becoming infected decreases the risk of death. Our system is part of the research on the detection and prediction of heart disease

A cardiovascular disease detection model has been developed using ML classification modelling techniques. This project predicts people with cardiovascular disease by extracting the patient medical history that leads to a fatal heart disease in a real time basis and from a dataset that includes patients' medical history such as chest pain, sugar level, blood pressure, etc.

This Heart Disease detection system assists a patient based on his/her clinical information of them been diagnosed with a previous heart disease by using these, computer aided techniques we can predict the patient fast and better and the cost can be reduced very much. The data is taken at real time for getting the updated medical values of the patient to diagnose the heart-based diseases in real time and with maximum accuracy

which swipe of a bunch of time and increases the chances of diagnosing the disease at an earlier stage and providing with its necessary treatment. There are also a number of medical databases that we can work on as these Machine learning techniques are better and provide ethnic support to the real time prediction and it will help to examine the patient's previous health records and they can predict better than a human being which helps the patient as well as the doctors.

Scalability of the Solution:

Scalability is a pretty straightforward concept.It is the property of a system to handle a growing amount of work by adding resources to the system.In order to stay at top positions in market the scalability and adaptability towards changing needs must be met.In our system , we focused on features that can make our application highly useful for the users.

1.Improved Accuracy on prediction:

Accuracy is one of the important factor that decides the success of the developed system.It is defined as the percentage of correct predictions for the test data.So we have analyzed many algorithms and improved the accuracy of our system by overcoming the cons of other algorithms that are used in heart disease prediction systems.We calculate it by dividing the number of correct predictions by the number of total predictions.Our System has the ability to handle non linear data as well with full efficiency.

2.Data security:

In today's world there are lots of data being generated in a period of time.The person with data are becoming richer.This in turn increases the data breaches across systems.The worst case can be exposure of a patients medical history in public.This can lead to lot of problems where medical sectors can misuse to promote their products to market in a short time.So to overcome that the data we collected and processed are encrypted with standards that cannot be broken.So this make our system to ensure privacy rights to the users.

3.Information sharing:

Making the data available in times need is important. We have enclosed a feature of sharing the patients record to authorized facilities in case of emergencies so that right medications can be given on time.This also ensures security of the patient's medical data.So these features make the system demanding in market.

4.User friendly dashboard:

The Dashboard is a dynamic User interface dashboard and this UI consist of several components which includes the details and real time medical values of the patients such as blood pressure, sugar levels, blood pumping rate, rate of blood flow etc. These values are taken and reflected in the dashboard at real time which makes it very efficient. The dashboard also provides statistics with the previous health records of the patient. The patient is provided with emergency support and specialized doctors to clarify their queries. The medical history of the patients, which includes the treatments done for the patient and the location of the treatment done and their dates and timings.

****