**MULTIPLE DISEASE PREDICTION**

**Abstract:**

Disease Prediction using Machine Learning is the system that is used to predict the diseases from the symptoms which are given by the patients or any user input. The system processes the symptoms provided by the user as input and gives the output as the probability of the diseases. The prediction system predicts chance of presence of a disease present in a patient on the basis of their symptoms.

**Introduction:**

Nowadays, humans face various diseases due to the current environmental condition and their living habits. The identification and prediction of such diseases at their earlier stages are much important, so as to prevent the extremity of it.so in this we are going to use some supervised machine learning algorithms to predict .

Machine learning model predictions allow business to make highly accurate guesses as to the likely outcomes of a question based on historical data, which can be about all kinds of things customer churn likelihood, possible fraudulent activity and more.

it gives enterprises a view of trends in customer behavior and business operational patterns, as well as supports the development of new products.

**Implementation:**

we are implemented the prediction of some diseases they are :

* Kidney-Disease prediction
* Liver-disease prediction
* Heart disease prediction
* Diabetes disease prediction
* Cancer disease prediction

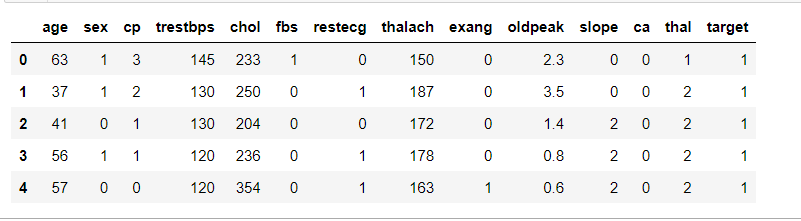
**Algorithms we used :**

* Logistic Regression
* Random forest Classification

we have used these two algorithms for predicting diseases and logistic regression for heart disease prediction and for remaining Diseases Random forest classification

1.Logistic regression:

Logistic regression is an supervised learning algorithm . It is used to calculate or predict the probability of yes/no event occurring we used it for heart disease prediction for following Dataset



2. Random Forest Classification

Random Forest is a supervised machine learning algorithm that grows and combines multiple decision trees to create a forest. It can be used for both classification and regression. Random forest algorithm builds a forest in the form of an ensemble of decision trees which adds more randomness while growing the trees. While splitting a node, the algorithm searches for the best features from the random subset of features which adds more diversity, thereby resulting in a better model.

We have applied random forest classification for remaining four diseases and one of the datasets we applied is

Graphical user interface, application

Description automatically generated

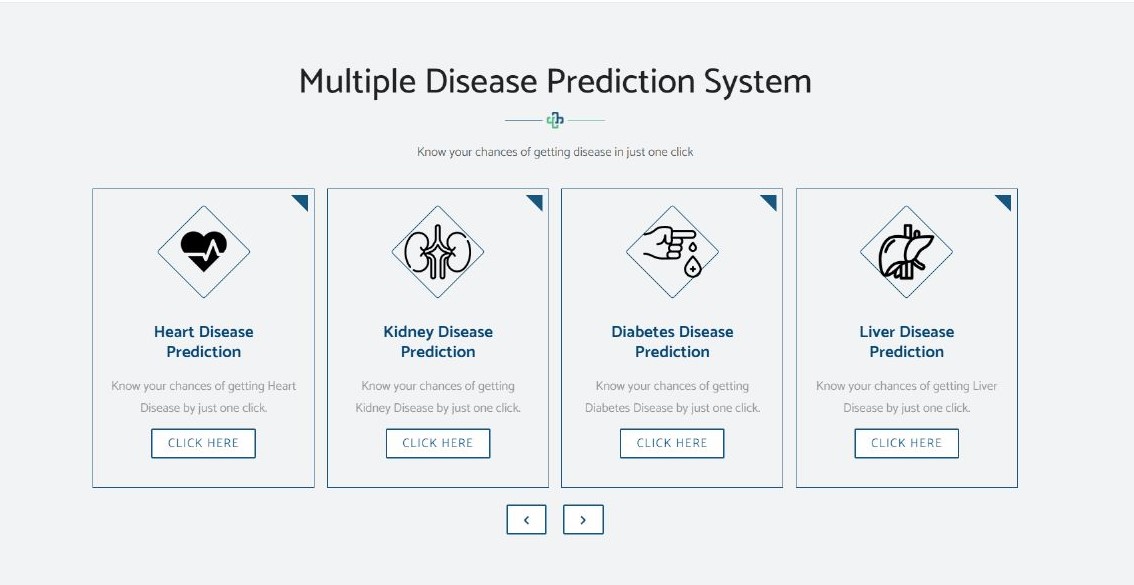
How we have dealing with data shown below

* Initially we imported data and print some rows
* Then checked for all the datatypes of attributes and null count by using info method and
* Data preprocessing :
* Checking for null values and outliers if any missing values is there did imputation by using some imputation techniques
* If it is numerical values we replaced by mean or median otherwise if it is categorical data replaced by mode of that attribute and
* Handled outliers and visualized by using boxplots and did data exploration analysis
* Separated Data,Splitting data into traing and test dataset
* And fitted the data applied techniques for it
* Predicted the output based on test data and

The process how the Predicting output by user input have shown below

**Homepage:**

Initially user input page look like this here the user can choose any prediction for disease ,if the person chosen Diabetes disease prediction it will goes to that diabetes disease page and next step shown below.

**API link:** https://multiple-disease-predictions.herokuapp.com 

**Input page:**

Here the user have to give some values or symptoms needed to predict that is Bloodpressure, insulin, Glucose level…and submit the details by hitting predict button shown below

Graphical user interface, text

Description automatically generated with medium confidence

**Output** : we will get message like this based on the input given by user.

A picture containing plant, bird

Description automatically generated

**Result:**

The accuracy got for every model is shown below

* Kidney-Disease prediction

Graphical user interface, text, application, email

Description automatically generated

* Liver-disease prediction

Graphical user interface, text, application, email

Description automatically generated

* Heart disease prediction

Graphical user interface, text, application, email

Description automatically generated

* Diabetes disease prediction



* Cancer disease prediction

Graphical user interface, text, application, chat or text message

Description automatically generated

**References:**

**1.Project guide:** Dr.P. Bhagat (IIT) , Associate Professor, Department of Computer science and Engineering ,LBRCE, Mylavaram.

**2.Textbook:** Machine learning , authors -saikatdutt ,Subramaninan chandramouli,Amit kumar Das Pearson india education services pvt.Ltd.