International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | Impact Factor: 7.488 |

|| Volume 8, Issue 5, May 2020 ||

Prediction of Cardiovascular Disease

Vaishali Baviskar[1], Bhakti Nalawade[2] ,Vaishnavi Patil[3] , Gururaj Rekhawar[4], Shubham Mahulkar[5]

Assistant Professor, Department of Computer Engineering, G.H Raisoni Institute of Engineering and Technology, Wagholi, Pune, Maharashtra, India[1]

B.E. Student, Department of Computer Engineering, G.H Raisoni Institute of Engineering and Technology, Wagholi, Pune, Maharashtra, India[2,3,4,5]

ABSTRACT: Healthcare is an inevitable task to be done in human life. Cardiovascular disease is abroad category for a range of diseases that are affecting heart and blood vessels. The early methods of forecasting the cardiovascular diseases helped in making decisions about the changes to have occurred in high -risk patients which resulted in the reduction of their risks. The health care industry contains lots of medical data, therefore machine learning algorithms are required to make decisions effectively in the prediction of heart diseases. Recent research has delved into uniting these techniques to provide hybrid machine learning algorithms. In the proposed research , data pre-processing uses techniques just like the removal of noisy data , removal of missing data , filling default values if applicable and classification of attributes for prediction and deciding at different levels . The performance of the diagnosis model is obtained by using methods like classification , accuracy , sensitivity and specificity analysis . This project proposes a prediction model to predict whether a people have a heart disease or not and to provide an awareness or diagnosis on that .

KEYWORDS: Machine learning, Neural network, Prediction, Classification

I. INTRODUCTION

It is difficult to identify heart disease because of several contributory risk factors such as diabetes, high blood pressure, high cholesterol, abnormal pulse rate and many other factors. Various techniques in data mining and neural networks have been employed to and out the severity of heart disease among humans- Genetic algorithm (GA), and Naive Bayes (NB). The nature of heart disease is complex and hence, the disease must be handled carefully. Not doing so may affect the heart or cause premature death. The perspective of medical science and data mining are used for discovering various sorts of metabolic syndromes. We have also seen decision trees be used in predicting the accuracy of events related to heart disease various methods have been used for knowledge abstraction by using known methods of data mining for prediction of heart disease. In this work, numerous readings have been carried out to produce a prediction model using not only distinct techniques but also by relating two or more techniques. These amalgamated new techniques are commonly known as hybrid methods. The dataset with a radial basis function network (RBFN) is used for classification, where 70% of the daily used for training and the remaining 30% is used for classification. We also introduce Computer Aided Decision Support System (CADSS) in there of medicine and research. In previous work, the usage of data mining techniques in the healthcare industry has been shown to take less time for the prediction of disease with more accurate results . We propose the diagnosis of heart disease using the GA. This method uses effective association rules inferred with the GA for tournament selection, crossover and the mutation which results in the new proposed fitness function For experimental validation, we use the well-known Cleveland dataset which is collected from a UCI machine learning repository. We will see later on how our results prove to be prominent when compared to some of the known supervised learning techniques. The most powerful evolutionary algorithm Particle Swarm Optimization (PSO) is introduced and some rules are generated for heart disease. pulse rate, sex, age, and many others. The ML algorithm with Neural Networks is introduced, whose results are more accurate and reliable. Healthcare is an inevitable task to be done in human life. Cardiovascular disease is a broad category for a range of diseases that are affecting heart and blood vessels.

II. RELATED WORK

[1]Heart disease is one of the most significant causes of mortality in the world today. Prediction of cardiovascular disease is a critical challenge in the area of clinical data analysis. Machine learning (ML) has been shown to be effective in assisting in making decisions and predictions from the large quantity of data produced by the healthcare industry. We have also seen ML techniques being used in recent developments in different areas of the Internet of Things (IoT). Various studies give only a glimpse into predicting heart disease with ML techniques. In this paper, we

propose a novel method that aims at finding significant features by applying machine learning techniques resulting in improving the accuracy in the prediction of cardiovascular disease. The prediction model is introduced with different combinations of features and several known classification techniques. We produce an enhanced performance level with an accuracy level of 88.7% through the prediction model for heart disease with the hybrid random forest with a linear model (HRFLM).

[2]In this paper, heart disease prediction modeled using partially observable markov decision process (POMDP) is proposed. In emergency, the patient is alerted through the doctor by fog computing. Ambulance sent to the location of patient at critical situations. The doctor gets the data through fog computing iFogSim. Fog computing in healthcare is a new area, which gains more attraction in research community. Many researches focus on cardiovascular disease i.e. heart disease. The important risk factor for cardiovascular disease is increase in blood viscosity. The highly viscous nature of blood does not allow the blood to flow creating a resistance in the blood flow. Heart disease risk factors are high blood pressure, obesity, diabetes, increased blood viscosity, etc. With the help of POMDP's states, observations, beliefs, probability transitions the patient health is noted. The POMDP model for heart disease prediction computes the policy approximation using states and timeslots. Rewards are tabulated using policy approximations over different iterations.

[3]The main severe reason for unnatural death over worldwide, as well as Asian nation especially in India is the coronary failure diseases associated potential by detection of any particular cardiovascular disease at an earlier stage can forestall these attacks. Those who are associated in practicing medical it's generate knowledge with a wealth of large and very useful data gift, and it's not used effectively for predictions. For this reason, the analysis converts the knowledge into a dataset for shaping exploitation totally different data processing algorithms. There's a demand for practitioners of medical to predict cardio diseases even way before when their patients is prone to any CVD. The options that increase the probabilities of the heart attacks are drinking and smoking, lack of physical activities, high level of vital sign, dangerous extent of cholesterol levels, unhealthy and unhygienic diet, damaging use of alcoholic beverage, and high sugar content level food. Cardio vascular diseases (CVD) constitutes coronary disease heart, vas or Stroke, and other hypertensive cardio disease, innate heart, other peripheral artery causes, rheumatic cardio disease, and inflammatory type cardio disease. Prophecies associated descriptions are principal goals of information mining; in observe Prediction of the processed data involves the attributes or the physiological variables in the data set to find a prediction probability or future state values of an alternative attributes. The Description will emphasize on the discovering a common recognizable patterns that describes that information that can be understood by humans.

III. PROPOSED SYSTEM

The proposed system is GUI-based, user-friendly, scalable, reliable and an expandable system. The proposed working model also can help in reducing treatment costs by providing Initial diagnostics in time. The model also can serve the aim of coaching tool for medical students and can be a soft diagnostic tool available for physician and cardiologist. General physicians can utilize this tool for initial diagnosis of cardio-patients.The Heart Disease Prediction application is an end user support and online consultation project. Here, we propose an internet application that permits users to urge instant guidance on their heart condition through an intelligent system online. The application is fed with various details and therefore the heart condition related to those details. The application allows user to share their heart related issues. It then processes user specific details to see for various illness that would be related to it. Here we use some intelligent data processing techniques to guess the foremost accurate illness that would be related to patient's details.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | Impact Factor: 7.488 |

|| Volume 8, Issue 5, May 2020 ||

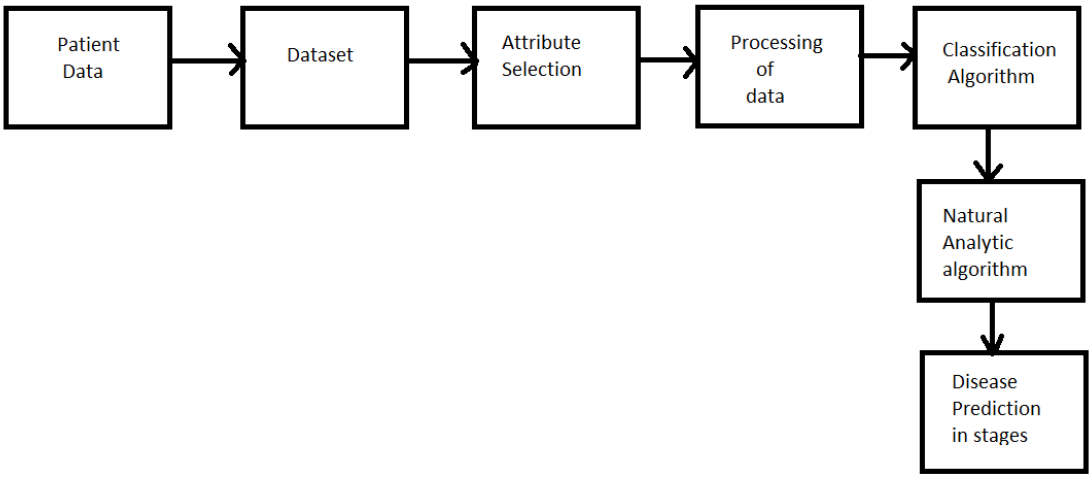

Figure 1: System architecture

IV. MATHEMATICAL MODEL AND ALGORITHM

1. Bayes' Theorem finds the probability of an event occurring given the probability of another event that has already occurred. Bayes' theorem is stated mathematically as the following equation:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

where A and B are events and P(B) ? 0.

- Basically, we are trying to find probability of event A, given the event B is true. Event B is also termed as **evidence**.
- P(A) is the **priori** of A (the prior probability, i.e. Probability of event before evidence is seen). The evidence is an attribute value of an unknown instance(here, it is event B).
- P(A|B) is a posteriori probability of B, i.e. probability of event after evidence is seen.

Now, with regards to our dataset, we can apply Bayes' theorem in following way:

$$P(y|X) = \frac{P(X|y)P(y)}{P(X)}$$

where, y is class variable and X is a dependent feature vector (of size *n*) where:
X=(x1, x2, x3,……., x_n)

Now, its time to put a naive assumption to the Bayes' theorem, which is, **independence** among the features. So now, we split **evidence** into the independent parts.
Now, if any two events A and B are independent, then,
P(A,B) = P(A)P(B)
Hence, we reach to the result:

$$P(y|x_1, ..., x_n) = \frac{P(x_1|y)P(x_2|y)...P(x_n|y)P(y)}{P(x_1)P(x_2)...P(x_n)}$$

2. Random forest is a supervised learning algorithm which is used for both classification as well as regression. But however, it is mainly used for classification problems. As we know that a forest is made up of trees and more trees means more robust forest. Similarly, random forest algorithm creates decision trees on data samples and then gets the prediction from each of them and finally selects the best solution by means of voting. It is an ensemble method which is better than a single decision tree because it reduces the over-fitting by averaging the result.
   Step 1 − First, start with the selection of random samples from a given dataset.
   Step 2 − Next, this algorithm will construct a decision tree for every sample. Then it will get the prediction result from every decision tree.
   Step 3 − In this step, voting will be performed for every predicted result.

Step 4 − At last, select the most voted prediction result as the final prediction result.

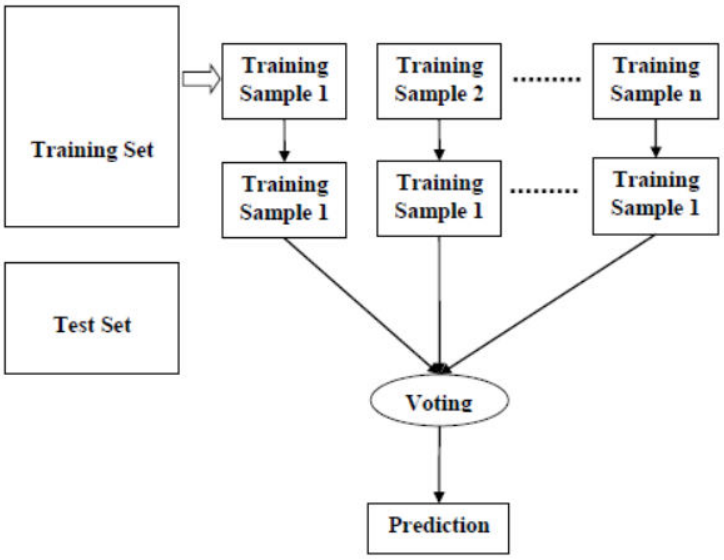

Fig 2. Random forest tree

V. RESULT

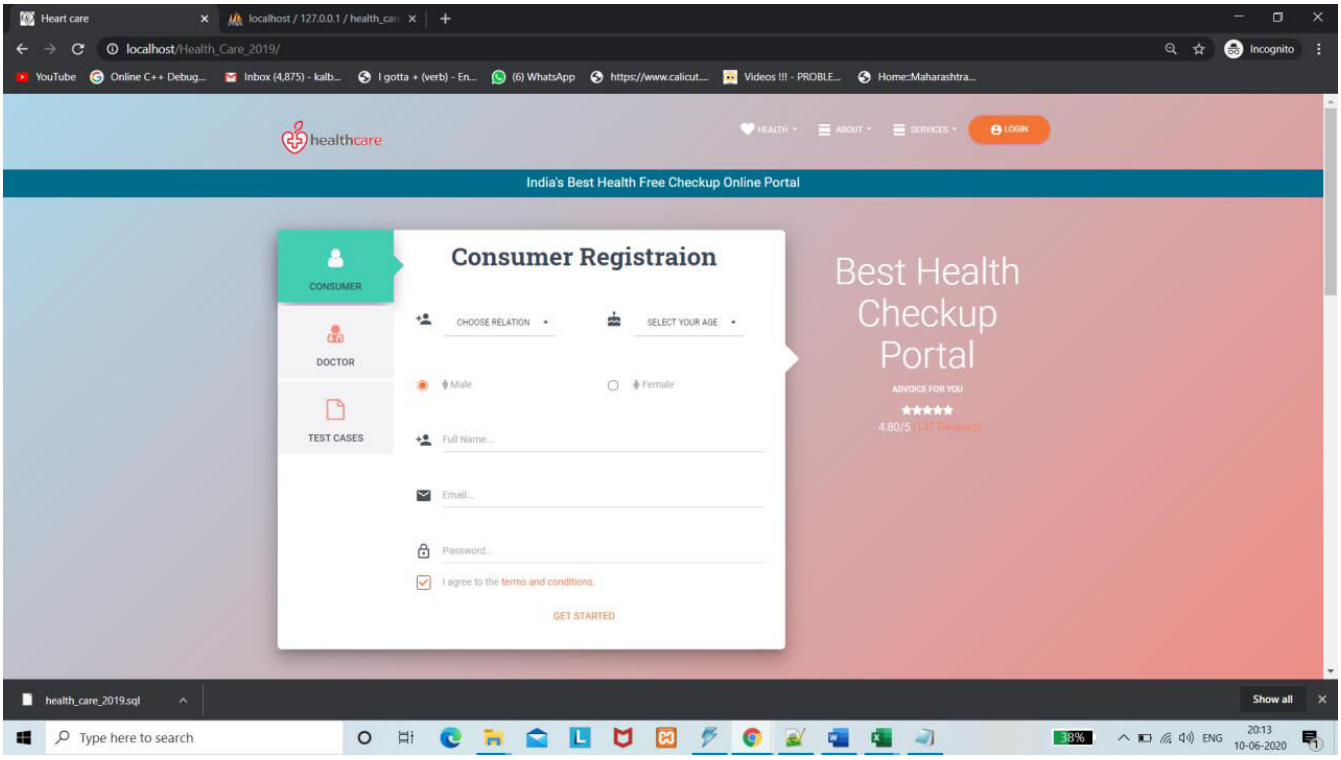

Fig 3: user registration

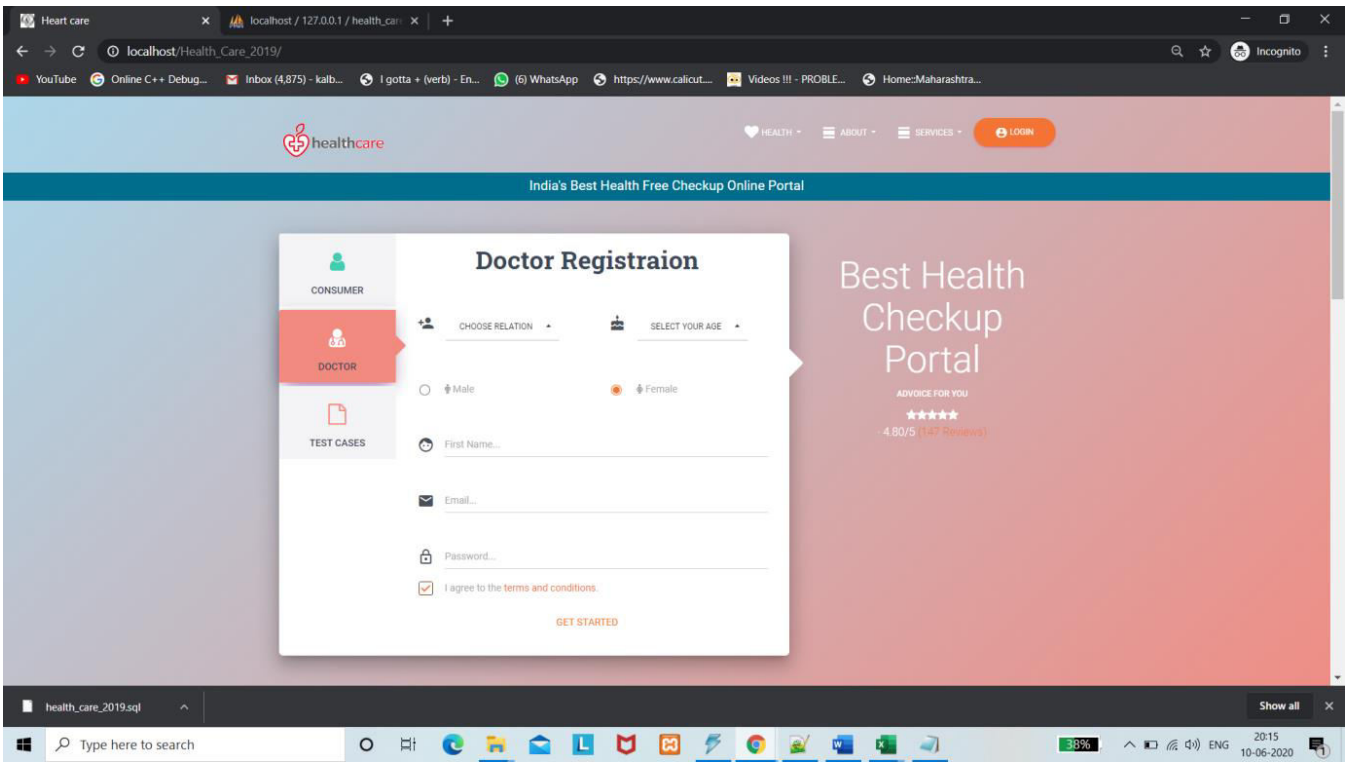

Fig 4: Doctor Registration

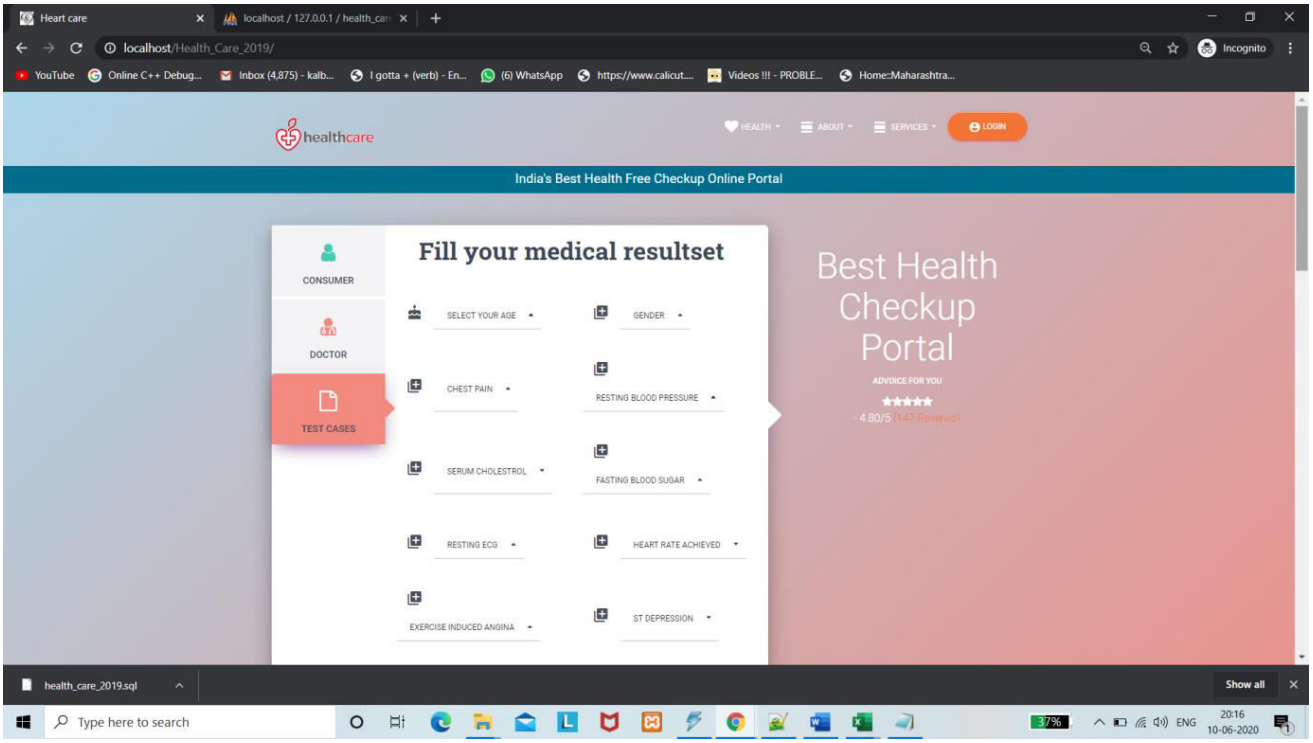

Fig 5: user fill the medical resultset to predict symptoms

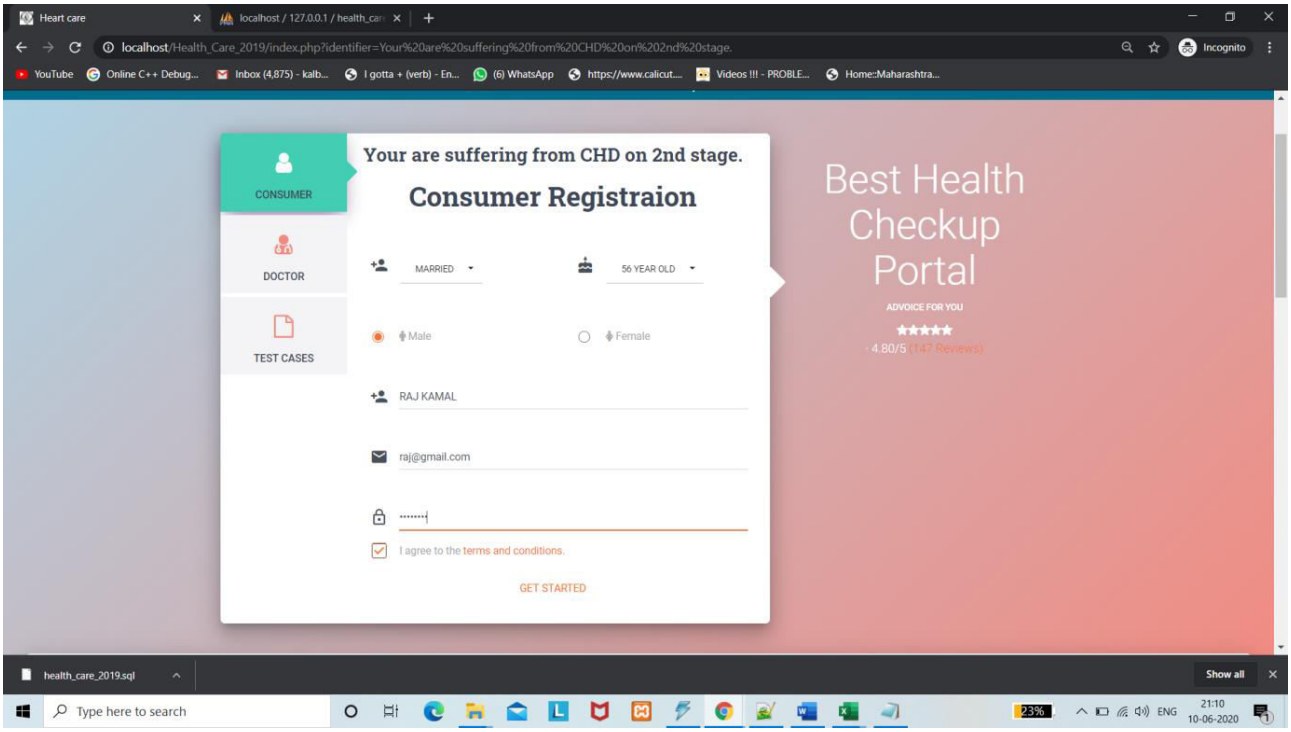

Fig 6: User registration after predicted symptoms for further evaluation

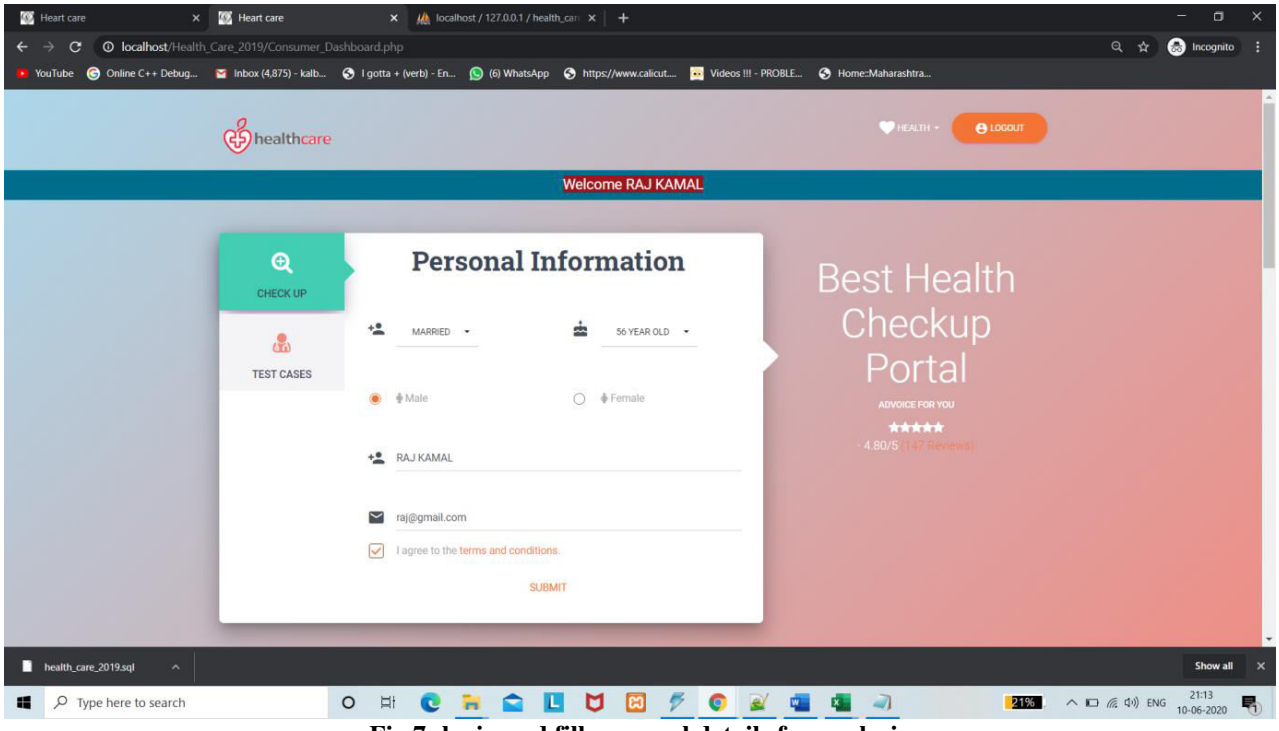

Fig 7: login and fill personal details for analysis

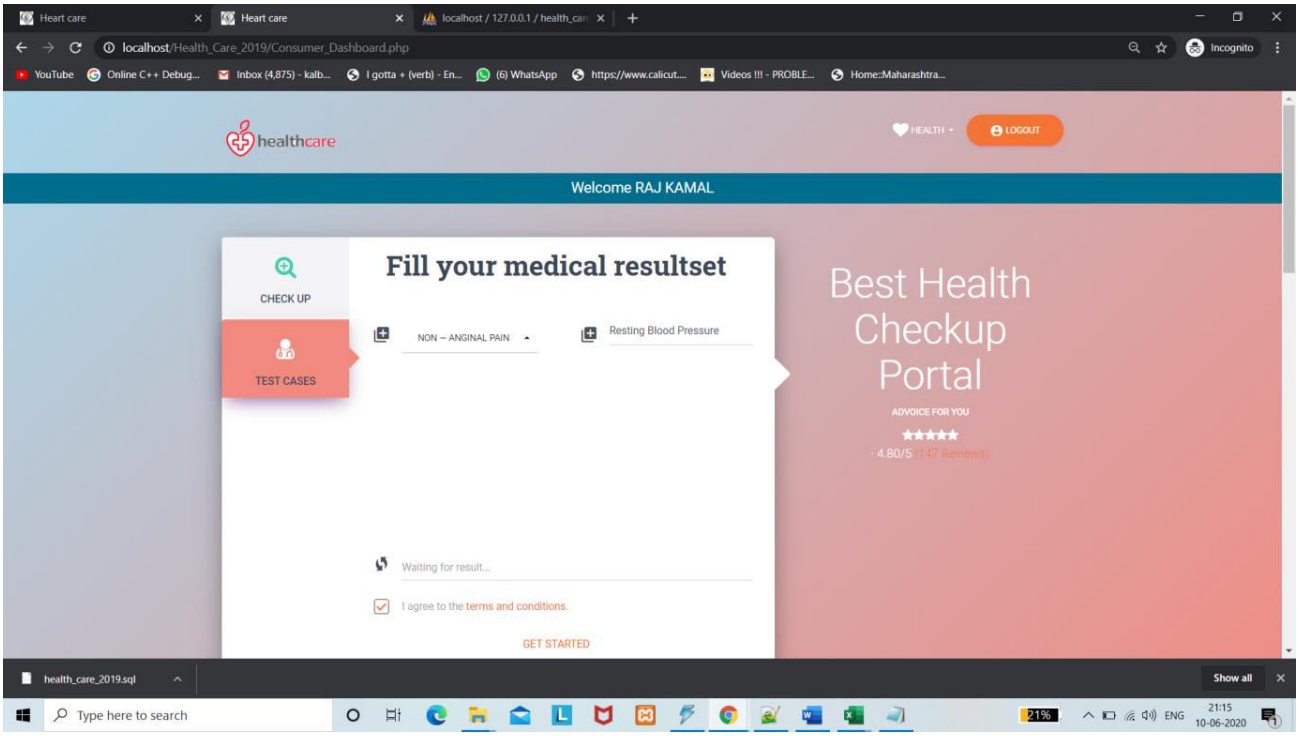

Fig 8: Feeding data into dataset and giving results accordingly

VI. CONCLUSION

Heart disease prediction is challenging and very important in the medical field. However, the mortality rate can be drastically controlled if the disease is detected at the early stages and preventative measures are adopted as soon as possible. Machine learning techniques were used in this work to process raw data and provide a new and novel discernment towards heart disease.

.

REFERENCES

1. Senthilkumar Mohan, "Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques"- 2019 IEEE Smart Caching, Communications, Computing and Cybersecurity for Information-Centric Internet of Things, 19 june 2019.
2. R Latha,"Blood Viscosity based Heart Disease Risk Prediction Model in Edge/Fog Computing", 2019 IEEE 11th International Conference on Communication Systems & Networks (COMSNETS), 13 may 2019.
3. Shivendra Kaura, "Heart disease-Sinus arrhythmia prediction system by neural network using ECG analysis", 2019 IEEE 2nd International Conference on Power Energy, Environment and Intelligent Control (PEEIC), 13 feb, 2019.
4. "Prediction of Cardiovascular Disease Using Machine Learning Algorithms", Kumar G Dinesh ; K Arumugaraj ; Kumar D Santhosh ; V Mareeswari 2018 International Conference on Current Trends towards Converging Technologies.
5. "Analysis of cardiovascular diseases using artificial neural network", Jyotismita Talukdar ; Bhupesh Kumar Dewangan 2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC)

BIOGRAPHY

Mrs. Vaishali Y. Baviskar, is working as an Assistant Professor in G. H. Raisoni Institute of Engineering and Technology, Wagholi, Pune. She has completed her B.E. in Computer Engineering from North Maharashtra University, M.E. in Computer Engineering from Savitribai Phule Pune University and currently pursuing her PhD from Bennett University, Greater Noida. Her Research area is Artificial Intelligence and Machine Learning. She has published 2 National and 6 International Journal Papers. She has an IETE Membership.