Heart Disease detection using machine learning algorithms

Aryan Kumar

11/06/2023+

Abstract:

In todays world the cases of heart disease are be4coming very common and most people are not aware of that which results in certain unfortunate incidents. Therefore detection of an arising heart disease becomes a very import thing from people's home in today's modern world where people do not have time to visit the doctors after every test.

So, in this product, I have proposed the use of **machine learning/AI** for detection of heart disease application. In this application people can enter their various details like age , regular blood pressure type of chest pain etc and using those details the application will be suggesting people if they should consult a doctor or not and based on that some revenue will be generated.

1. Introduction:

Machine learning technology is proving to be a major game changer in the realm of disease detection as it is able to detect may disease beforehand on the basis of some datasets. For one, machine learning based algorithms are able to analyse much larger data sets and take into account significantly more variables than is possible with traditional methods. Before, people used to go to make some time from their busy schedule to check with doctors if their reports are right or not bur with machine learning algorithms integrated with some user-friendly interface, they can do this job from their home and only visit doctors when required.

People are thereby able to use machine learning models to test their heart condition with their mobile phone only. They are able to do this fully automatically, much more precisely, and at a fraction of the effort.

Machine learning based heart disease detection tools are not only designed to learn, but over time they get better at finding the optimal solution for people.

2. Problem Statement:

The problem statement is to use AI to analyse various variables affecting heart condition of a person like person's Age , gender, blood pressure, cholesterol level etc. to predict if the

chance of a person to be suffering from any heart disease based on the prediction suggest people if they need to visit a doctor or not.

3. Costumer/Business Need Assessment:

There are many factors affecting the heart condition of a person like:

- Age
- Gender
- Chest Pain
- Rest Blood pressure
- cholesterol
- Maximum Blood Pressure

Using this model,
we aim to provide people with useful insights from the available data and provide good detection of their heart condition using these parameters along with few others.
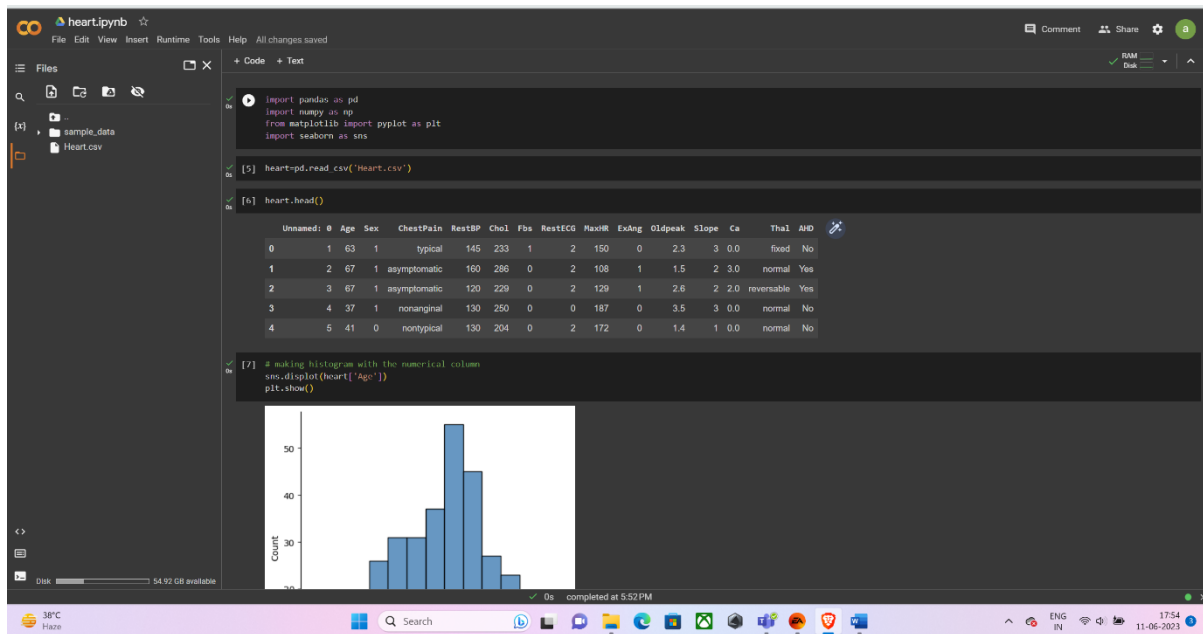
4. Target Specifications:

1- User Interface (UI): Design an intuitive and user-friendly interface that allows users to easily interact with the application. Consider using a clean and visually appealing design to enhance the user experience

2- User Registration and Authentication: Implement a registration and login system to ensure secure access to the application. This will allow users to create personal accounts and store their health data securely.

3- Data Input: Provide a means for users to input relevant health data such as age, gender, blood pressure, cholesterol levels, and any other relevant information. This can be achieved through manual data entry or integration with wearable devices or health apps that can automatically sync data.

4- Risk Assessment: Develop algorithms to analyse the user's health data and assess their risk of heart disease.

5- Results and Recommendations: Provide users with detailed results of the risk assessment, including the probability of heart disease and any identified risk factors.

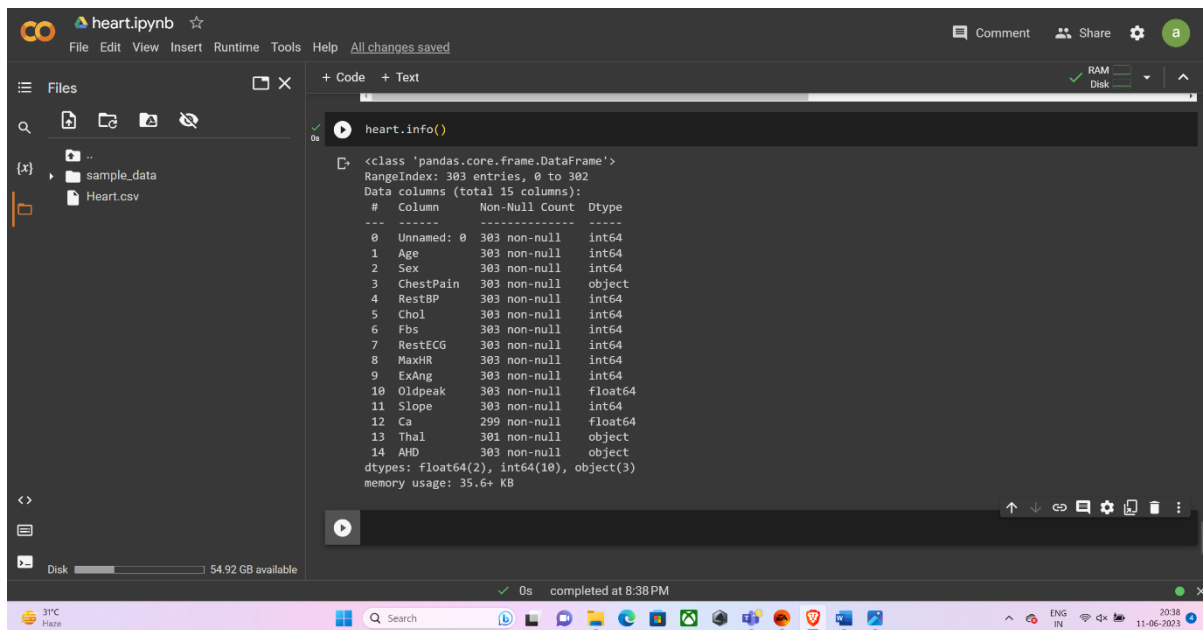## 5. External Searches:

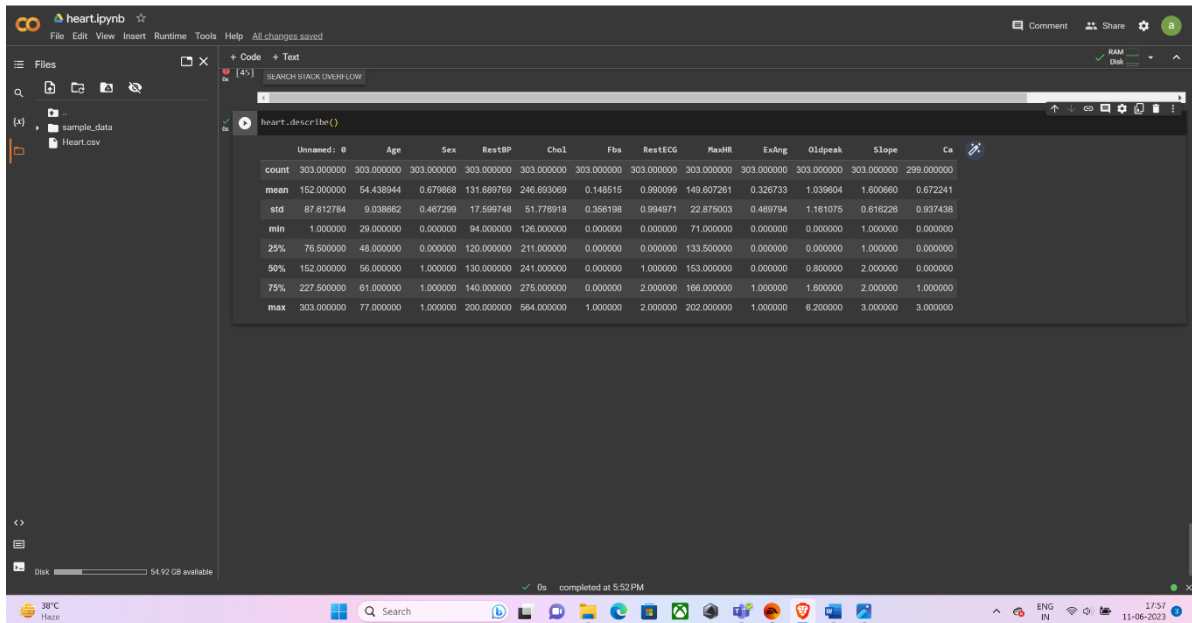I have used the heart disease dataset for reference.
The dataset can be found at: https://www.kaggle.com/datasets/johnsmith88/heart-disease-dataset.

Taking a look at the dataset:

Seeing the various stats of the dataset:

```python
heart.describe()
```

	Unnamed: 0	Age	Sex	RestBP	Chol	Fbs	RestECG	MaxHR	ExAng	Oldpeak	Slope	Ca
count	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	299.000000
mean	152.000000	54.438944	0.679868	131.689769	246.693069	0.148515	0.990099	149.607261	0.326733	1.039604	1.600660	0.672241
std	87.612784	9.038662	0.467299	17.599748	51.776918	0.356198	0.994971	22.875003	0.469794	1.161075	0.616226	0.937438
min	1.000000	29.000000	0.000000	94.000000	126.000000	0.000000	0.000000	71.000000	0.000000	0.000000	1.000000	0.000000
25%	76.500000	48.000000	0.000000	120.000000	211.000000	0.000000	0.000000	133.500000	0.000000	0.000000	1.000000	0.000000
50%	152.000000	56.000000	1.000000	130.000000	241.000000	0.000000	1.000000	153.000000	0.000000	0.800000	2.000000	0.000000
75%	227.500000	61.000000	1.000000	140.000000	275.000000	0.000000	2.000000	166.000000	1.000000	1.600000	2.000000	1.000000
max	303.000000	77.000000	1.000000	200.000000	564.000000	1.000000	2.000000	202.000000	1.000000	6.200000	3.000000	3.000000

6-Benchmarking:

Benchmarking a heart disease detection application's performance is critical for evaluating its effectiveness and comparing it to other existing options. When benchmarking a heart disease detection application, keep the following points in mind:

6.1-Collect a broad and representative dataset of patient medical records, including health data, diagnostic test results, and disease outcomes. Ascertain that the dataset includes a variety of demographics, risk factors, and illness prevalence.

6.2-Performance Metrics: Define relevant performance metrics to assess the accuracy and efficacy of the application. Sensitivity, specificity, positive predictive value, negative predictive value, accuracy, and area under the receiver operating characteristic curve (AUC-ROC) are all used indicators for detecting cardiac disease.

6.3-Baseline Models: Create baseline models or reference algorithms against which you may compare the performance of your application. Existing risk assessment models or other cutting-edge cardiac disease detection algorithms can be used.

6.4-Algorithm Evaluation: Run the heart disease detection algorithms in your application against the dataset and calculate the performance metrics. To evaluate the application's performance, compare these findings to the baseline models.

7.Business Opportunity:

This product has a very big market because as the age of a person increases, they are more likely to develop some diseases and it is very necessary to detect those disease at early stages so that they can be cured at early stages without causing much damage to the body to live a long and healthy life. For this purpose, most of the people are usually busy with their day to day life and cannot find time to visit to doctors therefore if their consultation can be done at their home they will be very much privileged.

8. Concept Generation:

Concept is simple, we want to inform the person before hand about his heart condition using the variables entered by him. For this purpose, we first need to train our model so that it had gone through enough number of data so that it can predict at its level best. So, for that purpose we need to divide our dataset into 2 parts training and testing and them use sklearn library to implement various machine learning algorithms.

The most feasible algorithm here will be naïve byes as far as I think and can be used with various variables to predict the heart condition.

9. Final Product Prototype:

The final product provides users with the risk percentage they might be suffering from a heart disease. In the first stage is must have a welcome page along with the app's logo, then a registration/login from should appear which should be mandatory for users to fill. After that they should be provided with various entry fields where they can enter various details such as their age , blood pressure chest pain type(if they feel any) , etc and based on these inputs the model will calculate the risk percentage of that purpose suffering from that particular disease. Based on the risk percentage he will be suggested if he should consult a doctor or not and that will be our revenue generation part.
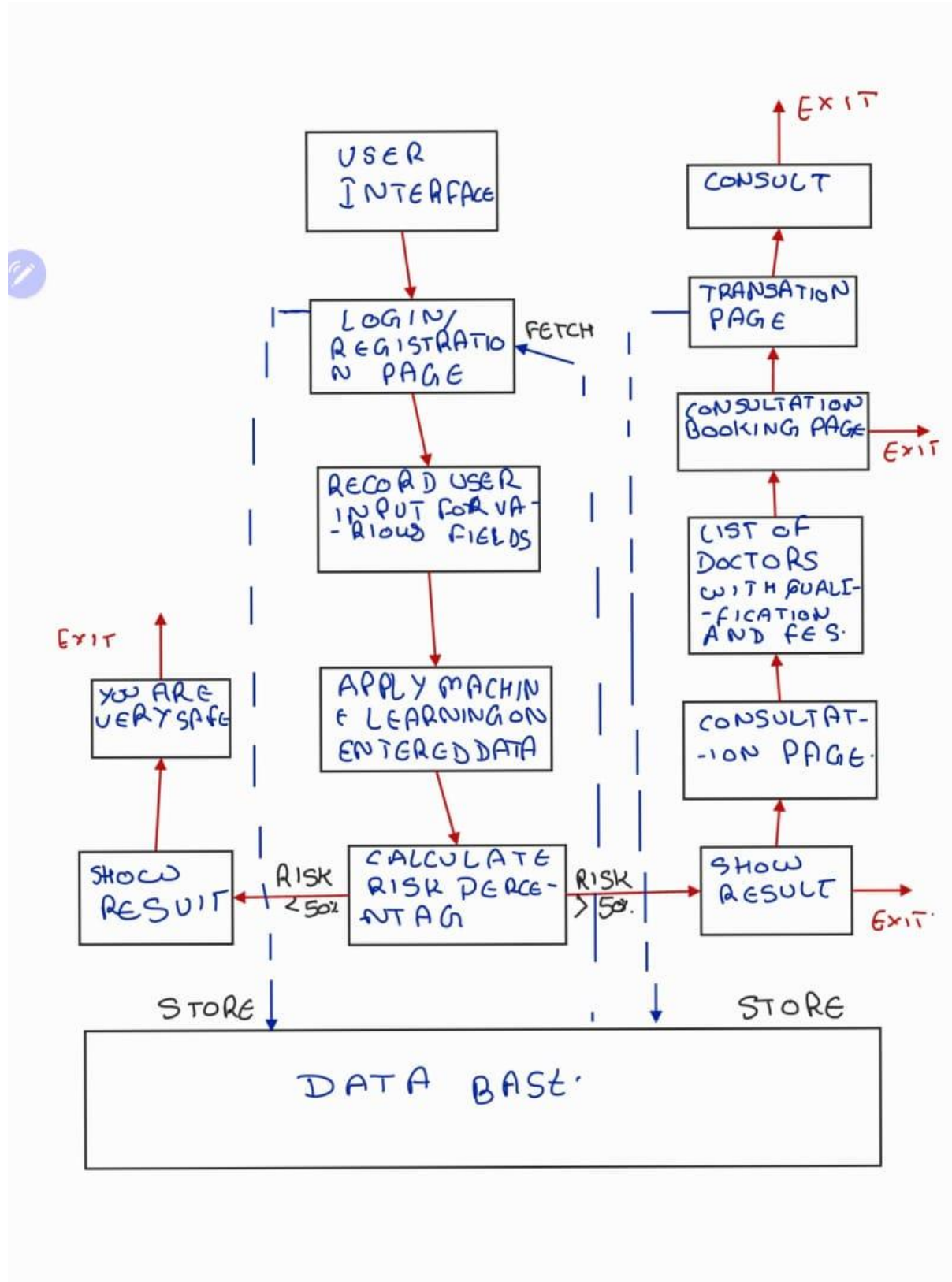
10.Product Details:

10.1 User Interface: This is basically a user-friendly environment where user registers himself and enters the various required data and get the required percentage.

10.2 Machine learning algorithm- Various classification algorithms can be used to determine the risk percentage it should be made clear that if the risk percentage is less than 50% then consultation should not be advised and if it is more than 50% then the consultation should be advised.

10.3 Consultation form- It is seen many times that people do not have time to visit a doctor or they have no idea which doctor to consult, in that case our application should advice people with a list of doctors and their qualifications and consultation fees so that people can choose whom to consult. the consultation offered should be both online and offline, in online they should upload their various test reports and in offline the appointment should be booked via app and people should be able to directly go and visit the doctor

10.4 Revenue generation: In the process of referring to doctors all the doctors registered in the app must pay the commission when the receive a patient through the application.

10.5 User data security: Users data must be secured properly and specially their transaction data.

11. Conclusion:

Disease detection is a very crucial job for any body and should be done properly and with proper accuracy as many people reply on apps for their health info therefore it is very crucial to train the model properly to get the accurate results while minimizing the error percentage.

12. References:

- https://github.com/Ravjot03/Heart-Disease-Prediction/blob/master/Heart_Disease_Prediction.ipynb

- Various articles on increasing heart attacks

- https://github.com/Ravjot03/Heart-Disease-Prediction/blob/master/Heart_Disease_Prediction.ipynb