



# International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

*(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)*



**Impact Factor: 8.699**

**Volume 14, Issue 4, April 2025**

# Heart Disease Predication Using ML

Gaurav Sutar<sup>1</sup>, Sakshi Jadhav<sup>2</sup>, Akshata Katkar<sup>3</sup>, Gayatri Pawar<sup>4</sup>, Dipali Khairnar<sup>5</sup>

U.G. Student, Department of Information Technology, DY Patil University, Ambi, Pune, Maharashtra, India<sup>1,2,3,4</sup>

Associate Professor, Department of Information Technology, DY Patil University, Ambi, Pune, Maharashtra, India<sup>5</sup>

**ABSTRACT:** We are living in the modern Technological era, and we are trying to solve the problem using technology for that we are striving for more knowledge with help of our technology, the more our lifestyle is changing. In the fast growing world everyone is neglecting their Health. Result of this Health diseases are increasing day by day due to our lifestyle and other hereditary reasons. Especially Heart Diseases have become a major issue these days. WHO report that over 11 million deaths are caused each year worldwide due to heart related complications. Using this we can understand that the condition is very serious. There is an urgent need for some technological solutions which can help in educating people and early diagnosis of these disease. This requires a string infrastructure including a large enough work force, which is a slow process. In the recent year machine learning has evolved a lot. To resolve this problem we can use Machine Learning that can help us in this. Sudden increase in the number of patients depict the need for scalability in our medical system. While work force may be limited, we can look for algorithmic solutions for screening and diagnosis stages, due to their correlation with Our project "Heart Health Classification and Early Diagnosis of Heart Disease" is a system that uses predictive capability of machine learning models based on similar data points collected in past. We are going to use the previous medical report data to predict the risk of heart disease. Using our system user can estimate the chance of being suffer from heart disease. Our system takes input data and runs 5Machine Learning Models on the data and gives the result. We are trying to provide the user friendly platform to user that can be access remotely to screening facilities that will predict the risk of heart disease and that will eliminate the load on the medical system.

**KEYWORDS:** KNN, Logistic Regression, Decision Tree, SVM.

## I. INTRODUCTION

Increasing population, changing life patterns, new diseases, pandemics, and change in the natural environment has surely changed the way we take care of our health. Maintaining good health is a major factor associated with any individuals efficiency and It has always been a big concern for all the countries of the world. Providing a good medical service across the whole country proves to be challenging for developing and poor countries. The condition of medical systems across the world is different for different countries. The ratio of healthcare workers to the population is a good way to determine the strength of any healthcare system. Developed countries (the USA and Western European) have this ratio to a good level (~25:10000), for Southeast Asian countries like India are way below the world average (India has a ratio of 12.21 to the 10000, whereas the world average is closer to 10:1000).

Medical doctors (per 10,000)

Last updated: 2022-01-24

EXPORT DATA IN CSV format: Right-click here & Save link

| Location                                             | Medical doctors (per 10,000) | Medical doctors (number) | Generalist medical practitioners (number) | Specialist medical practitioners (number) | Medical |
|------------------------------------------------------|------------------------------|--------------------------|-------------------------------------------|-------------------------------------------|---------|
| Togo                                                 | 0.17                         | 212                      | 145                                       | 35                                        | 357     |
| Togo                                                 | 0.45                         | 244                      | 219                                       | 294                                       | 244     |
| Tonga                                                | 3.03                         | 199                      | 39                                        | 12                                        | 39      |
| Trinidad and Tobago                                  | 10.08                        | 1,039                    | 4,595                                     | 1,098                                     | 1,039   |
| Tunisia                                              | 10.24                        | 10,554                   | 6,491                                     | 7,795                                     | 10,554  |
| Turkey                                               | 10.22                        | 100,853                  | 10,842                                    | 101,899                                   | 10,055  |
| Turkmenistan                                         | 22.13                        | 11,365                   | 4,156                                     | 6,851                                     | 20,051  |
| Tuvalu                                               | 10.31                        | 10                       | 11                                        | 3                                         | 0       |
| Uganda                                               | 0.82                         | 17,186                   | 17,007                                    | 179                                       | 2,299   |
| Ukraine                                              | 29.92                        | 134,986                  | 12,995                                    | 118,737                                   | 25      |
| United Arab Emirates                                 | 13.3                         | 11,630                   |                                           |                                           | 11,630  |
| United Kingdom of Great Britain and Northern Ireland | 16.2                         | 101,823                  | 28,220                                    | 101,214                                   | 185,43  |
| United Republic of Tanzania                          | 0.23                         | 1,481                    | 2,434                                     | 451                                       | 1,481   |
| United States of America                             | 24.39                        | 525,070                  |                                           |                                           | 525,070 |
| Uruguay                                              | 41.23                        | 12,394                   | 8,021                                     | 8,524                                     | 12,394  |
| Uzbekistan                                           | 29.74                        | 45,895                   | 10,868                                    | 97,279                                    | 65,895  |
| Venuela                                              | 1.14                         | 39                       |                                           |                                           | 39      |
| Venezuela (Bolivarian Republic of)                   | 17.3                         | 45,000                   |                                           |                                           | 45,000  |
| Viet Nam                                             | 5.24                         | 42,327                   |                                           |                                           | 42,327  |
| Yemen                                                | 2.39                         | 15,560                   | 5,412                                     |                                           | 5,414   |
| Zambia                                               | 0.33                         | 1,489                    | 1,791                                     | 325                                       | 1,489   |
| Zimbabwe                                             | 0.34                         | 1,054                    | 1,002                                     | 227                                       | 1,054   |

Fig 1. Ratio of Doctors per 10000 population in the USA

Upon investigating the need of medical professionals in current scenarios, which seems to be a long aim to achieve and deadliness and seriousness of heart disease this project is developed. The Heart Disease Prediction Using Machine Learning SOET, DYPUPU 2024-25 15 overall objective of this project is to be able to predict the presence or absence of heart disease by utilizing few important medical and common attributes. Attributes considered form the primary basis for tests and give accurate results more or less.

## II. LITERATURE SURVEY

Sushmita Roy Tithi et al discussed about ECG data analysis and heart disease prediction using machine learning algorithms. In this paper they have used 6 supervised machine learning algorithms to distinguish between normal and abnormal ECG. also they wanted to find a particular disease. They divided there dataset into 2 parts 75% for training and rest 25% for testing.They used- ECG, Machine Learning, Logistic Regression, Decision Tree, Nearest Neighbour, Naïve Bayes , Support Vector Machine, Artificial Neural Network, Right bundle branch block, Myocardial infarction, Sinus tachycardia, Sinus Bradycardia, Coronary Artery disease, Abnormal ECG.ECG provides us with series of sinus rhythm which defines the condition of heart. Used to detect certain kind of diseases. Bo Jin et al discussed about predicting the Risk of Heart Failure with EHR Sequential Data Modeling.Their aim was to provide heart patients an early diagnosis and treatments. Because now a heart failure is really common among people age of 65 , overweight people and those with previous heart attack. This paper develops a new approach to this vital task using and enhanced long short- term memory networks (LSTM) method and a data-driven framework. In this paper they proposed a novel method for diagnosis event modeling that includes one-hit encoding and word vectors and employs LSTM approach for this.

| Disease Name          | Best Algorithm | Score |
|-----------------------|----------------|-------|
| CAD                   | Naïve Bayes    | 94%   |
| Sinus Bradycardia     | Decision Tree  | 95%   |
| Sinus tachycardia     | All except NN  | 95%   |
| Myocardial infarction | Decision Tree  | 96%   |

## III. METHODOLOGY

### 3.1 Building Machine Learning Models

#### Data Collection

The dataset which we have used is available on Kaggle website and is available for public download. It is processed from UCI's dataset and contains valid tuples for further processing.

|                             |                            |
|-----------------------------|----------------------------|
| <u>Dataset Characters</u>   | Multivariate               |
| <u>Number of Tuples</u>     | 1025                       |
| <u>Number of Attributes</u> | 14                         |
| <u>Attribute Datatype</u>   | Categorical, Integer, Real |
| <u>Source</u>               | Kaggle                     |

Table 1. Dataset Detail

Next after downloading datasets is generally to clean the data if some missing values, noise in data is present. First the data is imported from downloaded csv files using pandas libraries of Python, to RAM in data type called data frame which is optimized adequately to handle two dimensional array data. The dataset does not have any null values. 13 of our attributes are the attributes which are used to predict the result, while the last attribute “target” is the result, i.e., whether or not the individual was suffering from heart disease.

The following are the result of the `heartData.info()` and `heartData.describe()` command, which reflects the statistics of Processed dataset.

```
success

In [7]: heartData=pd.read_csv("../dataset/heart_final_processed_by_Sudhansu.csv");
heartData.info()
heartData.describe()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1025 entries, 0 to 1024
Data columns (total 14 columns):
#   Column      Non-Null Count  Dtype
---  ---
0   age         1025 non-null   int64
1   sex         1025 non-null   int64
2   cp          1025 non-null   int64
3   trestbps    1025 non-null   int64
4   chol        1025 non-null   int64
5   fbs         1025 non-null   int64
6   restecg     1025 non-null   int64
7   thalach     1025 non-null   int64
8   exang       1025 non-null   int64
9   oldpeak     1025 non-null   float64
10  slope       1025 non-null   int64
11  ca          1025 non-null   int64
12  thal        1025 non-null   int64
13  target      1025 non-null   int64
dtypes: float64(1), int64(13)
memory usage: 112.2 KB
```

Fig 2. Attributes in Heart Dataset with datatype

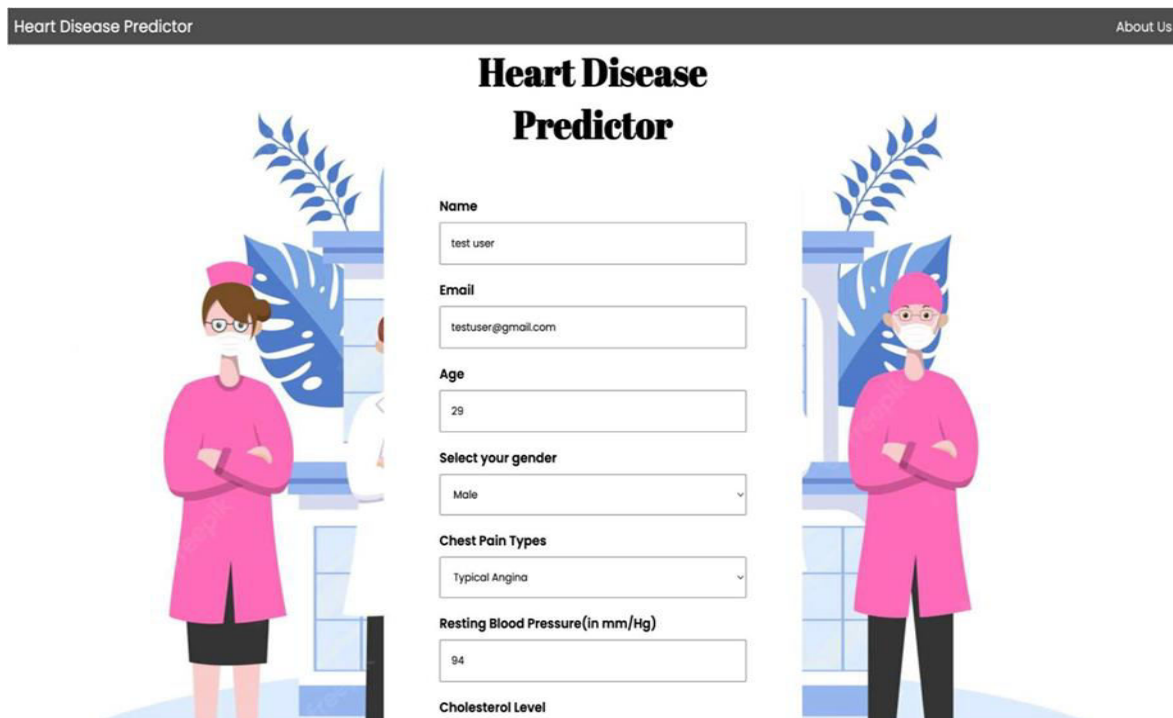
Out[7]:

|       | age         | sex         | cp          | trestbps    | chol        | fbs         | restecg     | thalach     |
|-------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| count | 1025.000000 | 1025.000000 | 1025.000000 | 1025.000000 | 1025.000000 | 1025.000000 | 1025.000000 | 1025.000000 |
| mean  | 54.434146   | 0.695610    | 0.942439    | 131.611707  | 246.000000  | 0.149268    | 0.529756    | 149.114146  |
| std   | 9.072290    | 0.460373    | 1.029641    | 17.516718   | 51.59251    | 0.356527    | 0.527878    | 23.005724   |
| min   | 29.000000   | 0.000000    | 0.000000    | 94.000000   | 126.000000  | 0.000000    | 0.000000    | 71.000000   |
| 25%   | 48.000000   | 0.000000    | 0.000000    | 120.000000  | 211.000000  | 0.000000    | 0.000000    | 132.000000  |
| 50%   | 56.000000   | 1.000000    | 1.000000    | 130.000000  | 240.000000  | 0.000000    | 1.000000    | 152.000000  |
| 75%   | 61.000000   | 1.000000    | 2.000000    | 140.000000  | 275.000000  | 0.000000    | 1.000000    | 166.000000  |
| max   | 77.000000   | 1.000000    | 3.000000    | 200.000000  | 564.000000  | 1.000000    | 2.000000    | 202.000000  |

#### IV. EXPERIMENTAL RESULT

The working front End has a page which contains form where user will enter all the required medical information

The form page screenshot is attached as:



Heart Disease Predictor

About Us

### Heart Disease Predictor

Name  
test user

Email  
testuser@gmail.com

Age  
29

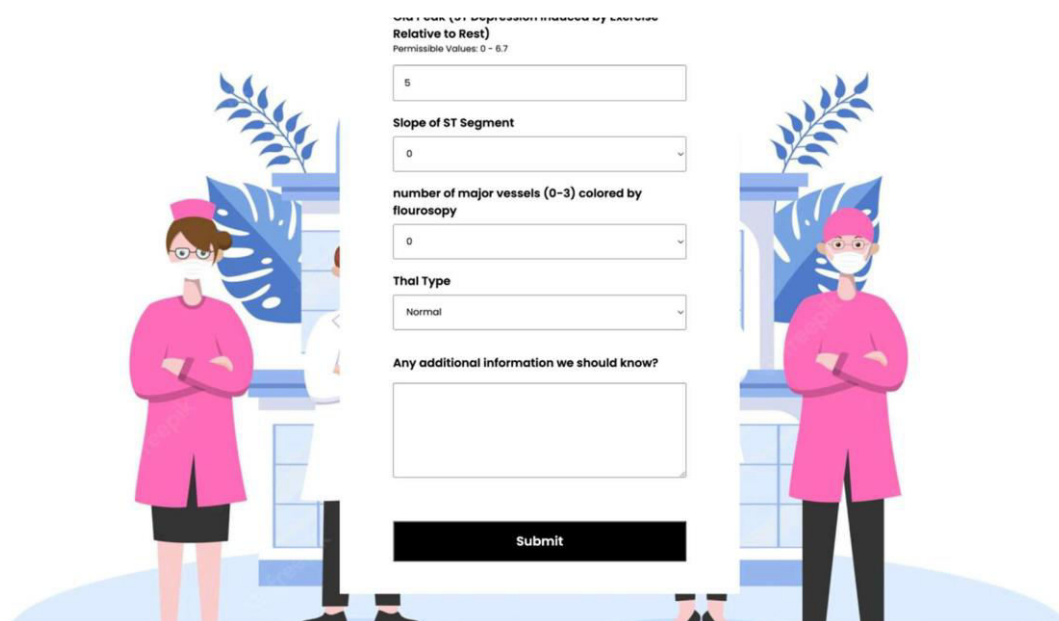
Select your gender  
Male

Chest Pain Types  
Typical Angina

Resting Blood Pressure(In mm/Hg)  
94

Cholesterol Level

Fig 3.Input Image



ST-T Segment (Relative to Rest)  
Permissible Values: 0 - 6.7  
5

Slope of ST Segment  
0

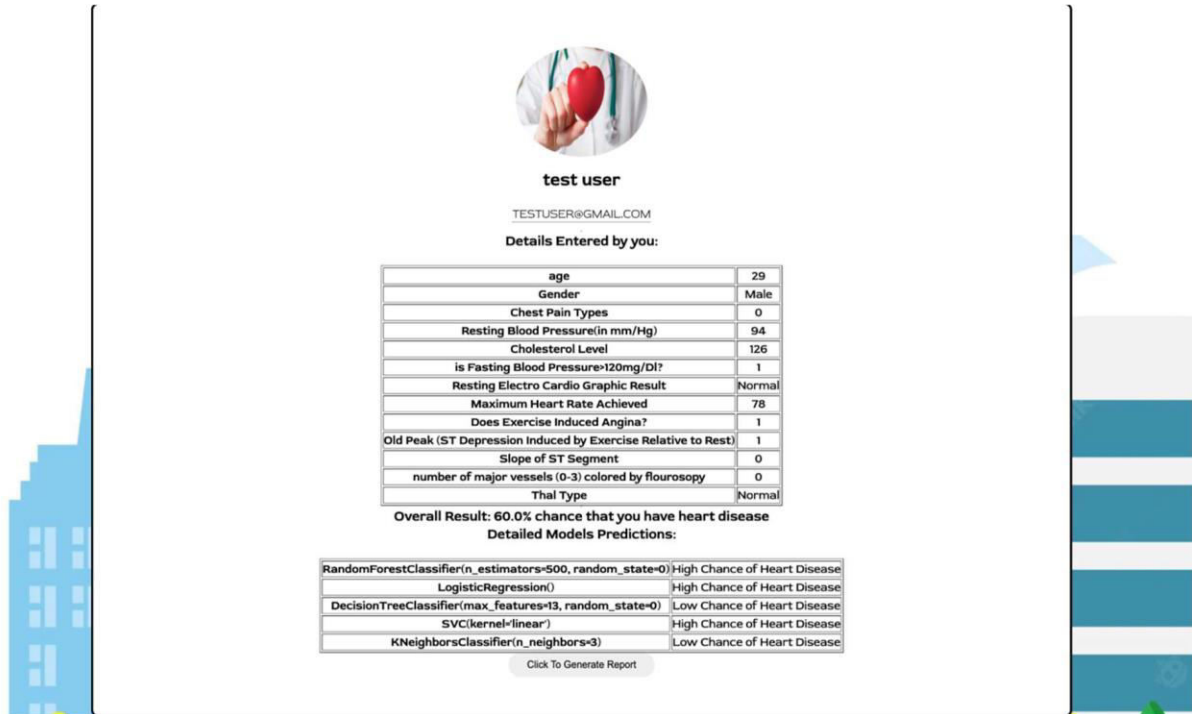
number of major vessels (0-3) colored by flourosopy  
0

Thal Type  
Normal

Any additional information we should know?

Submit

Fig 4.Input Image



**test user**  
TESTUSER@GMAIL.COM

Details Entered by you:

|                                                               |        |
|---------------------------------------------------------------|--------|
| age                                                           | 29     |
| Gender                                                        | Male   |
| Chest Pain Types                                              | 0      |
| Resting Blood Pressure(in mm/Hg)                              | 94     |
| Cholesterol Level                                             | 126    |
| is Fasting Blood Pressure>120mg/Dl?                           | 1      |
| Resting Electro Cardio Graphic Result                         | Normal |
| Maximum Heart Rate Achieved                                   | 78     |
| Does Exercise Induced Angina?                                 | 1      |
| Old Peak (ST Depression Induced by Exercise Relative to Rest) | 1      |
| Slope of ST Segment                                           | 0      |
| number of major vessels (0-3) colored by flourosopy           | 0      |
| Thal Type                                                     | Normal |

Overall Result: 60.0% chance that you have heart disease

Detailed Models Predictions:

|                                                          |                              |
|----------------------------------------------------------|------------------------------|
| RandomForestClassifier(n_estimators=500, random_state=0) | High Chance of Heart Disease |
| LogisticRegression()                                     | High Chance of Heart Disease |
| DecisionTreeClassifier(max_features=13, random_state=0)  | Low Chance of Heart Disease  |
| SVC(kernel=linear)                                       | High Chance of Heart Disease |
| KNeighborsClassifier(n_neighbors=3)                      | Low Chance of Heart Disease  |

Click To Generate Report

Fig 5. The result page screenshot

It also has a button to print the report generated so that user can have a record of data entered by him/her and the respective result.

## V.CONCLUSION

The heart disease prediction project harnesses the power of machine learning integrated with Flask to provide an efficient and accessible diagnostic tool. By leveraging patient medical histories and genetic data, the predictive model offers valuable insights into the likelihood of cardiovascular diseases. The deployment of this model through Flask ensures a user-friendly interface that can be easily accessed by healthcare professionals and patients alike. This project demonstrates the significant potential of combining data mining techniques with machine learning algorithms to enhance early detection and personalized treatment plans for heart disease. The integration of Flask as a web framework facilitates the practical application of the model, making it an invaluable resource in modern healthcare settings. The outcomes of this project not only highlight the importance of intelligent diagnostic systems but also pave the way for future advancements in predictive analytics within the medical field. By improving the accuracy and accessibility of heart disease predictions, this project contributes to better patient outcomes and more efficient healthcare delivery. The success of this approach underscores the transformative potential of technology in addressing critical health challenges.

## REFERENCES

- 1) <https://www.who.int/data/gho/data/themes/topics/health-workforce>
- 2) <https://www.mayoclinic.org/diseases-conditions/heart-disease/symptoms-causes/syc-20353118>
- 3) <https://www.who.int/healthtopics/cardiovascular diseases>
- 4) <https://www.cdc.gov/heartdisease/facts.htm>
- 5) [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds))  
[www.kaggle.com/datasets/johnsmith88/heart-disease-dataset](https://www.kaggle.com/datasets/johnsmith88/heart-disease-dataset)
- 6) Sushmita Roy Tithi ,AfifaAktar , Fahimul Aleem , Amitabha Chakrabarty ‘ECG data analysisand heart disease prediction using machine learning algorithms’. Proceedings of 2019 IEEERegion 10 Symposium

- 7) Bo Jin ,Chao Che, Zhen Liu, Shulong Zhang, XiaomengYin,AndXiaopeng Wei, “Predicting theRisk of Heart Failure WithEHR Sequential Data Modeling” ,IEEE Access Volume 6 2018.
- 8) Ashir Javeed, Shijie Zhou, Liao Yongjian, Iqbal Qasim,Adeeb Noor, Redhwan Nour4, SamadWali And Abdul Basit ,“An Intelligent Learning System based on Random SearchAlgorithm andOptimized Random Forest Model forImproved Heart Disease Detection” , IEEE Access2017.This work is licensed under a Creative Commons Attribution 4.0 License Volume 4 2016.
- 9) Muhammad, Y., Tahir, M., Hayat, M. *et al.* Early and accurate detection and diagnosis of heartdisease using intelligent computational model. *Sci Rep* 10, 19747 (2020). <https://doi.org/10.1038/s41598-020-76635-9>. [<https://www.nature.com/articles/s41598-020-76635-9>]



INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details