

HEART DISEASE PREDICTION SYSTEM

Salman Khan¹, Rajat Kumar Upadhyay², Amir Khan³, Param Goel

Department Of Computer Science & Engineering, United Institute Of Technology, Prayagraj, Uttar Pradesh

Abstract - Heart disease is the most common disease. But, unfortunately the treatment of heart disease is somewhat costly that is not affordable by common man. Hence, we can reduce this problem in some amount just by predicting heart disease before it becomes dangerous using **Heart Disease Prediction System Using Machine Learning and Data mining**. If we can find out heart disease problem in early stages then It becomes very helpful for treatment. **Machine Learning and Data Mining techniques** are used for the construction of **Heart Disease Prediction System**. In healthcare biomedical field, there is large use of health care data in the form of text, images, etc. but, that data is hardly visited and is not mined. So, we can avoid this problem by introducing **Heart Disease Prediction System**.

Key Words: app, Disease , Prediction, online

1.INTRODUCTION

The health care industries collect huge amounts of data that contain some hidden information, which is useful for making effective decisions. For providing appropriate results and making effective decisions on data, some advanced data mining techniques are used. In this study, a Heart Disease Prediction System (HDPS) is developed using Naive Bayes and Decision Tree algorithms for predicting the risk level of heart disease. The system uses 15 medical parameters such as age, sex, blood pressure, cholesterol, and obesity for prediction. The HDPS predicts the likelihood of patients getting heart disease. It enables significant knowledge.

2 . Body of Paper

Heart Disease Prediction System is a Machine Learning based web Application. This web application helps the patients to check about his Heart disease problem. The system can discover and extract hidden knowledge associated with diseases from a historical heart data set Heart disease prediction system aims to exploit data mining techniques on medical data set to assist in the prediction of the heart diseases.

After comparing different types of algorithms we have selected the best suited algorithms as the Naïve Bayes and Decision Tree algorithm for this project as our requirements are superior and it is a long developing project.

Naïve Bayes Algorithm :

Naïve bayes algorithm is a supervised learning algorithm which is based on Bayes theorem and used for solving classification problems.

- It is mainly used in text classification that includes a high-dimensional training dataset.

- It is a probabilistic classifier which means it predicts on the basis of the probability of an object .

Working of Naïve Bayes Classifier:

1. Convert the given dataset into frequency tables.
2. Generate likelihood table by finding the probabilities of given features.
3. Now, use Bayes Theorem to calculate the posterior probability.

Advantage Of Naïve Bayes Classifier :

We are using this Algorithm for the following advantages:

- Naïve Bayes is one of the fast and easy ML algorithms to predict a class of datasets.
- It can be used for Binary as well as Multi-class Classifications.
 - It performs well in Multi-class predictions as compared to the other algorithms. It is the most popular choice for text classification problems .

Decision Tree Algorithm:

- Decision Tree is a Supervised learning technique that can be used for both classification and Regression problems, but mostly it is preferred for solving Classification problems. It is a tree structured classifier, where internal node represents the features of a dataset, branches represents the decision rules and each leaf node represents the outcome.
- Decision Tree is a graphical representation for getting all the possible solutions to a problem/decision based on given conditions.

Working of Decision Tree Algorithm :

- **Step – 1 :** Begin the tree with root node, says S, which contains the complete dataset.
- **Step – 2:** Find the best attribute in the dataset using Attribute Selection Measure.
- **Step - 3:** Divide the S into subsets that contains possible values for the best attributes.
- **Step - 4:** Generate the decision tree node, which contains the best attribute.
- **Step – 5:** Recursively make new decision trees using the subsets of the dataset created in step – 3. Continue this process until a stage is reached where you cannot further classify the nodes and called the final node as a leaf node.

Advantage Of Decision Tree Algorithm :

- While utilizing a decision tree algorithms , it is not essential to standardize or normalize the data that has been collected . It can handle both continuous categorical variables.
- The execution of a Decision tree algorithm must be possible without having to scale the data as well.
- While utilizing the decision tree algorithm, it is not necessary to credit the missing values.
- Comprehensive rules are generated in a decision tree.

Requirement Analysis

Requirement Analysis, also known as Requirement Engineering, is the process of defining user expectations for a new software being built. We analyze, refine, and scrutinize the gathered requirements to make consistent and unambiguous requirements. This activity reviews all requirements and may provide a graphical view of the entire system. After the completion of the analysis, it is expected that the understandability of the project may improve significantly.

Developer Requirement

Following configurations are required for development:

Hardware Requirements

- Processor : Core i3, 2.30 GHz
- Operating System: Windows 10
- RAM:4GB

- Database: SQLITE (DATABASE)
- Frontend Application: HTML , CSS AND PYTHON BASICS
- Backend Application : PYTHON DJANGO, SQLITE (DATABASE)
- Any user having the following essentials can access this website: 11
- Internet connection
- UserI'D and password

3. CONCLUSIONS

It has been a great pleasure for me to work on this exciting and challenging project. This project proved good for me as it provided practical knowledge of not only programming in Python and Sqlite web based application. It also provides knowledge about the latest technology used in developing web enabled application and client server technology that will be great demand in future. This will provide better opportunities and guidance in future in developing projects independently.

ACKNOWLEDGEMENT

We would like to thank our principal Mr. Sanjay Srivastava and our HOD Mr. Abhishek Malviya for granting us the honour of being a member of project group, Heart Disease Prediction System. This project helped us in doing a lot of research and made us en-lighted with many new things.

We would like to thank Mr. Param Goel for his invaluable help and guidance. We would like to express our deep sense of gratitude to our seniors for their important help and suggestions requirements. We would like to thank all members of CSE lab who have directly or indirectly contributed in my project work and maintained a friendly atmosphere in the lab. Finally, we would like to thank our families and our friends for their constant moral support.

REFERENCES BOOKS

1. Two scoops of Django for 1.11 by *Daniel Greenfeld's and Audrey Greenfield*
- 2 Lightweight Django by *Elman and Mark Lavin*