**Project Synopsis**

On

**“ Heart Disease Prediction System ”**

Submitted in the partial fulfilment of the requirement for the degree

Of

**Bachelor of Technology**

In

**Computer Science**



**Arya Institute of Engineering and Technology**

**Kukas , Jaipur**

**Submitted to –**  **Submitted By -**

Ms. Dhara Upadhyay Aman Panchal

Assistant Professor Aakash Yadav

Department of Computer Science Aman Pandey

CS(7th Sem)

**Project Synopsis**

**Heart Disease Prediction System**

**Team :**

1. Aman Panchal - 17EAICS019
2. Aakash Yadav - 17EAICS001
3. Amam Pandey - 17EAICS020

**Email:** [panchalaman2000@gmail.com](mailto:panchalaman2000@gmail.com)

**Mob no. :** 9783253920

**Page no.**

1. **Abstract**
2. **Introduction**
3. **Why is the particular topic chosen**
4. **Project Objective  and Scope**
5. **Methodology**
6. **Existing System**
7. **Proposed System**
8. **Software & Hardware Requirement**
9. **Testing Technology used**
10. **Abstract :**

The idea of predicting heart disease by applying machine learning and pandas in python. Taking dataset from the medical background of different people (prime Indians dataset from UCI repository). This data set consists of information of the user whose age, sex, chest pain type, cholestoral, maximum heart rate , etc. related to diabetes. Design a testing and training set and predict what are the chances of patients having heart disease in the coming five years . Data is classified and shown in the form of different graphs.

1. **Introduction :**

Heart Disease Prediction System is playing an important part in predicting heart disease in the coming years. This prediction can give the option for doctors and sergeon to take steps in dealing with the patients having heart related problem. In this project prediction of heart disease is considered as the main problem.

1. **Why is the particular topic chosen ?**

As we know that in todays world heart related problems are growing rapidly. And there are many countries which do not have enough no. of doctors. So to solve this problem and to reduce time spent on each patient and predict the disease more accurately and precisely we are working on this project.

1. **Project Objective  and Scope :**

To predict heart disease by considering exiting the user’s data set and predict what are the chances of heart disease in the coming five years. Information is shown in the form of different graphs.

1. **Methodology :**

**1. Frame the problem and look at the picture :** First of all we start from understanding the problem in the project. To knew what the problem it actually is?

**2. Get the data :** We get data from kaggle(sea of dataset).

**3. Explore the data to gain insight :** Then we explore the data that which features really effect the result and which features are correlated to each other.

**4. Prepare the data for Machine Learning algorithms :** Before sending the data to machine learning algorithm we have to prepare the dataset according to the data which fit in algorithm properly. And we have remove the NAN values, remove the alphabets/string and convert them to categorical , etc.

**5. Explore the different models :** After preparing the datasets we fit this into various machine learning models. Just like Logistic regression, k-nearest neighbour, Support Vector machine, Decision Tree, etc and select that which give the best result.

**6. Fine tune model :** After selecting the best model we tune the model. This is done when we have limited computing resources and improve the model result. This can be done by adjusting learning rate, no. of leaves or depth of tree, no. of hidden layers in DNN, clusters in k-means.

**7. Present the solution :** After training and testing model we get the model which can detect disease in patient. Now we have to present the solution by using various graphs and charts using matplotlib, seaborn, etc.

1. **Existing System :**

There was no chance of prediction in existing studies it was just by manual analysis based on existing data but analyzing large amounts of datasets is not considered.

1. **Proposed System :**

Data analysis and machine learning libraries and algorithms are used for prediction on heart disease and information is shown in detail in the form of different types of graphs (histogram, density plots, box and whisker plots, and correlation matrix plots.

1. **SOFTWARE & HARDWARE REQUIREMENT :**

Software

* OS: Windows 7 or above
* Programming language: python 3.6
* IDE: Anaconda prompt
* Dependencies: numpy
* Libraries: panda, sklearn, matplotlib, seaborn

Hardware

* Processor: i5 or above
* RAM: 8 GB
* Hard Disk: 160 GB

1. **Testing Technology used:**

We tested our model on the 30% data which we have have separated from our original dataset. And we get 81.13% accuracy on training data and 87.91% accuracy on tested data.