SYNOPSIS

ON

“PRESERVE YOUR HEART”

Submitted in

Partial Fulfillment of requirements for the Award of Degree *of*

Bachelor of Technology

*In*

# Computer Science and Engineering

(Artificial Intelligence) By

**(Project Id: 22\_AI\_2A\_17)**

**Aman Bhadauriya (2101641520013)**

**Abhishek Swaroop (2101641520004)**

**Avantika Bajpai (2101641520036)**

**Abhinav Yadav (2101641520003)**

**Akhand Pratap Singh (2101641520010)**

Under the supervision of

Mr. S S Parihar

**(Professor)**



**Pranveer Singh Institute of Technology**.

Kanpur - Agra - Delhi National Highway - 19

Bhauti - Kanpur - 209305.

(Affiliated to Dr. A.P.J. Abdul Kalam Technical University)

# **1. Introduction :**

In recent years due to change in eating habits and life style of every human being the diseases has become next to normal.

Heart is an essential organ of one body and needs to be healthy and protected from diseases. It is that important part of the body which distinguishes a living body to a non-living body so it becomes the major duty to look after it.

Out of five one person is prone to heart disease in India and one of the most miserable thing is they are not aware of this disease. Due to unhealthy life style the people are having heart disease like :- blood pressure , cholesterol , cardiac arrest , anxiety attacks and many such types of diseases.

In early times disease related to heart used to pan out to people over 70yrs of age but now it has become common in youngsters also.

This project is build with the aim to predict the heart disease for every age group and alert them so that they can look after it and take required precautions and medications if needed.

**FIELD OF PROJECT :**

In this project we are making use of MACHINE LEARNING along with PYTHON. The project will be made on Google Colab.

**Google Colab :**

Colaboratory, or “Colab” for short, is a product from Google Research. Colab allows anybody to write and execute arbitrary python code through the browser, and is especially well suited to machine learning, data analysis and education. More technically, Colab is a hosted Jupyter notebook service that requires no setup to use, while providing access free of charge to computing resources including GPUs. Colab is it is a open source platform.free to use as

Google Colaboratory, popularly known as Colab, is a web IDE for python that was released by Google in 2017. Colab is an excellent tool for data scientists to execute Machine Learning and Deep Learning projects with cloud storage capabilities.

So, here we will be using colab for writing our projects as it is related to machine learning.

# **2. Project Objective :**

As already mentioned above the name of the project is related to the heart health. Themain aim of making this project is to predict the heart health condition of each and every individual and to make them aware of the consequences so that they can start taking care of themselves and their heart as well as take proper medication and precautions as required and instructed by their doctors.

A quiet Significant amount of work related to the diagnosis of Cardiovascular Heart disease using Machine Learning (M.L.) algorithms has motivated this work. This project contains a brief literature survey and also an attempt to build up a Machine Learning based Algorithm to let people know about the Heart-Diseases (if any) they are suffering from. An efficient Cardiovascular-Disease prediction has been made by using various algorithms some of them include Logistic Regression (L.R.), K-Nearest Neighbours (K.N.N.), Random Forest Classifier (R.F.C.) etc. It can be seen in results that each algorithm has its strength to register the defined objectives. The model incorporating Intelligent Heart Disease Prediction System (I.H.D.P.S.) had the ability to calculate the decision boundary using the previous and new model of machine learning and deep learning. It facilitated the important and the most basic factors / knowledge such as family history connected with any heart disease. But the accuracy that was obtained in such Intelligent Heart Disease Prediction System (I.H.D.P.S.) model was far more less than the new upcoming model such as detecting coronary heart disease using Artificial Neural Network (A.N.N.) and other algorithms of Machine Learning (M.L.) and Deep Learning (D.L.). The risk factors of coronary Heart-Disease or atherosclerosis is identified by McPherson, using the inbuilt implementation algorithm using uses some techniques of Neural Network and were just accurately able to predict whether the test patient is suffering from the given disease or not. Diagnosis and prediction of Heart-Disease and Blood Pressure along with other attributes using the aid of neural networks was introduced by R. Subramanian. A Deep Neural Network (D.N.N.) was Built incorporating the given attributes related to the disease which were able to produce a output which was carried out by the output perceptron and almost included 120 hidden layers which is the basic and most relevant technique of ensuring a accurate result of having Heart-Disease if we use the model for Test Dataset. The supervised network has been advised for diagnosis of heart diseases. When the testing of the model was done by a doctor using an unfamiliar data, the model used and trained from the previous learned data and predicted the result thereby calculating the accuracy of the given model.

**Main Objectives :**

The main objective of this project is to develop a Heart-Prediction System. 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.

**Specific Objectives :**

1. Provides new approach to concealed patterns in the data.
2. Helps avoid human biasness.
3. To implement Naïve Bayes Classifier that classifies the disease as per the input of the user.
4. Reduce the cost of medical tests.

# **3. Feasibility :**

**3.1 Operational Feasibility :**

We should make an estimation on the time require by the team members to gain the knowledge about the needs / requirements of the Project in order to make it a much more affective Project.

**3.2 Technical Feasibility :**

The main consideration is to be given to the study of available resources of the organization where the software is to be implemented. Here the system analyst evaluates the technical merits of the system giving emphasis on the performance, reliability, maintainability and productivity.

By taking the consideration before developing the proposed system, the resources availability of the system giving emphasis on the performance, reliability, maintainability and productivity.

**3.3 Economic Feasibility :**

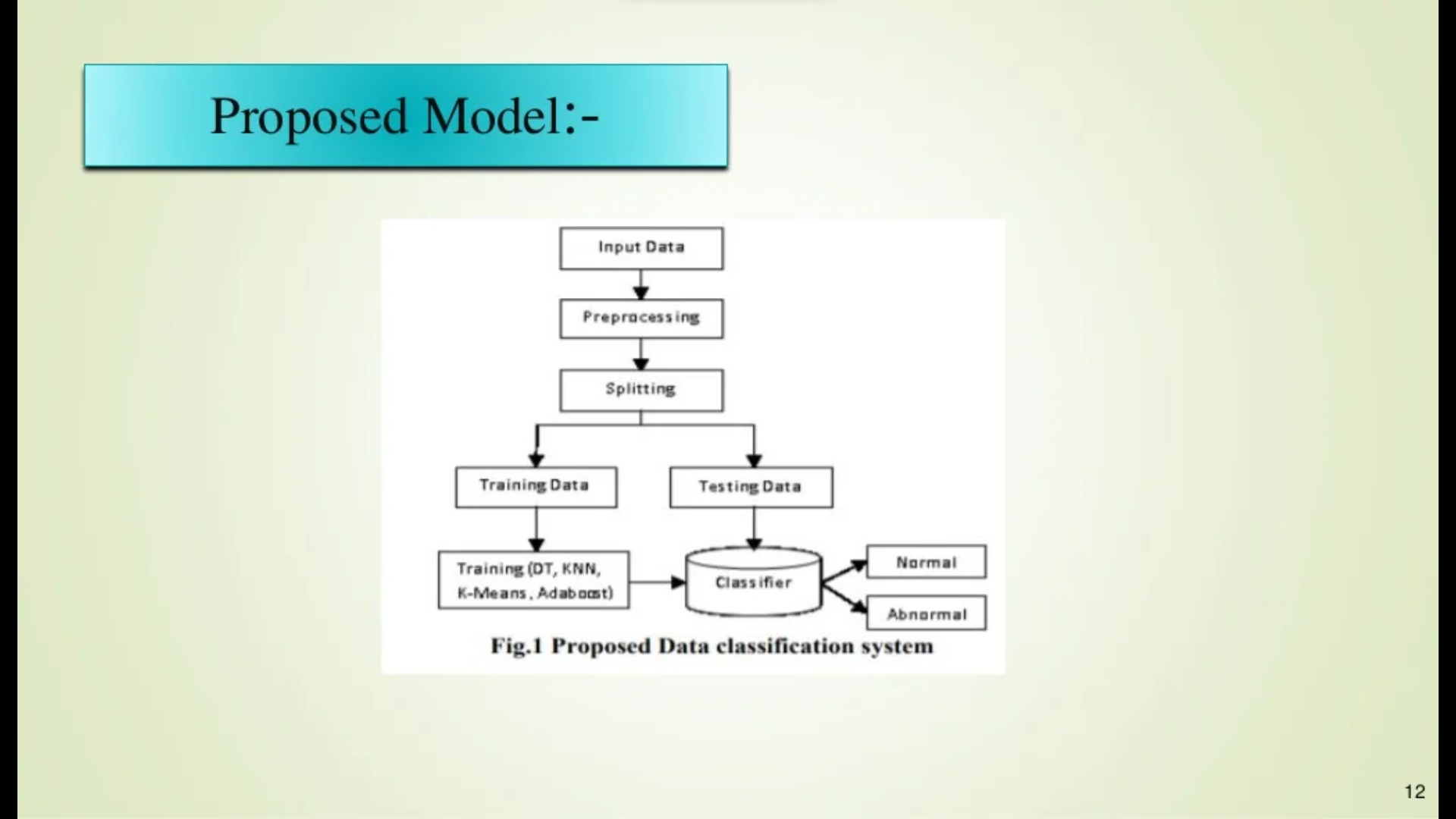
Economic feasibility is the most important and frequently used method for evaluating the effectiveness of the proposed system. It is very essential because the main goal of the proposed system is to have economically better result along with increased efficiency. Cost benefit analysis is usually performed for this purpose. It is the comparative study of the cost verses the benefit and savings that are expected from the proposed system. Since the organization is well equipped with the required hard ware, the project was found to be economically feasible.

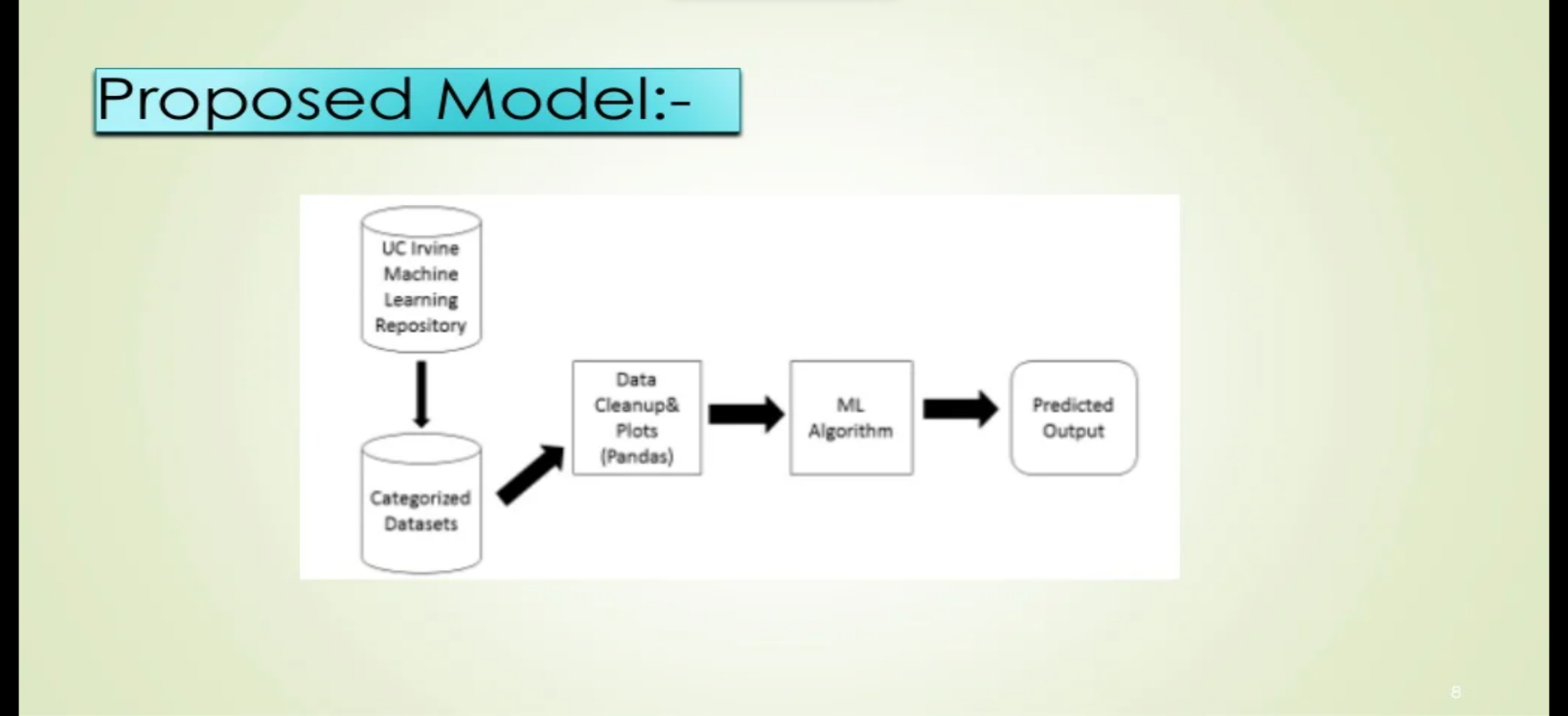
**3.4 Schedule Feasibility :**

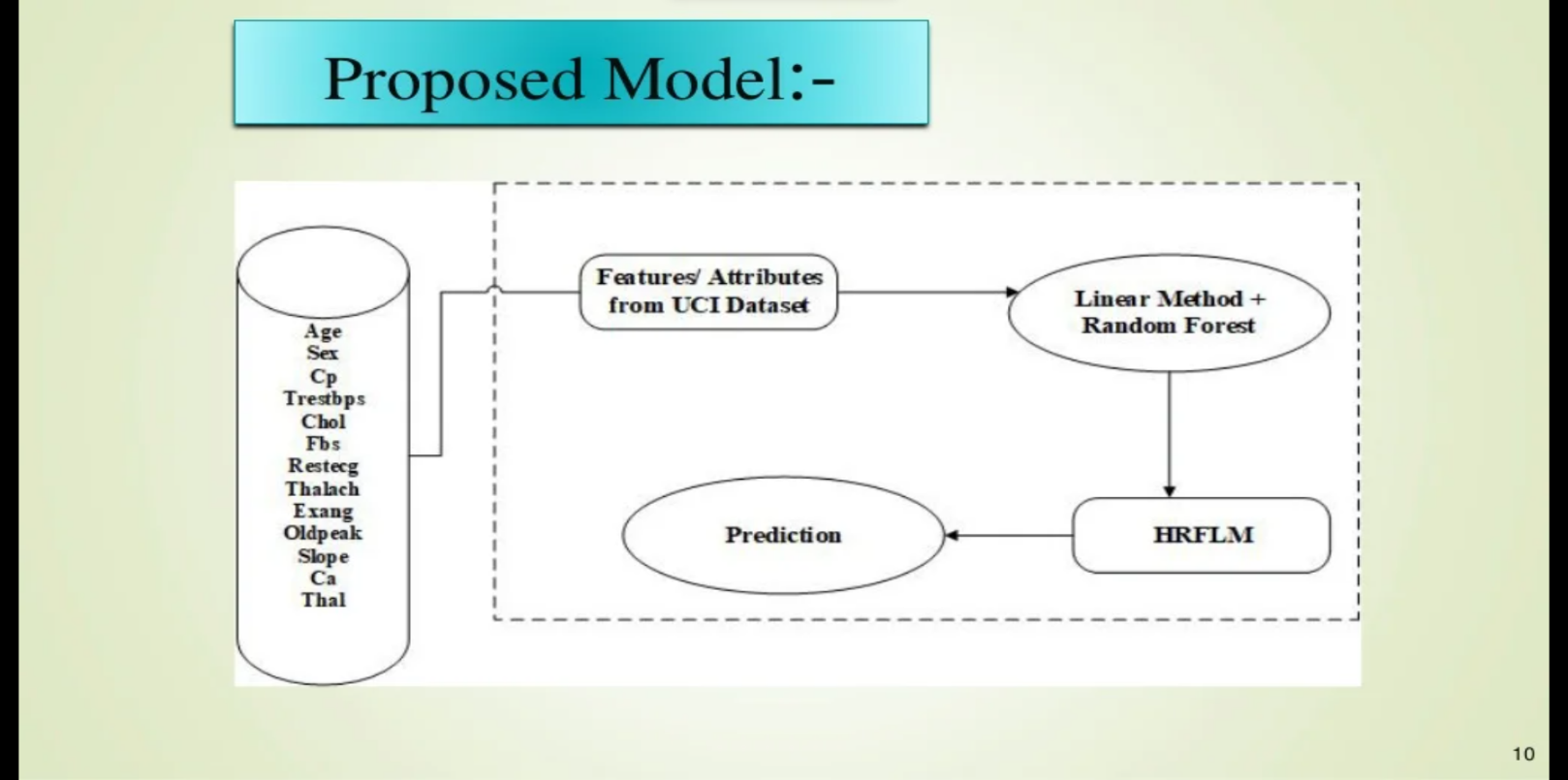
Time evaluation is most important consideration in the development of the project. Time schedule required for development of project is been provided in next page:-

**Feasibility Chart of the Project**









# **4.** **Methodology / Planning of Work :**

So the work flow that is going to be followed for this project development is as follows:

**Step 1 :-** The first and the foremost step is to get the heart data. This data contains several health parameters which corresponds to a person healthiness of the heart.

**Step 2 :-** Once we have our data set with us we need to process it . We cannot feed this raw data into our machine learning algorithm so we need to process this data set to make it fit and compatible for our machine learning algorithm.

**Step 3 :-** Once we process the data we need to split our data into training data and testing data .this is because we often train our machine learning algorithm with training data and we evaluate the performance of our model using the test data so this part is called Train Test Split where we will original data set into training data and test data.

**Step 4 :**- Once we are done with train test split process we will feed our training data to our machine learning model . In this case we are going to use a logistic regression model because this particular case has binary classification i.e here we are going to classify that whether the person has heart disease or not.

**Step 5 :-** When we will train this logic regression model with our trained data we will do some evaluations to check its efficiency and performance so after that we will get a trained logistic regression.

**Step 6 :-** When we will get the this trained logistic regression model we will feed new data to our model it can predict whether the person is having a diseased heart or the healthy one . So this is what we are going to do in our project.

This paper shows the analysis of various machine learning algorithms, the algorithms that are used in this paper are K-Nearest Neighbors (K.N.N.), Logistic Regression (L.R.) and Random Forest Classifiers (R.F.C.) which can be helpful for practitioners or medical analysts for accurately diagnose Heart-Disease. This paperwork includes examining the journals, published paper and the data of Cardiovascular disease of the recent times. Methodology gives a framework for the proposed model. The methodology is a process which includes steps that transform given data into recognized data patterns for the knowledge of the users. The proposed methodology includes steps, where first step is referred as the collection of the data than in second stage it extracts significant values than the third is the preprocessing stage where we explore the data. Data preprocessing deals with the missing values, cleaning of data and normalization depending on algorithms used. After pre-processing of data, classifier is used to classify the pre-processed data the classifier used in the proposed model are (K.N.N.), Logistic Regression (L.R.), Random Forest Classifier (R.F.C.). Finally, the proposed model is undertaken, where we evaluated our model on the basis of accuracy and performance using various performance metrics. Here in this model, an effective Heart Disease Prediction System (E.H.D.P.S.) has been developed using different classifiers. This model uses thriteen medical parameters such as Chest Pain, Fasting Sugar, Blood Pressure, Cholesterol, Age, Sex etc. for the prediction of the Heart Disease.

**5. Tools/Technology Used :**

5.1 Hardware Requirements

* Processor 🡪 Intel Core III or above
* Clock Speed 🡪 800MHZ
* System Bus 🡪 32/64 bits
* RAM 🡪 2 GB or above
* Hard disk 🡪 HDD/SSD 250GB or above

5.2 Software Requirements

* Operating System 🡪 Windows/Mac OS/Linux
* Browser 🡪 Any HTTP Browser
* Database layer 🡪 MySQL and Google Colaboratory
* Frontend scripting 🡪 HTML and CSS
* Backend scripting 🡪 Python (Django)

**5.3 Technology Specification**

The technologies used in this Project are based on Machine Learning Algorithms and various Web Frameworks like Django and Languages like Python are used for the development of this Project. This project also includes the usage of Google Colaboratory which is used for combining executable Python Codes.

**6. References :**

1. Machine Learning Engineer (Google Services) **🡪https://www.coursera.org/learn/gcp-big-data-ml-fundamentals**
2. Introduction to Programming with Python **🡪https://www.coursera.org/learn/python-programming-intro**
3. Fundamentals of Machine Learning for Healthcare **🡪 https://www.coursera.org/learn/fundamental-machine-learning-healthcare?**
4. <http://www.wikipedia.com/>
5. GeeksforGeeks **🡪** <http://www.geeksforgeeks.com/>
6. Google **🡪** <http://www.google.com/>
7. Google Colaboratory **🡪** <https://colab.research.google.com/>
8. Coursera **🡪** https://in.coursera.org/