# DETECTING HEART PROBLEMS USING LRA

A Project Report

Submitted By

## VISHNU IMMADI

### 210303126715

## ANVESH IMMADI

### 200303126097

## DINESH KONALA

### 210303126717

## K.V.KOTESWAR RAO

### 210303126716

in Partial Fulfilment For the Award of

the Degree of

BACHELOR OF TECHNOLOGY

COMPUTER SCIENCE AND ENGINEERING

Under the Guidance of

**Prof.(Dr.) Abhishek Kumar**

Associate Professor



**PARUL UNIVERSITY**

**VADODARA**

**October - 2023**

# PARUL UNIVERSITY

# CERTIFICATE

This is to Certify that Project - 2 -Subject code 203105400 of $7^{th}$ Semester entitled "DETECTING HEART PROBLEMS USING LRA" of Group No. PUCSE 291 has been successfully completed by

- VISHNU IMMADI – 210303126715

- ANVESH IMMADI – 200303126097

- DINESH KONALA – 210303126717

- K.V.KOTESWAR RAO - 210303126716

under my guidance in partial fulfillment of the Bachelor of Technology (B.TECH) in Computer Science and Engineering of Parul University in Academic Year 2023- 2024.

Date of Submission :-

**Prof. Abhishek Kumar**,                          Head of Department,

Project Guide                                     Dr. Amit Barve

                                                 CSE, PIET

Project Coordinator:-

**Prof. Yatin Shukla**,                          Parul University

External Examiner

# Acknowledgements

My humble thanks to the Almighty God.

We are pleased to present "DETECTING HEART PROBLEMS USING LRA" project and take this opportunity to express our profound gratitude to all those people who helped us in completion of this project. We express our deepest gratitude towards our project guide **Prof. Abhishek Kumar**, his valuable and timely advice during the various phases in our project. We would also like to thank him for providing us with all proper facilities and support as the project guide. We would like to thank him for support, patience and faith in our capabilities and for giving us flexibility in terms of working and reporting schedules.

We also express my deep sense of regards and thanks to **Dr. Amit Barve** (Head of Computer Science and Engineering Department). I feel very privileged to have had their precious advice, guidance and leadership.

We thank our college for providing us with excellent facilities that helped us to complete and present this project. We would also like to thank the staff members.

Finally, we would like to thank everyone who has helped us directly or indirectly in our project.

**PLACE: VADODARA**
**DATE:**

VISHNU IMMADI – 210303126715
ANVESH IMMADI – 200303126097
DINESH KONALA – 210303126717
K.V.KOTESWAR RAO - 210303126716

# Abstract

Heart diseases are a leading cause of mortality worldwide, emphasizing the need for effective early detection methods. This study presents a novel approach to detect heart problems using Linear Regression Analysis (LRA) on a dataset of clinical and diagnostic parameters. The objective is to develop a predictive model that can identify individuals at risk of heart problems based on easily obtainable medical information.

The dataset comprises a diverse range of features, including age, gender, blood pressure, cholesterol levels, and electrocardiogram (ECG) readings, collected from a large cohort of patients. After thorough preprocessing and feature selection, a linear regression model is trained to predict the likelihood of heart problems. The model's performance is evaluated using various metrics such as accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC-ROC).

The results indicate that LRA can effectively identify individuals at risk of heart problems with high accuracy. Furthermore, the model's interpretability allows healthcare professionals to gain insights into the importance of different risk factors in predicting heart problems. This approach holds promise for early detection and intervention, potentially reducing the burden of heart diseases on healthcare systems and improving patient outcomes.

In conclusion, this study demonstrates the feasibility and effectiveness of using Linear Regression Analysis as a tool for detecting heart problems. The model's transparency and predictive power make it a valuable addition to the arsenal of tools available for cardiovascular risk assessment and prevention. Further research and validation on larger and more diverse datasets will be essential to refine and generalize the approach for broader clinical use.

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

## 1.1 Introduction

According to the World Health Organization, 12 million deaths occur yearly due to heart disease. Load of cardiovascular disease is rapidly increasing all over the world in the past few years. Early detection of cardiac diseases can decrease the mortality rate and overall complications. However, it is not possible to monitor patients every day in all cases accurately and consultation with a patient for 24 hours by a doctor is not available since it requires more patience,time and expertise.

Our Heart Failure Prediction System is intended to assist patients in recognizing their heart state early and receiving treatment at an earlier stage, allowing them to avoid any serious conditions. We have designed this system using the Machine Learning model to predict the future possibility of heart disease by implementing the Logistic Regression algorithm.

The framework used in this project is Django. The Front End involves Html, CSS and JavaScript. The Back End involves MySQL Database. The Back End Language is Python.

## 1.2 Scope and Objective

The objective of a Heart Failure Prediction System is to identify individuals who are at high risk of developing heart failure or experiencing worsening symptoms. This system utilizes various algorithms and machine learning techniques to analyze a patient's medical history, lifestyle factors, and other relevant data to predict the likelihood of heart failure. By providing early identification and intervention, this system can help prevent or delay the onset of heart failure, improve patient outcomes, and reduce healthcare costs.

The scope of a Heart Failure Prediction System is broad, encompassing the entire patient population, from those who have no known risk factors to those who are already experiencing symptoms. The system can be integrated into various healthcare settings, such as hospitals, clinics,

and primary care physician offices, and can be utilized by healthcare professionals to support clinical decision-making.

## 1.3   Modules and their Description

The system comprises 1 major module with their sub-modules as follows:

**Register:**

- The user would need to register first to log in.

**Login:**

- The user can log in using their credentials.

**Prediction:**

The users have to provide the below inputs for the system to predict

if there is any heart failure or not.

1. Age: [years]

2. Sex: [M: Male, F: Female]

3. Chest Pain Type: [TA: Typical Angina, ATA: Atypical Angina, NAP: Non-Anginal Pain, ASY: Asymptomatic]

4. Resting Blood Pressure: [mm Hg]

5. Serum Cholesterol: [mm/dl]

6. Fasting Blood Sugar: [1: if FastingBS ¿ 120 mg/dl, 0: otherwise]

7. Resting Electrocardiogram Results: [Normal: Normal, ST: having ST-T wave abnormality (T wave inversions and/or ST elevation or depression of ¿ 0.05 mV), LVH: showing probable or definite left ventricular hypertrophy by Estes' criteria]

8. Maximum Heart Rate Achieved: [Numeric value between 60 and 202]

9. Exercise-Induced angina: [Y: Yes, N: No]

10. Oldpeak: = ST [Numeric value measured in depression]

11. ST Slope: The slope of the peak exercise ST segment [Up:upsloping, Flat: flat, Down: downsloping] The system will predict if there is any heart disease. The output class will be [1: heart disease, 0: Normal]

**Chatbot:** - The user can get to know the causes of heart failure and the diagnosis test required through the chatbot.

- It can also provide links to nearby hospitals/clinics that specialize in heart disease.

**View Free Checkup Camps:**

- The user can look for Free Check-up Camps.

**Login:**

- The admin can log in using their credentials.

**View Users:**

- The admin can view the list of users.

**Add free checkup camps details:**

- The admin can add free checkup camp details.

## 1.4 Existing System and Proposed System

**Problem with the current scenario**

• Presently, the major challenge of the medical industry is to predict cardiovascular diseases.

• There is a requirement for less expensive and more reliable methods to avoid the compounding effect of the disease in lowincome or developing countries.

**Drawbacks of the existing system**

• The traditional system did not always provide accurate predictions of heart failure.

• Some heart failure prediction systems do not provide clear explanations of how their algorithms work.

## 1.5 PROPOSED SYSTEM

Our Python-based Heart Failure Prediction System consists of 2 modules: User and Admin. The user would need to register first to log into the system. For the system to predict if there is heart failure or not, the user would require to give inputs. The parameters include Age, Sex, Chest Pain Type, Resting BP, Cholesterol, Fasting BS, Resting ECG, Maximum Heart Rate, Exercise-induced Agina, Oldpeak, and the slope of the peak exercise ST segment. After the user provides all these inputs, the system will detect if there is any heart disease. The chatbot in the system will inform the user about the causes of heart failure, and the diagnosis test required. It will also provide links to nearby hospitals/clinics that specializes in heart disease. The user can also check out some free checkup camps.

The admin can log in using their credentials. They can view the users using the system. They can also add free checkup camp details.

We have used Logistic Regression to develop this system. It is a significant machine learning algorithm because it can provide probabilities and classify new data using continuous and discrete datasets.
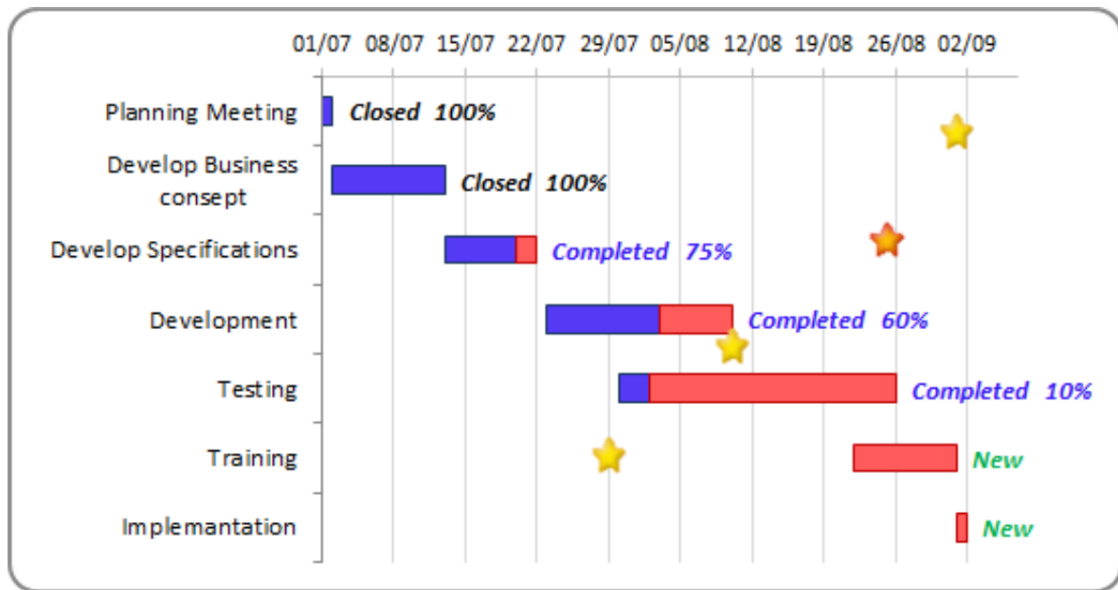
Figure 1.1: Gantt Chart

## 1.6   Project Lifecycle Details

**Description** The waterfall Model is a linear sequential flow. In which progress is seen as flowing steadily downwards (like a waterfall) through the phases of software implementation. This means that any phase in the development process begins only if the previous phase is complete.  The waterfall approach does not define the process to go back to the previous phase to handle changes in requirement. The waterfall approach is the earliest approach that was used for software development

**Waterfall Model**

Figure 1.2: Waterfall model

# Chapter 2

# Literature Survey
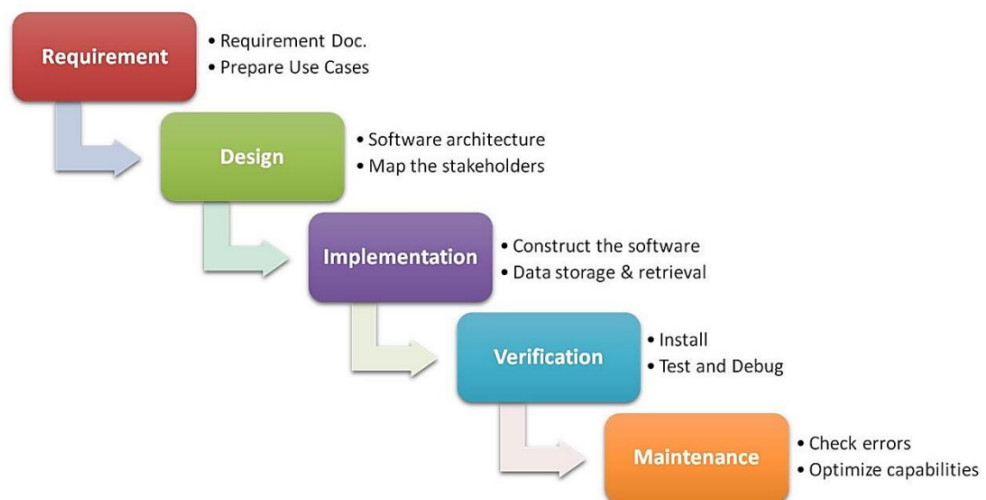
## 2.1 Paper 1

**TITLE:** Heart Disease Prediction Using Machine Learning Techniques

**AUTHOR:** Galla Siva Sai Bindhika

**PUBLISHED:** 04 Apr 2020

**SUMMARY:** Heart disease is one of the most significant problem that is arising in the world today. Cardiovascular disease prediction is a critical challenge in the area of clinical data analysis. Hybrid Machine learning (ML) has been showing an effective assistance in making decisions and predictions from the large quantity of data produced by the healthcare industries and hospitals. We have also seen ML techniques being used in recent developments in different areas of the Internet of Things (IoT). Various studies give only a glimpse in predicting heart disease with ML techniques. In this paper, we propose a narrative method that aims at finding significant features by applying machine learning techniques that results in improving the accuracy in the prediction of cardiovascular disease. The prediction model is proposed with combinations of different features and several classification techniques. We produce an enhanced performance level with an accuracy level of 92 through the prediction model for heart disease with the hybrid random forest with a linear model.

   **CONCLUSION:** In this paper, we proposed a method for heart disease prediction using machine learning techniques, these results showed a great accuracy standard for producing a better estimation result. By introducing new proposed Random forest classification, we find the problem of prediction rate without equipment and propose an approach to estimate the heart rate and condition. Sample results of heartrate are to be taken at different stages of the same subjects, we find the information from the above input via ML Techniques. Firstly, we introduced a support vector classifier based on

datasets

## 2.2 Paper 2

**TITLE:** Heart Disease Prediction

**AUTHOR:** Dr. Bal Krishna Bal

**PUBLISHED:** : 11th March 2020

**SUMMARY:** The major challenge in heart disease is its detection. There are instruments available which can predict heart disease but either they are expensive or are not efficient to calculate chance of heart disease in human. Early detection of cardiac diseases can decrease the mortality rate and overall complications. However, it is not possible to monitor patients every day in all cases accurately and consultation of a patient for 24 hours by a doctor is not available since it requires more sapience, time and expertise. Since we have a good amount of data in today's world, we can use various machine learning algorithms to analyze the data for hidden patterns. The hidden patterns can be used for health diagnosis in medicinal data.

**CONCLUSION:** The early prognosis of cardiovascular diseases can aid in making decisions on lifestyle changes in high risk patients and in turn reduce the complications, which can be a great milestone in the field of medicine. This project resolved the feature selection i.e. backward elimination and RFECV behind the models and successfully predict the heart disease, with 85was Logistic Regression. Further for its enhancement, we can train on models and predict the types of cardiovascular diseases providing recommendations to the users, and also use more enhanced models

## 2.3 Paper 3

**TITLE:** HEART DESEASE PREDICTION SYSTEM

**AUTHOR:** KENNEDY NGURE NGAR

**PUBLISHED:** January – April 2019

**SUMMARY:** Among all fatal disease, heart attacks diseases are considered as the most prevalent. Medical practitioners conduct different surveys on heart diseases and gather information of heart patients, their symptoms and disease progression. Increasingly are reported about patients with common diseases who have typical symptoms. In this fast moving world people want to live a very luxurious life so they work like a machine in order to earn lot of money and live a comfortable life therefore in this race they forget to take care of themselves, because of this there food habits change

their entire lifestyle change, in this type of lifestyle they are more tensed they have blood pressure, sugar at a very young age and they don't give enough rest for themselves and eat what they get and they even don't bother about the quality of the food if sick the go for their own medication as a result of all these small negligence it leads to a major threat that is the heart disease The term 'heart disease' includes the diverse diseases that affect heart. The number of people suffering from heart disease is on the rise (health topics, 2010). The report from world health organization shows us a large number of people that die every year due to the heart disease all over the world. Heart disease is also stated as one of the greatest killers in Africa. Data mining has been used in a variety of applications such as marketing, customer relationship management, engineering, and medicine analysis, expert prediction, web mining and mobile computing. Of late, data mining has been applied successfully in healthcare fraud and detecting abuse cases.

**CONCLUSION:** The proposed system is GUI-based, user-friendly, scalable, reliable and an expandable system. The proposed working model can also help in reducing treatment costs by providing Initial diagnostics in time. The model can also serve the purpose of training tool for medical students and will be a soft diagnostic tool available for physician and cardiologist. General physicians can utilize this tool for initial diagnosis of cardio-patients. There are many possible improvements that could be explored to improve the scalability and accuracy of this prediction system. As we have developed a generalized system, in future we can use this system for the analysis of different data sets. The performance of the health's diagnosis can be improved significantly by handling numerous class labels in the prediction process, and it can be another positive direction of research. In DM warehouse, generally, the dimensionality of the heart database is high, so identification and selection of significant attributes for better diagnosis of heart disease are very challenging tasks for future research.

## 2.4   paper 4

**TITLE:** Effective Heart Disease Prediction Using Machine Learning Techniques
**AUTHOR:** Chintan M. Bhat
**PUBLISHED:** 6 February 2023
**SUMMARY:**The diagnosis and prognosis of cardiovascular disease are crucial medical tasks to ensure correct classification, which helps cardiologists provide proper treatment to the patient. Machine learning applications in the medical niche have increased as they can recognize patterns

from data. Using machine learning to classify cardiovascular disease occurrence can help diagnosticians reduce misdiagnosis. This research develops a model that can correctly predict cardiovascular diseases to reduce the fatality caused by cardiovascular diseases. This paper proposes a method of k-modes clustering with Huang starting that can improve classification accuracy. Models such as random forest (RF), decision tree classifier (DT), multilayer perceptron (MP), and XGBoost (XGB) are used. GridSearchCV was used to hypertune the parameters of the applied model to optimize the result. The proposed model is applied to a real-world dataset of 70,000 instances from Kaggle. Models were trained on data that were split in 80:20 and achieved accuracy as follows: decision tree: 86.37cross-validation) and 86.53and 87.02(without cross-validation), multilayer perceptron: 87.28cross-validation). The proposed models have AUC (area under the curve) values: decision tree: 0.94, XGBoost: 0.95, random forest: 0.95, multilayer perceptron: 0.95. The conclusion drawn from this underlying research is that multilayer perceptron with cross-validation has outperformed all other algorithms in terms of accuracy. It achieved the highest accuracy of 87.28 **CONCLUSION:** The primary objective of this study was to classify heart disease using different models and a real-world dataset. The k-modes clustering algorithm was applied to a dataset of patients with heart disease to predict the presence of the disease. The dataset was preprocessed by converting the age attribute to years and dividing it into bins of 5-year intervals, as well as dividing the diastolic and systolic blood pressure data into bins of 10 intervals. The dataset was also split on the basis of gender to take into account the unique characteristics and progression of heart disease in men and women.

## 2.5  Paper 5

**TITLE:**Heart Disease Prediction System using Associative Classification and Genetic Algorithm

**AUTHOR:** M.Akhil jabbar

**PUBLISHED:** 2012

**SUMMARY:** Associative classification is a recent and rewarding technique which integrates association rule mining and classification to a model for prediction and achieves maximum accuracy. Associative classifiers are especially fit to applications where maximum accuracy is desired to a model for prediction. There are many domains such as medical where the maximum accuracy of the model is desired. Heart disease is a single largest cause of death in developed countries and one of the main contributors to disease burden in developing countries. Mortality data from the registrar general of India shows that heart disease are a major cause of death in India, and

in Andhra Pradeshcoronary heart disease cause about 30system for predicting heart disease of a patient. In this paper we propose efficient associative classification algorithm using genetic approach for heart disease prediction. The main motivation for using genetic algorithm in the discovery of high level prediction rules is that the discovered rules are highly comprehensible, having high predictive accuracy and of high interestingness values. Experimental Results show that most of the classifier rules help in the best prediction of heart disease which even helps doctors in their diagnosis decisions.

**CONCLUSION:** In the recent years India and other developing countries have witnessed a rapidly escalating epidemic of cardiovascular disease (CVD).It is predicted that by 2020 coronary heart disease will be leading cause of death in adult Indians, and Andhra Pradesh is in risk of more deaths due to CVD.The need to contain the epidemic of cardiovascular disease as well as combat its impact and minimize its toll on Andhra Pradesh is obvious and urgent. Hence a decision support system is proposed to identify a risk score for predicting the heart disease. In this paper, we proposed a system for heart disease prediction using data mining techniques. In our feature work we plan to reduce no. of attributes and to determine the attribute which contribute towards the diagnosis of disease using genetic algorithm.

## 2.6 paper 6

**TITLE:**Improving the Accuracy for Analyzing Heart Diseases Prediction Based on the Ensemble Method

**AUTHOR:** Xiao-Yan Gao

**PUBLISHED:** 2021

**SUMMARY:** Nowadays, the cardiac disease is one of the most critical problems relating to human safety. -e treatment of heart problems has recently been stated in a study that has received huge attention in the medical system worldwide. Cardiac diseases are one of the most principal causes of death worldwide. On median, 17.7 million deaths result from heart disease which counts for about 312016, according to World Health Organization (WHO) . -e cardiac cases number, as the focus of this study, shows that 82million are under 70 years of age and prone to noninfectious diseases, 6.7 million are affected by stroke, and 7.4 million people are suffering from heart disease (WHO, 2016) [2]. In the US and other developed countries, about half of all deaths are caused by heart disease; also, one-third of all people's deaths worldwide are related to heart disease. Cardiac disease affects not just people's health but the economies and costs of countries as well. -e most common cardiac

disorders are those of microvascular origin, primarily cardiac disorders and stroke. After several years of exposure to unhealthy lifestyles, cardiovascular disease clinically presents itself in early stages of life, as well as at an old age. -e main cardiac medical conditions include overweight, diabetes, family history, smoking, and high cholesterol.

**CONCLUSION:** In this paper, we developed the proposed system to predict heart disease. Ensemble methods (boosting and bagging) with feature extraction algorithms (PCA and LDA) are used to improve predicting heart disease performance. -e feature extraction algorithms are used to extract essential features from the Cleveland heart disease dataset. Comparison between ensemble methods (boosting and bagging) and five classifiers (KNN, SVM, NB, DT, and RF) is applied to selected features. -e experimental results showed that the bagging ensemble learning algorithm with DT and PCA feature extraction method had achieved the best performance.

## 2.7  Paper 7

**TITLE:** Heart Disease Prediction using Artificial Intelligence

**AUTHOR:** Zaibunnisa L. H. Malik

**PUBLISHED:** 2021

**SUMMARY:** Artificial Intelligence techniques have been widely used in clinical decision support systems for prediction and diagnosis of various diseases with good accuracy. These classifying techniques are very effective in designing clinical support systems due to their ability to get hidden patterns and relationships in medical data provided by medical professionals. One of the most important applications of such systems is in the diagnosis of heart diseases because it is one of the leading causes of deaths all over the world. Almost all systems that predict heart diseases using clinical dataset having parameters andinputs from complex tests conducted in labs. None of the systems predicts heart diseases supporting risk factors like age, case history, diabetes, hypertension, high cholesterol, tobacco smoking, alcohol intake, obesity or physical inactivity, etc. Heart disease patients have many of those visible risk factors in common which may be used very effectively for diagnosis. A system based on such risk factors would not only help medical professionals but it would give patients a warning about the probable presence of heart disease even before the patient visits a hospital or goes for costly medical check-ups. Hence this paper presents a technique for prediction of heart disease using major risk factors with help of different Classifying Algorithms. This technique involves four major classification algorithms such as K Neighbors, Support Vector, Decision Tree, Random Forest algorithms.

**CONCLUSION:** The proposed application uses Risk Factors, which need to be identified by Medical Professionals before using the application. The result may vary based on the identified Risk Factors. If the Risk Factors identified are less accurate or wrong, the application may give wrong results. The application may use different AI techniques to capture and correct response based on past experiences. The result of the application depends on the accuracy of the Classification Algorithms. If the accuracy is low, the result generated may be wrong or less accurate. Increasing the dataset, may result in more accurate results. We can build an intelligent system which could predict the disease using risk factors hence saving cost and time to undergo medical tests and check-ups and ensuring that the patient can monitor his health on his own and plan preventive measures and treatment at the earlystages of the diseases.

# Chapter 3

# PROJECT DESIGN AND PROJECT IMPLEMENTATION

## 3.1   PROJECT DESIGN

The Project is designed and developed in Django Framework. We used Django Framework for coding of the project. Created and maintained all databases into MySQL Server, in that we create tables, write query for store data or record of project.

Python is a powerful multi-purpose programming language created by Guido van Rossum. It has simple easy-to-use syntax, making it the perfect language for someone trying to learn computer programming for the first time. This is a comprehensive guide on how to get started in Python, why you should learn it and how you can learn it. However, if your knowledge of other programming languages and want to quickly get started with Python. Python is a general-purpose language. It has wide range of applications from Web development (like: Django and Bottle), scientific and mathematical computing (Orange, SymPy, NumPy) to desktop graphical user Interfaces (Pygame, Panda3D). The syntax of the language is clean and length of the code is relatively short. It's fun to work in Python because it allows you to think about the problem rather than focusing on the syntax. A simple language which is easier to learn Python has a very simple and elegant syntax. It's much easier to read and write Python programs compared to other languages like: C++, Java, C. Python makes programming fun and allows you to focus on the solution rather than syntax. If you are a newbie, it's a great choice to start your journey with Python
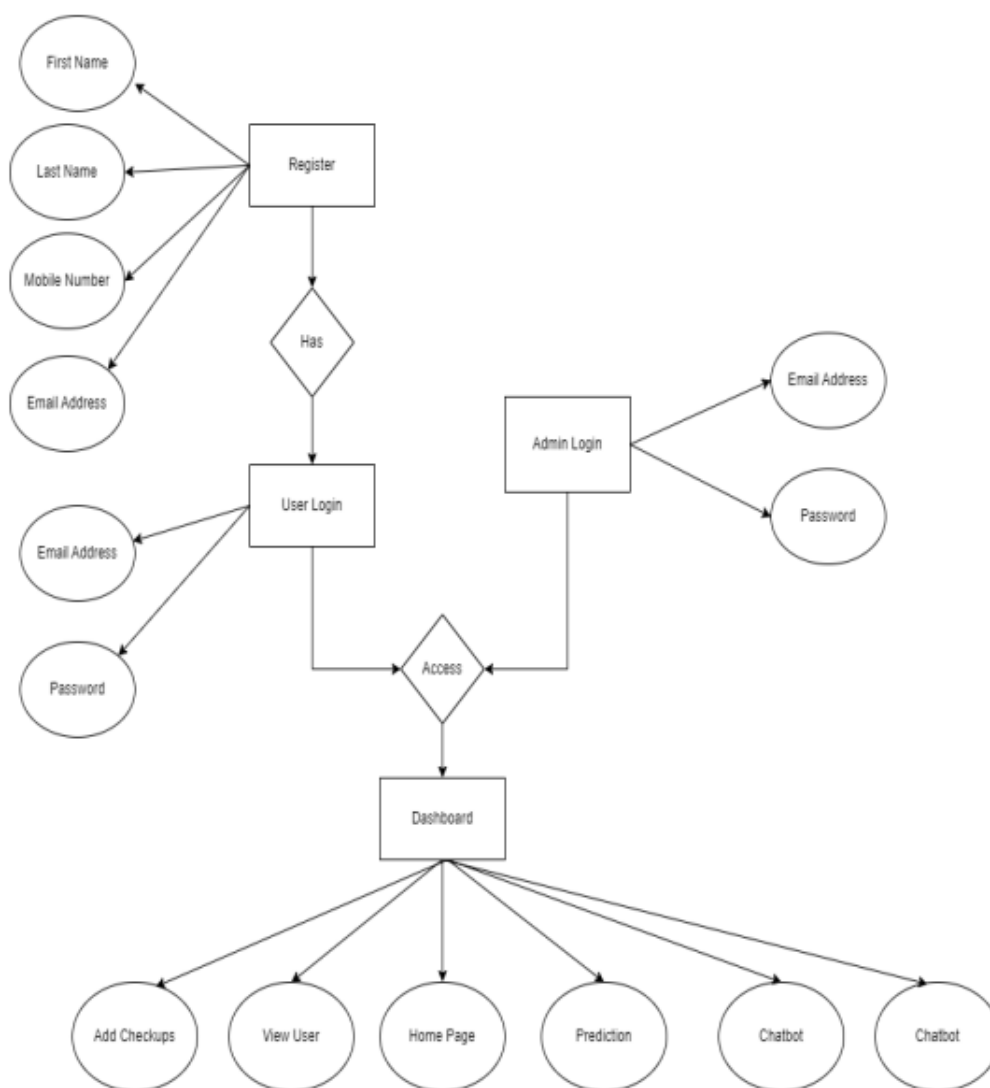
## PROJECT DESIGN
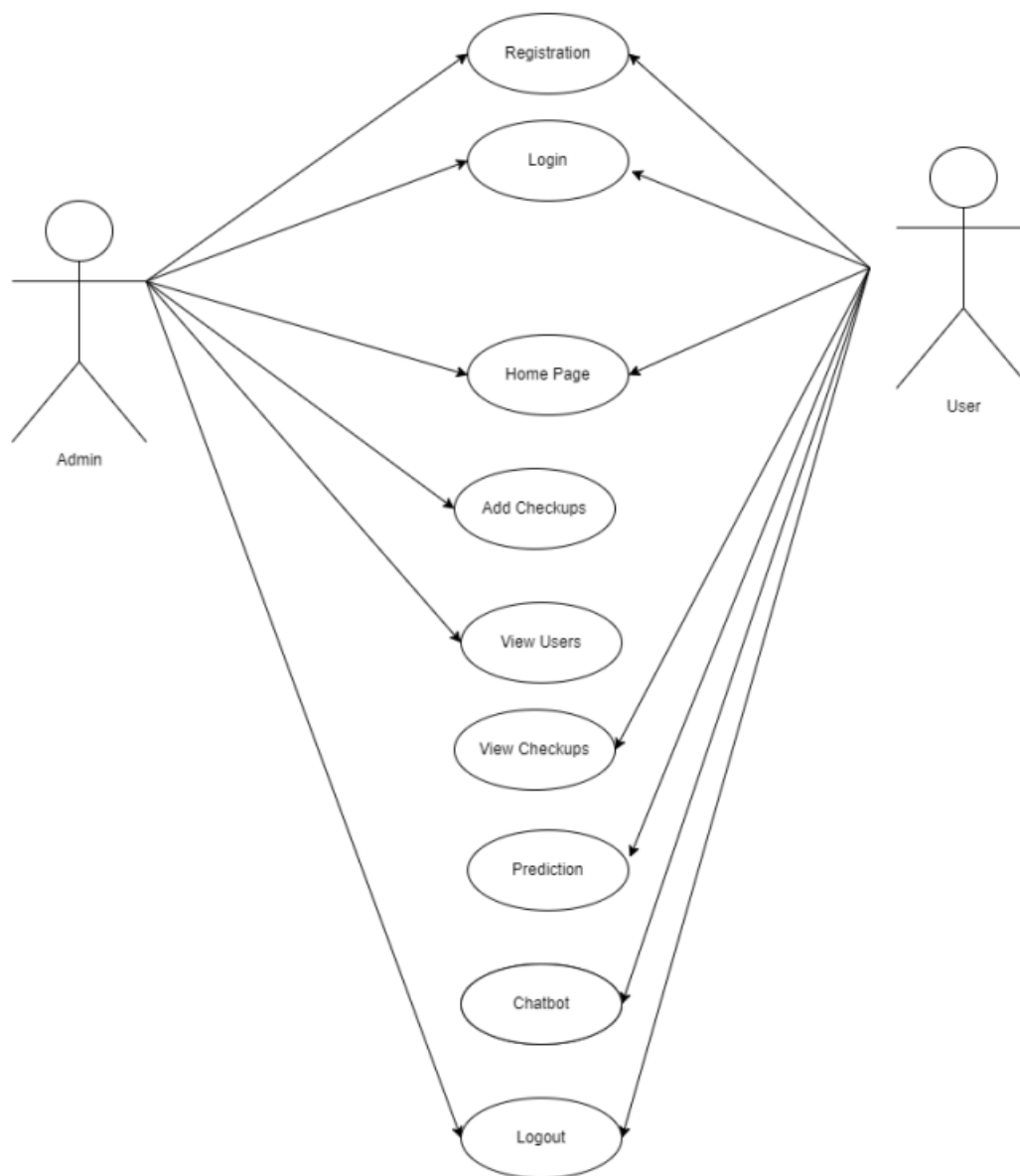
### E-R Diagram:



Figure 3.1: E-R Diagram:

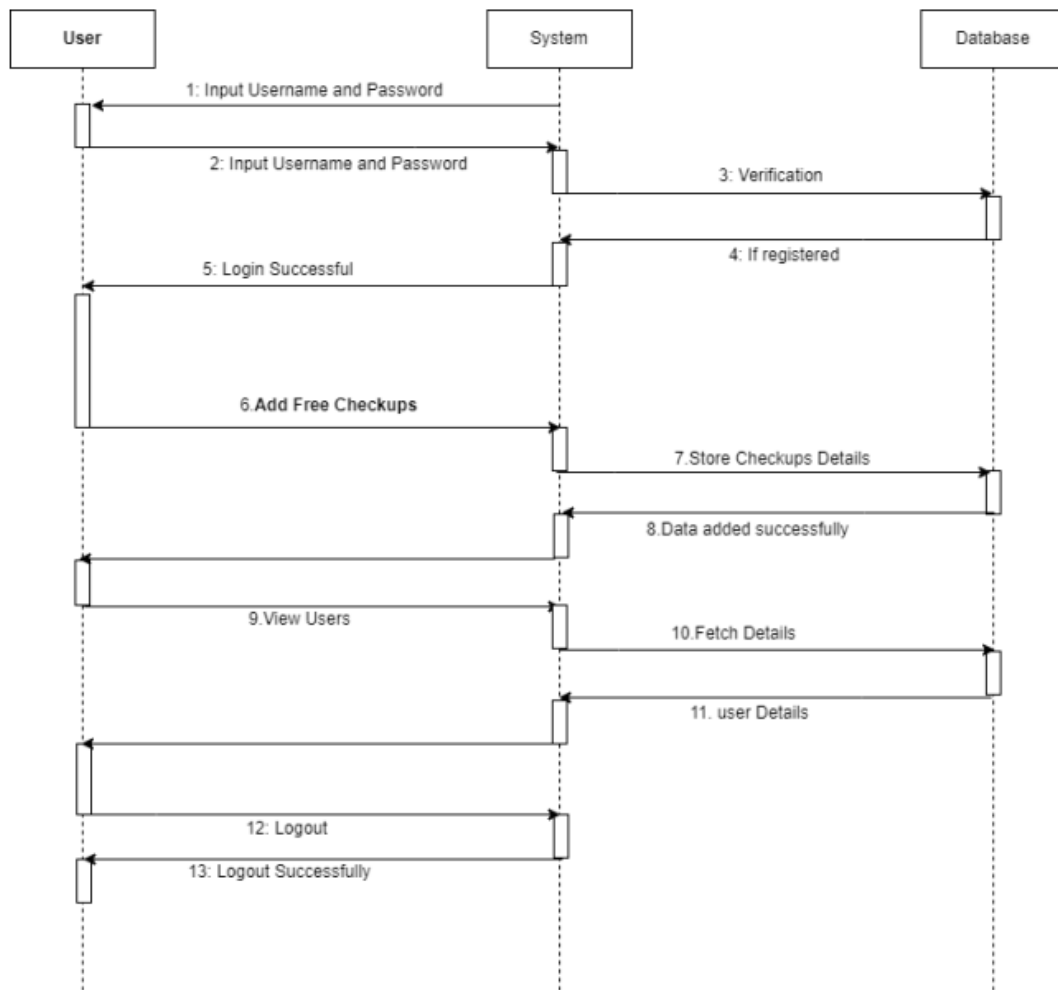## Use Case Diagram:



Figure 3.2: Use Case Diagram

## Sequence Diagram:


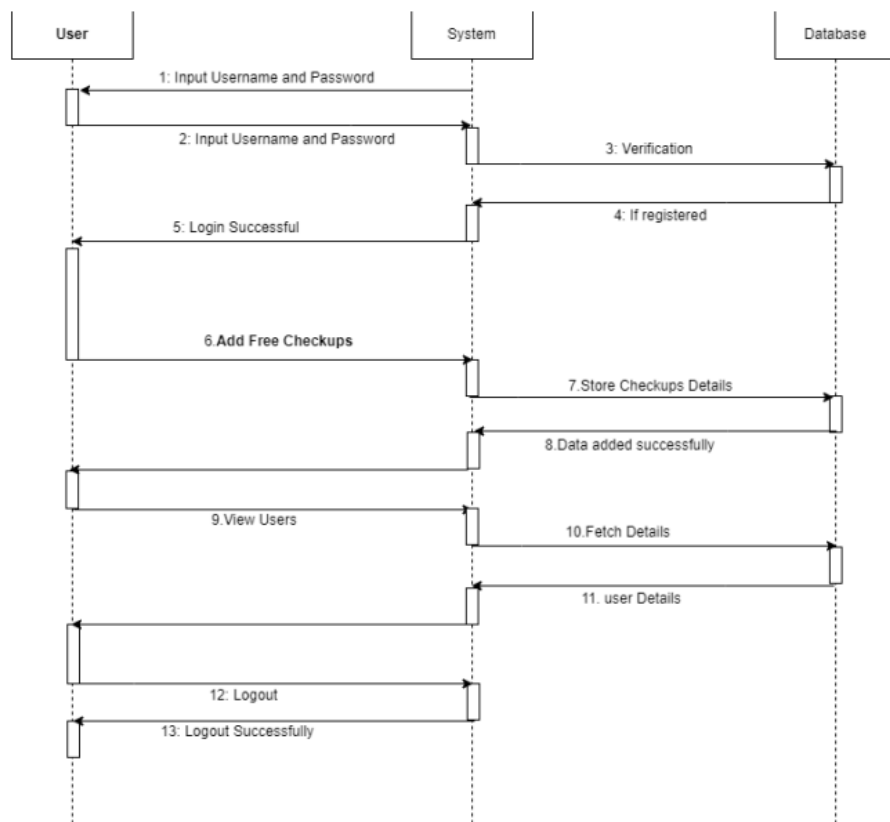
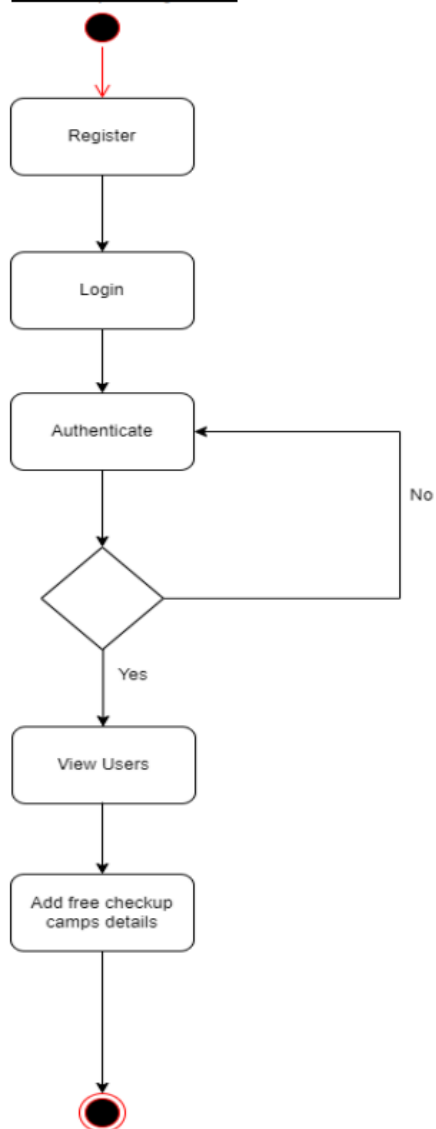Figure 3.3: Sequence Diagram of User

Figure 3.4: Sequence Diagram of Admin
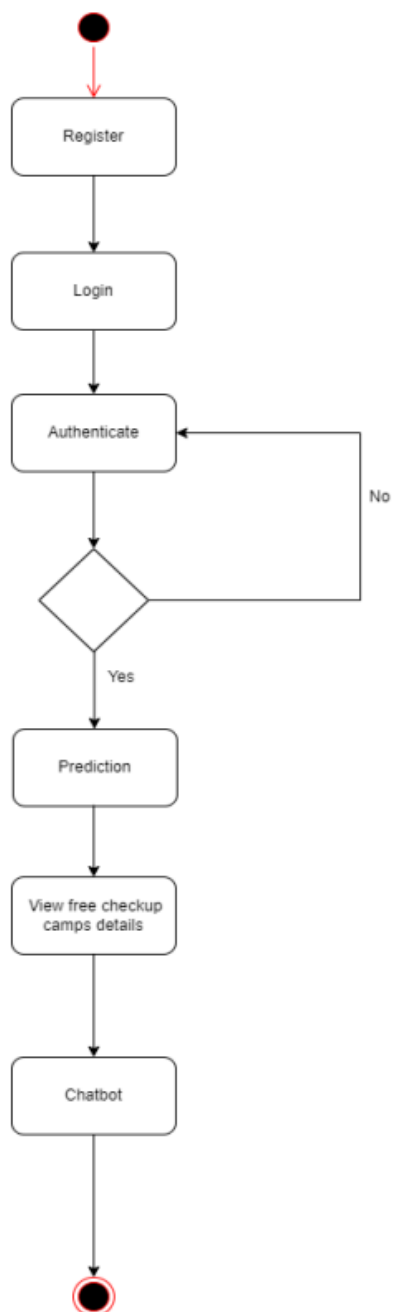
Activity Diagram:



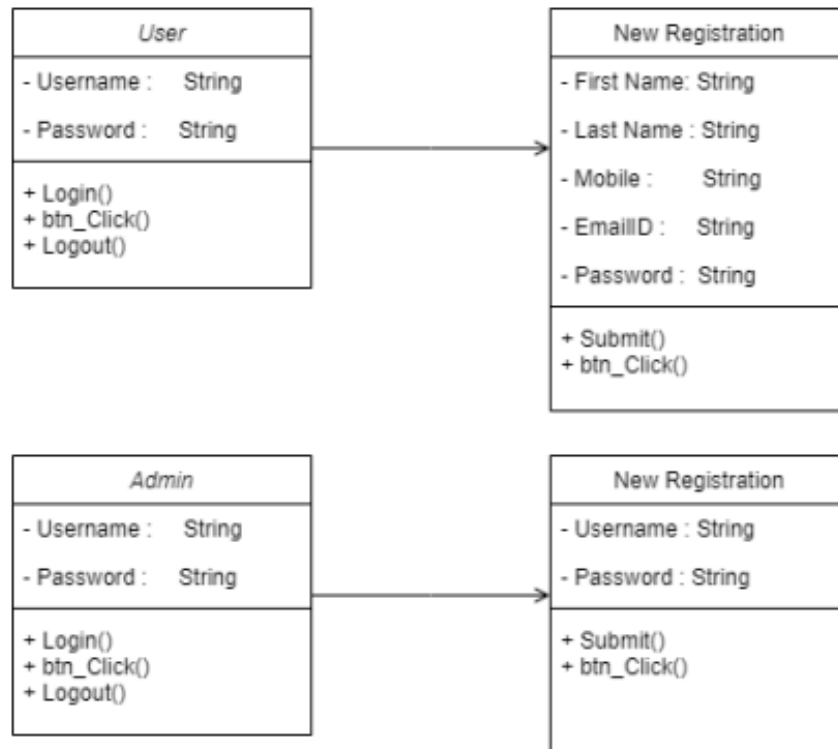Figure 3.5: Activity Diagram of Admin

Figure 3.6: Activity Diagram of user

## Class Diagram:

| User |
| --- |
| - Username : String |
| - Password : String |
| + Login() |
| + btn_Click() |
| + Logout() |

| New Registration |
| --- |
| - First Name: String |
| - Last Name : String |
| - Mobile : String |
| - EmailID : String |
| - Password : String |
| + Submit() |
| + btn_Click() |

| Admin |
| --- |
| - Username : String |
| - Password : String |
| + Login() |
| + btn_Click() |
| + Logout() |

| New Registration |
| --- |
| - Username : String |
| - Password : String |
| + Submit() |
| + btn_Click() |

Figure 3.7: Class Diagram

## 3.2   Data Flow Diagram (DFD's)

A data flow diagram is graphical tool used to describe and analyze movement of data through a system. These are the central tool and the basis from which the other components are developed. The transformation of data from input to output, through processed, may be described logically and independently of physical components associated with the system. These are known as the logical data flow diagrams. The physical data flow diagrams show the actual implements and movement of data between people, departments and workstations. A full description of a system actually consists of a set of data flow diagrams. Using two familiar notations Yourdon, Gane and Sarson notation develops the data flow diagrams. Each component in a DFD is labeled with a descriptive name. Process is further identified with a number that will be used for identification purpose. The development of DFD's is done in several levels. Each process in lower level diagrams can be broken down into a more detailed DFD in the next level. The lop-level diagram is often called context diagram. It consists a single process bit, which plays vital role in studying the current system. The process in the context level diagram is exploded into other process at the first level DFD.

The idea behind the explosion of a process into more process is that understanding at one level of detail is exploded into greater detail at the next level. This is done until further explosion is necessary and an adequate amount of detail is described for analyst to understand the process.

Larry Constantine first developed the DFD as a way of expressing system requirements in a graphical from, this lead to the modular design.

A DFD is also known as a "bubble Chart" has the purpose of clarifying system requirements and identifying major transformations that will become programs in system design. So it is the starting point of the design to the lowest level of detail. A DFD consists of a series of bubbles joined by data flows in the system.

**DFD SYMBOLS:** In the DFD, there are four symbols

1. A square defines a source(originator) or destination of system data

2. An arrow identifies data flow. It is the pipeline through which the information flows

3. A circle or a bubble represents a process that transforms incoming data flow into outgoing data

data

Process that transforms data flow.

Source or Destination of data
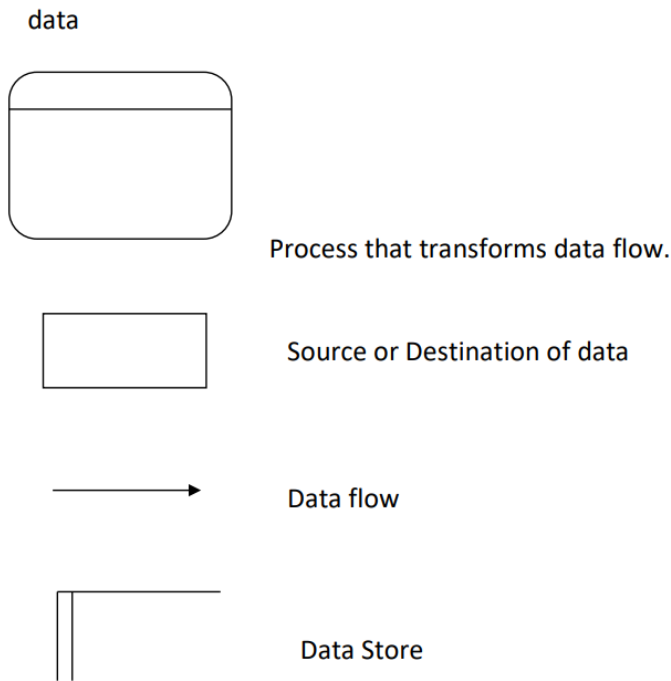
Data flow

Data Store

Figure 3.8: Data Flow Diagram

flows.

4. An open rectangle is a data store, data at rest or a temporary repository of data

**CONSTRUCTING A DFD:** Several rules of thumb are used in drawing DFD's:

1. Process should be named and numbered for an easy reference. Each name should be representative of the process

2. The direction of flow is from top to bottom and from left to right. Data traditionally flow from source to the destination although they may flow back to the source. One way to indicate this is to draw long flow line back to a source. An alternative way is to repeat the source symbol as a destination. Since it is used more than once in the DFD it is marked with a short diagonal.

3. When a process is exploded into lower level details, they are numbered

4. The names of data stores and destinations are written in capital letters. Process and dataflow names have the first letter of each work capitalized

A DFD typically shows the minimum contents of data store. Each data store should contain all the data elements that flow in and out. Questionnaires should contain all the data elements that flow

in and out. Missing interfaces redundancies and like is then accounted for often through interviews.

## 3.3 RULES GOVERNING THE DFD'S

**PROCESS**

1) No process can have only outputs.

2) No process can have only inputs. If an object has only inputs than it must be a sink.

3) A process has a verb phrase label.

**DATA STORE**

1) Data cannot move directly from one data store to another data store, a process must move data.

2) Data cannot move directly from an outside source to a data store, a process, which receives, must move data from the source and place the data into data store

3) A data store has a noun phrase label.

**SOURCE OR SINK**

The origin and /or destination of data.

1) Data cannot move direly from a source to sink it must be moved by a process

2) A source and /or sink has a noun phrase land

**DATA FLOW**

1) A Data Flow has only one direction of flow between symbols. It may flow in both directions between a process and a data store to show a read before an update. The later it usually indicated however by two separate arrows since these happen at different type.

2) A join in DFD means that exactly the same data comes from any of two or more different processes data store or sink to a common location.

3) A data flow cannot go directly back to the same process it leads. There must be at least one other process that handles the data flow produce some other data flow returns the original data into the beginning process.

4) A Data flow to a data store means update (delete or change).

5) A data Flow from a data store means retrieve or use.
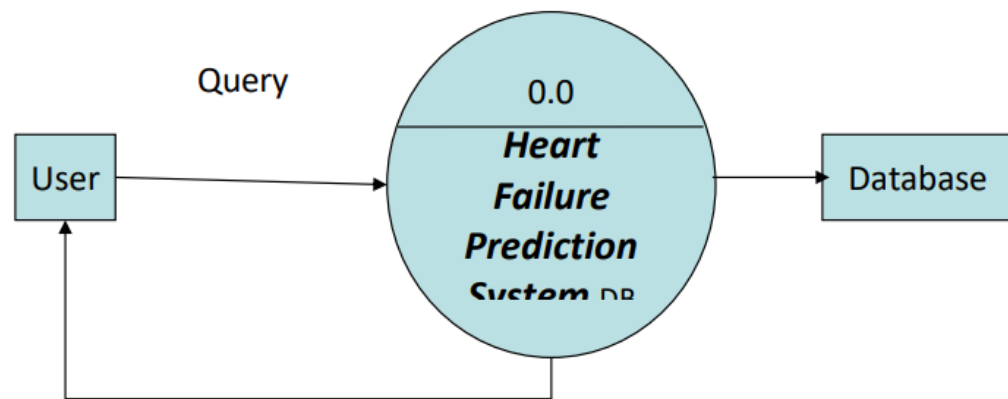
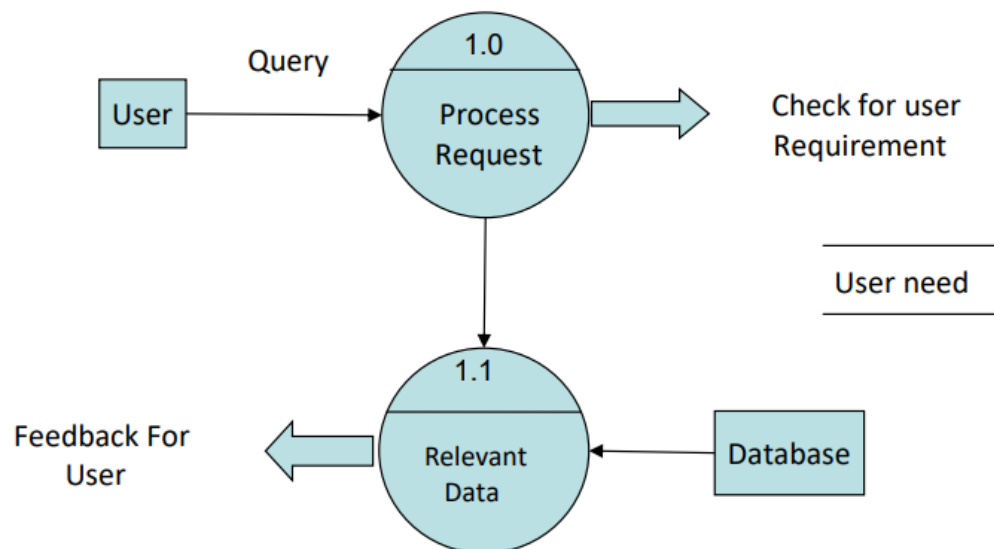**Data Flow Diagrams (DFD's)**

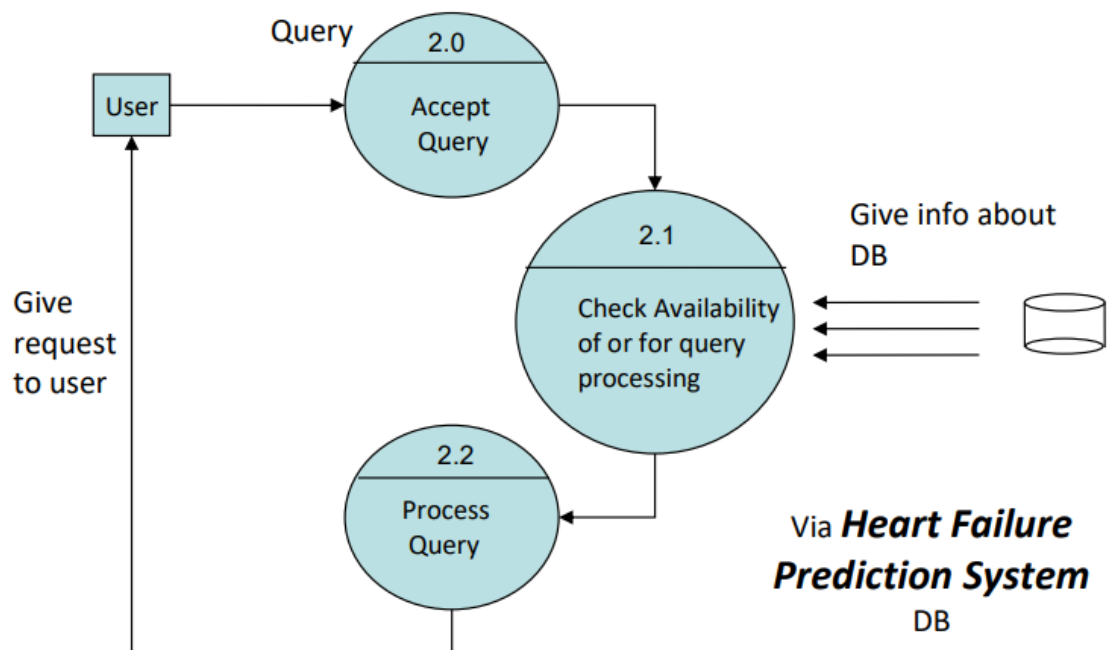Figure 3.9: DataBase Detail



Figure 3.10: Level 1 DFD
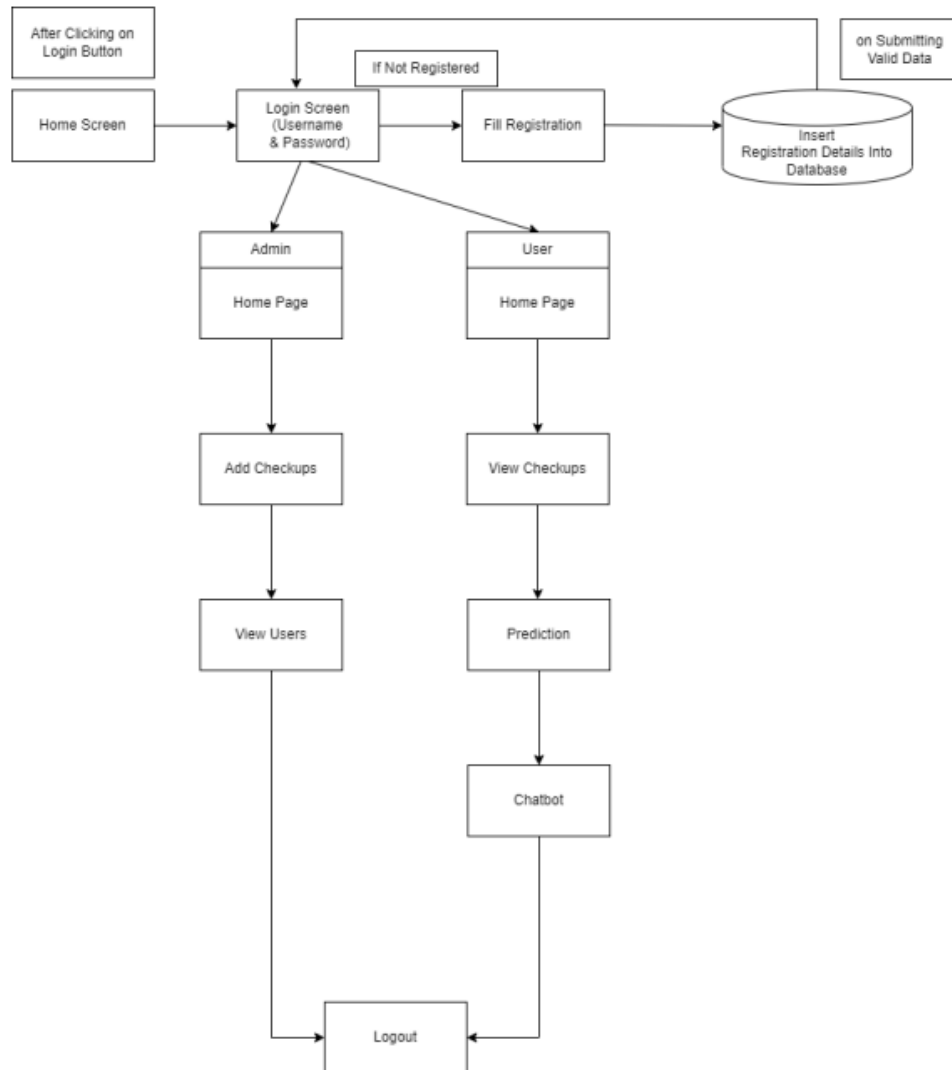
Figure 3.11: Heart Failure Prediction System DB

## System Architecture:



Figure 3.12: System Architecture

## 3.4 PROJECT IMPLEMENTATION

**Project Implementation Technology** The Project is designed and developed in Django Framework. We used Django Framework for coding of the project. Created and maintained all databases into MySQL Server, in that we create tables, write query for store data or record of project.

**Hardware Requirement**

Laptop or PC

1. Windows 7 or higher

2. I3 processor system or higher

3. 4 GB RAM or higher

3. 100 GB ROM or higher

**Software Requirement**

Laptop or PC

1. Python

2. Sublime text Editor

3. XAMP Server


## 3.5 OVERVIEW OF TECHNOLOGIES USED

**Features of Python Programming:**

A simple language which is easier to learn Python has a very simple and elegant syntax. It's much easier to read and write Python programs compared to other languages like: C++, Java, . Python makes programming fun and allows you to focus on the solution rather than syntax. If you are a newbie, it's a great choice to start your journey with Python.

Free and open-source You can freely use and distribute Python, even for commercial use. Not only can you use and distribute software's written in it, you can even make changes to the Python's source code. Python has a large community constantly improving it in each iteration.

**Django**

Django is a high-level Python Web framework that encourages rapid development and clean, pragmatic design. Built by experienced developers, it takes care of much of the hassle of Web development, so you can focus on writing your app without needing to reinvent the wheel. It's free and open source.

**Features of Django**

Rapid Development

Secure

Scalable

Fully loaded

Versatile

Open Source

Vast and Supported Community

**Rapid Development**

Django was designed with the intention to make a framework which takes less time to build web application. The project implementation phase is a very time taken but Django creates it rapidly.

**Open Source**

Django is an open source web application framework. It is publicly available without cost. It can be downloaded with source code from the public repository. Open source reduces the total cost of the application development.

**WAMP Server**

WAMP is a Windows OS based program that installs and configures Apache web server, MySQL database server, PHP scripting language, phpMyAdmin (to manage MySQL database's), and SQLiteManager (to manage SQLite database's). WAMP is designed to offer an easy way to install Apache, PHP and MySQL package with an easy to use installation program instead of having to install and configure everything yourself. WAMP is so easy because once it is installed it is ready to go. You don't have to do any additional configuring or tweaking of any configuration files to get it running.

There are usually two reasons why someone chooses to install WAMP. They are looking to install WAMP for development purposes or to run their own server.

**WAMP Server Contains**

Allows you to change or add users and for making new databases phpMyAdmin is a free software tool written in PHP, intended to handle the administration of MySQL over the World Wide Web. phpMyAdmin supports a wide range of operations with MySQL. The most frequently used operations are supported by the user interface (managing databases, tables, fields, relations, indexes, users, permissions, etc.), while you still have the ability to directly execute any SQL statement.
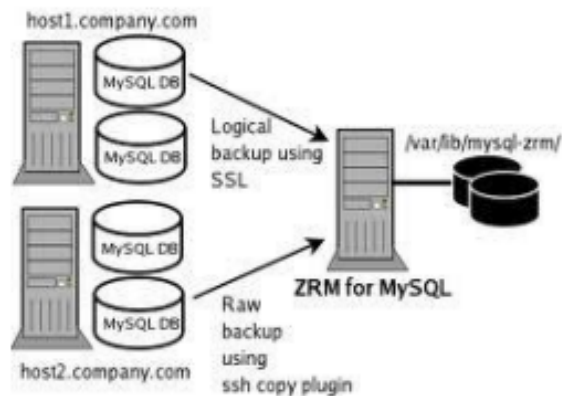
Figure 3.13: My SQL



Figure 3.14: SQL Server and Database System

**SQL Server and Database System**

SQL Server is a relational database management system from Microsoft that's designed for the enterprise environment. SQL Server runs on T-SQL (Transact -SQL), a set of programming extensions from Sybase and Microsoft that add several features to standard SQL, including transaction control, exception and error handling, row processing, and declared variables.

Generically, any database management system (DBMS) that can respond to queries from client machines formatted in the SQL language. When capitalized, the term generally refers to either of two database management products from Sybase and Microsoft. Both companies offer client-server DBMS products called SQL Server.

**Using WAMP as a Development Server**

You can use WAMP to develop and test websites locally on their own computer instead of having to get a web hosting account to develop with. Most people will be using WAMP for development purposes such as learning how to create websites with HTML, PHP, and MySQl

WARNING: WAMP was designed to be a testing and development server, not an actual production server. WAMP does not come with any real security in place so it offers no protection from any kind of attack. Any 10-yearold with access to the internet can easily hack your WAMP server.

If your website(s) have highly sensitive data (such as credit card numbers, social security numbers, user ids, passwords, etc.), you need to take this in consideration before your put this information online. Unless you are an experienced system administrator and can configure WAMP to be more secure, you should never user WAMP for a production server. MySQL Configuration

To begin MySQL installation, first download latest version of Essentials as an MSI package. During MySQL installation, select Typical installation and use default configuration values except for Sign-Up where you probably want to select Skip Sign-Up. When Setup Wizard is completed, make sure the option Configure the MySQL Server now is set. For MySQL Server Instance Configuration, select Standard Configuration. Next, you must set option Include Bin Directory in Windows PATH. This setting is crucial, otherwise a required library, libMySQL.dll, will not be found later during Apache startup.
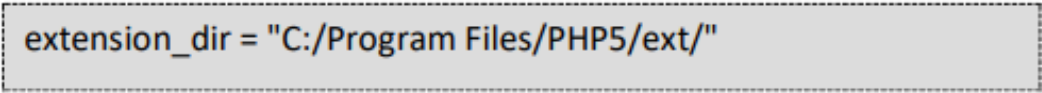
Finally, enter a proper root password. There is no need to neither enable remote root access nor create an Anonymous Account.

Please inspect messages during MySQL startup and verify that MySQL has been started successfully. Then, you must reboot the system. Otherwise, the required librarylibMySQL.dll will not be found during Apache startup when Apache is trying to load Apache's PHP module and Apache will, perhaps a bit confusingly, complain that it is unable to load the PHP's MySQL library, php mysql.dll. Therefore, it is necessary to reboot the system at this stage and then continue to PHP configuration.

**PHP Configuration**

PHP for Windows must be installed from the zip package, not using the installer because the installer does not work correctly when setting up the configuration files. Download the latest Windows binary version from the 5.x release series.

Create folder C: Program Files PHP5 and unzip the downloaded package there. Then, in folder C:

```
extension_dir = "C:/Program Files/PHP5/ext/"
```

Figure 3.15: PHP Configuration

```
extension=php_mysql.dll
```

Figure 3.16: PHP Configuration

Program Files PHP5 you need to copy the file php.in irecommended asphp.ini and make two changes into the php.ini file. Change extension dir to:

That is all what is needed for PHP configuration. Additionally, however, if you wish to run PHP from the command line it would be useful to add its installation directory to Windows PATH but for WAMP to operate it is not required. After Apache is installed and configured, also PHP configuration can be tested.

## 3.6   FEASIBILITY REPORT

Feasibility Study is a high level capsule version of the entire process intended to answer a number of questions like: What is the problem? Is there any feasible solution to the given problem? Is the problem even worth solving? Feasibility study is conducted once the problem clearly understood. Feasibility study is necessary to determine that the proposed system is Feasible by considering the technical, Operational, and Economical factors. By having a detailed feasibility study the management will have a clear-cut view of the proposed system.

The following feasibilities are considered for the project in order to ensure that the project is variable and it does not have any major obstructions. Feasibility study encompasses the following things:

1. Technical Feasibility

2. Economic Feasibility

3. Operational Feasibility

In this phase, we study the feasibility of all proposed systems, and pick the best feasible solution for the problem. The feasibility is studied based on three main factors as follows.

**Technical Feasibility**

In this step, we verify whether the proposed systems are technically feasible or not. i.e., all the technologies required to develop the system are available readily or not.

Technical Feasibility determines whether the organization has the technology and skills necessary to carry out the project and how this should be obtained. The system can be feasible because of the following grounds:

1. All necessary technology exists to develop the system.

2. This system is too flexible and it can be expanded further.

3. This system can give guarantees of accuracy, ease of use, reliability and the data security.

4. This system can give instant response to inquire.

Our project is technically feasible because, all the technology needed for our project is readily available

**Operating System** = Windows 7 or higher

**Languages**  = python

**Database System** = My SQL 5.6

**Documentation Tool** = MS - Word

**Economic Feasibility**

Economically, this project is completely feasible because it requires no extra financial investment and with respect to time, it's completely possible to complete this project in 6 months.

In this step, we verify which proposal is more economical. We compare the financial benefits of the new system with the investment. The new system is economically feasible only when the financial benefits are more than the investments and expenditure. Economic Feasibility determines whether the project goal can be within the resource limits allocated to it or not. It must determine whether it is worthwhile to process with the entire project or whether the benefits obtained from the new system are not worth the costs. Financial benefits must be equal or exceed the costs. In this issue, we should consider:

1. The cost to conduct a full system investigation.

2. The cost of h/w and s/w for the class of application being considered.

3. The development tool.

4. The cost of maintenance etc...

Our project is economically feasible because the cost of development is very minimal when compared to financial benefits of the application.

**Operational Feasibility**

In this step, we verify different operational factors of the proposed systems like man-power, time etc., whichever solution uses less operational resources, is the best operationally feasible solution. The solution should also be operationally possible to implement. Operational Feasibility determines if the proposed system satisfied user objectives could be fitted into the current system operation.

1. The methods of processing and presentation are completely accepted.

by the clients since they can meet all user requirements. 2. The clients have been involved in the planning and development of the system.

3. The proposed system will not cause any problem under any circumstances.

Our project is operationally feasible because the time requirements and personnel requirements are satisfied. We are a team of four members and we worked on this project for three working months.

## 3.7   TESTING

As the project is on bit large scale, we always need testing to make it successful. If each components work properly in all respect and gives desired output for all kind of inputs then project is said to be successful. So the conclusion is-to make the project successful, it needs to be tested.

The testing done here was System Testing checking whether the user requirements were satisfied. The code for the new system has been written completely using python as the coding language, Django as the interface for front-end designing. The new system has been tested well with the help of the users and all the applications have been verified from every nook and corner of the user.

Although some applications were found to be erroneous these applications have been corrected before being implemented. The flow of the forms has been found to be very much in accordance with the actual flow of data.
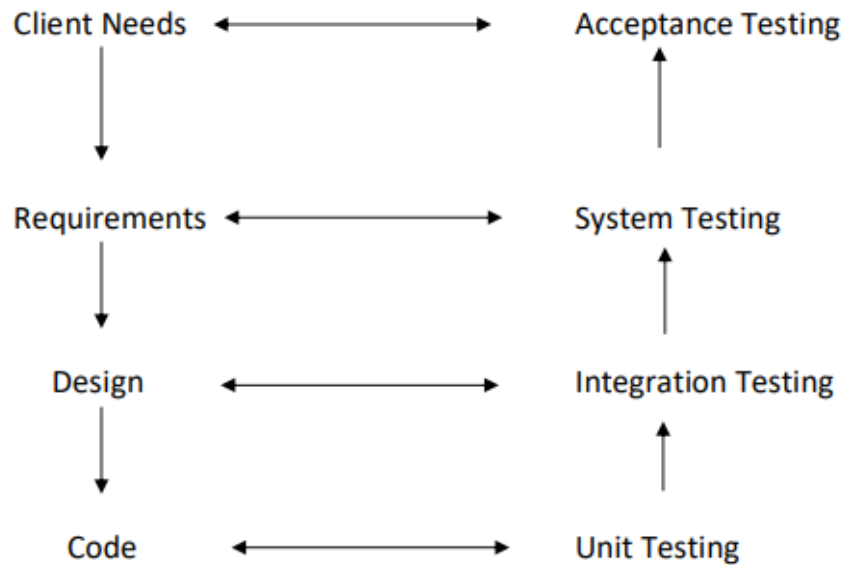
**Levels of Testing**

Figure 3.17: Levels of Testing

In order to uncover the errors present in different phases we have the concept of levels of testing. The basic levels of testing are:

A series of testing is done for the proposed system before the system is ready for the user acceptance testing

**Unit Testing**

Unit testing focuses verification efforts on the smallest unit of the software design, the module. This is also known as "Module Testing". The modules are tested separately. This testing carried out during programming stage itself. In this testing each module is found to be working satisfactorily as regards to the expected output from the module.

**Integration Testing**

Data can be grossed across an interface; one module can have adverse efforts on another. Integration testing is systematic testing for construction the program structure while at the same time conducting tests to uncover errors associated with in the interface. The objective is to take unit tested modules and build a program structure. All the modules are combined and tested as a whole. Here correction is difficult because the isolation of cause is complicate by the vast expense of the entire program. Thus in the integration testing stop, all the errors uncovered are corrected for the text testing steps.

**System testing**

System testing is the stage of implementation that is aimed at ensuring that the system works accurately and efficiently for live operation commences. Testing is vital to the success of the system. System testing makes a logical assumption that if all the parts of the system are correct, then goal will be successfully achieved.

**Validation Testing**

At the conclusion of integration testing software is completely assembled as a package, interfacing errors have been uncovered and corrected and a final series of software tests begins, validation test begins. Validation test can be defined in many ways. But the simple definition is that validation succeeds when the software function in a manner that can reasonably expected by the customer. After validation test has been conducted one of two possible conditions exists.

One is the function or performance characteristics confirm to specifications and are accepted and the other is deviation from specification is uncovered and a deficiency list is created. Proposed system under consideration has been tested by using validation testing and found to be working satisfactorily.

**Output Testing**

After performing validation testing, the next step is output testing of the proposed system since no system could be useful if it does not produce the required output in the specified format. Asking the users about the format required by them tests the outputs generated by the system under consideration. Here the output format is considered in two ways, one is on the screen and other is the printed format. The output format on the screen is found to be correct as the format was designed in the system designed phase according to the user needs.

For the hard copy also the output comes as the specified requirements by the users. Hence output testing does not result any corrections in the system.

**User Acceptance Testing**

User acceptance of a system is the key factor of the success of any system. The system under study is tested for the user acceptance by constantly keeping in touch with the prospective system users at the time of developing and making changes wherever required.

**Test Cases**

Registration: To begin with login, user need to register by filling up basic registration details. There are multiple fields in registration page and every field has to fill by user. User cannot use character in the login id field.

Login: Login id and password are kept compulsory fields, and if the id or password doesn't match then it will show an error message.

**VALIDATION CRITERIA**

1. In each form, no field which is not null able should be left blank.

2. All numeric fields should be checked for non-numeric values. Similarly, text fields like names should not contain any numeric characters.

3. All primary keys should be automatically generated to prevent the user from entering any existing key.

4. Use of error handling for each Save, Edit, delete and other important operations.

5. Whenever the user Tabs out or Enter from a text box, the data should be validated and if it is invalid, focus should again be sent to the text box with proper message.

## 3.8    ADVANTAGES OF PROJECT

**Advantages:**

1. The system is easy to maintain.

2. It is user-friendly.

3. Users can search for a doctor's help at any point in time.

4. Very useful in case of emergency.

5. They can also look for free check-up camps.

**Limitations:**

1. The system is not fully automated, it needs data from the user for a full diagnosis.

2. If a user enters any wrong input, the system will produce inaccurate results.

# Chapter 4

# Conclusion

This was our project of System Design about "Heart Failure Prediction System" developed in Django in Python programming language. The Development of this system takes a lot of effort from us. We think this system gave a lot of satisfaction to all of us. Though every task is never said to be perfect in this development field even more improvement may be possible in this application. We learned so many things and gained a lot of knowledge about the development field. We hope this will prove fruitful to us.

# References

1. A. H. M. S. U. Marjia Sultana, "Analysis of Data Mining Techniques for Heart Disease Prediction," 2018

2. Galla Siva Sai Bindhika, Munaga Meghana, Manchuri Sathvika Reddy, Rajalakshmi, 'Heart Disease Prediction Using Machine Learning Techniques', 04 Apr 2020

3. Zaibunnisa L. H. Malik, Momin Fatema, Nikam Pooja, Gawandar Ankita, "Heart Disease Prediction using Artificial Intelligence" 2021

4. Xiao-Yan Gao,1Abdelmegeid Amin Ali, Hassan Shaban Hassan, and Eman M. Anwar, "Improving the Accuracy for Analyzing Heart Diseases Prediction Based on the Ensemble Method" , 2021

5. Chintan M. Bhatt , Parth Patel , Tarang Ghetia and Pier Luigi Mazzeo , "Effective Heart Disease Prediction Using Machine Learning Techniques" 6 February 2023

6. ] M. I. K. ,. A. I. ,. S. Musfiq Ali, "Heart Disease Prediction Using Machine Learning Algorithms".

7. K. Bhanot, "towarddatascience.com," 13 Feb 2019. [Online]. Available: https://towardsdatascience.com/predicting-presence-of-heart-diseases-using-machinelearning-36f00f3edb2c. [Accessed 2 March 2020].

8. M. A. K. S. H. K. M. a. V. P. M Marimuthu, "A Review on Heart Disease Prediction using Machine Learning and Data Analytics Approach".