# Low Level Design (LLD)

## Heart Disease Diagnostic Analysis

Revision Number: 1.6
Last date of revision: 25/11/2022

Pranit Patil

## Document Version Control

| Date Issued | Version | Description | Author |
|---|---|---|---|
| 06/09/2022 | 1.0 | Abstract, Introduction | Pranit Patil |
| 13/09/2022 | 1.1 | Problem Statement | Pranit Patil |
| 27/09/2022 | 1.2 | Dataset Information | Pranit Patil |
| 5/10/2022 | 1.3 | Architecture | Pranit Patil |
| 12/10/2022 | 1.4 | Architecture Description | Pranit Patil |
| 17/10/2022 | 1.5 | Deployment | Pranit Patil |
| 25/11/2022 | 1.6 | Final Revision | Pranit Patil |

# Contents

# Abstract

Heart disease/ cardiovascular disease (CVD) is a class of diseases that involve the heart or blood vessels. It is a term covering any disorder of the heart. Heart diseases have become a major concern to deal with as studies show that the number of deaths due to heart diseases have increased significantly over the past few decades in India it has become the leading cause of death in India. A study shows that from 1990 to 2021 the death rate due to heart disease has increased by around 39% from 155.7 to 217 deaths per 1 lakh population in India. The underlying mechanisms vary depending on the disease. This may be caused by high blood pressure, smoking, diabetes mellitus, lack of exercise, obesity, high blood cholesterol, poor diet, and excessivealcohol consumption.

Thus, preventing heart disease has become more than necessary. It is estimated that up to 90% of CVD may be preventable. Good data-driven systems for predictingheart diseases can improve the entire research and prevention process, making surethat more people can live healthy lives. Detection of a CVD at an early stage leads to preventing more than 80% of potentially related deaths.

With the anecdotes of fairly and conventionally healthy people getting heart-related illnesses, it is no longer an old age disease as previously thought and needs dedicated research for prevention and potential cure.

# 1 Introduction

## 1.1 Why this Low-Level Design Document?

The purpose of this document is to present a detailed description of the heart disease diagnostic analysis. LLD describes the class diagrams with the methods and relations between classes and programs specs. It describes the modules so that the programmer can directly code the program from the document. This document is intended for both the stakeholders and the developers of the system and will be proposed to the higher management for its approval.

The LLD will be focusing on the below objectives:

1. Problem Understanding.
2. Data Acquisition.
3. Data Pre-Processing and Exploratory Analysis.
4. Dashboard report for important activities.

## 1.2 Scope

Low-level design (LLD) is a component-level design process that follows a step by step refinement process. The process can be used for designing data structures, required software architecture, source code and ultimately, performance algorithms. Overall, the data organization may be defined during requirement analysis and then refined during data design work.

## 1.3 Project Introduction

Heart disease is one of the most common disease nowadays, and an earlydiagnosis of such a disease is a crucial task for many health care providers to prevent their patients for such a disease and to save lives. 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 data analysis techniques need to be used. The data analysis predicts the likelihood of patients getting heart disease. It enables significant knowledge. E.g. Relationships between medical factors related to heart disease and patterns, to be established. The obtained results have illustrated that the analysis can effectively predict the risk level of heart disease.

### 1.4 Problem Statement

Health is real wealth in the pandemic time we all realized the brute effects of covid-19 on all irrespective of any status. You are required to analyze this health and medical data for better future preparation. A dataset is formed by taking into consideration some of the information of 303 individuals.

## 2 Technical specifications

### 2.1 Dataset

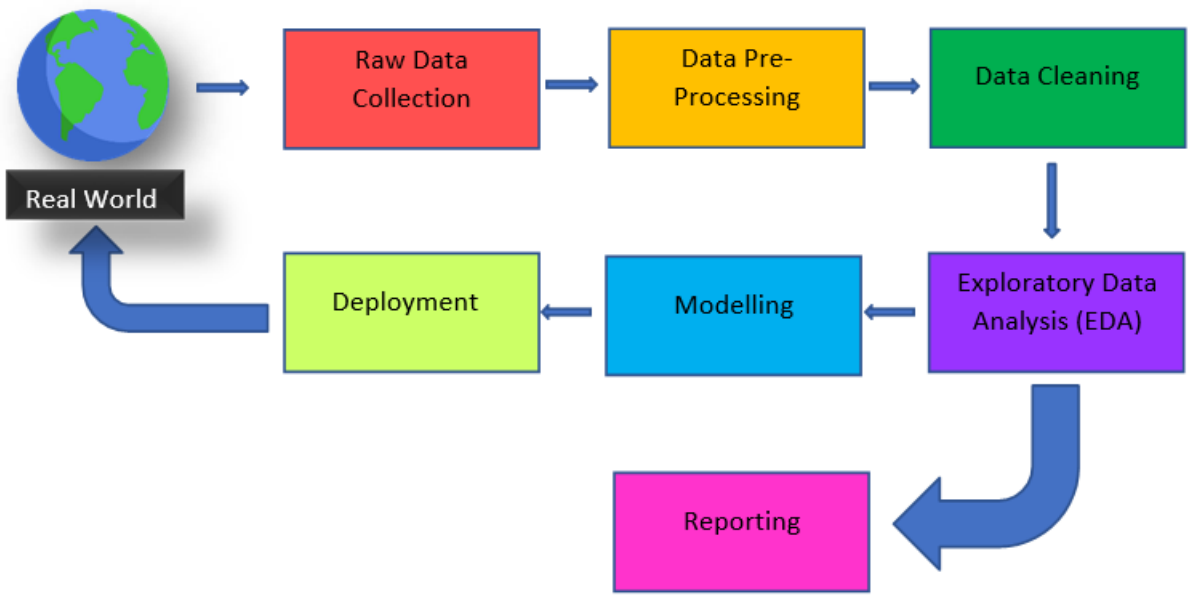
The dataset consists of 303 individuals data. There are 14 columns in the dataset, which are described below.

1. ***Age***: displays the age of the individual.
2. ***Sex***: displays the gender of the individual using the following format:
   1 = male
   0 = female
3. ***Chest-pain type***: displays the type of chest-pain experienced by the individual using the following format :
   0 = typical angina
   1 = atypical angina
   2= non - anginal pain
   3 = asymptotic
4. ***Resting Blood Pressure***: displays the resting blood pressure value of an individual in mmHg (unit).
5. ***Serum Cholesterol***: displays the serum cholesterol in mg/dl (unit).
6. ***Fasting Blood Sugar***: compares the fasting blood sugar value of an individual with 120mg/dl.
   If fasting blood sugar > 120mg/dl then: 1 (true) else : 0 (false)
7. ***Resting ECG***: displays resting electrocardiographic results
   0 = normal
   1 = having ST-T wave abnormality
   2 = left ventricular hypertrophy
8. ***Max heart rate achieved*** : displays the max heart rate achieved by an individual.
9. ***Exercise induced angina***:
   1 = yes
   0 = no

10. ***ST depression induced by exercise relative to rest***: displays the value which is an integer or float.
11. ***Peak exercise ST segment***:
    0 = upsloping
    1 = flat
    2 = downsloping
12. ***Number of major vessels (0–4) colored by flourosopy***: displays the value as integer or float.
13. ***Thal***: displays the thalassemia :
    0,1 = normal
    2 = fixed defect
    3 = reversible defect
14. ***Diagnosis of heart disease*** : Displays whether the individual is suffering from heart disease or not :
    0 = absence
    1 = present

# 3 Proposed Solution

## 3.1 Architecture



# 4 Architecture Description

## 4.1 Raw Data Collection:

The Dataset was taken from iNeuron's Provided Project Description Document.
Dataset link: -Heart-disease-diagnostic-analysis-Ineuron.ai/heart_disease_dataset.csv at main · Pp11112000/-Heart-disease-diagnostic-analysis-Ineuron.ai (github.com)
**Size of the Dataset:** 303 patients data with 14 features describing the presence/absence of disease.

## 4.2 Data Pre-Processing:

Before building any model, it is crucial to perform data pre-processing to feed the correct data to the model to learn and predict. Model performance depends on the quality of data feeded to the model to train. This Process includes:

- Handling Null/Missing Values
- Handling Skewed Data
- Outliers Detection and Removal

### 4.2.1 Data Cleaning

Data cleaning is the process of fixing or removing incorrect, corrupted,incorrectly formatted, duplicate, or incomplete data within a dataset.

- Remove duplicate or irrelevant observations
- Filter unwanted outliers
- Renaming required attributes

## 4.3 Exploratory Data Analysis

Exploratory Data Analysis refers to the critical process of performing initial investigations on data to discover patterns, spot anomalies, test hypothesis and to check assumptions with the help of summary statistics and graphical representations.

## 4.4 Reporting

Reporting is a most important and underrated skill of a data analytics field. Because being a Data Analyst you should be good in easy and self explanatory report because your model will be used by many stakeholders who are not from technical background.

a) High Level Design Document (HLD)
b) Low Level Design Document (LLD)
c) Architecture

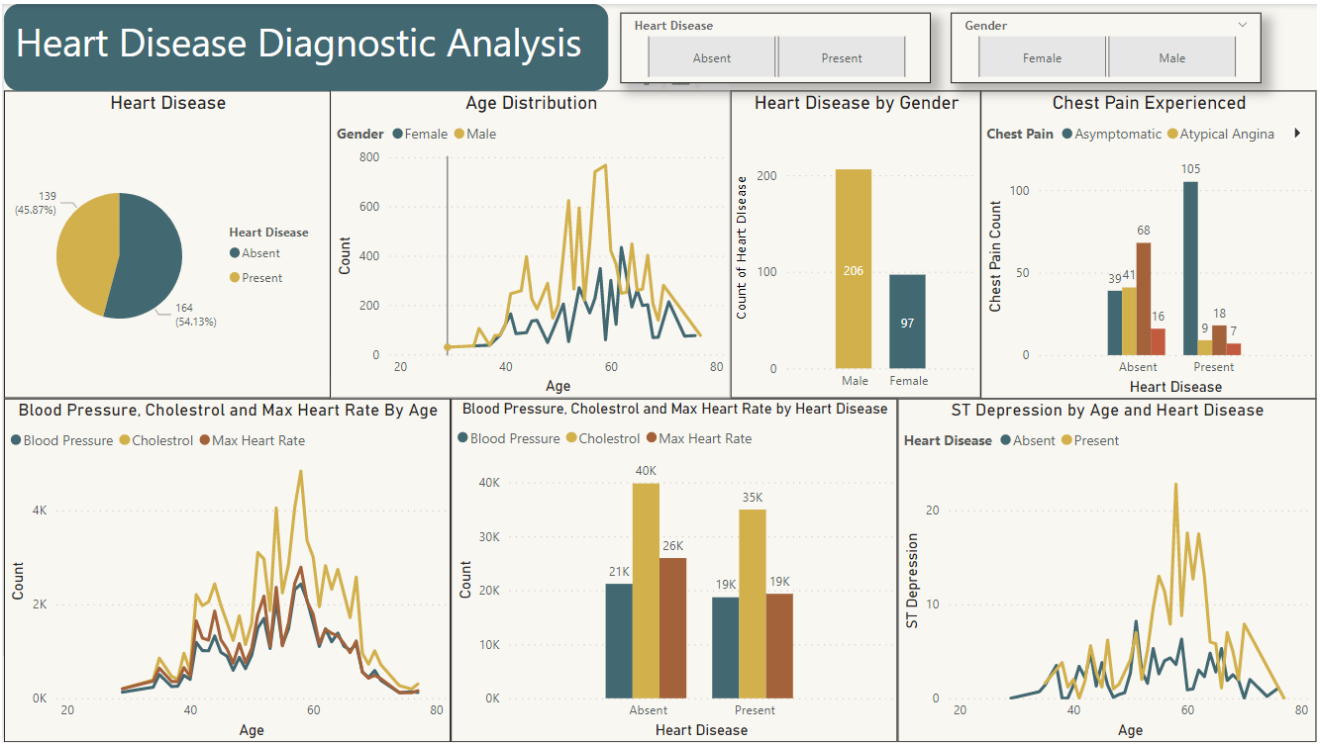
d) Wireframe
e) Detailed Project Report
f) Power Point Presentation

## 4.5 Modelling

Data Modelling is the process of analyzing the data objects and their relationship to the other objects. It is used to analyze the data requirements that are required for the business processes. The data models are created for the data to be stored in a database. The Data Model's main focus is on what data is needed and how we have to organize data rather than what operations we have to perform.

## 4.6 Deployment

We created a Power BI Dashboard and published it on Power BI Service.



# 5 Key performance indicators (KPI)

Key indicators display a summary of the heart disease analysis and its relationship with different metrics.

1. Percentage of People Having Heart Disease.
2. Age Distribution including Gender.
3. Gender Distribution Based on Heart Disease.
4. Chest Pain Experienced by People Suffering from Heart Disease.
5. Blood Pressure, Cholesterol Level and Maximum Heart Rate of People According totheir Age and Heart Disease Patients.
6. ST Depression Experienced by People According to their age and heart disease.