

Low Level Design

Heart Disease Diagnostic-Analysis

Revision Version: 1.0

Last Date of Revision:23/10/2022

Priyanka Tanpure

Document Version Control

Date Issued	Version	Description	Author
1 Aug 2022	1.0	Introduction	Priyanka Tanpure
7 Aug 2022	1.1	Problem Statement	Priyanka Tanpure
16 Aug 2022	1.2	Dataset Information	Priyanka Tanpure
25 Aug 2022	1.3	Architecture	Priyanka Tanpure
7 Sep 2022	1.4	Architecture Description	Priyanka Tanpure
23 Oct 2022	1.5	Deployment	Priyanka Tanpure

Contents

Document Version Control.....	2
Abstract.....	4
1. Introduction	5
1.1. Why this Low-Level Design Document?.....	5
1.2. Scope	5
1.3. Project Introduction.....	5
1.4. Problem Statement.....	5
2. Architecture.....	6
3. Architecture Description.....	7
3.1. Data Description.....	7
3.2. Data Pre-Processing.....	8
3.3. Data Cleaning.....	6
3.4. Exploratory Data Analysis (EDA).....	6
3.5. Reporting.....	6
3.6. Modelling.....	6
3.7. Deployment.....	6

1 Introduction

1.1 Why this Low-Level Design Document?

The goal of the LDD or Low-level design document (LLDD) is to give the internal logic design of the actual program code for the Heart Disease Diagnostic-Analysis dashboard. LLDD 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.

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

Health is real wealth in the pandemic time, and we all realized the brute effects of covid-19 on all irrespective of any status. You are required to analyse this health and medical data for better future preparation.

Heart Disease is a general term that includes many types of heart problems, and it has become major concern in India to deal with as there is around 34.3% increase seen in number of deaths due to cardiovascular problems from 2.26 million in 1990 to 4.77 million in 2020.

Thus, it has become necessary to take precautionary measures to prevent and reduce the heart diseases among people. Therefore, to make sure that the people live healthy lives, the needed research and the prevention process of heart disease can be improved by using good data driven systems for understanding the patterns.

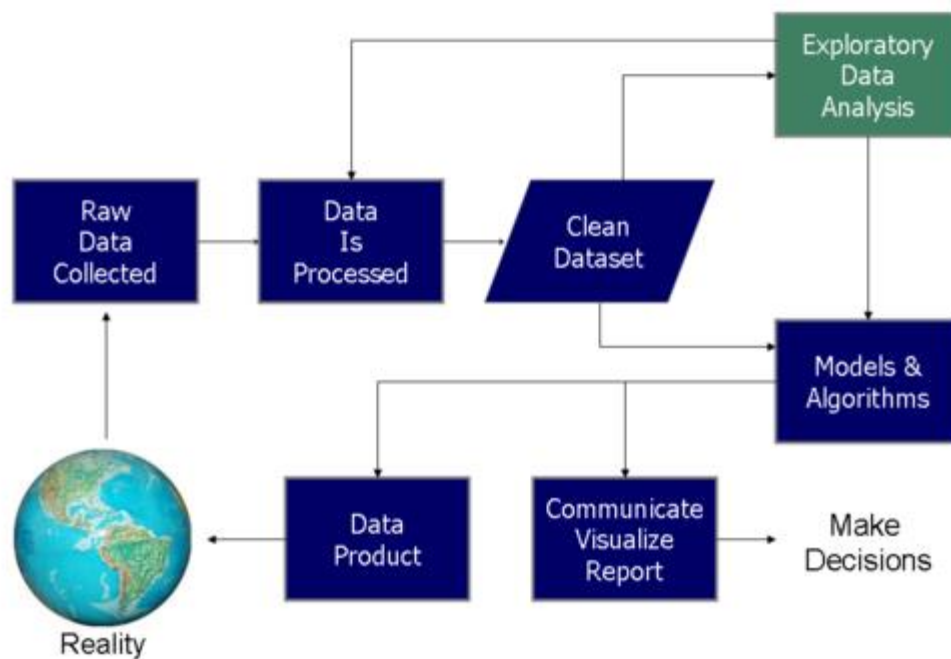
1.4 Problem Statement

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 do the research, analyze this health and medical data for better future preparation.

Business Objectives:

- Performing EDA through Python and find insights.
- Extract various information such as heart disease rates, Heart disease by - gender, by age.
- Compare attributes of the dataset to extract necessary information.
- Make Dashboard with extracted information from data.
- Find key metrics and factors and show relationship between attributes.

2 Architecture



3 Architecture Description

3.1. Data Description

The Dataset was taken from iNeuron's provided project description document.

Database list: heart_disease_dataset.csv

Link: <https://drive.google.com/drive/folders/165Pjmf9W9PGy0rZjHEA22LW0Lt3Y-Q8?usp=sharing>

Only 14 Attributes are used:

1. **age:** in years
2. **sex:** 2 Values (1 = male, 0 = female)
3. **chest pain type:** 4 values (Value 1: typical angina, Value 2: atypical angina, Value 3: non-anginal pain, Value 4: asymptomatic)
4. **trestbps:** resting blood pressure (in mm Hg on admission to the hospital)
5. **chol:** serum cholesterol in mg/dl
6. **fbs:** (fasting blood sugar > 120 mg/dl), (1 = true, 0 = false)
7. **restecg:** resting electrocardiographic results: (Value 0: normal, Value 1: having ST-T wave abnormality (T wave inversions and/or ST elevation or depression of > 0.05 mV, Value 2: showing probable or definite left ventricular hypertrophy by Estes' criteria)
8. **thalach:** maximum heart rate achieved
9. **exang:** exercise include angina (1 = yes, 0 = no)
10. **oldpeak:** ST depression induced by exercise relative to rest
11. **slope:** the slope of the peak exercise ST segment (Value 1: upsloping, Value 2: flat, Value 3: downsloping)
12. **ca:** number of major vessels (0-3) colored by flourosopy

13. **thal:** 3 = normal; 6 = fixed defect; 7 = reversable defect
14. **num:** the predicted attribute diagnosis of heart disease (angiographic disease status) (*Value 0: < 50% diameter narrowing, Value 1: > 50% diameter narrowing i.e. 0 = no (no presence of heart disease), 1 = yes (presence of heart disease)*)

3.2 Data Pre-Processing

Before finding insights from the given data, it is crucial to perform data pre-processing to feed the correct data to the model to learn and predict. Performance of the model depends on the quality of data used to train the model.

This Process includes:

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

3.3 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

3.4 Exploratory Data Analysis (EDA)

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.

3.5. Reporting

Reporting is a most important and underrated skill of a data analytics field. Because of 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.

- High-Level Design Document (HLD)
- Low-Level Design Document (LLD)
- Architecture
- Wireframe
- Detailed Project Report
- PowerPoint Presentation

3.6. 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.

3.7. Deployment

Created a Tableau dashboard.

