Low Level Design (LLD)
Heart Disease DiagnosticAnalysis

AAKASH BHUTE

Why this Low-Level Design Document?

This document aims to provide a comprehensive overview of the heart disease prediction analysis technique. It will outline the necessary steps that must be followed prior to conducting any analysis. Additionally, the document will elucidate the algorithms and techniques employed to predict the presence or absence of heart disease, accompanied by a comparative analysis of the results. It will also encompass low-level design (LLD) by illustrating class diagrams, depicting the methods and relationships between classes, and specifying program specifications. The document will describe the modules in a manner that enables programmers to directly code the program based on the information provided. Its intended audience comprises both stakeholders and developers of the system, and its ultimate goal is to secure approval from higher management.

The LLD will be focusing on the below objectives:

- Problem Understanding.
- Data Acquisition.
- Data Pre-Processing and Exploratory Analysis
- Development of models
- Auditing accuracy and retrain if require
- Finalizing the model
- Dashboard report for important activities

Scope

The low-level design (LLD) documentation provides a comprehensive representation of the internal structure of the heart disease prediction system, outlining the individual components and their interactions. The purpose of LLD is to offer a detailed logical design of the actual program code. This design is derived from the high-level design. The LLD documentation includes a thorough description of the employed model, including comparisons between the proposed model or library and an existing baseline model, based on a set of predefined metrics.

Project Introduction

In today's world, heart disease (HD) is a prevalent ailment, and early detection plays a critical role in the efforts of healthcare providers to safeguard their patients and save lives..

The healthcare industry gathers vast amounts of data that contain valuable hidden insights, which can aid in making informed decisions. To derive meaningful results and facilitate effective decision-making, the application of data science techniques becomes essential. Data analysis techniques can predict the probability of individuals developing heart disease, thereby facilitating the establishment of significant knowledge, such as identifying relationships between medical factors associated with heart disease and recognizing patterns. The results obtained from such analyses have demonstrated the effectiveness of the designed diagnostic system in accurately predicting the risk level of heart diseases.

Constraints

The analysis conducted utilizes a restricted dataset focusing on 14 specific features known to impact heart disease. It's important to note that external interventions, such as underlying diseases, medication types, lifestyle patterns, and BMI values, are not considered in this analysis. These factors have the potential to influence the occurrence and progression of heart disease. While the analysis provides insights based on the available dataset, it's crucial to acknowledge that a more comprehensive approach incorporating these external interventions would be necessary for a more accurate assessment and prediction of heart disease.

Risks

Document specific risks that have been identified or that should be considered.

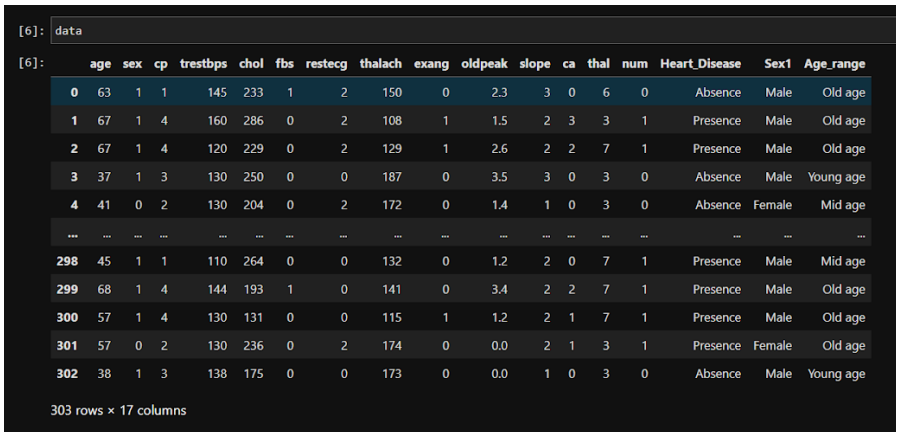
Out of Scope

Delineate specific activities, capabilities, and items that are out of scope for the project.

1.Technical specifications

Dataset

The Dataset was taken from iNeuron's provided Project Description Document.



[6]: data

[6]:

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	num	Heart_Disease	Sex1	Age_range
0	63	1	1	145	233	1	2	150	0	2.3	3	0	6	0	Absence	Male	Old age
1	67	1	4	160	286	0	2	108	1	1.5	2	3	3	1	Presence	Male	Old age
2	67	1	4	120	229	0	2	129	1	2.6	2	2	7	1	Presence	Male	Old age
3	37	1	3	130	250	0	0	187	0	3.5	3	0	3	0	Absence	Male	Young age
4	41	0	2	130	204	0	2	172	0	1.4	1	0	3	0	Absence	Female	Mid age
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
298	45	1	1	110	264	0	0	132	0	1.2	2	0	7	1	Presence	Male	Mid age
299	68	1	4	144	193	1	0	141	0	3.4	2	2	7	1	Presence	Male	Old age
300	57	1	4	130	131	0	0	115	1	1.2	2	1	7	1	Presence	Male	Old age
301	57	0	2	130	236	0	2	174	0	0.0	2	1	3	1	Presence	Female	Old age
302	38	1	3	138	175	0	0	173	0	0.0	1	0	3	0	Absence	Male	Young age

303 rows × 17 columns

Figure 1: Heart Disease Prediction dataset

The dataset consists of 303 individuals. 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 Cholestrol***: displays the serum cholesterol in mg/dl (unit)
6. ***Fasting Blood Sugar***: compares the fasting blood sugar value of an individual with 120 mg/dl.
   If fasting blood sugar > 120 mg/dl then : 1 (true) else
   : 0 (false)

7. ***Resting ECG*** : displays resting electrocardiographic results
   0 = norma1 = 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) coloured by fluoroscopy*** : 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

2. Problem Statement

In the midst of the pandemic, we have collectively recognized the significant impact of COVID-19 on people from all walks of life, regardless of their social or economic status. In order to enhance future preparedness, it is essential to analyze health and medical data related to this crisis. A dataset has been compiled, comprising information from 303 individuals, with the aim of gaining valuable insights and informing strategies for better preparation in the future.

3. Architecture Description

1. Raw Data Collection

The Dataset was taken from iNeuron's provided Project Description Document.
https://drive.google.com/drive/folders/165Pjmfb9W9PGy0rZjHEA22LW0Lt3Y-Q8

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

a) Handling Null/Missing Values
b) Handling Skewed Data
c) Outliers Detection and Removal

3. Data Cleaning

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

a) Remove duplicate or irrelevant observations
b) Filter unwanted outliers
c) Renaming required attributes

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.

5. 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 selfexplanatory report because your model will be used by many stakeholders who are not from a technical background.

a) High Level Design Document (HLD)
b) Low Level Design Document (LLD)
c) Architecture
d) Wireframe
e) Detailed Project Report
f) Powerpoint Presentation

6. Modeling

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.

7. Deployment

Power BI Dashboard