HIGH LEVEL DESIGN DOCUMENT FOR HEART DISEASE DIAGNOSTIC ANALYSIS

DOCUMENT VERSION CONTROL

Date Issue	Version	Description	Author
22 – Feb - 2023	1	Initial HLD – V 1.0	SINGAM RAHUL

CONTENT

- SCOPE
- GENERAL DESCRIPTION ● PROBLEM STATEMENT
- PRODUCT PERSPECTIVE
- PROPOSED SOLUTION
- TECHNICAL REQUIREMENTS
- TOOLS USED
- POWER BI INTRODUCTION
- DESIGN DETAILS
- OPTIMIZATION
- ARCHITECTURE DESCRIPTION
- POWER BI DASHBOARD INTRODUCTION
- HEART DISEASE DIAGNOSTIC VISUALIZATION DASHBOARD

ABSTRACT

Heart disease is the most major health issue that is suffered by many people all over the globe, some of the causes of heart diseases due to hypertension, diabetes, overweight, and an unhealthy lifestyle. This project of Healthcare Analysis on Heart Disease Data is aimed to explore the Heart Disease dataset .The objective is to analyze the various features and their relationship with each other and find out their contribution towards getting a heart disease. Various features such as Age, Sex, Chest pain type, Blood pressure, Cholesterol,Fasting Blood sugar, Rest ECG, Thalach, Exercise enduced Angina,Major vessels,oldpeak,slope,thal are present in the dataset.The goal of the project is to find all types of relationships between the features and come out with significant contributors to a heart disease.

INTRODUCTION

WHY THIS HIGH LEVEL DESIGN DOCUMENT ?

The purpose of this High Level Design (HLD) Document is to add the necessary detail to the current project description to represent a suitable model for coding. This document is also intended to help detect contradictions prior to coding, and can be used as a reference manual for how the modules interact at a high level.

THE HLD WILL -

- PRESENT ALL THE DESIGN ASPECTS AND DEFINE THEM IN DETAIL
- DESCRIBE THE USER INTERFACE BEING IMPLEMENTED
- DESCRIBE THE HARDWARE AND SOFTWARE INTERFACE
- DESCRIBE THE PERFORMANCE REQUIREMENT
- INCLUDE DEFINE FEATURE AND ARCHITECTURE OF THE PROJECT
- LIST AND DESCRIBE THE NON FUNCTIONAL ATTRIBUTES

    SECURITY

    RELIABILITY

    MAINTAINABILITY

    PORTABILITY

    REUSABILITY

    APPLICATION COMPATIBILITY

    RESOURCE UTILIZATION

    SERVICEABILITY

SCOPE

The HLD documentation presents the structure of the system, such as the database architecture, application architecture (layers), application flow (Navigation), and

technology architecture. The HLD uses non-technical to mildly-technical terms which should be understandable to the administrators of the system.

GENERAL DESCRIPTION

PROBLEMSTATEMENT

The Healthcare domain is one of the leading domains. Using Data analysis tools and techniques to build a solution for a healthcare domain problem can be counted as a significant contribution to this domain.

The main objective of the project includes –

- Analyzing raw dataset.
- Perform data cleaning and preprocessing operations on the raw data.
- Building visualizations on the cleaned features using a BI tool.
- Building Dashboard using Power BI

TECHNICAL REQUIREMENTS

The solution can be a cloud-based or application hosted on an internal server or even be hosted on a local machine. For accessing this application below are the minimum requirements:

- Good internet connection. ● Web Browser.

For training model, the system requirements are as follows:

- +4 GB RAM preferred
- Operation System: windows, Linux, Mac
  Jupyter Notebook / Google Colab
- S3 buckets for cloud storage for the collected data.

TOOLS USED

Python Libraries such as NumPy, Pandas, Matplotlib, Seaborn, and MS-Excel and Business Intelligence tools such as Colab / Jupyter Notebook and Python Programming Language are used to build the whole framework.

seaborn is a suite of business analytics tools developed by python libraries that allows users to analyze data and share insights. It includes a collection of software services, apps, and connectors that work together to turn data into interactive visualizations and business intelligence reports.

With seaborn , users can connect to a wide variety of data sources, including Excel spreadsheets, cloud-based and on-premises data sources, and even streaming data. Once the data is connected, users can create custom reports, dashboards, and visualizations that can be easily shared with others in their organization.

seaborn also includes features such as natural language query, which allows users to ask questions of their data using plain English, and machine learning capabilities, which can help users to identify trends and patterns in their data.

DESIGN DETAILS

DESIGN DETAILS

Layout

Layout: A well-designed dashboard should have a clear and consistent layout that is easy to navigate. Use a grid-based layout and align elements with a consistent margin to make your dashboard look organized and professional. Use headings, text boxes, and images to group and separate different sections of your dashboard.

Color

Color: Use color wisely to draw attention to important information and make your dashboard visually appealing. Use a limited color palette and choose colors that are easy to read and complementary. Avoid using too many bright or contrasting colors, which can be distracting or hard on the eyes.

Typography

Typography: Use typography to create hierarchy and improve readability. Choose a font that is easy to read and consistent throughout your dashboard. Use font size, weight, and color to differentiate between headings, subheadings, and body text.

Visualization

Visualization: Choose visualizations that effectively communicate your data and insights. Use charts, graphs, tables, and maps to present your data in a clear and compelling way.

Use appropriate chart types based on the data you are presenting and avoid using too many different types of visualizations on the same dashboard.

Interactivity

Interactivity – Use interactivity to make your dashboard more engaging and informative. Use filters, slicers, and drill-down capabilities to allow users to explore the data in more detail. Use tooltips and annotations to provide additional context and information.

Performance

Performance – Ensure that your dashboard is fast and responsive. Use optimized data models, avoid complex calculations or queries, and limit the amount of data you load into your dashboard. Use caching and pre-aggregation to improve performance, especially for large datasets.

Accessibility

Accessibility – Make your dashboard accessible to all users, including those with disabilities. Use alt text for images, provide captions for videos, and use a high contrast mode for text and visuals. Use data tables for tabular data and ensure that your dashboard is keyboard-friendly.

OPTIMIZATION

Data Strategy and Performance

- Removing duplicate records from the dataset.
- Handling null values.
- Performing feature encoding on the dataset.

ARCHITECTURE DESCRIPTION

Heart disease is a major health concern globally, and understanding the factors that contribute to this condition is critical for developing effective prevention and treatment strategies. In this project, we aim to analyze the relationships between different features

and identify the significant contributors to heart disease. Our end goal is to create a dashboard that can be used by healthcare professionals to gain insights into the factors that impact heart disease risk.

To achieve this goal, we will begin by collecting and pre-processing the dataset that will be used in our analysis. We will use Python data manipulation libraries like Pandas to clean and optimize the data. We will also perform table calculations and data filtering to obtain more granular results.

Next, we will import the cleaned dataset into a Business Intelligence (BI) tool such as Tableau. Using this tool, we will create visualizations and dashboards that will help us to explore the relationships between different attributes and identify key findings. We will also use filters to speed up the process of finding insights and create more granular reports.

In addition to creating visualizations and dashboards, we will also perform feature engineering to extract more useful information from the dataset. This can help us to identify the most important predictors of heart disease and improve the accuracy of our analysis. We will also build predictive models and validate them to ensure that they are accurate and reliable.

Finally, we will create a detailed project report that includes all the visual plots and key findings from the analysis. We will also host the dashboard on a cloud platform like Tableau Public, making it easily accessible to healthcare professionals around the world.

By following this comprehensive approach, we can gain valuable insights into the factors that contribute to heart disease and develop effective prevention and treatment strategies. We hope that our analysis will help healthcare professionals to identify patients who are at risk for heart disease and provide them with the care they need to stay healthy.