High-Level Design

Thyroid Disease Detection

Developed by	Saurabh Gupta
Version	1.0
Date	24-06-2023

Table of Contents

1   Document Change/History Control

Version	Date	Done By	Review By	Remarks
1.0	24-06-2023	Saurabh Gupta		

2   Reviews:

Version	Date	Reviewer	Remarks

3   Approval Status:

Version	Review Date	Reviewed By	Approved By	Remarks

4 Abstract

Machine Learning is a category of algorithms that allows software applications to become more accurate in predicting outcomes without being explicitly programmed. The basic premise of machine learning is to build models and employ algorithms that can receive input data and use statistical analysis to predict an output while updating outputs as new data becomes available. These models can be applied in different areas and trained to match the expectations of management so that accurate steps can be taken to achieve the organization's target. In this paper, Thyroid disease is a common cause of medical diagnosis and prediction, with an onset that is difficult to forecast in medical research. The thyroid gland is one of our body's most vital organs. Thyroid hormone releases are responsible for metabolic regulation. Hyperthyroidism and hypothyroidism are one of the two common diseases of the thyroid that releases thyroid hormones in regulating the rate of body's metabolism. The main goal is to predict the estimated risk on a patient's chance of obtaining thyroid disease or not.

5 Introduction

5.1   What are High-Level Design Documents?

The purpose of this High-Level Design (HLD) Documents is to add necessary details to the current project description to represent a suitable for coding. This document is also intended to help detect contradictions before coding. And can be used as a reference manual for how the modules interact at a high level.

The HLD will be:

- Present all of the design aspects and define them in detail.
- Describe the user interface being implemented.
- Describe the needed Python libraries for the coding.
- Describe the performance requirements.
- Include design features and the architecture of the project.
- List and describe the non-functional attributes like:
    - Security
    - Reliability
    - Maintainability
    - Portability
    - Reusability

- o Application Compatibility
- o Resource Utilization
- o Serviceability

5.2 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 and mildly-technical terms which should be understandable to the administrators of the system

5.3 Definition

TERM	DESCRIPTION
DB	Database, the cloud platform where the data will be stored. Can be considered cloud storage.
ML	Machine Learning
API OR APIS	Application Programming Interface can be considered a website link from there we can extract information.

6 General Description

6.1 Product Perspective

The Thyroid Detection Prediction is an ML-based Web Application that Is able to predict the estimated risk on a patient chance of obtaining thyroid disease or not. It will give the probability in yes/no.

6.2 Problem Statement

To build a system that will be able to take information about a patient test results and can predict or provide the probability of having thyroid disease. We have to build an application and that will be able to produce results in yes or no.

6.3 Proposed Solution

We will use EDA (Exploratory data analysis) to find the important relation between different attributes and will use a machine-learning algorithm to predict the thyroid disease. The patient will be filled the required feature as input get it from the test results and will get results through the web application. The system will get features and it will be passed into the backend where the features will be validated and preprocessed and then it will be passed to a hyperparameter tuned machine learning model to predict the final outcome.

6.4 Data Requirements

The data is required for the building of the project is already available on the dashboard.
Dataset Characteristics: Multivariate, Domain-Theory
Subject Area: Life

6.5 Tool Used

The programming language is Python that is used here, also we will use some other python-based libraries like, for ml, we will use Scikit-Learn library, for data manipulation we will use pandas, for numerical computation NumPy, for custom APIs creation Flask web frameworks. Visual Studio Code is used as python IDE for all modular coding and custom APIs creation. And storing all code files for publicly available we will use GitHub.
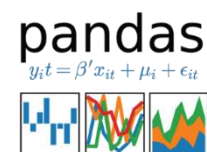
6.6 Constraints

The System should be user-friendly, the user should get all proper messages while using the web app. He/she also should get a proper error message if he/she has done something wrong on the web-app page. All the errors and results should be delivered in the easiest possible way and all the buttons are going to insert on the webpage should be labeled properly, so the user did not get confused to use the system.
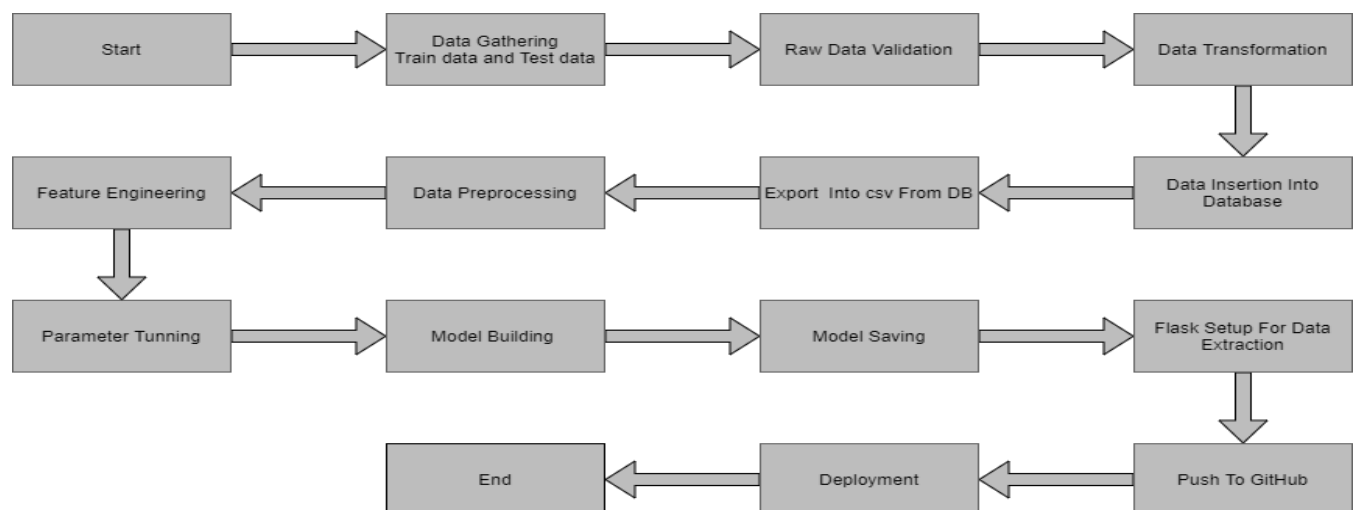
6.7 Assumptions

The main objective is to implement a system that will predict or provide the probability of having thyroid disease.
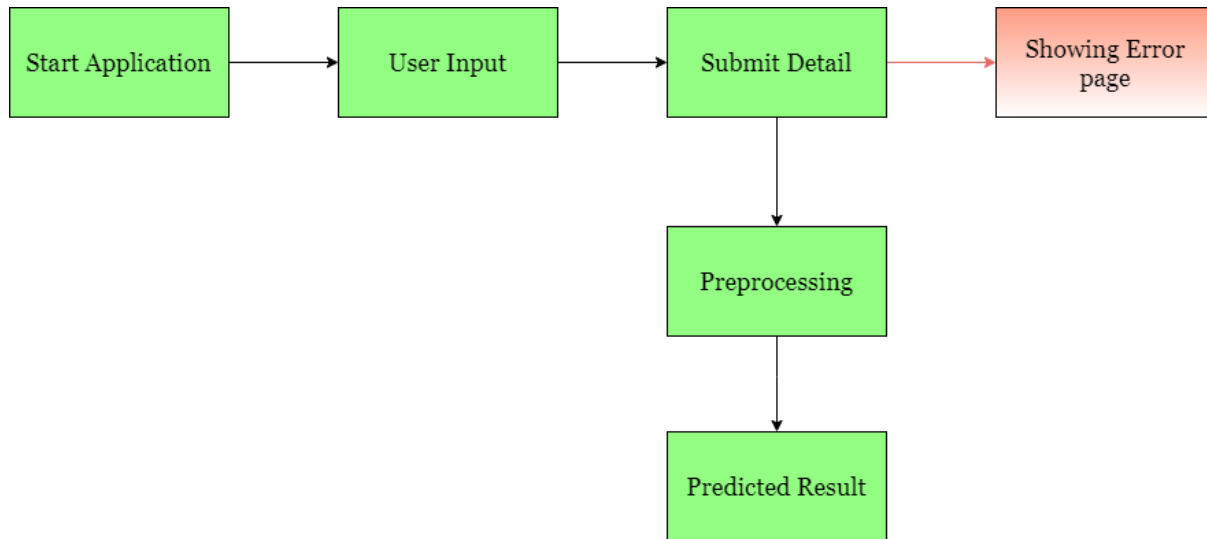
7 Design Details

7.1 Process Flow

We will be using following process flow for this project. The process will be based on modular coding i.e. use of oops concepts to build the entire project from start to end.

7.2 Deployment Process

7.3 Error Handling

If any error occurred in the processing way then the error message should be shown to the user in a completely non-technical way that can be understandable by any person. And Meaningful error message should be shown, so the user can spot his mistake and rerun the process with improvement. All the errors that are will occur should be handled properly. And we have to log every error for our application and have to manage the same.

8 Performance

The Thyroid Detection is dependent on machine-learning algorithms. We will train various ml algorithms and will find the best fitting algorithm for predicting the target. Our system performance will be based on the data we are going to feed to the algorithms. And the performance will depend on the finalized model. and the web application and the deployment server. With all of these components, our program should run properly.

8.1 Reusability

The code and the module are created during the time of building the project should maintain all coding guidelines and full project code is written in a Modular fashion. Our system should have the flexibility to work properly from any location. And it should handle any improper input

value from the user and should give a meaningful error message so the user can correct his/her mistake and enter valid input to get the result. And the system should be reusable in every manner with different types of inputs values that are all are it has been trained.

8.2 Application Compatibility

The different libraries and python programming languages are used to build the system. Every library has its own functionality and it should work properly with us fluctuate system. Flask will be used for making the web APIs and HTML/CSS will be used to make the web application. All the components of the application should work properly and it should produce a result without any interpretation.

8.3 Resource Utilization

Our application should utilize the given resource properly and it should use a minimal amount of internet to work and call the APIs on the Web page. Our system should not use much amount of computational resources hence it will make the application slow. Our application will be deployed cloud platform and it should utilize the resource given on the cloud and work properly.

9 Deployment

For the deployment process, we will be using AWS Elastic Bean Stack cloud platforms for hosting our application. The cloud platform will run the system and it will give the flexibility to use our application globally

10 Conclusion

Thyroid disease is a common cause of medical diagnosis and prediction, with an onset that is difficult to forecast in medical research. The thyroid gland is one of our body's most vital organs. Thyroid hormone releases are responsible for metabolic regulation. Hyperthyroidism and hypothyroidism are one of the two common diseases of the thyroid that releases thyroid hormones in regulating the rate of body's metabolism. The prediction of the estimated risk of thyroid disease will help the patient.

11 Reference

Google image for collection the logos and images.

Sketch diagram for drawing the diagrams.