High Level Design(HLD)

Thyroid Disease Detection

By-Arkoprovo Ghosh

Document Version Control

Date Issued	Version	Description	Author
29-07-2023	1	Initial HLD-V1.0	Arkoprovo Ghosh

Contents

Document Version Control 2

Abstract 4

1 Introduction 5

1.1 Why this High-Level Design Document 5

2. Scope 5
3. Definitions 5

2 General Description 6

1. Product Perspective 6
2. Problem Statement 6
3. Proposed Solution 6
4. Further Improvements 6
5. Data Requirements 7
6. Tools Used 8
7. Constraints 9
8. Assumptions 9

3 Design Details 9

1. Process Flow 9
   3.1.1 Model Training and Evaluation 10
   3.1.2 Deployment Process 10
2. Event log 11
3. Error Handling 11

2 Performance 12

1. Reusability 12
2. Application Compatibility 12
3. Resource Utilization 12
4. Deployment 12

3 Conclusion 13

4 References 13

Abstract

Thyroid disease is a prevalent health concern in India, affecting approximately one in ten individuals. It predominantly affects women between the ages of 17 and 54. At its extreme stage, thyroid disease can lead to various complications, including cardiovascular issues, high blood pressure, elevated cholesterol levels, depression, and reduced fertility. The thyroid gland produces two essential hormones, total serum thyroxine (T4) and total serum triiodothyronine (T3), which play a crucial role in regulating the body's metabolism, energy levels, protein synthesis, and body temperature.

Thyroid diseases are classified into three categories based on the levels of thyroid hormones: euthyroidism, hyperthyroidism, and hypothyroidism, representing normal, excessive, or deficient hormone levels, respectively. Euthyroidism signifies a normal production of thyroid hormones and normal levels at the cellular level. Hyperthyroidism, on the other hand, results from an excess of thyroid hormones circulating in the body. Hypothyroidism is mostly caused by inadequate thyroid hormone production or poor alternative therapy.

Effective diagnosis and timely treatment are paramount concerns for healthcare practitioners. Advanced diagnostic methods, along with symptom-based reports, are now available to aid in the accurate identification of thyroid disorders. Implementing machine learning techniques on healthcare data enables generating comprehensive medical reports, providing answers to questions like the causes of thyroid problems, the age groups affected, and suitable treatment options. By processing healthcare data with appropriate methodologies, this approach facilitates more efficient and accurate disease diagnosis, treatment, and decision-making, thereby reducing the risk of fatalities.

1 Introduction

1. Why this High-Level Design Document?

The High-Level Design (HLD) Document is created to provide the necessary details for coding, identify potential issues early on, and offer a clear overview of the system's architecture and module interactions. It serves as a reference manual for developers and ensures a solid foundation for project success.

The HLD will

- Present all of the design aspects and define them in detail
- Describe the user interface being implemented
- Describe the hardware and software interfaces
- Describe the performance requirements
- Include design feature and the architecture of the project
- List and describe the non-functional attribute like:
    - Security
    - Reliability
    - Maintainability
    - Portability
    - Reusability
    - Application compatibility
    - Resource utilization
    - Serviceability

2. Scope

The High-Level Design (HLD) document provides an overview of the system's organization, including details about the database architecture, application architecture (layers), application flow (navigation), and technology architecture. The HLD is written in a way that is accessible to system administrators, using language that is mostly non-technical but may include some mildly technical terms. The goal is to ensure that administrators can comprehend and grasp the system's structure without getting overwhelmed by technical jargon

.

2 General Description

1. Product Perspective

The Thyroid Disease Detection solution system is a data science- based machine learning model which help us to detect the thyroid disease in people and take necessary action.

2. Problem Statement

To create an AI solution for detecting thyroid disease and to implement the following use cases.

- To detect thyroid disease and its type in healthy person.
- To detect thyroid disease and its type in unhealthy person.

Here unhealthy person means person already affected by thyroid disease.

3. Proposed Solution

The proposed solution is a data science model built on machine learning that can be applied to address the mentioned use cases. In the first use case, we will input data from a healthy individual who does not have thyroid disease to assess the model's ability to correctly identify such cases. For the second use case, we will input data from an unhealthy individual already diagnosed with thyroid disease to evaluate the model's performance and ensure it can accurately detect the condition.

4. Further Improvements

The Thyroid Disease Detection solution can be expanded to include additional use cases within the healthcare domain. By integrating TDD with other healthcare solutions, it can provide an extra layer of confirmation for individuals who exhibit mild symptoms of thyroid disease. This synchronization enhances the overall health assessment process, offering greater confidence and accuracy in identifying potential thyroid issues in individuals with subtle symptoms.

2.5 Data Requirements

Data requirement completely depend on our problem statement.

we need data of people who have already gone with thyroid blood test to know whether they are suffering from thyroid disease or not. If yes then what kind of thyroid disease they are suffering from. We will be required these many attributes, in which some will be personal details and some will be attributes from blood test.

- Age: Because thyroid depend on age, older than 60, especially in women.
- Gender: A woman is about five to eight times more likely to be diagnosed with a thyroid condition than a man.
- People already on thyroxin treatment or not
- People already on anti thyroid medication or not
- Pregnancy if gender is female: Postpartum thyroiditis is a condition occurs in 5% to 9% of women after childbirth.
- Whether person is sick at the time of diagnosis.
- Iodine test: Excess and low amount both can cause thyroid disease.
- Lithium test: Lithium is concentrated by the thyroid and inhibits thyroidal iodine uptake
- Goitre test: A goitre can sometimes occur when your thyroid gland produces too much thyroid hormone (hyperthyroidism).
- Tumour test: Thyroid cancer occurs when cells in your thyroid undergo genetic changes (mutations). The mutations allow the cells to grow and multiply rapidly. The cells also lose the ability to die, as normal cells would. The accumulating abnormal thyroid cells form a tumour.
- TSH level measure: It supervise thyroid gland, TSH released by pituitary gland. Normal TSH range for an adult:
- 0.40 - 4.50 mIU/mL (milli-international units per litre of blood).
- T3 level measure: Hormone released by thyroid, should be in normal range.
- T4 level measure: Low T4 is seen with hypothyroidism, whereas high T4 levels may indicate hyperthyroidism. Normal T4 range for an adult: 5.0 – 11.0 ug/dL (micrograms per decilitre of blood).
- FTI(Free T4 or Free Thyroxine: The free T4 index (FTI) is a blood test used to diagnose thyroid disorders. The FTI is obtained by multiplying the (Total T4) times (T3 Uptake) to obtain an index. Normal FT3 range: 2.3 - 4.1 pg/mL (picograms per millilitre of blood).

2.5 Data Requirements

The data required for the Thyroid Disease Detection solution is crucial to address the problem statement effectively. We need information from individuals who have previously undergone thyroid blood tests to determine whether they are afflicted with thyroid disease and, if so, what specific type of thyroid condition they have. The following attributes are essential for our data collection:

1. Age: Age is an important factor as thyroid conditions may be more prevalent in individuals older than 60, particularly in women.

2. Gender: Gender is significant, as women are statistically more prone to being diagnosed with thyroid disorders compared to men.

3. Current Thyroxine Treatment: We need to know if individuals are already undergoing thyroxine treatment for their thyroid condition.

4. Current Anti-Thyroid Medication: The information on whether individuals are currently taking anti-thyroid medications is essential.

5. Pregnancy (for females): Pregnancy status is crucial, as postpartum thyroiditis can occur in a percentage of women after childbirth.

6. Health Condition during Diagnosis: We need to record whether individuals were sick or unwell at the time of diagnosis.

7. Iodine Test: Both excess and insufficient iodine levels can lead to thyroid disorders.

8. Lithium Test: Since lithium affects thyroidal iodine uptake, this test is significant.

9. Goitre Test: The presence of goitre indicates a potential issue with thyroid hormone production (hyperthyroidism).

10. Tumour Test: Identifying thyroid cancer requires analyzing genetic changes in thyroid cells.

11. TSH Level Measurement: TSH level monitoring helps assess thyroid gland function. The normal TSH range for adults is 0.40 - 4.50 mIU/mL.

12. T3 Level Measurement: T3 is a thyroid hormone that should be within the normal range.

13. T4 Level Measurement: T4 levels are relevant for diagnosing hypothyroidism (low T4) or hyperthyroidism (high T4). The normal T4 range for adults is 5.0 – 11.0 ug/dL.

14. FTI (Free T4 or Free Thyroxine Index): The FTI is calculated by multiplying Total T4 and T3 Uptake. It helps diagnose thyroid disorders. The normal FT3 range is 2.3 - 4.1 pg/mL.

15. Thyroxine-binding globulin (TBG): The TBG blood test measures the level of a protein that moves thyroid hormone throughout your body.

Collecting data on these attributes will enable the development of a robust Thyroid Disease Detection model, offering valuable insights for effective diagnosis and treatment decisions.

2.6 Tools used

Python programming language and frameworks such as NumPy, Pandas, Scikit-learn, Matplotlib, Plotly, Flask etc are used to build the whole model.

- Virtual Studio Code is a used as IDE
- For visualization of the plots, Matplotlib, Seaborn and Plotly are used.
- Cassandra database is used DB operations
- Python, Streamlit is used for backend development and deployment
- Github is used as Version Contol System.

7. Constraints

The Thyroid Disease Detection solution system must be correct enough that it not mislead any report and as automated as possible and users should not be required to know any of the workings.

8. Assumptions

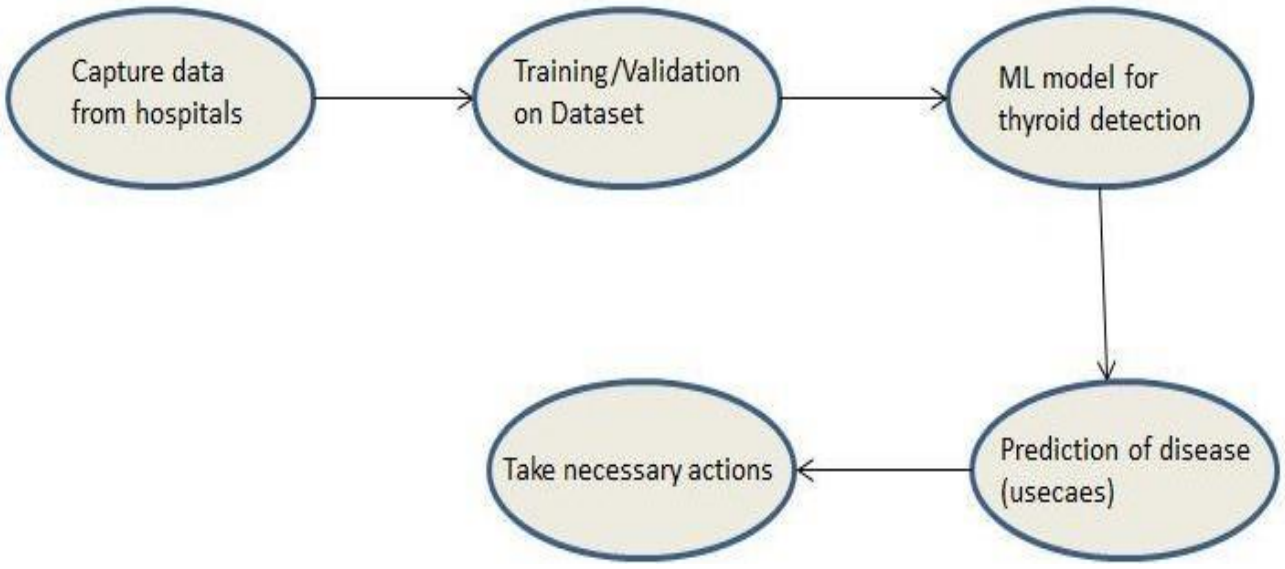
The main objective of the project is to implement the use cases as previously mentioned for new dataset that comes through Hospitals which has this solution install in their campus to capture people reports.

3 Design Details

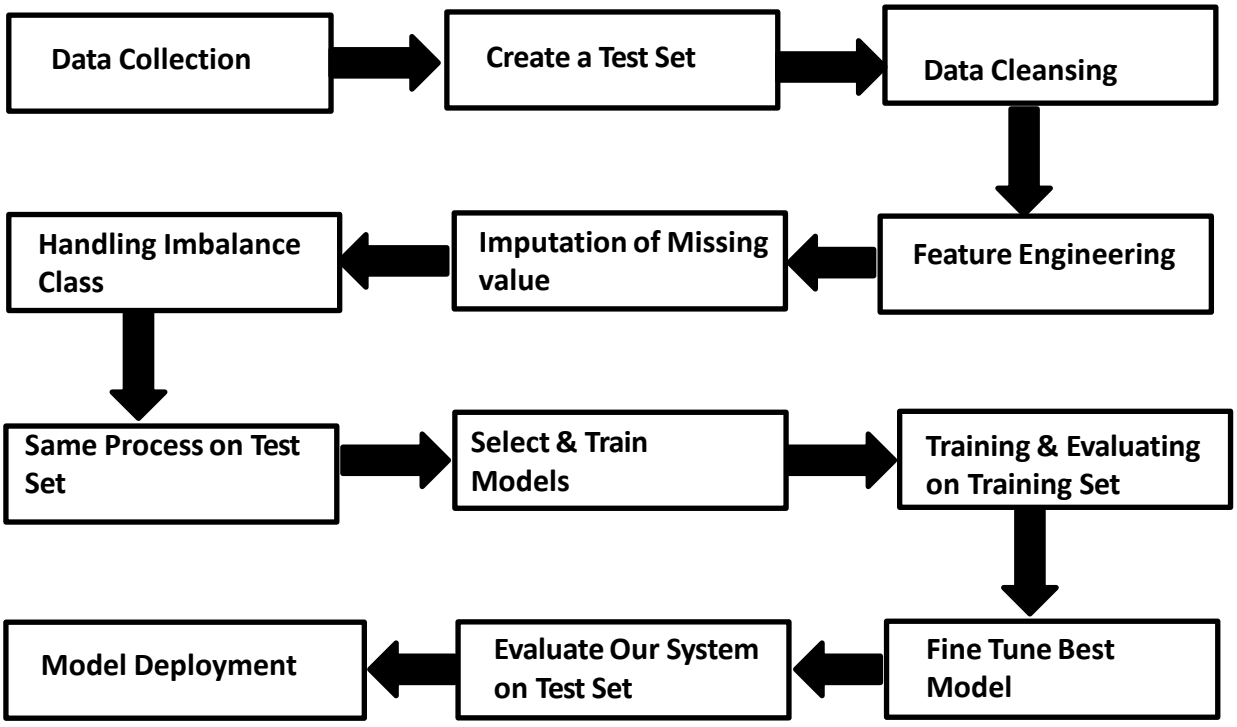
1. Process Flow

For detecting thyroid disease, we will use machine learning base model. Below is the process flow diagram is as shown below

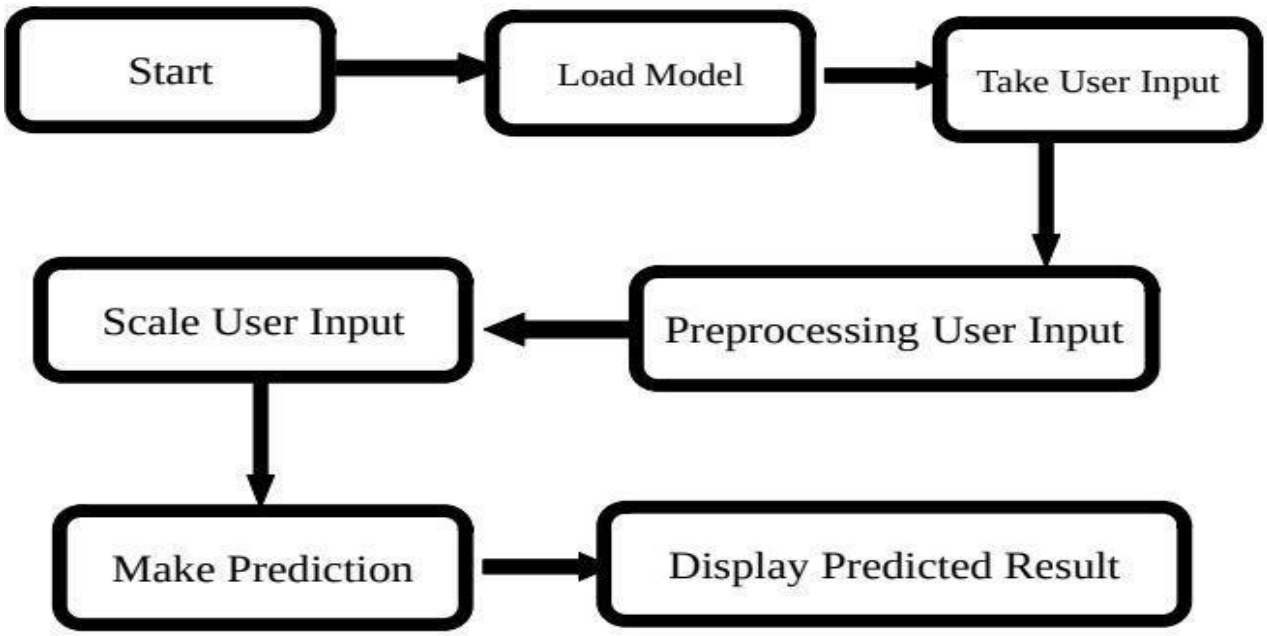
Proposed methodology



3.1.1 Model Training and Evaluation



3.1.2 Deployment Process



3.2 Event log

The system should log every event so that the user will know what process is running internally.

Initial Step-By-Step Description:

1. The System identifies at what step logging required.
2. The System should be able to log each and every system flow.
3. Developer can choose logging method. You can choose database logging/ File logging s well.
4. System should not hang even after using so many loggings. Logging just because we can easily debug issues so logging is mandatory to do.

3.3 Error Handling

Should errors be encountered, an explanation will be displayed as to what went wrong? An error will be defined as anything that falls outside the normal and intended usage.

4 Performance

We are developing a machine learning-based Thyroid Disease Detection solution for identifying thyroid diseases in patients who exhibit symptoms related to thyroid issues. The main objective is to facilitate early detection and prompt intervention. To ensure the system's effectiveness, regular model retraining will be implemented to continuously improve its performance and accuracy over time. This will enable us to take necessary actions promptly and provide appropriate medical attention to individuals diagnosed positively for thyroid disease.

1. Reusability

The code written and the components used should have the ability to be reused with no problems.

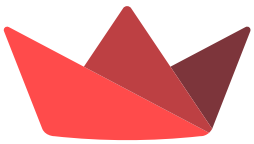
2. Application Compatibility

The different components for this project will be using python as an interface between them. Each component will have its own task to perform, and it is the job of the Python to ensure proper transfer of information.

3. Resource utilization

When any task is performed, it will likely use all the processing power available until that function is finished.

4. Deployment



Streamlit

Google Cloud

5 Conclusion

We propose to develop a Thyroid Disease Detection solution utilizing machine learning. The system will be trained on health-care domain data from patients who have undergone thyroid diagnosis. The model's performance will be evaluated across various use cases. Subsequently, we will utilize the trained model to predict the presence of thyroid disease in individuals exhibiting symptoms. In case of a positive prediction, the system will promptly alert the individuals, ensuring they receive immediate medical attention and treatment. Our primary focus is on achieving high accuracy to minimize the risk of generating misleading reports..

6 References

UCI Machine Learning Repository For Data Set

URL: https://archive.ics.uci.edu/ml/datasets/thyroid+disease