High Level Design (HLD)

THYROID DISEASE DEDECTION APPLICATION

Version: 1.0
Date: 09/10/2023

Contents

Document Version Control

Date Issued	Version	Description	Author
09/10/2023	1	Initial HLD – V1.0	N sri sai charan reddy

Abstract

Thyroid disease affects at least one in ten individuals in India. This condition primarily impacts women aged 17 to 54. In its severe stages, thyroid disease can lead to cardiovascular complications, elevated blood pressure, heightened cholesterol levels, depression, and reduced fertility. The thyroid gland produces two crucial hormones, total serum thyroxin (T4) and total serum triiodothyronine (T3), which play a vital role in regulating the body's metabolism, ensuring proper cellular and organ function, energy production, temperature regulation, and protein synthesis.

Thyroid disease is categorized into three main types: euthyroidism, hyperthyroidism, and hypothyroidism, representing normal thyroid hormone levels, excessive hormone levels, and deficient hormone production, respectively. Euthyroidism indicates the healthy production of thyroid hormones at both the systemic and cellular levels. Hyperthyroidism, on the other hand, results from an overabundance of thyroid hormones in circulation and within cells. Hypothyroidism is primarily caused by inadequate thyroid hormone production and suboptimal alternative therapies.

Effective disease management is a significant concern for healthcare professionals, and accurate and timely diagnosis is crucial for patients. Advanced diagnostic methods have recently enabled the generation of comprehensive medical reports, including symptom-based assessments. Questions such as the causes of thyroid disorders, the affected age groups, and suitable treatments can all be addressed using machine learning techniques applied to healthcare data. By processing healthcare data and employing specific methodologies, valuable information can be extracted to

enhance disease diagnosis and treatment, improve decision-making, and reduce the risk of mortality.

1. Introduction

1.1 Why this High-Level Design Document ?

The purpose of this High Level Design (HLD) Document is to add the necessary details 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 reference manual for how the modules interact at a high level.

The HLD

- 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

1.2 Scope

The HLD document presents the structure of the system, such as the database architecture, application architecture (layers), application flow (Navigation), and technology architecture. The HLD uses nontechnical to mildly-technical terms which should be understandable to the administrators of the system.

1.3 Definitions

- TDD – Thyroid Disease Detection

2. General Description

2.1 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.

2.2 Problem statement Solution

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

2.3 Problem Solution Further Improvements

The proposed solution involves the implementation of a data science model based on machine learning to address the mentioned use cases. In the first use case, we will input data from a healthy individual without thyroid disease and evaluate whether the solution can accurately detect their health status. In the second use case, we will input data from an unhealthy individual who is already suffering from thyroid disease and assess whether our solution performs correctly in identifying their condition.

Here's a rephrased version:

The proposed approach entails the deployment of a data science model utilizing machine learning to address the specified scenarios. In the initial use case, we will gather data from a disease-free individual to determine the effectiveness of the proposed solution in detecting their healthy status. Subsequently, in the second use case, we will collect data from an afflicted individual already diagnosed with thyroid disease to assess the accuracy of our solution in correctly identifying their condition.

2.4 Further Improvements

The Thyroid disease detection solution can be added with more use cases in health care domain. TDD solution can also be synchronized

with other health care domain solution to give one step extra confirmation of health to those people who has little symptoms of thyroid disease also.

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
- Goiter test: A goiter can sometimes occur when your thyroid gland produces too much thyroid hormone (hyperthyroidism).

- Tumor 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 tumor.

- TSH level measure: It supervise thyroid gland, TSH released by pituitary gland. Normal TSH range for an adult: 0.40 - 4.50 mIU/mL (milliinternational units per liter 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 Thyroxin): 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 milliliter of blood).

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

In all the above mentioned attributes if attribute is having binary answer then we need it in Boolean and for measures we need them in float values.

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.

- PyCharm is used as IDE.
- For visualization of the plots, Matplotlib, Seaborn are used.
- Azure is used for deployment of the model.
- Cassandra is used to retrieve, insert, delete, and update the database.
- Front end development is done using HTML/CSS, Flask is used for backend development and for API development.
- GitHub is used as version control system.

2.7 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.

2.8 Assumptions

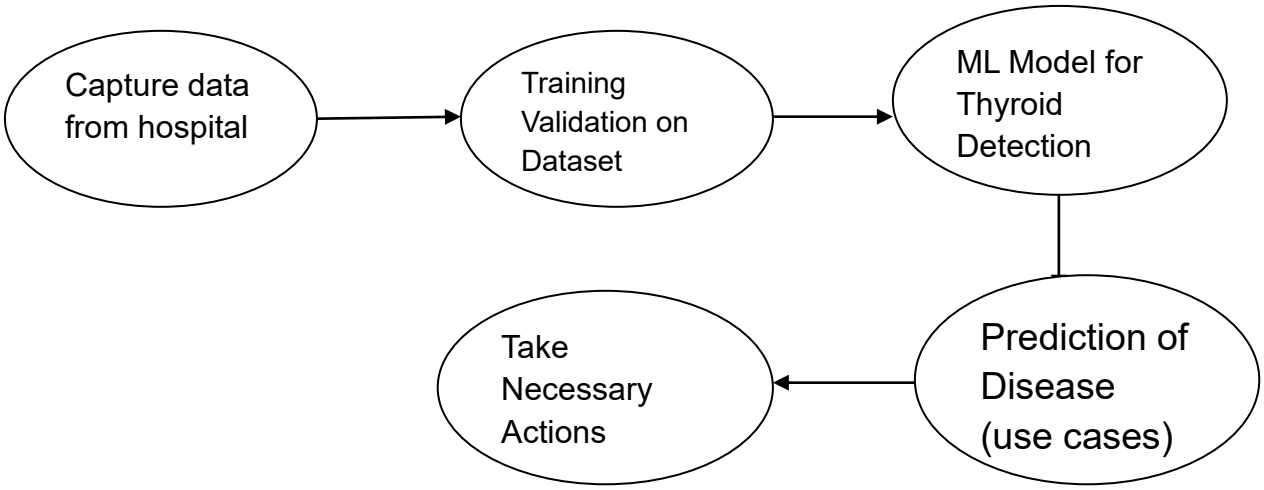
The main objective of the project is to implement the use cases as previously mentioned for new dataset that comes through Hospitals which has this solution install in their campus to capture people reports.

3 Design Details

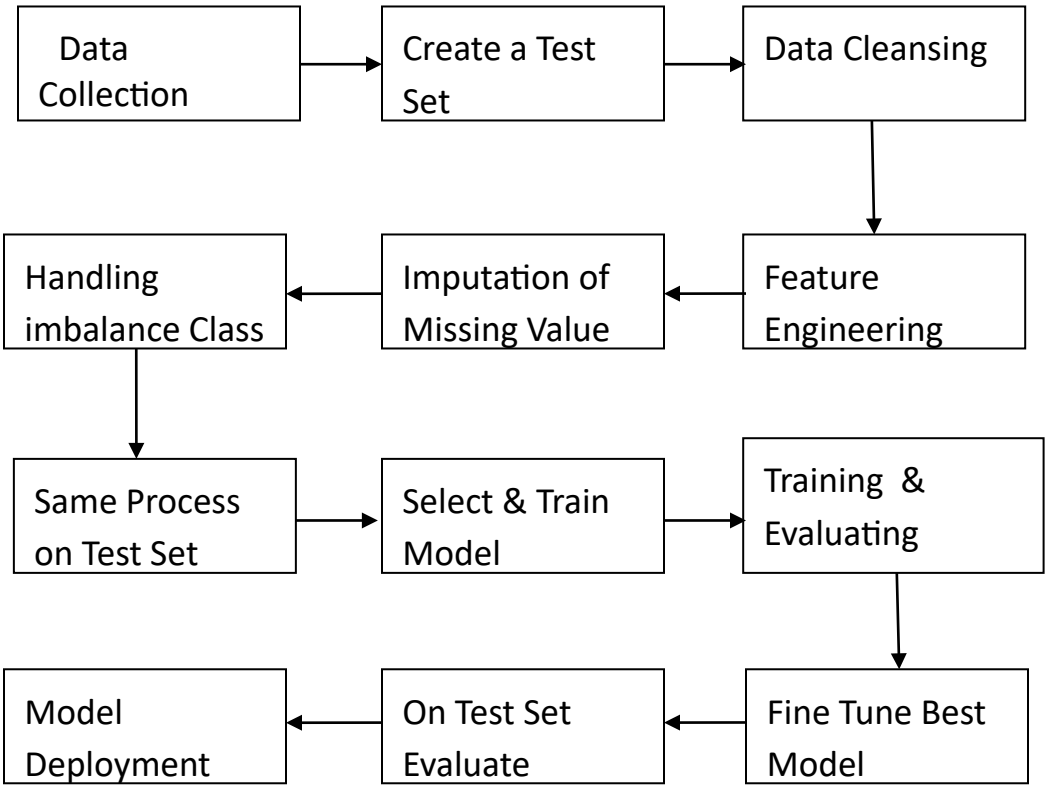
3.1 Process Flow

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

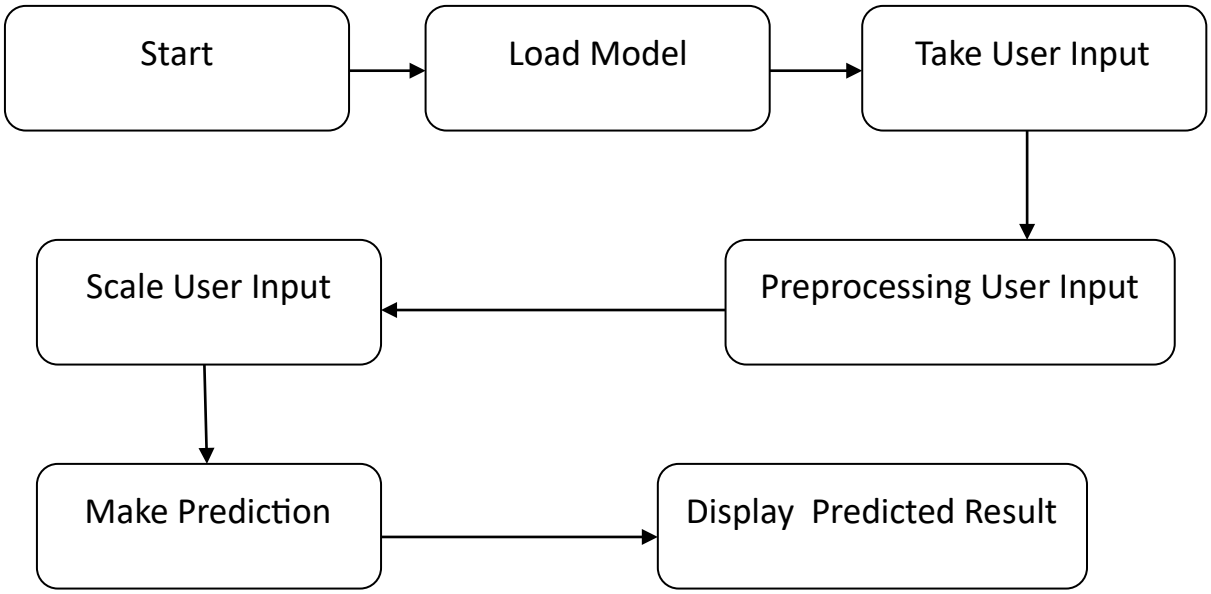
Proposed methodology



3.2 Model Training and Evaluation



3.3 Deployment Process



3.4 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.

3.5 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

The machine learning based Thyroid Disease Detection solution will used for detection of thyroid disease in patients having symptoms of thyroid. So that necessary action will be taken ASP. Also model retraining is very important to improve performance.

4.1 Reusability

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

4.2 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.

4.3 Resource utilization

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

4.4 Deployment

5 Conclusion

Thyroid Disease Detection solution will take health-care domain data of those patients who have undergone diagnosis for thyroid to train our machine learning model and will evaluate its performance over usecaes mentioned above. And then leverage its prediction to detect thyroid disease in people having symptoms of thyroid and able to alert people who is on positive side so that medical attention along with treatment will be given to that particular people as soon as possible. This solution should be as accurate as possible, so that chances of misleading reports will be taken good care of.

6 References

[UIC Machine Learning Repository For Data Set]

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