**HIGH LEVEL DESIGN**

Thyroid Disease Detection

By-Vinayak verma

# Abstract

Thyroid disease is a significant health issue in India, affecting around one in ten individuals, with a higher prevalence among women aged 17 to 54. Severe cases of thyroid disease can lead to complications such as cardiovascular problems, high blood pressure, elevated cholesterol levels, depression, and reduced fertility. The thyroid gland produces two key hormones, total serum thyroxine (T4) and total serum triiodothyronine (T3), which are essential for regulating metabolism, energy levels, protein synthesis, and body temperature.

Thyroid diseases are categorized based on hormone levels into three types: euthyroidism, hyperthyroidism, and hypothyroidism. Euthyroidism indicates normal hormone production and levels, hyperthyroidism results from excessive hormone levels, and hypothyroidism is due to insufficient hormone production or ineffective alternative therapies.

Effective diagnosis and timely treatment are crucial for managing thyroid disorders. Advanced diagnostic methods and symptom-based reports assist healthcare professionals in accurately identifying these conditions. Utilizing machine learning techniques on healthcare data allows for the generation of detailed medical reports, addressing questions about the causes of thyroid issues, the affected age groups, and appropriate treatments. This data-driven approach enhances the efficiency and accuracy of disease diagnosis, treatment planning, and decision-making, ultimately reducing the risk of fatalities



# 1 Introduction

**1.1 Why this High-Level Design Document?**

The High-Level Design (HLD) Document is crafted to furnish essential details for the coding phase, identify potential issues in advance, and deliver a comprehensive overview of the system's architecture and module interactions. It acts as a reference guide for developers and establishes a robust foundation for the project's 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



High

Level

Design

(

HLD

)

* Resource utilization
* Serviceability

# 1.2 Scope

The High-Level Design (HLD) document offers a comprehensive overview of the system's structure. It includes information about the database architecture, application architecture (layers), application flow (navigation), and technology architecture. The HLD is designed to be understandable to system administrators, using primarily non-technical language with some mildly technical terms. The aim is to ensure administrators can easily understand and grasp the system's framework without being burdened by complex technical jargon.



**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

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 ProposedSolution

The proposed solution is a data science model leveraging machine learning to address the specified use cases. In the first use case, data from a healthy individual without thyroid disease will be input to assess the model's ability to correctly identify such cases. In the second use case, data from an unhealthy individual already diagnosed with thyroid disease will be input to evaluate the model's performance and ensure it can accurately detect the condition.

# 2.4 Further Improvements



High

Level

Design

(

HLD

)

The Thyroid Disease Detection solution can be expanded to include additional use cases within the healthcare domain. By integrating this model with other healthcare solutions, it can provide an extra layer of confirmation for individuals exhibiting 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 requirements are entirely based on our problem statement. We need data from individuals who have already undergone thyroid blood tests to determine whether they have thyroid disease and, if so, what type of thyroid disease they have. The data will include both personal details and results from the blood tests. The necessary attributes are as follows:

* **Age**: Thyroid issues often depend on age, particularly affecting those over 60, especially women.
* **Gender**: Women are about five to eight times more likely to be diagnosed with a thyroid condition than men.
* **Current Thyroxin Treatment**: Indicates if the person is already undergoing thyroxin treatment.
* **Current Anti-Thyroid Medication**: Indicates if the person is on anti-thyroid medication.
* **Pregnancy (if female)**: Postpartum thyroiditis occurs in 5% to 9% of women after childbirth.
* **Current Illness**: Whether the person is sick at the time of diagnosis.
* **Iodine Test**: Both excess and deficiency in iodine can cause thyroid disease.
* **Lithium Test**: Lithium affects the thyroid by inhibiting thyroidal iodine uptake.
* **Goitre Test**: A goitre can occur when the thyroid gland produces too much thyroid hormone (hyperthyroidism).
* **Tumour Test**: Thyroid cancer is caused by genetic mutations in thyroid cells, leading to rapid growth and tumour formation.
* **TSH Level Measurement**: TSH, released by the pituitary gland, regulates the thyroid gland. The normal TSH range for an adult is 0.40 - 4.50 mIU/mL.
* **T3 Level Measurement**: Thyroid hormone that should be within the normal range.
* **T4 Level Measurement**: Low T4 is associated with hypothyroidism, while high T4 indicates hyperthyroidism. The normal T4 range for an adult is 5.0 – 11.0 ug/dL.
* **FTI (Free T4 or Free Thyroxine)**: The free T4 index (FTI) is used to diagnose thyroid disorders. It is calculated by multiplying Total T4 by T3 Uptake. The normal FT3 range is 2.3 - 4.1 pg/mL

## 2.6 Toolsused

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



High

Level

Design

(

HLD

)



* 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.



# 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

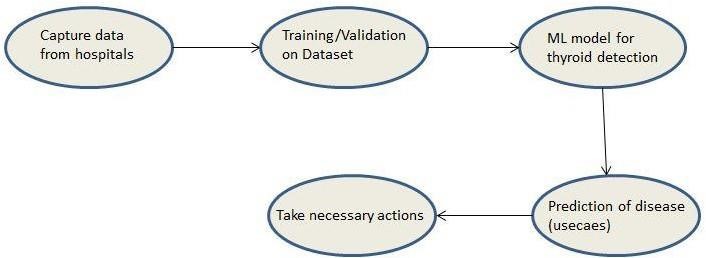
The main objective of the project is to implement the use cases as previously mentioned for new dataset that comes through Hospitals which has this solution install in their campus to capture people reports.

**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



09

### 3.1.1 Model Training and Evaluation

**Data**

**Collection**

**Create**

**a**

**Test**

**Set**

**Data**

**Cleansing**

**Feature**

**Engineering**

**Imputation**

**of**

**Missing**

**value**

**Handling Imbalance**

**Class**

**Evaluate**

**Our**

**System**

**on**

**Test**

**Set**

**Training**

**&**

**Evaluating**

**on**

**Training Set**

**Select**

**&**

**Train**

**Models**

**Same**

**Process**

**on**

**Test**

**Set**

**Fine**

**Tune**

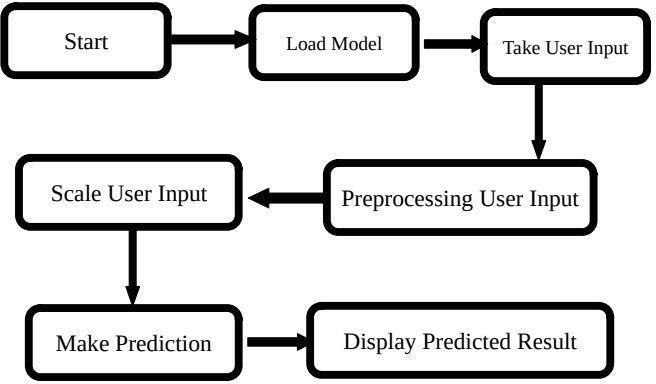
**Best**

**Model**

**Model**

**Deployment**

### 3.1.2 Deployment Process



## 3.2 Event log

The system should log every event so that the user is aware of the internal processes.

**Initial Step-By-Step Description:**

1. The system identifies the steps where logging is required.
2. The system should log every part of the system flow.
3. Developers can choose the logging method, either database logging or file logging.
4. The system should maintain performance and not hang despite extensive logging. Logging is essential for easy debugging, making it a mandatory practice.

## 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 to identify thyroid diseases in patients showing related symptoms. The primary goal is to facilitate early detection and prompt intervention. To ensure the system's effectiveness, regular model retraining will be conducted to continuously enhance its performance and accuracy. This approach enables timely actions and ensures that individuals diagnosed with thyroid disease receive appropriate medical attention promptly.

# 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 Conclusion

We propose to develop a Thyroid Disease Detection solution utilizing machine learning. The system will be trained on healthcare domain data from patients who have undergone thyroid diagnosis. The model's performance will be evaluated across various use cases. Subsequently, we will use the trained model to predict the presence of thyroid disease in individuals exhibiting symptoms. In the 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.

# 4.5 References

[UCI](https://archive.ics.uci.edu/) [Machine](https://archive.ics.uci.edu/) [LearningRepository For](https://archive.ics.uci.edu/) [Data](https://archive.ics.uci.edu/) [Set](https://archive.ics.uci.edu/)

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