Low-Level Design (HLD)

Thyroid Detection

Revision Number: 1.0

Last date of revision:

Document Version Control

|  |  |  |  |
| --- | --- | --- | --- |
| Date Issued | version | Descrition | Author |
|  | 1.0 | First Draft | Ritik Bhambid |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

Contents

Document Version Control……………………………………………………………………2

Abstract………………………………………………………………………………………..4

1. Introduction……………………………………………………………………………5
   1. Why this Low-Level Design Document?.................................................................5
   2. Scope………………………………………………………………………...........5
   3. Constraints………………………………………………………………………...5
   4. Risks……………………………………………………………………………….6
   5. Out of Scope……………………………………………………………………….6
2. Technical specifications……………………………………………………………...7

2.1 Dataset……………………………………………………………………………7

2.2 Dataset overview………………………………………………………………..…7

2.3 Input schema…………………………………………………………………..…..7

2.4 Logging………………………………………………………….…………….…..8

2.5 Database………………………………………………………….…………….….8

2.6 Deployment………..……………………………………………………………....8

3. Technology stack…………………………………………………..…………………..9

4. Proposed solution…………………………………………………………………….10

5. Model training/validation workflow…………………………………………………11

6. User I/O workflow…………………………………………………………………..12

7. Error Handling……………………………………………………………………….13

8. Test Cases……………………………………………………………………………14

9. Key performance indicators(KPI)……………………………..……………………..15

Abstract

Thyroid disease is a major cause of formation in medical diagnosis and in the prediction, onset to which it is a difficult axiomin the medical research. Thyroid gland is one of the most important organs in our body. The secretions of thyroid hormones are culpable in controlling the metabolism. 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. It is important to ensure a decent knowledge base that can be entrenched and used as a hybrid model in solving complex learning task, to predict the estimated risk on a patient's chance of obtaining thyroid disease.

1.Introduction

1.1 Why this Low-Level Design Document?

The purpose of this Low-Level Design (LLD) 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 main objective of the project is To assist doctors machine learning can help them in diagnosis of disease and reduces their burden.

* Predict high thyroid chances
* 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
* Application compatibility
* Resource utilization

1.2 Scope

The LLD documentation presents the structure of the system, such as the database architecture, application architecture (layers), application flow (Navigation), and technology architecture. The LLD uses non-technical to mildly-technical terms which should be understandable to the administrators of the system. This software system will be a Web application This system will be designed to detect Types of Thyroid.

1.3 Constraints

We will only be predicting Thyroid disease.

1.4 Risks

If the user deletes the storage or any specific services used while training and development from the cloud provider’s portal, the related history and detail stored in our portal will be destroyed completely.

1.5 Out of Scope

Thyroid cancer detection is out of scope for this project.

**2.Technical specifications**

2.1 Dataset

|  |  |  |
| --- | --- | --- |
| Dataset | Finalized | Type of model to be developed |
| Thyroid Disease | Yes | Classification |
|  |  |  |
|  |  |  |
|  |  |  |

2.2 Dataset overview

Thyroid Detection is the dataset used to Predicting thyroid disorder. It consists of multiple columns (30 columns) including the target output.

2.3 Input schem

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Feature Name | Validation sample | Data type | Size | Null/Required |
| Sample file name | hypothyroid\_0211198\_0102062112042 | CSV |  | Required |
| Length Of Date Stamp In File |  |  | 7 |  |
| Length Of Time Stamp In File |  |  | 12 |  |
| Number of Columns |  |  | 30 |  |
| Col Name | Age | int |  | Required |
| sex | Varchar |  | Required |
| on\_thyroxine | Bool |  | Required |
| on\_antithyroid\_medication | Bool |  | Required |
| pregnant | Bool |  | Required |
| thyroid\_surgery | Bool |  | Required |
| I131\_treatment | Bool |  | Required |
| query\_hypothyroid | Bool |  | Required |
| query\_hyperthyroid | Bool |  | Required |
| lithium | Bool |  | Required |
| goitre | Bool |  | Required |
| tumor | Bool |  | Required |
| hypopituitary | Bool |  | Required |
| psych | Bool |  | Required |
| TSH\_measured | Bool |  | Required |
| TSH | Float |  | Required |
| T3\_measured | Bool |  | Required |
| T3 | Float |  | Required |
| TT4\_measured | Bool |  | Required |
| TT4 | int |  | Required |
| T4U\_measured | Bool |  | Required |
| T4U | int |  | Required |
| FTI\_measured | Bool |  | Required |
| FTI | int |  | Required |
| TBG\_measured | Bool |  | Required |
| TBG |  |  | Required |
| referral\_source | Varchar |  | Required |
| Class | Varchar |  | Required |
|  |  |  |  |

2.4 Logging

We should be able to log every activity done by the incidents.

* The System identifies at what step logging required
* The System should be able to log each and every system flow.
* Developers can choose logging methods. You can choose database logging/ File logging as well.
* System should not be hung even after using so many loggings. Logging just because we can easily debug issues so logging is mandatory to do.

2.5 Database

System needs to store every request into the database and we need to store it in such a way that it is easy to retrain the model and debugging it.

1. The system stores each and every data given by the user or received on request to the database. Database can be SQLite.

1. User details and credentials are also stored in database. (The credentials will be stored in encrypted form).
   1. Deployment

This web application will be deployed over Microsoft Azure or Amazon Web Service based consider the cost and the feasibility of the system.

3.Technology stack

|  |  |
| --- | --- |
| Front End | HTML/CSS/JS |
| Backend | Python Flask |
| Database | SOLite |
| Deployment | Heroku |

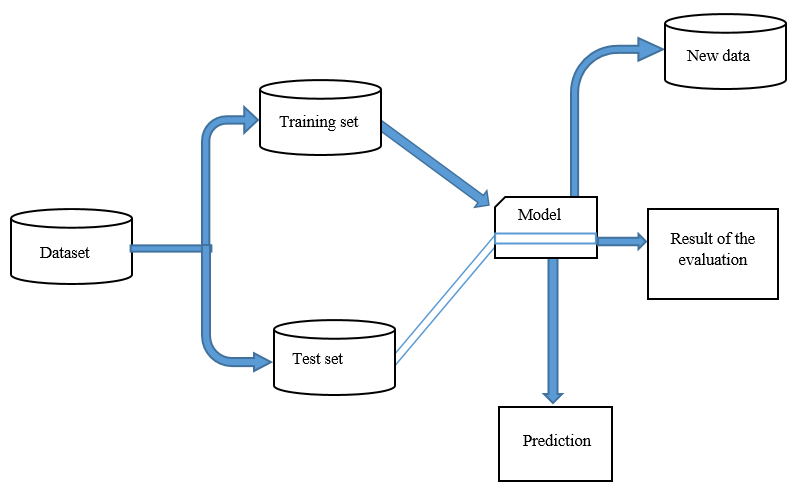
4. Proposed Solution

For predicting Thyroid disease analysing blood report is required to analyse and predict disease. Thyroid blood test data set analysis will be conducted using various supervised machine learning classifier techniques. Based on the accuracy of different algorithm, best accuracy algorithm will be chosen to fetch the result.

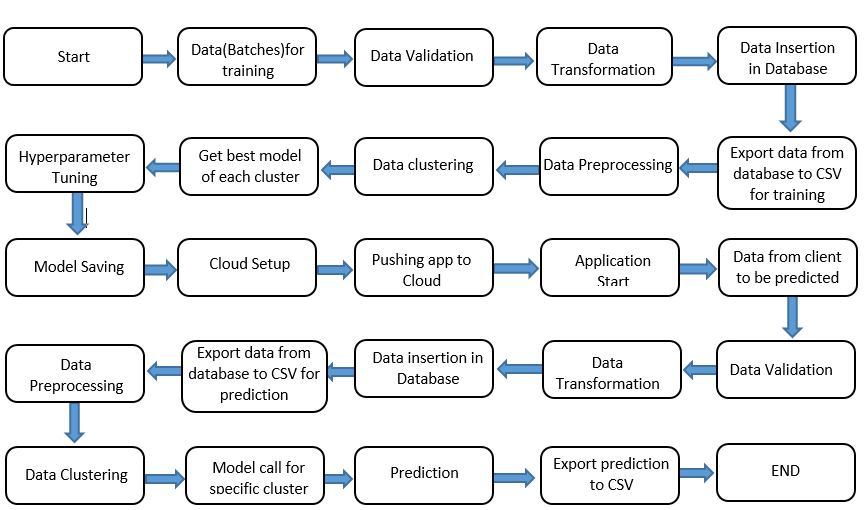
For first part, thyroid data set is taken from UCI repository. The dataset of hyperthyroidism and hypothyroidism is used where hyper and hypo are the two labels. These data set need to be checked before feeding it to training. There may be presence of null data or unnecessary data, this should undergo data cleaning to remove such data. Cleaned data is used as training data and test data. And make a clutters for better model accuracy and reduce the detection time.

The algorithm. Based on the feature extracted, probability will be generated for test data by comparing the features of both. Highest probability value will be classified to that particular label whether it is hyperthyroidism or hypothyroidism. Web app is developed using python where the chosen ML model will be linked with web app. The users blood test data will be entered in web app will process the data using model and result will be designed on prediction file.

5.Model training/validation workflow



6.User I/O workflow



7. 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.

8.Test cases

|  |  |  |  |
| --- | --- | --- | --- |
| Test case | Steps to perform test case | Module | Accuracy |
| 1. |  |  |  |

9.Key performance indicators (KPI)

* Time and workload reduction using the Thyroid detection model.
* Comparison of accuracy of model prediction and doctor’s prediction.