Architectural Document

TDPS (Thyroid Disease Prediction System)

Revision Number: 1.0

Last date of revision: 20/01/2024

K S Sujith

# Document Version Control

|  |  |  |  |
| --- | --- | --- | --- |
| Date Issued | Version | Description | Author |
| 20th Jan 2024 | 1.1 | First Draft | K S Sujith |
|  |  |  |  |

**Contents**

[Document Version Control 2](#_gjdgxs)

[**Abstract** 4](#_30j0zll)

[1](#_1fob9te) Introduction 5

[1.1](#_3znysh7) Why this Low-Level Design Document? 5

[1.2](#_2et92p0) Scope 6

[1.3](#_tyjcwt) Constraints 6

[1.4](#_3dy6vkm) 1.3 Risks 6

[1.5](#_1t3h5sf) 1.4 Out of Scope 7

[2](#_4d34og8) Technical specifications 7

[2.1 Predicting Disease 7](#_2s8eyo1)

[2.2 Logging 7](#_17dp8vu)

[3](#_26in1rg) Technology stack 8

[4](#_lnxbz9) Proposed Solution 9

[5](#_35nkun2) Model training/validation workflow 10

[6](#_1ksv4uv) Exceptional scenarios 13

**Abstract**

### Thyroid disease (TD) is one of the most progressive endocrine disorders in the human population today. Prediction of the endocrine disease is a critical task in the field of clinical data analysis. Machine Learning (ML) has shown effective results in the decision making and predictions from the enormous data generated by healthcare domain. Various studies in the prediction of thyroid disease have given only a glimpse using machine learning algorithms. In this project we have proposed three models based on the primary dataset collected from 3772 patients.

# Introduction

## Why this Low-Level Design Document?

The purpose of this document is to present a detailed description of the Thyroid Disease Prediction System Design. It will explain the purpose and features of the system. This document is intended for both the stakeholders and the developers of the system and will be proposed to the higher management for its approval.

There are following parameters as data:

* age
* sex
* query on thyroxine
* query on thyroxine
* on antithyroid
* medication
* sick
* pregnant
* thyroid
* surgery I131
* treatment
* query hypothyroid
* query hyperthyroid
* lithium goitre
* tumor
* hypopituitary
* psych
* TSH
* measured TSH
* T3
* measured T3
* TT4
* measured TT4
* T4U
* measured T4U
* binaryClass

This project shall be delivered in three phases:

Phase 1: All the functionalities with PyPi packages.

Phase2: Integration of UI to all the functionalities.

Phase3: Deploy in Cloud.

## Scope

This software system will be a Web application This system will be designed to detect the diseases at earliest for better disease management, improved interventions, and more efficient health-care resource allocation using. This system is designed to predict the diseases from patient information such as persona info, disease history, lab results, procedures and medications.

## Constraints

We will only be selecting a sample from a specific geographic area.

## Risks

Document specific risks that have been identified or that should be considered.

## Out of Scope

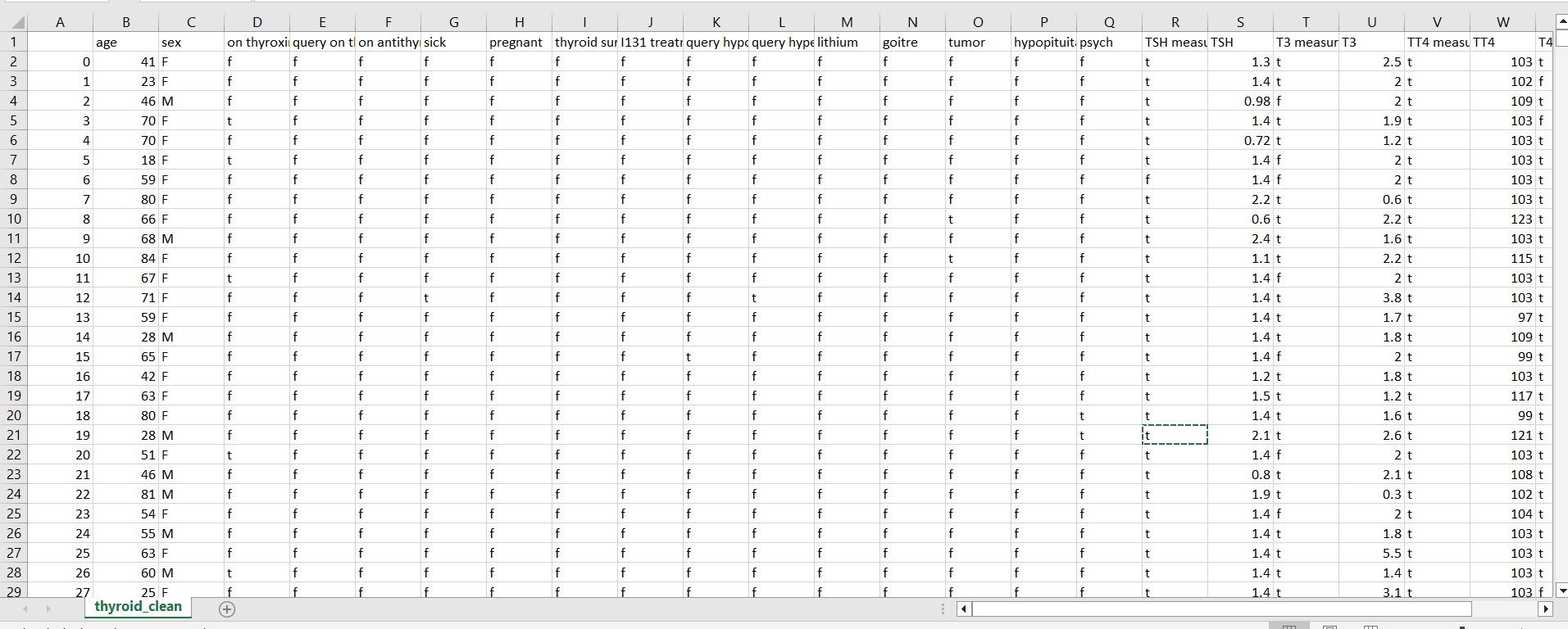
Delineate specific activities, capabilities, and items that are out of scope for the project.

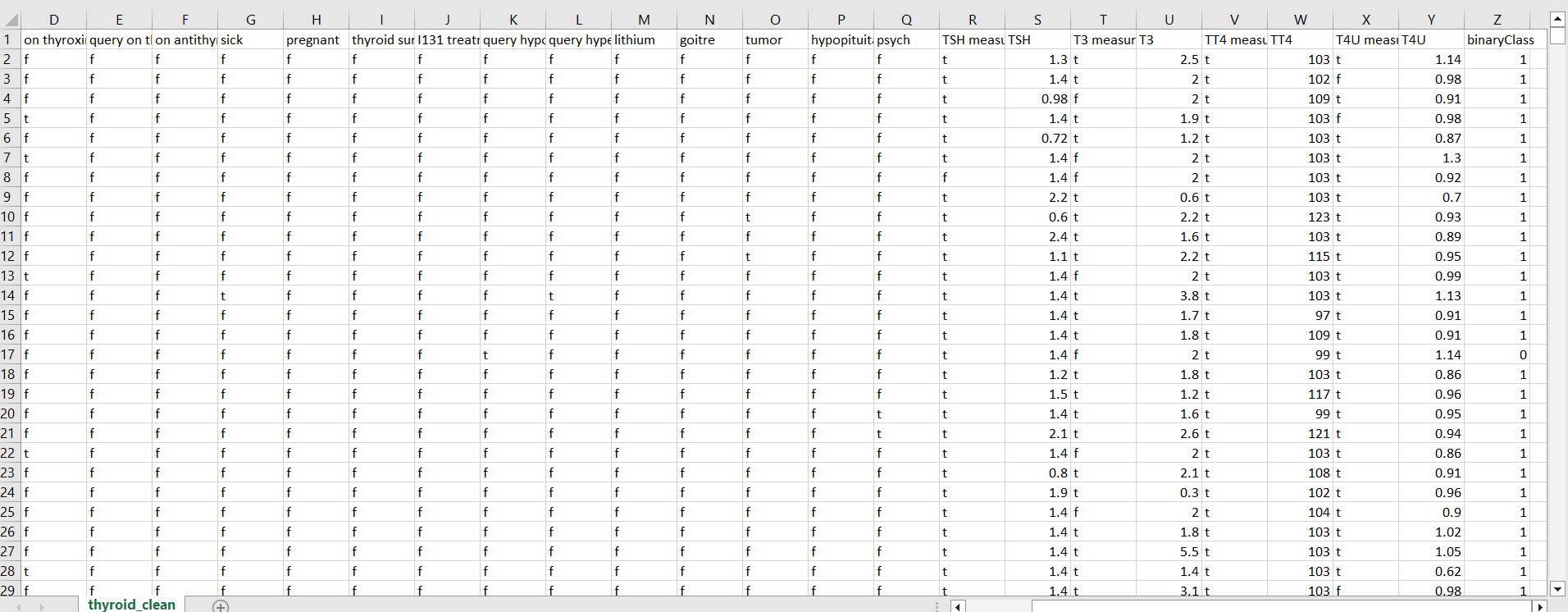
# Technical specifications

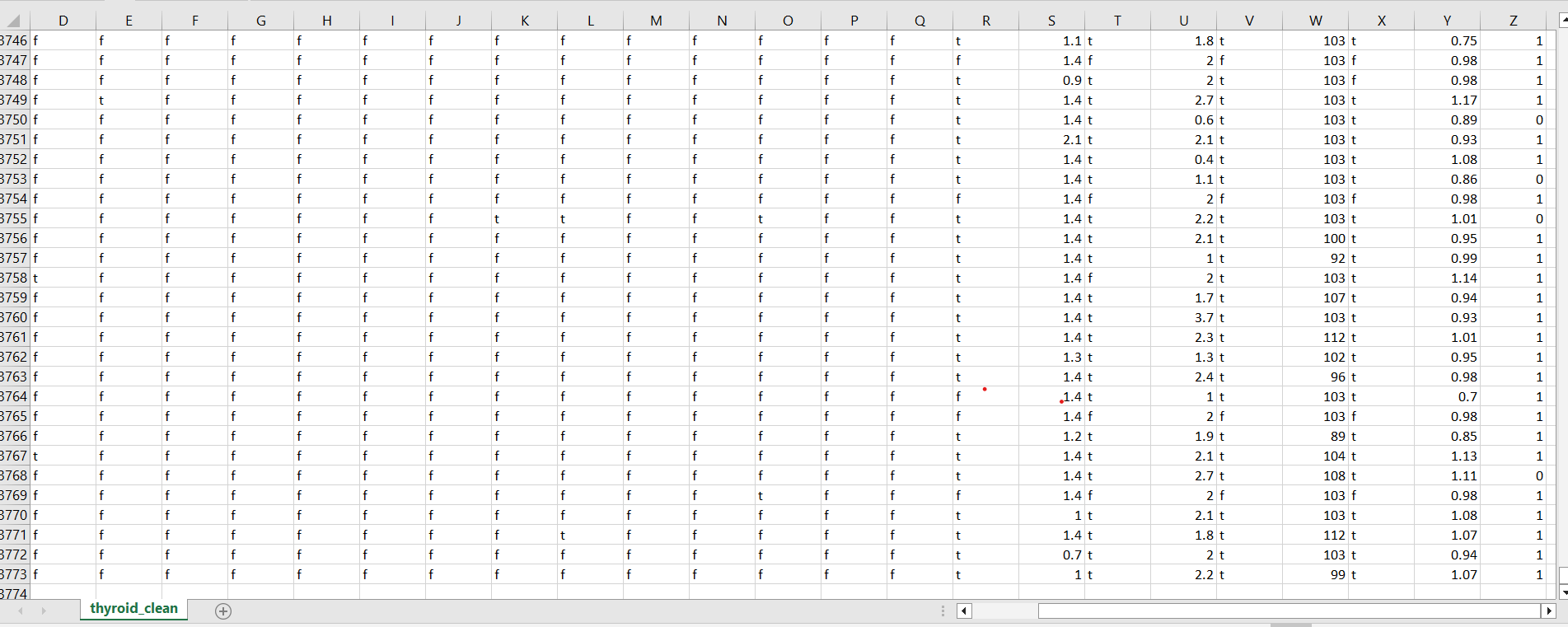
## 2.1 Dataset

## 2.1.1 Diabetes dataset overview

The dataset is as shown below.







## 2.2 Predicting Disease

* The system displays the fields to enter different input parameters for the model to analyse the data based on the highest accuracy model bult.
* The system should be able to predict chance of the disease based on the user information.

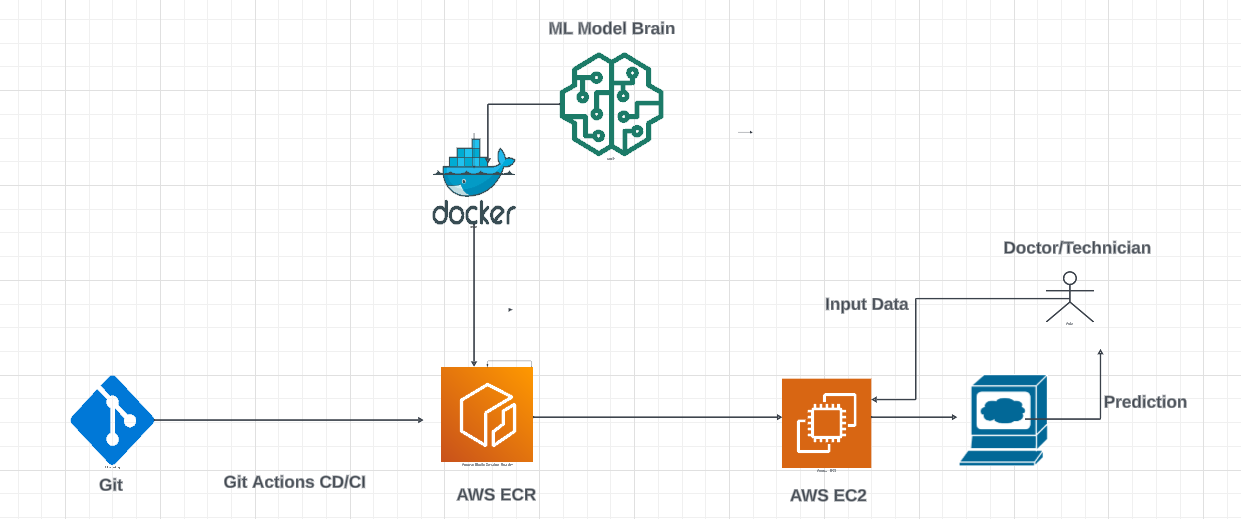
## 2.3 Logging

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

* 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 Deployment**

1. AWS cloud



# Technology stack

|  |  |
| --- | --- |
| **Front End** | HTML/CSS |
| **Backend** | Python Flask |
| **Deployment** | AWS |

# Proposed Solution

The Thyroid Disease Prediction System is a classification problem. So tried different ML Classification models like Logistic Regression, Ridge Classifier(),RandomForestClassifier,DecisionTree, KNeighborsClassifier,SVM,Naiave-Bayes,MLPClassifier,SupportVectorMachine .Of these the model having highest accuracy score is selected as base model to predict the outcomes.

.

# Model training/validation workflow

