Thyroid Disease Detection

ARCHITECTURE DESIGN DOCUMENTATION

Document Version Control

Date	Version	Description	Author
09-11-2022	1.0	Architecture Design	Archana Bora

Table of Contents

Abstract

Thyroid disease is a major cause of formation in medical diagnosis and in the prediction, onset to which it is a difficult axiom in the medical research. Thyroid gland is one of the most important organs in our body. This gland secretes two hormones which help in controlling the metabolism of the body. The two types of Thyroid disorders are Hyperthyroidism and Hypothyroidism. When this disorder occurs in the body, they release certain type of hormones into the body which imbalances the body's metabolism. The goal is to predict the probability of the disease based on patient's health history.

1. Introduction:

1.1 Why this Architecture Design Document?

The main objective of the Architecture design documentation is to provide the internal logic understanding of the Thyroid Disease Detection code. The Architecture design documentation is designed in such a way that the programmer can directly code after reading each module description in the documentation.

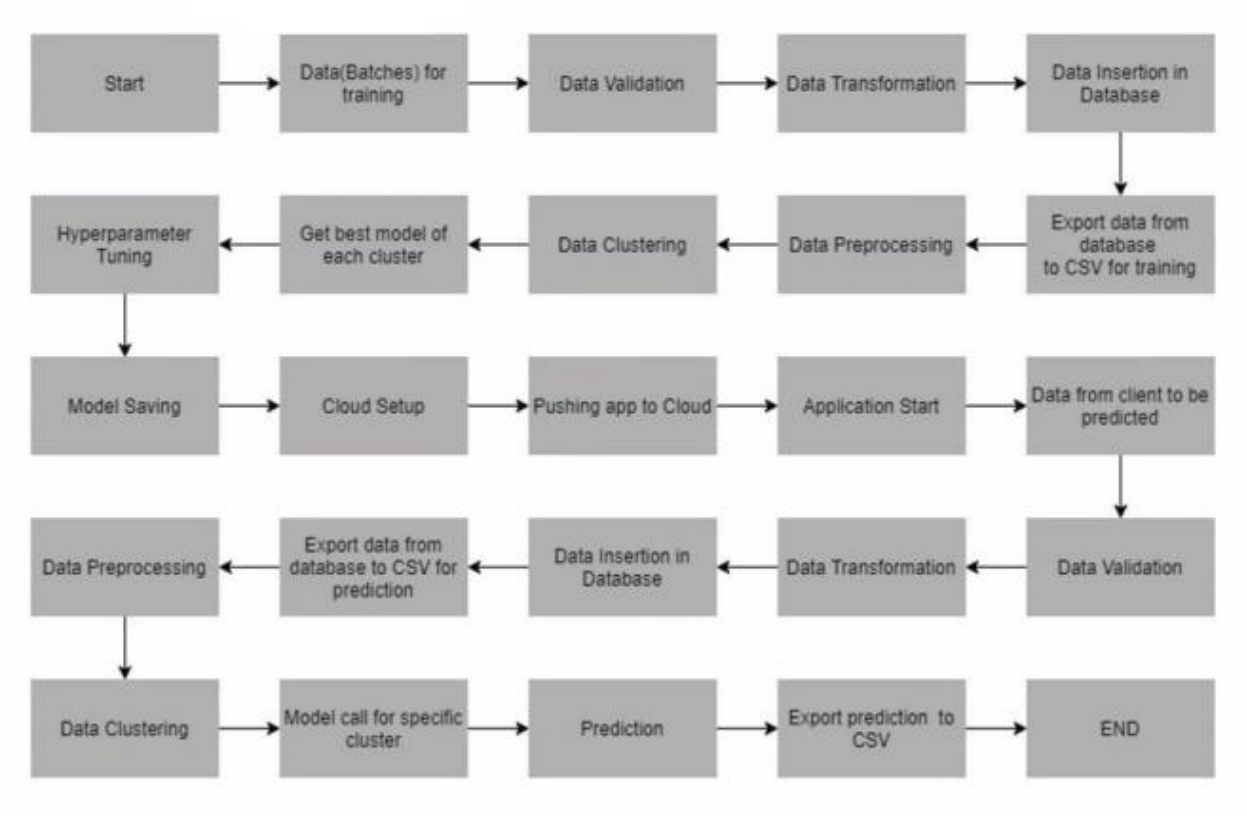
1.2 Scope

Architecture Design (AD) is a component-level design process that follows a step-by step refinement process. This process can be used for designing data structures, required software, architecture, source code, and ultimately, performance algorithms. Overall, the data organization may be defined during requirement analysis and then refined during data design work. And the complete workflow.

2. Dataset Information:

We will be using Thyroid Disease Dataset present in UCI Machine Learning Repository.
The Data set is satisfying our data requirement. Total 7200 instances are present in different Batches of data.

3. Architecture:



4. Architecture Design

4.1 DATA INGESTION

4.1.1 Data Validation

In this step we will do validation of data based on discussion with stack holder. In validation we will perform basic operations like file name validation, no of rows, no of columns in dataset etc…..

4.1.2 Data Transformation

In this step we need to transform data in order to store in our database. Basic operations like replacing nan values with null etc.…

4.1.3 Data insertion

After completion of data validation and transformation we need to insert our data into database. This step is required because we receive multiple batch files from user so that we need to consolidate those files.

4.2 CORE ML PIPELINE

4.2.1 Importing csv file

After insertion of data into database, now we need retire the data as csv file from database.

4.2.2 Data Preprocessing

ARCHITECTURE DESIGN

After importing csv file we need to start our machine leaning pipeline. As we know first step in ml pipeline is data preprocessing. In this we will convert raw data into good data.

4.2.3 Data Clustering

This step is required because when we train our data with single model, we have chance of getting high error. Instead of that creating multiple models will give better accuracy and less error compare with one model.

4.2.4 Hyper parameter tuning

As we know that we perform hyper parameter tunning to file best parameters for model. This step is required to create good model.

4.2.5 Model Saving

This is final step of ml pipeline. We need to save our model for future predictions.

4.3 DEPLOYMENT

4.3.1 Pushing code to github

To deploy our application on cloud first we need to push our local code into github repository.

4.3.2 Deploy using cloud

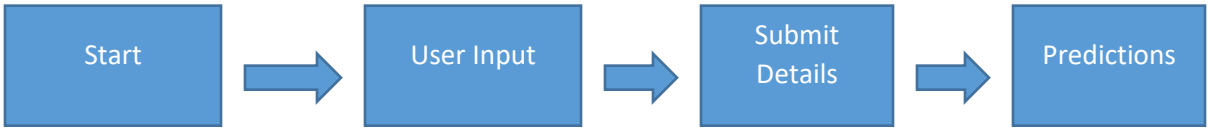
After pushing code repository into github. We can use any cloud deploy platforms to deploy our application.

4.4 PREDICTIONS

After successful deployment of application. Whenever the data comes from user, we need to perform all steps that are mentioned above and finally get predictions by using trained model.

ARCHITECTURE DESIGN

5. User Input / Output Flow:



6. Conclusion:

The project is designed in the flask; hence it is accessible to everyone. This study signifies machine learning and data mining techniques to benefit the medical field and healthcare system. According to the regular protocol, this study will help the doctors use this as a supplementary system, based on the model's predictions. The UI is made to be user-friendly so that the user will not need much knowledge of any tools but will just need the information for results.