Low level Design document (LLD)

Insurance Premium Prediction

Abstract

The Agriculture business domain is a vital part of the overall supply chain & is expected to highly evolve in the upcoming years with the introduction of cutting edge technologies like IOT, AI, drones etc.   
  
As sustainable consumption becomes increasingly important thus it became the basis of this project to initially analyze the crop production in INDIA so that at later stage advance technologies can be implemented to improve production with cost reduction.

Health Insurance is a type of insurance that covers medical expenses. A person who has taken a health insurance policy gets health insurance cover by paying a particular premium amount. There are a lot of factors that determine the premium of health insurance. The amount of the premium for a health insurance policy depends from person to person, as many factors affect the amount of the premium for a health insurance policy.

With Classical machine learning tasks, goal is to create a Model that predicts the premium of the person for health insurance.

# Introduction

## Why this High-Level Design Document?

The purpose of this High-Level Design (HLD) 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 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 features and the architecture of the project
    - List and describe the non-functional attributes like:
      * Security
      * Reliability
      * Maintainability
      * Portability
      * Reusability
      * Application compatibility
      * Resource utilization
      * Serviceability

## Scope

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

# General Description

## Product Perspective & Problem Statement

There are a lot of factors that determine the premium of health insurance. The amount of the premium for a health insurance policy depends from person to person, as many factors affect the amount of the premium for a health insurance policy.

The insurance.csv dataset contains 1338 observations (rows) and 7 features (columns). The dataset contains 4 numerical features (age, bmi, children and expenses) and 3 nominal features (sex, smoker and region) that were converted into factors with numerical value designated for each level.

## Proposed Solution

## The solution proposed here is a web application where user enters the details about age, bmi, children, sex, smoker & region & gets the premium in return. The purpose of this solution is to look into different features to observe their relationship, and plot a multiple linear regression based on several features of individual such as age, physical/family condition and location against their existing medical expense to be used for predicting future medical expenses of individuals that help medical insurance to make decision on charging the premium.

## 2.3 Tools used

Python programming Language with libraries as Pandas, Numpy & with Flask framework is used to build Web Application. For Deployment AWS Elastic Beanstalk is used to predict the Health Care Expenses of any individual.







# 3.Design Details

# 3.1 Process Flow

# For identifying different types of anomalies, we will use machine learning based model. Below is the process flow diagram

# C:\Users\Dell\Desktop\process flow.png

# 3.1.1 Model Training & Evaluation

# 

# 3.1.2 Deployment Process

# C:\Users\Dell\Desktop\deployment flow.png

# 3.2 Event Log



1.The system identifies at what step logging is required.

2.The system should be able to log each & every system flow.

3.Developer can choose logging method. You can choose database logging/File logging as well.

4.System should not hang even after so many loggings. Logging just we can easily debug issues so, logging is mandatory to do.

# 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 & intended usage.

# 4.Reusability



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

**4.3. Resource Utilization**

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

# 4.4 Deployment

# C:\Users\Dell\Desktop\deployment.png



**6.Conclusion**

Medical insurance premium prediction is an important area of research that has gained significant attention in recent years. With the increasing cost of healthcare, it is becoming increasingly important to accurately predict the insurance premiums that individuals should pay based on their medical history and other factors.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_End of Document\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Plav Tewari

Low level Design document (LLD)

Insurance Premium Prediction

**Change Control**

|  |  |  |  |
| --- | --- | --- | --- |
| **VERSION** | **DATE** | **AUTHOR** | **COMMENTS** |
| 0.1 | 15-Jun-- 2023 | Pallavi | Introduction and architecture defined |
| 0.2 | 16 - Jn - 2023 | Pallavi | Architecture & Architecture description appended and  updated. |
|  |  |  |  |
|  |  |  |  |

**Approval Status:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **VERSION** | **REVIEW**  **DATE** | **REVIEWED BY** |  | **APPROVED BY** | **COMMENTS** |
|  |  |  |  |  |  |

1. [Introduction 1](#_bookmark0)
   1. [What is Low-Level design document? 1](#_bookmark1)
   2. [Scope 1](#_bookmark2)
2. [Architcecture 2](#_bookmark3)
3. [Architecture Description 3](#_bookmark4)
   1. [Data Description 3](#_bookmark5)
   2. [Web Scrapping 3](#_bookmark6)
   3. [Data Transformation 3](#_bookmark7)
   4. [Data Insertion into Database 3](#_bookmark8)
   5. [Export Data from Database 3](#_bookmark9)
   6. [Data Pre-processing 3](#_bookmark10)
   7. [Data Clustering 3](#_bookmark11)
   8. [Model Building 4](#_bookmark12)
   9. [Data from User 4](#_bookmark13)
   10. [Data Validation 4](#_bookmark14)
   11. [User Data Inserting into Database 4](#_bookmark15)
   12. [Data Clustering 4](#_bookmark16)
   13. [Model Call for Specific Cluster 4](#_bookmark17)
   14. [Recipe Recommendation & Saving Output in Database 4](#_bookmark18)
   15. [Deployment 4](#_bookmark19)
4. [Unit Test Cases 5](#_bookmark20)

# Introduction

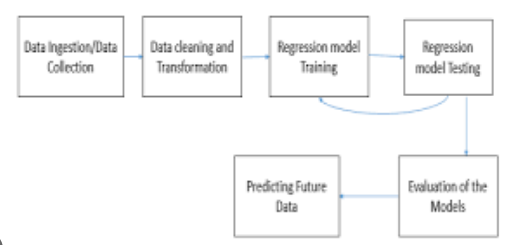
## What is Low-Level design document?

The goal of the LDD or Low-level design document (LLDD) is to give the internal logic design of the actual program code for the Crop Production in India Prediction dashboard. LDD describes the class diagram with the methods and relations between classes and programs specs. It describes the modules so that the programmer can directly code the program from the document.

## Scope

Low-level design (LLD) is a component-level design process that follows a step-by-step refinement process. The 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

# Architecture



# 3.Architecture Description

## 3.1 Data Description

The insurance.csv dataset contains 1338 observations (rows) and 7 features (columns). The dataset contains 4 numerical features (age, bmi, children and expenses) and 3 nominal features (sex, smoker and region) that were converted into factors with numerical value designated for each level.

3.2 Data Transformation

Integrated data is not ready to visualize data because the data should be transformed. To transform the data, it should be cleaned or pre-processed. For example, redundant or missing values are removed from the data sets. After data is pre-processed or cleaned, business rules are applied to transform the data. After processing the data, it is loaded into the data warehouse.

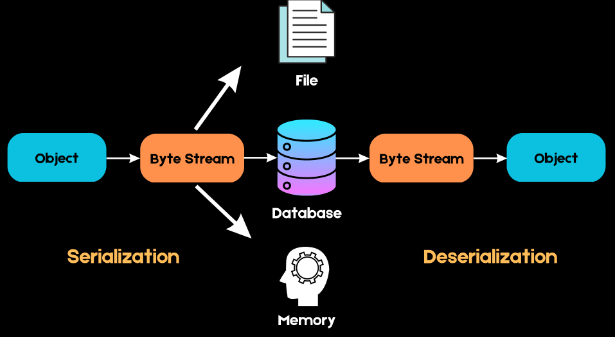
3.2 Model Building

In this step, first dataset Splitting is done. After that model is trained on different Machine Learning Algorithms such as:

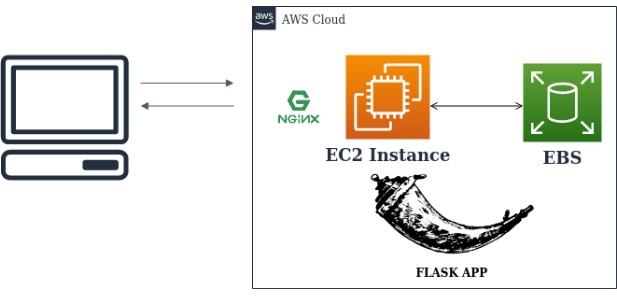
* Linear Regression
* Decision Tree Regressor
* Random Forest Regressor
* Gradient Boosting Regressor
* XGBoost Regressor
* KNN

## 3.3Model Selection: Tested all the models to check the RMSE & R-squared.

3.4 Pickle File: Selected model as per best RMSE score & R-squared and created pickle file using pickle library.



3.5 Webpage &Deployment:



Created a web application using Flask that takes all the necessary inputs from the user & shows the output. For Deployment AWS Elastic Beanstalk is used to predict the Health Care Expenses of any individual.

------------------------------End of Document------------------------------------