High Level Design (HLD)

High Level Design

Insurance Premium Prediction

Revision Number: 2.0

Last date of revision: 15/09/2021

# Document Version Control

|  |  |  |  |
| --- | --- | --- | --- |
| **Date Issued** | **Version** | **Description** | **Author** |
| 15/09/2021 | 1 | Initial HLD - V1.0 | Mayuri |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**Contents**

[Document Version Control 2](#_TOC_250021)

[Abstract 4](#_TOC_250020)

1. [Introduction 4](#_TOC_250019)
   1. [Why this High-Level Design Document? 4](#_TOC_250018)
2. [General Description 4](#_TOC_250015)
   1. [Project Perspective 5](#_TOC_250014)
   2. [PROPOSED SOLUTION 5](#_TOC_250012)
   3. [FURTHER IMPROVEMENTS 5](#_TOC_250011)
   4. [Data Requirements 5](#_TOC_250009)
   5. [Tools used 5](#_TOC_250008)
   6. [Constraints 6](#_TOC_250007)
   7. [Assumptions 7](#_TOC_250006)
3. [Design Details 7](#_TOC_250005)
   1. [Process Flow 7](#_TOC_250004)
   2. [Deployment 8](#_TOC_250001)
   3. [Event log 8](#_TOC_250000)
   4. Error Handling 9
4. Performance 9
   1. Reusability 9
   2. Application Compatibility……………………………………………….9
   3. Resource Utilization…………………………………………………….9
   4. Deployment………………………………………………………………9
5. Conclusion 10

# Abstract

Health insurance or medical insurance is a type of insurance that covers the whole or a part of the risk of a person incurring medical expenses. As with other types of insurance is risk among many individuals. By estimating the overall risk of health risk and health system expenses over the risk pool, an insurer can develop a routine finance structure, such as a monthly premium or payroll tax, to provide the money to pay for the health care benefits specified in the insurance agreement. The benefit is administered by a central organization, such as a government agency, private business, or not-for-profit entity. By using machine learning algorithm we are predicting the amount that to be paid by the user.

# 1.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 user 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

o Reusability

o Application compatibility

* Resource utilization
* Serviceability

|  |  |
| --- | --- |
|  |  |

# 2.General Description

## 2.1 Project Perspective

The insurance premium prediction project is a machine learning project which will help us to predict the amount to be paid based on given information.

### **PROPOSED SOLUTION**

The insurance premium prediction project estimates the amount that a person should pay based on some information like Age, Sex, Bmi, Children, Region, Smoker or not

**2.3 FURTHER IMPROVEMENTS**

We have given few inputs to estimate premium, but we can add more information about health diseases and many more.

## 2.4 Data Requirements

Data requirement completely depend on our problem statement.

1)We have taken 1339 records of information.

2)Data consists of 7 columns

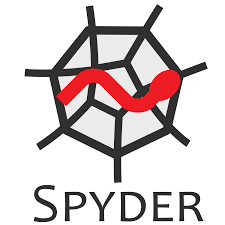
3)The last column is dependent variable that we need to estimate, the remaining six columns are those we must feed to the algorithm.

4)The data is fully balanced and there is no missing data.

5)We require 80 to 20 % of split for training and test data.

**2.5 Tools used**

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



* Spyder is used as IDE.
* For visualization of the plots, Matplotlib, Seaborn are used.
* Heroku is used for deployment of the model.
* Front end development is done using HTML/CSS
* Python Flask is used for backend development.
* GitHub is used as version control system.

## 2.6 Constraints

The Insurance Premium Prediction must be user friendly , as automated as possible.

## 2.7 Assumptions

The main objective of the project is to predict the amount for new dataset that comes through user interface.

**3.Design Details**

**3.1 Process Flow**

Data Cleaning

Data collection

Feature Engineering

Handling Categorical Variables

Feature Selection

Train test split

Model Training

Model Testing

Hyper Parameter Tuning

Model Deployment

Model Testing

# 3.2 Deployment Process

Take User Input

Load Model

Start

Scale User Input

Preprocessing User Input

Display Predicted Result

Make Prediction

# 3.3 Event Log

# In this Project we are logging every process so that the user will know what process is running internally.

# Step-By-Step Description:

# • In this Project we defined logging for every function, class.

# • By logging we can monitor every insertion, every flow of data in database.

# • By logging we are monitor every step which may create problem or every step which is important in file system.

# • We have designed logging in such a way that system should not hang even after so many logging's, so that we can easily debug issues which may arises during process flow.

# 3.4 Error Handling

# We have designed this project in such a way that, at any step if error occur then our application should not terminate rather it should catch that error and display that error with proper explanation as to what went wrong during process flow.

# 4.Performance

The insurance premium prediction is used to predict the amount based on the information given by the user through user interface.

## 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 Deployment**



**5.Conclusion**

The Insurance Premium Prediction will predict the insurance amount based on various anomalies data used to train our algorithm.