High Level Design (HLD)

High Level Design (HLD)

[**Insurance Price\_Prediction**](https://github.com/mohiteyashprogrammer/Adult_income_Prediction/tree/main)

Revision Number – 1.0

Last Date of Revision – 29-06-2023

**ILMA ARSHAD**

1

High-Low Level Designing (LLD)(HLD)

2

High Level Design (HLD)

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Contents** | | | | | | | | | | | | | | | | | | | | | |  |  |  |  |
| [**Abstract**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.gjdgxs) | | | | | |  | | | | | | | | | | | | | | | | [4](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.gjdgxs) | | | |
|  |  |  |  | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| [**INTRODUCTION**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.30j0zll) | | | | | | | | | |  | | | | | | | | | | | | [5](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.30j0zll) | | | |
|  |  |  |  | |  | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | [**Why this HLD documentation?**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.1fob9te) | | | | | | | | | | | | | | | | | | | | | [5](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.1fob9te) | | |  |
|  |  |  |  | |  | |  |  |  | |  |  |  |  |  |  |  |  |  |  |  |  |  | |  |
| [**1 Description**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.2et92p0) | | | | | | | |  | | | | | | | | | | | | | | [5](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.2et92p0) | | | |
|  |  | |  |  |  | |  |  |  | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | [**1.1**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.tyjcwt) | |  |  | [**Problem Perspective**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.tyjcwt) | | | | | | | | | | | | | | | | | [5](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.tyjcwt) | | |  |
|  |  | |  | |  | |  | |  | |  |  |  |  |  |  |  |  |  |  |  |  |  | |  |
|  | [**1.2 Problem Statement**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.3dy6vkm) | | | | | | | | | | | | | | |  | | | | | | [5](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.3dy6vkm) | | |  |
|  |  | | | | | |  | |  | |  |  |  |  |  |  |  | |  |  |  |  |  | |  |
|  | [**1.3 Proposed Solution**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.1t3h5sf) | | | | | | | | | | | | | | |  | | | | | | [5](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.1t3h5sf) | | |  |
|  |  | | | | | |  | |  | |  |  |  |  |  |  |  | |  |  |  |  |  | |  |
|  | [**1.4 Solution Improvements**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.4d34og8) | | | | | | | | | | | | | | | | | | | | | [5](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.4d34og8) | | |  |
|  |  | | | | | |  | |  | |  |  |  |  |  |  |  | |  |  |  |  |  | |  |
|  | [**1.5 Technical Requirements**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.2s8eyo1) | | | | | | | | | | | | | | | | | | |  | |  | [6](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.2s8eyo1) | |  |
|  | [**1.6 Data Requirements**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.17dp8vu) | | | | | | | | | | | | | | |  | | | | | | [6](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.17dp8vu) | | |  |
|  |  | | | | | |  | |  | |  |  |  |  |  |  |  | | | |  |  |  | |  |
|  | [**1.7 Tools Used**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.3rdcrjn) | | | | | | | | | |  | | | | |  | | | | | | [7](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.3rdcrjn) | | | |
|  |  | |  |  | | |  | |  | |  |  |  |  |  | |  | | | |  |  |  |  | |
|  | [**1.8**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.26in1rg) | |  | [**Constraints**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.26in1rg) | | | | | | | |  | | | | | | | | | | [7](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.26in1rg) | | | |
|  |  | | | | | |  | |  | | |  |  |  |  | |  | | | |  |  |  | | |
|  | [**1.9 Assumptions**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.lnxbz9) | | | | | | | | | | | | | | | | | | | | | [7](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.lnxbz9) | | | |
|  | |  | | | | |  | |  | | |  |  |  |  | |  | | | |  |  |  | | |
| [**2 Design Flow**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.35nkun2) | | | | | | | | |  | | | | | | | | | | | | | [7](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.35nkun2) | | | |
|  |  | | | | | |  | |  | | |  |  | |  | |  | | | |  |  | |  |  |
|  | [**2.1 Modelling Process**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.1ksv4uv) | | | | | | | | | | | | | | | | | | | | | [8](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.1ksv4uv) | | |  |
|  |  | | | | | |  | | | | |  |  | |  | |  | | | |  |  |  | |  |
|  | [**2.2 Deployment Process**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.44sinio) | | | | | | | | | | | | | | | | | | | | | [8](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.44sinio) | | |  |
|  | |  | |  | | |  | | | | |  |  | | | |  | | | |  |  |  | |  |
| **2.3** | | | | **Logging** | | | | | | | | | | | | | | | | | | 8 | | |  |
| **2.2** | | | | **Error Handling** | | | | | | | | | | | | | | | | | | 8 | | |  |
| [**3 Performance Evaluation**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.4i7ojhp) | | | | | | | | | | | | | | | | |  | | | | | [9](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.4i7ojhp) | | |  |
|  | | |  |  | | |  | | | | |  |  | | | |  | | | |  |  |  | |  |
|  | [**3.1**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.2xcytpi) | |  | [**Reusability**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.2xcytpi) | | | | | | | |  | | | | |  | | | | | [9](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.2xcytpi) | | | |
|  |  | |  |  | | |  | | | | |  |  | | | |  | | | |  |  |  | |  |
|  | [**3.2**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.1ci93xb) | |  | [**Application Compatibility**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.1ci93xb) | | | | | | | | | | | | | | | | | | [9](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.1ci93xb) | | |  |
|  | [**3.3**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.3whwml4) | |  | [**Resource Utilization**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.3whwml4) | | | | | | | | | | | | |  | | | |  |  | [9](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.3whwml4) | |  |
|  |  | | | | | |  | | | | | |  | | | |  | | | | |  |  | |  |
|  | [**3.2 Deployment**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.2bn6wsx) | | | | | | | | | | | | | | | |  | | | | | [9](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.2bn6wsx) | | | |
|  |  | | | | | |  | | | | | |  | | | | | | | | |  |  | | |
| [**Conclusion**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.qsh70q) | | | | | | | | | | | | | | | | | | | | | | [9](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.qsh70q) | | | |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

3

High Level Design (HLD)

**Abstract**

This high-level design (HLD) abstract outlines a project aimed at providing individuals with personalized estimates of their healthcare expenses based on their individual health situations. The goal is to empower customers to make informed decisions when selecting health insurance plans by considering projected costs derived from our study. By predicting a person's future medical expenses using their past records, including age, sex, BMI, number of children, smoking status, and region of residence, individuals can focus on the health aspects of an insurance policy rather than being burdened by ineffective coverage. The HLD emphasizes the need for efficient data collection and preprocessing, feature engineering, model development, and deployment. The system aims to leverage advanced machine learning techniques to accurately estimate future medical expenses for individuals, providing them with valuable insights to guide their health insurance decisions and ensure appropriate coverage for their needs.

4

High Level Design (HLD)

**1 Introduction**

**1.1 Why this High-Level Design Document?**

The main purpose of this HLD documentation is to feature the required details of the project and supply the outline of the machine learning model and also the written code. This additionally provides the careful description on however the complete project has been designed end-to-end.

**1.2 Description**

**Problem Perspective**

Predicting insurance price is a complex task that involves analyzing various factors related to an individual's age, sex, bmi, number of children they have, and other relevant variables. While I can provide you with some general information, please note that predicting income accurately for specific individuals can be challenging due to the wide range of factors involved and the inherent uncertainty of future outcomes.

**1.3 Problem Statement**

The aim of this project is to develop a system that provides individuals with personalized estimates of their required health insurance coverage based on their individual health situations. The problem at hand is that individuals often struggle to determine the appropriate level of coverage they need, resulting in ineffective or inadequate health insurance plans. By leveraging past records, including age, sex, BMI, number of children, smoking status, and region of residence, the system must accurately predict a person's future medical expenses.

The main problem lies in the lack of accessible tools for individuals to estimate their healthcare costs, leading to uninformed decisions and potential financial burden. Additionally, the complex nature of healthcare expenses necessitates the utilization of advanced data analytics and machine learning techniques to derive accurate predictions.

By providing individuals with a reliable estimate of their healthcare expenses, the proposed solution aims to assist customers in navigating various health insurance carriers and their plans and perks, enabling them to make informed decisions. This empowers individuals to focus on the health-related aspects of an insurance policy rather than getting entangled in ineffective or insufficient coverage. Ultimately, the problem statement revolves around developing a predictive model that utilizes historical data to estimate future medical expenses, offering individuals valuable insights to guide their health insurance choices and ensuring they receive appropriate coverage tailored to their specific needs.

**1.4. Project Solution**

Project requires the desired input of user from the created interface and method all the provided information to satisfy the wants of the machine learning model and at last show the expected output.

5

High Level Design (HLD)

**1.6 Technical needs**

There are not any hardware needs needed for victimization this application, the user should have an interactive device that has access to the web and should have the fundamental understanding of providing the input. And for the backend half the server should run all the package that's needed for the process and provided information to show the results.

**1.7 Information needs**

The info demand is totally supported the matter statement. and also, the information set is accessible on the Kaggle within the type of standout sheet(.xlsx), because the main theme of the project is to induce the expertise of real time issues, we have a tendency to once more mercantilism {the information into the prophetess data base and commerce it into csv format.

**1.8Tools Used**

* Python 3.7 is employed because the programming language and frame works like numpy, pandas, sklearn and alternative modules for building the model.
* VScode is employed as IDE.
* For visualizations seaborn and components of matplotlib are getting used.

6

High Level Design (HLD)

* For information assortment prophetess info is getting used.
* Front end development is completed victimization HTML/CSS.
* Flask is employed for each information and backend readying.
* GitHub is employed for version management.
* NETLIFY is employed for deployment.

**1.9 Constraints**

The objective of this project is to develop a predictive model that can accurately estimate the expense level of people based on various sex, age, bmi, number of children,smoking habits and region. The model should analyze a given individual's characteristics and provide a prediction of how their medical expenses be in the future.

**1.10 Assumptions**

The most objective of the project is to implement the utilization cases as for the new dataset that user provides through the programme. Machine learning model is employed for process the on top of computer file. It's additionally assumed that each one aspects of this project have the flexibility to figure along within the approach as the designer is expecting.



7

High Level Design (HLD)

**2.1 and 2.2 Design Flow and Deployment Process**



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



**2.3 Logging**

Each step is being logged within the system that runs internally, that shows the date time and therefore the processed that has been performed, work is completed in several layers as information, DEBUG, ERROR, WARNINGS. this provides US the perceive of the logged info.

**2.4 Error Handling**

Once ever a slip is occurred, the reason are logged in its several log file, in order that the developer will rectify the error.

High Level Design (HLD)

**3 Performance analysis**

**3.1 Reusability**

Elements of the code written is accustomed different applications and therefore the rest is changed and be reused.

**3.2 Application Compatibility**

The various parts for this project are exploitation python as associate interface between them. every element can have its own tasks to perform, and it's the work of the python to make sure correct transfer of data.

**3.3 Resource Utilization**

Once any task is performed, it'll doubtless; use all the process power offered till that performs is finished.

**3.4 Deployment**

The model is being deployed on AWS. Here’s the link:

<http://insurancedeploy-env.eba-zvcbn3vk.us-east-1.elasticbeanstalk.com/>

**Conclusion**

In conclusion, this project provides personalized estimates of health insurance coverage based on individual health situations and past medical expenses. By accurately predicting future medical expenses using historical data, individuals can make informed decisions. It enables customers to work with any health insurance carrier while considering the projected cost from our study. This empowers individuals to focus on the health side of an insurance policy, ensuring effective coverage. Advanced data analytics and machine learning techniques are employed for accurate predictions. The project aims to improve decision-making, prioritize health needs, and enhance financial security.

9