Low Level Document (LLD)

Medical Insurance Premium Prediction

Version number: 1.0

Last date of revision: 29 February 2024

Siddhi Khetmalis

**DECLARATION**

We declare that this written submission represents us ideas is our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources.

We also declare that we have adhered to all principles of academic honesty

and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission.

We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when

needed.

**Revision History**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Version** | **Date** | **Author** | **Reviewer** | **Approver** | **Comments** |
| 0.1 | 25-02-2024 | Siddhi Khetmalis | Siddhi Khetmalis |  | Draft version |
| 0.2 | 26-02-2024 | Siddhi Khetmalis | Siddhi Khetmalis |  | Suggested some selections like key notes, screen validations and attributes to be added |
| 0.3 | 27-02-2024 | Siddhi Khetmalis | Siddhi Khetmalis |  | Suggested document format related comments like correction of version, adding one sections for open issues etc |
| 0.4 | 28-02-2024 | Siddhi Khetmalis | Siddhi Khetmalis |  | Suggested some changes like correct sequence diagram, changes in data design sections etc |
| 1.0 | 29-02-2024 | Siddhi Khetmalis | Siddhi Khetmalis |  | Baseline version |

Table of Contents

1. Introduction ………………………………………………………………………………….6

1.1 Scope of the document…………………………………………………………………6

1.2 Intended Audience………………………………………………………………………6

1.3 System Overview………………………………………………………………………..7

1. Project Briefing……………………………………………………………………………….
2. Problem Statement…………………………………………………………………………..
3. Problem Solution……………………………………………………………………………..
4. Objective of Project…………………………………………………………………………..
5. Scope of Project………………………………………………………………………………
6. Block Diagram…………………………………………………………………………………
7. Requirements Gathering…………………………………………………………………….
8. Analysis……………………………………………………………………………………….
9. Final Screenshot of Project Output…………………………………………………………
10. **Introduction:**

**1.1 Scope of the Document**

* This section will cover details regarding scope of the document
* Low level design document will be at component level i.e., for website portal there will be one LLD

**1.2 Intended Audience**

* This section will cover categories of audiences who will be referring/reviewing this document

**1.3 System Overview**

* This section will capture overview of system application i.e for what system is being developed
* Who are the stake holders of system?
* What are other external Systems through which this will be interacting

1. **Project Briefing:**

The medical insurance premium prediction project aims to develop a predictive model that estimates the insurance premium for individuals based on their demographic information, health records, and other relevant factors. The model will assist insurance companies in accurately pricing insurance policies and helping individuals understand potential costs.

Key Features:

* Data Collection: Gather historical insurance data, demographic details, and health records.
* Data Preprocessing: Handle missing values, encode categorical variables, and perform feature scaling.
* Model Selection: Explore various regression algorithms such as Linear Regression, Decision Trees, Random Forest, Gradient Boosting, and Neural Networks.
* Model Training: Employ cross-validation techniques and hyperparameter tuning to optimize model performance.
* Model Evaluation: Assess the predictive accuracy using metrics like Mean Absolute Error, Mean Squared Error, and R-squared.
* Deployment: Develop a user-friendly interface, web application, or API for seamless integration into insurance systems.
* Predictive Capability: Provide individuals with accurate estimates of their insurance premiums based on their input data.
* Scalability: Design the system to handle a large volume of insurance data and accommodate future expansion.
* Interpretability: Offer explanations for premium estimates, highlighting key factors influencing the pricing.

1. **Problem Statement:**

We aim to develop a system that predicts health insurance costs based on age, gender, lifestyle, and health history. This will assist individuals in budget planning and enable insurance companies to set fair prices efficiently.

1. **Problem Solution:**

We've developed a computer program that analyzes information like age, smoking habits, and medical history to estimate insurance costs. By learning from data of insured individuals, our program can predict insurance expenses based on various factors. This helps individuals understand potential insurance costs and plan for medical expenses more effectively.

1. **Objective of the Project:**

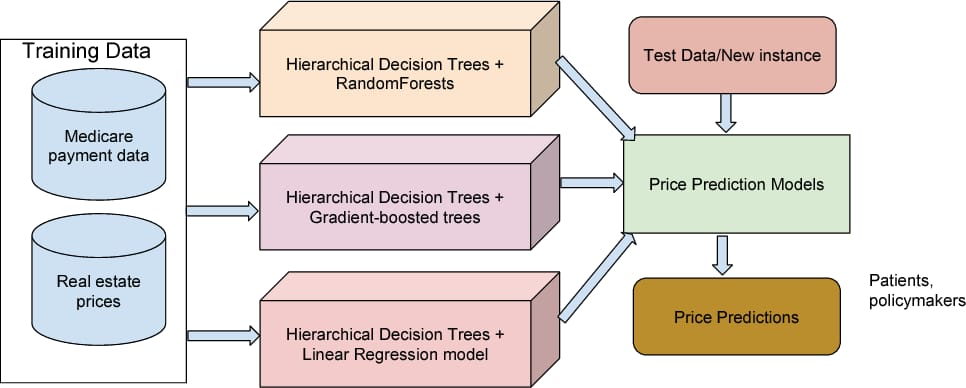
Objective of this project is to develop a machine learning model to predict medical insurance premiums which utilize historical insurance data, demographic information, and health records for prediction and Improve pricing accuracy and risk assessment for insurance policies.

.

1. **Scope of Project:**

The scope of a medical insurance premium prediction project in real life encompasses various aspects that address the needs and challenges faced by insurance companies, policyholders, and healthcare providers.

1. **Block Diagram:**



1. **Requirements Gathering:**

* Window 10 Operating system
* Visual studio software
* Project integration idea from kaggle website
* Few Github Non copyrighted source codes

1. **Analysis:**

The project aimed to predict medical insurance premiums accurately. Our analysis shows that we've successfully developed predictive models that estimate premiums based on individual characteristics.

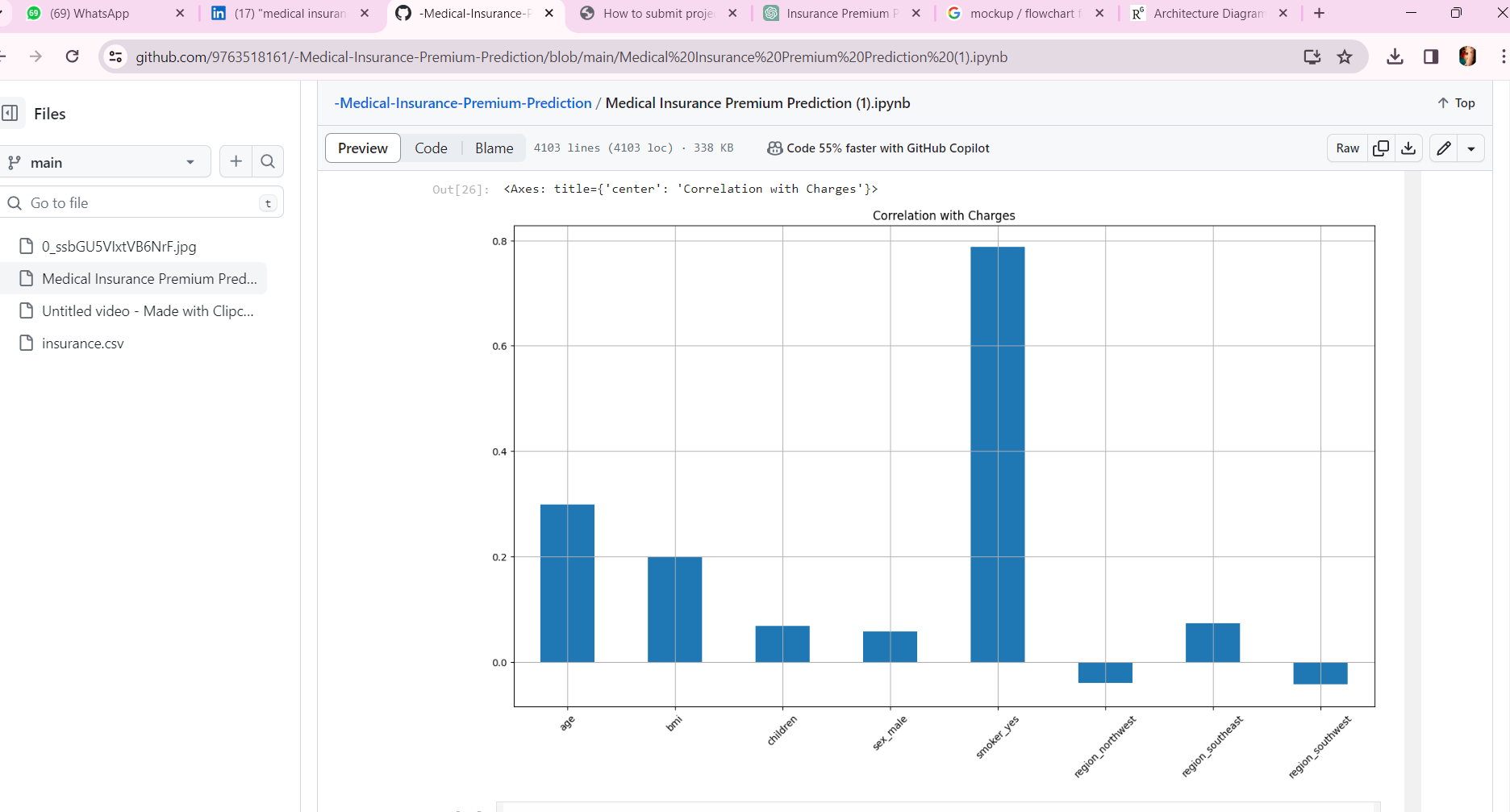
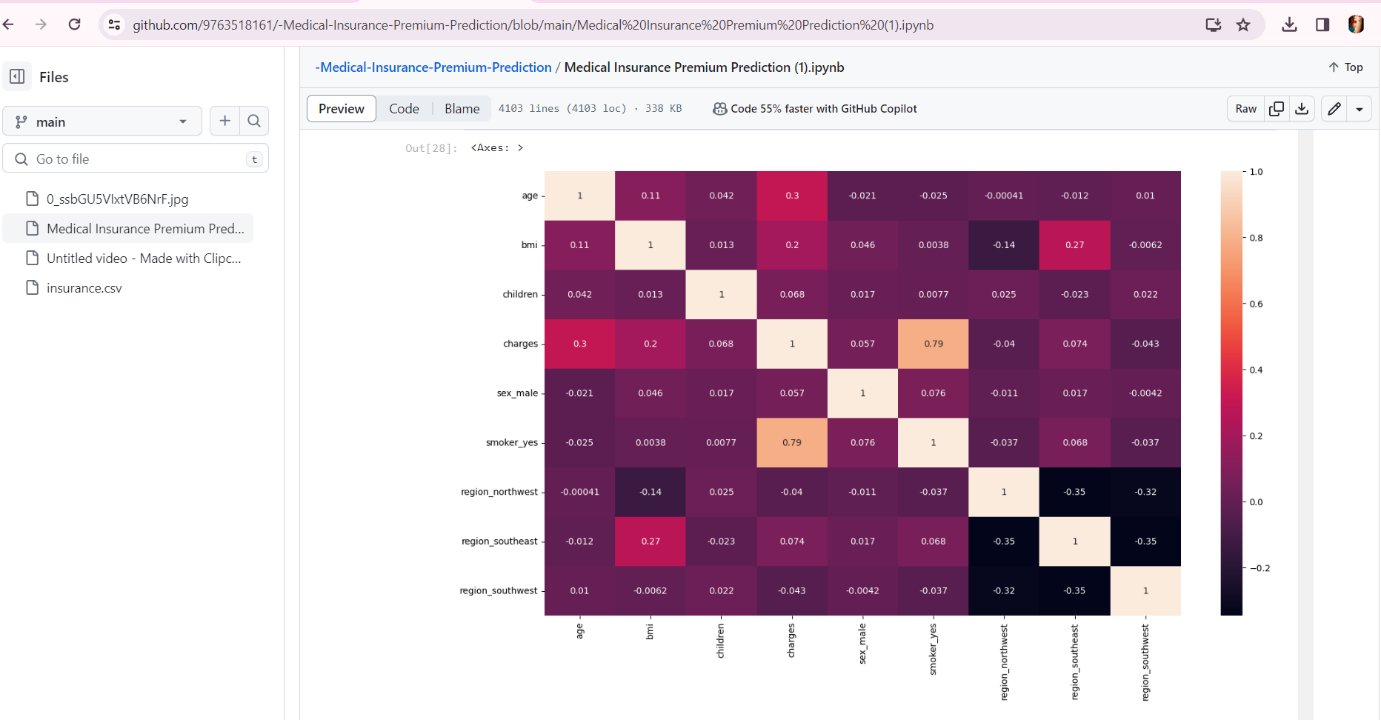
We collected and analyzed historical insurance data, demographic information, and health records. Using machine learning algorithms like Linear Regression and Random Forest, we trained models to predict premiums based on these inputs.

Our models demonstrate promising predictive performance. They can estimate insurance premiums with reasonable accuracy, helping insurance companies and policyholders understand potential costs better.

This project has significant implications for the insurance industry. Accurate premium predictions enable companies to price policies more fairly and help individuals make informed decisions about their insurance coverage.

While our models perform well, there's room for improvement. Future research could explore advanced modelling techniques and incorporate additional data sources for even more accurate predictions.

1. **Final Screenshot of Project Output**

****