VIEHGROUP

Low Level Document (LLD)
Medical Insurance Prediction

Version number: 1.0
Last date of revision: 05 May 2024

Team Members:

Nandana B
Ashwin Sabu
Nambiar Krishnendu Radhakrishnan
Rithin Chand V
M Reshma

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.

VIEHGROUP

Revision History

Version	Date	Author	Reviewer	Approver	Comments
0.1	26-04-2024	M Reshma	Nandana B		Draft version
0.2	27-04-2024	Nambiar Krishnendu Radhakrishnan	Nandana B		Suggested some selections like key notes, screen validations and attributes to be added
0.3	28-04-2024	Ashwin Sabu	Nandana B		Suggested document format related comments like correction of version, adding one sections for open issues etc
0.4	03-05-2024	Rithin Chand	M Reshma		Suggested some changes like correct sequence diagram, changes in data design sections etc
1.0	10-05-2024	Nandana B	M Reshma		Baseline version

VIEH GROUP

Table of Contents

VIEHGROUP

1. 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

- Categories of audiences who will refer to/review 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 interacti

2. Project Briefing:

The medical insurance premium prediction project aims to develop a machine learning model capable of accurately predicting insurance premiums for individuals. This project is crucial for both insurance companies and individuals seeking insurance coverage. By leveraging demographic and health-related attributes, the model aims to provide personalized and precise premium predictions.

3. Problem Statement:

In the healthcare industry, accurately predicting insurance premiums is crucial for both insurance companies and individuals.

Traditional methods of calculating premiums may not take into account individual health factors, leading to inaccuracies.

The project addresses this problem by developing a machine learning model that leverages demographic and health-related attributes to predict premiums more accurately.

VIEHGROUP

4. Problem Solution:

The proposed solution involves developing a machine learning model that takes into account demographic and health-related attributes to predict insurance premiums. By analyzing these attributes, the model can provide more personalized and accurate premium predictions for individuals.
The user-friendly interface allows individuals to input their information easily and obtain predicted insurance premiums quickly.

5. Objective of the Project:

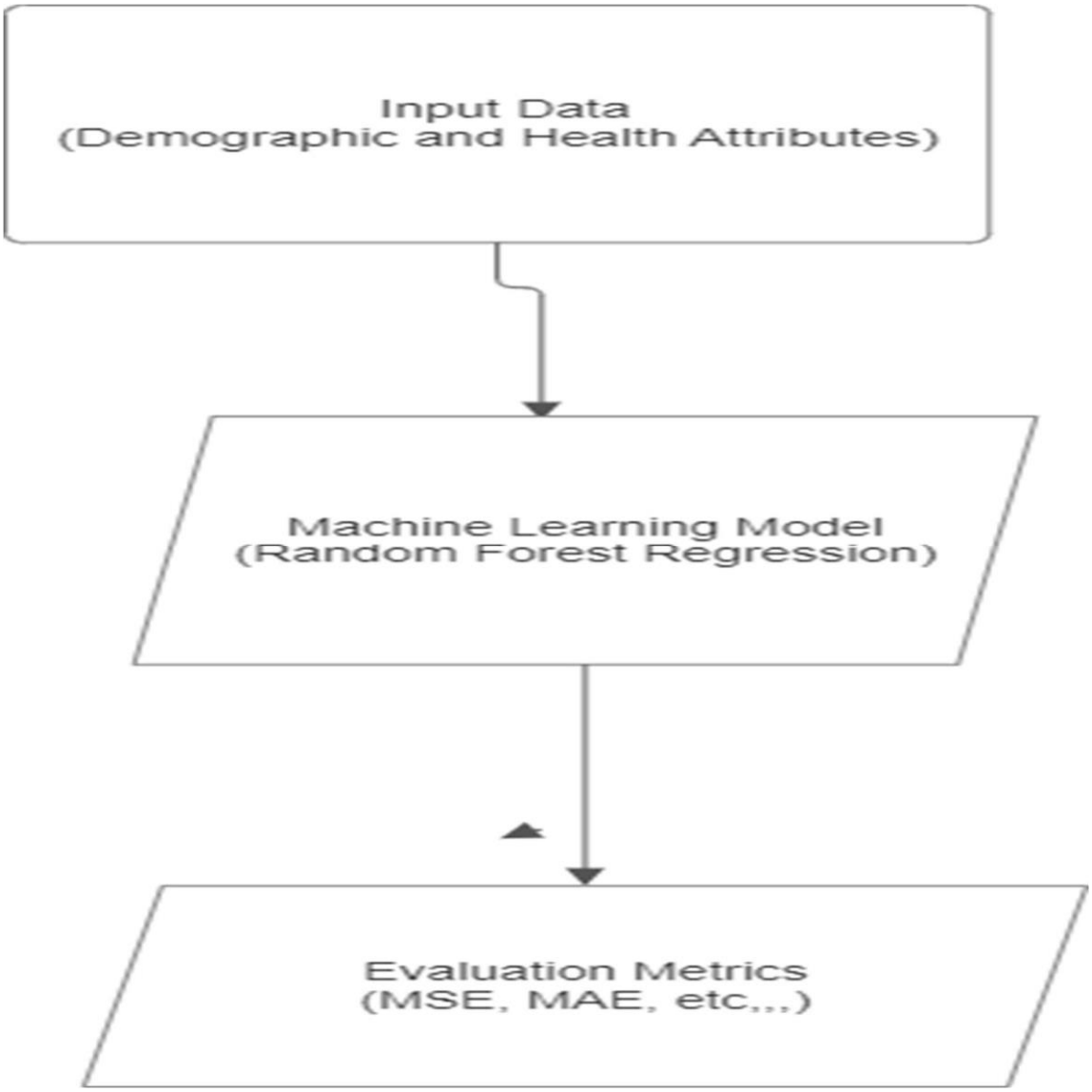
Develop a machine learning model for predicting medical insurance premiums. Utilize demographic attributes such as age, gender, and location, along with health-related attributes including BMI, smoking status, and pre-existing conditions. Provide accurate and personalized insurance premium predictions for individuals. Enhance the efficiency of premium calculations for insurance companies. Create a user-friendly interface for individuals to input their information and obtain predicted insurance premiums.

6. Scope of Project:

The scope of the project encompasses several key aspects:

Data Collection: Gather demographic and health-related data from individuals to build the training dataset.
Data Preprocessing: Clean and preprocess the data to handle missing values, encode categorical variables, and normalize numerical features.
Model Development: Explore various machine learning algorithms suitable for regression tasks and train the model on the preprocessed dataset. Model Evaluation: Assess the performance of the trained model using appropriate evaluation metrics such as mean squared error (MSE) or Rsquared.
Interface Design: Develop a user-friendly interface where individuals can input their demographic and health-related information and obtain predicted insurance premiums.
Integration: Integrate the machine learning model with the user interface to provide seamless prediction capabilities.
Testing and Validation: Conduct thorough testing and validation of the entire system to ensure accuracy, reliability, and usability.

7. Block Diagram:



8. Requirements Gathering:

- Window 10 Operating system
- Visual studio software
- 5 Team members for the research part
- Project integration idea from IEEE website
- Few Github Non copyrighted source codes

9. Analysis:

Analysis for Medical Insurance Prediction System

In the medical insurance prediction system, we are not utilizing bioinformatics algorithms like BLAST-SSAHA hybridization. Instead, we are employing machine learning techniques to predict insurance premiums based on demographic and health-related attributes. Therefore, the analysis for the medical insurance prediction system will focus on the data, modeling techniques, and evaluation metrics relevant to machine learning.

1. Data Collection and Preprocessing:

Data Collection: Gather demographic information such as age, gender, and location, as well as health-related attributes like BMI, smoking status, and pre-existing conditions from reliable sources or user input forms.
Data Preprocessing:
Clean the data by handling missing values, encoding categorical variables, and normalizing numerical features to prepare the dataset for model training.

2. Model Selection and Training:

Model Selection: Experiment with various regression algorithms suitable for predicting insurance premiums, such as linear regression, decision trees, random forests, or gradient boosting models.
Model Training: Train the selected model(s) on the preprocessed dataset using techniques like cross-validation and hyperparameter tuning to optimize performance.

3. Evaluation Metrics:

Mean Squared Error (MSE): Measures the average squared difference between the predicted and actual insurance premiums. Lower MSE indicates better model performance.
Mean Absolute Error (MAE): Measures the average absolute difference between the predicted and actual insurance premiums. Lower MAE indicates better model performance.
R-squared (R2): Measures the proportion of the variance in the insurance premiums that is predictable by the model. Higher R2 values indicate better model fit.

4. User Interface:

Design and implement a user-friendly interface using frameworks like Streamlit or Flask to allow individuals to input their demographic and health-related information. Ensure that the interface provides clear instructions and feedback to users and can handle various input formats and data validations.

5. Integration and Deployment:

Integrate the trained model with the user interface to enable seamless prediction functionality.
Deploy the system to a suitable platform, such as a web application or mobile app, to make it accessible to users.

6. Testing and Validation:

Conduct thorough testing and validation of the entire system to ensure its accuracy, reliability, and usability.
Validate the predictions against real-world insurance premiums to verify the model's performance and utility.

Conclusion:

The medical insurance prediction system utilizes machine learning techniques to accurately predict insurance premiums based on demographic and health-related attributes. By providing personalized premium predictions and a user-friendly interface, the system aims to enhance the efficiency and accuracy of insurance premium calculations for both insurance companies and individuals.

VIEHGROUP

10. Final Screenshot of Project Output: