

# Low Level Design (LLD)

## Insurance Premium Predictor

Written By	Vibhu Manikpuri
Document Version	0.1
Last Revised Date	08 September, 2021

# Document Control

## Change Record:

Version	Date	Author	Comments
0.1	08 September, 2021	Vibhu Manikpuri	Introduction & Architecture Description Defined
0.2	09 September, 2021	Vibhu Manikpuri	Document Control and Version Control added

## Reviews:

Version	Date	Reviewer	Comments
0.2	09 September, 2021	Vibhu Manikpuri	Document Control and Version Control added

## Approval Status:

Version	Review Date	Reviewed By	Approved By	Comments

# Contents

Introduction .....	4
1.1 What is Low-Level Design Document? .....	4
1.2 Scope.....	4
Architecture .....	5
3.1 Data Description .....	6
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.....	6
3.2 Data Transformation.....	6
3.3 Data Preprocessing .....	6
3.4 Machine Learning Algorithms.....	6
3.5 Evaluation Metrics .....	6
3.6 Model Building .....	6
3.7 Data from User.....	6
3.8 Data Validation .....	7
3.9 User data feeding into Machine Learning model .....	7
3.10 Insurance Premium Prediction .....	7
3.11 Deployment .....	7

# Introduction

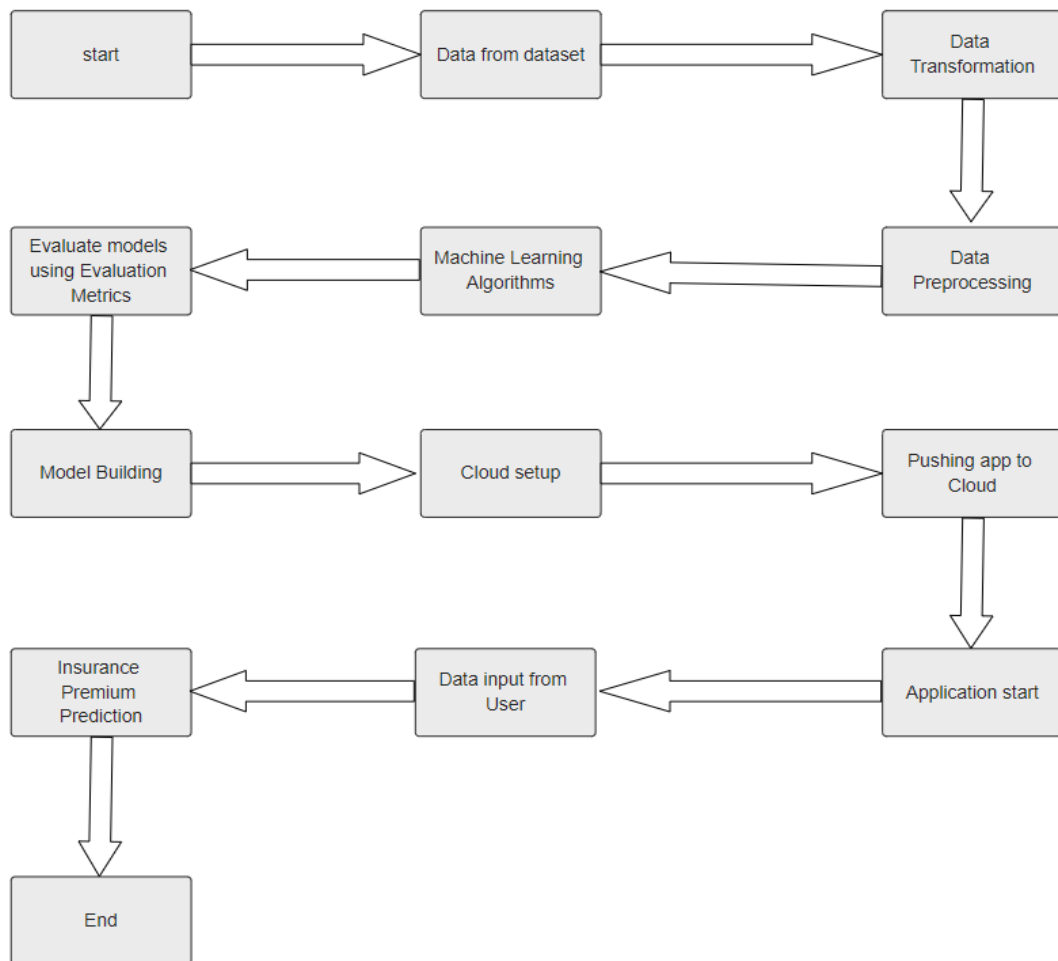
## 1.1 What is Low-Level Design Document?

The goal of LLD or a low-level design document (LLDD) is to give the internal logical design of the actual program code for Insurance Premium Predictor. LLD describes the class diagrams with the methods and relations between classes and program specs. It describes the modules so that the programmer can directly code the program from the document.

## 1.2 Scope

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



### 3.1 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

In the Transformation Process, we will convert the categorical features into numerical features.

### 3.3 Data Preprocessing

In data preprocessing steps we rid the dataset of Null values.

### 3.4 Machine Learning Algorithms

We predict the values using Machine Learning algorithms including Linear Regression, Polynomial Regression, Ridge Regression, Decision Tree Regression, Random Forest Regressor.

### 3.5 Evaluation Metrics

We check the performance of each Machine Learning model using the standard Evaluation Metrics: R2 score, k-fold Cross Validation and Root Mean Square Error value.

### 3.6 Model Building

The machine learning model is built using the best performing algorithm and training it with the entire dataset. The model is serialized (pickled).

### 3.7 Data from User

Here we will collect data from user by prompting for numerical features and giving options for categorical features.

### 3.8 Data Validation

Here Data Validation is done, given by user.

### 3.9 User data feeding into Machine Learning model

The data collected from the User is input into the Machine Learning model.

### 3.10 Insurance Premium Prediction

The machine learning model operates on the input data and provides a prediction for that data.

### 3.11 Deployment

We will be deploying the model to Heroku.