Low Level Design (LLD)

FLIGHT FARE PREDICTION

Revision Number – 1.0

Last Date of Revision – 16-05-2023

**SHAHIN ANJUM**

**K M SUMANTH**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **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 LLD 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 Architecture**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.3znysh7) | | | | | | | | | | | | | | | | [**3**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.3znysh7) | | | |
|  |  | |  |  |  |  |  |  |  |  |  |  | |  |  |  |
| [**2 Architecture Design**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.3dy6vkm) | | | | | | | |  | | | | | | | | [**6**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.3dy6vkm) | | | |
|  |  | | |  |  |  |  |  |  |  |  |  | |  |  |  |
|  | [**2.1 Data Gathering from Main Source**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.1t3h5sf) | | | | | | | | | | | | | | | [**6**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.1t3h5sf) | | |
|  |  | | |  |  |  |  | |  |  |  |  | |  |  |  |
|  | [**2.2 Tools Used**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.17dp8vu) | | | |  | | | | | | | | | | |  | |
|  |  | | |  |  |  |  | |  |  |  |  | |  |  |  |
|  | [**2.3 Data Description**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.3rdcrjn) | | | | | | | | | | | | | | | [**7**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.3rdcrjn) | | |
|  |  | | |  |  |  |  | |  |  |  |  | |  |  |  |
|  | [**2.4 Import Data into Database**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.26in1rg) | | | | | | | | | | | | |  | | [**7**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.26in1rg) | | |
|  |  | | |  |  |  |  | | |  |  |  | |  |  |  |
|  | [**2.5 Export Data from Database**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.lnxbz9) | | | | | | | | | | | | | |  | [**8**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.lnxbz9) | | |
|  |  | | |  |  |  |  | | |  |  |  | | |  |  |
|  | [**2.6 Data Pre-Processing**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.35nkun2) | | | | | | | | | |  | | | |  | [**8**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.35nkun2) | | |
|  |  | | |  |  |  |  | | |  |  |  | | | |  |
|  | [**2.7 Modelling**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.1ksv4uv) | | | | | | | | | |  | | | | | [**8**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.1ksv4uv) | | | |
|  |  | | |  |  |  |  | | |  | |  | | | |  |
|  | [**2.8 UI Integration**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.44sinio) | | | | |  | | | | | | | | | | [**8**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.44sinio) | | | |
|  |  | | | |  |  |  | | |  | |  | | | |  |
|  | [**2.9 Data From User**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.2jxsxqh) | | | | | |  | | | | | | | | | [**8**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.2jxsxqh) | | | |
|  |  | | | |  | |  | | |  | |  | | | |  |
|  | [**2.10 Data Validation**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.z337ya) | | | | | | | | |  | | | | | | [**8**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.z337ya) | | |
|  |  | | | |  | |  | | |  | |  | | | |  |
|  | [**2.11 Rendering the Results**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.3j2qqm3) | | | | | | | | | | |  | | | | [**8**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.3j2qqm3) | | |
|  |  | | | |  | |  | | | | |  | | | |  |
|  | [**3 Deployment**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.1y810tw) | | | |  | | | | | | |  | | | |  |
|  | [**3.1 Unit Test Cases**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.4i7ojhp) | | | | | |  | | | | | | | | | [**9**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.4i7ojhp) | | | |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

**Abstract**

Buildings are responsible for 40% of energy consumption and 33% of greenhouse gas emissions. We should design buildings with improve energy efficiency that way we can contribute to the environment.

When it comes to efficient building design, the computation of the heating load and the cooling load is required to determine the specifications of the heating and cooling equipment needed to maintain comfortable indoor air conditions. To estimate the required cooling and heating capacities, architects and building designers need information about the characteristics of the building and of the conditioned space (for example occupancy and activity level). For this reason, we will investigate the effect of eight input variables: relative compactness, surface area, wall area, roof area, overall height, orientation, glazing area, and glazing area distribution, to determine the output variables heating load and cooling load of residential buildings.

**Introduction**

**Why this Low-Level Design Document?**

The main purpose of this LLD 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 Architecture**

A picture containing text, screenshot, font, diagram

Description automatically generated

2 Architecture design

This project is to create an interface for the user to know the heating cooling amount, in addition to this, in need of getting the real time project experience we are importing the gathered data into our own database and then start the project from the scratch.



2.1 Data gathering from main source

The data for the current project is being gathered from UCI Machine Learning Repository , the link to the data is:

https://archive.ics.uci.edu/ml/datasets/energy+efficiency#

**2.2 Tools Used**

Python 3.8 is employed because the programming language and frame works like numpy, pandas, sklearn and alternative modules for building the model.

* For visualizations seaborn and components of matplotlib are getting used
* Vscode is employed as IDE(integrated development environment).
* 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.

2.3 Data description

This data is created by studying 12 different buildings shape simulated in ecolect. The dataset comprises 768 samples and 8 features, aiming to predict two real valued responses. It can also be used as a multi-class classification problem if the response is rounded to the nearest integer.

The dataset contains eight attributes (or features, denoted by X1...X8) and two responses (or outcomes, denoted by y1 and y2). The aim is to use the eight features to predict each of the two responses.

Specifically:

* X1 Relative Compactness
* X2 Surface Area
* X3 Wall Area
* X4 Roof Area
* X5 Overall Height
* X6 Orientation
* X7 Glazing Area
* X8 Glazing Area Distribution
* y1 Heating Load
* y2 Cooling Load

2.4 Import data into Cassandra

Created an api for the upload of the data into the Cassandra database, steps performed are:

* Connection is made with the database.
* Created a database with name Energy Efficiency.
* Cqlsh command is written for creating the data table with required parameters.
* And finally, a cqlsh command is written for uploading the dataset into data table by bulk insertion.

2.5 Export data from database

In the above created api, the download url is also being created, which downloads the data into a csv file format.

2.6 Data pre-processing

Steps performed in pre-processing are:

* First rename all columns
* The data types are being checked and found all columns are of type integer.
* Checked for null values as there are no null values.
* Checking outliers as looking box plot we don’t need to worry about outliers.
* Scaling is performed for required data.

And the data is ready for passing to the machine learning algorithm

2.7 Modelling

The pre-processed data is then visualized, and all the required insights are being drawn. Although from the drawn insights, the data is randomly spread but still modelling is performed with different machine learning algorithms to make sure we cover all the possibilities. And finally, as expected random forest regression performed well.

2.8 UI integration

Both CSS and HTML files are being created and are being integrated with the created machine learning model. All the required files are then integrated to the app.py file and tested locally. Note I copy the code from internet and make the CSS and HTML File .

2.9 Data from user

The data from the user is retrieved from the created HTML web page.

2.10 Data validation

The data provided by the user is then being processed by app.py file and validated. The validated data is then sent for the prediction.

2.11 Rendering the results

The data sent for the prediction is then rendered to the web page.

3 Deployment

I didn’t deploy.