Low Level Design (LLD)

Flight Fare Prediction

Revision Number: 1.2

Last date of revision: 02 October 2022

Adil Anwar

# Document Version Control

|  |  |  |  |
| --- | --- | --- | --- |
| Date Issued | Version | Description | Author |
| 05 – 09 - 2022 | 1.0 | Abstract  Introduction  Architecture | Adil Anwar |
| 21 – 09 - 2022 | 1.1 | Architectural Design | Adil Anwar |
| 02 – 10 - 2022 | 1.2 | Deployment Unit Test Cases | Adil Anwar |

**Contents**

[Document Version Control 2](#_Toc92035717)

[1 Introduction 3](#_Toc92035718)

[1.1 Why this Low-Level Design Document? 3](#_Toc92035719)

[1.2 Scope 3](#_Toc92035720)

[2 Architecture 5](#_Toc92035724)

[3.](#_Toc92035725) Architecture Design 5

[3.1 Data Collection 5](#_Toc92035726)

[3.2 Data Descritption 5](#_Toc92035727)

[3.3 Importing data into Database 5](#_Toc92035728)

[3.4 Exporting Data from Database 5](#_Toc92035729)

[3.5 Data Pre-processing 6](#_Toc92035726)

[3.6 Model Creation 6](#_Toc92035727)

[3.7 Data from User 7](#_Toc92035728)

[3.8 Data Validation 7](#_Toc92035729)

3.9 Rendering the Results 7

[4. Technology stack 7](#_Toc92035730)

4.1 Unit Testing 8

Abstract

The recent changes in the international market had a large impact on the Aviation sector because of the several reasons. These impact the two class folks, the first is Business perspective and second is Customer perspective. The major reason of such impact is the governments around the world amended totally different rules to their various Airline firms. Taking of these factors in thought the value of the flight tickets has vary from one place to another. Booking a flight ticket its price tag has split into two, one is online bookings and other is offline bookings. Each of these have their various criteria for value of the price, one such example is that the server load and therefore the range of booking requests. During this machine learning implementation, we are going to see numerous factors that impact the price of the flight ticket and predict the acceptable price of the ticket.

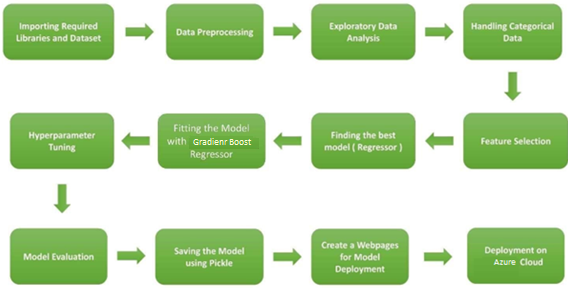
# Introduction

## Why this Low-Level Design Document?

The main goal of the LLD document is to give the internal logic design of actual code implementation and supply the outline of the machine learning model and its implementation. Additionally, it provides the description how our project will designed end - to - end. The main objective of the project is to predict if a person is having compensated hypothyroid, primary hypothyroid, secondary hypothyroid or negative(no thyroid).

* 1. 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



2. Architecture



3. Architecture Design

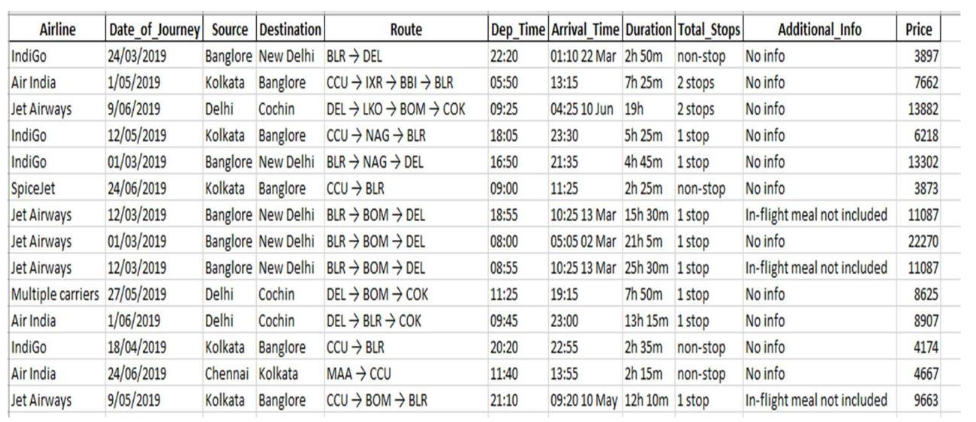
This project is designed to make an interface for the User to predict the its approximate flight ticket price.

3.1 Data Collection

The data for this project is collected from the Kaggle Dataset, the URL for the dataset is: kaggle.com/datasets/nikhilmittal/flight-fareprediction-mh

3.2 Data Description

Flight Fare Prediction is 10K+ dataset publicly available on the Kaggle. The information in the dataset is present in two separated excel files named as train.xlsx and test.xlsx. Dataset contains 10683 rows which shows the information such Date of Journey, Source, Destination, Arrival Time, Departure Time, Total stops, Airlines, Additional Info and Price. The glance of the Dataset is :



3.3 Importing data into Database

Created associate API for the transfer of the info into the Cassandra info, steps performed are :

• Connection is created with the info.

• Created a info with name FlightInfo.

• cqlsh command is written for making the info table with needed parameters.

• And finally, a cqlsh command is written for uploading the Knowledge Set into data

table by bulk insertion.

3.4 Exporting Data from Database

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

3.5 Data Pre-processing

• Checked for info of the Dataset, to verify the correct datatype of the Columns.

• Checked for Null values, because the null values can affect the accuracy of the

model.

• Converted all the desired columns into Datetime format.

• Performed One – Hot encoding on the desired columns.

• Checking the distribution of the columns to interpret its importance.

• Now, the info is prepared to train a Machine Learning Model.

3.6 Model Creation

The Pre - processed info is now envisioned and drawn insights helps us to select the feature that improves the accuracy of the model. The info is randomly used for modelling with different machine learning algorithms to create a model to predict the Flight ticket price. After performing on different algorithms, we use Random Forest Regression to create a model and then also perform Hyperparameter Tuning to improve the accuracy of the model.

3.7 Data from User

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

3.8 Data Validation

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

3.9 Rendering the Results

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

4. Deployment

The tested model is then deployed to Azure. So, users can access the project from any internet devices.

## 4.1 Unit Test Cases

|  |  |  |
| --- | --- | --- |
| **Test Case Description** | **Pre - Requisites** | **Expected Results** |
| Verify whether the  Webpage is accessible to  the User or not. | Webpage URL should be  defined | Webpage should be  accessible to the User |
| Verify whether the  Webpage is completely  loads for the User or not | 1. Webpage URL is  accessible.  2. Webpage is  deployed. | The Webpage should be  completely loads for the  User when it is accessed. |
| Verify whether the User is  able to enter data in input  fields or not. | 1. Webpage URL is  accessible.  2. Webpage is  deployed.  3. Webpage  input fields are editable. | The User is able to enter  data in input fields. |
| Verify whether the User is  able to submit details or  not. | 1. Webpage URL is  accessible.  2. Webpage is  deployed.  3. Webpage  input fields are editable. | The User is able to submit  details to process. |
| Verify whether the User  gets recommended  results on submitting the  details or not. | 1. Webpage URL is  accessible.  2. Webpage is  deployed.  3. Webpage  input fields are editable. | The User gets  recommended results on  submitting the details. |