Low Level Design (LLD)

FLIGHT FARE PREDICTION

Revision Number – 1.0

Last Date of Revision – 14/12/2022

ANWESHA DAS

Table of Contents

[ABSTRACT 3](#_Toc122127651)

[**1 Introduction** 3](#_Toc122127652)

[1.1 Why this Low-Level Design Document? 3](#_Toc122127653)

[1.2 Architecture 3](#_Toc122127654)

[2. Architecture Design 4](#_Toc122127655)

[**2.1. Data Gathering** 4](#_Toc122127656)

[2.2. Tool Used 4](#_Toc122127657)

[2.3 Data Description 5](#_Toc122127658)

[2.6 Data Pre-processing 5](#_Toc122127659)

[2.7 Modeling 5](#_Toc122127660)

# ABSTRACT

Travelling through flights has become an integral part of today’s lifestyle as more and more people are opting for faster travelling options. The flight ticket prices increase or decrease every now and then depending on various factors like timing of the flights, destination, and duration of flights various occasions such as vacations or festive Season. Therefore, having some basic idea of the flight fares before planning the trip will surely help many people save money and time. The main goal is to predict the fares of the flights based on different factors available in the provided dataset.

# **1 Introduction**

# 1.1 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.

# 1.2 Architecture



# 2. Architecture Design

This project is to make associate interface for the user to grasp their approximate flight price ticket worth, additionally to the present, in would like of obtaining the important time project expertise we have a tendency to square measure mercantilism the gathered information into our own information then begin the project from the scratch.

# **2.1. Data Gathering**

The data for the current project is being gathered from ineuron dataset

# 2.2. Tool Used

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

* PyCharm/vscode is employed as IDE.
* For visualizations seaborn and components of matplotlib are getting used
* For information assortment prophetess info is getting used version

Management.

# 2.3 Data Description

There are about 10k+ records of flight information such as airlines, data of journey, source, destination, departure time, arrival time, duration, total stops, additional information, and price.

# 2.6 Data Pre-processing

Steps performed in pre-processing are:

* Checked for null values as there square measure few null values, those rows square measure born.
* Converted all the desired column into the date time format.
* Performed one-hot cryptography for the desired columns.
* Scaling is performed for needed information.
* And, the info is prepared for passing to the machine learning formula

# 2.7 Modeling

We have used Random forest Regressor to do the specific task