Architecture

FLIGHT FARE PREDICTION

Revision Number – 1.0

Last Date of Revision – 14/12/2022

ANWESHA DAS

Table of Contents

[ABSTRACT 3](#_Toc121913246)

[1 INTRODUCTION 4](#_Toc121913247)

[2 Architecture design 5](#_Toc121913248)

[2.1 Data gathering from main source 5](#_Toc121913249)

[2.2 Data description 5](#_Toc121913250)

[2.3 Data pre-processing 5](#_Toc121913251)

[2.4 Modelling 5](#_Toc121913252)

[2.5 UI integration 6](#_Toc121913253)

[2.6 Data from user 6](#_Toc121913254)

[2.7 Data validation 6](#_Toc121913255)

[2.8 Rendering the results 6](#_Toc121913256)

[2.9 Deployment 6](#_Toc121913257)

# 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.

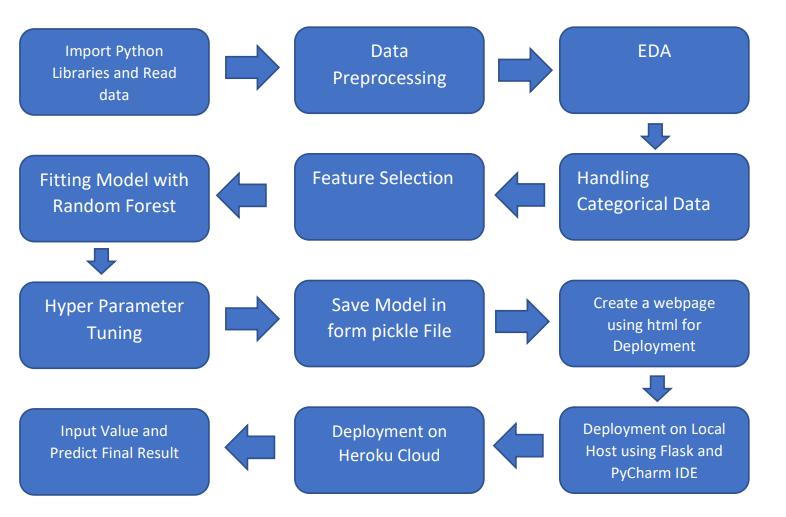
The main goal is to predict the fares of the flights based on different factors available in

The provided dataset.

# 1 INTRODUCTION

Why this Architecture Design documentation?

* The main objective of the Architecture design documentation is to provide the internal logic understanding of the flight fare prediction code. The Architecture design documentation is designed in such a way that the programmer can directly code after reading each module description in the documentation.



# 2 Architecture design

* This project is to create an interface for the user to know their approximate flight ticket price, 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 ineuron internship portal dataset.

# 2.2 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.

# 2.3 Data pre-processing

Steps performed in pre-processing are:

* First the data types are being checked and found only the price column is of type integer.
* Checked for null values as there are few null values, those rows are dropped.
* Converted all the required column into the date time format.
* Performed one-hot encoding for the required columns.
* Scaling is performed for required data.
* And, the data is ready for passing to the machine learning algorithm.

# 2.4 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 and further hyper parameter tuning is done to increase the model’s accuracy.

# 2.5 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.

# 2.6 Data from user

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

# 2.7 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.

# 2.8 Rendering the results

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

# 2.9 Deployment

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