High Level Design

Flight Fare Prediction

Revision number: 1.2

Last date of revision: 02/10/2022

# Document Version Control

|  |  |  |  |
| --- | --- | --- | --- |
| **Date Issued** | **Version** | **Description** | **Author** |
| 05/09/2022 | 1.0 | Initial HLD-V1.0 | Adil Anwar |
| 21/09/2022 | 1.1 | Design Flow | Adil Anwar |
| 02/10/2022 | 1.2 | Performance Evaluation Conclusion | Adil Anwar |

**Contents**

[Document Version Control 1](#_Toc16093)

[Abstract 2](#_Toc16094)

[1. Introduction 3](#_Toc16095)

[1.1. Why this High-Level Document? 3](#_Toc16096)

[1.2. Scope 3](#_Toc16097)

[1.3. Definitions 4](#_Toc16098)

[2. General Description 4](#_Toc16099)

[2.1 Product Perspective 4](#_Toc16100)

[2.2 Problem Statement 4](#_Toc16101)

[2.3 Proposed Solution 4](#_Toc16102)

[2.4 Data Requirements 4](#_Toc16103)

[2.5 Tools Used 4](#_Toc16104)

[2.6 Constraints 4](#_Toc16105)

[2.7 Assumptions 5](#_Toc16106)

[3. Design Details 6](#_Toc16107)

[3.1 Process flow 6](#_Toc16108)

[3.1.1 Model Training & Evaluation 6](#_Toc16109)

[3.1.2 Deployment Process 6](#_Toc16110)

[3.2 Event log 7](#_Toc16111)

[3.3 Error Handling 7](#_Toc16112)

[4. Performance 7](#_Toc16113)

[4.1 Reusability 7](#_Toc16114)

[4.2 Application Capability 7](#_Toc16115)

[4.3 Resource Utilization 8](#_Toc16116)

[4.4 Deployment 8](#_Toc16117)

[5. Conclusion 8](#_Toc16118)

[6. References 9](#_Toc16119)

# 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 High-Level Document?

The purpose of this High-Level Design (HLD) Document is to add necessary details to the current project description to represent a suitable model for coding. This model is also intended to help detect contradictions prior to coding and can be used as a reference manual for how the modules interact at a high level.

**The HLD will:**

* Present all the design aspects and define them in detail
* Describe the user interface being implemented
* Describe the hardware and software interfaces
* Describe the performance and requirements
* Include design features and the architecture of the project
* List and describe the non-functional attributes like:
  + Security
  + Reliability
  + Maintainability
  + Portability
  + Reusability
  + Application compatibility
  + Resource utilization
  + Serviceability

## Scope

The HLD documentation presents the structure of the system, such as the database architecture, application architecture, application flow (Navigations), and technology architecture. The HLD uses non-technical to miIdIy-technical term which should be understandable to the administrator of the system.

## Definitions

**TERM DESCRIPTION**

|  |  |
| --- | --- |
| **Database** | Collection of all the information monitored by this system |
| **IDE** | Integrated Development Environment |
| **AZURE** | Microsoft Azure |

# General Description

## Product Perspective

The flight fare prediction may be a machine learning model that helps users to predict the price of the flight tickets and help them to understand the price of their journey.

## Problem Statement

After amendment of the new rules, there is changes in the flight fare price from one location to another. The main goal of the system is to create a model to predict the price of their flight fare on the basis of bound input provided by user like date of journey, Source, Destination and many more.

## Proposed Solution

To solve the problem, we have created a User interface for taking the input from the user to predict the flight fare price using our trained ML model after processing the input and at last the output ( predicted value ) from the model is communicated to the User.

## Further Improvements

We also analysis the data used for training the ML model by considering different occasions such as Weekday, Season or any Social reasons, considering different angle of business. If we method such information and predict the discounted flight fare price, it will bring some loss to the airline companies but user can get benefit from that. If we develop these using Business perspective of Airline, this technique isn’t thought - about.

## Technical Requirements

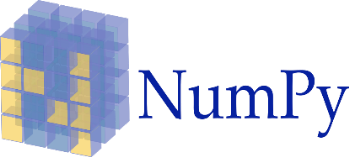
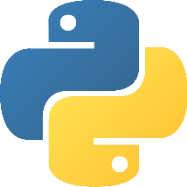
As technical requirements, we doesn’t need any specialized hardware for virtualization of the application. The user should have the device that has the access to the web and the fundamental understanding of providing the input. And for the backend, we need server to run all the required packages to process the input and predict the desired output ( predicted flight fare price ).

## Data Requirements

The Data requirements is totally supported the matter statement and also the dataset is accessible on the Kaggle within the file format of (.xlsx). Because the main theme of the project is to induce the expertise of real time issues, we have a tendency to transform the information into the prophetess database and commerce it into csv format.

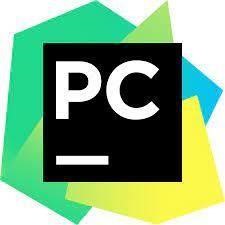
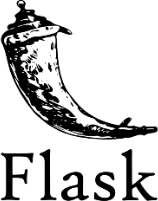
## Tools used

Python programming language and frameworks such as NumPy, Pandas, Scikit-learn, Flask, Azure, Git.









* PyCharm is used as IDE.
* For visualization of the plots, Matplotlib, Seaborn and Plotly are used.
* Azure is used for deployment of the model.
* Frontend development is done using HTML/CSS on Sublime Text.
* Python Flask is used for backend development.
* GitHub is used as version control system.

## Constraints

The flight fare prediction answer should be user friendly, as automatic as attainable and also the user should not be needed to understand any of the operating.

## Assumptions

The main objective of the project is to implement the utili cases as for the new dataset that provides the user the ability to predict Flight Fare Price. Machine learning model is employed for process the user input for prediction. It additionally assumed that each one aspects of this project have the flexibility to figure along within the approach the designer is expecting.

# Design Details

## Process Flow

For predicting the fare price prediction, we will use regression model. Below is the process flow diagram is as shown below.

Proposed Methodology

Capture the data

from the web app

Training /

validation on

dataset

ML model for

Regression

Prediction of the

output

### Model Training and Evaluation

**Data Pre-processing and Feature Engineering**

**Data collection**

**from source**

**Split the data**

**into train & test**

**Feature**

**selection**

**Train the**

**model**

**Test the**

**model**

**Improve the**

**model**

### Deployment Process

Data

ML model

training

Model

prediction

Flask

HTTP /

Rest API

WSGI

HTTP

server

Azure

User

## Event Log

The system should log every event so that the user will know that process is running internally.

**Initial Step-By-Step Description:**

* The system identifies at what step logging required.
* The system should be able to log each and every system flow.
* Developer can choose logging method. You can choose database logging / File logging as well.
* System should not hang even after using loggings. Logging just because we can easily debug issues, so logging is mandatory to do.

## Error Handling

Should error be encountered, an explanation will be displayed as to what went wrong? An error will be defined as anything that falls outside the normal and intended usage

# Performance

## Reusability

The code written and the components used should have the ability to be reused with no problems.

## Application Compatibility

The different components for this project will be using as an interface between them. Each component will have its own task to perform, and it is the job of the python to ensure proper transfer of information.

## Resource Utilization

When any task is performed, it will likely use all the processing power available until that function is finished.

## Deployment

The model can be deployed using the cloud service Microsoft Azure.



# Conclusion

The Flight Fare Prediction system will predict the price for helping the customers with the trained knowledge with set of rules. The user can use this system to recognize the approximate value of its flight fare for his or her journey.

# References

1. www.kaggle.com
2. A Framework for Airfare Price Prediction: A Machine Learning Approach

Tianyi Wang, Samira Pouyanfar, Haiman Tian, Yudong Tao, Miguel Alonso Jr., Steven Luis and Shu-Ching Chen. School of Computing and Information Sciences, Florida International University.

1. Airfare prices prediction using machine learning techniques

K. Tziridis; Th. Kalampokas; G. A. Papakostas; K. I. Diamantaras. Department of Computer and Informatics Engineering, Eastern Macedonia and Thrace Institute of Technology, Kavala, Greece.

1. <https://scikit-learn.org/stable/> Scikit-Learn: Machine Learning Python.