High-Level Design (HLD)

Flight Ticket Price Prediction

Revision Number: 2.0

Last Date of revision: 03/2/2023

Document Version control

Date issued	Version	Descriptions	Author
20/12/2022	1.0	Abstract, Introduction	Varun Salunkhe
3/1/2023	1.1	General Description	Saurabh Jumnalkar
15/1/2023	1.2	Design Details	Saurabh Jumnalkar
28/1/2023	1.3	KPLs, Conclusion	Sourabh Hawale

Contents

Document Version Control

Abstract

Abstract

For purchasing an airplane ticket, the traditional purchase approach is to buy a ticket far in advance of the flight's departure date to avoid the risk that the price may increase quickly before the date of departure. However, this is not always the case; if airline corporations wish to increase sales, they can lower prices. Airlines employ a variety of factors to decide flight ticket rates, including whether the trip is around the holidays, the quantity of available seats on the plane, and even the month. Some of the variables can be seen, while others are hidden. In this context, customers are attempting to discover the best day to purchase a ticket, while airline firms, on the other hand, are attempting to maximize overall revenue.

1. Introduction

1.1 Why this High-Level Design?

The purpose of this High-Level Design (HLD) Document is to add the necessary detail to the current project description to represent a suitable model for coding. This document 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 of the design aspects and define them in detail Describe the user interface being implemented
- Describe the hardware and software interfaces
- Describe the performance 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

1.2 Scope

The HLD documentation presents the structure of the system, such as the database architecture, application architecture (layers), application flow (Navigation), and technology architecture. The HLD uses non-technical to mildly-technical terms which should be understandable to the administrators of the system.

2. General Description

2.1 Problem Statement

For purchasing an airplane ticket, the traditional purchase approach is to buy a ticket far in advance of the flight's departure date to avoid the risk that the price may increase quickly before the date of departure. However, this is not always the case; if airline corporations wish to increase sales, they can lower prices. Airlines employ a variety of factors to decide flight ticket rates, including whether the trip is around the holidays, the quantity of available seats on the plane, and even the month. Some of the variables can be seen, while others are hidden. In this context, customers are attempting to discover the best day to purchase a ticket, while airline firms, on the other hand, are attempting to maximize overall revenue.

2.2 Data Content

Flight ticket prices can be something hard to guess, today we might see a price, check out the price of the same flight tomorrow, it will be a different story. We might have often heard travelers saying that flight ticket prices are so unpredictable. As data scientists, we are gonna prove that given the right data anything can be predicted. Here you will be provided with prices of flight tickets for various airlines between the months of March and June of 2019 and between various cities.

2.3 About this File

- Size of test set: 2671 records
- FEATURES: Airline: The name of the airline.
- Date_of_Journey: The date of the journey
- Source: The source from which the service begins.
- Destination: The destination where the service ends.
- Route: The route taken by the flight to reach the destination.
- Dep_Time: The time when the journey starts from the source.
- Arrival_Time: Time of arrival at the destination.
- Duration: Total duration of the flight.
- Total_Stops: Total stops between the source and destination.
- Additional_Info: Additional information about the flight

- Price: The price of the ticket

2.4 Tools Used

2.4.1 Hardware Requirements

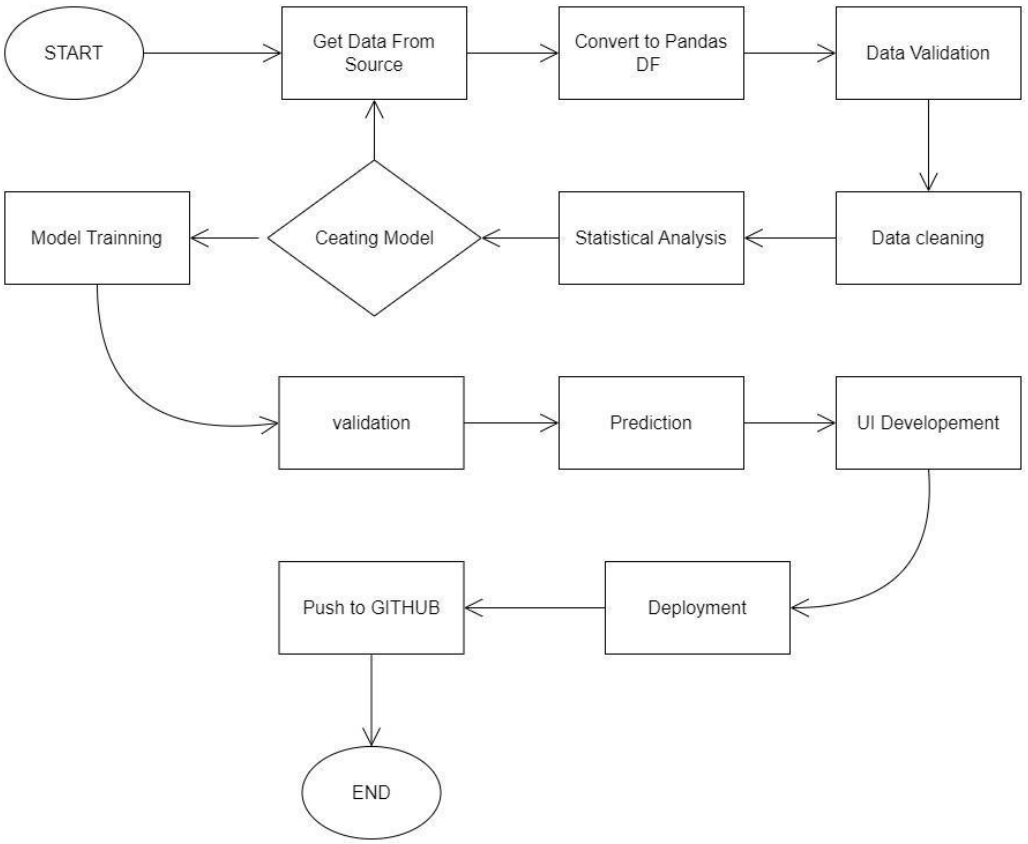
- All Operating System
- Min 500MB ram
- I3 intel

2.4.2 Software Requirements

- Python
- Flask Framework
- Vs Code
- Numpy
- Pandas
- Matplotlib
- Jupyter

3. Design Details

3.1 Process Flow



3.2 Event log

We should be able to log every activity done by the user.

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

3.3 Error Handling

In Python, A traceback is a report containing the function calls made in your code at a specific point i.e when you get an error it is recommended that you should trace it backward (traceback). Whenever the code gets an exception, the traceback will give the information about what went wrong in the code. Python traceback contains great information that can help you find what is going wrong in the code. These tracebacks can look a little wearisome, but once you break them down to see what it's trying to show you, they can be very helpful.

3.4 Performance

Performance is quite good, as you can see below accuracy score for training set is 95.78% and accuracy score for test set is 79.65%.

```
# Model Performance

print('Model Performance on Training Set:', round(model.score(train_x, train_y)*100,2))
print('Model Performance on Test Set:', round(model.score(test_x, test_y)*100,2))
```

```
Model Performance on Training Set: 95.78
Model Performance on Test Set: 79.65
```

4. Dashboards

4.1 KPLs

Thus, by using the Wrapper method and Filter method, the Key Attributes that contributed to the better Price of flight ticket. The attributes that have been found to be the best ones from both the attribute selection methods are compared. It is found that both the attribute selection methods almost gave the same results as the output. Hence by using these attributes as the key attributes, there will be better accuracy in the Price of flight ticket.

5. Conclusion

This project the methods of pre-processing, steps to identify the key attributes that help in the better Price of flight ticket for a given set of travel details, like: departure date, arrival date, departure city, arrival city, stoppages, and the airline carrier.