High Level Design (HLD)
Flight Ticket Price Prediction

Revision Number: 2.0

Last date of revision: 05 March 2023

Mohammed Saad K

Document Version Control

Date Issued	Version	Description	Author
05/03/2023	1.0	Added Introduction and General description	Mohammed Saad K
06/03/2023	1.1	Added Design details and abstract	Mohammed Saad K
07/03/2023	1.2	Created and organized whole document	Mohammed Saad K
07/03/2023	1.3	Updated Design Details	Mohammed Saad K
09/03/2023	1.4	Checked and added Performance	Mohammed Saad K
10/03/2023	1.5	Added Conclusion and References	Mohammed Saad K
11/03/2023	2.0	Updated diagrams in Design Details	Mohammed Saad K

Table of Content

Abstract

A lot of factors that affect the overall price of airline tickets, including the airline, the date of travel, source, destination, route, duration, and so on. Each provider seems to have its own unique set regulations and methods for determining pricing. Recent breakthroughs in Artificial Intelligence (AI) and Machine Learning (ML) allow for the inference of such principles as well as the modelling of price volatility. This article is a study conducted on predicting flight prices. Utilizing two datasets for testing and training, this study analyses various machine learning methods for predicting flight prices.

Introduction

1. Why this High-Level Design Document?

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 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

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.

General Description

1. Product Perspective

Travelling through flights has become an integral part of today's lifestyle as more and more people are opting for faster traveling 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.

2. Problem Statement

Predict the price of the flight based on various factors.

3. Problem Solution

A system / web-applications which can predict the price of the flight based on various affecting factors. The price should be calculated based on the time, source, destination and number of stops given by user. You can add multiple factors to decide the price.

4. Further Improvement

The future aim is to work more on the feature selection and model accuracy. We can also consider various other crucial features that affect the flight ticket prices like public holidays, number of luggage, number of hours till departure, crude oil price, etc. in order to get best results.

5. Data Required

For training the model I need the data that consist of Airline, Date_of_Journey, Source, Destination, Route, Dep_Time, Arrival_Time , Duration, Total_Stops, Additional_Info, Price.

Data is completely depending upon our problem statement

6. Tools Used

- Python programming language and frameworks such as NumPy, Pandas, Scikit-learn, Matplotlib, Seaborn are used to build the whole model.
- Visual Studio Code is used as IDE.
- For visualization of the plots, Matplotlib and Seaborn are used.
- Front end development is done using HTML, CSS and Bootstrap.
- GitHub is used as version control system.

7. Constraints

The Flight Ticket Price Prediction website should be user friendly. Distance, Peak season, Flight timings, travel type etc, affects the Prediction.

8. Assumptions

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 dataset.

Design Details

1. Process Workflow

For identifying the different types of anomalies, we will use a machine learning model. Below is the process flow diagram.
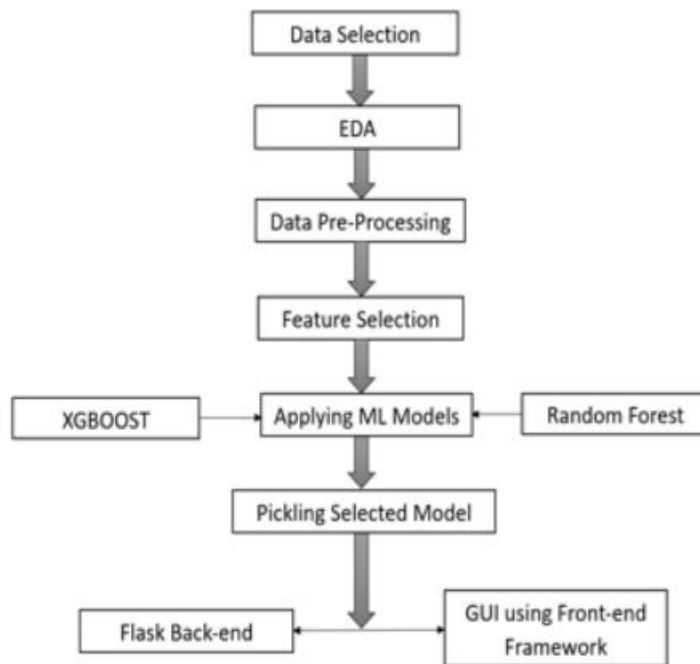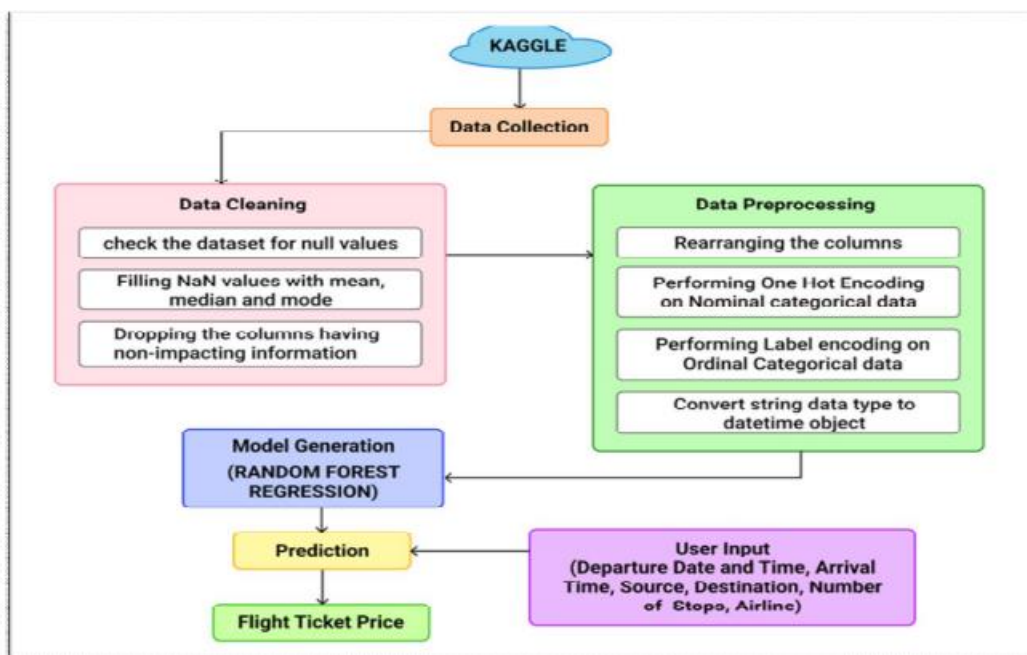

Fig. 1 Proposed System Diagram

Model Training and Evaluation

2. Error Handling

Initially I got an error displaying the graphs and multiple text were not solving.
All of the above errors have been resolved.

Performance

1. Reusability

Price Prediction of the Flight Tickets should be as accurate as possible, so that it will not mislead the user. Best possible model will be used to predict prices. Documentation is available, my project follows reusability.

2. Application compatibility

Since I am using python and it is compatible with any platform, I follow Application compatibility

3. Resource utilization

At the initial stage, I was using high space to create the model. Once the model is created, my system only needs at least of 2GB RAM and 1 GB of storage to run the application smoothly. Whenever user tries to predict the price of the flight tickets, system uses less than 10% of the processing power.

4. Deployment

The code is deployed in GitHub.

Conclusion

This project proposes the machine learning model for Flight Ticket Price Prediction. This model can be used for Travelers to get the fare prediction handy using which it's easy to decide the airlines. It also saves time in searching / deciding for airlines.

References

1. https://www.ijraset.com/research-paper/flight-price-prediction
2. https://medium.com/geekculture/flight-fare-prediction-93da3958eb95
3. https://www.analyticsvidhya.com/blog/2022/01/flight-fare-prediction-using-machine-learning/
4. https://numpy.org/doc/
5. https://seaborn.pydata.org/examples/regression_marginals.html
6. https://seaborn.pydata.org/examples/scatterplot_matrix.html
7. https://matplotlib.org/
8. https://pandas.pydata.org/docs/