High Level Design(HLD)

FLIGHT FARE PREDICTION

Submitted by :

Umesh Prajapati

Ramdas Prajapati

Contents

Abstract

Traveling 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, duration of flights. various occasions such as vacations or festive seasons. Therefore, having some basic idea of the flight fares before planning the trip will surely help many people save money and time. In the proposed system a predictive model will be created by applying machine learning algorithms to the collected historical data of flights. This system will give people the idea about the trends that prices follow and also provide a predicted price value which they can refer to before booking their flight tickets to save money. This kind of system or service can be provided to the customers by flight booking companies which will help the customers to book their tickets accordingly.

1 Introduction

1.   Why this High-Level Design Document?

The main purpose of this HLD 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 a careful description on how the complete project has been designed end-to-end.

2. Description

Problem Perspective

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.

3.   Problem Statement

The objective of this project is to predict flight prices given the various parameters. Data used in this project is publicly available at Kaggle. This will be a regression problem since the target or dependent variable is the price (continuous numeric value).

1.4. Project Solution

Nowadays, the number of people using flights has increased significantly. It is difficult for airlines to maintain prices since prices change dynamically due to different conditions. That's why we will try to use machine learning to solve this problem. This can help airlines by predicting what prices they can maintain. It can also help customers to predict future flight prices and plan their journey accordingly.

5. Answer enhancements

We will even predict the price of the price tag considering whether or not it is a weekday, season or alternative social reasons. However, considering from the angle of business, if we have a tendency to method such information and predict the price of the discounted price tag it'll bring some loss to the airline company. Therefore, this technique isn't thought-about.

6. Technical needs

There are not any hardware needs needed for victimization of this application, the user should have AN interactive device that has access to the web and should have the fundamental understanding of providing the input. And for the backend half the server should run all the package that's needed for the process the provided information and to show the results.

7. Information needs

The info demand totally supported the matter statement. and also, the information set is accessible on the Kaggle within the type of standout sheet(.xlsx). Because the main theme of the project is to induce the expertise of real time issues, we have a tendency to ar once more mercantilism {the information into the prophetess database and commerce it into csv format.

8. Tools Used

- Python 3.9 is employed because of the programming language and frameworks like numpy, pandas, sklearn and alternative modules for building the model.
- PyCharm is employed as an IDE.
- For visualizations seaborn and components of matplotlib are getting used.

- For information assortment prophetess info is getting used.
- Front end development is completed victimization HTML/CSS.
- Flask is employed for each information and backend readying.
- GitHub is employed for version management.
- Heroku is employed for deployment.

9. 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 operations.

10. Assumptions

The most objective of the project is to implement the utilization cases for the new dataset that the user provides through the programme. Machine learning model is employed to process the on top of a computer file. it's additionally assumed that each one aspect of this project has the flexibility to figure along within the approach the designer is expecting.

2.1 and 2.2 Design Flow and Deployment Process
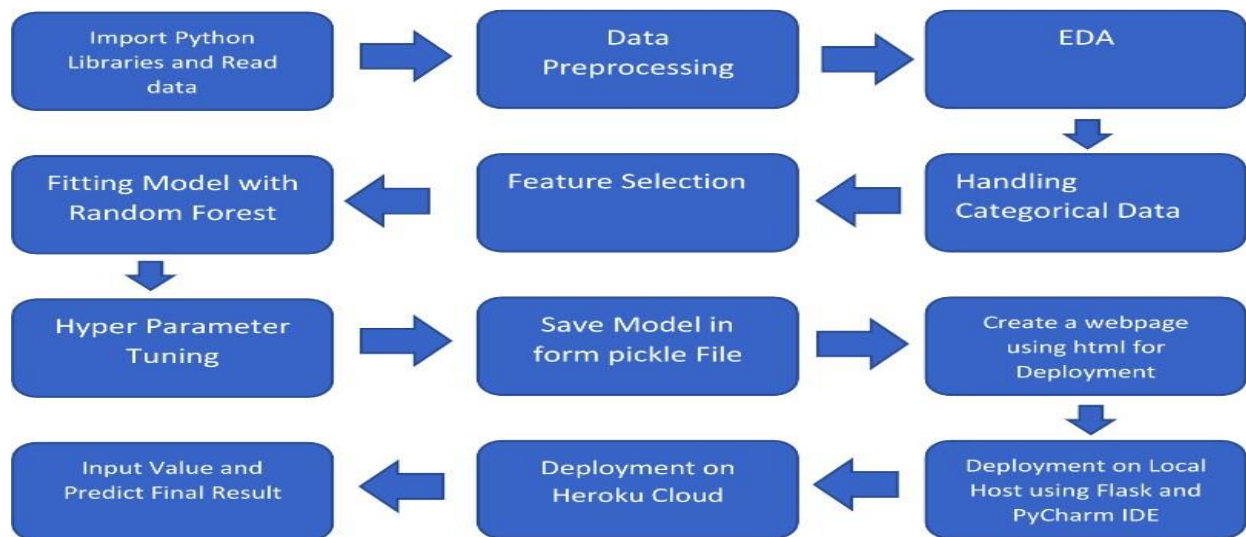
3. Logging

Each step is being logged within the system that runs internally, that shows the date time and therefore the process that has been performed, work is completed in several layers as information, DEBUG, ERROR, WARNINGS. This provides the US the perception of the logged info.

4. Error Handling

Once a slip has occurred, the reason is logged in its several log files, in order that the developer will rectify the error.

3 Performance analysis

1. Reusability

Elements of the code written are accustomed to different applications and therefore the rest is changed and reused.

2. Application Compatibility

The various parts for this project are exploitation python as an associate interface between them. every element can have its own tasks to perform, and it's the work of the python to make sure the transfer of data.

3. Resource Utilization

Once any task is performed, it'll doubtless use all the process power offered till that performance is finished.

4. Deployment

The model is being deployed on Heroku.

Conclusion

The flight fare prediction will predict the worth of the trained knowledge set within the rule. therefore, the user will recognize the approximate value for his or her journey