High Level Design (HLD)

Flight Fare Prediction

Pritam Bhakta

Abstract

Introduction

Need for High Level Document (HLD)?

1. Description
   1. Problem Perspective
   2. Problem Statement
   3. Proposed Solution
   4. Solution Improvement
   5. Technical requirements
   6. Data requirements
   7. Tools used
   8. Constraints
   9. Assumptions
2. Design flow
   1. Process Modelling
   2. Process Deployment
   3. Logging
   4. Error Handling
3. Performance Evaluation
   1. Reusability
   2. Application compatibility
   3. Resource utilization
   4. Deployment
4. Conclusion

Abstract

Travelling has become the integral part of human lives. Some travel for some purposes, some of them travel for passion or for some other purposes. In today’s fast paced life each and every one is opting for the faster travel options to travel form one place to other. In this arena fare of the flights keep on fluctuating every now and then, it majorly depends on the seasons, weekends, yearend, discounts and many more cases. So having some basic idea about the fare of the ticket will surely help every individual which will be travelling in the coming future will surely help him to manage his expenses up to some extent. In this case what if we design a predictive model which will predict the fare of the ticket based on the features provided such as source, destination, departure time and arrival time along with the date of the journey. This kind of the predictive systems aviation companies can integrate to provide their consumers next level of booking experience.

1. INTRODUCTION

1.1 Need of High-Level Design?

The main purpose of the HLD document is to feature the required details of the project and give the overview of the machine learning model and written code as well.

* 1. Description

Problem Perspective

Flight ticket is something which is very hard to predict, we often here passengers saying flight ticket was something different yesterday than what it is today. As working the data science industry, we know that if we are providing the right data than we can predict anything related to that dataset. Here we are provided the flight tickets and their dataset for year 2019 based on that we have to build the model which will be predicting the flight fare of tickets with provided input data by user.

* 1. Problem Statement

The main objective of this project is to make the interactive web application which will predict the fare of the prices by taking some parameters. The dataset which is used in this problem statements is publicly available on the Kaggle platform.

* 1. Project Solution

These days the number of people who are travelling using flights is increasing tremendously and for aviation companies it is very difficult to change the prices of the flights dynamically based on certain conditions, that’s why they will be using machine learning to solve this problem this can help them to predict the prices on the other hand it also helps consumers to predict the prices of upcoming journey and prepare themselves accordingly which also helps in saving money.

* 1. Answer enhancement

While building this machine learning model we will also consider the flight price based on the festive seasons, weekdays, weekends, seasons and discounts which is provided by various aviation industries but we are not considering it because price of flight considering all these parameters may vary from company to company so using this particular approach may not be good and may lead to loss for various industry.

* 1. Technical needs

There is no hardware requirements as such for this project, the only requirements which is needed that is user should have an interactive device which will be able to connect to internet and at the same time user should have the understanding of giving inputs to the web application

* 1. Information needs

The information which is needed for this project is available on the kaggle.com website all the information about this dataset is available on this website.

* 1. Tools used
* Python 3.8 is used in this project along with that some python libraries like NumPy, pandas, matplotlib, seaborn, scikit learn is been used.
* Jupyter notebook and Visual studio code is the IDE which is used for coding purpose.
* Front end web application is designed using HTML and CSS.
* FASTAPI is used as the framework in the backend.
* Git hub is used for version control.
* GCP is used for the deployment.
  1. Constraints

The web application should be user friendly so that without knowing any technical information he should be able to use our predictive system.

* 1. Assumptions

The main objective of the project is to utilize the data which is provided by the user and to predict the flight fare. The web application should be accessible from every system which is connected to the internet.

1. FLOW

2.1 Design flow and deployment process

* 1. Logging

Each step is been logged within the system that runs internally, it basically shows us the data time of each process which is done with our system. It provides us the logging information.

2.3 Error handling

The logging which we have done in the above process it helps us to handle the error because the error is being logged in several log file so that developer can rectify it

1. PERFORMANCE ANALYSIS

3.1 Reusability

The elements of the code is written in such a way that it can be written and reused again.

* 1. Application Compatibility

The elements of the project are written in python, it acts as the interface between model to the user.

* 1. Resource Utilization

Once the task is assigned to the model doubtlessly it will use all the resource which is allocated until the task is finished.

* 1. Deployment

This model is deployed on GCP which is the cloud platform by google.

1. CONCLUSION

Flight share prediction will predict the flight fare once you enter the parameters.