High Level Design (HLD)
Flight Price Prediction

Revision Number : 1.2
Last Date of revision : 23/06/2023

Document Version Control

Date Issued	Version	Description	Author
15-06-2023	1.0	Initial HLD – V1.0	ANISH MAURYA
18-06-2023	1.1	Design Flow	ANISH MAURYA
23-06-2023	1.2	Performance Evaluation Conclusion	ANISH MAURYA

Contents

Abstract

Flight price prediction is a challenging task that involves forecasting the prices of airline tickets based on various factors such as historical data, seasonality, demand, and other external factors. Accurate price prediction can be beneficial for travelers, airlines, and travel agencies in making informed decisions related to ticket purchasing, revenue management, and pricing strategies.

This project aims to develop a flight price prediction model using machine learning techniques. The dataset used for training and testing the model consists of historical flight data, including features such as departure and arrival locations, date and time of travel, airline carrier, and other relevant information. The dataset is preprocessed to handle missing values, outliers, and feature engineering to extract useful insights.

Different machine learning algorithms such as regression, time series analysis, and ensemble methods are implemented and evaluated to identify the most suitable approach for flight price prediction. The models are trained on a subset of the dataset and evaluated using appropriate evaluation metrics such as Accuracy Score

In addition to traditional machine learning techniques, advanced methods like deep learning and neural networks may also be explored to capture complex patterns and dependencies in the data. Transfer learning and feature selection techniques can be employed to enhance the model's performance and reduce computational complexity.

The trained flight price prediction model can be integrated into a user-friendly application or website where travelers can input their desired travel details and receive accurate price estimates. The model's performance can be continuously monitored and refined using additional data to improve its accuracy over time.

Overall, this project aims to contribute to the field of travel and tourism by developing a reliable and accurate flight price prediction model that can assist travelers and industry professionals in making informed decisions related to flight bookings and pricing strategies.

1. Intíoduction

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 referencemanual 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 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

Scope

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

Definition

Term	Description
BSP	Flight Price Prediction
Jupyter – Notebook	It is an interactive computational environment, in which we can combine code execution, text, plots rich media.
Azure	It is a platform as a service (pass) that enables developers to build, run and operate application in the cloud.

2. General Description

Problem Perspective

The Flight Price Prediction is a machine learning model that helps companies to predict how many fare of between to city based on some input data

Problem Statement

The main goal of this model is to predict flight price based on the input features

Pioposed Solution

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

Technical Requirements

As technical requirements, we don't need any specialized hardware for virtualization of the application. The user should have a device that has the access to the web and the fundamental understanding of providing the input. And for the backend, we need a serverto run all the required packages to process the input and predict the desired output.

Data Requirements

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

Tools used

Python programming language and frameworks such as Numpy, Pandas, Scikit-learn, VSCode, git and Azure are used to build the whole model

- VS Code is used as IDE.
- For visualization of the plots, Matplotlib, Seaborn and Plotly are used.
- Azure is used for deployment of the model.
- Front end development is done using Azure.
- Python Flask is used for backend development.
- GitHub is used as version control system.

Constrains

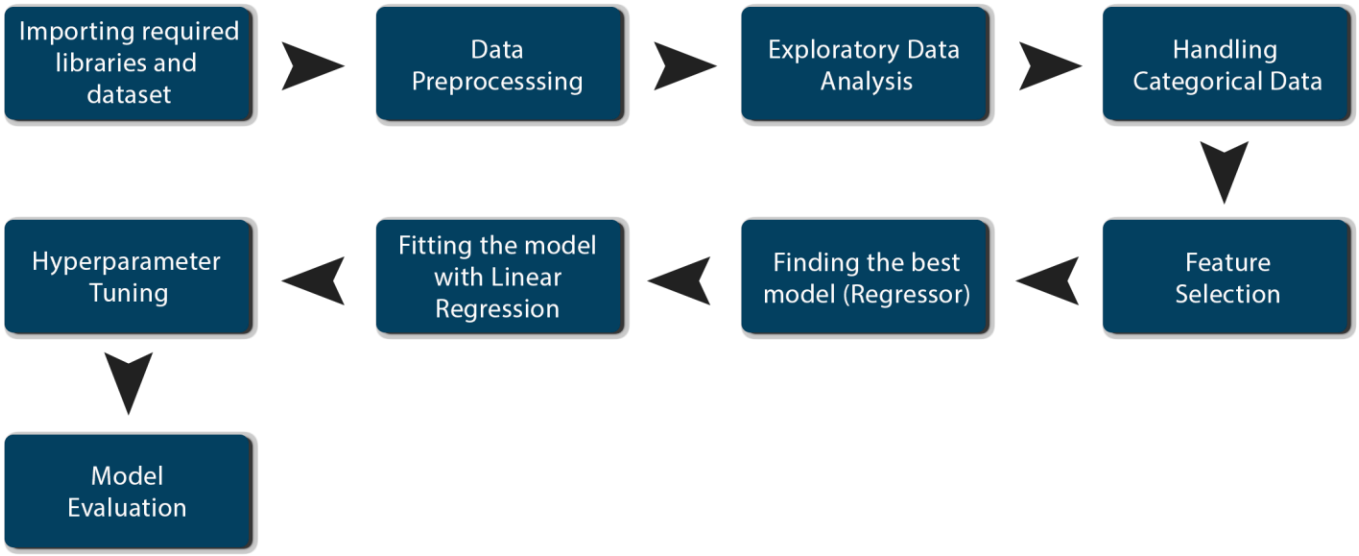
The system must be user friendly, as automated as possible and users should not berequired to know any of the workings.

Assumptions

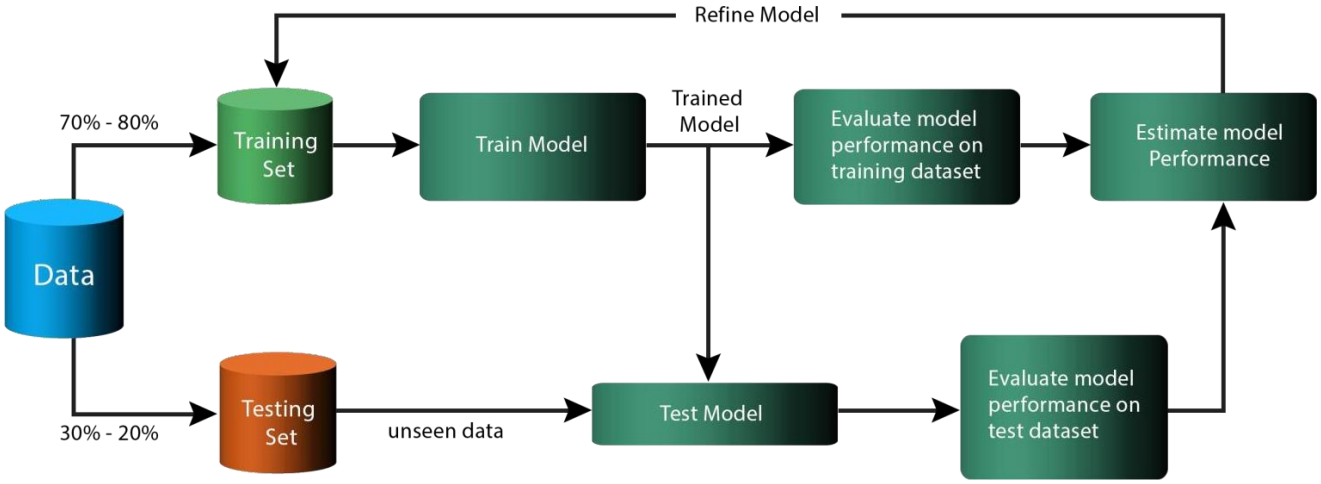
The main objective of the project is to implement the use cases as previously mentioned (2.2 Problem Statement) for new dataset that comes through source. Machine Learning based model is used for detecting the above-mentioned use cases based on the input data.It is also assumed that all aspects of this project have the ability to work together in the waythe designer is expecting

3. Design Details

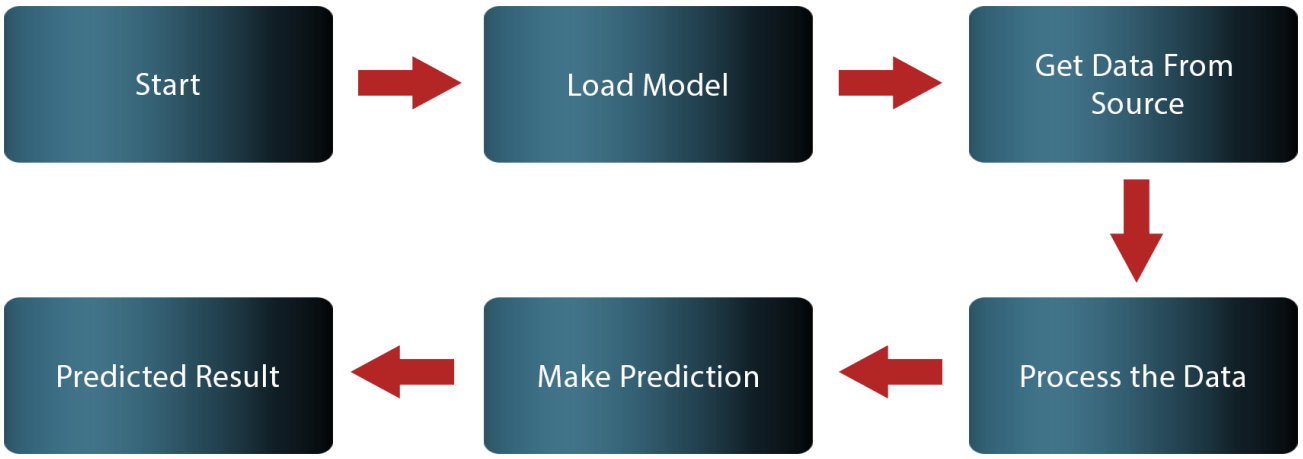
Process Flow



Model Training and Evaluation



Deployment Process



Logging

In logging, each time an error or an exception occurs, the event is logged into the system log file with reason and timestamp. This helps the developer to debug the system bugs andrectify the error.

Error Handling

Should errors 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

4. Performance Evaluation

The machine learning-based flight price Prediction project predicts flight prices between two cities.

Reusability

The code written and the components used should have the ability to be reused with noproblems.

Application Compatibility

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

Resource Utilization

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

Deployment

Conclusion

Flight price prediction is a complex task that requires the integration of various factors and machine learning techniques. In this project, we developed a flight price prediction model using historical flight data and machine learning algorithms. The dataset was preprocessed and features were engineered to extract useful insights.

By developing a reliable and accurate flight price prediction model, this project contributes to the field of travel and tourism. It empowers travelers and industry professionals with valuable insights for making informed decisions regarding flight bookings and pricing strategies.