

# High-Level Document (HLD)

## Flight Price Prediction Project

### Introduction:

The Flight Price Prediction project aims to develop a system that can accurately predict flight ticket prices based on historical data and various influencing factors. The system will assist travelers in making informed decisions about their flight bookings by providing estimates of future ticket prices.

### Objectives:

- Develop a machine learning model that predicts flight ticket prices with high accuracy.
- Gather and analyze historical flight data to identify significant factors affecting ticket prices.
- Build a user-friendly interface for users to input their travel preferences and receive price predictions.
- Enhance the system's performance and accuracy through continuous model training and improvement.

### Scope:

The Flight Price Prediction system will focus on the following aspects:

- Domestic flights within a single country.
- Economy class ticket prices for major airlines.
- Prediction of ticket prices for up to six months in advance.

### System Architecture:

The Flight Price Prediction system will consist of the following components:

### Data Collection:

- Retrieve historical flight data from reliable sources such as airline websites, travel agencies, or APIs.
- Extract relevant features, including departure/arrival locations, dates, flight duration, airline, and ticket price.

### Data Preprocessing:

- Clean and transform the collected data to ensure consistency and accuracy. Handle missing values, outliers, and data inconsistencies.

- Perform feature engineering to derive additional relevant features.

## **Machine Learning Model:**

- Select a suitable machine learning algorithm for regression, such as Random Forest, Gradient Boosting, or Neural Networks.
- Train the model using the preprocessed historical data.

Evaluate and optimize the model's performance using appropriate metrics and techniques.

## **User Interface:**

- Design a user-friendly interface for users to input their travel preferences.
- Collect user input, including departure/arrival locations, travel dates, and other relevant information.
- Display the predicted ticket price based on the user's input and the trained machine learning model.

## **Model Updates:**

- Periodically retrain the machine learning model using updated historical data to improve prediction accuracy.
- Implement a mechanism to automatically update the model and its predictions on a regular basis.

## **Assumptions and Constraints:**

- The accuracy of flight price prediction may vary based on the availability and quality of historical data.
- The prediction model's performance may be influenced by factors such as economic conditions, seasonal variations, and unforeseen events.
- The system will not consider additional factors like seat availability, baggage fees, or ancillary services that may affect the final ticket price.

## **Risks and Mitigation:**

- Insufficient or unreliable historical data: Mitigate by collecting data from multiple sources and implementing data quality checks.
- Inaccurate predictions: Regularly evaluate and refine the machine learning model to improve accuracy.
- Changing market conditions: Monitor external factors and update the model periodically to adapt to market changes.

## **Timeline:**

- Data collection and preprocessing: 2 weeks

- Model development and training: 4 weeks
- User interface design and implementation: 2 weeks
- Model updates and maintenance: Ongoing

## **Future Enhancements:**

- Include more advanced features, such as weather conditions, holidays, or promotional events.
- Expand the system to cover international flights and premium class ticket prices.
- Provide recommendations for alternative travel dates or destinations based on predicted price fluctuations.