

FareSight: Leveraging Machine Learning to Forecast Flight Prices

Aarzoo
(2022008)

Anushka Srivastava
(2022086)

Nandini Jain
(2022316)

Suhani Kalyani
(2022511)

Project Proposal

1 Motivation

Flights have become the preferred choice for long distance travels. However, the cost of flight tickets varies a lot, increasing one day and dropping the very next. This leaves many travelers confused about the factors affecting these variations which is an important financial consideration.

Like many travelers, we have personally experienced the inconsistencies in prices while booking flight tickets. Therefore, we plan to build a predictive model that can help predict the prices of flight tickets. This will help travelers make more informed decisions in their travel plans, and allow us to contribute to making air travel more accessible and affordable.

2 Related Work

2.1 Dynamic Flight Price Prediction Using Machine Learning Algorithms

The paper describes a machine learning model capable of predicting flight prices. The study reveals that most airline ticket prices fluctuate daily. The dataset contains flight prices of various airlines between March and June 2019 across different cities, along with other relevant attributes. Three machine learning models are implemented, including Linear Regression, Random Forest, and Decision Trees, achieving an accuracy of 80%.

2.2 Flight Price Prediction for Enhanced Recommendations via Machine Learning Web Application

The paper describes a dependable machine learning model that can accurately estimate flight prices. Six different machine learning algorithms are thoroughly evaluated to ensure the model's accuracy and reliability. Performance metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE) etc. are used for comparison. Among the algorithms tested, the Random Forest Regressor outperformed the others, achieving the highest R^2 score.

2.3 A Prediction of Flight Fare Using K-Nearest Neighbors

The study identifies the factors driving airplane price fluctuations, how they influence price changes. The study employed the K-Nearest Neighbors algorithm to model the factors. Data visualization determined the most impactful factors on airplane prices. The model was evaluated using various metrics. The results show that the K-Nearest Neighbors algorithm achieved a solid accuracy of 81.77%.

3 Timeline

1. **First week of September:** Data Retrieval and Analysis
2. **September:** Statistical analysis of data, feature analysis, data pre-processing
3. **First week of October:** Training models and analyzing the result of each model
4. **Mid-October:** Feature engineering and hyperparameter tuning to optimize model performance
5. **Last week of October:** Selecting the best model on the basis of evaluation metrics
6. **November:** Documenting the results

4 Individual Tasks

Each team member is expected to contribute to the above mentioned tasks to the best of their ability, and the expected contribution from each member is equal for all tasks.

5 Final Outcome

By analyzing the dataset and training ML models using feature engineering and hyperparameter tuning, we aim to identify the patterns and trends in fare fluctuations and the factors affecting fares to achieve a robust and accurate model. This could lead to practical applications in optimizing travel planning.

Apart from forecasting flight fares, we also aim to offer additional insights, such as the best times to book flights and how airline choices impact fares, helping travelers to decide better.

References

- [1] "Dynamic flight price prediction using machine learning algorithms," IEEE Conference Publication — IEEE Xplore, Dec. 16, 2022. Paper 1
- [2] "Flight price prediction for enhanced recommendations via machine learning web application," IEEE Conference Publication — IEEE Xplore, Jul. 10, 2024. Paper 2
- [3] "A prediction of flight fare using K-Nearest neighbors," IEEE Conference Publication — IEEE Xplore, Apr. 28, 2022. Paper 3