

# Flight Price Prediction Using Machine Learning

*G D Vijaya Lakshmi<sup>1</sup>, T S Ritvik<sup>2</sup>, Thota Gayathri<sup>3</sup>*

*<sup>1</sup>, Assistant Professor, Department of IT, GRIET, Hyderabad, Telangana, India*

*<sup>2,3</sup>UG Student, Department of IT, GRIET, Hyderabad, Telangana, India.*

**Abstract:** The subject of airfare is examined in this paper. As a result, a collection of factors that characterize a typical flight are selected under the presumption that they have an impact on airline ticket costs. The price of a plane ticket is influenced by the length of the trip, the location, the schedule, and a number of other factors, like holidays or vacations. Therefore, many people will surely save time and effort by having a basic awareness of airline expenses prior to making trip arrangements. The performance of the seven different machine learning (ML) models used to anticipate the price of airline tickets is compared after three datasets were analyzed to acquire insight into airline fares. The objective is to investigate the factors that influence flight prices. The data can then be used to build a system that can predict how much a flight will cost.

## 1. Introduction

To amplify benefits, aircrafts today end over towards control cost about flight tickets. By far most who fly reliably know best times towards buy unassuming tickets. Nevertheless, various clients who are awful at booking tickets fall into markdown trap set by association, causing them towards spend their money. While clients are searching for best arrangement, aircraft organizations' essential goal is benefit. Clients a large part about time intend towards purchase tickets a longways in front about flight date towards hinder cost increases as departure date moves close.

## 2. Problem Statement

The authors of [1] used a linear quantile mixed regression model to estimate the lowest ticket price before departure. In order to predict the lowest price, all observations are used to create one lowcost quintile. 2,271 flights out of a total of 126,412 recordings on a single route made up the data used, which was collected 60 days prior to shipment. Amount, departure date, statement date, list of dates before to departure, and weekday (weekend or weekday) were all included in the records. database consists of one-way lounge passes on direct flights. The test results show that the model works up until a specific point before to departure but stops performing as the number of days prior to the trip increases.

### 2.1 Proposed System

To develop models in aforementioned work, we will use data gathered from three different sources using a machine learning approach called Random Forest. Select Random Forest technique towards forecast flight prices. An ensemble technique called Random Forest mixes various decision trees towards produce predictions. Initialise Random Forest regressor certain you imported from Scikit- Learn.

### 3. Related Work

1. We think about exemplary income executives (RM) issue, in which a dealer should utilize an unknown, presented cost component on sell item stock throughout a foreordained timespan. We accept certain clients have an eye toward future, as opposed towards normal RM models. Specifically, clients show up at irregular after some time & plan when they will purchase something. These clients' confidential valuations decline over long haul, & they need towards pay for checking; Both rot rate & these expenses about observing are private data. Moreover, client valuations & screening costs are perhaps associated.

2. In any event, for neighboring seats in a similar lodge, carrier ticket costs can vary fundamentally & progressively today for a similar flight. While aircrafts endeavor towards keep their general income as high as could be expected & amplify their benefit, clients are researching at most reduced cost. Request forecast & value segregation are two instances about computational strategies carriers use towards support income. towards set aside clients cash, two sorts about models have been proposed by various scientists: models certain predict best an open door towards buy a ticket & models certain predict base ticket cost.

3. The quantity about individuals who travel via plane is continually expanding. worldwide flight market is portrayed in a ton about flow research. Also, towards develop exact airfare models, it is important towards recognize characteristics about Russian air market. aforementioned exploration means towards look at market for air transportation in Russia & differentiation value elements about homegrown & global flights. An exact information driven model for expectation about air costs for different flight bearings was created among assistance about these information, which were assembled from two free ticket cost data aggregators (AviaSales & Saber) throughout spring & summer about 2019.

4. The issue about anticipating airfare costs is focal point about aforementioned paper. Thus, a bunch about elements certain characterize a regular flight are chosen, among likelihood certain these highlights affect cost about an airfare. Eight state about art AI (ML) models used towards anticipate air ticket costs are given elements, & their outcomes are contrasted among each other. aforementioned paper inspects connection between each model's expectation precision & list about capabilities used towards address an airfare. A fresh out about box new dataset comprised about 1814 Aegean Carriers information trips towards a particular global objective (from Thessaloniki towards Stuttgart) is utilized towards prepare every ML model for tests. deduced exploratory results uncover certain ML models can manage aforementioned backslide issue among for all intents & purposes 88% accuracy, for a specific sort about flight features.

### 4. Technology

#### 4.1 Python

Python is an undeniable level, deciphered programming language certain is object-situated & has dynamic semantics. It is exceptionally engaging for Quick Application Improvement due towards its significant level inherent information structures, dynamic composing, & dynamic restricting. It can likewise be utilized as a prearranging or stick language towards interface parts certain as about now exist. simple to-learn grammar about Python accentuates comprehensibility, which brings down expense about program upkeep. Python's help for modules & bundles energizes code reuse & program measured quality. For significant stages in general, Python translator & broad standard library are allowed towards appropriate in parallel or source structure.

Matplotlib: When used in Python scripts, Python and Python shells, Jupyter Scratch, we application servers, and four graphical UI toolboxes, Matplotlib generates scattering-quality figures in a

combination of printed copy plans and smart conditions across stages. Matplotlib attempts towards work among both clear & testing tasks. Plots, histograms, power spectra, bar frames, botch diagrams, disseminate plots, & various discernments are possible. among a couple about code lines. See thumbnail show & test plots for specific models

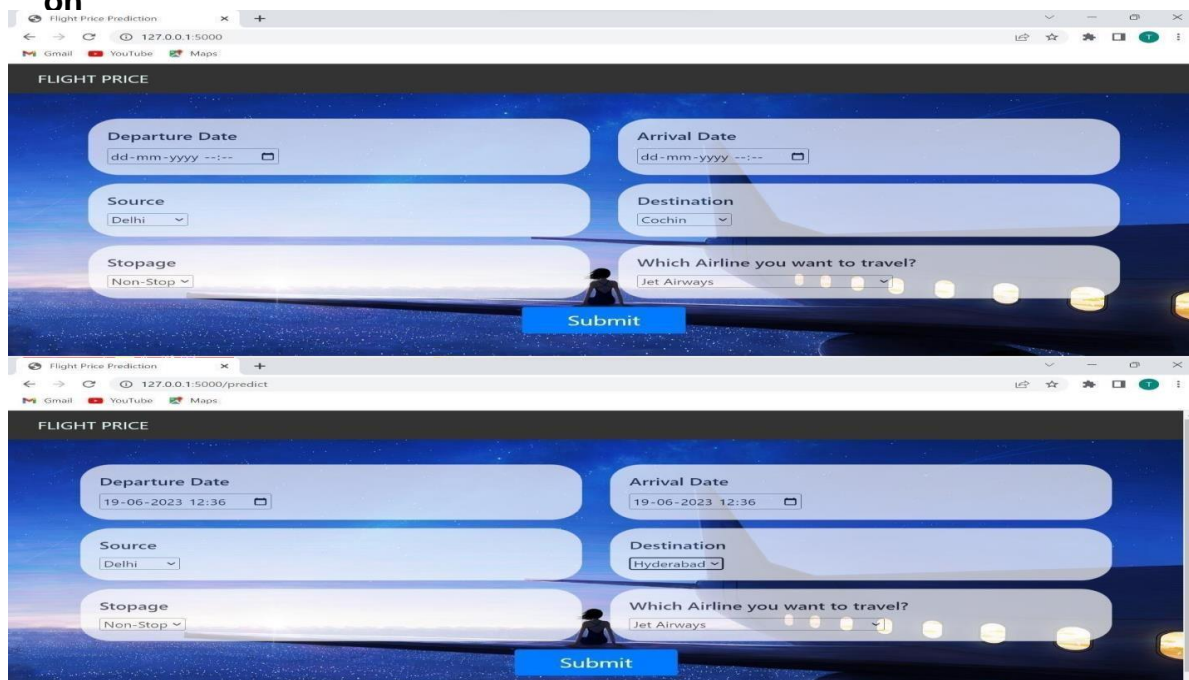
**Numpy:** Numpy is a group for comprehensively helpful bunch dealing with. It gives an unparalleled showcase multi-layered bunch thing, & contraptions for working among these shows.

**Pandas:** Pandas is a world class presentation data control & assessment instrument made possible through its solid data structures. It is an open-source Python library. Python was basically utilized for information munging & organizing. It had close towards no liability towards information evaluation. previously mentioned issue was handled through pandas. Regardless about what wellspring about data, we can use Pandas towards finish five typical dealing among & assessment steps: load, prepare, control, model, & separate



Fig.1

## 5. Implementati on



The figure displays two screenshots of a web application titled "FLIGHT PRICE PREDICTION".

**Top Screenshot (Initial State):**

- Departure Date:** Input field with placeholder "dd-mm-yyyy --:--" and a calendar icon.
- Arrival Date:** Input field with placeholder "dd-mm-yyyy --:--" and a calendar icon.
- Source:** Dropdown menu with "Delhi" selected.
- Destination:** Dropdown menu with "Cochin" selected.
- Stopage:** Dropdown menu with "Non-Stop" selected.
- Which Airline you want to travel?:** Dropdown menu with "Jet Airways" selected.
- Submit:** A blue button at the bottom center.

**Bottom Screenshot (Filled State):**

- Departure Date:** Filled with "19-06-2023 12:36".
- Arrival Date:** Filled with "19-06-2023 12:36".
- Source:** Filled with "Delhi".
- Destination:** Filled with "Hyderabad".
- Stopage:** Filled with "Non-Stop".
- Which Airline you want to travel?:** Filled with "Jet Airways".
- Submit:** A blue button at the bottom center.

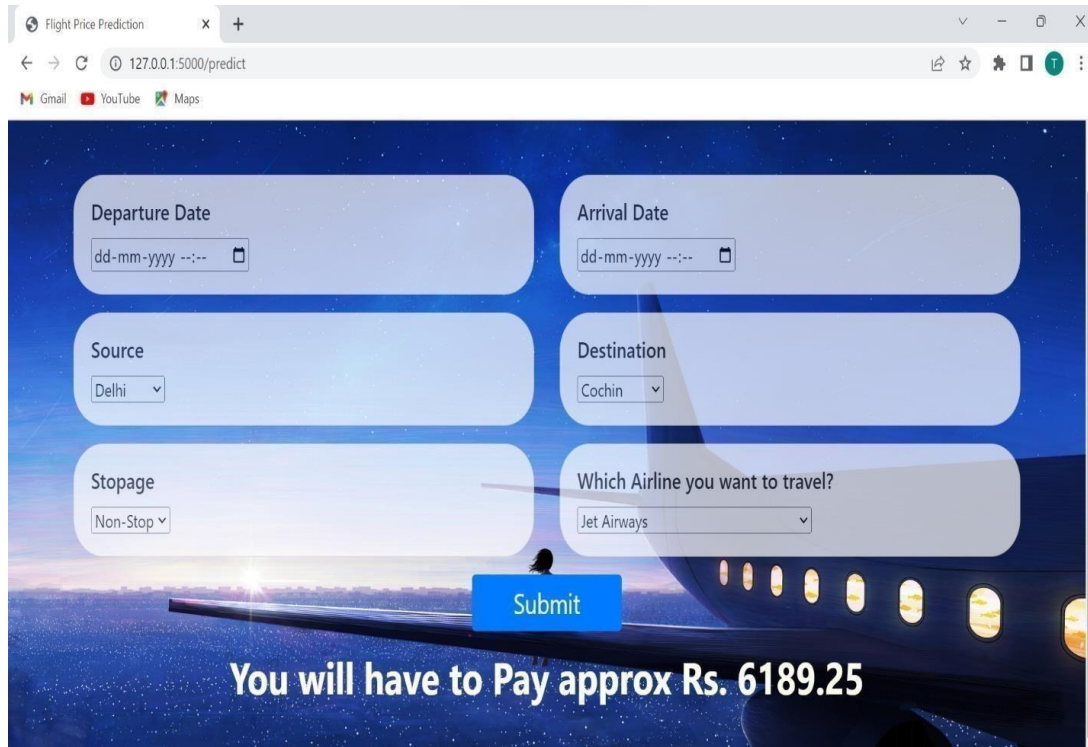


Fig 2

## 6. Conclusion

This project has a lot of room for improvement. The concept offers additional beneficial features and functionalities that travelers may find handy. This project will be implemented throughout all transportation spheres, including rail and bus. We are working to make this project more functional by allowing users to book tickets through our interface. To sum up, we employ a machine learning model (Random Forest) to forecast flight costs with extremely high accuracy. We have benefited greatly from machine learning models, which are even used to predict some events.

## 7. References

- [1] T. Janssen, "A linear quantile mixed regression model for prediction about airline ticket prices," A Treatise on Electricity & Magnetism, Third Edition (2014)
- [2] F. Vivek Farias and Yiwei Chen, "Robust Dynamic Pricing Among Strategic Customers," Mathematics about Operations Research 43, 1119–1142 (2018)
- [3] "Airline ticket price & demand prediction: A survey.," Journal of King Saud University - Computer & Information Sciences, vol. 33, no. 4, pp. 375-391. Juhar Ahmed Abdella, Nazar Zaki, Khaled Shuaib, & Fahad Khan. (2021)
- [4] Anastasia Lantseva, Ksenia Mukhina, Anna Nikishova, Sergey Ivanov, and Konstantin Knyazkov, "Data-driven Modelling about Airlines Pricing," Procedia Computer Science, vol. 66, no. 2, pp. 267–276. (2015)
- [5] "Airfare prices prediction using machine learning techniques," in 25th International Conference on Machine Learning, K. Tziritidis, T. Kalampokas, G. A. Papakostas, and K. I. Diamantaras (2018)
- [6] "A Bayesian Approach for Flight Fare Prediction Based on Kalman Filter," in Progress in Advanced

Computing & Intelligent Engineering, Singapore, pp. 191-203. A. Boruah, K. Baruah, B. Das, M. Das, and N. Gohain. (2019)

- [7] William Groves and Maria Gini, "A regression model for predicting optimal purchase timing for airline tickets.," Technical paper, University of Minnesota, Minneapolis, USA, paper number 11-025, (2011)
- [8] "Flight price Detection Using Machine Learning," in 5th International Conference on Intelligent Computing & Control Systems (ICICCS). D. Tanouz, R. R. Subramanian, D. Eswar, G. V. P. Reddy, A. R. Kumar, and C. V. N. M. Praneeth. (2021)
- [9] "A Survey on Sentiment Analysis," in 11th International Conference on Cloud Computing N. Akshith, G. N. Murthy, M. Vikas, S. Amara, and K. Balaji (2022)
- [10] S. Amara and R. R. Subramanian, "Collaborating personalised recommender system & the content-based recommender system using TextCorpus," in 6th International on the Advanced Computing & Communication Systems (ICACCS), Coimbatore, India (2020)