

Flight Ticket Price Predictor Using Python

Vivekanand P. Thakare, Ankita Sanjay Murraya, Roshani Bandu Gawade, Mrunali Mukundrao Sawarkar, Trupti Khemraj Shende, Ujjwala Kamlesh Badole

Department of Computer science & Engineering, GWCET, Nagpur, Maharashtra, India

Corresponding Author

E-Mail Id: nationalelectronics1100@gmail.com

ABSTRACT

Boxer timing for airline ticket purchasing from the consumer's perspective is challenging principally because buyers have insufficient information for reasoning about future price movements. In this project, we mainly directed to uncover underlying drifts of flight prices in India using historical data and also to advise the best time to buy a flight ticket. Remarkable, the trends of the price are highly sensitive to the route, month of departure, day of departure time of the day is a holiday and airline carrier. Highly competitive routes like most business routes hand a non-decreasing trend where prices increased as day to departure decreased

Keywords: *Feature selection, airfare price, prediction model, random forest, pricing models*

INTRODUCTION

For this project, we have gathered data from 18 routes all over India while the data of 4 routes were broadly used for the analysis due to the sheer volume of data collected over 4 months bring about 5.28 lakh data points each across the Mumbai-Delhi and Delhi-Mumbai route and 1.05 lakh data points crosswise the Delhi-Guwahati and Guwahati-Delhi route [1-4].

The project implements is the validations or contradictions towards myths regarding the airline industry, a comparison study among various models in predicting the optimal time to buy the flight ticket and the amount that can be saved if done so, A customized model which included [8] a combination of ensemble and statistical models have been implemented with the best accuracy of above 90% for a few routes, mostly from Tier to metro cities [7]. These models have led to substantial savings and formed average positive savings on every transaction.

BACKGROUND

Nowadays airline ticket prices can vary dynamically and significantly for some

flights even for nearby seats within the same cabin [6,9], customers are seeking to get the lowest price while airlines are trying to keep their overall revenue as high as possible and maximize their profit [3].

Airlines use various kinds of computational techniques to increase their revenue such as demand prediction and price discrimination. On the customer side [10], two kinds of models are proposed by different researchers to save money for customers models that predict the optimal time to buy a ticket and models that predict the minimum ticket price [4,5].

Time of purchase patterns[11] (making sure last-minute purchases are expensive)
2. Keeping the flight as full as they want it (raising prices on a flight that is filling up to reduce sales and hold back inventory for those expensive last-minute expensive purchases)

If we could inform the travelers about the boxer time to buy their flight tickets based on the historic data and show them various trends in the airline industry [11,12], we could help them save money on their

travels. This would be a practical implementation of data analysis, statistics [13], and machine learning techniques to solve a daily problem faced by travelers.

LITERATURE SURVEY

- K. Tziridis, T. Kalampokas It is hard for the client to buy an air ticket at the most reduced cost. For this few procedures are explored to determine the time and date to grab air tickets with a minimum fare rate [12,]. The majority of these systems are utilizing the modern computerized system known as Machine Learning. To determine the ideal purchase for a flight ticket.
- A. Sharma, since the deregulation of the airline industry, airfare pricing strategy has developed into a complex structure of sophisticated rules and mathematical models that drive the pricing strategies of airlines.
- Hoseinzade E, Haratizadeh S. (2019) Cnnpred: CNN-based stock market prediction using a diverse set of variables. Expert Syst Appl 129(SEP.):273–285 Traditional variables such as distance, although still playing a significant role, are no longer the sole factor that dictates the pricing strategy. Elements related to economic, marketing, and societal trends have played increasing roles in dictating airfare prices.
- T. Wang et al., "A Framework for Airfare Price Prediction: A Machine Learning Approach. A Linear Quantile Mixed Regression Model for Prediction of Airline Ticket Price [4].

ADVANTAGE

- Travelers get the fare prediction handy using which it's easy to decide on the airline.
- Save time in searching / deciding for the airline.

- Always available for making the reservation.
- Cancellation and changes of travel.

PROPOSED MODEL

- After we need data, we need to clean & prepare the Data according to the model's requirement.
- We used various statistical techniques & logic and implemented those using built-in packages.
- In this paper [4], random forest and decision tree algorithms are to find solutions for flight price problems in machine learning tasks, and a hybrid method is formulated from Chi-square, ANOVA, and Correlation tests are performed, departure and arrival features play an important role in predicting the price. Running the random forest and decision tree models, grouping the maximum price of airlines. Next, perform the feature engineering and calculate the accuracy
- In this article, we will be analyzing the flight fare prediction using a Machine Learning dataset using essential exploratory data analysis techniques then will draw some predictions about the price of the flight based on arrival time, what is the departure time, what is the duration of the flight, source, destination and more.
- Flight ticket prices can be something hard to guess, today we might see a price, check out the price of the same flight tomorrow, and it will be a different story.
- To solve this problem, we have been provided with prices of flight tickets for various airlines between January and March of 2021 and between various cities, using which we aim to build a model which predicts the prices of the flights using various input features.

METHOD

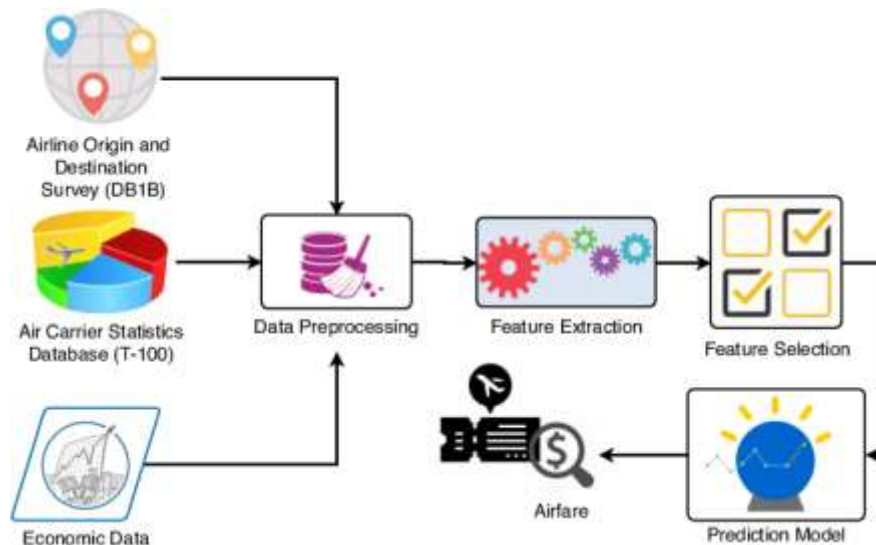


Fig. 1: Overview of the model.

AIMS & OBJECTIVE

A machine learning framework was developed to predict the quarterly average airfare price at the market segment level. Several features were extracted from the datasets and combined with macroeconomic data, to model the air travel market segments. With the help of the feature assortment methods, our proposed model can forecast the quarterly average airfare price with an adjustment. The cheapest available ticket for a given date gets more or less expensive over time.

APPLICATIONS

RavelSaathi is an AI-enabled Web Platform-based application. It serves as the day-to-day companion for Airport Authority individuals. It can be useful for both aviation professionals and their customers.

They can intelligently set the Fare price of flight tickets, can compare with competitive industries, and can also provide environmentally friendly and energy-saving solutions. More specifically[8], this system is designed to allow the Aviation industry to set competitive fare prices for their Flight service and obtain beneficiary margins;

along with monitoring Airport premises and conserving Energy.

This web App is cross-language compiled software [11]. Its machine learning algorithm and API are written in Python Language and its various Frameworks and Libraries.

These days, Python is covering almost all aspects of the IT industry Software production. HTML, CSS, and JavaScript have been used for the front-end UI interface and back-end connectivity has been handled by PHP language. Flask has been used for API building, and Wamp Server to host WEB APP on local windows PC.

FUTURE SCOPE

- More routes can be added and the same analysis can be expanded to major airports and travel routes in India.
- The analysis can be done by increasing the data points and increasing the historical data used that will train the model better giving better accuracies and more savings.
- More rules can be added to the Rule-based learning based on our

understanding of the industry also incorporating the offer periods given by the airline.

- Developing a more user-friendly interface for various routes giving more flexibility to the users.

CONCLUSION

From our thorough analysis of each of the 18 routes, we can define the following. Flight prices almost always remain constant or increase between the major cities.

Tourist routes that offer services involving Tier-2 cities of the country have uneven trends related to the increase and decrease of airline ticket prices. Routes with data collected over a longer duration of time tend to facilitate much more accurate predictions in the model and thus lead to higher average savings.

We were successfully able to analyze each route and generalize the entire project based on terms. We have also successfully busted some of the typical myths and misconceptions related to the airline industry and backed them up with data and analysis.

REFERENCES

1. Groves, W., & Gini, M. (2011). A regression model for predicting optimal purchase timing for airline tickets.
2. Liu, T., Cao, J., Tan, Y., & Xiao, Q. (2017, December). ACER: An adaptive context-aware ensemble regression model for airfare price prediction. In *2017 International Conference on Progress in Informatics and Computing (PIC)* (pp. 312-317). IEEE.
3. Boruah, A., Baruah, K., Das, B., Das, M. J., & Gohain, N. B. (2019). A bayesian approach for flight fare prediction based on kalman filter. In *Progress in Advanced Computing and Intelligent Engineering* (pp. 191-203). Springer, Singapore.
4. Wang, T., Pouyanfar, S., Tian, H., Tao, Y., Alonso, M., Luis, S., & Chen, S. C. (2019, July). A framework for airfare price prediction: a machine learning approach. In *2019 IEEE 20th international conference on information reuse and integration for data science (IRI)* (pp. 200-207). IEEE.
5. Papakostas, G. A., Diamantaras, K. I., & Papadimitriou, T. (2017). Parallel pattern classification utilizing GPU-based kernelized Slackmin algorithm. *Journal of Parallel and Distributed Computing*, 99, 90-99.
6. Francis, G., Fidato, A., & Humphreys, I. (2003). Airport–airline interaction: the impact of low-cost carriers on two European airports. *Journal of Air Transport Management*, 9(4), 267-273.
7. Koopmans, C., & Lieshout, R. (2016). Airline cost changes: To what extent are they passed through to the passenger?. *Journal of Air Transport Management*, 53, 1-11.
8. Lee, S., Seo, K., & Sharma, A. (2013). Corporate social responsibility and firm performance in the airline industry: The moderating role of oil prices. *Tourism management*, 38, 20-30.
9. Nair, V., & Hinton, G. E. (2010, January). Rectified linear units improve restricted boltzmann machines. In *Icml*.
10. Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.
11. Hong, D. H., & Hwang, C. (2003). Support vector fuzzy regression machines. *Fuzzy sets and systems*, 138(2), 271-281.
12. Gerardi, K. S., & Shapiro, A. H. (2009). Does competition reduce price dispersion? New evidence from the airline industry. *Journal of Political Economy*, 117(1), 1-37.

13. Guo, X., Zhao, Q., Zheng, D., Ning, Y., & Gao, Y. (2020). A short-term
14. Load forecasting model of multi-scale CNN-LSTM hybrid neural network considering the real-time electricity price. *Energy Reports*, 6, 1046-1053.