IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 5, May 2023

DOI: 10.17148/IJARCCE.2023.125140

Predicting The Price Of A Flight Ticket With The Use Of Machine Learning Algorithms

Pravar Umesh Ved (72170275K)[1], Abhishek Anant Waghmare (72170284J)[2], Nikita Navnath Thorat (72170285G)[3]

G H Raisoni College of Engineering & Management, Ahmednagar ·Gate No.: 1030, Village Chass, Nagar Pune Road, Ahmednagar, Maharashtra 414008[1-3]

Abstract: Flight ticket prices fluctuate based on factors such as flight timing, destination, and duration. To address the challenge of determining the optimal time to purchase tickets, this proposed system aims to develop a predictive model using machine learning algorithms. By analyzing historical flight data, our project focuses on identifying underlying price trends in India and providing recommendations for the best time to buy tickets.This project seeks to validate or debunk myths surrounding the airline industry, comparing different models to predict the optimal timing for ticket purchases and potential cost savings. Notably, price trends vary significantly depending on the route, month, day, time of departure, whether it's a holiday, and the airline carrier. For highly competitive routes like major business destinations (e.g., Mumbai-Delhi), prices tend to increase as the departure date approaches. However, other routes, such as tier 1 to tier 2 cities like Delhi-Guwahati, have specific time frames when prices are at their lowest. Additionally, the collected data reveals two distinct categories of airline carriers in India: the economical group and the luxurious group. In most cases, the lowest-priced flights belong to the economical group. Furthermore, the data confirms that certain periods of the day are associated with higher expected prices. Expanding the scope of this project to cover various routes can lead to significant savings when purchasing domestic flight tickets in the Indian market..

Keywords: Flight ticket, Optimal timing, historical data, competitive routes, Indian Domestic Airline market

I. INTRODUCTION

The flight ticket purchasing process often suggests buying tickets well in advance of the departure date to avoid higher prices. However, airlines may not always adhere to this practice. They may lower prices to stimulate market demand or increase prices when tickets become scarce. Therefore, ticket prices can depend on various factors. This project utilizes AI to model the price patterns of flight tickets over time.Airlines have the flexibility to modify ticket prices at any given moment. Travelers can save money by booking tickets at the lowest available prices. Frequent flyers are typically aware of the price fluctuations. Airlines employ sophisticated Revenue Management policies to implement various pricing strategies. These strategies involve adjusting fares based on factors such as time, season, and holidays. It is not uncommon for airlines to modify the header or footer on subsequent pages to reflect these changes.The primary objective of airlines is to generate profits, while customers seek the lowest possible rates. Customers often attempt to purchase tickets well in advance of the departure date to avoid inflated airfares as the date approaches. However, in reality, this may not always hold true. Customers may end up paying more than necessary for the same seat.

II. LITERATURE REVIEW

1) K. Tziridis T. Kalampokas G.Papakostas and K. Diamantaras "Airfare price prediction using machine learning techniques" in European Signal Processing Conference (EUSIPCO), DOI: 10.23919/EUSIPCO .2017.8081365L. Li Y. Chen and Z. Li" Yawning detection for monitoring driver fatigue based on two cameras" Proc. 12th Int. IEEE Conf. Intel. Transp. Syst. pp. 1-6 Oct. 2009.

In the study titled "Airfare Price Prediction Using Machine Learning Techniques" , the researchers focused on predicting airfare prices by employing machine learning methods. They gathered a dataset consisting of 1814 flight records from Aegean Airlines, which was used to train their machine learning models. To explore the impact of feature selection on model accuracy, they experimented with different combinations of features. Several machine learning algorithms were utilized in their study, including Multilayer Perceptron (MLP), Generalized Regression Neural Network, Extreme Learning Machine (ELM), Random Forest Regression Tree, Regression Tree, Bagging Regression Tree, Regression SVM (Polynomial and Linear), and Linear Regression (LR). Each algorithm produced distinct outputs and results when applied to the dataset.

DOI: 10.17148/IJARCCE.2023.125140

The researchers conducted extensive training and testing of various models by incorporating and removing features to observe the effects on prediction accuracy. Their aim was to identify the most effective model for airfare price prediction. Further details and insights regarding their experimental setup and findings can be found in the referenced study.

2) William Groves and Maria Gini "An agent for optimizing airline ticket purchasing" in proceedings of the 2013 international conference on autonomous agents and multi-agent systems

In the case study conducted by William Groves [2], an agent is introduced with the capability to optimize the timing of ticket purchases on behalf of customers. The study utilizes the technique of Partial Least Square (PLS) regression to build a predictive model. Initially, various techniques for feature selection are employed, including Feature Extraction, Lagged Feature Computation, Regression Model Construction, and Optimal Model Selection. The experiments conducted in the study aim to estimate the real-world costs associated with using the prediction models. The lag scheme approach is found to work well with many machine learning algorithms, but in the specific domain of flight fare prediction, PLS regression is identified as the most effective technique. The improved performance of PLS regression can be attributed to its inherent ability to handle collinear and irrelevant variables, which are commonly encountered in this domain..

3) J. Santos Dominguez-Menchero, Javier Rivera and Emilio Torres Manzanera "Optimal purchase timing in the airline market".

In this paper, the researchers conducted a study on the general patterns in airline pricing behavior and developed a methodology for analyzing different routes and carriers. The main objective was to provide customers with relevant information to help them determine the optimal time to purchase a ticket, considering both cost savings and any time constraints they may have. The study highlights the effectiveness of non-parametric isotonic regression techniques in analyzing airline pricing. These techniques prove to be particularly useful compared to standard parametric methods. By utilizing these techniques, the researchers were able to determine the timeframe within which consumers can delay their ticket purchase without experiencing significant price increases. Additionally, they were able to quantify the economic loss incurred for each day of delay and identify instances where waiting until the last day to make a purchase is preferable. Overall, the study contributes valuable insights into understanding airline pricing dynamics and provides practical information to assist customers in making informed decisions about their ticket purchases.

4) Supriya Rajankar, Neha sakhrakar and Omprakash rajankar “Flight fare prediction using machine learning algorithms”
International journal of Engineering Research and Technology (IJERT) June 2019.

In the survey conducted by Supriya Rajankar, the focus was on flight fare prediction using machine learning algorithms. The dataset used in the study consisted of flights between Delhi and Bombay. Various machine learning algorithms were applied, including K-nearest neighbors (KNN), linear regression, and support vector machine (SVM), to obtain different outcomes and analyze their performance. To predict flight ticket prices, several machine learning algorithms were implemented, such as Support Vector Machine (SVM), Linear Regression, K-Nearest Neighbors, Decision Tree, Multilayer Perceptron, Gradient Boosting, and Random Forest. These models were implemented using the scikit-learn library in Python. Performance evaluation of these models was conducted using parameters like R-square, MAE (Mean Absolute Error), and MSE (Mean Squared Error). Based on the evaluation, the results indicated that the Decision Tree algorithm yielded the best performance among the tested models. Further details and insights about the experimental setup and findings can be found in Supriya Rajankar's journal.

5) Tianyi wang, samira Pouyanfar, haiman Tian and Yudong Tao "A Framework for airline price prediction: A machine learning
approach"

In the paper authored by Tianyi Wang, Samira Pouyanfar, Haiman Tian, and Yudong Tao [5], a framework is proposed that combines two databases along with macroeconomic data to model the average ticket price based on source and destination pairs. The researchers employed machine learning algorithms, specifically Support Vector Machine (SVM) and XGBoost, to perform the modeling task. The framework presented in the study achieved a high prediction accuracy of 0.869, as measured by the adjusted R-squared performance metric. Additionally, the researchers reported the lowest error rate of 0.92 when utilizing the XGBoost algorithm.

DOI: 10.17148/IJARCCE.2023.125140

6) T. Janssen "A linear quantile mixed regression model for prediction of airline ticket prices"

In this paper, the researchers focused on predicting the optimal time to purchase flight tickets. They employed a variety of machine learning algorithms, including linear regression, Decision Tree, Random Forest, K-Nearest Neighbor, Multilayer Perceptron (MLP), gradient boosting, support vector machine (SVM), Naïve Bayes, and Stacked Prediction Model. To develop the desired model, the researchers utilized the Linear Quantile Blended Regression methodology for the San Francisco–New York route. They obtained daily airfare data from an online website. Two features were considered in the model development: the number of days until departure and whether the departure date fell on a weekend or weekday.

7) Wohlfarth, T.clemencon, S.Roueff “A Data mining approach to travel price forecasting” 10th international conference on machine learning Honolulu 2011. In the research paper authored by Wohlfarth, T. Clemencon, and S. Roueff , the focus is on the development of a flight fare prediction system using the technique of yield management in the air travel industry. The researchers employ various data mining techniques in their study. The primary objective of the paper is to design decision-making tools that take into account the fluctuating travel prices from the customer's perspective. The paper introduces terms such as machine techniques or algorithms, specifically mentioning clustering.

8) Vinod Kimbhaune, Harshil Donga, Ashutosh Trivedi, Sonam Mahajan and Viraj Mahajan research paper on flight fare prediction system.

In the research paper authored by Vinod Kimbhaune, Harshil Donga, Ashutosh Trivedi, Sonam Mahajan, and Viraj Mahajan , the focus is on developing a flight fare prediction system using various machine learning algorithms. The algorithms employed in the study include Random Forest, Decision Tree, and Linear Regression. The objective is to determine the ideal time for purchasing flight tickets. The researchers aimed to create an application that predicts flight prices for different flights using a machine learning model. The techniques utilized in the study are specifically mentioned as Linear Regression, Decision Tree, and Random Forest. Performance metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE) were used to evaluate the performance of the models. The outcome of their project was not entirely accurate; however, the researchers suggest that incorporating more real-time datasets would yield more precise results.

9) W. Groves and M. Gini, ―An agent for optimizing airline ticket purchasing, ‖ 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2013), St. Paul, MN, May 06 - 10, 2013, pp. 1341-1342.
In the case study conducted by William Groves , an agent is introduced with the capability to optimize the timing of ticket purchases on behalf of customers. The study utilizes the technique of Partial Least Square (PLS) regression to build a predictive model. Initially, various techniques for feature selection are employed, including Feature Extraction, Lagged Feature Computation, Regression Model Construction, and Optimal Model Selection. The experiments conducted in the study aim to estimate the real-world costs associated with using the prediction models. The lag scheme approach is found to work well with many machine learning algorithms, but in the specific domain of flight fare prediction, PLS regression is identified as the most effective technique. The improved performance of PLS regression can be attributed to its inherent ability to handle collinear and irrelevant variables, which are commonly encountered in this domain.

10) Viet Hoang Vu, Quang Tran Minh and Phu H. Phung,‖An Airfare Prediction Model for Developing Markets‖, IEEE paper 2018.
In this paper, the researchers propose a novel model that enables buyers to predict price trends in the absence of official information from airlines. Despite the lack of key features such as the number of unsold seats on flights, the model leverages publicly available airfare data to accurately predict trends and actual changes in airfare up to the departure dates. The researchers also identify the features that have the most significant impact on airfare changes, providing valuable insights into the pricing dynamics. They introduce a ticket purchasing time improvement model that incorporates various preprocessing techniques, data mining frameworks (including course of action and grouping), and statistical analysis methods. The proposed framework aims to transform multiple value-added arrangements into a cohesive value heading, facilitating more informed decision-making. This value heading is constructed based on comparative pricing behavior, and the framework utilizes a tree-based analysis to select the optimal timing group for comparison. The researchers then analyze the price change patterns within the selected group.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 5, May 2023

DOI: 10.17148/IJARCCE.2023.125140

PROPOSED SYSTEM

Following is the basic proposed system:

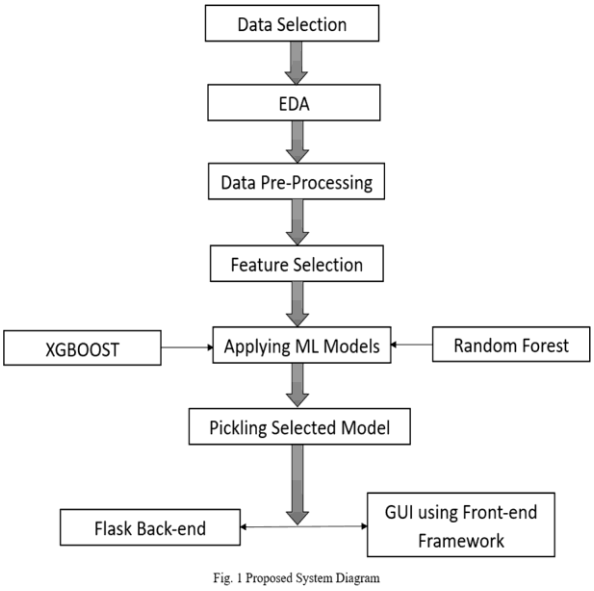

Fig. 1 Proposed System Diagram

CONCLUSION

The dataset is utilized to apply machine learning algorithms for predicting the dynamic fares of flights. This enables the estimation of flight ticket prices at the minimum cost. Data is sourced from websites that offer flight ticket sales, limiting the available information. The accuracy of the model is determined by the R-squared values obtained from the algorithm. Access to additional data in the future, such as current seat availability, would enhance the accuracy of the predictions. Ultimately, the entire process of predicting airline ticket prices has been developed, and the validity of these predictions is supported by previous trends and our own predictive analysis.

REFERENCES

[1] K. Tziridis T. Kalampokas G.Papakostas and K. Diamantaras "Airfare price prediction using machine learning techniques" in European Signal Processing Conference (EUSIPCO), DOI: 10.23919/EUSIPCO .2017.8081365L. Li Y. Chen and Z. Li" Yawning detection for monitoring driver fatigue based on two cameras" Proc. 12th Int. IEEE Conf. Intel. Transp. Syst. pp. 1-6 Oct. 2009.

[2] William Groves and Maria Gini "An agent for optimizing airline ticket purchasing" in proceedings of the 2013 international conference on autonomous agents and multi-agent systems.

[3] J. Santos Dominguez-Menchero, Javier Rivera and Emilio TorresManzanera "Optimal purchase timing in the airline market".

[4] Supriya Rajankar, Neha sakhrakar and Omprakash rajankar "Flight fare prediction using machine learning algorithms" International journal of Engineering Research and Technology (IJERT) June 2019.

[5] Tianyi wang, samira Pouyanfar, haiman Tian and Yudong Tao "A Framework for airline price prediction: A machine learning approach"

[6] T. Janssen "A linear quantile mixed regression model for prediction of airline ticket prices"

[7] Wohlfarth, T.clemencon, S.Roueff "A Dat mining approach to travel price forecasting" 10th international conference on machine learning Honolulu 2011.

[8] Vinod Kimbhaune, Harshil Donga, Ashutosh Trivedi, Sonam Mahajan and Viraj Mahajan research paper on flight fare prediction system.

[9] W. Groves and M. Gini, ―An agent for optimizing airline ticket purchasing, ‖ 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2013), St. Paul, MN, May 06 - 10, 2013, pp. 1341-1342.

[10] Viet Hoang Vu, Quang Tran Minh and Phu H. Phung,‖An Airfare Prediction Model for Developing Markets‖, IEEE paper 2018.