e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:10/October-2023 Impact Factor- 7.868 www.irjmets.com

AIRFARE PRICE PREDICTION USING MACHINE LEARNING

Ajit Phulmante[*1], Avishkar khankare[*2], Pratik Kulkarni[*3],
Komal Gosavi[*4], Prof. Pragati Deole[*5]

[*1,2,3,4,5]Department Of Computer Engineering, Smt. Kashibai Navale College Of Engineering,
Affiliated To Savitribai Phule Pune University, Pune, Maharashtra, India.

DOI : https://www.doi.org/10.56726/IRJMETS45709

ABSTRACT

Airfare price prediction is a critical application of machine learning in the realm of travel and aviation. The dynamic and often unpredictable nature of airline ticket pricing presents challenges for travelers seeking cost-effective options. This study explores the use of machine learning techniques to predict airfare prices, offering travelers the ability to make informed decisions about their flight bookings. Leveraging historical pricing data, this research employs advanced algorithms to model airfare fluctuations, taking into account factors such as booking time, route, airline, and external variables. By constructing accurate predictive models, this approach empowers travelers to optimize their travel plans and potentially save money. This abstract provides an overview of the methodology and key findings in airfare price prediction, highlighting its relevance in enhancing the travel experience.

Keywords: Machine Learning, Analysis, Prediction, Airfare, Research.

I. INTRODUCTION

Air travel is an essential part of modern life, connecting people and businesses across the globe. However, the cost of airfare can fluctuate significantly, making it challenging for travelers to find the best deals and plan their journeys efficiently. Airfare price prediction, powered by machine learning, has emerged as a powerful tool to address this challenge. This introduction sets the stage for understanding the significance of airfare price prediction and its relevance in the travel industry. It highlights the key elements of airfare prediction, the benefits it offers to both travelers and airlines, and the role of machine learning in revolutionizing this field.

The Challenge of Airfare Pricing : The pricing of airline tickets is a complex and dynamic process influenced by a multitude of factors. These factors include demand, seasonality, route popularity, airline competition, and external events like holidays or global crises. As a result, airfare prices can change rapidly, and travelers often struggle to determine when and where to book their flights to secure the most cost-effective options.

The Role of Airfare Price Prediction : Airfare price prediction addresses this challenge by leveraging historical pricing data, advanced algorithms, and machine learning techniques to forecast future ticket prices. The primary goal is to empower travelers with insights that enable them to make informed decisions about their flight bookings. This includes identifying the optimal time to purchase tickets to secure the best fares.

Benefits of Airfare Price Prediction:

1. Cost Savings: By predicting airfare fluctuations, travelers can avoid overpaying for flights, potentially saving significant amounts of money on their journeys.

2. Booking Confidence: Travelers can have greater confidence in their booking decisions, knowing they are making data-driven choices.

3. Flexible Planning: Predictive models enable travelers to plan their trips more flexibly, choosing the most favorable times and routes.

4. Airline Competitiveness: Airlines can use price prediction models to stay competitive by adjusting their fares in response to market dynamics. The Role of Machine Learning plays a pivotal role in airfare price prediction. These algorithms analyze historical pricing data, along with a myriad of variables such as booking time, route, and airline, to identify patterns and trends. This data-driven approach enables the creation of accurate predictive models capable of estimating future ticket prices. Structure of this Guide In this guide, we will explore the methodologies, techniques, and tools used in airfare price prediction using machine learning. We will delve into the data sources, feature engineering, model selection, and evaluation metrics employed to build robust

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:10/October-2023 Impact Factor- 7.868 www.irjmets.com

predictive models. Additionally, we will discuss ethical considerations and practical implementation aspects of these models. As airfare price prediction continues to evolve, it offers travelers and airlines alike the opportunity to navigate the complex world of air travel with greater confidence and efficiency. By harnessing the power of machine learning, we can unlock the potential for smarter, more cost-effective airfare decisions, ultimately enhancing the travel experience for all.

II. METHODOLOGY

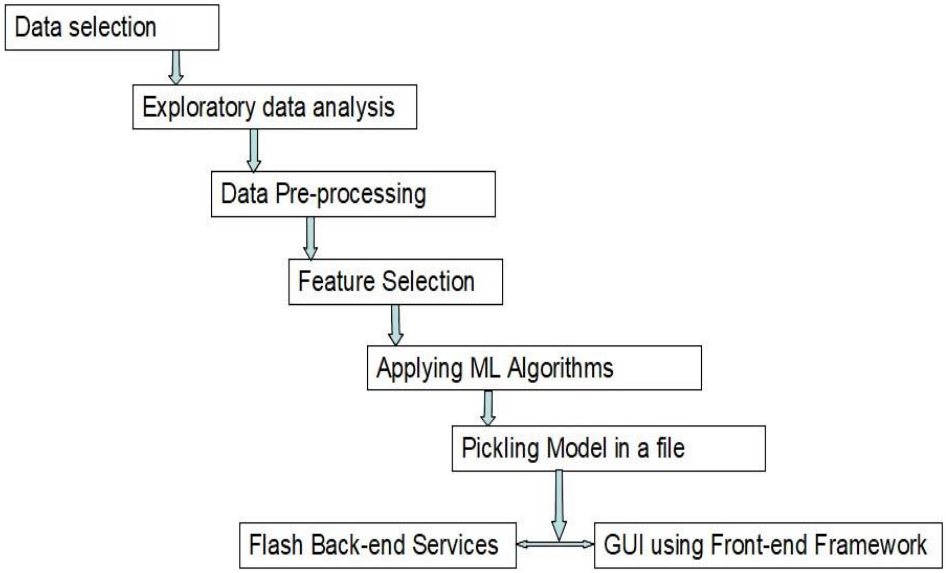

Figure 1: METHODOLOGY

Dataset

To build an airline ticket price model at the market segment level, we need data related to aircraft traffic and the number of passengers in each market segment. For our training and evaluation, we are utilizing data from the year 2018. The dataset used for this purpose was generously provided by Kaggle, a freely accessible platform catering to data scientists and machine learning enthusiasts.

Data selection

The initial phase involves the acquisition of historical flight data, which is essential for training the algorithm used in price prediction. Our dataset encompasses a rich repository of more than 10,000 data points pertaining to various flight details and corresponding pricing information. This comprehensive dataset includes key attributes such as origin, destination, departure date, departure time, stopovers, arrival time, prices, and several other relevant parameters. In the exploratory data analysis phase, we diligently performed data cleansing operations by eliminating duplicate entries and handling null values. The removal of these discrepancies is crucial to maintaining the model's accuracy and reliability. Furthermore, as part of our data exploration efforts, we delved into the distribution of the dataset and incorporated additional informational insights.

Exploratory data analysis

All the attributes within our dataset are readily available for analysis. Leveraging the mathematical capabilities of libraries, we explore the interrelationships among the data points. We employ libraries that offer various visualization tools such as scatter plots, histograms, and other graphical representations to gain insights into the connections between the independent and dependent variables. These visualizations aid in illustrating the rationale behind our selection of independent and dependent variables in the analysis.

Feature selection

In the process of feature selection, we aim to pinpoint the essential characteristics that bear a stronger correlation with pricing. Some attributes, such as extraneous details and routing information, are deemed redundant and have the potential to undermine the accuracy of the models deployed for the group's benefit. Once we've identified the attributes closely linked to pricing, the subsequent step entails constructing a model through the utilization of machine learning algorithms. Given that our dataset comes equipped with labeled data, our

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:10/October-2023 Impact Factor- 7.868 www.irjmets.com

approach will involve the application of supervised machine learning techniques. Additionally, since our dataset comprises continuous feature values, we'll be harnessing support vector machines, random forests, and decision tree algorithms in our analysis.

Applying ML algorithm

Within our project scope, we are conducting comprehensive testing of multiple algorithms to ascertain their respective levels of efficiency. Specifically, we are putting the support vector machine, random forest algorithm, and decision tree algorithm to the test.

III. METHODS AND MATERIAL

Data Collection:

Describe the data sources you used to gather information relevant to airfare prediction. This could include historical flight pricing data, airline schedules, booking records, and external factors such as economic indicators or weather data.

Specify the time period covered by your dataset and explain why you chose this timeframe.

Data Preprocessing:

Explain the steps you took to clean and prepare the data for analysis. This might involve handling missing values, outlier detection and treatment, and data normalization.

Detail any transformations or feature engineering you conducted to create relevant variables for modeling.

Feature Selection:

Describe the process of selecting the features (variables) you used to build your prediction model. Explain the rationale for including or excluding specific features.

Mention any domain-specific knowledge that influenced your feature selection.

Modeling Approach:

Explain the machine learning or statistical techniques you employed to predict airfare prices. Common methods include regression models, time series analysis, or machine learning algorithms like Random Forests or Neural Networks.

If you used a specific library or software for modeling (e.g., scikit-learn, TensorFlow, or R), mention it here.

Model Training:

Detail how you split your data into training and testing sets to train and evaluate your model's performance.

Discuss any hyperparameter tuning or cross-validation procedures you used to optimize your model.

Evaluation Metrics:

Specify the metrics you used to assess the accuracy and performance of your airfare prediction model. Common metrics include Mean Absolute Error (MAE), Root Mean Square Error (RMSE), or R-squared (R^2).

Explain why you chose these metrics and their significance in the context of airfare prediction.

Software and HardwareTools:

Software Requirements Specification

Coding Language: Python

Operating System: Windows 10

Hardware Requirements Specification

Processor: Pentium-IV

RAM: 512 MB(min)

Hard Disk: 40 GB

Key Board: Standard Windows Keyboard

Mouse: Two or Three Button Mouse

Monitor: LCD/LED

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:10/October-2023 Impact Factor- 7.868 www.irjmets.com

IV. RESULTS AND DISCUSSION

The "Results and Discussion" section of a research paper or report on airfare price prediction using machine learning is where you present the outcomes of your study and interpret their significance. It's a critical part of your research, as it helps readers understand the effectiveness of your predictive models and the implications of your findings. Below, I'll provide a structured outline for this section: Performance Evaluation of Predictive Models: Present the results of your machine learning models for airfare price prediction. Include relevant evaluation metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), R-squared (R^2), or any other metrics you used. Provide a summary table or visualization to compare the performance of different models if applicable. Accuracy and Precision: Discuss the accuracy and precision of your predictive models in forecasting airfare prices. Analyze any discrepancies between predicted and actual prices. Explain the implications of the model's accuracy for travelers and the travel industry. Model Interpretation: If your machine learning model allows for feature importance analysis, discuss the key factors driving airfare price predictions. Highlight which features or variables have the most significant impact on pricing. Discuss how this information can be valuable for both travelers and airlines. Temporal Analysis: If relevant, present a temporal analysis of airfare predictions over time. Identify any patterns, trends, or seasonality in airfare prices. Discuss the practical implications of temporal variations in airfare pricing. Case Studies or Examples: Provide real-world case studies or examples of successful airfare price predictions made by your model. Share stories of travelers who benefited from your predictions or scenarios where your model's insights proved valuable. Discussion of Ethical Considerations: Reflect on any ethical considerations related to airfare price prediction, such as the responsible use of travelers' data or potential biases in the model. Discuss steps taken to address ethical concerns and ensure fair and transparent predictions. Limitations and Future Research: Acknowledge the limitations of your study, including any challenges or constraints you encountered. Suggest areas for future research and improvements to enhance airfare price prediction models. Comparison with Existing Methods: If applicable, compare the performance of your machine learning approach with traditional methods or existing commercial solutions. Highlight the advantages and disadvantages of your approach. Practical Implications: Discuss how the results of your study can be practically applied in the travel industry or by travelers themselves. Offer recommendations for travelers on how to use airfare price predictions effectively. Conclusion: Summarize the key findings of your study and their significance in the context of airfare price prediction. Conclude with a concise statement about the overall success and contributions of your research. In the "Results and Discussion" section, it's essential to provide a clear and comprehensive analysis of your findings and their implications. This section should help readers understand the practical value of your machine learning models in the domain of airfare price prediction.

V. CONCLUSION

Airfare prediction using machine learning holds the promise of enhancing travel experiences for both individuals and airlines. By leveraging historical data and advanced algorithms, these models can provide insights into future airfare trends, aiding travelers in making informed decisions and airlines in optimizing their revenue strategies. As machine learning techniques continue to advance, the accuracy of airfare predictions is expected to improve, leading to a more efficient and predictable air travel industry

VI. REFERENCES

[1] Machine Learning-Assisted Flight Ticket Price Prediction ISSN 2278-3091 Volume 10, No.2, March - April 2021. Jibin Joseph1, Abhijith P2, Aryasree S3, Jinsu Anna Joseph4, Meghana Sara Oommen5, Abin T Abraham6

[2] Dhief, I., Wang, Z., Liang, M., Alam, S., Schultz, M., and Delahaye, D., “Predicting Aircraft Landing Time in Extended-TM Using Machine Learning Methods,” Proceedings of 9th International Conference for Research in Air Transportation (ICRAT), Tampa, FL, USA, 2020.

[3] Machine Learning Techniques for Predicting Airfare Prices Based on Reviews Ruchi K. Oza, Aakanksha V. Jain, Aanal S.Raval ISSN: 2319-8753 ISSN: 2347-6710 Vol. 9, Issue 3, March 2020

[4] S. Sutthithatip, S. Perinpanayagam, and S. Aslam, "(Explainable) artificial intelligence in aerospace safety-critical systems," in Proc. IEEE Aerosp. Conf. (AERO), Mar. 2022, pp. 1–12.

[5] R. R. Subramanian, M. S. Murali, B. Deepak, P. Deepak, H. N. Reddy, and R. R. Sudharsan, "Airline fare prediction using machine learning algorithms," in Proc. 4th Int. Conf. Smart Syst. Inventive Technol. (ICSSIT),Jan. 2022, pp. 877–884

[6] A Machine Learning Approach to Airfare Price Prediction Tianyi Wang, Samira Pouyanfar, Haiman Tian, Yudong Tao, Miguel Alonso Jr., Steven Luis, and Shu-Ching Chen.

[7] Prediction of Airline Prices Using Machine Learning Algorithms Abhilash1, Ranjana Y2, Shilpa S3, Zubeda A Khan4 are the authors. ISSN 2321-2004 (Online) Vol. 7, Issue 2, February 2019 ISSN (Print) 2321-5526 IJIREEICE copyright DOI 10.17148/IJIREEICE.2019.7203 14.

[8] Prediction of airline ticket prices and demand: A survey Juhar Ahmed Abdella, NM Zaki, Khaled Shuaib, and Fahad Khan King Saud University Journal of Computer and Information Sciences 33 (2021) 375–391.

[9] A Survey on Machine Learning for Flight Pricing Prediction, International Journal of Engineering Research & Technology (IJERT) http://www.ijert.org ISSN: 2278-0181 IJERTV8IS060672 Vol. 8 Issue 06, June-2019 Supriya Rajankar, Neha Sakharkar.