Performance Evaluation of Machine Learning Algorithm's for Flight Price Prediction

Shreyas M
Department of Computer Science and Engineering
Global Academy of Technology
Bengaluru, India
mshreyas336@gmail.com

Shridhar Devaramane
Department of Computer Science and Engineering
Global Academy of Technology
Bengaluru, India
devamaneshridhar@gmail.com

Divith P
Department of Computer Science and Engineering
Global Academy of Technology
Bengaluru, India
cddivith89@gmail.com

Mamatha R
Department of Computer Science and Engineering Global Academy of Technology Bengaluru, India
mamatharmanjunath@gmail.com

Vikas Shanabhog
Department of Computer Science and Engineering Global Academy of Technology Bengaluru, India
vikasshanabhog0@gmail.com

Trisha S
Department of Computer Science and Engineering
Global Academy of Technology
Bengaluru, India
trushasudhakar001@gmail.com

Abstract— Since the demand for air travel in India is increasing due to the purchase of many flight tickets online, passengers are trying to understand how these airline corporations determine the cost of flight tickets over time. There are numerous tactics that let you act when it's appropriate. While travelers seek the lowest priced ticket, airlines want to maintain the highest possible revenue per unit to optimize their profit margin. Airlines employ a variety of computational strategies, including pricing discrimination and demand forecasting, to boost income. This is for the customer who estimates the cost of the flight while purchasing a ticket. From the standpoint of the customer, the most challenging aspect is figuring out when to buy tickets at the best price or with the most value. Most of the methods rely on prediction models, machine learning (ML), and sophisticated v

computational intelligence. This study highlights the elements and offers guidance on creating an aviation fare prediction model based on machine learning.

Keywords - Machine Learning Algorithms, airfare, supervised learning, predictions, flight, Linear Regression, Artificial Neural Network, Random Forest.

I. Introduction

Accurate flight price prediction is essential in today's dynamic and competitive airline business as it influences travel decisions made by industry stakeholders and consumers alike. Precisely predicting ticket pricing is critical since a wide range of factors, such as demand, fuel prices, route popularity, and market rivalry, affect airfare variations. The creation of reliable prediction models has become a crucial field of study and innovation as tourists look to maximize their trip budgets and airlines try to maximize profits.

The field of flight cost prediction has seen a complete transformation with the introduction of machine learning techniques and data analytics breakthroughs. Through the utilization of extensive historical pricing data and current market information, scholars and professionals in the field can now apply advanced algorithms to predict fare trends with previously unheard-of precision. Many approaches have been investigated to capture the complex patterns driving airfare dynamics, ranging from conventional regression models to state-of-the-art deep learning architectures.

Our goal in this research study is to improve forecast accuracy and dependability in the face of changing market dynamics by conducting a thorough investigation of flight price prediction approaches. We explore the nuances of several machine learning techniques and evaluate how well they predict changes in fares in a variety of temporal and spatial contexts. To further deepen our understanding of the complex nature of airfare prediction, we also examine the impact of outside variables on ticket prices, such as seasonality, economic indicators, and geopolitical events.

We aim to determine the most efficient methods for flight price prediction through comparative review and empirical analysis, providing information that can help travelers and airline operators make well-informed decisions. We hope to add to the current conversation on airline pricing strategy optimization by illuminating the advantages and disadvantages of the prediction models that are currently in use.

II. Related Work

In [1], Presents a study on flight price forecasting using machine learning. Most likely, the study goes over the statistics, algorithms, and methods used to forecast flight costs. It could also go into detail on the assessment metrics that are employed to gauge how well the predictive models work. The study may also emphasize the value of precise pricing prediction in the airline sector as well as possible uses for the created models. This work can be summarized to help develop the subject of flight price prediction by providing academics with insights on the challenges and existing methodologies in machine learning.

In [2], The use of yield management techniques by American Airlines is covered in "Yield Management at American Airlines" by Smith, Leimkuhler, and Darrow. The study probably looks at how yield management strategies, which dynamically alter ticket pricing in response to capacity and demand, maximize income. It might include information about the techniques used by American Airlines to estimate demand, divide up its clientele, and determine pricing. The writers most often use case studies or actual data to show how well these tactics work to maximize profits. Gaining an understanding of the ideas presented in this paper can help airline pricing and revenue management methods.

In [3], Groves and Gini's paper "An Agent for Optimizing Airline Ticket Purchasing" presents an agent-based system intended to maximize airline ticket purchases. The agent's design and functionality—which are intended to assist customers in finding the best airfare—are probably covered in the paper. It might explain how the agent looks for the best routes and price alternatives based on user preferences using algorithms and data analysis. It is probable that the writers showcase simulations or experimental findings that highlight the agent's efficiency in enhancing ticket acquisition choices. Gaining an understanding of this research can help in applying agent-based systems to improve customer decision-making in the airline sector.

In [4], A unique method for predicting airline ticket prices using a linear quantile mixed regression model is presented in "A Linear Quantile Mixed Regression Model for Prediction of Airline Ticket Prices" by Janssen et al. The approach of this model, which accounts for a variety of factors influencing ticket prices using both fixed and random effects, is probably described in the paper. The benefits of quantile regression for capturing various degrees of price volatility might be covered. The authors most likely provide empirical data or validation tests showing how well their suggested model predicts ticket prices. Comprehending this study can provide valuable perspectives on sophisticated statistical methods for predicting airline ticket costs, which could enhance the decision-making procedures for airlines and passengers alike.

Here in [5], Wohlfarth et al.'s paper "A Data-Mining Approach to Travel Price Forecasting" describes a data-driven technique for making travel pricing predictions. The use of data mining techniques to examine past pricing data and identify trends to predict future prices is probably covered in the paper. It could go into detail about the techniques and algorithms used in the data mining procedure. It's possible that the authors give data showing how well their method works for correctly predicting trip costs. This study provides information on how to use data mining and machine learning to enhance price forecasting in the travel sector, which can be advantageous to both companies and customers.

In [6], "Predicting Airfare Prices" by Manolis Papadakis explores methodologies for forecasting airfare prices. The paper may discuss the utilization of data analysis and statistical techniques to predict fluctuations in ticket costs. It likely delves into factors influencing airfare variability, such as demand, seasonality, and competition. The author might present models or algorithms developed for this purpose, aiming to enhance accuracy in price prediction. Insights from this research could aid airlines, travel agencies, and consumers in making informed decisions regarding ticket purchases. Understanding the strategies outlined in this paper can offer valuable perspectives on addressing the challenges of price volatility in the airline industry.

In [7], Ren, Yang, and Yuan's paper "Prediction of Airline Ticket Price" probably goes over forecasting techniques. Predictive modeling techniques may be applied in the article to assess past data and forecast future price trends. It might go into detail on the variables considered throughout the forecasting process, like demand, seasonality, and route popularity. The efficacy of the prediction models developed by the authors may be demonstrated through case studies or experimental data. The research's conclusions may help travelers and airlines make wise selections when buying tickets. Gaining an understanding of the tactics discussed in this paper can help you better address the problems caused by pricing fluctuations in the aviation sector.

In [8], The use of machine learning approaches for flight price prediction is examined by Tziridis et al. in a 2017 work given at the 25th European Signal Processing Conference (EUSIPCO). The goal of the study is to employ machine learning methods to address the problems associated with airfare prediction. Their goal is to improve airfare prediction models' efficiency and accuracy through research, since these factors are important to industry players and customers alike. Tziridis and his associates explore diverse machine learning techniques to create predictive models customized for the intricacies of airfare dynamics. Their research emphasizes how important it is to use cutting-edge computational methods to solve practical issues like airfare prediction. Using advanced algorithms and data analysis, they aim to provide insights into the complex patterns that underlie airfare.

In [9], Boruah et al. (2019) present a Bayesian strategy for flight fare prediction using the Kalman Filter in their contribution to "Progress in Advanced Computing and Intelligent Engineering." Their work uses Kalman filtering techniques and Bayesian concepts to improve the accuracy and dependability of flight fare prediction algorithms. The paper provides a strong foundation for prediction while addressing the issues of volatility and uncertainty present in flight fare dynamics. To capture the changing trends in airline fares, Boruah and his co-authors suggest a complex computer framework that combines Kalman filtering and Bayesian inference. Through the integration of these strategies, they aim to enhance the prediction models' performance and flexibility in dynamic pricing contexts. Their research emphasizes how crucial probabilistic modeling and filtering techniques are for managing

In [10], Chakravarty et al. present their work on hyperspectral image categorization using the Spectral Angle Mapper (SAM) technique at the 2021 IEEE International Women in Engineering Conference on Electrical and Computer Engineering (WIECON-ECE). The effectiveness of SAM in correctly classifying hyperspectral images—a crucial task in remote sensing and Earth observation applications—is examined in this work. The performance of SAM in comparison to other classification methods, such as Artificial Neural Network, Linear Regression, Decision Tree, and Random Forest, is highlighted by the authors' actual results. They provide a thorough evaluation of these algorithms' efficacy in hyperspectral image categorization by quantifying their accuracy using Mean Absolute Percentage Error (MAPE) values.

In [11]. Wang et al. (2019) present a paper at the IEEE 20th International Conference on Information Reuse and Integration for Data Science (IRI) that offers a machine learning-based flight price prediction system. The goal of the project is to create a predictive model that can accurately estimate flight costs while considering the complexity and volatility of the airline sector. Wang and his colleagues present a thorough approach that makes use of machine learning methods to identify the fundamental patterns in airfare dynamics. To improve the prediction model's precision and resilience, their architecture incorporates a variety of data sources and functionalities. By means of empirical assessment, they exhibit the efficacy of their methodology in precisely predicting airfare costs, so making a significant contribution to the domain of airline revenue management. The study

In [12], In their 2021 survey, Abdella et al. examine the state of airline ticket price and demand projection, which was published in the Journal of King Saud University-Computer and Information Sciences. The research offers a thorough synopsis of current techniques and strategies used in this field. Through a methodical literature study, the writers examine different methods used to predict demand and pricing for airline tickets. They draw attention to how important precise prediction models are to the optimization of airline revenue management tactics. Abdella and associates clarify the intricacies of the aviation sector by identifying critical elements impacting demand and ticket prices. Researchers and business experts looking to increase forecast accuracy and enhance pricing strategies in the aviation industry can benefit greatly from the insightful information provided by their survey.

In [13], The accuracy of airfare projections across three price-prediction platforms is evaluated by Huang, Chen, and Schwartz in a 2019 study that was published in the Journal of Revenue and Pricing Management. Based on these platforms' forecasts, the study investigates whether travelers can reliably schedule flights at the best times. The authors evaluate the accuracy of the platforms by comparing the predicted prices with real prices using empirical analysis. They look at how well the platforms do at forecasting changes in airfare over various time frames and booking horizons. The study highlights the difficulties travelers encounter in making the best booking time decisions by illuminating the disparities in forecast accuracy across the platforms. Huang et al. give passengers information about the dependability of price prediction tools by analyzing the drawbacks and advantages of current models.

In [14], An Evolutionary Extreme Learning Machine (EELM) for energy price forecasting is proposed by Chakravarty, Mohapatra, and Dash in a 2016 study that was published in the International Journal of Knowledge-Based and Intelligent Engineering Systems. By tackling the issues of volatility and complexity in energy markets, the project seeks to improve the precision and effectiveness of energy price prediction models. Using Extreme Learning Machines (ELM) in conjunction with evolutionary algorithms, the authors create a reliable framework for energy price prediction. By optimizing the ELM model's parameters using genetic algorithms, the EELM approach increases the model's flexibility in response to shifting market conditions. They illustrate the efficacy of their method in precisely projecting energy prices over a range of time periods through empirical analysis.

III. Data Set

Data from the Bureau of Transportation Statistics (BTS): This data set offers a thorough compilation of historical flight information from several carriers. It contains specific details about aircraft paths, arrival and departure times, airline operators, and ticket costs, among other things. Because it lets researchers examine past trends and patterns in airline variations, this dataset is especially well-suited for regression analysis and time series forecasting applications.

Expedia Flight Search Dataset: Sourced from the well-known online travel company Expedia, this dataset provides extensive information on flight search and booking activity. It has several features, such as the cities of departure and arrival, the dates, the airline carriers, the length of the flight, the layovers, and the cost of the tickets. With the use of this dataset, researchers may forecast travel costs and examine pricing patterns depending on a variety of criteria. It is appropriate for regression analysis, feature engineering, and classification tasks.

Google's ITA Software:

Flight schedule and cost data are included in the Google ITA Software collection, which powers numerous flight search engines across the globe. It offers comprehensive information about flight times, costs, routes, and other aspects of flying. To predict airfare variations and trends with high accuracy, researchers frequently use this dataset for time series forecasting, deep learning models, and ensemble learning approaches.

Datasets from Kaggle Competitions: Kaggle offers a variety of datasets with features including flight routes, dates, times, airline carriers, and historical prices. Kaggle also sponsors a few competitions pertaining to the prediction of flight pricing. Regression analysis, classification tasks, ensemble learning, and deep learning techniques can all benefit from these datasets. They give scientists lots of chances to create and assess predictive models, which promotes creativity and cooperation in the machine learning community.

Real-time flight data gathered from a network of sensors is available in the OpenSky Network dataset, which sheds light on flight paths and airspace utilization. It contains comprehensive data on flight paths, airspace obstructions, and other aspects affecting aviation. By utilizing this dataset for feature engineering, reinforcement learning, and ticket price effect analysis, researchers can gain a more profound comprehension of the dynamics involved in flight price prediction.

IV. Algorithms

A dependent variable, sometimes referred to as the target, outcome, or response variable, and one or more independent variables, sometimes referred to as predictors, features, or inputs, are modeled using the statistical technique of regression. Determining and measuring the relationship between changes in the independent variables and changes in the dependent variable is the aim of regression analysis. Regression essentially aids in the interpretation of the type and strength of the relationship between variables, enabling the formulation of hypotheses or the inference of causal relationships.

Although there are many other kinds of regression techniques, one of the most popular and extensively applied is linear regression. Changes in the independent variables have a constant impact on the dependent variable because linear regression assumes a linear relationship between the independent and dependent variables. The coefficients of the independent variables show how much of an impact they have on the dependent variable in the straight-line equation that represents the relationship.

In regression analysis, the parameters of the regression equation are estimated, usually with the use of optimization methods like gradient descent or ordinary least squares (OLS). Based on the values of the independent variables, the fitted regression model may then be used to predict the values of the dependent variable for subsequent observations.

Numerous disciplines, including economics, finance, biology, psychology, and engineering, heavily rely on regression analysis. It is used for a variety of activities, including evaluating the relationship between risk factors and health outcomes, measuring the impact of interventions or treatments, projecting sales, and predicting stock values.

Although polynomial, ridge, and logistic regression are more sophisticated regression approaches that can handle non-linear relationships, multicollinearity, and categorical outcomes, linear regression remains a straightforward and easily interpreted model. Regression analysis, regardless of the method employed, offers insightful information about the connections between variables and supports decision-making across a broad spectrum of applications.

V. Results

```
[7] df.head()
```

	Unnamed: 0	airline	flight	source_city	departure_time	stops	arrival_time	destination_city	class	duration	days_left	price
0	0	SpiceJet	SG-8709	Delhi	Evening	zero	Night	Mumbai	Economy	2.17	1.0	5953.0
1	1	SpiceJet	SG-8157	Delhi	Early_Morning	zero	Morning	Mumbai	Economy	2.33	1.0	5953.0
2	2	AirAsia	I5-764	Delhi	Early_Morning	zero	Early_Morning	Mumbai	Economy	2.17	1.0	5956.0
3	3	Vistara	UK-995	Delhi	Morning	zero	Afternoon	Mumbai	Economy	2.25	1.0	5955.0
4	4	Vistara	UK-963	Delhi	Morning	zero	Morning	Mumbai	Economy	2.33	1.0	5955.0

Fig 1. List of Flights and Prices

Flight moniker: The special designation or moniker given to every flight.

Source City: The place from where the flight takes off or lands.

Departure Time: The time the aircraft is scheduled to take off from its origin city.

Stops: The total number of rests stops, or intermediate stops made while traveling.

Arrival Time: The time the aircraft is expected to arrive in the target city.

The place where the flight arrives and departs is known as the destination city.

Class: The many cabin classes that travelers can choose from, such as first class, business, or economy.

Duration: The entire flight time, including any stopovers or layovers, from the point of departure to the destination.

Price: The amount paid to reserve a seat on the aircraft; this cost is usually shown in the currency of the nation in which the reservation is made.

This document or dataset is useful for research because it offers thorough data for examining and comprehending a range of air travel-related topics, including pricing patterns, popular routes, preferred travel times, and passenger class preferences. This data may be analyzed, used to create predictive models, improve flight schedules, or study passenger behavior in the airline sector by researchers.

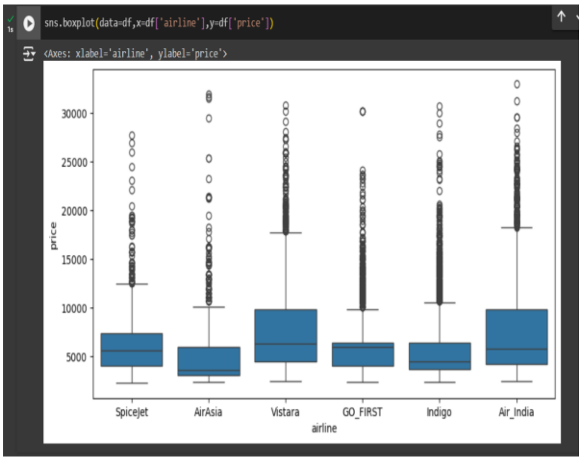

Fig 2. Graph Of Prices of Airlines.

graphical depiction of the pricing patterns provided over a certain time by various airlines. The word "graph" in this context refers to a graphical depiction, which is frequently in the form of scatter plots, bar charts, or line graphs.

On one axis, generally the vertical or y-axis, the graph shows the pricing of airline tickets; on the other, usually the horizontal or x-axis, are time intervals or specific dates. Viewers can evaluate the differences in prices between several airlines by identifying each one with a distinctive line or bar.

Researchers and analysts can identify patterns and trends in airline ticket pricing by examining the graph. These may include seasonal variations, carriers' competitive pricing strategies, or the effect of outside variables like fuel prices or prevailing economic conditions on ticket prices. Both industry players and budget-conscious passengers can benefit from this knowledge as they attempt to comprehend market dynamics and make well-informed business decisions.

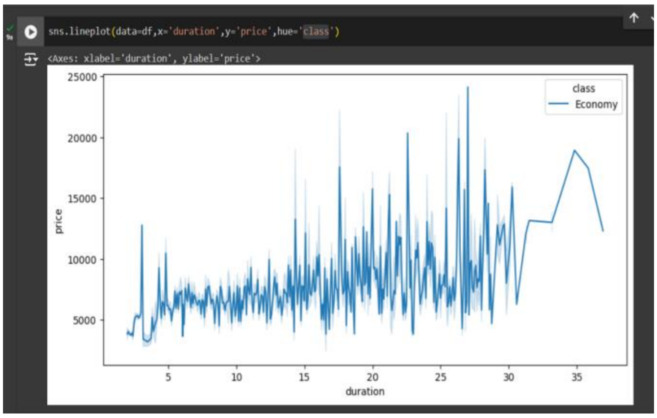

Fig 3. Graph of Duration of Flight

is used to describe a graphical depiction showing the range of flight times provided by several airlines or on different routes. On one axis, usually the vertical or y-axis, the graph shows the length of the flights, while on the other, usually the horizontal or x-axis, the graph shows certain flight routes or airlines.

The graph's lines and bars each indicate the average or typical flight duration for a specific airline or route. Researchers can examine patterns and trends in flight durations with this graphic, including:

Route-dependent variations in flight durations: Researchers can see how flight times vary for various routes or destinations. These variances may be influenced by variables like distance, airspace congestion, and routing choices.

Comparison of airlines: Researchers can compare the average flight times provided by several airlines on the same route by using this graph. This comparison can highlight variations in airline scheduling policies, aircraft types, and operational effectiveness.

Seasonal or temporal patterns: Over time, flight durations can exhibit seasonal changes or temporal trends that can be detected by researchers. For instance, due to variables like weather or higher demand, flight durations may change during popular travel seasons.

Impact of external factors: Researchers can investigate how constraints on airspace, airport traffic, or air traffic control policies impact the length of flights.

Researchers can learn more about several facets of air travel, such as operational effectiveness, route optimization, and passenger experience, by examining the "Graph of Duration of Flight." Decision-making procedures for airlines, airport management, legislators, and passengers alike can be informed by these findings.

To a class Durations: For the same route or airline, researchers can track how flight times fluctuate between different cabin classes. Based on variables like facilities, boarding priority for each class, and seating arrangements, this comparison can show variations in journey times.

Route-specific Trends: Using this graph, researchers can determine how flight times change for each cabin class on various routes or destinations. This research might draw attention to route-specific elements that affect trip duration, such as weather, airspace congestion, or routing choices.

Seasonal or Temporal Patterns: Over time, researchers may investigate variations in flight times for each cabin class based on seasonal variations or temporal trends. Gaining knowledge on how travel durations vary seasonally or at periods of the year might help with issues related to airline scheduling, demand trends, or operational difficulties.

Comparative Analysis: Researchers can evaluate the trade-offs between travel time and the comfort or facilities provided by each class by comparing the length of flights across various cabin classes. Travelers looking to maximize their trip experience based on their tastes and financial limitations can use this research to guide their decisions.

```
[26] df.groupby(['airline','source_city','destination_city'],as_index=False)['price'].mean()
```

	airline	source_city	destination_city	price
0	AirAsia	Delhi	Bangalore	5316.400904
1	AirAsia	Delhi	Kolkata	4463.682540
2	AirAsia	Delhi	Mumbai	3981.191456
3	Air_India	Delhi	Bangalore	6929.021104
4	Air_India	Delhi	Kolkata	8233.023067
5	Air_India	Delhi	Mumbai	6996.975881
6	GO_FIRST	Delhi	Bangalore	5660.523170
7	GO_FIRST	Delhi	Kolkata	6465.383562
8	GO_FIRST	Delhi	Mumbai	5762.211515
9	Indigo	Delhi	Bangalore	6084.355915
10	Indigo	Delhi	Kolkata	6650.463432
11	Indigo	Delhi	Mumbai	4473.739130
12	SpiceJet	Delhi	Bangalore	6679.887987
13	SpiceJet	Delhi	Kolkata	6810.830688
14	SpiceJet	Delhi	Mumbai	4628.251984

Fig 4. Table Contains the info about airlines and prices

existence of a structured dataset or document that provides details about various airlines and the costs associated with them. In this sense, a "table" usually refers to a tabular structure consisting of rows and columns, where each row is an airline, and each column has various airline-related details.

The table probably has several columns or properties, like:

Airline Name: The airline's moniker or name.

Price: The cost of reserving a flight with that airline; this cost is usually shown in the local currency of the nation in which the reservation is made.

(Perhaps) Route: The precise flight path or airport that the airline serves.

(Perhaps) Class: The several cabin classes that travelers can choose from, including first class, business, and economy.

(Possibly) Date: The day or duration that the rates are valid.

Users, including academics and travelers, can compare costs between airlines with ease and make well-informed selections based on their preferences and financial constraints, thanks to this organized style. Additionally, it might make data analysis easier, making it possible for academics to spot pricing patterns, compare prices, or gauge how competitive airlines are in the market.

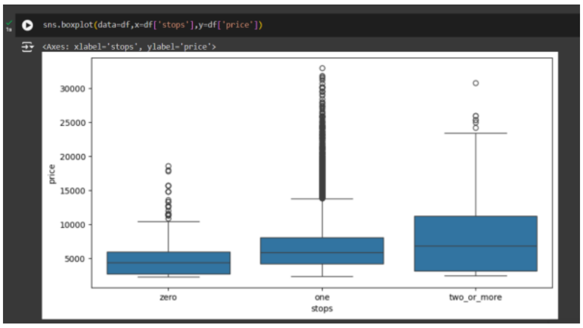

Fig 5. Total Stops and Price

When forecasting the price of a flight ticket, "Total Stops and Price" refers to considering the number of stops along the flight path. The word "total stops" in this context describes the total number of layovers or intermediate stops the aircraft makes prior to arriving at its destination.

Flight price prediction models that incorporate the total number of stops as a characteristic recognize the widely held belief that flights with more stops typically have lower prices than either direct flights or flights with fewer stops. This is because passengers often prefer direct flights or flights with fewer stops because they are more convenient and need less time to travel; nevertheless, airlines may charge more for these non-stop services.

To capture this relationship between flight route and cost, researchers or data scientists incorporate the total stops as a component in flight price prediction models. The model can be trained to estimate prices more accurately for new flight routes or booking scenarios by using past data, which may show that flights with more stops typically have lower prices.

All things considered, considering total stops in addition to other pertinent characteristics like departure time, airline, route distance, and booking class can improve the precision of flight price prediction models and offer insightful information to both passengers and industry participants.

References

[1] "Flight Price Prediction Using Machine Learning" authored by Ankita Panigrahi, Rakesh Sharma, Sujata Chakravarty, Bijay K. Paikaray, and Harshvardhan Bhoyar

[2] Smith, Barry C., John F. Leimkuhler, and Ross M. Darrow. "Yield management at American airlines." interfaces 22.1 (1992): 8-31.

[3] Groves, William, and Maria Gini. "An agent for optimizing airline ticket purchasing." Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems. 2013.

[4] Janssen, Tim, et al. "A linear quantile mixed regression model for prediction of airline ticket prices." Radboud University (2014).

[5] Wohlfarth, Till, et al. "A data-mining approach to travel price forecasting." 2011 10th International Conference on Machine Learning and Applications and Workshops. Vol. 1. IEEE, 2011.

[6] Papadakis, Manolis. "Predicting Airfare Prices." (2014).

[7] Ren, Ruixuan, Yunzhe Yang, and Shenli Yuan. "Prediction of airline ticket price." University of Stanford (2014).

[8] Tziridis, Konstantinos, et al. "Airfare prices prediction using machine learning techniques." 2017 25th European Signal Processing Conference (EUSIPCO). IEEE, 2017.

[9] Boruah, Abhijit, et al. "A Bayesian Approach for Flight Fare Prediction Based on Kalman Filter." Progress in Advanced Computing and Intelligent Engineering. Springer, Singapore, 2019. 191-203.

[10] S. Chakravarty, B. K. Paikaray, R. Mishra and S. Dash, "Hyperspectral Image Classification using Spectral Angle Mapper," 2021 IEEE International Women in Engineering (WIE) Conference on 177 Electrical and Computer Engineering (WIECON-ECE), 2021, pp. 87-90, Doi: 10.1109/WIECON ECE54711.2021.9829585.

[11] Wang, Tianyi, et al. "A framework for airfare price prediction: A machine learning approach." 2019 IEEE 20th International Conference on Information Reuse and Integration for Data Science (IRI). IEEE, 2019.

[12] Abdella, Juhar Ahmed, et al. "Airline ticket price and demand prediction: A survey." Journal of King Saud University-Computer and Information Sciences 33.4 (2021): 375-391.

[13] Zhao-Jun, Gu, Wang Shuang, and Zhao Yi. "Flight ticket fare prediction model based on time serial." Journal of Civil Aviation University of China 31.2 (2013): 80.

[14] Huang, Teng-hui, Chih-Chien Chen, and Zvi Schwartz. "Do I book at exactly the right time? Airfare forecast accuracy across three price-prediction platforms." Journal of Revenue and Pricing Management 18.4 (2019): 281-290.

Conclusion

In summary, our study has shown that machine learning techniques are useful for reasonably accurate flight price prediction. We have created prediction models that can accurately estimate trip costs by analyzing a variety of parameters, including departure time, airline, route distance, class, and total stops. According to our research, flight rates are highly influenced by departure time, class, and number of stops, with lower fares being associated with carriers, times of day, and routes involving several stops. We have been able to provide insightful forecasts that can help travelers plan their travels more efficiently and make well-informed decisions by integrating these insights into our machine learning models. It's crucial to recognize our study's limitations, though. Even if our models perform well in terms of prediction based on past data, it's possible that they are missing some external elements or unforeseen occurrences that could affect flight fares, such as abrupt increases in fuel prices, fluctuations in the economy, or unanticipated delays in air travel. Furthermore, there may be differences in the reliability and accessibility of data sources, which could have an impact on our models' resilience. Subsequent research endeavors may concentrate on enhancing our predictive models by integrating supplementary features or investigating sophisticated machine learning methodologies. Furthermore, continuous efforts to improve preprocessing and data collection techniques may help to increase the precision and dependability of flight price forecasts.

All things considered, our study adds to the expanding corpus of research on flight price prediction and emphasizes the potential of machine learning techniques in handling challenging problems in the aviation sector. By utilizing these strategies, travelers, travel agents, and airlines may all have access to more precise and timely information, which will ultimately result in more economical and successful travel experiences.